DOCTORAT ECOLOGIE
BRETAGNE GEOSCIENCES
LOIRE AGRONOMIE ALIMENTATION

Oniris
École Nationale Vétérinaire, Agroalimentaire et de l'Alimentation Nantes Atlantique
*Santé et alimentation au cœur de la vie*

# THÈSE DE DOCTORAT DE

ONIRIS

ÉCOLE DOCTORALE N° 600
*Écologie, Géosciences, Agronomie et Alimentation*
Spécialité : *Epidémiologie, évaluation des risques*

Par
**Mathilde MERCAT**

# Conception et évaluation d'une méthode d'estimation d'une probabilité d'infection d'un troupeau à partir de données hétérogènes : contribution au développement d'une surveillance épidémiologique basée sur la comparabilité des résultats

**Thèse présentée et soutenue à Nantes, le 07 Juillet 2021**
**Unité de recherche : UMR INRAE 1300 Bioepar, Oniris, Nantes**

**Rapporteurs avant soutenance :**

| | | |
|---|---|---|
| Pascal HENDRIKX | | Inspecteur général de santé publique vétérinaire, Ecole Nationale des services vétérinaires-FVI |
| Karine CHALVET-MONFRAY | | Professeur, VetAgro Sup |
| Président : | Prénom NOM | Fonction et établissement d'exercice *(à préciser après la soutenance)* |
| Examinateurs : | Claude SAEGERMAN | Professeur à la Faculté de Médecine Vétérinaire, Université de Liège |
| | Timothée VERGNE | Maître de conférence à l'Ecole Nationale Vétérinaire de Toulouse (ENVT) |
| | Nathalie BAREILLE | Professeur à Oniris |
| Dir. de thèse : | Christine FOURICHON | Professeur à Oniris |
| Co-dir. de thèse : | Aurélien MADOUASSE | Maître de conférence à Oniris |

# Remerciements

A Karine Chalvet-Monfray et Pascal Hendrikx, merci d'avoir accepté de vous investir en tant rapporteur de mon jury de thèse.

A Timothée Vergne et Claude Saegerman, merci d'avoir accepté de faire partie de ce jury de thèse en tant qu'examinateurs.

A Nathalie Bareille, merci d'avoir accepté d'être membre de ce jury et pour ta bonne humeur au quotidien au sein de l'unité.

A Christine Fourichon, ma directrice de thèse, merci de m'avoir accompagné durant ce projet de thèse. Merci pour les discussions scientifiques et pour ta pédagogie. Merci d'avoir toujours pris le temps d'expliquer tes commentaires, tant sur la forme que sur le fond. Merci pour ta rigueur scientifique et d'avoir su te rendre disponible quand c'était nécessaire.

A Aurélien Madouasse, mon co-directeur de thèse, merci pour ton encadrement et ton accompagnement au quotidien. Merci pour les discussions scientifiques toujours stimulantes, parfois longues. Merci, aussi pour tes relectures et corrections approfondies.

To all the members of the STOC free project, thank you for welcoming me into your project with such simplicity. Thank you for allowing me to understand all the ins and outs of such a project, both organisational and scientific. Thank you for the annual face to face meetings. As much for the interminable but very enriching scientific discussions, which always allowed me to take a step back from my work, as for all the other moments of conviviality spent together.

To annika, thank you for being a great partner in this project, as a PhD student of the STOC free project. Luck was on our side, and especially on mine, thank you very much for all the support you gave me in this project. And also for the express visit of Amsterdam, I'll come back... and I'm waiting for you in the south.

Aux membres de l'ED EGAAL, et tout particulièrement à Anne Huet et Hervé Prévost, merci pour votre soutien et votre écoute. Merci aussi à Benjamin Rondeau pour ton écoute et tes conseils avisés en tant que tuteur.

Aux membres de l'équipe peps, merci de m'avoir accueillie dans votre équipe et pour la bonne humeur quotidienne dans le couloir.

A l'équipe dynamo, merci pour l'accueil. Merci pour ces fameux échanges lors des repas, tout aussi passionnant que parfois futile. Merci de m'avoir permis de réviser l'histoire française, les présidents de la cinquième république, les b.a-ba de la cafetière à piston... Et tout le reste. Je pars avec le regret d'avoir loupé la discussion sur la transcendance de pi.

A Michel, Sylvie, Juliette, Cécile et Erwann, merci pour votre appui au quotidien, pour tout (et parfois n'importe quoi). Merci d'avoir toujours été disponibles.

A mes compagnons de fortunes : Axelle, Marie, Hélène, Vianney, Romain, Georges, Nikki, Adèle, Guillaume, Caroline, qu'elle chance d'avoir croisé autant de doctorants pendant ma thèse ! Merci pour les restau, qui nous faisaient oublier nos salaires de doctorants, les discussions et le soutien, surtout au quotidien. Marie, ma cops de bureau, merci pour tout (les discussions et les fous rires) et pour Dadju. Axelle, merci pour ton soutien, même après ton départ. Marie, Adèle, Georges, Caroline (et un peu Hélène) merci pour les footings pour vider les méninges... Bref merci à tous.

A mes parents, merci de nous avoir toujours poussés et donné la possibilité de faire nos propres choix. Merci, d'avoir toujours cru en nous, quelque soit le domaine. Quelques lignes ici ne suffiront pas à vous dire à quel point je suis reconnaissante de la famille que vous nous avez construite.

A ma mère, la mère, merci pour ton soutien et ton aide, tant psychologique que logistique. Merci pour tes engagements aussi et pour ton courage, avec pour preuve ta reprise d'étude à l'aube de la retraite.

A mon père, merci pour la curiosité dont tu fais preuve sur tout, pour tout et tout le temps, et qui a surement développé la mienne. Merci à toi et à Nicole pour les chouettes moments partagés durant cette thèse.

Aux Walter et Ramognino, doux mélange de méditerranée et d'Alsace. Merci de m'avoir accueillie dans votre famille et un grand merci pour le soutien, pour les bricolages, pour les bons plats de Mamily (et de Marie-Christine aussi), et surtout pour la bonne humeur à chaque moments partagés.

A ma sœur, Gueuthe, merci d'être aussi givrée que moi et merci surtout de toujours croire en ce que je fais, quoique ce soit d'ailleurs.

A mon frère, Ben, (quoique, toi qui était passé par là, je ne sais pas si je dois te remercier) mais merci quand même pour m'avoir montré le chemin et de me redonner confiance, et à Perrine, merci pour les moments sportifs ou non et pour le soutien, parfois même à distance, pendant ces années de thèse.

A mes neveux et nièces, Eliott et Manon, chouchou et choupette, juste merci d'être des petits rayons de soleil pleins de vie et d'énergie.

A MC, merci pour ces années d'études partagées (et ces projets absurdes finis à des heures nocturnes), mais surtout pour l'amitié indéfectibles. Tu fais partie de ces amis qui font partie de la famille. Alors merci et vivement la suite.

Et à tout le reste de ma grande famille merci (vous citer tous serait bien trop long) mais merci.

A Alicia, *the last but not the least*, merci pour le soutien (et tout le reste) que tu m'apporte en toute circonstance et au quotidien, à 1000km de là pendant 3 ans, et puis juste a coté depuis quelques mois, et encore plus ces dernières semaines. Ce n'est que le début, mais ma foi je finirai par aimer ton sud.

# Liste des abréviations

## Français

— BVD : Diarrhée bovine virale
— BVDV : Virus de la diarrhée bovine virale
— IPI : Infecté immunotolérant permanent

## English

— Ab : Antibody
— Ag : Antigen
— BTM : bulk tank milk
— BVD : Bovine Viral Diarrhea
— BVDV : Bovine Viral Diarrhea virus
— No. : Number
— RF : Risk factor
— Se : Sensitivity
— Sp : Specificity
— CP : Control programme
— TI : Transiently infected
— PI : Persistently infectide

# TABLE DES MATIÈRES

# TABLE DES FIGURES

# Liste des tableaux

Chapitre 1

# Introduction générale

Maîtriser la propagation des maladies infectieuses animales endémiques est un enjeu majeur de santé animale et parfois de santé publique. Le caractère endémique se définit comme la présence et la persistance pendant une longue période et sur un territoire donné de l'infection, avec une prévalence stable ou qui varie relativement lentement. Les infections endémiques se transmettent durablement dans les populations animales au sein d'un territoire. Leurs conséquences, à la fois économiques et sur le bien-être animal, peuvent durer dans le temps. La gestion des maladies endémiques peut également être essentielle en santé publique pour contrôler le risque zoonotique lié à certains agents pathogènes. L'épidémie de fièvre Q aux Pays Bas de 2007 à 2010 en est un exemple (Van der Hoek *et al.*, 2010). La fièvre Q est une zoonose causée par *Coxiella burnetii* et endémique au sein des élevages bovins.

Diverses actions de maîtrise existent afin de lutter contre les maladies infectieuses endémiques. Des mesures instaurées au cas par cas à l'initiative des éleveurs pour limiter les conséquences de l'infection au sein de leur troupeau sont des mesures dites non systématiques (Moennig *et al.*, 2005; Lindberg & Houe, 2005). Bien qu'apportant des bénéfices non négligeables au sein de l'exploitation, les mesures non systématiques ne permettent pas de maîtriser ou d'éradiquer l'infection à l'échelle d'un territoire. Le caractère transmissible des maladies infectieuses et la gestion au cas par cas ne permettent pas de maîtriser le risque de transmission des troupeaux infectés vers les autres troupeaux. De plus, des mesures dites systématiques peuvent être mises en place à l'échelle d'une population. Il s'agit de programmes de maîtrise collectifs, à l'échelle régionale ou nationale, coordonnant les efforts avec un ou plusieurs objectifs communs tels que la maîtrise ou l'éradication de l'infection à l'échelle du territoire.

Les programmes de maîtrise systématiques classent généralement les troupeaux en fonction de leur statut vis à vis de l'infection, en mettant en place une surveillance régulière de ces statuts basée sur un dispositif de dépistage. Cette surveillance vise à détecter les troupeaux *infectés*, mais également à identifier les troupeaux *indemne d'infection* (Houe

*et al.*, 2014). Des mesures de contraintes et ou d'élimination de l'infection peuvent être appliquées aux troupeaux classés *infectés*, notamment des mesures d'interdiction de vente d'animaux. En complément, la certification des troupeaux *indemne d'infection* peut lever ces contraintes, tout en limitant les risques de transmission d'agents pathogènes par le commerce d'animaux entre troupeaux.

Pour les maladies infectieuses dites réglementées définies par l'OIE [1], l'Union Européenne prescrit des standards de surveillance, identiques pour tous les territoires [2]. La définition d'un statut *indemne d'infection* d'un troupeau se base alors sur les mêmes modalités, notamment sur une stratégie d'échantillonnage et de tests de dépistage identiques. Pour les maladies dites non réglementées, les programmes collectifs de maîtrise sont des initiatives régionales ou nationales. Du fait de l'impact non négligeable des maladies animales infectieuses non réglementées, de nombreux programmes collectifs ont été mis en place, notamment en élevage bovin, pour la diarrhée virale bovine (BVD) (van Roon *et al.*, 2020b), pour la paratuberculose (Whittington *et al.*, 2019) ou pour la rhinotrachéite infectieuse bovine (Raaperi *et al.*, 2014). Chacun de ces programmes mis en place a sa propre définition du statut *indemne d'infection*. Pour ces infections, il n'existe donc pas de standard de surveillance à l'échelle de l'Union Européenne, créant une hétérogénéité de définition de statuts *indemne d'infection* entre les territoires.

Différents territoires sont en contact via le commerce d'animaux, ce qui peut représenter un risque d'introduire ou de ré-introduire certaines maladies infectieuses. Si on considère un territoire indemne d'une infection, l'achat d'un bovin issu d'un territoire non indemne peut représenter un risque d'introduction. Acheter un bovin dans un troupeau classé *indemne d'infection* peut sembler réduire le risque, cependant, une incertitude existe toujours autour du statut *indemne d'infection*. En effet, le statut se base principalement sur un dispositif de dépistage appliqué à un moment donné, souvent sur un échantillon d'animaux, et reposant sur l'usage de tests de diagnostic imparfaits. Selon les programmes, l'incertitude du statut peut varier car des tests différents peuvent être appliqués sur des catégories et des effectifs de bovins différents, et à une fréquence qui varie entre les programmes. Chaque programme de maîtrise des maladies non réglementées définit ainsi ses propres standards pour qu'un troupeau soit classé *indemne d'infection*, pour adapter au mieux le dispositif de surveillance au contexte du territoire. Ces différences qui

1. Organisation mondiale de la santé animale.

2. Les plans de maîtrise peuvent s'appliquer à un département une région ou un pays. La notion de territoire est définit ici comme la zone géographique délimitée où, pour une infection donnée, un seul plan de maîtrise est en place.

visent à optimiser un dispositif de surveillance dans un contexte donné rendent impossible la comparaison des probabilités qu'un troupeau soit classé à tort *indemne d'infection* entre deux programmes, car les valeurs prédictives négatives sont inconnues.

Cette introduction générale présentera dans un premier temps la diversité des programmes de maîtrise, puis dans un second temps, la nécessité de développer de nouveaux outils pour la surveillance dite *output-based*[3].

# 1 Diversité des plans de maîtrise des maladies non réglementées en Europe

A l'échelle européenne, divers programmes de maîtrise ont été mis en place pour différentes infections (Figure 1.1). Ces programmes présentent une structure générale similaire (Figure 1.2). Cependant, l'objectif et l'historique, la participation des exploitations, ainsi que la définition du statut *indemne d'infection* de chaque programme varient d'un programme à un autre, créant une grande hétérogénéité de situations.

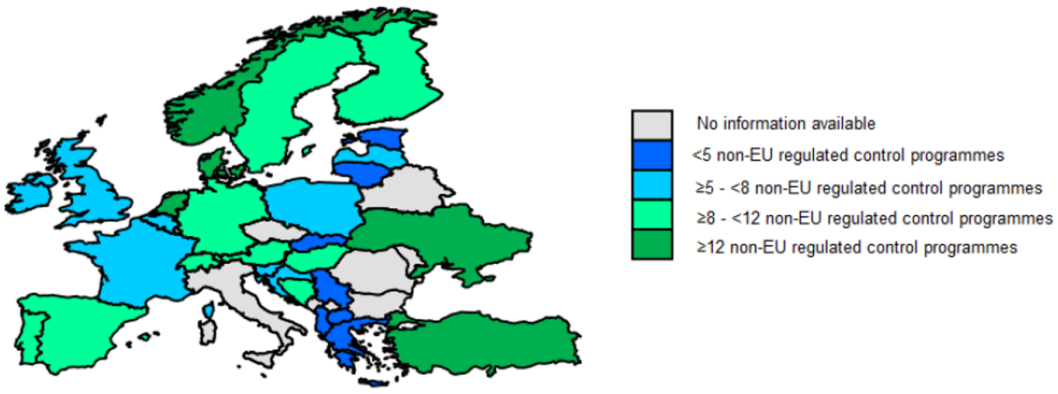



FIGURE 1.1 – Nombre de programmes de maîtrise implémentés pour les maladies non règlementées par pays participant au projet SOUND-Control COST ACTION (*Standardizing OUtput-based surveillance to control Non-regulated Diseases in the EU*) d'après Costa *et al.* (2020).

3. Le terme anglais *output-based* est gardé, car c'est un concept relativement récent pour lequel il n'existe pas, à ma connaissance, de traduction française acceptée.

Plan de surveillance

**Objectifs :**

- identifier les troupeaux infectés
- identifier les changements de statuts vis à vis de l'infection
- certifier des troupeaux indemnes d'infection

**Moyens :**
Dispositif de dépistage, adapté au contexte et objectifs, définis par :

- ⋆ un ou des tests de diagnostics
- ⋆ une fréquence de réalisation des tests
- ⋆ un échantillon d'animaux qui se veut représentatif du troupeau
- ⋆ une règle de classement des troupeaux basée sur le dispositif

Plan d'assainissement :

**Objectifs :**
Éliminer l'infection dans le troupeau

**Moyens :**
Dispositif de dépistage pour identifier les animaux infectés

- ⋆ un ou des tests de diagnostics
- ⋆ sur des sous-échantillons du troupeau (groupe, individus)
- ⋆ sur une période donné

→ Élimination des animaux infectés(cas IPI BVD)

Mesures de biosécurité :

**Objectifs :**
Limiter l'introduction de l'infection

**Moyens :**

- ⋆ Identifier les facteurs de risques d'introduction
- ⋆ Mise en place de mesure de contrôle des facteurs de risques

FIGURE 1.2 – Modèle général des programmes de surveillance : objectifs et moyens des différentes composantes.

## 1.1 Objectifs et historiques

Un programme collectif de maîtrise est construit autour d'un objectif qui peut être soit la maîtrise, soit l'éradication de l'infection à l'échelle d'un territoire. La maîtrise de l'infection a généralement pour objectif de réduire la prévalence de l'infection à un niveau relativement faible, bien que la transmission soit suffisamment fréquente pour empêcher

son éradication (Andrews & Langmuir, 1963), et vise parfois à limiter seulement les pertes associées à l'infection. L'éradication vise à diminuer la prévalence d'une infection spécifique jusqu'à l'absence continue de transmission dans un territoire déterminé (Andrews & Langmuir, 1963).

Les programmes de maîtrise des maladies non réglementées sont des initiatives locales ou nationales et n'ont pas le même historique selon les territoires. Dans certains territoires, les programmes de maîtrise ont débuté il y a longtemps et leur objectif respectif est atteint, ou au moins bien avancé, alors que pour d'autres territoires, les mesures de maîtrise systématiques ont été mises en place bien plus récemment. L'historique de la mise en place de plan de maitrise de la BVD dans 6 pays européens en est un exemple (Figure 1.3).

Ainsi, aujourd'hui à l'échelle de l'Europe, certains territoires ont éradiqué certaines maladies infectieuses alors que d'autres n'en sont encore qu'aux prémices. Cette hétérogénéité de circulation de l'infection rend certains territoires à risque de réintroduction. La Suède qui a atteint l'éradication pour 13 maladies infectieuses endémiques non réglementées (Hodnik *et al.*, 2020) est ainsi particulièrement à risque.

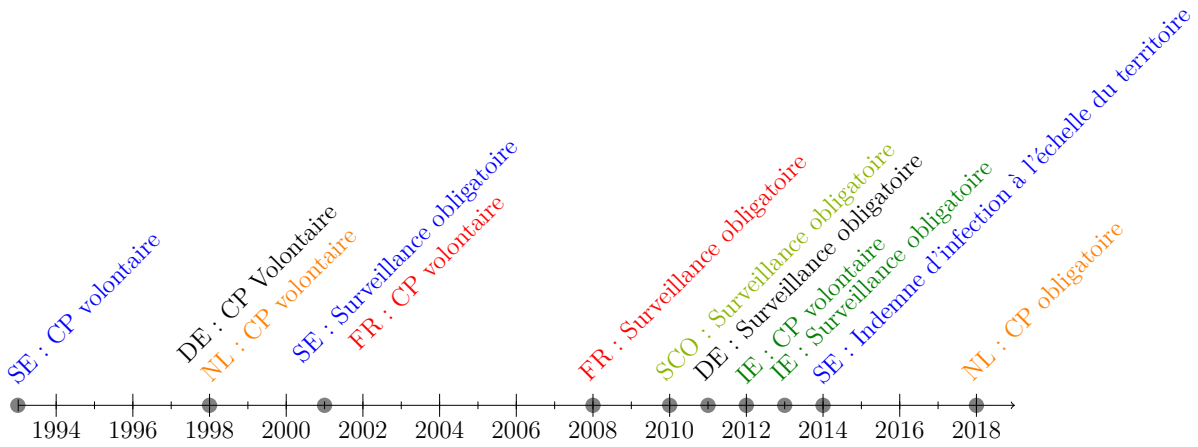



FIGURE 1.3 – Frise chronologique de la mise en place des programmes de maîtrise (CP) de la diarrhée virale bovine dans 6 pays européens (Allemagne (DE) en noir, Ecosse (SCO) vert clair, France (FR) en rouge, Irlande (IE) en vert foncé, Pays-Bas (NL) en orange et Suède (SE) en bleu) d'après van Roon *et al.* (2020b). La France ayant des programmes régionaux ou départementaux, est présentée ici la mise en place de programmes en Bretagne.

## 1.2 Participation volontaire ou obligatoire des exploitations

La participation des exploitations à un programme de maîtrise peut être basée sur le volontariat ou rendue obligatoire. La participation sur la base du volontariat maintient une hétérogénéité au sein du territoire, avec une partie seulement des exploitations qui visent à maîtriser l'incidence de l'infection. Le manque d'informations sur la circulation de l'infection dans les exploitations ne participant pas au programme, peut être un risque pour les exploitations voisines qui participent au programme. Le caractère volontaire de la participation au programme de maîtrise peut évoluer avec le temps (van Roon *et al.*, 2020b) et devenir obligatoire, en mobilisant des instruments réglementaires. De plus, au sein d'un même programme de maîtrise, certaines composantes seulement peuvent être rendues obligatoires. Par exemple, en Bretagne, la connaissance du statut troupeau vis à vis de l'infection est obligatoire depuis 2008 par un arrêté préfectoral (Bernard, 2018). Par contre, la participation au plan d'assainissement de l'infection à l'échelle troupeau restait, elle, la décision de l'éleveur. En 2019, une réglementation nationale a rendu le dépistage de la BVD obligatoire dans tous les troupeaux bovins en France [4].

## 1.3 Diversité des définitions du statut *indemne d'infection*

La définition du statut *indemne d'infection* repose sur un dispositif de dépistage spécifique qui varie d'un programme de maîtrise à un autre. Le dispositif de dépistage est définit par l'utilisation d'un ou plusieurs tests de diagnostics réalisés à intervalle de temps régulier sur un animal ou un groupe d'animaux, sélectionnés pour représenter le statut du troupeau. La sensibilité du dispositif est sa capacité à détecter les troupeaux infectés. La spécificité du dispositif est sa capacité à ne détecter que les troupeaux infectés. On distingue deux types d'erreurs du dispositif de dépistage : un troupeau infecté peut être classé à tort *indemne d'infection* (faux négatifs) et un troupeau indemne peut être classé à tort *infecté* (faux positifs).

Les caractéristiques du test de dépistage tel que sa cible biologique, sa spécificité et sa sensibilité, ainsi que les animaux échantillonnés, peuvent impacter les performances du dispositif de dépistage. Par exemple, le test utilisé peut être très sensible et spécifique pour détecter la présence d'anticorps dans un échantillon. Mais, la présence d'anticorps détectés à un temps donné ne reflète pas forcement la présence actuelle de l'infection,

4. D'après l'Arrêté du 31 juillet 2019, "Fixant des mesures de surveillance et de luttle contre la maladie des muqueuses/diarhée virale bovine (BVD)", JORF n$^o$0177 du 1 août 2019.

TABLE 1.1 – Définition des statuts troupeaux *indemnes d'infection* vis à vis de la BVD pour 6 programmes de maîtrise en Europe d'après van Roon *et al.* (2020b).

| Participation territory | Definition of freedom at herd level |
|---|---|
| Germany | All cattle are virus negative for 24 mo ; no contact with non-free farms[1] |
| France (Brittany) | Dairy : At least "A" status[2] after 3 consecutive tests with results 000, 010 or 100[3]<br>Beef breeding : 2 consecutive negative tests for all animals tested in the screening spot test (A status) |
| Ireland | > 3-yr participation ; all animals with known negative status (direct or indirect) and no persistently infected infected animal (PI) present ≥ 1yr |
| The Netherlands | No virus positive cattle during 10 mo |
| Sweden[2] | National level : Surveillance is designed to reach, annually, a 0.99 probability of freedom (design prevalence 0.2% at herd-level, 99% confidence) |
| Scotland | No free designation ; farms are either negative or not negative after testing |

[1] All cattle in the herd are free from clinical signs suggestive of BVDV infection. All cattle born in the herd have been tested for BVDV within 30 d after birth using a method described in the official set of methods with a negative result. Only BVDV-unsuspected cattle have been added to the herd. The cattle of the herd have not been in contact with cattle outside the herd that are BVDV-suspect. The cattle of the herd may be inseminated only with seeds of BVDV-unsuspected bulls or in case of natural breeding, only BVDV- unsuspected bulls have been used.
[2] Herds are classified based on 3 consecutive bulk milk tests. Status "A" means that a herd had 3 results of 0 in the bulk milk screening. All dairy and dry cattle receive this status. After more test rounds, herds can receive "super A" status followed by "super A+" status and eventually "A+ 90 d/180 d" status.

[3] The results of 3 consecutive tests. For example, 000 means 3 consecutive tests with test result 0. Test result 0 : <10% of the cows are positive ; test result 1 : 10–30% of the cows are positive ; test result 2 : >30% of the cows are positive ; test result n : very low percentage of cattle.

notamment si l'immunité dure toute la vie de l'animal. La répétition des tests à intervalles réguliers permet l'interprétation : la succession d'un résultat négatif puis positif met en évidence une séroconversion du troupeau et peut indiquer que le troupeau a été infecté entre les deux tests (sauf si vaccination ou achat d'animaux immunisés). L'interprétation du résultat du dispositif dépend également des animaux testés qui doivent représenter au mieux le troupeau. Si l'agent pathogène ne circule que dans un groupe d'animaux, sans contact avec les autres, et que le test est effectué sur un autre groupe d'animaux alors il ne sera pas détecté. Enfin, la fréquence à laquelle le test de dépistage est réalisé peut entraîner un retard de détection, qui augmente avec l'augmentation de l'intervalle de temps entre deux tests. Un troupeau qui s'infecte après un test de diagnostic reste classé *indemne d'infection* jusqu'au prochain test, au moins.

Chaque programme de maîtrise ayant ses propres stratégies de surveillance, l'incertitude autour du vrai statut d'un troupeau classé *indemne d'infection* varie en fonction du programme. Une comparaison qualitative de 6 programmes de maîtrise de la BVD en Europe a illustré la diversité des définitions de statuts *indemne d'infection* (van Roon *et al.*, 2020b) (Table 1.1). La diversité des dispositifs de dépistages de la BVD a aussi été mis en évidence par Duncan *et al.* (2016). Cette diversité ne permet pas de comparer directement les dispositifs de dépistage entre eux et ainsi, le risque qu'un troupeau infecté soit classé *indemne d'infection* par erreur quel que soit le programme dans lequel il est inclut.

# 2 Le développement de méthodes pour l'*output-based surveillance*

## 2.1 Définition et limites de l'*input-based surveillance*

Au sein de l'Union Européenne, la maîtrise des maladies réglementées repose sur une obligation de moyens, c'est la surveillance dite *input-based*. Les standards européens prescrivent les mesures à mettre en œuvre pour atteindre le statut *indemne d'infection* . Chaque pays est contraint de mettre en place le même dispositif de surveillance (tests, animaux testés...).

Cependant, les méthodes *input-based* présentent plusieurs limites liées au fait que, prescrire les mêmes mesures de surveillance pour des territoires différents implique de considérer les territoires comme homogènes. Or, de part la variété des contextes, la prévalence et l'incidence de l'infection varient entre les territoires. Les troupeaux issus de territoires différents n'ont donc pas la même probabilité d'être infectés. Ainsi, la probabilité d'être classé *indemne d'infection* à tort n'est pas la même d'un territoire à un autre, malgré un dispositif de dépistage identique. En effet, la valeur prédictive négative varie en fonction de la prévalence. Il en va de même pour des troupeaux d'un même territoire, où le risque d'être infecté peut varier en fonction du contexte et des pratiques. De plus, du fait des pratiques d'élevages, telles que la taille des troupeaux ou la structuration des lots, les mêmes modalités d'échantillonnage et l'utilisation du même test peuvent entrainer des résultats différents en terme de sensibilité et spécificité du dispositif de dépistage. Ainsi, les mesures prescrites peuvent être excessives pour les troupeaux (ou territoires) à faible risque, ou insuffisantes dans les troupeaux (ou territoires) où le risque est plus élevé

(Cameron, 2012). Les limites de *l'input-based* surveillance ont été mis en évidence dans le cadre de maladies dîtes réglementées (on Animal Health & (AHAW), 2012).

## 2.2 La surveillance fondée sur le risque

Afin de prendre en compte la diversité des contextes entre et au sein des territoires, la surveillance fondée sur le risque s'est grandement développée. La surveillance fondée sur le risque prend en compte l'information sur la probabilité d'occurrence et l'ampleur des conséquences, biologiques et/ou économiques, des risques sanitaires pour planifier, et concevoir et/ou interpréter les résultats obtenus par les systèmes de surveillance (Hoinville *et al.*, 2013). Une des composantes principales est la conception d'une stratégie d'échantillonnage fondée sur le risque. Un effort supplémentaire est fourni pour aller chercher l'infection où elle est le plus probable d'être et ainsi augmenter la sensibilité du système de surveillance. Les troupeaux ayant la plus grande probabilité d'être infecté seront d'autant plus échantillonnés. L'actuel programme de surveillance de la BVD en Suède, qui est *indemne d'infection* depuis 2014 (Norström *et al.*, 2014), catégorise les troupeaux en fonction de leur probabilité d'être infecté chaque année (élevé, moyenne, faible), ce qui définit la fréquence des tests (van Roon *et al.*, 2020b). Ce mode de surveillance peut aussi être appliqué dans des territoires non indemne d'infection. On peut alors catégoriser les troupeaux soit en fonction de leur probabilité d'être infecté soit en fonction des conséquences si un troupeau n'est pas détecté à temps (Cameron, 2012).

Diverses méthodes d'évaluations des programmes de surveillances ont alors été développées pour améliorer leurs performances. L'évaluation peut être complexe, car les programmes de surveillance ont de multiples composantes. Calba *et al.* (2015) ont identifié les critères pour développer une approche complète d'évaluation des systèmes de surveillance et ont également mis en avant le manque de considération des aspects économiques et sociologiques dans l'évaluation des programmes de surveillance. Le consortium RISKSUR a élaboré une approche d'évaluation de la surveillance intégrant notamment une analyse coûts-bénéfices (Peyre2019).

## 2.3 Définition, objectifs et limites des outils disponibles pour l'*output-based surveillance*

Pour les maladies endémiques non règlementées, des programmes de maîtrise ont été développés en fonction des spécificités de chaque territoire. La construction des pro-

grammes de maîtrise a été à la fois adaptée au contexte épidémiologique, et aux moyens, techniques et/ou financiers, de chaque territoire. La diversité des territoires a entrainé le développement de programmes divers, aux composantes variées ne permettant pas de comparer directement les résultats des programmes, tel que, le risque qu'un troupeau infecté soit classé *indemne d'infection* par erreur quel que soit le programme dans lequel il est inclut. De part la diversité des programmes mis en place, notamment à l'échelle européenne, il est nécessaire d'instaurer des méthodes de surveillance dîtes *output-based* (More *et al.*, 2009; Cameron, 2012; Norström *et al.*, 2014; Schuppers *et al.*, 2012; Foddai *et al.*, 2016). La surveillance *output-based* prescrit les résultats à atteindre, et non les moyens à mettre en place (Cameron, 2012), et nécessite le développement de méthodes permettant de comparer les résultats des programmes de maîtrise.

### 2.3.1 Estimation de la probabilité d'absence d'infection à l'échelle d'un territoire avec les arbres de décisions

Martin *et al.* (2007) ont développé une méthode basée sur la modélisation des *scenario trees* pour démontrer l'absence d'infection à l'échelle d'un territoire. En effet, une longue série de résultats de tests négatifs sur l'ensemble d'un territoire ne peut suffire à démontrer l'absence d'infection, à cause de l'imperfection des tests, de l'échantillonnage et de la probabilité de ré-introduction de l'infection. La méthode de Martin *et al.* (2007) estime la probabilité de ne pas avoir détecté l'infection si elle était présente à une prévalence de référence faible, appelée *design prevalence*, en prenant en compte l'imperfection des tests et de l'échantillonnage. Elle permet d'intégrer à la fois les résultats de tests ainsi que la probabilité d'infection. Cette méthode a notamment permis de démontrer l'absence d'infection de la Suède par la BVD et de soutenir que la Suède est bien indemne de la BVD (Norström *et al.*, 2014).

### 2.3.2 Estimation de la probabilité d'absence d'infection à l'échelle du troupeau dans les territoires non indemnes

Pour les territoires n'ayant pas atteint l'éradication, démontrer l'absence d'infection à l'échelle du troupeau peut permettre de sécuriser le commerce d'animaux. Le troupeau est souvent l'échelle à laquelle le statut vis à vis de l'infection est surveillé au sein des programmes. Il existe donc des données de résultats de tests à l'échelle troupeau pouvant permettre d'estimer la probabilité qu'un troupeau soit indemne. Cependant, la stratégie de test (fréquence, type de test, animaux testés) peut faire varier la proportion de faux

négatifs, c'est-à-dire la probabilité qu'un troupeau soit classé *indemne d'infection* par erreur. Développer des méthodes pour permettre d'estimer l'incertitude autour du statut *indemne d'infection* en prenant en compte les différentes modalités des programmes de surveillance et leur contexte, est donc un enjeu majeur pour réduire le risque d'introduction via le commerce d'animaux.

# 3 Contexte de la thèse

Ces travaux de thèse s'inscrivent dans le cadre du projet Européen STOC free [5] financé par l'EFSA [6] ayant pour objectif de développer et de valider une méthode pour mettre en place une approche de surveillance *output-based* et permettre une comparaison de l'incertitude autour du statut *indemne d'infection* entre différents territoires. Ce projet réunit des chercheurs de 6 pays Européens différents : l'Allemagne, la France [7], l'Ecosse, l'Irlande, les Pays-Bas et la Suède.

# 4 Objectifs de la thèse

Les objectifs de ces travaux de thèse sont de contribuer au développement et d'évaluer une méthode statistique permettant d'estimer une probabilité (d'absence) d'infection à l'échelle du troupeau, à partir de données longitudinales issues de programmes de surveillance. Cette méthode statistique permet d'intégrer une grande diversité d'informations : séries de résultats de tests, caractéristiques du test utilisé, informations sur la dynamique d'infection, notamment les facteurs de risque d'introduction de l'infection.

# 5 Stratégie de la thèse

Afin de répondre à ces objectifs, les travaux de cette thèse s'articulent en 4 chapitres, suivis d'une discussion générale.

Le deuxième chapitre vise à identifier et organiser les informations disponibles pouvant être utilisées dans une approche de surveillance *output-based*. En se basant sur l'exemple

5. Surveillance analysis Tool for Outcome-based Comparison of the confidence of FREEdom (STOC free). Pour plus d'information : `https://www.stocfree.eu`
6. European Food Safety Authority
7. En France, c'est l'UMR BIOEPAR, qui est partenaire du projet.

de l'infection des bovins par le virus de la diarrhée virale bovine (BVDV), cette partie présente dans un premier temps un modèle conceptuel qui représente le processus dynamique de l'infection et y relie les différents types d'informations potentiellement disponibles. Dans un second temps, la disponibilité des informations d'intérêt et la possibilité d'extraire certaines informations de la littérature y sont explorées.

Le troisième chapitre décrit le modèle statistique développé dans le projet STOC free. Ce modèle permet d'intégrer les différents types d'informations identifiés dans le premier chapitre.

Le quatrième chapitre évalue les performances du modèle statistique dans différents contextes, pouvant représenter différents programmes de maîtrise, à partir de données simulées. Les simulations représentent différents scénarios en termes de caractéristiques de test, de dynamiques d'infection, de fréquence et d'associations entre facteurs de risque et probabilité de nouvelle infection. L'utilisation des données simulées permet de connaitre le statut vrai de chaque troupeau et de quantifier les performances du modèle à détecter les troupeaux infectés.

Le cinquième chapitre, présente une méthode pour catégoriser les troupeaux en *indemne/infecté* à partir des distributions de probabilités d'infection prédites par le modèle statistique. Il étudie notamment en quoi la manière de résumer cette distribution postérieure et le choix du seuil de catégorisation peuvent être influencés par certains éléments de contexte.

Enfin dans une discussion plus générale, nous discuterons de l'ensemble des résultats obtenus.

Chapitre 2

# Description and organisation of available data for the estimation of freedom from infection - Application on BVDV control

---

Ce chapitre se base sur différents documents rédigés auxquels j'ai contribué. Une sélection des principaux éléments utiles pour l'approche développée dans ma thèse est présentée dans le chapitre. Ces documents sont deux rapports techniques à destination de l'EFSA (Mercat *et al.*, 2018b,a), organisation ayant financé le projet, et une revue de la littérature publiée dans une revue scientifique internationale (van Roon *et al.*, 2020a) disponible en annexe de cette thèse.

Dans Mercat *et al.* (2018b), j'ai conçu le plan du rapport, recueilli les informations bibliographiques sur la diarrhée virale bovine utilisées dans le rapport, rédigé et corrigé le rapport en collaboration avec les co-auteurs.

Dans Mercat *et al.* (2018a), j'ai contribué à la conception du plan du rapport, à l'élaboration d'un fichier de recueil de donnée et a son évaluation ainsi qu'à la rédaction du rapport en collaboration avec Annika van Roon.

Dans van Roon *et al.* (2020a), j'ai participé à la conception du protocole, à la recherche et l'évaluation de l'éligibilité des articles, à l'extraction des données, à l'analyse des résultats, à la rédaction et la révision du manuscrit.

To develop an output-based surveillance method, data that can be used in the estimation of a herd probability of being infected, have to be identified. Such data are heterogeneous and can give information about the herd status against infection in different ways. It is important to list and to understand how such information indicate herds infectious status. Moreover, data diversity must be appreciated as well as their availability and sources.

Herd infection is a continuous dynamic process that is never directly continuously observed. Herd status against infection can change over time at any time, because of pathogen introduction into a naive herd, pathogen spread and disease effects on infected animals, pathogen persistence or clearance from the herd. It is not possible to directly observe this continuous process as it would require to monitor the status of each animal very frequently using a perfect test, which is not possible in the field. To monitor herd status, imperfect diagnostic tests are repeated over time. Diagnostic test results are observations of the system, determined by the herd status against infection and test performances. Testing modalities and frequency can vary between CPs impacting the performance of the monitoring scheme. Therefore, the surveillance of herd status, be it input-based or output-based, rely on discrete and imperfect observations of a continuous process which can be complemented with other information.

Risk factors to be infected can bring additional information about the probability of a herd to be infected. RFs influence the herd status against infection and indicate the risk of being infected. RFs at herd level can influence either the risk of introduction into the herd or the risk of circulation within the herds. RFs are disease specific, and related to the pathogen and its route of transmission (direct or indirect). Information on disease-specific RFs could be available in CPs, given that action on RFs is used to prevent the introduction and the spread of infections (Lindberg & Houe, 2005).

The continuous process over time and the discrete information related to a given time constitute a complex system. RFs and test results inform differently on the herd status against infection. Both information types refer to a given time. Test result and some RFs, purchase for example, are one-off events giving information at a specific date. Some RFs can have an impact over a period, as neighborhood related risk for example.

Infection with BVDV has been chosen as disease case study because it is a widely spread endemic disease (Scharnböck *et al.*, 2018), for which many CPs are implemented. BVDV infection can lead to significant economic losses, especially when the virus is introduced in a naive herd. Moreover, infected and infectious animals can be asymptomatic,

which represent a high risk for disease introduction through purchase. To get around this issue, many CP have been developed against BVDV infections, which has led to many definitions of *free from infection* herd status (van Roon *et al.*, 2020b).

The objective of this chapter is to determine the different types of information on the infectious process that can be integrated to estimate a probability of freedom from infection. We first built a conceptual model that represents the course of infection at different levels as a dynamic system, allowing to map the different types of discrete observations providing information about the state of the system onto the continuous dynamic of infection. Then, we appreciated the availability and the sources of discrete information through (i) a questionnaire on availability of data of interest within the partner countries of the STOC free project and (ii) a literature review on RF of BVD.

# 1 BVD conceptual model mapping discrete information onto the infection process at different levels

## 1.1 Aims and strategy

The aim of the conceptual model is to represent the infection dynamics in a simplified way and to establish the link between the underlying dynamic process of the biological system and the information issued from observation of this system.

Three levels of epidemiological unit are considered : animal, herd and territory. The animal level represents the level at which the infectious process in the host occurs. The herd level takes into account the transmission of the pathogen between hosts and events that can occur at this scale which can impact the dynamic of the disease. The herd is also often the level at which infection is monitored within CP. The territory level is the level of application of a CP. For each level, the first step of the conceptual model is the representation of the biological features of infection. The representation has to include the biology of the disease, the dynamics of infection and transmission, and the characteristics of the pathogen of interest. This work requires a good overview of the host population, the infectious agent, the disease and its specificity (Victora *et al.*, 1997).

Status against infection are defined at each level to represent the dynamics of infection. Status of interest depend on the disease characteristics. The simplest model includes a susceptible and an infected status for units which have not been exposed to the infection and units in which the pathogen can be detected, respectively. Other status can include

removed and immune for units that cleared infection and are protected against a new infection.

Finally, discrete information are listed and connected to the dynamics of infection. They can take the form of aggregate information like prevalence for a territory, test at herd level or on a group of animals.

## 1.2 Conceptual model for BVD

### 1.2.1 Animal level

#### 1.2.1.1 Epidemiological statuses

We consider 4 main statuses against BVDV infection at the animal level : susceptible (S), transiently infected (TI), immune (R) and persistently infected individuals (PI) (Figure 2.1).

**Susceptible animals (S).** Susceptible animals have not been infected with BVDV and have not developed antibodies. Hence, they are naïve (not immune) and can get infected.

**Transiently infected animals (TI).** TIs are susceptible animals which get infected by BVDV and develop a transient infection. A transient viremia starts approximately 3 days after the infection (Pedrera *et al.*, 2012) until immunity develops around 2 weeks later (Meyling *et al.*, 1990). Compared to PIs, TIs shed lower amounts of virus(Niskanen & Lindberg, 2003).

**Recovered animals (R).** Recovered are TIs that have cleared infection and developed an immunity. Animals are assumed to remain immune for the rest of their lives.

**Persistently infected animals (PI).** PIs are animals infected in utero, while their immune system is immature (McClurkin *et al.*, 1984). As a consequence, they are immunotolerant (they do not produce antibodies against homologous virus) and become persistently infected. PI animals shed large amounts of virus throughout their lives (Brownlie *et al.*, 1987).

Two additional animal statuses can be identified : vaccinated animals (V) and animals under maternal immunity conferred by transfer of antibodies (M). In both statuses animals show a transient immunity which does not reflect a direct contact with the virus. V and M animals can have positive results for tests targeting antibodies.

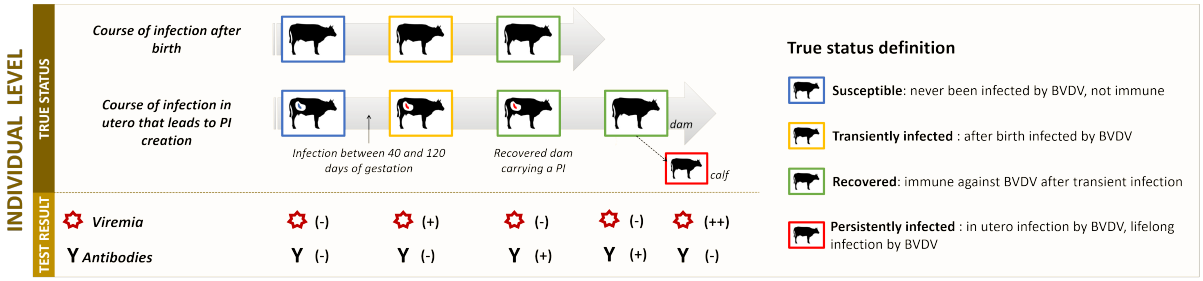



FIGURE 2.1 – Representation of course of infection at animal level and test results according to animal level status. Viremia represents test targeting virus presence (ACE or RT-PCR) while Antibodies represents tests targeting the immune response (Ab ELISA).

#### 1.2.1.2 Course of infection

BVDV transmission occurs from different sources and through different routes of infection. There are two types of BVDV infections : infection after birth (i.e. horizontal) and in utero infection (i.e. vertical which result in a different course of infection).

**Course of infection after birth**

Susceptible animals that are infected after birth become TIs. After immunity has developed, around two weeks after, they become recovered (Figure 2.1).

**Course of infection in utero**

Susceptible pregnant cows can get infected during pregnancy. The dam becomes TI. The virus multiplies in the dam and can infect the foetus. The impact of the infection on the foetus depends on the stage of gestation. Between 30 and 120 days of gestation before the development of the immune system in the foetus, infection can lead to the birth of PI calves (Brownlie *et al.*, 1998) (Figure 2.1). Infection before or after the susceptible window for PI creation can lead to embryonic death in early pregnancy, to no effect, teratogenic effects, foetal death or abortion at later stages (Moennig & Liess, 1995). Foetuses that are immunologically competent at the time of infection can be born either TI or immune. A calf born to a PI cow is always a PI.

#### 1.2.1.3 Information on animal status against infection

**Diagnostics test**

Available tests can be divided into two groups : (i) tests that detect an on-going infection through the detection of the virus or viral antigens (Ag) and (ii) tests that detect the

immune response against the virus through the detection of circulating antibodies (Ab). Both are used for the diagnosis of BVDV infection. However, detection of Ag indicates a current infection with BVDV while detection of Ab indicates a past infection.

To monitor BVDV infection within CP, three tests are mainly used : (i) Ab detection using Antibody enzyme linked immunosorbent assay (Ab ELISA), (ii) Ag detection using Antigen capture ELISA (ACE) and (iii) virus genome detection using Reverse transcriptase polymerase reaction (RT-PCR).

**- Detection of virus**

**ACE.** Antigen-capture ELISA tests (ACE) detect infected animals that shed the virus : TIs and PIs. TIs can be challenging to detect as they shed low amounts of virus during a short time period (Hanon *et al.*, 2014). However, once an Ag ELISA returns a positive result, interpretation of the state of animal is TI or PI without any other information (Hanon *et al.*, 2014). Repeating the test two weeks/ 1 month later can clarify whether the animal is TI, if the second test is negative, or PI, if the second test is positive.

**RT-qPCR.** Reverse transcriptase polymerase reaction (RT-PCR) is able to detect infected animals (TI and PI) by targeting viral RNA. Quantitative RT-PCR (qRT-PCR) has been developed for BVDV diagnosis, as there exists a relationship between threshold cycle (CT), cycle number at which the fluorescence generated is higher than the threshold, and the quantity of viral RNA present (Bhudevi & Weinstock, 2001). qRT-PCR can be used to make a distinction between TI and PI in term of CT, knowing that PI sheds a larger quantity of virus. However, like ACE a second test can be used to confirm the diagnostic of PI.

**- Detection of specific antibody**

**Ab ELISA.** Antibody enzyme-linked immunosorbent assay (Ab ELISA) is an immune-enzymatic technique that allows the detection of specific antibodies. A positive Ab-ELISA can be associated with either an immune state resulting from a natural infection, the presence of maternal antibodies in calves under 6 months or with vaccination. A single test result may not be able to distinguish those three categories, in the absence of additional information. However, repeated testing can clarify the true BVD status in that maternal and vaccination derived antibodies decrease in time.

### 1.2.2 Herd level

#### 1.2.2.1 Herd statuses against infection

We consider four different statuses at the herd level depending on the situation of the herd regarding BVDV infection.

**Virus free and seronegative herds.** Naïve free herds are herds that are not currently infected and that have not been recently (in the past +/- 10 years) infected by BVDV. They are composed of susceptible cattle that are not immune against BVDV.

**Herd infected with at least one TI animal.** Herds in this status are infected by at least one transiently infected animal. They are composed of S and TI animals and as the herd infection progresses the proportion of S and TI decline and R cattle arise. No PI are present, either alive or in the fetal stage.

**Herd infected with at least one PI animal.** Persistently infected herds contain at least one PI animal alive or to be born (the dam of a PI fetus is called "Trojan cow",). They are composed of S, TI and at least one PI animal and as the herd infection progresses by an increasing number of immune cattle (R).

**Virus free and partly seropositive herd with at least one seropositive animal.** Herds can be in this status : (i) when all infectious animals (PI, TI) are removed from the herd (by death, sale, conversion to recovered animals) and there are still animals with antibodies (R) present ; (ii) after vaccination campaign of a part or the whole herd ; (iii) by a combination of both. Herds in this state can become *virus free and seronegative herd* once all the immune and vaccinated animals have left the herd.

#### 1.2.2.2 Course of infection

The course of infection within a herd starts with the introduction of BVDV (Figure 2.2. Different routes can lead to BVDV introduction within a herds (see **Risk factor for BVDV introduction**). Once an animal is infected it sheds the virus and infects other susceptible animals. Newly infected animals in turn infect other susceptible animals. The proportion of cattle newly infected within the herd can vary, depending on the mean of introduction and factors that can influence the spread within the herd (see **Risk factor for BVDV spread within herd**). If some infected cows are pregnant between they can give birth to a PI calf in the herd. As a PI sheds a high level of virus (Brownlie *et al.*, 1987) throughout its entire life, spread in the herd can occur quickly, continuing whilst the PI animal remains in the herd. If nothing is done to limit the infection, the virus can

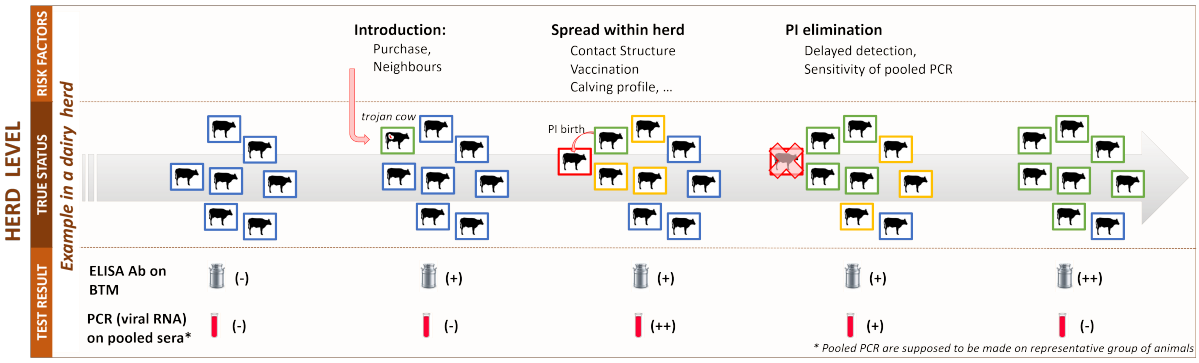



FIGURE 2.2 – Representation of herd level status, courses of infection and example of test result. For the seek of representation herd test results are represented when herd status change occurred. In real CP test are done at regular time interval which can occur before, during or after a status modification. Blue squared represents susceptible animals, yellow squared represents TIs animals, green squared immune animals and red squared PIs animals. Viremia represents test targeting virus presence (ACE or RT-PCR) while Antibodies represents tests targeting the immune response (Ab ELISA).

continue to spread within the herd with a negative impact on reproduction (e.g. abortion). After a while, a large proportion of cows within the herd become immune.

Two types of risk factors that can influence the course of infection at the herd level can be considered : (i) risk factor for introduction of BVDV in a herd ; and (ii) risk factor for virus spread within the herd once BVDV has been introduced. Depending on the biosecurity measure in place, the relative importance of the different RFs can vary in time.

**Risk factor for BVDV introduction**

**Introduction of cattle.** Introduction of infected animals, which can be either PIs, TIs or Trojan cows, in the herd are an important source of introduction of disease (van Roon *et al.*, 2020b). Introduction of PIs animals, that shed a large amount of virus, are the main source of introduction of BVDV in a free herd in the absence of control measure. Buying pregnant cows can be at risk of introduction as it can be Trojan cows (Reardon *et al.*, 2018). TIs animals can also be a source of introduction of BVDV into a herd. The relative importance of TIs in (re-) introduction of BVDV in a herd is under discussion : some argue that TIs are unlikely to be a source of infection (Niskanen *et al.*, 2002; Sarrazin *et al.*, 2014) while others suggest that BVDV can be maintained in a herd without presence (or at least identification) of PIs (Moennig *et al.*, 2005).

**Neighbourhood Risk.** Direct nose-to-nose contact with infected cattle from another herd is also a possible means of introduction of BVDV (Lindberg & Alenius, 1999). Such contacts can occur through shared grazing or adjacent herd pasturing areas. It is especially at risk when the infected cattle come in contact with susceptible cattle in early pregnancy, leading to a risk of producing a PI.

**Farm Management and Biosecurity.** Farm management and biosecurity measures influence the risk of introduction by indirect transmission. Introduction can occur through contaminated persons, when they have contact with animals (e.g. veterinarian, farmers, claw cutters, inseminators) or contaminated products or materials (Meyling & Jensen, 1988; Rikula *et al.*, 2008; Gunn, 1993; Niskanen & Lindberg, 2003) .Participation to cattle shows or market can also lead to introduction of BVDV (Lindberg & Alenius, 1999). Compared to direct contact with infected cattle, indirect routes may play a minor role in transmission. However, towards or at the end of an eradication programme, when introduction of BVDV through cattle introduction and neighbors is controlled, indirect transmission can become relatively more important (Hult & Lindberg, 2005).

**Risk factor for BVDV spread within herd**

**Herd contact structure.** In cattle herds, animals can live in separate groups which can have more or less contact depending on the type and structure of the herd. In dairy herds, calves and dams are quickly separated and there are usually groups of calves, heifers and lactating cows. In beef herds, calves stay with their dam until weaning at up to 9 months. This results in PI calves being in close contact with the breeding herd for much longer in beef than in dairy herds. From a herd to another the separation between groups can be quite different, groups can be or not kept apart in different barns or on different pastures.

**Farm Management.** Some farm management practices are of major importance in the dynamics of BVD. The calving distribution can be either seasonal or not. Seasonal calving, i.e. all calvings grouped in a short period (e.g. 3 months), is associated with most pregnant cows being in the window of susceptibility for the formation of a PI calf at the same time. Conversely, year-round calving and breeding means that a PI born at any time of the year may have the opportunity to be in contact with a pregnant cow in the window of susceptibility. The replacement rate, which determines the proportion of female calves born on the farm that are kept to replace breeding cows, can also impact the within-herd spread. The lower the replacement rate, the higher the probability that a present PI calf is sold rather than kept as replacement. Vaccination can modify the course of infection within a herd by limiting the production of PIs.

#### 1.2.2.3 Information on herd statuses against infection

At the herd level, two types of information can be available to evaluate the herd status against infection :

— Information related to RFs influencing the herd status against infection

— Information related to the imperfect observation of the herd status against infection using diagnostic test.

**Information related to RF**

Presence of risk factors for a herd can be estimated with different variables (Table 2.1).

TABLE 2.1 – Examples of variables that can be used to estimate the different risk factors impact of introduction and spread of BVDV within a a herd from van Roon *et al.* (2020a)

| Risk factors | Example of variables |
|---|---|
| Risk factors for BVDV introduction | |
| Introduction of cattle | Purchase (Yes/No), Number of purchase, Age of animal purchase, Participation to cattle shows and market (Yes/No) |
| Neighbourhood | Pasture (Yes/No), BVD-positive neighbour herds (prevalence within an area), Cattle density (herd/$km^2$), Contact with other animal species (Yes/No) |
| Farm management and biosecurity | People on farm, Shared equipment(Yes/No), Housing, Vaccination (Yes/No), Mixed beef and dairy cattle (Yes/No), Quarantine (Yes/No) |
| Risk factors for BVDV spread within a herd | |
| Herd contact structure | Size of the herd, Age at which calves are separated from their dam, Age at first calving, Replacement rate |
| Herd management | Calving distribution, Biosecurity, Vaccination (Yes/No) |

**Information related to diagnostics**

**Detection of BVDV in a herd.** Monitoring the herd status can rely on testing individual animals or groups of animals. Depending on the context, like prevalence, type of herds and practices, and the objective of the CP different testing strategies can be chosen. Table 2.2 presents the main screening strategies used in CP for BVD and their limitations.

**Delayed detection of BVDV in a herd.** Delayed detection corresponds to the time interval between the introduction of BVDV in a herd and its detection. Delayed detection is a key concern for surveillance as a herd can keep its status "free from infection" while

being infected during this delay. Delayed detection can vary between CP as it results from the design of the testing procedure and the test performance. Features of the CP design influencing delayed detection are the time interval between tests, the target animals for serial sampling and the type of test used. Test performance can contribute to delayed detection are diagnostic are never perfect. In general, test sensitivity is quite high for BVDV diagnostic tests at the animal level. However, test sensitivity at the herd level is not the same as at the animal level, because it depends on the sampling scheme in the herd. Herd level sensitivity is the capacity of the test to detect an infected herd. For example, the sensitivity of an ELISA test on BTM (Bulk Tank Milk) to detect antibody in the lactating dairy cow herd is considered high (Beaudeau *et al.*, 2001b) but the sensitivity of this type of test to detect a herd infected by a PI at a given time can be much lower, and positive results can occur several months after introduction of the virus in the herd (Ducrot *et al.*, 2010).

### 1.2.3 Territory level

#### 1.2.3.1 Territory statuses against infection

A territory is here defined as an area where herds follow the same control programme and where information is available. Territory can be either a region or a country. As an example, within the STOCfree consortium, BVDV CP for the Netherlands, Sweden, Ireland and Scotland, are applied at country level ; while for Germany and France, CPs are applying respectively at Federal States and at region or department levels. Each territory has a BVDV CP which can be based on different components (e.g. different component for dairy and beef herds).

**Infection free and seronegative territory.** An infection-free territory is defined as a territory composed of seronegative herds that are currently not infected by BVDV and where all cattle are susceptible.

**Territory with infected herds.** An infected territory is defined as a territory with at least one infected herd(s) meaning that the infection is present or spreading within the territory. In this defined territory, herds can be naïve and infection free, currently infected or seropositive (some or all animals). The proportion of herds in each state depends on the prevalence of BVDV infection and the control measure in place (endemic territory versus on-going eradication programme). Over time and depending on the contact between herds within the territory and the actions taken to trace and eradicate infected animals these

TABLE 2.2 – Diagnostic strategy mainly used and it interpretation within the surveillance component of a BVDV CP.

| Test level | Test | Sample | Animals sampled | Usual frequency | Interpretation of one positive results [1] | Limits |
|---|---|---|---|---|---|---|
| Herd | Ab ELISA | BTM | Dairy cows | Twice a year | Actual, recent or old infection with BVDV (infected or not infected with immune cattle herd) | Interpretation of one single test result is not possible<br>Interpretation depends on herd test results history<br>Introduced immune cattle participating to BTM can not be exclude of the test |
| Herd | Ab ELISA | Pooled serum | Young cattle | Once a year | Recent or Actual infection with BVDV | Selection of animal to test<br>Every year testing increase delayed detection risk |
| Animal | PCR | Ear notch | New born calves | As birth arises | Presence of infected animal within the herd (PI or TI) (infected herds) | Delay between introduction and a birth of a PI<br>Interpretation at herd level : one positive results classifies a herd as *infected*<br>Many negative test (number of calves tested or over period) are needed to classify a herd as *not infected* |

[1] For each test there is a possibility to be a false postive results depending on specificity of the test.

proportions can change.

**Infection free and seropositive territory (post-eradication territory).** A post-eradication territory does not have any infected herds but is composed of seropositive and seronegative herds.

#### 1.2.3.2 Course of infection

In a territory where infection has been eradicated, the course of infection starts with the re-introduction of BVDV, which can occur through different routes. Once infection is present, the proportion of infected herds vary depending on factors influencing the spread of BVDV between herds. Two types of RF influence the course of infection at the territory level : (i) RFs for introduction of BVDV in a territory and (ii) RFs for the spread of BVDV in the territory.

**Risk factor for BVDV introduction**

**Cattle movement.** As for herd level, cattle movement through purchase and market outside of a territory can be sources of (re-) introduction of BVDV into a territory. As PI animals are the main source of (re-)introduction of BVDV, purchasing young animals or pregnant dams (with a chance of being a Trojan cow) is particularly risky.

**Infection prevalence in neighboring territories.** Infection prevalence in neighboring territories can also be a risk factor for introduction of BVD within territories, when cattle are moved to/through or grazed in the neighboring territory.

**Wildlife (reservoir).** BVDV have been reported for over 40 different species, including domestic and wildlife species (Nelson *et al.*, 2016) and can induce persistent infections in 8 other species than cattle(Terpstra & Wensvoort, 1997; Scherer *et al.*, 2001; Duncan *et al.*, 2008; Nelson *et al.*, 2008; Bachofen *et al.*, 2013). Sources of infection for non-bovine species can be a spillover from cattle population by sharing environment or through direct contact (Nelson *et al.*, 2016). Nevertheless, infection through wildlife is not considered a major cause of introduction.

**Risk factor for spread within territory**

Important territory characteristics that can vary from one territory to another and influence BVDV spread once BVDV has been introduced. The proportion of beef and dairy herds can modify the spread as their practices differ. Herd density and degrees of fragmentation of farms may influence the contact structure and potential contact between herds within the territory. The intensity of these contacts between herds can influence

the spread within the territory. The purchasing dynamics within the territory can also have an impact. A high number of exchanges between herds increases the potential transmission between herds. Finally, a high prevalence in the territory increases the potential transmission between herds either through purchases or contacts at the pastures.

#### 1.2.3.3 Information on territory statuses against infection

At territory level, data on two types of information can be available to evaluate territory level state against infection :

— Information related to RFs influencing the status of the territory against infection
— Information derived from CP informing on both RF and territory status against infection

**Information on RF linked to territory structure**

RF for BVD introduction and spread can be approximate using territory structure variables, such as :

— the number of herd within the territory (No. of herds)
— the density of herds within the territory (No. of herds/km$^2$)
— the proportion of dairy and beef herds (%)
— the infection prevalence of neighboring territories (prevalence)
— the number of cattle purchased from outside of the territory and their source (No. of cattle, No. of sources)
— the participation in market/trade shows either inside or outside of the territory with participant from everywhere (yes/no questions, No. of cattle show attended)

**Information related to CP**

Information derived from CP Status observation at territory level is derived from aggregation of observations at herd level. Information about the presence or absence of infection within the territory can be derived from CP, like an approximation of prevalence. Delayed detection of BVDV at territory level is the time period between re-introduction of BVDV and it detection and depends on the number of herd participated in the CP.

# 2 Availability and sources of information to estimate a probability of infection at herd level.

## 2.1 Introduction

The conceptual model has identified information that can be used to estimate the herd status against infection. It explores the different nature of the information which are (i) related to the monitoring of status that varies from a CP to another, and (ii) RFs that may influence the status against infection that varies from a herd or a territory to another. In both cases, the conceptual model identifies potential variables that can be used on field.

To include RF related information in an output-based surveillance method, the strength of association between the RF occurrence and infection is required as well as the data on presence or absence of the RF.. RF studies enable to estimate these values. Specific RF studies are not available for all territories. In this case, meta-analysis can be a source of information to appreciate the variation of the strength of RFs between territories.

It is necessary to assess the availability of information from different sources before considering its inclusion in an output-based surveillance method. To appreciate the availability of data of interest, we first used a questionnaire within 6 European countries, asking for the availability and sources of identified variables within the conceptual model. Then, we described the type of information on RF and its limitation, that can be derived from a literature review on BVD risk factor (van Roon *et al.*, 2020a).

## 2.2 Material and Methods

### 2.2.1 Inventory of data availability

The questionnaire was sent to the six countries partner of the STOCfree project (France, Germany, Ireland, the Netherlands, Scotland and Sweden). Each country filled one questionnaire corresponding to one CP (national or regional). The questionnaire asked for the availability of data, whether quantitative or qualitative, their sources and their potential strength and limitations for a given territory. The template of the questionnaire is available in appendix. Data are mostly given at herd level, but some are aggregated at the territory level. The questionnaire includes :

— Demographical related data, which describe cattle herds demographics data within

the territory for the most recent full calendar year.
— Control programme related data, which describe data from the monitoring of herds status by year, and data availability of the previous 5 years.
— Herd management related data, which describe management data related to RF of introduction and spread.

Data availability were requested for all cattle, dairy, non-dairy and a subsequently relevant subset of non-dairy, beef breeding. For the sake of representation, only the availability of data for all cattle are presented is the results.

### 2.2.2 Information derived from literature review

The systematic literature review of (van Roon *et al.*, 2020a) assessed the importance of most frequently studied RF for BVD, and, depending on study quality and the availability of quantitative data, performed a meta-analysis. We first explored the possibility to refine the selection of RF and to export quantitative data that can be used within an output-based surveillance method. In a second step, we focused on the neighborhood risk factor at the herd level, defined has the risk of introduction of BVDV in a herd through direct contact at pasture. This RF relies on both the presence of infected animals in neighbor herds, and on the possibility to have direct nose to nose contact at pasture. This RF is difficult to observe directly, because of many unknown parameters (fields proximity, animal release times) and many proxies have been used.

## 2.3 Results

### 2.3.1 Inventory of availability of data within 6 European countries

The availability of data within the six territories varies widely depending on the type of data (Figure 2.3). The majority of demographical related data was available in most countries. The majority of data related to CP were available in most countries. Data related to RF were very sparse, and varied between RF. Data related to RF of introduction through purchasing were mainly available. Data related to the neighboring were mainly not available. None of the data related to herd management (breeding, housing) and biosecurity were available in any territory.

Sources of the data varied from a type of data to another (results not shown). Demographical related data were available from national database, as were purchase data. CP

TABLE 2.3 – Pooled estimates odds for risk factors and number of studies include in the meta-analysis reproduced from (van Roon *et al.*, 2020a)

| Risk factor | Factors | No. of studies | OR[1] |
|---|---|---|---|
| Herd caracteristics | Herd type | 6 | 1.63 [1.06 - 2.50] |
| | Herd size | 3 | 1.004 [1.002-1.006] |
| Cattle movement | Participation in cattle shows or markets | 5 | 1.45 [1.10 - 1.91] |
| | Introduction of cattles | 8 | 1.41 [1.18 - 1.69] |
| Neighborhoods | Pasture | 3 | 1.10 [0.62 - 1.87] |
| | Contact between cattle at pasture | 6 | 1.32 [1.07 - 1.63] |

[1] Odd-ratio

related data were mostly available for specific CP database. Sources of RF related data, other than purchases, varied a lot from national database to literature or expert opinion.

### 2.3.2 Availability of data from literature review

#### 2.3.2.1 Selection of RF of interest

In van Roon *et al.* (2020a), we have identified 5 groups of RF studied for BVDV infection : herd and animal characteristics, cattle movement, reproduction, neighborhood and farm management and biosecurity. RFs related to reproduction and farm management and biosecurity were studied in a very low number of studies or variables used were not comparable enough to perform a meta-analysis. However, some of the variables related to these RF had a significant impact. As an example, in one study, herds with a year-round calving pattern had a significantly higher risk to become infected than herds with a seasonal calving (Williams & Winden, 2014). In contrast most variables related to farm management and biosecurity measures were found not significant in the different studies.

#### 2.3.2.2 Estimates from literature review

Only some variables related to herd and animal characteristics, cattle movements and neighboring RFs were candidate to meta-analysis (Table 2.3). For these variables, a wide variability has been observed.

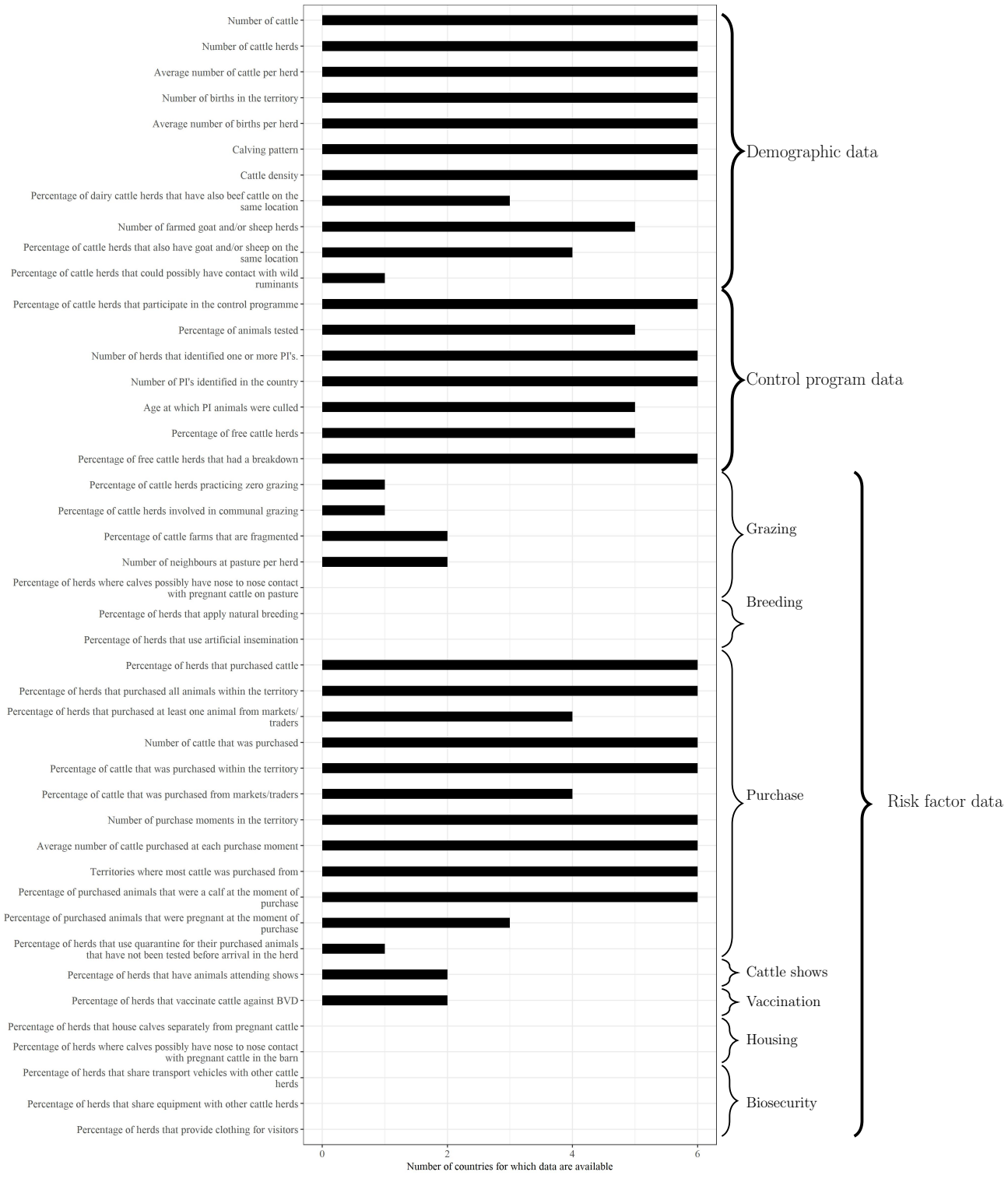



FIGURE 2.3 – Availability of data related to the risk of BVDV infection for all cattle herds within six European territories.

TABLE 2.4 – Overview of the number of the risk factor studies that included neighborhood variables and the availability of quantitative data, reproduced from van Roon *et al.* (2020a)

| Factor | No. of studies | No. of variables | No. of variables with quantitatives data |
|---|---|---|---|
| Farm fragmentation | 1 | 1 | 1 |
| Environnement | 1 | 4 | 4 |
| Cattle density | 6 | 9 | 7 |
| BVD-positive neighbor herds | 3 | 11 | 8 |
| Contact with other animal species | 5 | 10 | 8 |
| Pasture | 8 | 20 | 14 |

#### 2.3.2.3 Selection of variables of interest, using Neighborhood RF as an example.

Six different factors related to neighbor RF haves been studied in the review in 12 different papers (Table 2.4). Variables describing pasture were included most frequently. Most of the variable were built for the specific purpose of the study with different definition.

Charoenlarp *et al.* (2018), was the only study to look at environment related variables which described the land type present around herds (like natural grassland or forests), obtained using ArcGIS software. Graham *et al.* (2013) was the only study looking at farm fragmentation related variables, i.e. the number of individual non-contiguous parcels of land associated with the herd, which was available from a specific national database.

Three studies looked at BVD-positive neighbor factors which requires to defined both neighbor and BVD-positive. Such definitions vary a lot between studies (Graham *et al.*, 2016; Charoenlarp *et al.*, 2018; Ersbøll *et al.*, 2010).

Pasture variables were studied enough and in a comparable way to do meta-analysis. However, different aspect of pasture was studied : whether cattle had access to pasture, the possibility of contact with cattle from others herds at pasture and the use of shared pasture (van Roon *et al.*, 2020a). Such variables were mainly derived from specific questionnaire built for the specific purpose of the study (Amelung *et al.*, 2018; Gates *et al.*, 2013; Hanon *et al.*, 2018; Houe *et al.*, 1995; Valle *et al.*, 1999), and less frequently from national database (Presi *et al.*, 2011; Charoenlarp *et al.*, 2018). The meta-analysis found a not significant combined effect for BVDV infection for herds that had access to pasture compared with

herds that do not (Table 2.3). A significant effect indicating that contact at pasture had a higher odds of BVD infection was founded, only when variables indicated shared pasture and whether contact between cattle at pasture could occur were taken together (Table 2.3).

## 2.4 Discussion

Demographical and purchase related data were mainly available because they are related to mandatory animal identification and traceability . EU legislation details the rules for the identification and registration of bovine animals (Regulation (EC) 1760/2000),. It allows for the traceability and the localization of the animals from birth to death. Demographical characteristics and purchase related variables have been identified as significant RF for BVDV (herd size, herd type and purchase) (van Roon *et al.*, 2020a) and pooled estimates are available and can be used as inputs for an output-based method. However, variability of odds among territories highlights that the RF strength of association must be re-estimated specifically for each territory.

CP related data were mainly available in every territory. The herd status against infection defined within CP is often based on consecutive tests results (van Roon *et al.*, 2020b; Bernard, 2018), for which test dates and results have to be stored over periods. . Thus, longitudinal data is available, and such historic must be considered in an output-based method as well as knowledge of test characteristics to take into account potential mistake.

Data related to neighborhood RF were available in a limited number of territories (in 1 or 2). Variables requested within the questionnaire required to know both the fragmentation of the fields for each herd, and the period of grazing in each field during the year. If some of these variables are available from specific database (for Sweden and Ireland), such data are mainly not available and can be at the best, approximated. Pooled estimate for pasture was not significant (van Roon *et al.*, 2020a), which can be explain by the fact that risk of grazing is influenced by many factors like the prevalence of BVD within territory and cattle density (Houe *et al.*, 1995). Cattle density and BVD-positive neighbor herds, has been studied in many study but in such different way that no comparison was possible. However, such variables can be built from available data : cattle density can be estimated from demographic related data and neighbor positive herd can be estimated using CP related data to include such RF in a output-based method.

Other data related to RF were less available, especially because they do not depend

on a specific legislation and depend on the goodwill and the means of farmers. Biosecurity and housing practices vary a lot from a herd to another, as well as from a territory to another. From the literature review, biosecurity factors were mainly not significant which can be largely explained by the way they are measured than in the fact that they do have no real impact. The questionnaire design often uses yes-no questions, answered by the farmer himself, forcing him to choose one option and he may want to give the socially acceptable answer. Vaccination policies that consider BVDV vary from a territory to another. Territory free from infection or within an eradication programme can banish the use of BVD vaccination, as it can modify the results of monitoring testing (Austria and Scandinavian countries).

Finally, despite the fact that a lot of data of interest exist and are stored in a standardized manner, their access can be limited to some specific organizations or companies. Condition of accessibility must be assessed beforehand.

Chapitre 3

# The STOC free model : assumptions and implementation

---

Ce chapitre se base sur des descriptions du modèle STOC free rédigées à différentes étapes de son développement. Un rapport technique préliminaire a été rédigé comme délivrable pour le projet STOC free (Madouasse *et al.*, 2019) et disponible en annexe. Une description plus récente et complète peut être trouvée dans un article qui est actuellement en révision (Madouasse *et al.*, 2021). Le travail décrit dans les prochains chapitres a contribué au développement et à l'évaluation du modèle.

# 1 Introduction

The aim of the STOC free model is to predict herd-level probabilities of (freedom from) infection at a given point in time given all the information available up to that time. Because of the important heterogeneity in disease CPs, the model needs to be able to take heterogeneous input data. However, the model also needs to have a structure that reflects what is common to all CPs against infectious diseases in order to be able to use these heterogeneous inputs to produce an output that is comparable. Therefore, the modelling framework must have a structure that reflects what is common to all CPs while including what differs as data and parameters.

Before the STOC free model, scenario tree modelling was developed to substantiate freedom from disease at the territory level while accounting for the fact that diagnostic tests are imperfect and performed on a sample of animals. The principle of scenario tree modelling is to estimate the probability of not having detected the infection with the surveillance programme in place if it were present at a certain prevalence, called the *design prevalence*. Therefore, this approach offers the possibility of quantifying the probability of being (almost) free from infection. The method allows the use of multiple and complex data sources (Martin *et al.*, 2007). Using this framework, it is possible to include both test results and relative risks of infection associated with risk factors. However, it is usually applied at the territory level, in territories where eradication is assumed. Applying the scenario tree method to territories where the infection is still present would require the prediction of probabilities of freedom of infection for all herds in the data in order to identify infected and uninfected herds.

Furthermore, in territories where the infection is still present, information from infected and newly infected herds could be used to predict probabilities of infection in the population. This is not possible with scenario trees, which rely entirely on stochastic simulations. With scenario trees, hypotheses regarding disease and test characteristics must be retrieved from the literature or from expert opinion. This method cannot learn from the data. A method able to learn from data collected in both infected and uninfected herds, i.e. to perform inference, could improve the prediction of herd-level probabilities of infection in contexts where the infection is still present.

From the conceptual model we identified information that can be used for the estimation of herd-level probabilities of infection. Two types of informations were identified : (i) information related to an increase of the probability of infection (risk factors of infection :

**RF**) and (ii) information related to status evaluation through diagnostic testing.

As part of the STOC project, we developed a Bayesian latent class model that predicts herd-level probabilities of infection using data from CPs and knowledge about disease dynamics and risk factors. The aims of this chapter are to describe the main hypotheses that led to the choice of the model developed in the STOC free project and to provide a brief description of the model and its implementation.

# 2 Modelling hypotheses

The hypotheses considered in designing the STOC free model are :

— The infection status of a herd has a small probability of changing over time.

— Infection modifies some biological parameters that can be measured using biological tests.

— In territories where the infection is still present and where there is a CP in place, herd statuses are evaluated at regular times using biological tests. These CPs generate herd level longitudinal data.

— Biological tests sometimes provide inaccurate results by being negative in infected units (lack of sensitivity) or being positive in uninfected units (lack of specificity).

— Information on the probability of changing status, i.e. infection dynamics, (incidence/ cure rates) is often available or can be estimated from CP data and can be incorporated into the prediction.

— Information on risk factors acting on infection dynamics, such as risk factors of new infection, is often available or can be estimated and can then be incorporated into the prediction.

These hypotheses, summarised in Figure 3.1, describe the system of interest and need be included in the model.

# 3 Description of the STOC free model

The STOC free model is a Hidden Markov Model (**HMM** ; See Zucchini *et al.* (2017) for an overview), in which herd statuses regarding infection are considered to be imperfectly observed Markov processes. This status is modelled in discrete time steps. The Markov property means that the status at a given time step only depends on the status at the previous time step. For a herd, the status transition between consecutive time steps

depends on two probabilities : the probability of becoming infected and the probability of remaining infected. The prediction of the probability of becoming infected can be improved by incorporating knowledge about RF for introduction of infection. Finally, test results are included as imperfect observations of the herds status, which makes this status latent. Test characteristics, sensitivity and specificity, define how good are the tests used at evaluating the latent status.

The model outcome is the predicted herd-level probability of infection on the last month where data are available. Data collected before the month of prediction are used for parameter estimation. Parameter estimation is carried out in a Bayesian framework. Bayesian parameter estimation allows the introduction of existing knowledge or hypotheses about the different parameters using prior distributions.

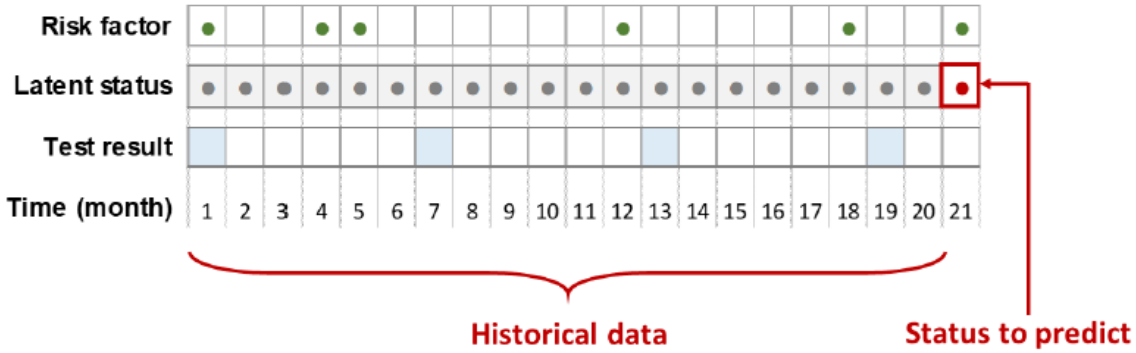



FIGURE 3.1 – Conceptual representation of the implementation of a surveillance programme within a herd. The focus of the model is the latent status regarding infection, which is modelled at the herd-month level. This status partly depends on risk factors and determines test results. In this diagram, risk factors are represented as green dots when present and available test results as blue shaded squares. The model predicts a herd-level probability of infection for the most recent month in the surveillance programme using all the data collected for the estimation of model parameters. From Madouasse *et al.* (2021).

## 3.1 Latent status dynamics

The status regarding infection is modelled in discrete time steps. In the STOC free project, when the model was applied to BVDV infections, it was important that all countries used the same time step in order for the model parameters to have the same unit, regardless of the CP. A monthly time step was used, which means that a status was mo-

delled for each month, regardless of test availability. The month was chosen to take into account the infection dynamics, to limit and ignore double state transition between time step.

Two different statuses regarding infection were considered : not infected (negative herd status at time $t : S_t = 0$) and infected (positive herd status at time $t : S_t = 1$). Within the STOC free model, the latent status of a herd at time $t$ follows a Bernoulli distribution :

$$S_t \sim \ Bernoulli(\pi_t)$$

with $\pi_t$ being the probability of being infected at time $t$.

### 3.1.1 Latent statuses at $t = 1$

At the first time step, no transition can be modelled because there is no previous status to transition from. Therefore, the probability of being infected must be modelled directly. The probability of being infected at time $t = 1$ ($\pi_{t=1}$) was modelled using a beta prior distribution, which represents the initial infection prevalence :

$$\pi_{t=1} \sim \ Beta(\alpha_\pi, \beta_\pi)$$

.

The parameters of this Beta distribution can be chosen so as to reflect the initial infection prevalence. However, when the test used has high sensitivity and specificity, the model can estimate this initial prevalence from the proportion of test positives, even with weakly informative priors.

### 3.1.2 Latent statuses between $t = 2$ and time of prediction

From the second time step onwards, the herd status at time $t$ is modelled as a function of the herd status at time $t - 1$. A herd that is not infected at time $t - 1$ can either (i) become infected at $t$ with a probability of new infection noted $\tau_1$ or (ii) remain uninfected with probability one minus the probability of new infection $(1-\tau_1)$. A herd that is infected at $t - 1$ can either (i) remain infected at $t$ with a probability of remaining infected noted $\tau_2$ or (ii) eliminate the infection with probability one minus the probability of remaining remain infected $(1 - \tau_2)$. Herd status transitions are represented in (Figure 3.2).

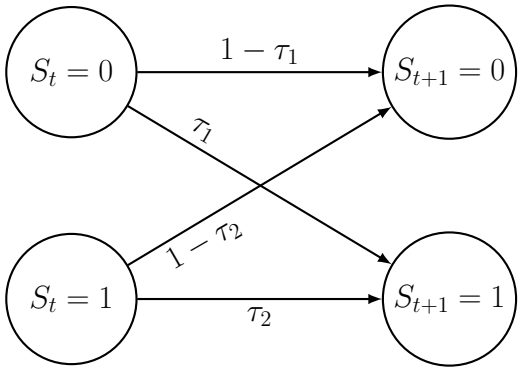



FIGURE 3.2 – Possible transitions in herd status between $t$ and $t + 1$, with $S = 1$ representing infection and $S = 0$ absence of infection. $\tau_1$ is the probability of new infection and $\tau_2$ is the probability of remaining infected.

Thus, the probability of infection at time $t$ ($\pi_t$) depends on both the latent status at time $t - 1$ and the transition probabilities :

$$\pi_t = \begin{cases} \tau_1 & \text{if } S_{t-1} = 0 \\ \tau_2 & \text{if } S_{t-1} = 1 \end{cases}$$

with $\tau_1$ being the probability of new infection and $\tau_2$ the probability of remaining infected. As both $\tau_1$ and $\tau_2$ are probabilities beta prior are used :

$$\tau_1 \sim \ Beta(\alpha_{\tau_1}, \beta_{\tau_1})$$

$$\tau_2 \sim \ Beta(\alpha_{\tau_2}, \beta_{\tau_2})$$

.

## 3.2 Incorporation of information on risk factors

The STOC free model takes RFs of new infection into account for the prediction of probabilities of infection. It is expected that the incorporation of RFs can improve the prediction of the probability of infection as compared to when using test results alone. Considering BVDV infection, purchase has often been identified as a RF of introduction of the infection, but the estimated strength of association between this RF and the probability of new infection varies a lot between studies (van Roon *et al.*, 2020a). The model can estimate this strength of association in a CP by incorporating both previous estimations as priors and historical test result data.

In the model, in each herd $h$ at each time step $t$, $\tau_{1,h,t}$ is modelled as a function of one or several RFs using logistic regression. This regression replaces the prior distribution on $\tau_1$ used when no RF is included. When RFs are included, $\tau_{1,h,t}$ is modelled as :

$$ln\Big(\frac{\tau_{1,h,t}}{1-\tau_{1,h,t}}\Big) = X_{h,t}\theta$$

where $X_{h,t}$ is the matrix of RF, with $h$ denoting herd and $t$ denoting time. $\theta$ is the vector of coefficients in the logistic regression, $\theta_1$ being the intercept and other $\theta$s the log-odds new infection associated with the risk factors. Normal priors are used for logistic regression coefficients :

$$\theta_i \sim \ Normal(\mu_i, \tau_i)$$

Several RFs can be included in the logistic regression. However, parsimony principles must be applied as each additionally included parameter will increase the number of parameters to estimate.

## 3.3 Incorporation of information on test results

The model includes test results as imperfect measures of the latent status. Test results depend on both the latent status regarding infection and test characteristics. These test characteristics are the test sensitivity and specificity, and are considered at the herd level. Test sensitivity is the probability of having a positive test result when the latent status is positive (Figure 3.3). Test specificity is the probability of having a negative test result when the latent status is negative (Figure 3.3).

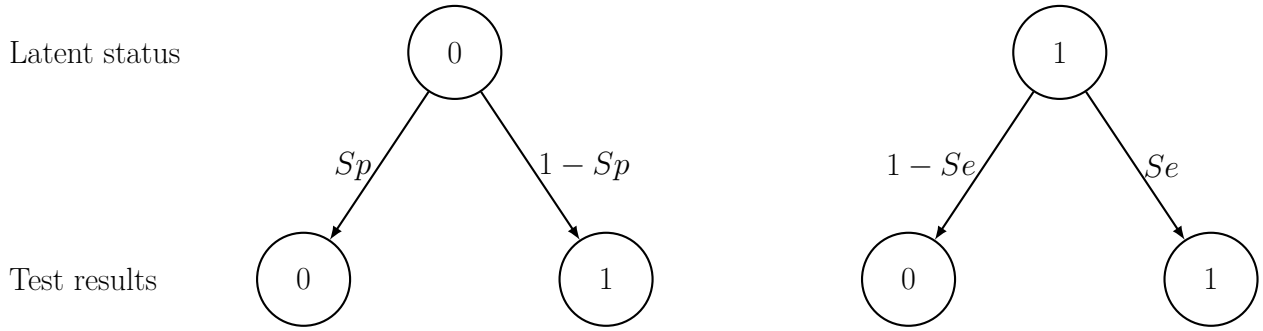



FIGURE 3.3 – Representation test results depending on latent status and herd level sensitivity ($Se$) and specificity ($Sp$).

Test results are modelled as following a Bernoulli distribution :

$$T_{ht} \sim \ Bernoulli(p(T^+_{ht}))$$

with $p(T^+_{ht})$ being the probability of being test positive. This probability depends on the latent status at $t$ as well as on the herd test sensitivity and specificity.

$$p(T^+_{ht}) = \begin{cases} 1 - Sp & \text{if } S_{ht} = 0 \\ Se & \text{if } S_{ht} = 1 \end{cases}$$

with $Se$ being the test sensitivity and $Sp$ the test specificity. Beta priors are used for these test characteristics :

$$Se \sim \ Beta(\alpha_{Se}, \beta_{Se})$$

$$Sp \sim \ Beta(\alpha_{Sp}, \beta_{Sp})$$

One potential difficulty when using the model could be the construction of priors for sensitivity and specificity because the model sensitivity and specificity relate to the detection of the herd latent status, while these characteristics are usually defined at the individual animal level. These characteristics could be widely different depending on the latent status of interest. Test characteristics can vary depending on the level of sampling (individual vs group sample). Moreover, test characteristics must be refined to represent its ability to detect the modelled latent status. As an example, Ab ELISA on BTM are often used within BVD CP in dairy cattle. If the infected latent status is the presence of at least one PI, the prior on sensitivity has to reflect the probability of the test being positive when there is at least one PI in the herd and not its sensitivity for the detection of antibodies in individual animals.

## 3.4 Prediction of the herd status at the last time step

The model predicts the herd level probability of being infected at the last time step using the status estimated on the previous month, the estimated infection dynamic parameters, and the estimated test specificity and sensitivity. First, the model predicts the probability of being herd status positive (noted $\tilde{\pi}^*_{h,t}$) depending on previous predicted status ($\hat{S}_{t-1}$) and estimated infection dynamics parameters ($\hat{\tau}_{1,h,t}, \hat{\tau}_2$) :

$$\tilde{\pi}^*_{h,t} = p(S_{h,t} = 1 | \hat{S}_{h,t-1}, \hat{\tau}_{1,h,t}, \hat{\tau}_2)$$

with

$$\hat{\tau}_{1,h,t} = logit^{-1}(X_{h,t}\hat{\theta})$$

Then, when a test result is available at the final time step, this predicted probability of being herd status positive ($\tilde{\pi}^*_{h,t}$) is updated with the test result to compute the final predicted probability of infection (noted $\tilde{\pi}_{h,t}$). Table 5.1 represents the probability of each test result depending on the predicted probability of being herd status positive ($\tilde{\pi}^*_{h,t}$) and test characteristics. The formula used to update the predicted probabilities of infection with the test results were derived from this table :

$$p(\tilde{S}_{h,t} = 1 | T_{h,t}, \tilde{\pi}^*_{h,t}) = \begin{cases} \frac{Se.\tilde{\pi}^*_{h,t}}{Se.\tilde{\pi}^*_{h,t} + (1-Sp).(1-\tilde{\pi}^*_{h,t})} & \text{if } T_t = 1 \\ \frac{(1-Se).\tilde{\pi}^*_{h,t}}{(1-Se).\tilde{\pi}^*_{h,t} + Sp.(1-\tilde{\pi}^*_{h,t})} & \text{if } T_t = 0 \end{cases}$$

with $T_t$ being test results at the final step time, and $Se$ and $Sp$ being test characteristic parameters estimated by the model.

TABLE 3.1 – Probability of test results depending on the estimated probability of being herd status positive at the last time step.

| | | Herd status at time to predict | |
|---|---|---|---|
| | | 1 | 0 |
| Test | 1 | $Se.\tilde{\pi}^*_{h,t}$ | $(1-Sp)(1-\tilde{\pi}^*_{h,t})$ |
| | 0 | $(1-Se)\tilde{\pi}^*_{h,t}$ | $Sp(1-\tilde{\pi}^*_{h,t})$ |

# 4 Implementation of the STOC free model

The STOC free model has been implemented in the JAGS computer programme (Plummer, 2003). JAGS performs Bayesian inference using Gibbs sampling. To interface R and JAGS the R runjags package was used (Denwood, 2016).

Chapitre 4

# Capacity of a Bayesian model to detect infected herds using disease dynamics and risk factor information from surveillance programmes : A simulation study

---

M. Mercat[1], A.M. van Roon[2], I. Santman-Berends[2,3], G. van Schaik[2,3], M. Nielen[2], D. Graham[4], S.J. More[5], M. Guelbenzu-Gonzalo[4], C. Fourichon[1], A. Madouasse[1]

**Keywords : output-based surveillance, Hidden Markov Model, herd-level probability of infection, freedom from infection, longitudinal data, repeated testing.**

[1] INRAE, Oniris, BIOEPAR, Nantes 44300, France

[2] Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, PO Box 80151, 3508, TD Utrecht, the Netherlands

[3] Royal GD, PO Box 9, 7400 AA, Deventer, the Netherlands

[4] Animal Health Ireland, Unit 4/5, The Archways, Bridge St., Carrick-on-Shannon, Co. Leitrim N41 WN27, Ireland

[5] Centre for Veterinary Epidemiology and Risk Analysis, UCD School of Veterinary Medicine, University College Dublin, Belfield, Dublin D04 W6F6, Ireland

Ce chapitre a été soumis à *Preventive veterinary medecine* et est actuellement en cours de révision.

# 1 Abstract

Control programmes against non-regulated infectious diseases of farm animals are widely implemented. Different control programmes have different definitions of *freedom from infection* which can lead to difficulties when trading animals between countries. When a disease is still present, in order to identify herds that are safe to trade with, estimating herd-level probabilities of being infected when classified *free from infection* using field data is of major interest. Our objective was to evaluate the capacity of a Bayesian Hidden Markov Model, which computes a herd-level probability of being infected, to detect infected herds compared to using test results only. Herd-level risk factors, infection dynamics and associated test results were simulated in a population of herds, for a wide range of realistic infection contexts and test characteristics. The model was used to predict the infection status of each herd from the simulated risk factor and test result longitudinal data. Two different indexes were used to categorize herds from the probability of being infected into a herd predicted status. The model predictive performances were evaluated using the simulated herd status as the gold standard. The model detected more infected herds than a single final test in 85% of the scenarios which converged. The proportion of infected herds additionally detected by the model, compared to test results alone, varied a lot depending on the context. It was higher in a context of a low herd test sensitivity. On average 20% to 50% of the newly infected herds undetected by the test were classified as infected by the model. Model convergence did not occur for 39% of the scenarios, mainly in association with low herd test sensitivity. Detection of additional newly infected herds was always associated with an increased number of false positive herds (except for one scenario). The number of false positive herds was lower for scenarios with low herd test sensitivity and moderate to high incidence and prevalence. These results highlight the benefit of the model, in particular for control programmes with infection present at an endemic level in a population and reliance on test(s) of low sensitivity.

# 2 Introduction

Various control programmes (CPs) against infectious diseases of farm animals are implemented in Europe. In order to control or eradicate these diseases, CPs typically focus on the identification of infected units (animals or herds) using diagnostic tests performed at regular time intervals. Testing schemes can vary in terms of type and performance of

the test used, the cohorts and numbers of animals tested, and the time interval between tests. CPs may be deployed across a territory, from regional to national scale. These differences have been documented for some endemic cattle diseases, including infections by bovine viral diarrhoea virus (van Roon *et al.*, 2020b), *Mycobacterium avium* subspecies *paratuberculosis* (Whittington *et al.*, 2019), and bovine herpesvirus 1 (Raaperi *et al.*, 2014).

Heterogeneity in CPs may lead to difficulties when trading animals between different regions or countries, as each CP has its own definition of *freedom from infection* which cannot be directly compared. These definitions of a *free status* are usually based on one, or a combination of several, diagnostic test result(s). Limitations in test performance, and time interval between tests lead to uncertainty around these statuses. Imperfections in the testing schemes lead to two types of error. Firstly, a lack of specificity means that some uninfected herds are wrongly categorized as infected, i.e. false positives. Secondly, a lack of sensitivity leads to some infected herds being wrongly categorized as free from infection, i.e. false negatives. The time interval between tests may result in a delay between the times of infection and detection. For herds classified as *free from infection*, those that become infected between two consecutive tests will remain classified as *free from infection* until a next test event. Hence, as each CP has its own surveillance strategy, the confidence and associated uncertainty in the true status of a herd classified as *free from infection* may vary depending on the CP. Currently when purchasing an animal from a herd classified as *free from infection* under different CPs, it is not possible to assess the probability of infection for that animal. As trade can be an opportunity for infectious diseases to spread, confidence in *free status* is a key point to support international trade.

There is a need for the development of methods that enable a CP-level comparison of confidence of herd-level *freedom from infection*. Traditionally, input-based surveillance has been implemented, prescribing how surveillance for a given disease should be performed in terms of the tests used and the proportion of herds and animals tested. However, input-based surveillance does not take into account the diversity of contexts in which CPs are applied (van Roon *et al.*, 2020b) and can be expensive to run, while not being adapted to the specific context of each CP (Cameron, 2012). Alternatively, output-based surveillance may be used, which is not prescriptive in terms of the elements of the programme, but rather in the degree of confidence associated with a free status that must be achieved.

Imperfect testing regimes lead to misclassification errors, as highlighted above. To

account for this, known risk factors (RFs) for the introduction of infection could be included in the calculation of probability of freedom, as predictors of either current or new infection. Data on such disease-specific RFs should be available for many CPs, given that action on these RFs is used as a way to prevent the introduction of infection. Disease-specific RFs for introduction depend on the pathogen as well as the route of transmission (direct or indirect transmission). For many diseases, animal purchase is a common RF for introduction of infection into herds (Rangel *et al.*, 2015; van Roon *et al.*, 2020a). In the European Union, where cattle identification and the recording of cattle movements between holdings are mandatory, these data could be used to predict (new) infections through purchase, thus contributing to improved estimation of the infection-free status of a herd.

In Madouasse (Madouasse *et al.*, 2021), a modelling framework was described, called the STOC free (Surveillance analysis Tool for Output based Comparison of the confidence of FREEdom from infection) model, that estimates the herd-level probabilities of infection, using data from CP and taking RF occurrence into account. The model estimates the probability of infection at the last time-step for each herd (in a series of sequential test results). Model inputs include repeated test results and the presence of RFs for each herd as measured regularly within the surveillance programme. The framework incorporates knowledge at the population level on infection dynamics, test characteristics and the effect of RFs when estimating probability of infection.

In order to evaluate the capacity of the STOC free model to detect infected herds, a gold standard is required. Gold standard is the true herd status. In the context of the STOC free model, an infected herd is defined as the presence of at least one infected animal. To measure the true status of the herd, it would be necessary to test all the animals within a herd using a perfect test. However, no such data exist in the real-world. Up to now, the STOC free model has only been applied to a single French dataset, which included test results and RFs but no gold standard (Madouasse *et al.*, 2021). An evaluation of the performance of this model under different circumstances is therefore lacking.

The use of simulated data is the most effective way to evaluate the predictive accuracy of the STOC free model given the absence of gold standard information in real-world surveillance data. Data simulation allows a simplified system to be created where the true herd status is known. Simulated surveillance data, i.e. test results and RF information collected at regular intervals, can be used as input for the STOC free model as an alternative

to real-world surveillance data. The performance of the model can be then evaluated by looking at errors in herd status classification, by comparing true herd status to the status predicted by the model. Furthermore, compared to real data, using simulated data enables a wide range of epidemiological situations and surveillance modalities to be evaluated. It makes it possible to investigate the potential of the model to be used for different diseases where performance of CPs differs.

The objective of this work was to evaluate the capacity of the STOC free model, which takes account of both dynamics of testing and risk factor information, to improve the detection of infected and newly infected herds compared to test results alone (i.e. the added value of the model in sensitivity). Among infected herds, newly infected are the ones which were not infected at the previous test event. We assume that the added value of the model in terms of the detection of newly and previously infected herds could be different depending on the epidemiological context. Simulated data were used to generate a wide range of realistic CPs (different CP corresponding either to different diseases or to the results of different testing strategies for a disease in different contexts). We quantified the number of additional infected herds detected by the STOC free model compared to test results.

# 3 Material and methods

## 3.1 Overall design strategy

Firstly, a dynamic model was developed to simulate herd-level infection and surveillance data under a wide variety of CP scenarios corresponding either to different diseases or to the results of different testing strategies for a disease in different contexts. The simulated surveillance data were then used as input for the STOC free Bayesian Hidden Markov Model, which was run to generate outputs on model parameters estimates and predicted herd status for probability of infection at the last time-step. Finally, model and test performance were compared. The overall design strategy is presented in Figure 4.1.

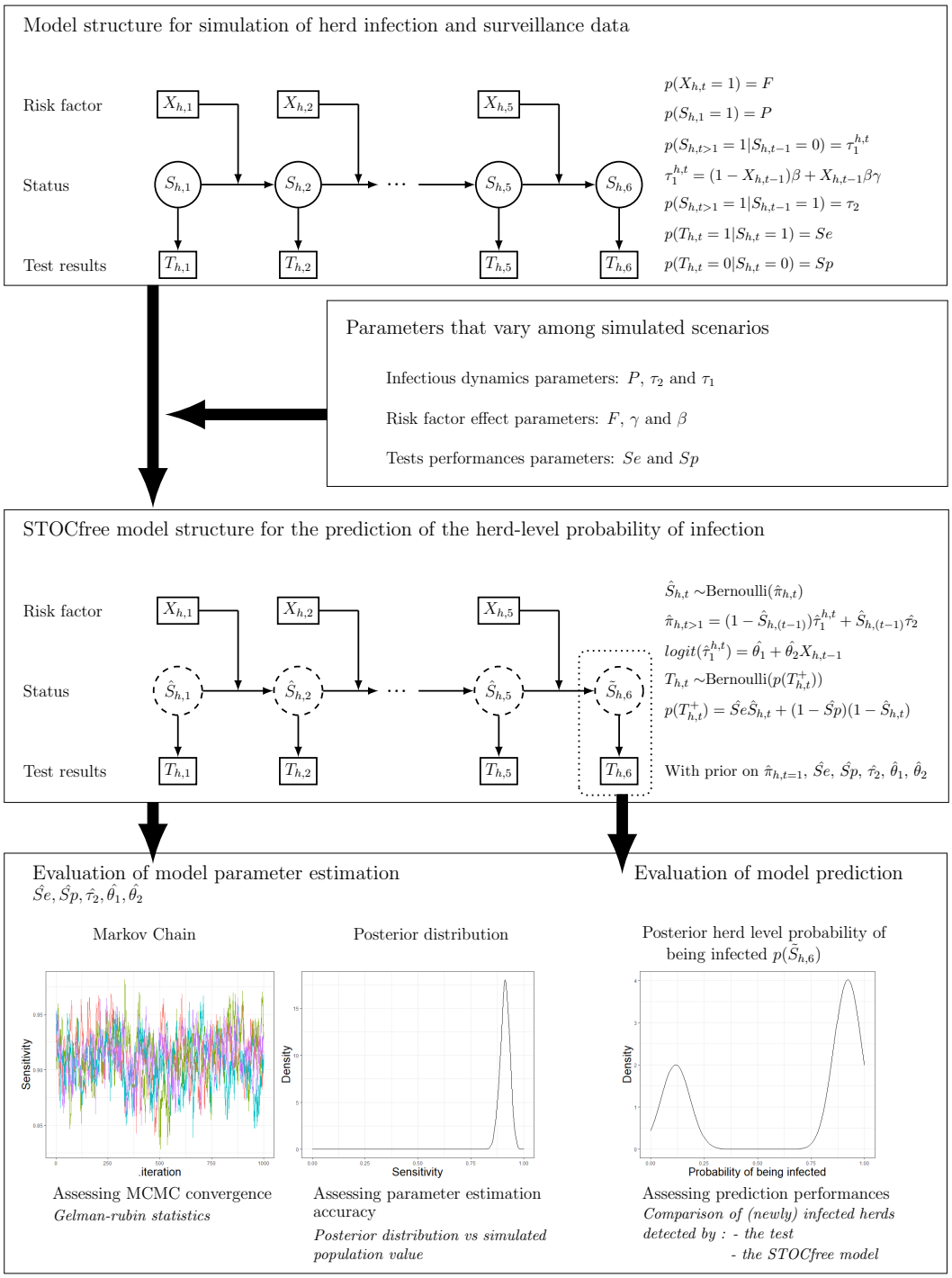



FIGURE 4.1 – Representation of the design strategy. Variables in rectangles represent observational data (risk factor and test results). Variables in circles represent herd infection statuses : true simulated status in solid line and latent estimated/predicted status in dashed line. Observational data simulated using the simulation model are used as input for the STOC free model. Herd statuses predicted by the STOC free model on the 6th time-step are compared to the corresponding simulated statuses, considered as the gold standard.

## 3.2 Simulation of herd infection and surveillance data model

We simulated the dynamics of herd infection status in the presence of a single RF associated with an increased probability of introduction, generating data on herd status and test results at each time-step. Initially, RF presence/absence was simulated. Herd infection status at the first time-step was based on the chosen simulated prevalence of infection. Then, at a given time-step, non-infected herds could become infected according to a probability of new infection between time-steps, which varied depending on RF occurrence. The probability that an infected herd would remain infected between two sequential time-steps was determined by a simulation parameter that represented this probability (of infection not being resolved between two different time-steps). Infection status for a given herd at a given time-step determined the result of a test, assuming a given herd-level test sensitivity and specificity.

### 3.2.1 Simulation of herd status at each time-step

Infection dynamics were simulated by herd status change. Herd status was simulated as a binary event, with 0 and 1 denoting absence and presence of infection, respectively. Herd status was assumed to undergo Markovian dynamics with status at time $t$ depending on status at time $t-1$ and RF occurrence. In each scenario, the overall herd infection prevalence was held constant over the time-steps to evaluate the STOC free model in different situations over a short period. Keeping the prevalence constant prevents the infection of either dying out or rapidly increasing over the number of time-steps and allows a comparable number of infected herds to be detected. For consistency, the probability of new infection between time-steps was a function of both overall herd infection prevalence and the probability of a herd remaining infected between time-steps to allow overall prevalence to remain constant over time.

Status simulation can be described by the following set of equations. In herd $h$ at time $t$, the infection status $S_{h,t}$, was sampled from a Bernoulli distribution :

$$S_{h,t} \sim Bernoulli(\pi_{h,t}) \,,$$

with $\pi_{h,t}$ being the probability of being infected at time-step $t$ for herd $h$. For a given

herd at time $t = 1$, the probability of infection was :

$$\pi_{h,t=1} = P\,,$$

with $P$ being the herd infection prevalence for that scenario. For a given herd $h$ at time $t > 1$, $\pi_{h,t}$ depended on previous status and infection dynamics parameters :

$$\pi_{h,t} = (1 - S_{h,t-1})\tau_1^{h,t} + S_{h,t-1}\tau_2\,,$$

with $S_{h,t-1}$ being the status of herd h at the previous time-step, $\tau_2$ being the probability of remaining infected between time-steps (fixed variable in each scenario), and $\tau_1^{h,t}$ the probability of new infection between time-steps which was defined as a function of herd-level risk factor exposure and defined as :

$$\tau_1^{h,t} = (1 - X_{h,t-1})\beta + X_{h,t-1}\beta\gamma\,,$$

where $\beta$ was the probability of new infection when the risk factor was absent, i.e. $X_{h,t-1} = 0$ and $\beta\gamma$ was the probability of new infection when the RF was present, i.e. $X_{h,t-1} = 1$. Thus, $\gamma$ was the relative risk of new infection in herds exposed to the RF. Exposure to the RF ($X$) was considered a random dichotomous variable simulated as :

$$X_{h,t} \sim Bernoulli(F)\,,$$

with $F$ being the RF frequency in the data set.

Assuming an endemic situation with a constant prevalence over time-steps, at each time-step in each scenario the average number of newly infected herds was constrained to be equal to the average number of herds eliminating the infection. Therefore, the following condition had to be met :

$$E\left(\tau_1^{h,t}\right)(1 - P) = (1 - \tau_2)P\,,$$

where $E\left(\tau_1^{h,t}\right)$ was the expectation for the probability of new infection. This amounts to applying the following constraint on the overall probability of new infection :

$$E\left(\tau_1^{h,t}\right) = \frac{(1 - \tau_2)P}{1 - P}\,.$$

From the definition of $\tau_1^{h,t}$ and the frequency of the RF, $F$, at a given time-step, the expected probability of new infection was :

$$E\left(\tau_1^{h,t}\right) = (1-F)\beta + F\beta\gamma\,,$$

where $\gamma$ was the relative risk of new infection in herds exposed to the RF and $\beta$ the probability of new infection in herds that were not exposed to the RF. The frequency of the RF ($F$) and the relative risk of new infection in herds exposed to the RF ($\gamma$) are inputs in the simulation. The probability of new infection in herds that were not exposed to the RF ($\beta$) can be computed as :

$$\beta = \frac{E\left(\tau_1^{h,t}\right)}{1+F(\gamma-1)}\,.$$

### 3.2.2 Simulation of test results

A test result was simulated for each herd at each time-step as a function of the simulated herd status, the herd-level test sensitivity and specificity. Test result in herd $h$ at time $t$ was sampled from a Bernoulli distribution :

$$T_{h,t} \sim Bernoulli\left(p\left(T_{h,t}^{+}\right)\right)\,,$$

with $p\left(T_{h,t}^{+}\right)$ being the probability of being tested positive defined by :

$$p\left(T_{h,t}^{+}\right) = (1-S_{h,t})Se + S_{h,t}(1-Sp),$$

with $Se$ and $Sp$ being respectively herd-level test sensitivity (probability for an infected herd to be tested positive) and specificity (probability for an uninfected herd to be tested negative).

## 3.3 Input scenario : differing infection and epidemiological situation

We simulated various scenarios to represent different diseases in different contexts and different tests performances for which STOC free model could be used. Different range of values for the 10 different parameters of the data simulation are presented in Table 4.1. For all scenarios, the number of simulated herds was set at 5,000 and the number of simulated time-steps to 6. At each time-step, test results and RF information were available. The choice of parameter values was based on knowledge and discussion with a group of infectious disease experts, from different countries involved in the STOC free consortium, to represent variation in context for different endemic situations.

Various epidemiological situations were simulated to represent various endemic infections and contexts. We simulated two prevalence values, 0.3 and 0.1, representing territories in the beginning of their CP and territories already in an advanced stage of control, respectively. The probability of remaining infected depends on the effectiveness of herd-level eradication measures in the CP. We consider high values, from 0.75 to 0.9, consistent with endemic infection dynamics. For consistency with a constant prevalence of infection, the probability of becoming infected ($\tau_1$) was calculated for all combinations of $P$ and $\tau_2$ values (4 values).

Various effect of RFs on infection dynamics has been simulated to account for variability between CP. We simulated low to high RF frequency setting a maximum frequency of 0.5 considering that a more frequent risk factor would not be discriminatory between herds. In contrast, we have set a minimum frequency at 0.1 because a very rare RF (below 0.1) will only bring information for a small number of herds. The relative risk of new infection in herds exposed to the RF ($\gamma$) ranged from 1.5 to 5, given that RF association may be variable depending on the infection and territory (van Roon *et al.*, 2020a).

Test characteristic parameters represent herd-level sensitivity and specificity for the complete testing process to detect herd-level infection. These parameters depend on specific test characteristics, the number of animals tested and within-herd prevalence Christensen & Gardner (2000). Therefore, herd-level sensitivity and specificity can differ from specific test characteristics. We simulated herd-level sensitivity from 0.4 to 0.9 and herd-level specificity from 0.8 to 0.95. Low herd-level sensitivity values represent infections for which highly sensitive tests are not available, e.g. paratuberculosis (Nielsen & Toft, 2008). We considered a sensitivity of 0.9 as the maximum value. In case of higher sensitivity, we

hypothesized that there would be limited added value from the STOC free model. After taking into account the complete testing process, which often includes retesting of positive herds, high values of specificity were considered appropriate. Low diagnostic specificity is less common in CPs.

Combinations of parameters values represented the simulation of 216 different scenarios. Simulation of herd infection and surveillance data model were implemented in R software (R *et al.*, 2017).

TABLE 4.1 – Parameter values for scenario simulation.

| Parameter | Description | Value | Condition |
|---|---|---|---|
| $nherds$ | Number of herds | 5000 | - |
| $nTests$ | Number of test times per herd | 6 | - |
| $Se$ | Herd-level sensitivity | 0.4, 0.7, 0.9 | - |
| $Sp$ | Herd-level specificity | 0.8, 0.95 | - |
| $P$ | Prevalence of infection | 0.1, 0.3 | - |
| $\tau_2$ | Probability of remaining infected | 0.75, 0.9 | - |
| $\tau_1$ | Probability of becoming infected | 0.011, 0.028, 0.043, 0.107 | Depends on $\tau_2$ and $P$ value |
| $\gamma$ | Relative risk associated with $X$ | 1.5, 2, 5 | - |
| $F$ | Frequency of $X$ | 0.1, 0.25, 0.5 | - |
| $\beta$ | Probability of new infection for an uninfected herd without $X$ | 0.004, 0.005, 0.007, 0.008, 0.009, 0.010, 0.014, 0.019, 0.020, 0.021, 0.022, 0.025, 0.027, 0.029, 0.031 0.034, 0.036, 0.038, 0.039, 0.041, 0.054, 0.071, 0.076, 0.086, 0.086, 0.095, 0.097, 0.102 | Depends on $\tau_1$, $\gamma$ and $F$ value |

## 3.4 Description and use of the STOC free model

The model described by Madouasse *et al.* (2021), represents infection presence at herd level as a latent status over time-steps. The latent status is evaluated at regular time intervals through testing. Tests may be imperfect, i.e. with a sensitivity and a specificity less than 1. The variable of interest (the latent status) has a Markovian dynamic : the latent status at a given time-step depends on both the latent status at the previous time-step and actions taken or RF occurrence since the previous time-step. Risk factors is incorporated as predictors for new infection. The model predicts the probability of infection in the final time-step for each herd in the CP. Data collected before the final time-step are used as historical data for the estimation of the different model parameters, including previous latent statuses. Parameters estimation and prediction are performed in a Bayesian framework.

### 3.4.1 Model Structure

To describe the STOC free model and explain how predictions were performed, we use the following notation : $\hat{\beta}$ is the estimated value of $\beta$ and $\tilde{y}$ is the predicted value of $y$.

**Latent state.** We consider two latent states : 0 for uninfected herds and 1 for infected herds. For a given herd $h$ at a given time $t$, status $\hat{S}_{h,t}$ follows a Bernoulli distribution :

$$\hat{S}_{h,t} \sim Bernoulli\left(\hat{\pi}_{h,t}\right) ,$$

with $\hat{\pi}_{h,t}$ being the probability of being infected. At $t = 1$, a beta prior is used for $\hat{\pi}_{h,t=1}$, representing initial prevalence :

$$\hat{\pi}_{h,t=1} \sim Beta(\alpha_\pi, \beta_\pi)) .$$

**Infection dynamics.** From the second time-step on, the probability of being infected at t depends on the latent state at $t-1$. Herds that were uninfected at $t-1$ (i.e. $\hat{S}_{h,(t-1)} = 0$) can become infected with probability of new infection $\hat{\tau}_1^{h,t}$. Infected herds remain infected with a probability of remaining infected $\hat{\tau}_2$ :

$$\hat{\pi}_{h,t} = \left(1 - \hat{S}_{h,(t-1)}\right) \hat{\tau}_1^{h,t} + \hat{S}_{h,(t-1)} \hat{\tau}_2.$$

A beta prior is used for the probability of remaining infected, which is constant over time and herds :

$$\hat{\tau_2} \sim Beta\left(\alpha_{\tau_2},\, \beta_{\tau_2}\right) .$$

**Probability of new infection.** The probability of new infection $\tau_{i,t}^1$ is modelled as a function of the presence or absence of the RF $X_{h,t-1}$ using a logistic regression :

$$\text{logit}\left(\hat{\tau}_1^{h,t}\right) = \hat{\theta}_1 + \hat{\theta}_2 X_{h,t-1} .$$

Normal priors are used for logistic regression parameters $(\hat{\theta}_1,\, \hat{\theta}_2)$ :

$$\hat{\theta}_1 \sim Normal(\mu_1, \sigma_1) ,$$

$$\hat{\theta}_2 \sim Normal(\mu_2, \sigma_2) .$$

**Test results.** Test results are considered as an imperfect measure of the latent status. We consider two herd-level test results : positive or negative (discrete). Each result follows a Bernoulli distribution with a probability $p\left(T^+\right)_{h,t}$ of being positive :

$$T_{h,t} \sim Bernoulli\left(p\left(T_{h,t}^+\right)\right) ,$$

with $p\left(T^+\right)_{h,t}$ depending on estimate latent status at $t$ and test characteristics : herd-level sensitivity $(\hat{Se})$ and specificity $(\hat{Sp})$ :

$$p\left(T^+\right)_{h,t} = \hat{Se}\hat{S}_{h,t} + (1 - \hat{Sp})(1 - \hat{S}_{h,t}) .$$

Beta priors are used for test characteristics parameters :

$$\hat{Se} \sim Beta(\alpha_{Se}, \beta_{Se}) ,$$

$$\hat{Sp} \sim Beta(\alpha_{Sp}, \beta_{Sp}) .$$

### 3.4.2 Predicting the probability of infection

The model predicts the herd-level probability of being infected at the last time-step using status prediction from the previous month, estimated infection dynamic parameters, and estimated test specificity and sensitivity.

First, the model predicts the probability of being herd status positive $\left(\text{noted}\, p\left(\tilde{S}^{+*}_{h,t}\right)\right)$ depending on previous predicted status $\left(\hat{S}^{+}_{h,t-1}\right)$ and estimated infection dynamics parameter $(\tilde{\tau}^{1}_{h,t}, \hat{\tau}_2)$ :

$$p\left(\tilde{S}^{+*}_{h,t}\right) = p\left(\tilde{S}^{+*}_{h,t} | p\left(\hat{S}^{+}_{h,t-1}, \tilde{\tau}^{h,t}_{1}, \hat{\tau}_2\right)\right) ,$$

with

$$\tilde{\tau}^{h,t}_{1} = \text{logit}^{-1}\left(\hat{\theta}_1 + \hat{\theta}_2 X_{h,t-1}\right) .$$

Then, it combines this prediction to test results to compute the final predicted probability of being infected $\left(\text{noted}\, p\left(\tilde{S}^{+}_{h,t}\right)\right)$ :

$$\begin{aligned} p\left(\tilde{S}^{+}_{h,t} | T^{+}_{h,t}, \tilde{S}^{+*}_{h,t}\right) &= T^{+}_{h,t} . \frac{\hat{Se}.p\left(\tilde{S}^{+}_{h,t-1}\right)}{\hat{Se}.p\left(\tilde{S}^{+}_{h,t-1}\right) + \left(1-\hat{Sp}\right).\left(1-p\left(\tilde{S}^{+}_{h,t-1}\right)\right)} \\ &+ \left(1-T^{+}_{h,t}\right) \frac{\left(1-\hat{Se}\right).p\left(\tilde{S}^{+}_{h,t-1}\right)}{\left(1-\hat{Se}\right).p\left(\tilde{S}^{+}_{h,t-1}\right) + \hat{Sp}.\left(1-p\left(\tilde{S}^{+}_{h,t-1}\right)\right)} , \end{aligned}$$

with $T^{+}_{h,t}$ being test results at final step time, and $\hat{Se}$ and $\hat{Sp}$ being test characteristics parameters estimated by the model. The way to estimate these predicted probability and test results is presented in supplementary materials.

### 3.4.3 Choice of prior distribution

The STOC free model requires prior distributions for six different parameters : $\hat{Se}$, $\hat{Sp}$, $\hat{\tau_2}$, $\hat{\pi}_{h,t=1}$, $\hat{\theta}_1$ and $\hat{\theta}_2$. We used Beta distributions for parameters bounded between 0 and 1. Parameter $\alpha$ and $\beta$ of these distributions can be computed using the mean and variance. In our model, a Beta prior was used for the probability of being infected at time-step 1 ($\hat{\pi}_{h,t=1}$), test characteristics $\left(\hat{Se}\,\text{and}\,\hat{Sp}\right)$ and for the probability of remaining negative ($\hat{\tau}_2$). We used true input parameter values as the means. We used a Normal prior for the logistic regression parameter $\left(\hat{\theta}_1\,\text{and}\,\hat{\theta}_2\right)$ centred on the true value. Types of priors and distribution parameters used are summarized in Table 4.2. Example of the 95% intervals are displayed in the supplementary material.

TABLE 4.2 – Prior distribution for the model parameters.

| Parameter | Description | Distribution | Mean | Variance |
|---|---|---|---|---|
| $\hat{Se}$ | Herd-level test sensitivity | *Beta* | True value | $0.05^2$ |
| $\hat{Sp}$ | Herd-level test specificity | *Beta* | True value | $0.05^2$ |
| $\hat{\tau_2}$ | Probability for an infected herd not to eliminate the infection | *Beta* | True value | $0.05^2$ |
| $\hat{\theta_1}$ | Intercept (risk factor) | *Normal* | True value | 1 |
| $\hat{\theta_2}$ | Coefficient (risk factor) | *Normal* | True value | 1 |
| $\hat{\pi_{i,1}}$ | Probability of being infected at time 1 | *Beta* | Prevalence true value | $0.15^2$ |

## 3.5 Evaluation of STOC free model output

For each scenario, the STOC free model produced different outputs. The model returns Markov Chain Monte Carlo (MCMC) samples from the posterior distributions model parameters and probabilities of being infected at the last time-step. Model parameters include parameters related to infection dynamics, association between RF and probability of new infection and test characteristics. Estimations of these model parameters are performed from historical data on test results and RFs (in our case, data from the first five time-steps) as well as from the prior distributions for the different model parameters. First, we evaluated the convergence of the MCMC chains as well as the consistency between estimated model parameters and the parameters used for simulating the data. Then, from the posterior distributions of the herd-level probabilities of infections, rules were defined to categorize herds as infected or uninfected. Error rates of the STOC free model were computed and compared to test results to enable computation of model performance.

### 3.5.1 Evaluation of model parameter estimation

#### 3.5.1.1 Assessing MCMC convergence

The STOC free model were implemented in the JAGS computer programme (Plummer, 2003). The model was applied to each scenario, running 4 chains in parallel. We removed the first 1,000 iterations as burn-in. Then 5,000 more iterations were run, of which one in five iterations was stored for analysis, to reduce the size of the output file. For each parameter, the posterior distribution was built with 4000 iterations (1000 for each chain). We used the Gelman-Rubin statistics ($\hat{r}$) to assess convergence of the chains (Gelman *et al.*, 1992). This statistic was computed for the five parameters estimated by the model ($\hat{Se}$, $\hat{Sp}$, $\hat{\tau}_2$, $\hat{\theta}_1$ and $\hat{\theta}_2$). We considered that scenarios with $\hat{r}$ values less than 1.05 had converged. Scenarios that did not reach convergence using 1,000 iterations of burn-in were run again using 5,000 iterations of burn-in. Convergence was assessed using Gelman-Rubin statistics. Scenarios that did not reach convergence after this second step were excluded for the rest of the analysis. To again run these scenarios with more iterations would have been too time consuming.

#### 3.5.1.2 Verification of parameter estimation

Parameter estimation was verified by comparing the posterior distributions to the parameter values within the simulated populations. In the simulation process, some events are driven by probabilities of occurrence. Therefore, the value of a parameter can differ between the chosen value for simulating a scenario and the resulting simulated population value. We verified that the distribution of the parameter estimates issued from the STOCfree model was consistent with the simulated population value.

### 3.5.2 Evaluation of model prediction performances

The STOC free model returns distributions of the predicted posterior probability of being infected for the 5000 herds at the last time-step (Figure 4.1). In order to evaluate the performance of the model for the prediction of true infection status, these probability distributions were discretised into *predicted infected* or *predicted uninfected* status. First,

each herd posterior probability of being infected was summarised. The median probability per herd was used as the summary value as it was the variable that best discriminated between uninfected and infected herds (results not shown). Then, a cut-off value was applied to the summary values to classify herds as predicted infected or uninfected. The general framework of the prediction performances analysis is presented in Figure 4.2.

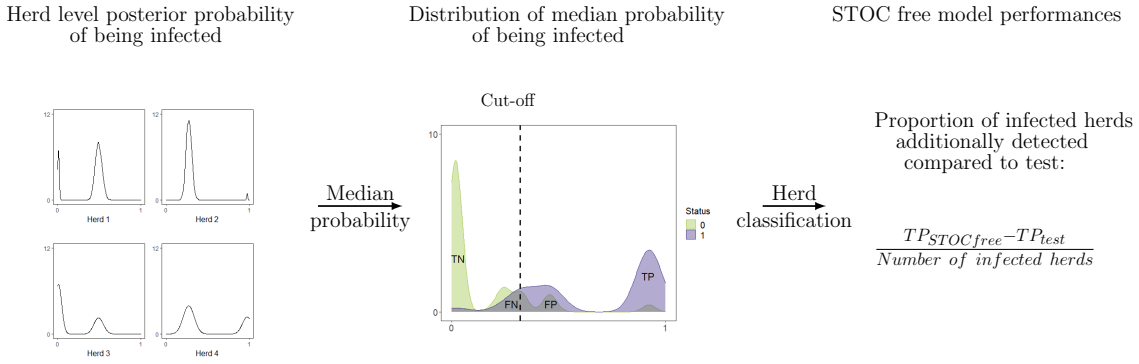



FIGURE 4.2 – Representation of the STOC free model prediction performance analysis. At first stage posterior herd probability of being infected are summarized using the median value. Then, categorization of herds is done by applying a cut-off to the distribution of posterior median. Cut-off determination is based on two different indexes. Finally, the performance of the STOC free model is obtained by comparing the number of true positives using the STOC free model and the number of true positives obtained using test information alone.

Two different indices were used to select the cut-off value, corresponding to two different objectives and are described below. Those two methods are based on knowledge of true herd status. In our study, we used the simulated herd status as the gold standard, which allowed the number of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) to be calculated. TP herds are infected herds classified infected, FN are infected herds classified uninfected, FP are uninfected herds classified infected and TN are uninfected herds classified uninfected. We computed them using the STOC free model or the test, represented by corresponding subscript (e.g. $TP_{STOCfree}$ and $TP_{test}$).

#### 3.5.2.1 Identification of cut-off value using Youden's index

Firstly, we used all herd predictions at the last time-step to estimate the cut-off value which minimized classification error (i.e. false positive and false negative). The cut-off choice was determined using the criterion below (Youden, 1950), noting that it is a trade-off between sensitivity and specificity :

$$Youden's\,index = max(Se + Sp),$$

with

$$Se = \frac{TP_{STOCfree}}{TP_{STOCfree} + FP_{STOCfree}},$$

and

$$Sp = \frac{TN_{STOCfree}}{TN_{STOCfree} + FN_{STOCfree}}.$$

We ran this analysis using pROC packages in R software.

We compared STOC free model performances to test performances. We firstly compared the number of accurately classified herds (TN+TP) by the STOC free model and by the test. Then, we explored the impact of the simulation parameter values on the additional number of infected herds (TP) detected by the STOC free model compared to test results.

We applied this cut-off value to a sub-group of the population, specifically only herds that were not infected at the step before prediction (i.e. candidate herds for new infection), using true simulated herd status, to allow us to distinguish between herds remaining uninfected and newly infected herds. We compared STOC free model performances to test performances by doing the same analysis as described above.

#### 3.5.2.2 Alternative cut-off optimizing detection of newly infected herds

We explored an alternative method to choose a cut-off value designed to evaluate the performances of the model for detection of newly infected herds compared to testing. We selected herds that were candidates to be newly infected. With this approach, we firstly constrained the cut-off value to detect at least one more newly infected herd compare to

test results :

$$Number\,of\,additional TP = TP_{STOC free} - TP_{test} > 0\,.$$

For cut-off values that verified this condition, we computed the associated additional number of false positive (FP) :

$$Number\,of\,additional FP = FP_{STOC free} - FP_{test}\,.$$

Finally, we computed the NewI cost index. This index was based on a trade-off between the additional numbers of true positive herds and of false positive herds in the STOC free model compared to test results :

$$NewI\,cost\,index = \frac{Number\,of\,additional\,FP}{Number\,of\,additional\,TP}\,.$$

We chose the cut-off value with the lowest value of NewI cost index. This NewI cost index represents the additional number of false positive for each additional true positive. When the NewI cost index is negative, there are fewer FP and more TP using the STOC free model results compared to the test results. A NewI cost index of 1 implies that using the STOC free model we had one additional FP for each additional TP. When the NewI cost index is positive (and more than one), there is more than one additional FP for each additional TP using the STOC free model.

In addition, cut-off values selected with both methods (Youden index and NewI cost index) were compared.

# 4 Results

## 4.1 Evaluation of model parameter estimation

### 4.1.1 Assessing MCMC convergence

Of the 216 scenarios, 131 had a $\hat{r} < 1.05$ for all parameters $\left(\hat{Se},\, \hat{Sp},\, \hat{\tau}_2,\, \hat{\theta}_1 \text{ and } \hat{\theta}_2\right)$ which confirmed convergence. For the 85 other scenarios, at least one of the five estimated parameters had a $\hat{r} > 1.05$. For most of these scenarios (71/85), $\theta_1$ chains did not converge. There were fewer scenarios where $Se$, $Sp$, $\tau_2$ and $\theta_2$ chains did not converge (21, 28, 43

and 37 of 85 scenarios, respectively). These scenarios were re-run using a greater number of iterations during burn-in. From those 85 scenarios, 41 subsequently converged.

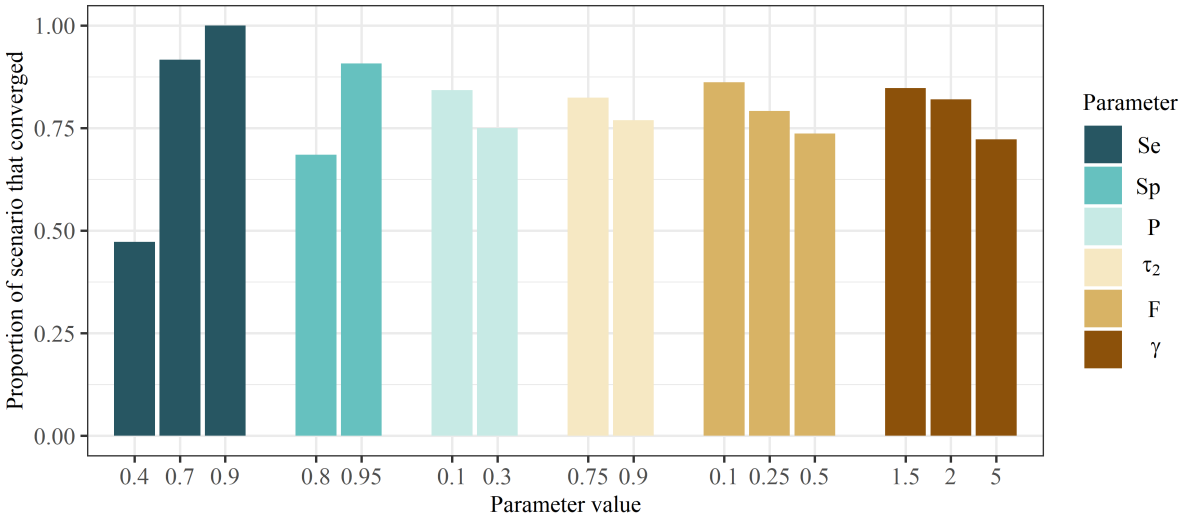



FIGURE 4.3 – Proportion of scenarios that converged for each simulation parameter value. Six of the seven simulated parameters are represented : $Se$ (test sensitivity), $Sp$ (test specificity), $P$ (prevalence), $\tau_2$ (probability of remaining infected), $F$ (frequency of the risk factor) and $\gamma$ (relative risk associated with the risk factor).

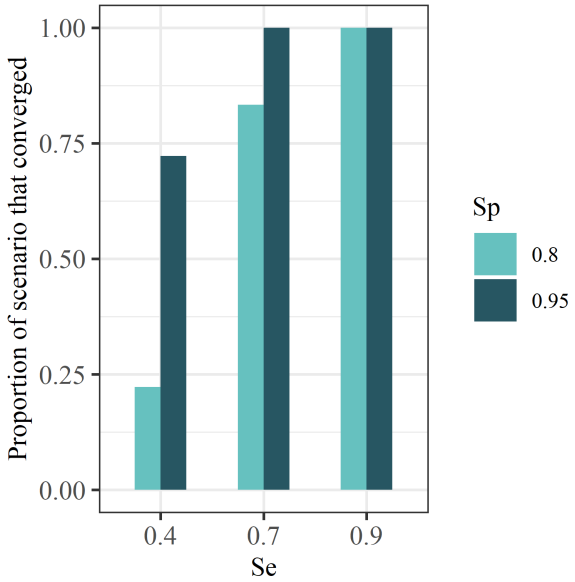



FIGURE 4.4 – Proportion of scenarios that converged for all combinations of $Se$ (test sensitivity) and $Sp$ (test specificity) values.

The proportion of scenarios that finally converged (with either 1,000 or 5,000 iterations) varied between values of the simulation parameters (Figure 4.3). About half of the scenarios (38/72) with a test sensitivity of 0.4 did not converge, and about a third of the scenarios (34/108) with a test specificity of 0.8 did not converge. The values of these two simulation parameters ($Se$ and $Sp$) had the biggest impact on convergence (Figure 4.3). Considering both parameters, it appears that higher specificity values helped the model to converge for lower and medium, but to a lesser extent with sensitivity values of 0.4 and 0.7. However, it did not make any difference for scenarios with higher sensitivity values (Figure 4.4).

### 4.1.2 Checking parameters estimation

Of the 172 scenarios for which model convergence was validated, the credibility interval for at least one parameter did not include the true parameter value in 13 scenarios. Depending on the scenario, the parameter for which this was the case varied. The gap between the 95% confidence interval of the posterior distribution and the population value was low in each of the 8 scenarios (supplementary material).

## 4.2 Evaluation of model prediction performances

Performances of the model were analysed for the 172 scenarios that did converge. Table 4.3 summarizes the number of scenarios for each simulation parameter value remaining at this step.

### 4.2.1 Ability to detect infected herds in the whole population

With the cut-off based on the Youden index to select the "best" cut-off to classify the whole population, the model accurately classified more infected herds in 152 of the 172 scenarios compared to test results alone (Figure 4.5). The difference between the model and test results varied from 125 fewer to 509 additional infected herds detected. On average the model detected an additional 105 truly infected herds. This represented a proportion of infected herds additionally detected by the STOC free model from -0.085 to 0.358, with a mean value of 0.110, corresponding to the added value in sensitivity of the surveillance scheme provided by the model (Figure 4.6). For all scenarios with herd test

sensitivity ($Se$) of 0.4 and 0.7, the STOC free model detected more infected herds than the test results (Figure 4.6). For 12 out of 34 scenarios with low sensitivity, the STOC free model detected an additional 0.3 proportion of infected herds than the test, with a mean value of 0.258. Conversely, when sensitivity was high (0.9) the mean value of additional proportion of infected herds was 0.022. Additionally, for all but two scenario with a herd test specificity ($Sp$) of 0.95, the STOC free model detected more infected herds than the test (Figure 4.6). The proportion of herds additionally detected was similar whatever the values of the infection dynamics parameters (prevalence ($P$), incidence rate ($\tau_1$), and probability to remain infected ($\tau_2$)) and RF link parameters (frequency ($F$) and relative associated risk ($\gamma$)) (Figure 4.6).

TABLE 4.3 – Number of scenarios that converged depending on each value of the simulated parameters.

| Parameter | Value | Initial number of scenarios | Number of scenarios that converged |
|---|---|---|---|
| $Se$ | 0.4 | 72 | 34 |
| | 0.7 | 72 | 66 |
| | 0.9 | 72 | 72 |
| $Sp$ | 0.8 | 108 | 74 |
| | 0.95 | 108 | 98 |
| $P$ | 0.1 | 108 | 91 |
| | 0.3 | 108 | 81 |
| $\tau_1$ | 0.0111 | 54 | 45 |
| | 0.0278 | 54 | 46 |
| | 0.0429 | 54 | 38 |
| | 0.1071 | 54 | 43 |
| $\tau_2$ | 0.75 | 108 | 89 |
| | 0.9 | 108 | 83 |
| $F$ | 0.1 | 72 | 62 |
| | 0.25 | 72 | 57 |
| | 0.5 | 72 | 53 |
| $\gamma$ | 1.5 | 72 | 61 |
| | 2 | 72 | 59 |
| | 5 | 72 | 52 |

### 4.2.2 Classification of uninfected herds

With the cut-off based on the Youden index, the number of herds classified as false positives increased in 126 scenarios with the model (Figure 4.5). Only 27 of the 172 scenarios had a higher number of both infected and uninfected herds that were accurately classified. They were mainly associated with medium and high values of sensitivity (0.7, 0.9), the lowest value of specificity (0.8) and the highest value of probability of remaining infected ($\tau^2$) (0.9).

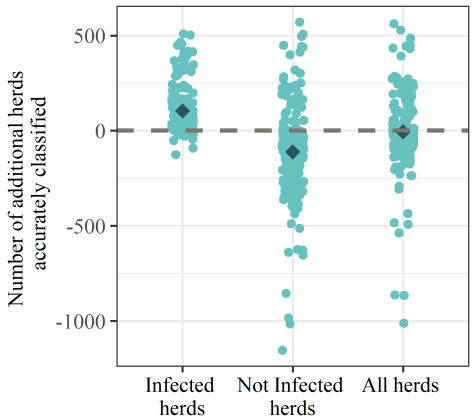



FIGURE 4.5 – Difference between the number of herds accurately classified by the STOC free model and the number of herds accurately classified using test results for infected herds only, for uninfected herds and for all herds. Dark blue diamond represents the mean of each distribution. At the dashed grey line, the STOC free model and test results accurately classified the same numbers of herds.

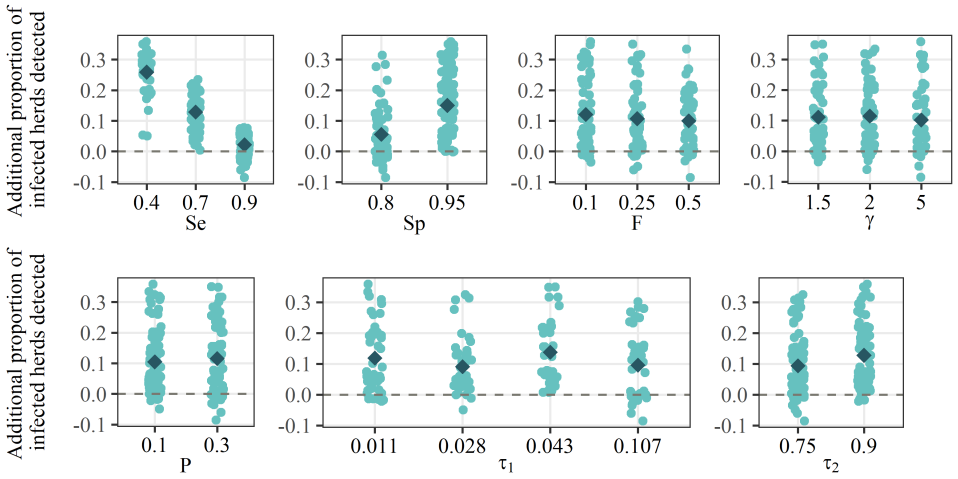



FIGURE 4.6 – Additional proportion of infected herds accurately classified by the STOC free model relative to test results, among the total number of infected herds, depending on simulated parameter values, using cut-off found applying Youden index. The seven simulated parameters are represented : $Se$ (test sensitivity), $Sp$ (test specificity), $F$ (frequency of the risk factor), $\gamma$ (relative risk associated with the risk factor), $P$ (prevalence), $\tau_1$ (probability of being newly infected) and $\tau_2$ (probability of remaining infected). Dark blue diamond represents the mean of each distribution. At the dashed grey line, the STOC free model and test results accurately classified the same numbers of herds.

### 4.2.3 Ability to detect newly infected herds among candidates to new infection

#### 4.2.3.1 Using Youden index

With the cut-off based on the Youden index, the STOC free model accurately classified more newly infected herds in 65 scenarios compared to the test results (Figure 4.7). The difference between the model and test results varied from 82 fewer to 88 more newly infected herds detected. On average, the model detected 5 fewer herds than the test. This corresponded to a proportion of newly infected herds additionally detected by the STOC free model from -0.603 to 0.370, with a mean value of -0.046 (Figure 4.8). Interestingly, for all scenarios with herd test sensitivity of 0.4, the STOC free model detected more newly infected herds than the test results, with the additional proportion of newly infected herds detected ranging from 0.008 to 0.370 (Figure 4.8). For 48 of the 98 simulated scenarios with a herd test specificity of 0.95, the model detected more truly newly infected herds than the test alone.

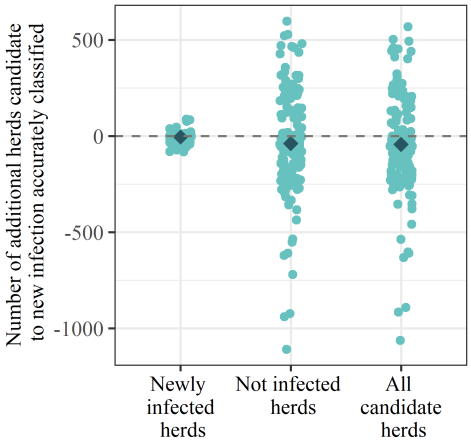



FIGURE 4.7 – Difference between the number of herds accurately classified by the STOC free model and the number of herds accurately classified using test results only for herds which were candidates for new infection at the final time-step (i.e. herds that were uninfected at the previous step time) for newly infected herds, uninfected herds and all herds, using cut-off found applying Youden index. Dark blue diamond represents the mean of each distribution. At the dashed grey line, the STOC free model and test results accurately classified the same numbers of herds.

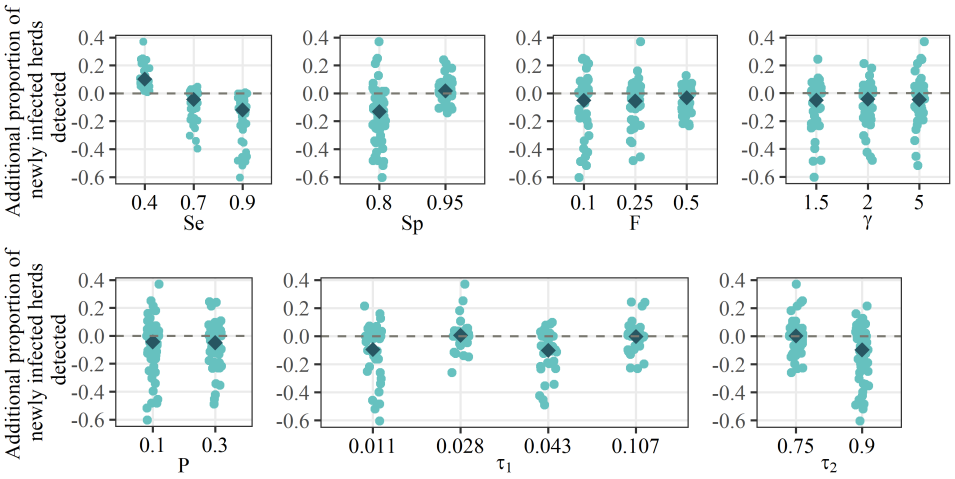



FIGURE 4.8 – Additional proportion of newly infected herds detected by the STOC free model relative to test results, among the total number of newly infected herds, depending on simulated parameter values, using cut-off found applying Youden index. The seven simulated parameters are represented : $Se$ (test sensitivity), $Sp$ (test specificity), $F$ (frequency of the risk factor), $\gamma$ (relative risk associated with the risk factor), $P$ (prevalence), $\tau_1$ (probability of being newly infected) and $\tau_2$ (probability of remaining infected). Dark blue diamond represents the mean of each distribution. At the dashed grey line, the STOC free model and test results accurately classified the same numbers of herds.

#### 4.2.3.2 Using NewI cost index

We developed a new index to select cut-off values, with the constraint to detect at least one more newly infected herd compared to the test. For 13 of the 172 scenarios, no cut-off value allowed the detection of at least one additional newly infected herd. For all the 159 remaining scenarios, using this index allowed the detection of an additional proportion of newly infected herds, ranging from 0.003 to 0.429, with a mean value of 0.071 (Figure 4.9). This corresponded to the detection of 1 to 156 additional newly infected herds with a mean value of 14 herds. In 24 scenarios, the proportion of additional newly infected herds that were detected was higher than 0.15 (Figure 4.9). By construction, the test sensitivity value limits the potential number of additional newly infected herds that can be detected by the model (e.g. with a sensitivity of 0.9, the maximum potential proportion of newly infected herds additionally detected is 0.1). On average, the model captured proportions increased by 0.125, 0.076, and 0.034 for sensitivity values of 0.4, 0.7 and 0.9, respectively (Figure 4.9).

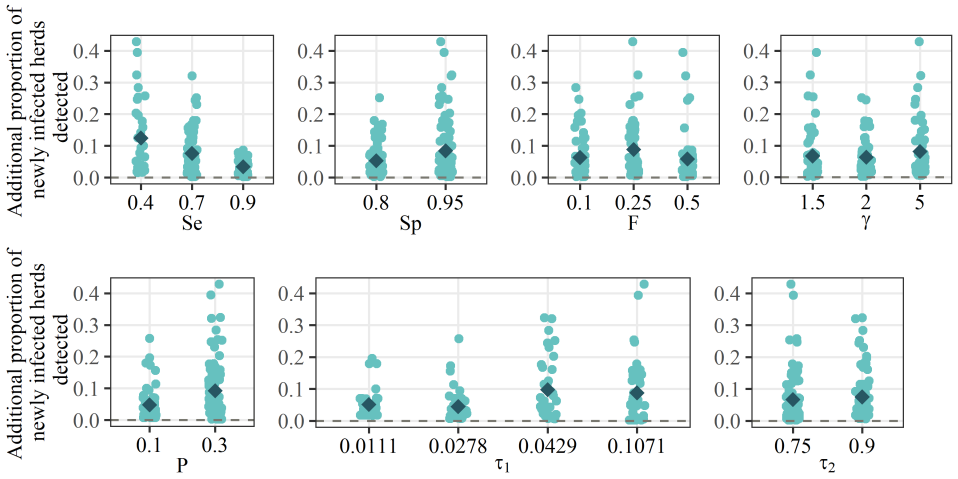



FIGURE 4.9 – Additional proportion of newly infected herds detected by STOC free model relative to test results, among the total number of newly infected herds, depending on simulated parameter values, using cut-off found applying NewI cost index. The seven simulated parameters are represented : $Se$ (test sensitivity), $Sp$ (test specificity), $F$ (frequency of the risk factor), $\gamma$ (relative risk associated with the risk factor), $P$ (prevalence), $\tau_1$ (probability of being newly infected) and $\tau_2$ (probability of remaining infected). Dark blue diamond represents the mean of distribution. At the dashed grey line, the STOC free model and test results accurately classified the same numbers of herds.

Using the NewI cost index, the cut-off value allows systematically for a better detection of newly infected herds compared to test results but is associated with a cost in false positives. Only 3 scenarios had a negative cost index, whereby it was able to detect more newly infected herds while having less false positives (Figure 4.10). For all the other scenarios, the additional detection of newly infected herds was always associated with a positive NewI cost index, i.e. a number of additional false positives for each additional true positive detected (Figure 4.10). This NewI cost index ranged from - 266 to 1055. On average, the cost index value was 98 meaning that for each additional newly infected herd detected, there were an additional 98 false positive herds compared to test results. NewI cost index was <100 for 73% of the scenarios (116/159) (Figure 4.10). Extremely high values of the cost index (above 500) were associated with a sensitivity of 0.9 for 5 scenarios (Figure 4.10). These extreme values were also associated with lower proportions of additionally detected newly infected herds (Figure 4.11.A). When the proportion of herds additionally detected was above 0.1, the cost index was <100 except in three (Figure 4.11.A). All scenarios (43) with a high number of newly infected herds (corresponding to $\tau_1$=0.107) had a NewI cost index below 100 (Figure 4.10 and Figure 4.11.B).

### 4.2.4 Comparison of cut-off values

The cut-off values varied substantially between scenarios for both indexes (Figure 4.12). Use of the Youden index resulted in higher cut-off values (mean cut-off equal 0.14 against 0.05 for cost index) (Figure 4.12). No association between input parameter values (test characteristics, disease dynamics and risk factors parameters) and selection of a cut-off value was found (supplementary material).

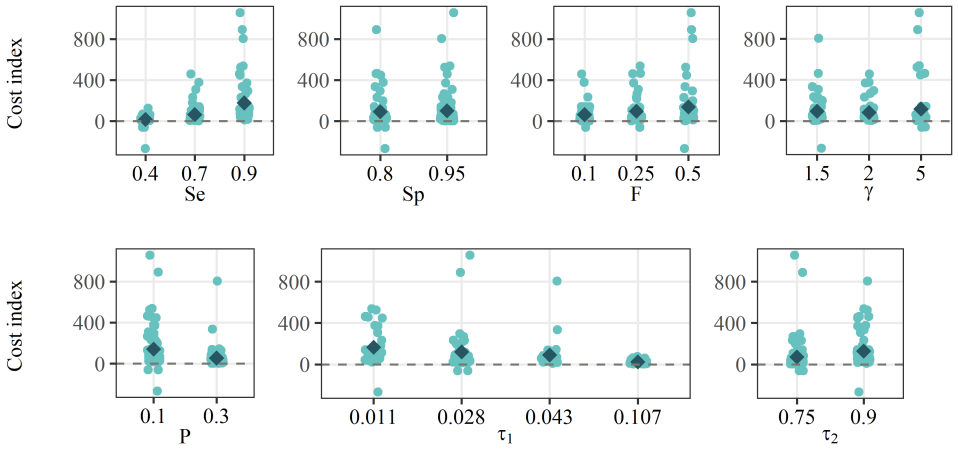



FIGURE 4.10 – Cost index value, i.e. the number of additional false positive herds for each additional true positive herds by the STOC free model relative to test results, depending on simulated parameter values, using cut-off found applying NewI cost index. The seven simulated parameters are represented : $Se$ (test sensitivity), $Sp$ (test specificity), $F$ (frequency of the risk factor), $\gamma$ (relative risk associated with the risk factor), $P$ (prevalence), $\tau_1$ (probability of being newly infected) and $\tau_2$ (probability of remaining infected). Dark blue diamond represents the mean of distribution. Under the dashed grey line cost is negative meaning that STOC free model do detect more newly infected for less false positive herds compared to test results.

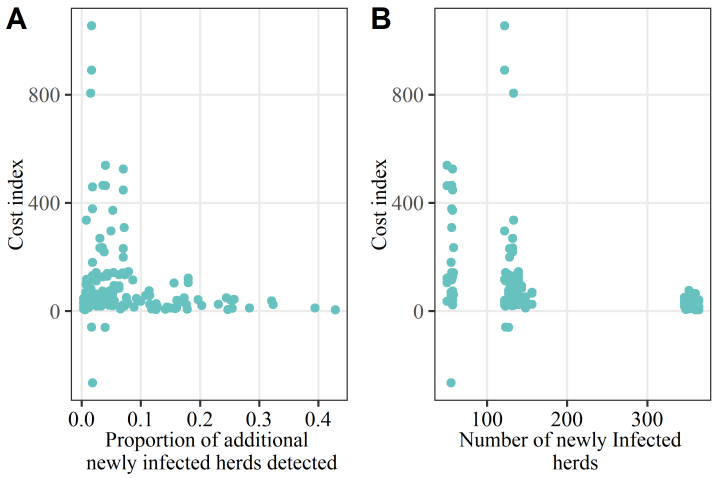



FIGURE 4.11 – Cost index value, i.e. the ratio of additional false positive herds on the additional true positive herds, using cut-off found applying NewI cost index, depending on the proportion of additional newly infected herds detected (A) and the number of newly infected herds which depend on the four possible values of the probability of become infected ($\tau_1$).

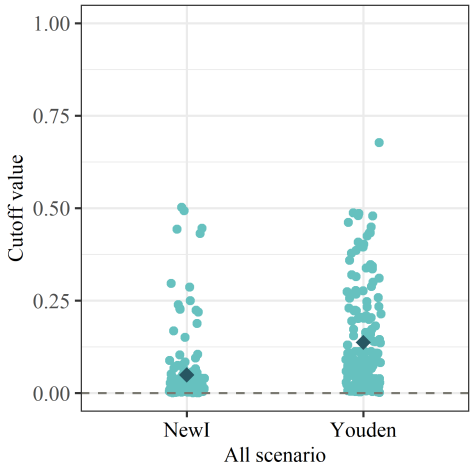



FIGURE 4.12 – Distribution of cut-off values for herd status classification depending on the index used. "NewI" criterion is based on a trade-off between the additional number of true positive herds and the additional number of false positive herds. "Youden" criterion is based on maximizing sensitivity and specificity based on classification of all herds. Dark blue diamond represents the mean of the distribution.

# 5 Discussion

Our simulation study illustrates the added value of a Bayesian Hidden Markov model, the STOC free model, compared to test results alone to detect infected herds in many different contexts. This model was able to predict herd-level probabilities of infection in about 80% of the investigated scenarios. Situations in which the model did not converge and therefore could not provide estimates of the probabilities of infection were mainly related to low sensitivity values. When it converged, the model detected more infected herds compared to test alone in 152 of the 172 scenarios and detected more newly infected herds in only 65 of 172 scenarios. In these scenarios, the STOC free model sensitivity was higher than the herd-level test sensitivity.

Test sensitivity had a great impact on the added value of the STOC free model. Indeed, following a test, the total number of infected and newly infected herds still to be detected (false negatives) increases as test sensitivity decreases. The STOC free model was able to detect an important proportion of these undetected infected herds. On average, the model detected around 25% more infected herds when sensitivity was low (0.4) and around 2% more infected herds when the sensitivity was high (0.9), i.e. around 40% and 20% of the

herds still to be detected in our simulations (as assumed for the given levels of sensitivity). The range of herd-level test sensitivities evaluated in this study covers the known range of sensitivities for endemic diseases for which control programmes are in place.

An increase in the number of newly infected herds detected by the STOC free model was associated with an increase in the number of false positive herds detected in all but one scenario. We quantified the proportion of additional false positives for each additional newly infected herd detected using a cost index. This cost index increased with high test sensitivity and low prevalence corresponding to small numbers of false test negatives. In five scenarios with a high herd-level test sensitivity, the cost index was substantial (above 500, i.e. 500 false positives for each additional true positive herd detected by the model). This tends to advise against using the STOC free model when test sensitivity is high. On the other hand, the cost index was lower (below 100) with low test sensitivity and high incidence, i.e. when the number of newly infected herds still to be detected was high. For decision support, the level of acceptability in terms of extra false positives would differ according to the consequences in a given control programme, and to the possibilities and resources necessary to confirm a herd status with complementary testing.

Different reasons could explain the fact that the STOC free model did not reach convergence in a number of scenarios. In this study, we limited the number of burn-in and sampling iterations to reduce computing time (around 3.5 hours per scenario). For scenarios that did not meet our convergence criterion, re-running the model with more burn-in iterations allowed convergence in around 50% of cases. Adding more iterations could address the remaining convergence issues. A larger population (number of herds) would increase available data (especially in terms of numbers of infected herds) to estimate parameters values. We did not further investigate these hypotheses due to computing time constraint for both simulation and analysis. Low test performance also led to convergence issues. Indeed, as test sensitivity and specificity decrease, the contribution of test results to defining the latent status decreases whereas the contribution of model parameters accounting for new infection and elimination of infection increases. In our study, given the relatively wide prior distributions put on the association between the risk factor and the probability of new infection, this association was estimated from the data. This means that in the scenarios in which test performance was poor, the contribution of surveillance data to estimation and prediction could be expected to be small, which could have made it more difficult for the model to converge. Such estimation issues have already been described in state-space models, when measurement error is high (Auger-Méthé *et al.*, 2016). In such

cases, increase the sample size (e.g. the number of herds) or adding prior information could reduce this issue. In our study, informative priors were used for measurement error parameters (sensitivity and specificity) assuming that relevant epidemiological quantities would be known beforehand. To decrease convergence issue, it could be hypothesised that a good knowledge of the strength of association between risk factors and the probability of new infection facilitates convergence by reducing uncertainty around latent statuses. This knowledge would need to be translated into narrow prior distributions.

The frequency and strength of the risk factor did not influence the STOC free model performances, contrary to our assumptions. The inclusion of RFs was expected to improve the detection of newly infected herds when they strongly contribute to the risk of new infections (high strength of association). This added value was especially expected to be important when test sensitivity is poor, because knowing that a RF is present could compensate for the lack of sensitivity. Here, the added value that was evidenced didn't show variations associated with strength or frequency of the RFs, in the range of values investigated which was chosen according to range values for known RF (van Roon *et al.*, 2020a). In our study, only one RF was included to establish its influence on model performance. More RFs can easily be added to the logistic regression if necessary. The choice of RFs to be included must be based on specific knowledge of infection dynamics within the CP.

A cut-off value is needed to classify herds as infected or uninfected from the distributions of probabilities of infection predicted by the STOC free model. The cut-off value varied depending on the method of selection and the simulated context. In the field, the "best" cut-off value would also depend on the objective of the CP. The Youden index equally values sensitivity and specificity without other constraint (i.e. separates at best infected versus non infected herds), while our NewI cost index ensures the detection of a higher number of newly infected herds than the test alone. For most scenarios, the cut-off value identified with the cost index was lower than the cut-off value identified with the Youden index. Indeed, given an endemic situation, the probability of becoming infected is lower than the probability of remaining infected. The specific detection of newly infected herds, that have not been detected by the test, requires a lower cut-off value compared to the cut-off value selected without this constraint. To compute our cost index associated with the detection of a higher number of newly infected herds we gave the same weight to false positives and false negatives. These two types of misclassification have different consequences : in the context of cattle trade, introducing a false negative into a disease

free herd is more damaging than not allowing a false positive to be introduced. Whatever the method of selection used, cut-off values were highly variable between the simulated contexts. According to our study, it does not seem possible to determine a cut-off value directly from CP characteristics. However, we can argue that low cut-off values should be favoured where the objective is safe trade, i.e. limiting false negative herds. In real data where no gold standard is available, the choice of the cut-off value has to rely on another method. This point is an important question when this framework is applied to real data and it needs more exploration.

Applying the STOC free model to real CPs also requires previous knowledge about the distributions of the model parameters. The choice of prior distributions will be crucial because when the prior distributions deviate too much from the true parameter values, this may lead to convergence issues or bias in the posterior distributions. In this simulation study, true parameter values were known, allowing prior distributions to be centred on the true parameter values. In the context of real CPs, test characteristics are almost always assessed before designing the CP. However, even if information is often available, its interpretation must be made in relation to the targeted latent status which may differ from the definition used in the literature and can be challenging (Duncan *et al.*, 2016). Test characteristics may change depending on the latent status of interest. Information on risk factor of introduction is often available as controlling them is a key measure in CPs to reduce the spread of infection between herds (Lindberg & Houe, 2005). Quantitative data can be derived from the literature (e.g. risk factor study, meta-analysis) but are highly variable between territories and not always available for a specific territory (*blue*van Roon et al., 2020a). The model makes it possible to use more or less precise priors according to the available information in the population of interest.

Within a CP, the dynamics of the infection (incidence and clearance of infection) as well as the contribution of risk factors are expected to change over time given that the majority of CPs generally act on both preventing new infections and eliminating the pathogen from infected herds. Depending on the CP, these changes may be observed over different periods of time. Example of CPs against BVDV have shown that the decrease in prevalence and incidence in European countries occurred over different time lapses (Houe *et al.*, 2014; Presi *et al.*, 2011; Joly *et al.*, 2001). Risk factor contribution (frequency and strength) may also change during a CP. For example, neighbourhood risk of introduction is linked to infection prevalence in the area. When the prevalence decreases in the territory, the strength of association between having contact with neighbouring herds and

becoming infected will decrease, while the frequency of contacts between herds remains the same. In our study, infection dynamics and the contribution of the risk factors remained stable over time to simplify parameter estimations. The changes in infection dynamics and contribution of RFs to new infections could be accommodated by running the model over short time periods (e.g. 1 to 3 years), using the parameter posterior distributions for one period as the prior distributions for the next one.

The decrease of infection prevalence and incidence with time during a CP can influence performances of the STOC free model. Here, the cost index was higher when incidence and prevalence were low, reflecting a lower positive predictive value, when the number of true positive herds decreases in a population (similarly to surveillance based on tests only). Therefore, we speculate that the use of the STOC free model will be more interesting with disease present at an endemic level in a population rather than when CP results in decreased prevalence close to eradication.

# 6 Conclusion

This simulation study demonstrated the capacity of a Hidden Markov Model using disease dynamics and risk factor information from surveillance programmes to detect more infected herds and newly infected herd than test results alone. The added value of the model depends on the context in which a control programme is conducted. It was greatest in situations with low sensitivity tests. However, these situations were also the ones in which the convergence of the model was the most difficult. The added value of the model did not depend on the strength and frequency of the risk factor. The use of the model is likely to be beneficial especially in the early stages of a control programme (when prevalence and incidence are at moderate level) rather than close to eradication.

Chapitre 5

# Evaluation of rules for the categorisation of herds as *infected* from probability distributions predicted by the STOC free model

---

# 1 Introduction

In CPs, decisions regarding individual herds are made based on whether they are considered *infected* or *not infected*. The STOC free model was designed to permit such a categorisation into *infected* or *not infected* from surveillance data. The model outputs needed for this categorisation are posterior distributions of infection at the herd-level (Madouasse *et al.*, 2021). Based on these posterior distributions, some rules are required for the categorisation. Such rules involve summarizing each posterior distribution with a single value and then to apply an appropriate cut-off value on this summary. Herds with a summary value below the chosen cut-off will be categorised as not infected and those with a summary value above the cut-off as infected. If a gold standard were available, these rules could be determined so as to achieve some acceptable level of error such as a pre-specified probability of wrongly categorising an infected herd as infection, i.e. one minus the negative predictive value. However, a gold standard is usually not available to define the categorisation rules.

In the study presented in the previous chapter, a gold standard was available in the form of a simulated herd status. In this study, herds were classified as *infected*/*not infected* using the median value of each herd posterior distribution as a summary. The median was used because it appeared to be the variable that discriminated best uninfected from infected herds. However, when we explored the use of different percentiles to categorise herds, we observed variation between scenarios. The cut-off values were chosen using the classical Youden index (Youden, 1950) or a cost index specifically built for the purpose of the study. Both of these indexes are based on the availability of a gold standard. To apply the STOC free model on field data for which true herd statuses is never known, the choice of the cut-off value is challenging.

The aim of this preliminary study was to evaluate the effect of choosing different rules for the categorisation of herds as *infected* / *not indected* in different contexts on the performance of the categorisation. A secondary aim was to determine whether such rules could be determined from the characteristics of the disease or of the test used in the absence of a gold standard.

# 2 Materials and Methods

## 2.1 Design strategy

The effect of using different combinations of distribution summaries and cut-off values on categorisation performance was evaluated. Simulated data and model results obtained for some scenarios from the previous study (ChapterIV) were used. These scenarios represented different test performances or disease dynamics. The principles of the categorisation were the following : first, each posterior distribution was summarised using a percentile. Then, a cut-off value was selected. If, for a given herd, the distribution percentile was below the cut-off, the herd was categorised as *no infected*. If the distribution percentile was above the cut-off value, the herd was categorised as *infected*. From the true (simulated) status, each herd could then be labelled as a true/false negative/positive which allowed computing different performance measures. The process was repeated for a wide range of combinations of percentiles and cut-off values.

## 2.2 Categorisation of herds from predicted probabilities of infection

In the previous chapter, we saw that the posterior distributions for the herd-level probabilities of infection predicted by the STOC free model were mainly bi-modal (Figure 5.1). This bi-modality is due to the fact that, for each time step before the time of prediction ($t < T$), the infection status is represented as a binary event. Then, at the last time step $T$, infection is predicted as a probability that can take any value between 0 (certain absence of infection) and 1 (certain presence of infection) depending on the infection status at the previous time step. If the status predicted at the time step before the time of prediction ($T-1$) was *not infected*, the herd had a certain probability of new infection modelled with $\tau_1$. If the status predicted at the time step before the time of prediction was *infected*, the herd had certain probability of remaining infected modelled with $\tau_2$.

As we focused on endemic diseases, the probability of remaining infected ($\tau_2$) was greater than the probability of becoming infected ($\tau_1$). Thus, the peak on the left of the posterior distribution of the probability of being infected on the figure 5.1 (closest to zero) corresponded mainly to iterations for which the estimated status at $T-1$ was not

infected. While the peak on the right (closest to one) of the distribution corresponded to iterations for which the estimated status at $T-1$ was infected (Figure 5.1).

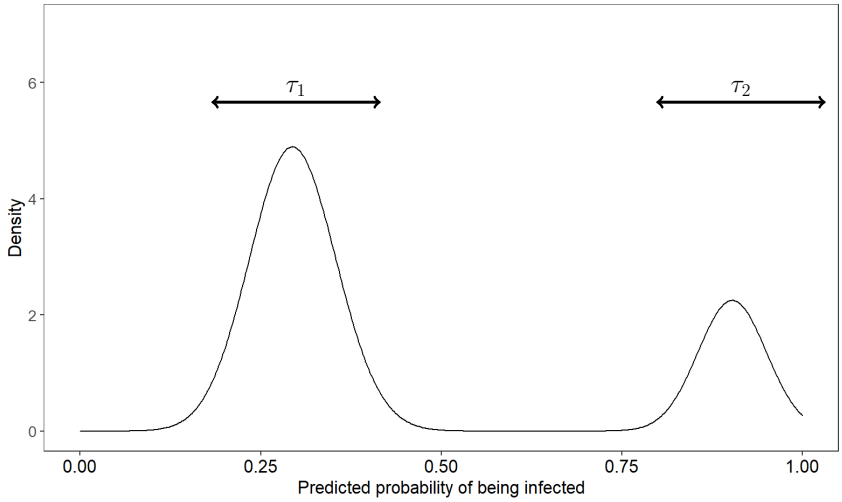



FIGURE 5.1 – Example of a posterior distribution of the probability of being infected for one herd at time $T$. $\tau_1$ represents the probability of becoming infected when the herd is not infected at $T-1$, $\tau_2$ represents the probability of remaining infected when the herd is infected at $T-1$.

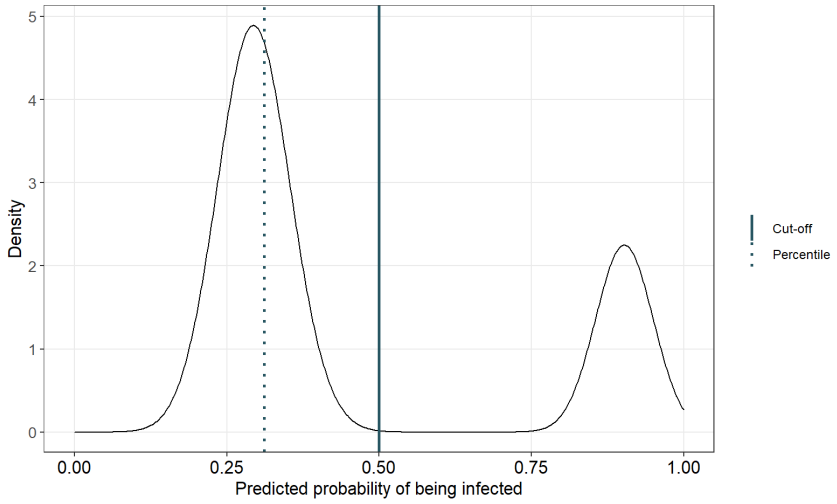



FIGURE 5.2 – From the posterior distribution of being infected to herd status.Posterior distribution of being infected for one herd, dashed line represent the percentile 50 of the distribution and solid line the cut-off value of 0.5. In this example the herd is classified *not infected* by the model.

In the evaluation, a sequence of 101 percentiles equally spaced between 0 to 100 were used (0, 1, 2..., 100) as well as a sequence of 101 cut-off values equally spaced between 0 and 1 (0,0.01..., 1). A grid of all possible combinations of percentiles and cut-off values was constructed. Figure 5.2 shows how a herd is classified as infected or not infected from its predicted posterior probability of infection, a percentile and a cut-off value.

## 2.3 Performance of the categorisation

The performance of the model and subsequent categorisation were computed for all combinations of percentiles and cut-off values and for every tested scenario. For each combination, the added value of the model compared to using test results only was computed by calculating :

- the difference between the number of infected herds detected (i.e. true positives) by the model and by the test,
- the difference between the number of uninfected herds wrongly classified as infected (i.e. false positives) by the model and by the test.

Then, we selected all the combinations of percentiles/cut-off values, for which :

- the number of infected herds detected by the model was higher than the number detected by the test,
- at least one herd was categorised as uninfected.

For these selected combinations, we computed a cost index as :

$$cost\,index = \frac{FP_{STOCfree} - FP_{test}}{TP_{STOCfree} - TP_{test}}$$

where the numerator was the difference in the number of false positives predicted by the model and the number of test false positives to the test, and the denominator was the difference in the number of true positives predicted by the model and the number of true positives to the test.

## 2.4 Selected scenarios

In the previous study, the distribution of cut-off values selected by the two different methods seemed to vary with some simulation parameters : test characteristics (sensitivity ($Se$) and specificity ($Sp$)) and disease dynamics (prevalence ($P$), probability of becoming

infected ($\tau_1$) and probability of remaining infected ($\tau_2$) (Chapter IV, supplementary material). On the contrary, risk factor related parameters (relative risk ($\gamma$) and frequency ($F$)) did not seem to have an impact on the cut-off value.

For this study, scenarios were selected from the simulations run in Chapter IV to explore the impact of the parameters of interest. Scenarios with different test sensitivities ($Se$), different test specificities ($Sp$) or different prevalences ($P$) were selected. $P$ was selected to represent the impact of disease dynamics on the choice of cut-off and the percentile.

Moreover, prevalence, sensitivity and specificity values have a great impact on the number of infected and uninfected herds as well as on the number of herds accurately classified by the test (Table (5.1 and Figure 5.3). In our simulation study, the number of herds was set at 5000 for all scenarios. Prevalence and test sensitivity values determined the number of infected herds yet to be detected (i.e. false negatives) (Figure 5.3). Prevalence and test specificity values determined the number of uninfected herds that tested positive (i.e. false positives).

According to these criteria, we selected :

— 3 scenarios with 3 different sensitivity values, with $Sp = 0.95, prev = 0.3, \tau_1 = 0.107, F = 0.1, \gamma = 1.5$
— 2 scenarios with 2 different specificity values, with $Se = 0.7, prev = 0.3, \tau_1 = 0.107, F = 0.1, \gamma = 1.5$
— 2 scenarios with 2 different prevalence values, with $Se = 0.7, Sp = 0.95, F = 0.1, \gamma = 1.5$

A high prevalence was preferred to have the highest number of infected herds possible. A medium sensitivity was chosen as a compromise. Indeed, with a very high sensitivity, the model performance was limited (Chapter IV) but a low test sensitivity is not frequent in CPs. A high specificity was chosen because test specificity is usually good since positive herds are usually re-tested. Finally, parameter values for RFs were chosen arbitrarily and fixed for every scenario, as these parameters did not seem to have an impact on the performances of the model in Chapter IV.

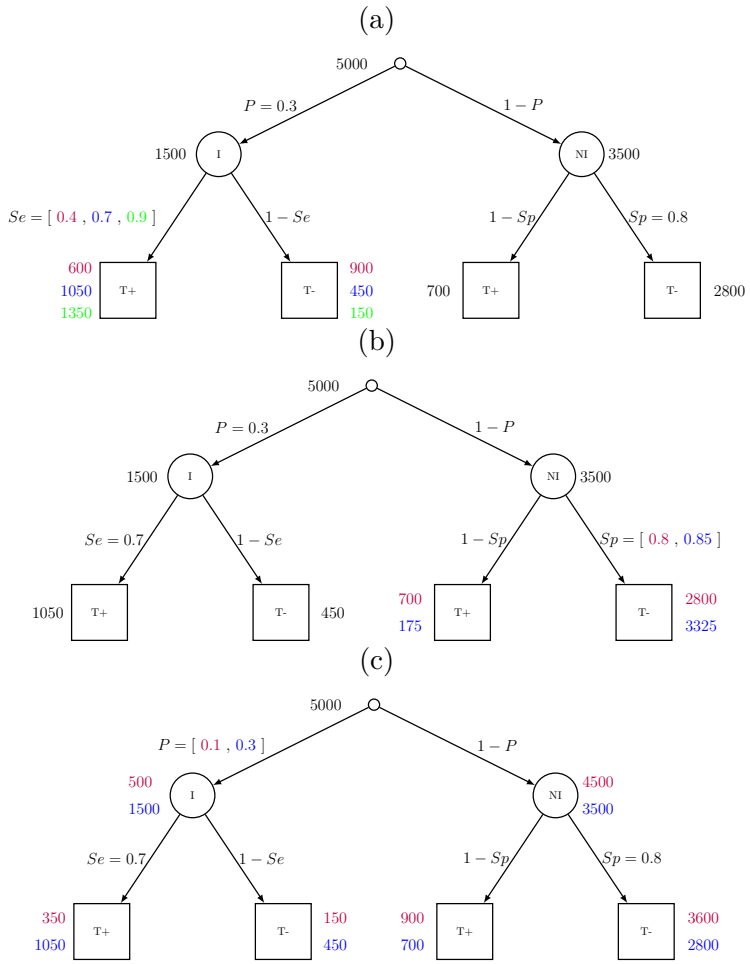



FIGURE 5.3 – Probability trees for different simulated scenarios with the numbers of infected (I) and uninfected (NI) herds, with a positive (T+) or a negative (T-) test result depending on the prevalence (P), the herd level specificity (Sp) and sensitivity (Se) with different values of sensitivity (a), specificity (b), and prevalence (c).

TABLE 5.1 – Probability of test results depending on herd status, with $P$ being the infection prevalence, $Se$ the test sensitivity and $Sp$ the test specificity.

| | | *Herd status* | |
|---|---|---|---|
| | | *Infected* | *Not infected* |
| *Test* | + | $(Se)P$ | $(1-Sp)(1-P)$ |
| | − | $(1-Se)P$ | $Sp(1-P)$ |

# 3 Results

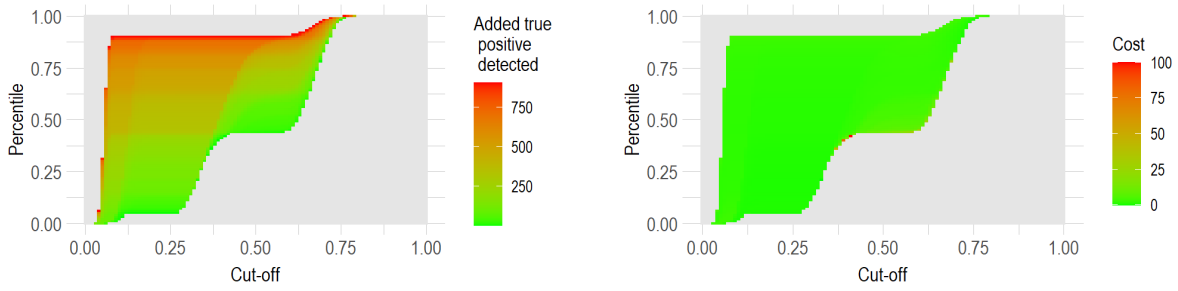
## 3.1 Sensitivity

The proportion of percentile/cut-off combinations for which the number of infected herds detected was higher with the model than when using test results alone, without classifying all herds as infected, varied depending on test sensitivity (Figure 5.4). Around 40% of all combinations met both criteria when sensitivity was low or medium (respectively 4081/10201 and 4027/10201 combinations for sensitivity values of 0.4 and 0.7), but only 20% (2056/10201) when sensitivity was high (0.9). Although the patterns of excluded combinations varied depending on sensitivity value (grey area in Figure 5.4), it can be noted that the range of percentile values associated with a higher number of infected herds detected and a smaller cost was larger than the corresponding range of cut-off values.

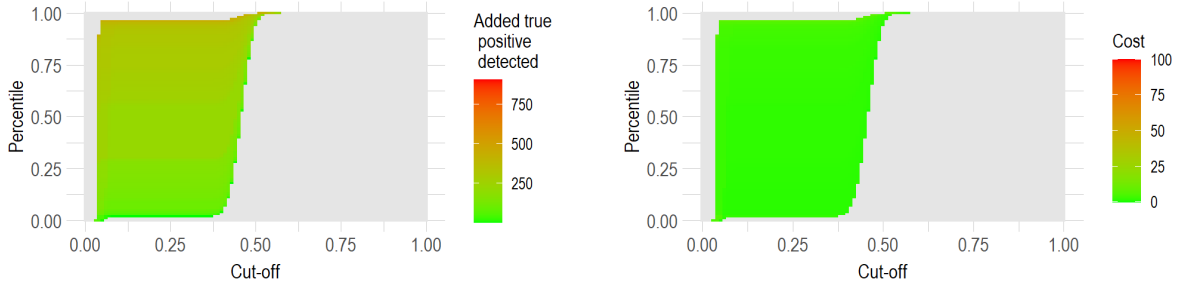
The median number of additionally detected infected herds decreased with sensitivity, respectively 341, 222 and 85 for sensitivities of 0.4, 0.7, and 0.9 (Figure 5.5a and Figure 5.4a, 5.4c and 5.4e). The maximum number of additional infected herds detected are observed for lower sensitivity. The maximum number of herds additionally detected by the model number was respectively 905, 434 and 122 for low, medium and high sensitivity. For low sensitivity, maximum number of additional infected herds detected were associated with percentiles and cut-off values along the upper border of exclusion (upper grey area Figure 5.4a).

The median cost in false positive of each additional infected herd detected in comparison to the test increases with the sensitivity value, respectively 1.70, 1.83 and 5.50 for a sensitivities of 0.4, 0.7, and 0.9 (Figure 5.5b). Only one combination lead to a cost value above 100, when sensitivity was low.

(a) $Se = 0.4$ (b) $Se = 0.4$




(c) $Se = 0.7$ (d) $Se = 0.7$




(e) $Se = 0.9$ (f) $Se = 0.9$

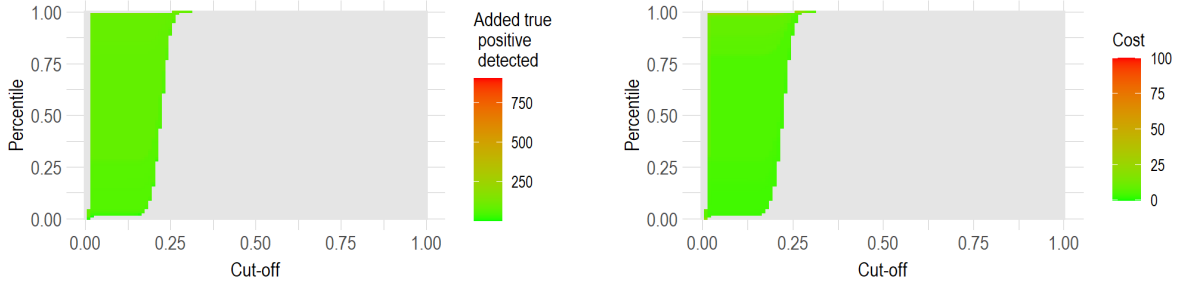



FIGURE 5.4 – Number of additional infected herds detected by the model compared to the test results alone (a, c and e) and cost associated values depending on percentile and cut-off values (b, d and f) depending on test sensitivity values : (a and b) $Se = 0.4$, (c and d) $Se = 0.7$ and (e and f) $Se = 0.9$. *Cost* is the number of additional false positive herds for each additional true positive herd detected by the model compared to test results alone. Shaded areas represent excluded combinations. The green to red gradient represents the cost values. For representation purposes, cost values equal or above 100 are represented in the same colour (bright red).

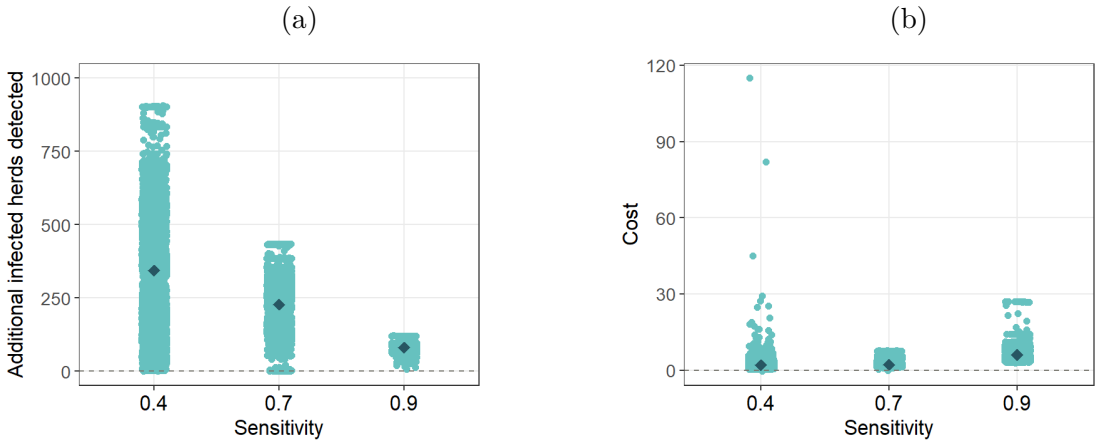



FIGURE 5.5 – Distribution of (a) the number of additional infected herds detected by the model and (b) cost values depending on sensitivity values. *Cost* is defined as the number of additional false positive herds for each additional true positive herd detected by the model relative to test results. Dark blue diamonds represent the median value.

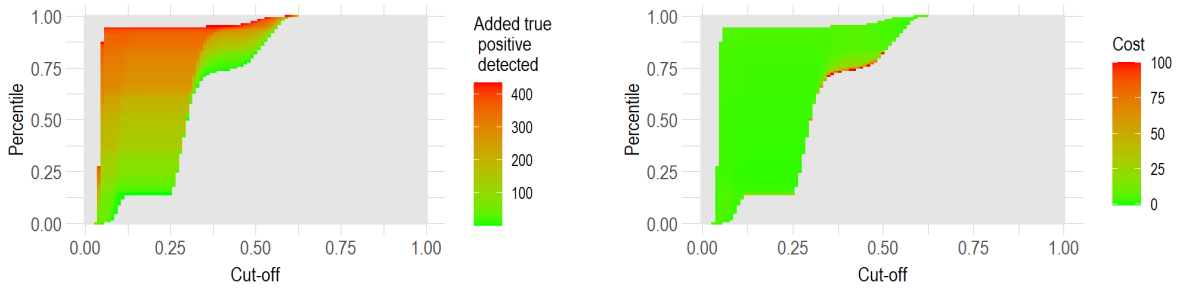
## 3.2 Specificity

The proportion of percentile/cut-off combinations for which the number of infected herds detected was higher with the model than when using the test alone, without classifying all herds as infected, varied depending on test specificity (Grey area in Figure 5.6). For the low specificity value (0.8) 26% (2614/10201) of the combinations were included, while for the high specificity value (0.95), 43% (4388/10201) of the combinations were included.

The median number of additionally detected infected herds was similar for both specificity values, respectively 188 and 222 for specificities of 0.8 and 0.95 (Figure 5.7a). Depending on the percentile used, the additional number of infected herds detected varied a lot for both scenario (Figure 5.7a and 5.7c). Higher percentiles were associated with higher numbers of infected herds detected.

The median cost index value was higher for low specificity, respectively 3.12 and 1.83 for a specificities of 0.8 and 0.95 (Figure 5.7b). Extreme value of cost are observed with lower specificity. They were mainly associated with percentiles around 0.75 and on the border with the excluded combination (Figure 5.6b).

(a) $Sp = 0.8$ (b) $Sp = 0.8$




(c) $Sp = 0.95$ (d) $Sp = 0.95$

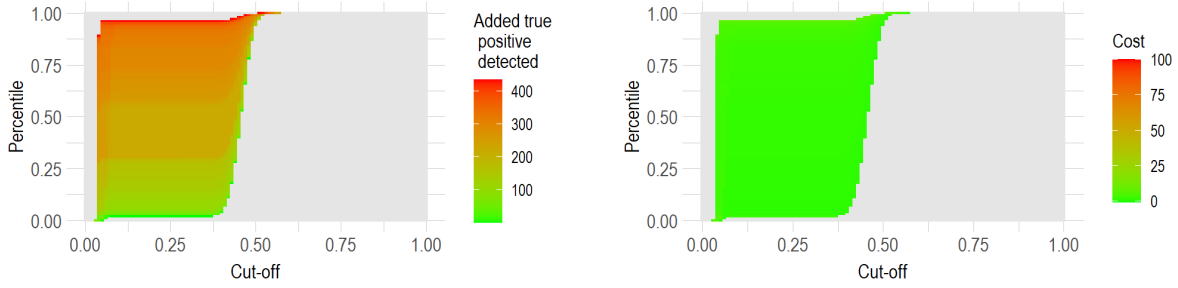



FIGURE 5.6 – Number of additional infected herds detected by the model compared to the test results alone (a and c) and cost associated values depending on percentile and cut-off values (b and d) depending on test specificity values : (a and b) $Sp = 0.8$ and (c and d) $Sp = 0.95$. *Cost* is the number of additional false positive herds for each additional true positive herd detected by the model compared to test results alone. Shaded areas represent excluded combinations. The green to red gradient represents the cost values. For representation purposes, cost values $\geq 100$ are represented in the same colour (bright red).

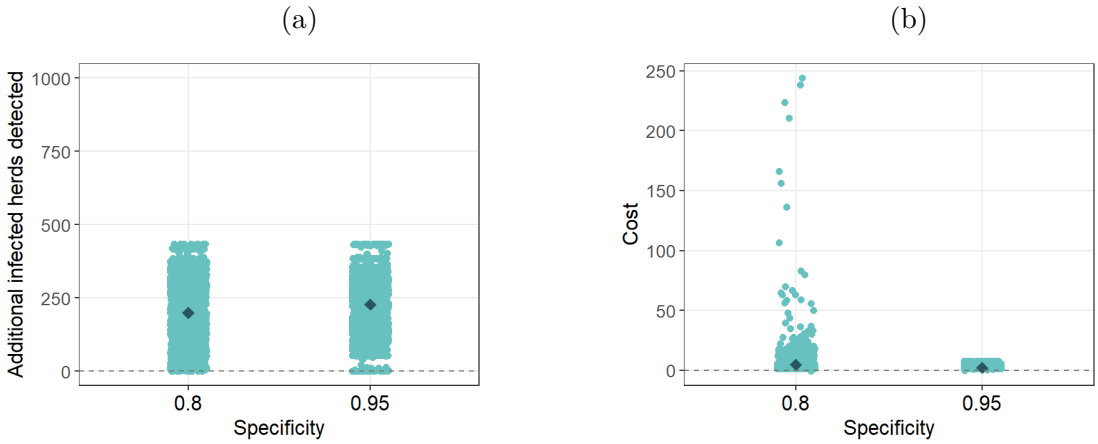



FIGURE 5.7 – Distribution of (a) the number of additional infected herds detected by the model and (b) cost values depending on specificity values. *Cost* is defined as the number of additional false positive herds for each additional true positive herd detected by the model compared to test results. Dark blue diamonds represent the median value.

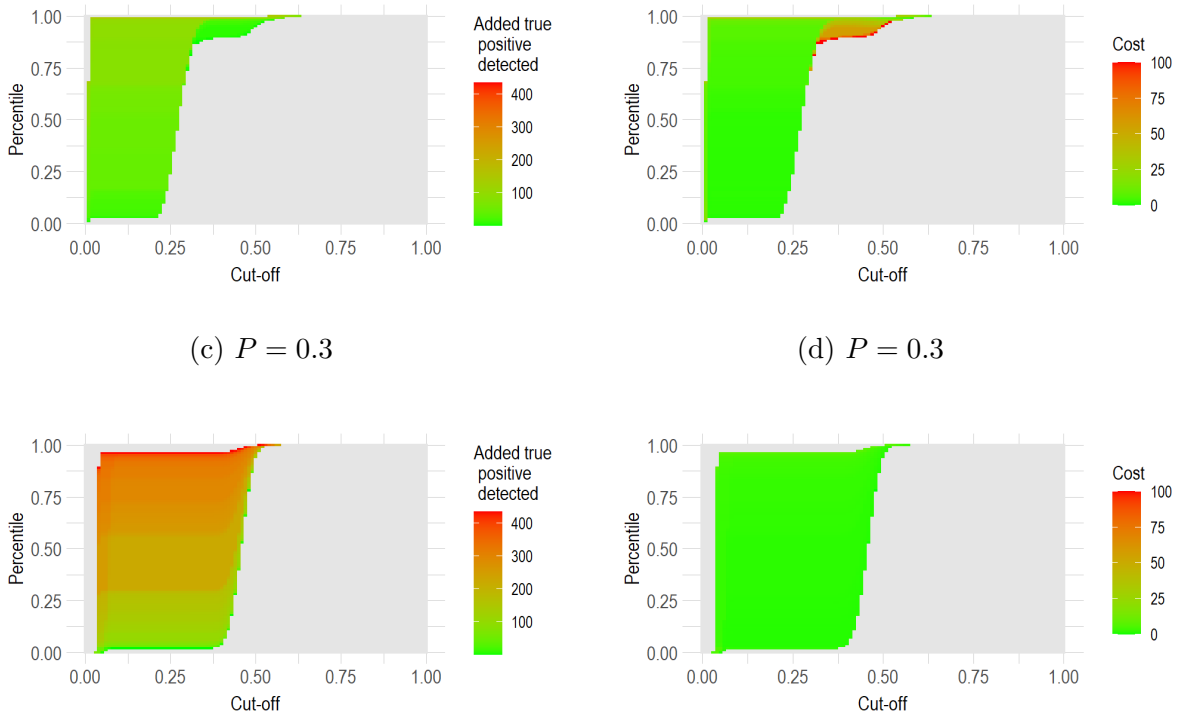
## 3.3 Prevalence

The proportion of percentile/cut-off combinations for which the number of infected herds detected was higher with the model than the test without classifying all herds as infected, varied depending on prevalence values (Grey area in Figure 5.8). For the low prevalence (0.1) 27% (2808/10201) of the combinations were included, while for higher prevalence (0.3), 40% (4027/10201) of the combinations were included.

The median number of additional infected herds detected compared to when using the test alone varied depending on prevalence, respectively 53 and 222 for a prevalence of 0.1 and 0.3 (Figure 5.9a). The additional number of infected herds detected varied a lot with a prevalence of 0.3 depending on combination (Figure 5.8c). Higher additional number of infected herds detected were associated with combination with high percentiles.

The median cost index value was lower for high prevalence, respectively 2.91 and 1.82 for a prevalence value of 0.1 and 0.3 (Figure 5.9b). Extreme cost values were only observed for the lower prevalence.

(a) $P = 0.1$ (b) $P = 0.1$




(c) $P = 0.3$ (d) $P = 0.3$




FIGURE 5.8 – Number of additional infected herds detected by the model compared to the test results alone (a and c) and cost associated values depending on percentile and cut-off values (b and d) depending on prevalence values : (a and b) $P = 0.1$ and (c and d) $P = 0.3$. Cost is the number of additional false positive herds for each additional true positive herds detected by the model compared to test results alone. Shaded areas represent excluded combinations. The green to red gradient represents the cost values. For representation concerns, cost values equal or above 100 are represented in the same colour (bright red).

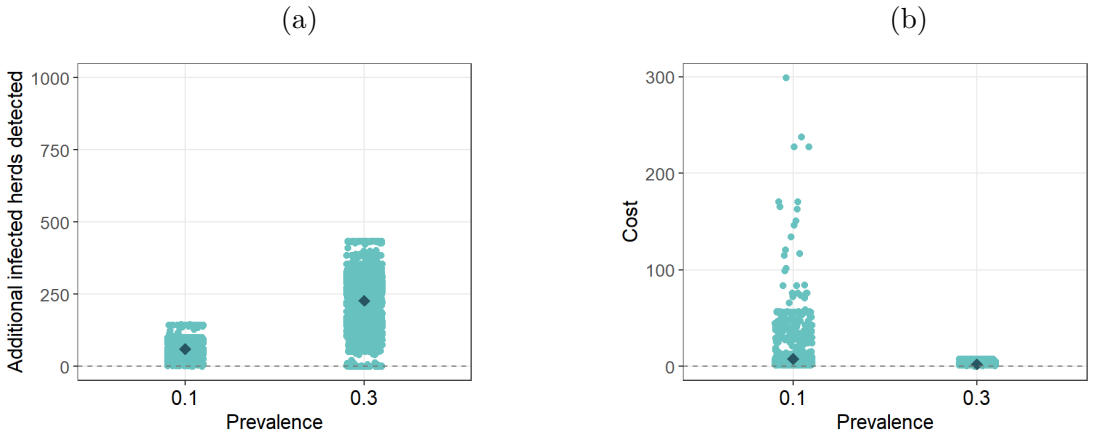



FIGURE 5.9 – Distribution of (a) the number of additional infected herds detected by the model and (b) cost values depending on prevalence values. *Cost* is defined as the number of additional false positive herds for each additional true positive herd detected by the model compared to test results. Dark blue diamonds represent the median value.

# 4 Discussion

In this chapter, the application of different combinations of percentiles and cut-off values to the posterior probability distributions predicted by the STOC free model were evaluated on their ability to correctly categorise herds as *infected/uninfected*. In general, cut-off values above 0.5 were mostly excluded, as in these cases the model did not detect more infected herds than the test. Percentiles greater then 90 were also mostly excluded as they either did not detect any additional herds or all herds were classified as infected. Finally, extremely low cut-off values were also generally excluded, as they resulted in all herds being infected. Moreover, exclusion of combination also vary depending on context parameters values. The proportion of combination excluded was higher when sensitivity was high, specificity was low and prevalence was low. Within such context, the number of herd still to detect is low, limiting the additional value of the model. The added number of infected herds compared to test vary a lot depending on percentile and cut-off value. Higher number of additional infected herds detected of number of infected herds were observed with high percentiles. Different patterns were observed depending on parameter values. Cut-off values around 0.2 associated with a high percentiles appears to be combinations that can be used in many case. However, these study results need to be deepened to get a better understanding on the impact of context parameters on selection of percentile and cut-off.

Chapitre 6

# Discussion générale

Dans l'Union Européenne, de nombreux programmes de maîtrise des maladies infectieuses non règlementées ont été mis en place. De la grande diversité d'objectifs et de moyens de ces programmes, résulte une grande hétérogénéité de situations vis-à-vis des infections ciblées. Chaque programme ayant sa propre définition du statut *indemne d'infection*, basée sur des tests de diagnostic différents et imparfaits (van Roon *et al.*, 2020b), la probabilité d'être classé *indemne d'infection* à tort est différente selon les programmes. Les méthodes actuelles ne permettant pas d'estimer la probabilité d'être classé *indemne d'infection* à tort, quelles que soient les modalités de surveillance et le contexte épidémiologique, estimer une probabilité d'infection à l'échelle troupeau est donc nécessaire.

Dans cette thèse, nous avons contribué au développement et à l'évaluation d'une méthode statistique permettant d'estimer une probabilité d'absence d'infection à l'échelle troupeau, à partir de données hétérogènes longitudinales issues de programmes de maîtrise. La méthode développée s'applique sur une population de troupeaux participant à un même programme de maîtrise. Le modèle permet d'intégrer des données longitudinales relatives à chaque troupeau. Les données utilisables sont à la fois les données issues de la surveillance du statut de chaque troupeau, c'est-à-dire les résultats du dispositif de dépistage mis en place dans un programme de maîtrise, et des données relatives aux facteurs de risques d'introduction de l'infection à l'échelle du troupeau. Le modèle intègre également des informations relatives à la dynamique de l'infection à l'échelle du territoire, incluant le rôle de facteurs de risques sur cette dynamique, ainsi que des informations sur les performances du dispositif de surveillance.

La description, l'organisation, ainsi que l'évaluation de la disponibilité des différentes informations sur le processus infectieux étaient une étape importante pour créer le cadre général d'une méthode d'estimation à partir de données hétérogènes. La diarrhée virale bovine est une maladie infectieuse à la fois très répandue (Scharnböck *et al.*, 2018), très

étudiée et avec de nombreux programmes de maîtrise en place depuis plusieurs années dans différents pays (van Roon *et al.*, 2020b; Ståhl & Alenius, 2012). Les étapes du processus infectieux chez un animal et le processus de transmission, d'un animal ou d'un troupeau à un autre, ainsi que la diversité des dispositifs de dépistage ont été largement décrits (Houe, 1995; Lindberg & Houe, 2005; Houe *et al.*, 2006; Meyling *et al.*, 1990). Ainsi, nous avons pu identifier l'échelle du troupeau comme étant l'échelle la plus pertinente pour développer le modèle. Le dispositif de dépistage mis en place dans les programmes de maîtrise est établi pour suivre le statut des exploitations. Ainsi, à cette échelle, il existe des données longitudinales de résultats de tests répétés, mais aussi des données sur les facteurs de risques d'introduction de l'agent infectieux. La méthode développée vise particulièrement à détecter les troupeaux classés *indemne d'infection* à tort. Les facteurs de risque d'introduction permettent de prendre en compte la probabilité que l'agent pathogène ai été introduit dans le troupeau, mais pas encore détecté par le dispositif de dépistage. De plus, nous avons mis en évidence la disponibilité de certaines informations à l'échelle du territoire, telles que, les performances du dispositif de dépistage en place et le rôle des facteurs de risques dans la dynamique. Les facteurs influençant l'élimination de l'infection à l'échelle du troupeau, sont très dépendants des mesures mises en place dans le programme de maîtrise, et ont été considérés comme homogènes pour un programme de maîtrise donné, et donc homogènes dans le territoire.

Le modèle STOC free que nous avons développé permet d'estimer une probabilité d'être infecté à l'échelle du troupeau en intégrant différents types d'informations. Les hypothèses du modèle ont été construites pour permettre d'intégrer la diversité des informations disponibles identifiées dans le chapitre II. Le cadre bayésien du modèle permet de prendre en compte les connaissances sur la dynamique de changement du statut d'un troupeau vis-à-vis de l'infection, au cours du temps et au sein du territoire. Plus précisément, les probabilités d'un troupeau de devenir ou de rester infecté et entre deux pas de temps sont estimées par le modèle, et sont communes à tous les troupeaux du territoire. Le modèle intègre également des données spécifiques à chaque troupeau pour ajuster le risque de devenir infecté, en prenant en compte la présence des facteurs de risques d'introduction au cours du temps. Il intègre par ailleurs les résultats de tests issus des programmes de maîtrise, selon le dispositif de dépistage mis en œuvre de façon répétée au cours du temps dans chaque troupeau. Les connaissances sur les performances du dispositif de dépistage quant à sa capacité à correctement classer un troupeau, sont intégrées et constantes pour tout le territoire. Ce modèle intègre donc à la fois les don-

nées spécifiques à chaque troupeau et des connaissances, à priori communes, à tous les troupeaux d'un même territoire.

Nous avons mis en évidence, avec des données simulées, l'impact de certains éléments de contexte sur les performances du modèle STOC free. Les performances du modèle ont été définies comme sa capacité à détecter plus d'exploitations infectées que le simple résultat de test. Le nombre troupeaux infectés détectés en plus a été rapporté au nombre de troupeaux indemnes classés comme *infectés* à tort par le modèle, pour définir une notion de coût. La sensibilité du dispositif de dépistage est apparue comme le paramètre le plus influent sur la capacité du modèle à détecter plus de troupeaux infectés, mais également sur sa capacité à estimer les autres paramètres du modèle. Quand les performances du dispositif de dépistage sont faibles, la contribution des données issues de ce dispositif pour définir le statut latent de l'exploitation est faible. De mauvaises performances du dispositif de dépistage peuvent entrainer des difficultés à estimer les autres paramètres du modèle. D'un autre côté, plus le dispositif de dépistage est performant, moins la valeur ajoutée du modèle est élevée, car la majorité des troupeaux infectés sont déjà détectés.

Cette étude n'a en revanche pas mis en évidence l'importance du poids et de la fréquence des facteurs de risques sur les performances du modèle dans les contextes étudiés. Dans cette étude le facteur de risque a été distribué de manière aléatoire dans chaque troupeau, ce qui a pu limiter l'identification de l'influence des facteurs de risques. Etudier les performances du modèle avec des exploitations où la fréquence du facteur de risque serait plus élevée, et des exploitations avec une plus faible fréquence du facteur de risque, pourrait permettre de mieux identifier le rôle des facteurs de risques.

Nous avons identifié la difficulté de classer les troupeaux dans les différents statuts vis à vis de l'infection, à partir de la distribution de probabilité d'être infect, produite par le modèle. En effet, le modèle STOC free produit pour chacun des troupeaux une distribution de probabilité d'être infecté. Classer un troupeau comme *infecté* ou *indemne d'infection* en fonction de cette distribution nécessite à la fois, de résumer cette distribution de probabilité, mais également de sélectionner une valeur seuil en dessous de laquelle les troupeaux sont classés *indemne d'infection*, et au-dessus de laquelle ces troupeaux sont classés *infecté*. Nous avons mis en évidence que le percentile utilisé pour résumer la distribution postérieure de la probabilité d'être infecté pour un troupeau et la valeur du seuil, pouvaient impacter grandement les performances du modèle, dans des contextes variés. Cependant, nous n'avons pas pu identifier des règles claires pour guider le choix du

percentile et de la valeur du seuil en fonction du contexte. Des travaux supplémentaires sont nécessaires pour développer une règle de décision de classement des troupeaux à partir de la distribution de probabilité d'être infecté.

L'étude des performances du modèle STOC free a été réalisée sur des données simulées. Le choix des données simulées a permis d'avoir un statut vrai pour chacune des exploitations, utilisé comme *gold standard*, pour établir les performances du modèle. Les données simulées dans cette étude étaient relativement simples, avec un statut troupeau dépendant de la dynamique de l'infection dans le territoire et de la présence du facteur de risque à l'échelle du troupeau. Les résultats de tests du troupeau étaient ensuite simulées en fonction du vrai statut et des performances de dispositif de dépistage. Simuler des données plus complexes, pourrait permettre d'affiner les performances du modèle STOC free. Afin de réaliser une évaluation sur des données simulées plus réalistes, une solution pourrait-être d'utiliser des modèles de simulation mécaniste existants, qui permettent d'incorporer plus de détails sur la dynamique intra-troupeau et sur les performances de dépistages.

# 1 Définition des statuts vis à vis de l'infection

Pour une maladie infectieuse donnée, il existe plusieurs définitions possibles du statut *infecté* au niveau troupeau. Les statuts doivent être définis pour bien caractériser le processus dynamique comme décrit dans le chapitre II, et pour permettre l'interprétation des résultats des dispositifs de dépistage et des facteurs de risques . Les connaissances sur la BVD ont permis d'identifier les statuts d'intérêts. Nous avons, par exemple, distingué les troupeaux *infectés par des animaux infectés transitoires (TI)*, des troupeaux *infectés par des animaux infectés permanents immuno-tolérant (IPI)*. Les IPI ayant un rôle primordial dans la transmission de l'infection (Lindberg & Houe, 2005), ils ont une grande importance du point de vue de la surveillance. Les spécificités de chaque maladie infectieuse doivent donc être prises en compte pour définir les statuts vis-à-vis de l'infection. A titre d'exemple, pour les maladies infectieuses dont l'agent persiste durablement dans l'environnement, tel que l'agent de la paratuberculose, *Mycobacterium avium* subsp. *paratuberculosis*, hautement résistant dans l'environnement (Whittington *et al.*, 2004), la définition des statuts d'intérêts devra prendre en compte la contamination dans l'environnement, et pas seulement la présence d'animaux infectés. Dans le modèle STOC free

seulement deux statuts *infecté* et *indemne d'infection* sont définis pour l'instant. En effet, le modèle vise à estimer la probabilité qu'une exploitation soit infectée, pour limiter le risque de commercialiser des animaux infectés.

Les statuts sont aussi définis en fonction de la modification de paramètres biologiques, tels que la présence d'anticorps ou d'antigènes spécifiques, pouvant être mesurés via des tests de dépistage et pouvant indiquer la présence de l'agent infectieux. Par exemple, nous avons définis à la fois un statut *indemne* d'infection et un statut *immunisé* pour la BVD. Dans les troupeaux *indemnes* et les troupeaux *immunisés*, aucun animal infecté n'est présent. Cependant, un test de dépistage basé sur la détection d'anticorps réalisé dans ces deux types de troupeaux donnera des résultats différents : la recherche d'anticorps anti-BVD dans le troupeau *immunisé* aura un résultat positif alors que le troupeau *indemne* aura un résultat négatif. Dans le modèle STOC free seulement deux statuts *infecté* et *indemne d'infection* sont définis. Si le statut *infecté* représente les troupeaux avec au moins un animal infecté, alors les troupeaux *indemnes* et *immunisés* devraient tout les deux être classés comme *indemnes d'infection*, et dans ce cas, la spécificité du dispositif de dépistage basée sur la recherche d'anticorps devrait être faible, car l'immuntié des animaux vis à vis de la BVD est considérée comme persistante toute la vie du bovin.

## 2 Interprétation des dispositifs de dépistage

Pour appliquer le modèle STOC free sur des données réelles, il est nécessaire de rassembler des connaissances sur les performances du dispositif de dépistage pour la détermination du statut troupeau. La sensibilité et spécificité telles que définies dans le modèle, sont la capacité du dispositif à détecter les troupeaux infectés et non infectés. C'est ce que nous avons considéré dans notre étude sur données simulées, où la sensibilité et la spécificité du test représentent les capacités du test à détecter les troupeaux *infectés* et les troupeaux *indemnes*. Dans un plan de surveillance, il n'est pas simple de transposer les performances d'un test habituellement défini à l'échelle de l'animal, en performances à l'échelle troupeau. Par exemple, la sensibilité et la spécificité des tests pour détecter des anticorps contre le virus de la BVD dans le lait de tank sont généralement très bonnes (Houe *et al.*, 2006; Beaudeau *et al.*, 2001a,b). Cependant, les performances du même test sont moins bonnes si on cherche à détecter un troupeau infecté par un animal IPI (Houe *et al.*, 2006). Il est essentiel de construire les priors des performances du dispositif de

dépistage au regard de la définition des statuts vis-à-vis de l'infection que l'on cherche à identifier, et non en fonction de leur(s) cible(s) biologique(s) (anticorps, ou agent infectieux). Bien que les tests de diagnostics du virus de la BVD soient très performants à l'échelle animale, l'interprétation des tests dans le sens des statuts vis à vis de l'infection à l'échelle troupeau peut être complexe.

## 3 Intégration des facteurs de risques

Le modèle STOC free permet d'intégrer des informations sur les facteurs de risques d'introduction de l'infection, pour lesquels il est nécessaire de sélectionner des variables qui décrivent, pour chaque troupeau, l'exposition au risque au cours du temps. Les variables descriptives sont souvent appelées *proxy* car elles mesurent la présence de pratiques connues comme étant associées à une possibilité d'introduire le virus dans un troupeau, sans que chaque évènement puisse être mesuré. Une revue de la littérature permet de sélectionner au mieux les facteurs de risques principaux, mais également d'en extraire des informations sur le risque relatif associé à chaque facteur de risque pour aider à construire les priors. En l'absence de données spécifiques pour un territoire donné, les valeurs ainsi estimées peuvent servir de base à la construction des priors.

La prise en compte de certains facteurs de risques, non observables ou non renseignés, peut-être envisagée en s'appuyant sur d'autres variables dont on fait l'hypothèse qu'elles sont liées au facteur de risque. Ainsi, bien que ces variables ne correspondent pas directement à une mesure approximative du mécanisme en jeu, elles permettent de renseigner le facteur de risque. Par exemple, les études observationnelles réalisées n'ont pas permis de mettre en évidence l'impact des mesures de bio-sécurité sur la présence de la BVD dans un troupeau (van Roon *et al.*, 2020a). La manière dont les mesures de bio-sécurité étaient renseignées (par questions fermées ou déclaratives), ont pu limiter l'identification de ces mesures comme facteur de risque. Les variables utilisées sont principalement des variables déclaratives, qui présentent de multiples biais (questions fermées), et ne reflètent sans doute pas précisément si les pratiques réduisent effectivement la probabilité d'introduction du virus. Pour intégrer le potentiel effet bénéfique de mesures de bio-sécurité sur le statut d'un troupeau, d'autres variables pourraient être intégrée au modèle, telles que, par exemple la durée depuis laquelle le troupeau est testé négatif. Un troupeau dans lequel l'infection n'a pas été détectée depuis longtemps est peut-être moins à risque que d'autres,

car cela met en avant une bonne maîtrise des risques à l'échelle du troupeau. On peut même envisager d'intégrer le statut du troupeau vis à vis d'autres maladies infectieuses. En effet, certains facteurs de risques d'introduction dans un troupeau sont communs à différentes maladies infectieuses et une bonne maîtrise des risques pour une infection donnée peut permettre d'inclure ces mêmes risques pour l'infection d'intérêt.

# 4 Perspectives

Les travaux présentés dans cette thèse nécessitent des développements ultérieurs pour que la méthode de surveillance *output-based* développée puisse être mise en place sur le terrain. En effet, il reste à définir en quoi la probabilité d'être infecté produite par le modèle STOC free permet de comparer la probabilité d'acheter un animal infecté entre deux troupeaux issus de territoires différents. L'estimation des valeurs prédictives négatives, c'est-à-dire la proportion de troupeaux réellement indemnes d'infection parmi les troupeaux classés *indemne d'infection*, pourrait être associée à la probabilité d'être infecté du troupeau, et représenter l'incertitude autour de ce statut. De plus, pour une mise en place sur le terrain de cette démarche il sera essentiel de communiquer auprès des acteurs de terrain et de les former à l'utilisation de tels outils.

# BIBLIOGRAPHIE


Amelung, Silke, Hartmann, Maria, Haas, Ludwig, & Kreienbrock, Lothar. 2018. Factors associated with the bovine viral diarrhoea (BVD) status in cattle herds in Northwest Germany. *Veterinary microbiology*, **216**, 212–217.

Andrews, Justin M, & Langmuir, Alexander D. 1963. The philosophy of disease eradication. *American Journal of Public Health and the Nations Health*, **53**(1), 1–6.

Auger-Méthé, Marie, Field, Chris, Albertsen, Christoffer M, Derocher, Andrew E, Lewis, Mark A, Jonsen, Ian D, & Flemming, Joanna Mills. 2016. State-space models' dirty little secrets : even simple linear Gaussian models can have estimation problems. *Scientific reports*, **6**(1), 1–10.

Bachofen, Claudia, Vogt, Hans-Rudolf, Stalder, Hanspeter, Mathys, Tanja, Zanoni, Reto, Hilbe, Monika, Schweizer, Matthias, & Peterhans, Ernst. 2013. Persistent infections after natural transmission of bovine viral diarrhoea virus from cattle to goats and among goats. *Veterinary Research*, **44**(1), 1–10.

Beaudeau, F, Assie, S, Seegers, H, Belloc, C, Selial, E, & Joly, A. 2001a. Assessing the within-herd prevalence of cows antibody-positive to bovine viral diarrhoea virus with a blocking ELISA on bulk tank milk. *Veterinary Record*, **149**(8), 236–240.

Beaudeau, F, Belloc, C, Seegers, H, Assie, S, Sellal, E, & Joly, A. 2001b. Evaluation of a blocking ELISA for the detection of bovine viral diarrhoea virus (BVDV) antibodies in serum and milk. *Veterinary microbiology*, **80**(4), 329–337.

Bernard, Coralie. 2018. *Description de la diversité des plans de Lutte contre la diarrhée virale bovine en France.* Thèse de doctorat vétérinaire, Oniris : Ecole Nationale Vétérinaire, Agroalimentaire et de l'Alimentation, Nantes, France.

Bhudevi, Bodreddigari, & Weinstock, Daniel. 2001. Fluorogenic RT–PCR assay (TaqMan) for detection and classification of bovine viral diarrhea virus. *Veterinary microbiology*, **83**(1), 1–10.

Brownlie, J, Mc, Clarke, Cj, Howard, & Dh, Pocock. 1987. Pathogenesis and epidemiology of bovine virus diarrhoea virus infection of cattle. *Pages 157–166 of : Annales de Recherches Veterinaires*, vol. 18.

Brownlie, J, Hooper, LB, Thompson, I, & Collins, ME. 1998. Maternal recognition of foetal infection with bovine virus diarrhoea virus (BVDV)—the bovine pestivirus. *Clinical and diagnostic virology*, **10**(2-3), 141–150.

Calba, Clementine, Goutard, Flavie L, Hoinville, Linda, Hendrikx, Pascal, Lindberg, Ann, Saegerman, Claude, & Peyre, Marisa. 2015. Surveillance systems evaluation : a systematic review of the existing approaches. *BMC public health*, **15**(1), 1–13.

Cameron, AR. 2012. The consequences of risk-based surveillance : Developing output-based standards for surveillance to demonstrate freedom from disease. *Preventive veterinary medicine*, **105**(4), 280–286.

Charoenlarp, W, Frankena, K, Strain, SAJ, Guelbenzu-Gonzalo, M, Graham, J, & Byrne, AW. 2018. Spatial and risk factor analysis of bovine viral diarrhoea (BVD) virus after the first-year compulsory phase of BVD eradication programme in Northern Ireland. *Preventive veterinary medicine*, **157**, 34–43.

Christensen, Jette, & Gardner, Ian A. 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. *Preventive veterinary medicine*, **45**(1-2), 83–106.

Costa, Lina, Duarte, Elsa L, Knific, Tanja, Hodnik, Jaka Jakob, van Roon, Annika, Fourichon, Christine, Koleci, Xhelil, van Schaik, Gerdien, Gunn, George, Madouasse, Aurélien, *et al.* 2020. Standardizing output-based surveillance to control non-regulated cattle diseases : Aspiring for a single general regulatory framework in the European Union. *Preventive Veterinary Medicine*, **183**, 105130.

Denwood, Matthew J. 2016. runjags : An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of statistical software*, **71**(1), 1–25.

Ducrot, Christian, Pécaud, Dominique, Petit, Étienne, Krebs, Stéphane, Viet, Anne-France, Durand, Benoit, Biteau-Coroller, Fabienne, Beaudeau, François, Frappat, Brigitte, Calavas, Didier, *et al.* 2010. Farm Animal Health Qualification and Risk Repre-

sentations by the Stakeholders and Scientific Disciplines. *Natures Sciences Societes*, **18**(1), 3–13.

Duncan, AJ, Gunn, GJ, & Humphry, RW. 2016. Difficulties arising from the variety of testing schemes used for bovine viral diarrhoea virus (BVDV). *Veterinary Record*, **178**(12), 292–292.

Duncan, Colleen, Ridpath, Julia, Palmer, Mitchell V, Driskell, Elizabeth, & Spraker, Terry. 2008. Histopathologic and immunohistochemical findings in two white-tailed deer fawns persistently infected with Bovine viral diarrhea virus. *Journal of Veterinary Diagnostic Investigation*, **20**(3), 289–296.

Ersbøll, Annette Kjær, Ersbøll, Bjarne Kjær, Houe, Hans, Alban, L, & Kjeldsen, AM. 2010. Spatial modelling of the between-herd infection dynamics of bovine virus diarrhoea virus (BVDV) in dairy herds in Denmark. *Preventive veterinary medicine*, **97**(2), 83–89.

Foddai, Alessandro, Stockmarr, Anders, & Boklund, Anette. 2016. Evaluation of temporal surveillance system sensitivity and freedom from bovine viral diarrhea in Danish dairy herds using scenario tree modelling. *BMC veterinary research*, **12**(1), 1–12.

Gates, MC, Woolhouse, MEJ, Gunn, GJ, & Humphry, RW. 2013. Relative associations of cattle movements, local spread, and biosecurity with bovine viral diarrhoea virus (BVDV) seropositivity in beef and dairy herds. *Preventive Veterinary Medicine*, **112**(3-4), 285–295.

Gelman, Andrew, Rubin, Donald B, *et al.* 1992. Inference from iterative simulation using multiple sequences. *Statistical science*, **7**(4), 457–472.

Graham, DA, Clegg, TA, Lynch, M, & More, SJ. 2013. Herd-level factors associated with the presence of bovine viral diarrhoea virus in herds participating in the voluntary phase of the Irish national eradication programme. *Preventive Veterinary Medicine*, **112**(1-2), 99–108.

Graham, DA, Clegg, TA, Thulke, H-H, O'sullivan, P, McGrath, G, & More, SJ. 2016. Quantifying the risk of spread of bovine viral diarrhoea virus (BVDV) between contiguous herds in Ireland. *Preventive veterinary medicine*, **126**, 30–38.

Gunn, HM. 1993. Role of fomites and flies in the transmission of bovine viral diarrhoea virus. *The Veterinary Record*, **132**(23), 584–585.

Hanon, J-B, Van der Stede, Yves, Antonissen, Anne, Mullender, Cedric, Tignon, Marylène, van den Berg, Thierry, & Caij, Brigitte. 2014. Distinction Between Persistent and Transient Infection in a B ovine V iral D iarrhoea (BVD) Control Programme : Appropriate Interpretation of Real-Time RT-PCR and Antigen-ELISA Test Results. *Transboundary and emerging diseases*, **61**(2), 156–162.

Hanon, Jean-Baptiste, De Baere, Miet, de la Ferté, Camille, Roelandt, Sophie, Guillot, Gilles, Van der Stede, Yves, & Cay, Brigitte. 2018. Serological monitoring on milk and serum samples in a BVD eradication program : A field study in Belgium showing antibody ELISA performances and epidemiological aspects. *Preventive veterinary medicine*, **160**, 136–144.

Hodnik, Jaka Jakob, Santman-Berends, Inge, Correia-Gomes, Carla, Strain, Sam, Gethmann, Jörn, Gunn, George, & members, WG1. 2020. *SOUND CONTROL, Overview of non-regulated cattle diseases in the European Union for which control programmes (CPs) are in place within member states.* techreport. COST ACTION.

Hoinville, Linda, Ronello, A, Alban, L, *et al.* 2013. Animal health surveillance terminology final report from Pre-ICAHS workshop. *Health*, **2013**, 1–26.

Houe, H, Baker, JC, Maes, RK, Lloyd, JW, & Enevoldsen, C. 1995. Comparison of the prevalence and incidence of infection with bovine virus diarrhoea virus (BVDV) in Denmark and Michigan and association with possible risk factors. *Acta Veterinaria Scandinavica*, **36**(4), 521–531.

Houe, H, Lindberg, A, & Moennig, V. 2006. Test strategies in bovine viral diarrhea virus control and eradication campaigns in Europe. *Journal of Veterinary Diagnostic Investigation*, **18**(5), 427–436.

Houe, Hans. 1995. Epidemiology of bovine viral diarrhea virus. *Veterinary Clinics of North America : Food Animal Practice*, **11**(3), 521–547.

Houe, Hans, Nielsen, Liza Rosenbaum, & Nielsen, Søren Saxmose. 2014. *Control and eradication of endemic infectious diseases in cattle.* College Publications.

Hult, L, & Lindberg, A. 2005. Experiences from BVDV control in Sweden. *Preventive veterinary medicine*, **72**(1-2), 143–148.

Joly, A, Beaudeau, F, & Seegers, H. 2001. Evaluation de la prévalence et de la dynamique de l'infection BVD en Bretagne à l'aide d'un test ELISA sur lait de grand mélange. *Epidémiologie et santé animale*, **40**, 7–14.

Lindberg, A, & Houe, H. 2005. Characteristics in the epidemiology of bovine viral diarrhea virus (BVDV) of relevance to control. *Preventive veterinary medicine*, **72**(1-2), 55–73.

Lindberg, Ann LE, & Alenius, Stefan. 1999. Principles for eradication of bovine viral diarrhoea virus (BVDV) infections in cattle populations. *Veterinary microbiology*, **64**(2-3), 197–222.

Madouasse, A., Mercat, M., van Roon, A., Santman, I., Nielen, M., van Schaik, G., More, S.J., Graham, D., Guelbenzu, M., Frössling, J., Ågren, E., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., & Fourichon, C. 2019. *STOC free : WP1, Deliverable 1.3Description of the STOC free model.* techreport.

Madouasse, A., Mercat, M., van Roon, A., Graham, D., Guelbenzu, M., Berends, I. Santman, van Schaik, G., Nielen, M., Frössling, J., Ågren, E., Humphry, R. W., Eze, J., Gunn, G. J., Henry, M. K., Gethmann, J., More, S. J., Toft, N., & Fourichon, C. 2021. A modelling framework for the prediction of the herd-level probability of infection from longitudinal data. *Under review.*

Martin, P. A. J., Cameron, A. R., & Greiner, M. 2007. Demonstrating freedom from disease using multiple complex data sources. **79**, 71–97.

McClurkin, AW, Littledike, ET, Cutlip, RC, Frank, GH, Coria, MF, & Bolin, SR. 1984. Production of cattle immunotolerant to bovine viral diarrhea virus. *Canadian journal of comparative medicine*, **48**(2), 156.

Mercat, M., van Roon, A., Santman, I., Nielen, M., van Duijn, L., van Schaik, G., More, S.J., Graham, D., Frössling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., Fourichon, C., & Madouasse, A. 2018a. *STOC free : WP1 and WP2 , Deliverable 1.2Guidelines for the identification and sources of data.* techreport. EFSA.

Mercat, M., van Roon, A., Santman, I., Nielen, M., van Duijn, L., van Schaik, G., More, S.J., Graham, D., Frössling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., Fourichon, C., & Madouasse, A. 2018b. *STOC free : WP1, Deliverable 1.1Guidelines for the design of conceptual models.* techreport. EFSA.

Meyling, A, & Jensen, A Mikél. 1988. Transmission of bovine virus diarrhoea virus (BVDV) by artificial insemination (AI) with semen from a persistently-infected bull. *Veterinary Microbiology*, **17**(2), 97–105.

Meyling, A, Houe, H, & Jensen, AM. 1990. Epidemiology of bovine virus diarrhoea virus. *Revue scientifique et technique (International Office of Epizootics)*, **9**(1), 75–93.

Moennig, Volker, & Liess, Bernd. 1995. Pathogenesis of intrauterine infections with bovine viral diarrhea virus. *Veterinary Clinics of North America : Food Animal Practice*, **11**(3), 477–487.

Moennig, Volker, Houe, Hans, & Lindberg, Ann. 2005. BVD control in Europe : current status and perspectives. *Animal Health Research Reviews*, **6**(1), 63.

More, Simon J, Cameron, Angus R, Greiner, Matthias, Clifton-Hadley, Richard S, Rodeia, Sandra Correia, Bakker, Douwe, Salman, Mo D, Sharp, J Michael, De Massis, Fabrizio, Aranaz, Alicia, *et al.* 2009. Defining output-based standards to achieve and maintain tuberculosis freedom in farmed deer, with reference to member states of the European Union. *Preventive veterinary medicine*, **90**(3-4), 254–267.

Nelson, Danielle D, Dark, Michael J, Bradway, Daniel S, Ridpath, Julia F, Call, Neill, Haruna, Julius, Rurangirwa, Fred R, & Evermann, James F. 2008. Evidence for persistent Bovine viral diarrhea virus infection in a captive mountain goat (Oreamnos americanus). *Journal of Veterinary Diagnostic Investigation*, **20**(6), 752–759.

Nelson, Danielle D, Duprau, Jennifer L, Wolff, Peregrine L, & Evermann, James F. 2016. Persistent bovine viral diarrhea virus infection in domestic and wild small ruminants and camelids including the mountain goat (Oreamnos americanus). *Frontiers in microbiology*, **6**, 1415.

Nielsen, Søren Saxmose, & Toft, Nils. 2008. Ante mortem diagnosis of paratuberculosis : a review of accuracies of ELISA, interferon-$\gamma$ assay and faecal culture techniques. *Veterinary microbiology*, **129**(3-4), 217–235.

Niskanen, Rauni, & Lindberg, Ann. 2003. Transmission of bovine viral diarrhoea virus by unhygienic vaccination procedures, ambient air, and from contaminated pens. *The Veterinary Journal*, **165**(2), 125–130.

Niskanen, Rauni, Lindberg, A, & Tråvén, M. 2002. Failure to spread bovine virus diarrhoea virus infection from primarily infected calves despite concurrent infection with bovine coronavirus. *The Veterinary Journal*, **163**(3), 251–259.

Norström, Madelaine, Jonsson, Malin E, Åkerstedt, Johan, Whist, Anne Cathrine, Kristoffersen, Anja Bråthen, Sviland, Ståle, Hopp, Petter, & Wahlström, Helene. 2014. Estimation of the probability of freedom from Bovine virus diarrhoea virus in Norway using scenario tree modelling. *Preventive veterinary medicine*, **116**(1-2), 37–46.

on Animal Health, EFSA Panel, & (AHAW), Welfare. 2012. Scientific Opinion on the use of a gamma interferon test for the diagnosis of bovine tuberculosis. *EFSA Journal*, **10**(12), 2975.

Pedrera, M, Gómez-Villamandos, JC, Molina, V, Risalde, MA, Rodríguez-Sánchez, B, & Sánchez-Cordón, PJ. 2012. Quantification and Determination of Spread Mechanisms of Bovine Viral Diarrhoea Virus in Blood and Tissues from Colostrum-Deprived Calves During an Experimental Acute Infection Induced by a Non-Cytopathic Genotype 1 Strain. *Transboundary and emerging diseases*, **59**(5), 377–384.

Plummer, M. 2003. JAGS : A program for analysis of Bayesian graphical models using Gibbs sampling. *In :* Hornik, K., Leisch, F., & Vines, K. (eds), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003).*

Presi, Patrick, Struchen, Rahel, Knight-Jones, Theodore, Scholl, Sabrina, & Heim, Dagmar. 2011. Bovine viral diarrhea (BVD) eradication in Switzerland—experiences of the first two years. *Preventive veterinary medicine*, **99**(2-4), 112–121.

R, Core, & Team. 2017. R : A language and Computing.

Raaperi, Kerli, Orro, Toomas, & Viltrop, Arvo. 2014. Epidemiology and control of bovine herpesvirus 1 infection in Europe. *The Veterinary Journal*, **201**(3), 249–256.

Rangel, Saray J, Paré, Julie, Doré, Elizabeth, Arango, Juan C, Côté, Geneviève, Buczinski, Sebastien, Labrecque, Olivia, Fairbrother, Julie H, Roy, Jean P, Wellemans, Vincent,

*et al.* 2015. A systematic review of risk factors associated with the introduction of Mycobacterium avium spp. paratuberculosis (MAP) into dairy herds. *The Canadian Veterinary Journal*, **56**(2), 169.

Reardon, Fiona, Graham, David A, Clegg, Tracy A, Tratalos, Jamie A, O'Sullivan, Padraig, & More, Simon J. 2018. Quantifying the role of Trojan dams in the between-herd spread of bovine viral diarrhoea virus (BVDv) in Ireland. *Preventive veterinary medicine*, **152**, 65–73.

Rikula, Ulla, Nuotio, L, Laamanen, UI, & Sihvonen, L. 2008. Transmission of bovine viral diarrhoea virus through the semen of acutely infected bulls under field conditions. *Veterinary Record*, **162**(3), 79–81.

Sarrazin, Steven, Dewulf, Jeroen, Mathijs, Elisabeth, Laureyns, Jozef, Mostin, Laurent, & Cay, Ann Brigitte. 2014. Virulence comparison and quantification of horizontal bovine viral diarrhoea virus transmission following experimental infection in calves. *The Veterinary Journal*, **202**(2), 244–249.

Scharnböck, Bettina, Roch, Franz-Ferdinand, Richter, Veronika, Funke, Carsten, Firth, Clair L, Obritzhauser, Walter, Baumgartner, Walter, Käsbohrer, Annemarie, & Pinior, Beate. 2018. A meta-analysis of bovine viral diarrhoea virus (BVDV) prevalences in the global cattle population. *Scientific reports*, **8**(1), 1–15.

Scherer, CFC, Flores, EF, Weiblen, R, Caron, L, Irigoyen, LF, Neves, JP, & Maciel, MN. 2001. Experimental infection of pregnant ewes with bovine viral diarrhea virus type-2 (BVDV-2) : effects on the pregnancy and fetus. *Veterinary microbiology*, **79**(4), 285–299.

Schuppers, ME, Stegeman, JA, Kramps, JA, & Stärk, KDC. 2012. Implementing a probabilistic definition of freedom from infection to facilitate trade of livestock : putting theory into praxis for the example of bovine herpes virus-1. *Preventive veterinary medicine*, **105**(3), 195–201.

Ståhl, Karl, & Alenius, Stefan. 2012. BVDV control and eradication in Europe—an update. *Japanese Journal of Veterinary Research*, **60**(Supplement), S31–S39.

Terpstra, C, & Wensvoort, G. 1997. A congenital persistent infection of bovine virus diarrhoea virus in pigs : clinical, virological and immunological observations. *Veterinary Quarterly*, **19**(3), 97–101.

Valle, PS, Martin, SW, Tremblay, R, & Bateman, K. 1999. Factors associated with being a bovine-virus diarrhoea (BVD) seropositive dairy herd in the Møre and Romsdal County of Norway. *Preventive veterinary medicine*, **40**(3-4), 165–177.

Van der Hoek, W, Dijkstra, F, Schimmer, B, Schneeberger, PM, Vellema, P, Wijkmans, C, Ter Schegget, R, Hackert, V, & Van Duynhoven, Y. 2010. Q fever in the Netherlands : an update on the epidemiology and control measures. *Eurosurveillance*, **15**(12), 19520.

van Roon, A. M., Mercat, M., van Schaik, G., Nielen, M., Graham, D. A., More, S. J., Guelbenzu-Gonzalo, M., Fourichon, C., Madouasse, A., & Santman-Berends, I. M. G. A. 2020a. Quantification of risk factors for bovine viral diarrhea virus in cattle herds : A systematic search and meta-analysis of observational studies. *Journal of dairy science*, **103**, 9446–9463.

van Roon, AM, Santman-Berends, Inge MGA, Graham, D, More, Simon John, Nielen, M, van Duijn, L, Mercat, M, Fourichon, C, Madouasse, A, Gethmann, J, *et al.* 2020b. A description and qualitative comparison of the elements of heterogeneous bovine viral diarrhea control programs that influence confidence of freedom. *Journal of dairy science*, **103**(5), 4654–4671.

Victora, C G, Huttly, S R, Fuchs, S C, & Olinto, M T. 1997. The role of conceptual frameworks in epidemiological analysis : a hierarchical approach. *International Journal of Epidemiology*, **26**(1), 224–227.

Whittington, Richard, Donat, Karsten, Weber, Maarten F, Kelton, David, Nielsen, Søren Saxmose, Eisenberg, Suzanne, Arrigoni, Norma, Juste, Ramon, Sáez, Jose Luis, Dhand, Navneet, *et al.* 2019. Control of paratuberculosis : who, why and how. A review of 48 countries. *BMC veterinary research*, **15**(1), 1–29.

Whittington, Richard J, Marshall, D Jeff, Nicholls, Paul J, Marsh, Ian B, & Reddacliff, Leslie A. 2004. Survival and dormancy of Mycobacterium avium subsp. paratuberculosis in the environment. *Applied and environmental microbiology*, **70**(5), 2989–3004.

Williams, D, & Winden, SV. 2014. Risk factors associated with high bulk milk antibody levels to common pathogens in UK dairies. *Veterinary Record*, **174**(23), 580–580.

Youden, William J. 1950. Index for rating diagnostic tests. *Cancer*, **3**(1), 32–35.

Zucchini, W., MacDonald, I. L., & Langrock, R. 2017. *Hidden Markov models for time series : an introduction using R.* Chapman and Hall/CRC.

# ANNEXES

# Annexe 1 : STOC free : WP1, Deliverable 1. Guidelines for the design of conceptual models representing the infectious process at different levels, from animal to region, with an application to BVD, 04/2018.

M. Mercat, A. van Roon, I. Santman, M. Nielen, L. van Duijn, G. van Schaik, S. More, D. Graham, J. Frössling, A. Lindberg, J. Gethmann, C. Sauter-Louis, G. Gunn, C. Gomes, M. Henry, C. Fourichon , A. Madouasse.

Surveillance analysis Tool for Outcome-based Comparison of the confidence of FREEdom

European Food Safety Authority

**STOC free: WP1, Deliverable 1**

# Guidelines for the design of conceptual models representing the infectious process at different levels, from animal to region, with an application to BVD

**M. Mercat[1], A. van Roon[2], I. Santman[2,3], M. Nielen[2], L. van Duijn[3], G. van Schaik[2,3], S. More[4], D. Graham[5], J. Frössling[6], A. Lindberg[6], J. Gethmann[7], C. Sauter-Louis[7], G. Gunn[8], C. Gomes[8], M. Henry[8], C. Fourichon[1] , A. Madouasse[1]**

[1] BIOEPAR, INRA, Oniris, Université Bretagne Loire, 44307, Nantes, France, [2]Urecht University, the Netherlands, [3]GD Animal Health, the Netherlands, [4]University College Dublin, Ireland, [5]Animal Health Ireland, [6]Swedish National Veterinary Institute, Sweden, [7]FLI, Germany, [8]SRUC, Scotland

April 2018

**GA/EFSA/AFSCO/2016/01-03**

**This study was awarded a grant by EFSA and was co-financed by public organisations in the countries participating in the study.**

# TABLE OF CONTENT

# 1 Conceptual models aims and design strategy

## 1.1 Definition of the conceptual model

The aim of the STOC free project is to design and validate a framework that enables a transparent and standardized estimation of a probability of freedom from infection and its associated uncertainty from heterogeneous information. An essential step in this estimation will be to connect the available information to a probability of infection.

The conceptual model will map the different types of information that exist for a given infectious disease onto the true status regarding infection. It is conceptual in the sense that it connects:

- The biological system: the true status regarding infection which is of interest for different levels of analysis: animal, herd and territory.
- Information that is extremely diverse. Conceptually, 2 types of information that are different in nature can be distinguished:
    - Information generated and collected to specifically detect the infection or the pathogen of interest such as test results from control programmes
    - Information associated with an increased probability of pathogen presence such as risk factors for infection

The conceptual model will be made of diagrams and text explanations. It will eventually be used to design statistical models that will integrate different pieces of information (data) for the estimation of probabilities of being in each single state of interest (outcome) at different levels.

## 1.2 Motivations for the conceptual model

For some non-regulated endemic infectious diseases of cattle, control programmes have been developed and implemented in several territories. These programmes can differ in terms of objective and means. As an example, bovine viral diarrhoea (BVD) is a disease against which several control programmes exist worldwide. Some of these programmes, at least initially, aim at the control of the infection while others aim at its eradication. Some use bulk tank milk testing for disease screening, while others resort to individual screening methods such as serum or tissue tag testing. Some operate at a regional scale while others are conducted at a national level. Some are mandatory while others are voluntary. Some only test once every year while others test on multiple moments (bi-annually, quarterly or monthly). This heterogeneity makes estimations of probabilities of freedom from infection hard to compare between territories operating under different control programmes, because information used or available to determine these probabilities are different. This complicates animal trade between regions when buyers need to evaluate the risk taken when purchasing an animal or a group of animals from another region. Avoiding the introduction of an infectious disease can be important, especially in regions that have managed to successfully control or eradicate the disease.

In this context, the aim of the STOC free project is to design and validate a framework that enables a transparent and standardized estimation of a probability of freedom from infection and its associated uncertainty from heterogeneous information. This is a type of surveillance, known as outcome-based surveillance, in which the focus is on the surveillance outcome regardless of the

means used to arrive at this outcome. Therefore, we will use heterogeneous information to estimate a probability of freedom from infection which will be the outcome. Information that can be used to estimate a probability of freedom from infection depends on the type of data that are collected and available and varies between regions, types of production (notably beef vs. dairy) and control programmes.

The available information can relate to either consequences or potential causes of the infection. Usually, the status of animals, herds or territories regarding an infection is evaluated by performing biological tests. These tests measure consequences of infection. But risk factors relating to the probability of introducing the infection could also be included. These risk factors are conceptually different from biological test results because they are associated with the cause of the infection, which may have occurred or not, while test results are associated with a possible infection which is a past event.

Information can be obtained on different biological phenomena associated with the infection process. For instance, biological tests can seek to identify antibodies, antigens or nucleic acids. For the same biological phenomenon, the information obtained can be different either because the test used is different or because it is performed on a different matrix, such as blood, milk, faeces or skin tissue.

Regarding the level at which the information is available, biological tests can be performed at the individual level or on pooled samples such as bulk milk. Therefore, there is a methodological challenge in being able to estimate an outcome that is comparable, i.e. a probability of freedom from infection and associated uncertainty, regardless of the inputs, that may be extremely variable. In this work, this process will be completed for BVD. The first step will consist of representing the features of the infection that do not vary between countries such as course of infection in a bovine and then to connect to these features to the different types of information that can be used to estimate a probability of freedom from infection by BVDV.

The representation of the (true) states regarding infection and their connection to available observations **is what we call a conceptual model.** Depending on the level of interest (animal, herd or territory), there can be several conceptual models for the same disease. Each level is composed of two layers: the first layer is the representation of the different possible (true) states of the system and the second layer represents the different types of observations that can be used to determine the state of the system. This model will serve as the basis to construct the statistical models that will integrate all the available information on all three levels in order to estimate the probability of freedom from infection and associated uncertainty.

The first layer of the conceptual model that is independent from observations is the representation of the BVDV biological system. At animal level this is the infection process. The infection process first involves the course of infection within an animal: how an infectious agent is transmitted to an animal, how the animal responds to the infection and how it recovers from infection. This course of infection is agent dependent ( in this case BVDV), it will be linked to infectious agent characteristics: the routes of transmission, the clinical disease associated to the infection and the ability of hosts to become immune (lifelong or not). If vaccination is available, the course of infection in animals can be modified by vaccinating the animal. At herd level, the infection process is impacted by herd

husbandry. Depending on agent characteristics, the structure and the management of the herd can enhance or reduce the probability of (re-) introduction and/or spread and/or the delayed detection of the virus within the herd. Finally, at the territory level, contact structure (both within and between territories), prevalence, control programmes and policies can impact the introduction and the spread of the agent through a territory.

The second layer of the model will represent the available information or observations and how it can be connected to the infection process. Direct observations of the evolving infection process are generally not available. So, we have to use available information or observations that can inform or enhance our understanding of the actual process, the state of the system. Those can be for example diagnostic test results, demographic/geographic parameters and surveillance or control programme information.

This conceptual model, mapping information onto the true status regarding infection, will enable us to have a better understanding of available information and how to interpret them. The next sections describe the step-by-step design of the conceptual model. In section 2, 3 and 4 the developed conceptual model for BVDV infection at animal, herd and territory level respectively. The last section will discuss how the conceptual models will be translated into the statistical model.

## 1.3 Step for the design of the conceptual model

Three levels are considered: animal, herd and territory. For each level, the first step of the conceptual model is the representation of the biological features of infection. The work is based on bibliographic research. The representation has to include disease biology, dynamics of infection and transmission, and characteristics of the pathogen of interest (survival rate in the environment for example). This work requires a good overview of the susceptible population, the infectious agent, the disease and associated risk factors (Victora *et al.*, 1997). Then, quantitative information about the infection (such as duration of infection, duration of shedding …) will be added.

The next step will be the description and connection of all the possible available observations to the different states of our systems. Observations can be either causes of infection, such as risk factors for introduction or transmission; or consequences like diagnostic test results. They can take the form of aggregate observations like prevalence for a territory, a herd, a group of animals.

In non-regulated livestock diseases, the susceptible population will be farmed animals. To focus on the infection dynamics, we can study it at different levels. The first level will be the animal level, which explains the transmission of infectious agent and the course of infection in an animal. This level will be based on good knowledge of the disease biology and its progression, after the animal is infected by the pathogen. It describes different types of infection (i.e. vertical or horizontal), of disease (i.e. clinical, subclinical or mucosal disease in persistently infected animals), possible ways of transmission (i.e. direct or indirect) and the development of immunity. In this part we also include the possibility of vaccination and its efficiency and possibly interference with testing results.

Then, we will focus on herd level. Herd structure, herd dynamics and specific farming practices can influence (re-)introduction, virus spread within the herd and delayed detection of infectious agent in the herd. For example, if the transmission of infection can occur through direct contact between animals, the density of animals in a herd will influence the transmission of virus within the herd. Finally, we can also consider the territory level. In fact, this is often the level where the programmes

are practically applied. This means that herds within a territory, even if they can be different, will have common surveillance practices and measurements. Therefore, the second level of the model will be to explain the dynamics of infection at each level, including connection with risk factors.

Finally, in parallel with the description of the infection dynamics at each chosen level, the conceptual model must list, describe and link all the possible available observations of our system to each level. Observations can be either diagnostic test results, geographic parameters, like density of animals or density of herds. They can also take the form of aggregate observations like prevalence for a territory, a herd or a group of animals.

# 2 Conceptual model for BVD at the individual animal level

In this section we first describe the epidemiological states of individual animals regarding infection with BVDV. We then describe how the animals can move between these states, i.e. the course of infection. The last part details how biological test results can be used to elucidate the epidemiological status of an animal.

## 2.1 Epidemiological states of the system

In order to quantify the probability that an animal is free from infection with BVDV, 4 mutually exclusive categories of animals are considered and described: persistently infected, (PI), transiently infected (TI), immune (R) and susceptible individuals (S).

### 2.1.1 Persistently infected (PI)

Persistently infected animals are the most important source of BVDV infection. PIs are infected in utero, between 30 and 120 days of gestation, while their immune system is immature. As a consequence, they are immunotolerant (they do not produce antibodies against homologous virus), become persistently infected and shed large amounts of virus throughout their lives. A calf born to a PI cow will always be PI but if a cow has a non-PI calf she cannot be a PI.

At birth, PI calves can appear either clinically healthy or small, weak and ill-thrifty (Baker, 1995) and may show stunted growth and chronic ill thrift (Voges *et al.*, 2006). Furthermore, PI animals are regularly reported to be particularly susceptible to secondary infections (Voges *et al.*, 2006), suggesting poor immune function. This results in the fact that PI animals have a poor survivability rate (Houe, 1993). Only PI calves can develop mucosal disease, which is inevitably fatal. This disease appears after the acquisition of a cytopathogenic strain of BVDV that can occur with a mutation of a non-cytopathogenic BVDV biotype circulating in the PI animal or through infection by a cytopathogenic strain (Brownlie, Clarke and Howard, 1984).

### 2.1.2 Transiently infected (TI)

Animals that are infected by BVDV after birth or during the last trimester, when the immune system is able to fight the infection, develop a transient infection. A transient viremia will start approximately 3 days after the infection (Pedrera *et al.*, 2012) until immunity develops around 2 weeks later (Meyling, Houe and Jensen, 1990). The transient infection is most of the time subclinical but usually comes with a transient immunosuppression, especially in calves. After a transient infection, the immunity developed against the BVDV is considered to be lifelong.

### 2.1.3 Immune post infection

After infection by BVDV, all animals apart from PIs remain immune for the rest of their lives. After obtaining immunity, cows cannot produce PI calves anymore. It is worth adding that non-PI female cows that give birth to PI calves are always immune (seroconversion occur during gestation) and that immune female from natural infection before insemination will not give birth to PI calves.

### 2.1.4 Susceptible

Susceptible animals are animals that haven't been infected with BVDV and that have not developed antibodies. Hence, they are naïve (not immune). These animals can get infected and pregnant females can give birth to PIs if they are infected during the gestation window of susceptibility for development of PIs.

## 2.2 Course of infection

BVDV transmission can occur from different sources and through different routes of infection. There are two types of BVDV infections: infection after birth (i.e. horizontal) and in utero infection (i.e. vertical). This part will describe all the aspects of the course of infection by BVDV.

### 2.2.1 Sources and routes of infection

The BVD virus is shed through a wide range of body fluids: nasal discharge, urine, milk, semen, saliva, tears and foetal fluids (Meyling, Houe and Jensen, 1990). Faeces appears to be a poor source of virus but can be infectious (Brownlie *et al.*, 1987).

The most common means of transmission is from nose to nose contact with a permanently infected (PI) individual as they shed large amounts of virus. Although, they shed lower amounts of virus, transiently infected animals can also be involved in the transmission (Niskanen and Lindberg, 2003).

### 2.2.2 Infection after birth

Susceptible animals that are infected after birth become transiently infected. After immunity has developed, after around two weeks, they become immune.

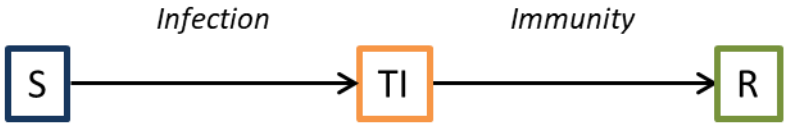



**Figure 1: Representation of the course of infection thought horizontal transmission: susceptible animal (S) can be infected by BVDV thought different means of infection and become a transiently infected (TI) bovine. Around 2 weeks post infection, as a result of immunity development this bovine will become immune to the virus and become recovered (R).**

### 2.2.3 In utero infection

When susceptible pregnant cows are infected they become transiently infected: the virus multiplies in the cow and can infect the foetus. The impact of the infection on the foetus depends on the stage of gestation. Usually, during the first 30 days post conception, embryonic infection leads to embryonic death (Moennig and Liess, 1995). Between 30 and 120 days of gestation (susceptible window for PI creation), before the development of the immune system in the foetus, infection can lead to the birth of persistently infected (PI) calves (Brownlie *et al.*, 1998). Later in pregnancy the effect of foetal infection is variable from no effects to teratogenic effects, foetal death and abortion. Foetuses that are immunologically competent at the time of infection can be born either transiently infected or immune. Recent work show a long term impact of pre-natal infection with many possible congenital defects in the central nervous system (Givens and Marley, 2013).

During their entire life, PI animals will shed large amounts of virus in all excretions and secretions: milk, semen, saliva, nasal secretion, urine, faeces, blood and aerosol (Brownlie *et al.*, 1987; Nettleton and Entrican, 1995). Only a small proportion of female PI calves reaches adulthood and gets pregnant. However, calves born of PI cows are also PI.

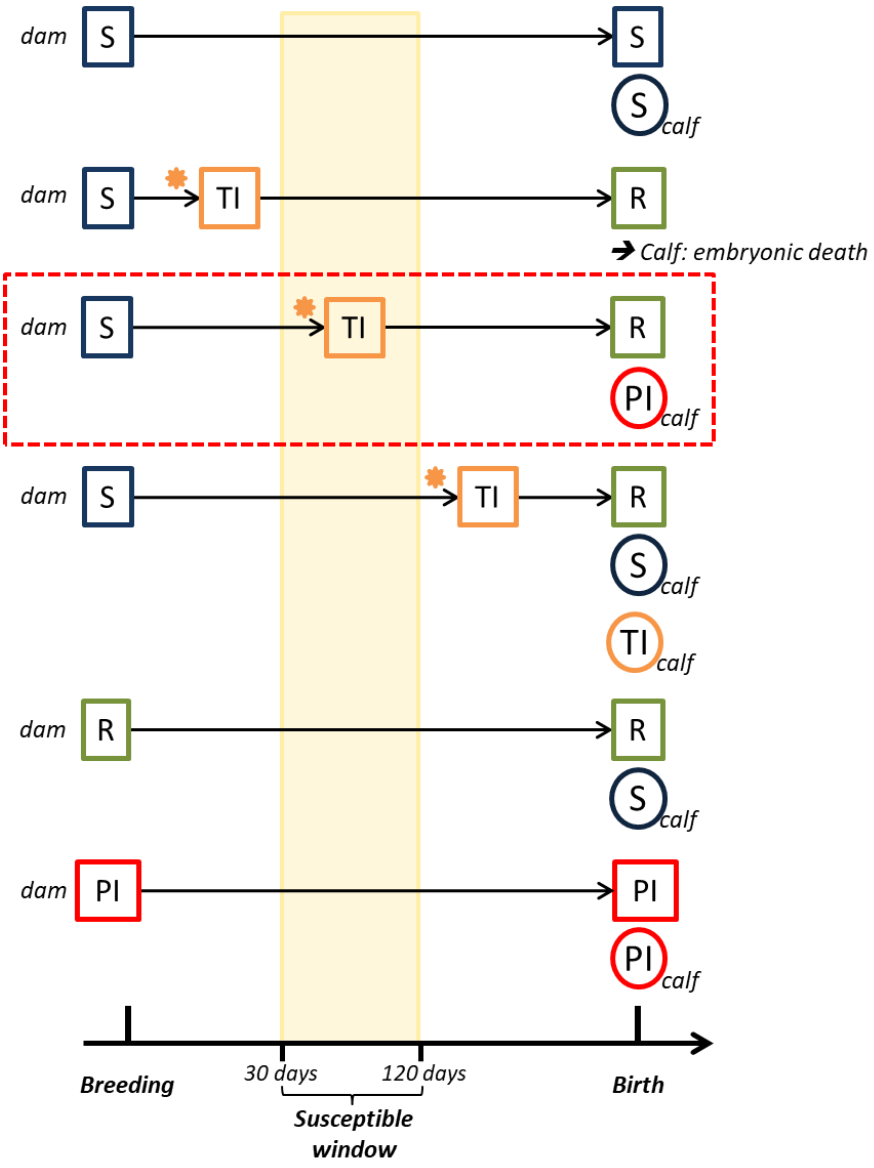



**Figure 2: Impact of in utero infection on the status of the calf at birth. Squares represent dams' statuses and circles calves' statuses at birth. The calf's status at birth depends on whether its dam got infected during the gestation and on the stage of gestation at which the infection occurred. No seroconversion during gestation leads to the birth of a susceptible calf, either while the dam is either S or R. Transient Infection during gestation can lead to different calf states at birth depending on the stage of gestation when infection occurs. Only transient infection of the dam that occurs during the windows of susceptibility (30 to 120 days of gestation) leads to PI calf. S: susceptible, TI: transiently infected, R: recovered, PI: persistently infected.**

### 2.2.4 Maternally derived immunity

New-born calves can acquire passively derived immunity against BVDV through serum antibodies present in colostrum (Moerman *et al.*, 1994; Chamorro *et al.*, 2015). The duration of this immunity can vary depending especially on the amount of antibodies ingested and absorbed (Fulton *et al.*, 2004) and can last for 3 to more than 6 months (Fux and Wolf, 2012; Fulton, 2013). The decline in maternally derived immunity over this period will increase the susceptibility of calves to acute infections. It is worth noting that passive maternally-derived immunity can modify diagnostic test results of PI calves as it can create false negative results particularly when testing blood samples for presence of BVDV by ELISA (Fux and Wolf, 2012).

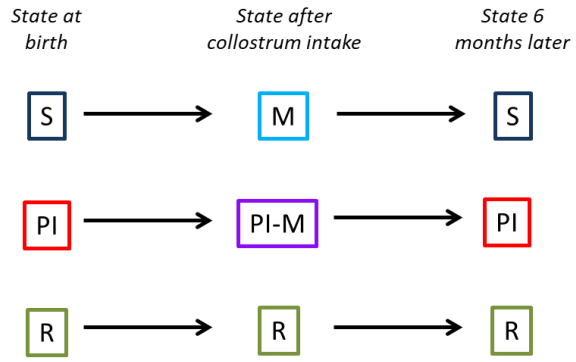



**Figure 3: Evolution of the calf's status at birth, after colostrum intake and 6 months later. After colostrum intake, susceptible (S) calves will be protected by maternal antibodies (M) for around 6 months and then will be S again. Persistently infected (PI) calves will stay PI after colostrum intake but with maternal antibodies (PI-M) for around 6 months. Finally in recovered (R) calves, colostrum intake will not change the status of the calf.**

### 2.2.5 Vaccination

In individual animals, the course of infection can be modified by vaccination. Vaccination against BVDV is mainly used to prevent transplacental infection of the foetus and thus to reduce the formation of more PIs (Frey et al., 2002; Patel et al., 2002; Meyer et al., 2012). Vaccines that contain both BVDV1 and BVDV2 strains are available. Two types of vaccines have been developed: inactivated and modified-live viral (MLV) vaccines. MLV vaccines lead to higher and quicker onset of immunity with a more consistent antibodies response and usually need only one dose for immunization, but some have the potential to create PIs if used in pregnant cattle.

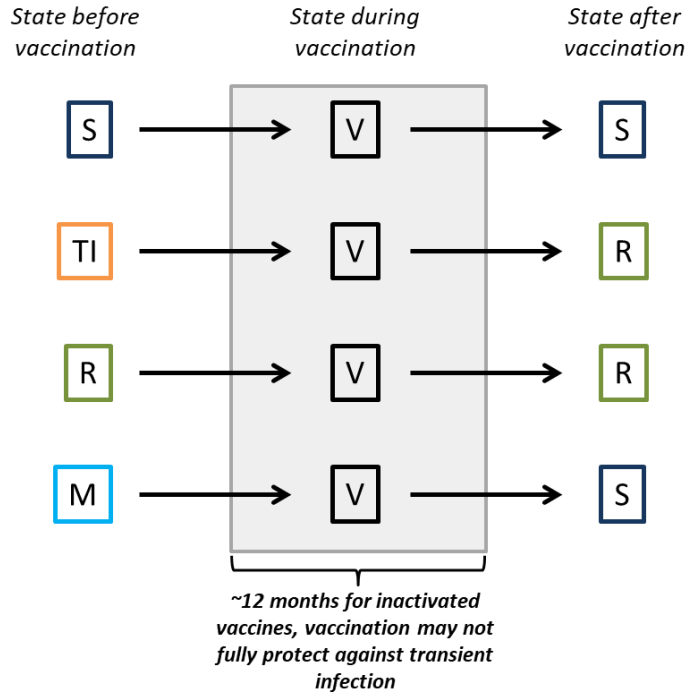



**Figure 4: Vaccination of cattle and results of vaccination in bovines from different statuses. After vaccination susceptible (S), transiently infected (TI) and immune (R) and protect with maternal antibodies (M) bovine become vaccinated (V). Vaccinated transiently infected and immune bovine will become immune when susceptible and protect with maternal antibodies bovine will become susceptible once duration of protection thanks to vaccination is over. Vaccination of dams with a specified vaccine should prevent transplacental infection and production of PI cattle by vaccinated dams should be limited under the period of protection.**

## 2.3 Available information at individual animal level to observe the system

Available tests can be divided into two groups: tests that detect an on-going infection through the detection of the virus (viral antigens: Ag ELISA or viral RNA: PCR) and tests that detect an immune response against the virus through the detection of circulating antibodies (Ab ELISA). Both can be used for the diagnosis of BVDV depending on the purpose and context.

Sensitivity and specificity of Ag ELISA and PCR are relative to virus isolation as the gold standard. Sensitivity and specificity of Ab ELISA are relative to serum neutralization test as the gold standard.

### 2.3.1 Antigen detection tests: ACE

Antigen-capture ELISA tests (ACE) need to target a highly conserved Ag across BVD strains. Two tests have been developed against two BVDV proteins: NS3 (formerly p80) and $E^{RNS}$ (formerly $E^{0}$). The approved samples that can be tested with ACE are: serum (plasma), tissue (skin biopsy, ear notch) and individual milk samples.

As an ACE detects viral antigens, this test is able to detect infected animals that shed the virus: TIs and PIs. TIs can be challenging to detect as they shed lower amounts of virus during a short time period. Using RT-qPCR as a reference test, an Ag ELISA test was able to detect only 10 out of 57 TIs but correctly detected 17 out of 17 PIs (Hanon *et al.*, 2014). However, once an Ag ELISA returns a positive result, interpretation of the state of animal, without any other information will be TI or PI as this single test is not able distinguish between them (Hanon *et al.*, 2014) but the test value can be predictive of the state of infection. Repeating the test three weeks/ 1 month later will clarify whether the animal is TI (negative Ag-ELISA) or PI (again positive Ag-ELISA).

### 2.3.2 Nucleic acid detection: RT-qPCR

Reverse transcriptase polymerase reaction (RT-PCR) (Hertig *et al.*, 1991) is widely used for BVDV diagnosis. A wide range of samples can be used in these test: blood, milk, saliva and tissue (Bhudevi and Weinstock, 2003; Kim and Dubovi, 2003; Kliučinskas *et al.*, 2008). Moreover, some RT-PCR tests can distinguish BVDV type I and type II (Letellier *et al.*, 1999). Quantitative RT-PCR (qRT-PCR) has been developed for BVDV diagnosis, as there exists a relationship between threshold cycle (CT), cycle number at which the fluorescence generated is higher than the threshold, and the quantity of viral RNA present (Bhudevi and Weinstock, 2001). qRT-PCR can be used to make a distinction between TI and PI in term of CT, knowing that PI will shed a larger quantity of virus.

### 2.3.3 Antibody ELISA (enzyme-linked immunosorbent assay) (Ab ELISA)

Ab ELISA is an immune-enzymatic technique that allows the detection of antibodies in a sample. Sample antibodies will bind specifically to an antigen present on a surface and the binding will be visualised following an enzymatic coloured reaction. Relative to the SNT, specificity and sensitivity of Ab ELISA for BVDV detection is high: up to 99% and 98% respectively (Cho *et al.*, 1991; Kramps *et al.*, 1999; F. Beaudeau *et al.*, 2001) . Both serum and milk can be used as matrices. This test is also able to detect, but not differentiate colostrum-derived antibodies in suckling calves (Fux and Wolf, 2012).

A positive Ab-ELISA can be associated with either an immune state resulting from a natural infection, the presence of maternal antibodies in calves under 6 months or with vaccination. A single test result may not be able to distinguish those three categories. However, repeated testing can clarify the true BVD status in that maternal and vaccination derived antibodies will decrease with time.

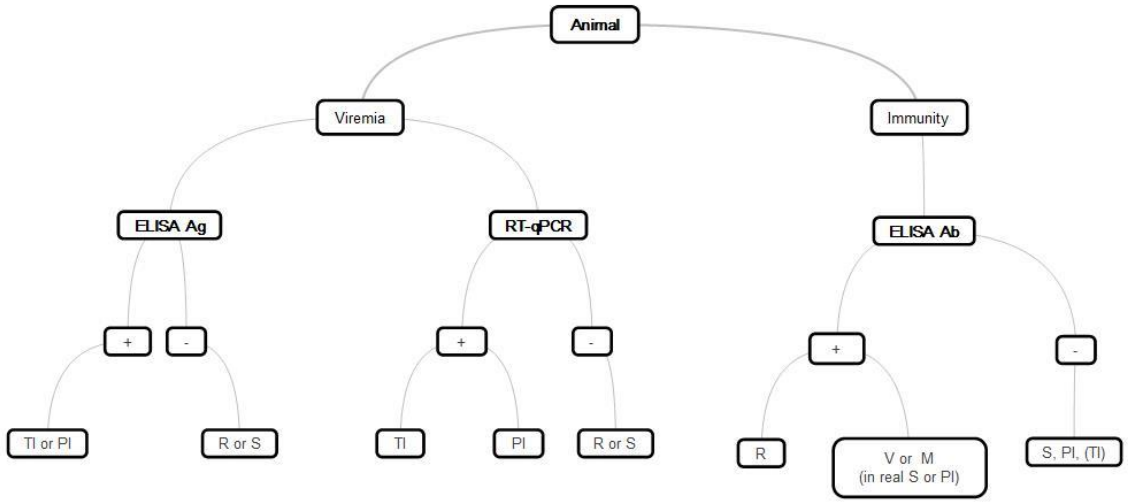



**Figure 4: Representation of test result interpretation at the animal level with a single test.**

### 2.3.4 Test combinations

Test combinations can be defined as: different diagnostic tests at the same time or several identical tests at different times on the same animal, or both. Different diagnostic tests can inform both the immunological and viremia state of the animal. Several identical tests can inform about the evolution of the animal's state. For example, two ELISA Ag test results with an interval of 3 weeks can discriminate between TI and PI. Two positive Ag ELISA results identify/indicate a PI animal while a positive and a negative result identify/indicate a TI.

# 3 Conceptual model for BVD at the herd level

The probability of a herd selling an infectious (PI or TI) animal depends on both the probability of this herd having introduced the infection as well as on the within herd dynamics of the infection once it has been introduced and on the ability to detect the change. BVDV introduction can occur through different routes (i.e. purchase, contact at boundaries fences ...). Once the infection has been introduced, the within herd dynamics depends on herd demographic and contact structures as well as on herd management. Important differences exist between beef and dairy herds which need to be taken into account. These differences can be represented in terms of herd structure and herd management.

## 3.1 Epidemiological state at the herd level

There are, at least four different states at herd level depending on the situation of the herd regarding BVDV infection.

### 3.1.1 Virus free and seronegative herds

Naïve free herds are herds that are not currently infected and that have not been recently (in the past +/- 10 years) infected by BVDV. They are composed of susceptible cattle that are not immune against BVDV (S=100%).

### 3.1.2 Herd infected with at least one transiently infected animal (absence of persistently infected animal, either alive or in the foetal stage)

Herds in this category are infected by at least one transiently infected animal. They are composed of S and TI animal and as the herd infection progresses the proportion of S and TI declines and R cattle will rise. In general, immune animals (R>0) will be found in these farms.

### 3.1.3 Herd infected with at least one persistently infected animals (either alive or in a foetal stage)

Persistently infected herds contain at least one persistently infected animal alive or to be born (Trojan cow). They are composed of susceptible, transiently infected and at least one persistently infected animal and as the herd infection progresses by an increasing number of immune cattle. In general, immune animals (R>0) and transiently infected animals (TI>0) will be found in these farms.

### 3.1.4 Virus free and partly seropositive herd (at least one animal is seropositive)

This state occurs:

- When all infectious animals (PI, TI) are removed (by death, sale, converted to recovered animals) and there are still animals with antibodies (R) present. (R>0; TI=0; PI=0);
- after vaccination of parts of the herd (V>0; TI=0; PI=0)
- by a combination of both (R>0 & V>0; TI=0; PI=0).

Herds in this state can become “virus free and seronegative herd” once all the immune and vaccinated animals have left the herds.

## 3.2 Course of infection at herd level

### 3.2.1 Risk factors for introduction of BVDV in a herd

BVDV introduction into a herd can occur through different routes. Those possible routes are described and quantified in terms of probability of transmission between herds in table in Annex I. We can separate them into 3 categories: introduction of infectious animals, contact with infected animals direct from another herd and indirect transmission through contaminated material (biological or equipment).

#### *3.2.1.1 Purchase and introduction of infectious animals*

##### 3.2.1.1.1 Persistently infected

Introduction of a PI to a herd (directly or through a Trojan cow, see below) is the main source of introduction of BVDV in a herd in endemic situations in the absence of control measures. As a PI will shed a high level of virus (Brownlie *et al.*, 1987) throughout its entire life, transmission/infection of the herd can occur quickly, continuing whilst the PI animal remains in the herd.

##### 3.2.1.1.2 Trojan cow

Trojan cows are non PI, immune cows that have been transiently infected by BVDV during their first semester (day 40 to 120) of pregnancy. As a TI individual, the dam will clear the infection in approximately two weeks but will carry a PI foetus. These cows can be sources of new infections in a herd at the birth of the PI calf. As the dam is healthy and immune, Trojan cows are a high risk for introduction or reintroduction of BVDV.

##### 3.2.1.1.3 Transiently infected

Transiently infected animals can be a source of introduction of BVDV into a herd. Nevertheless their relative role as a source of infection is much lower than the role of the PI animal. In fact, they shed lower amounts of virus and the period of shedding is short (around 2 weeks). The relative importance of TI in (re-) introduction of BVDV in a herd is under discussion: some argue that TIs are unlikely to be a source of infection (Niskanen, Lindberg and Tråvén, 2002; Sarrazin *et al.*, 2014) while others suggest that BVDV can be maintained in a herd without presence (or at least identification) of PIs (Moen, Sol and Sampimon, 2005).

#### *3.2.1.2 Contact with animals from neighbouring/other herds*

Direct contact with infected cattle from another herd is also an important means of introduction of BVDV. These contacts can occur through shared grazing or adjacent herd pasturing areas, animal shows and markets. Annex I Tables I and II list animal contact on pasture or across boundaries as a risk factors for introduction of BVDV, especially when the susceptible cattle comes in contact in early pregnancy (at risk of producing a PI). A survey in Danish dairy herds showed that contact with cattle from another herd and pasturing within 5m were positively associated with seroconversion to BVDV (Houe, 1999).

#### *3.2.1.3 Person contacts*

Introduction of BVDV in a herd can also occur through indirect transmission by contaminated persons, when they have contact with animals (e.g. veterinarian, farmers, claw cutters, inseminators). It is essential that persons that have contact with the animals follow strict hygiene rules.

#### 3.2.1.4 Contaminated materials and products

Introduction of BVDV in a herd can also occur through indirect transmission through contaminated products or materials. Compared to direct contact with infected cattle, indirect routes may play a minor role in transmission. However, towards or at the end of an eradication programme, when introduction of BVDV through purchases and contacts is limited/rare, indirect transmission can become relatively more important (Hult and Lindberg, 2005). BVD virus can be preserved in cryopreserved semen of infected bulls, so artificial insemination with contaminated semen of PI and TI can lead to dam infection (Meyling and Mikél Jensen, 1988; Rikula *et al.*, 2008). Other products like contaminated vaccine or contaminated veterinary materials like needles and tongs can also lead to new infections (Gunn, 1993; Niskanen and Lindberg, 2003). Finally, sharing equipment e.g. trailers during transport can also be a source of infection.

### 3.2.2 Within herd dynamics of BVDV

Within-herd spread can be influenced by herd management once BVDV has been introduced. First, the course of infection within a herd after introduction of BVDV will be presented. Then, herd factors that can influence spread will be described.

#### 3.2.2.1 Representation of the course of infection within a herd

The course of infection within a herd will start with the introduction of BVDV through different routes (cf. part below: Introduction risk factor for BVDV). Depending on the route of introduction, the proportion of bovines newly infected in the herd can vary. Once an infectious animal is introduced it will shed virus and infect other animals. Then, those newly infected animals will also shed virus and in turn will infect other susceptible animals. If some infected cows are pregnant between 30-120 days of gestation, the infected foetus will become a PI calf. If nothing is done to limit the infection, the virus will continue to spread within the herd with a negative impact on reproduction. After a while, a large proportion of cows within the herd will become immune against BVDV and will no longer be susceptible to BVDV.

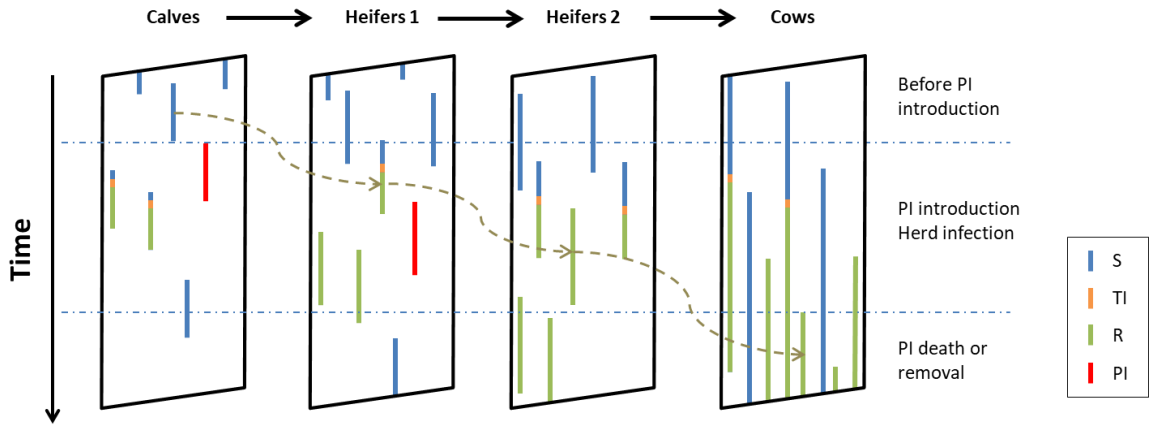



**Figure 5: Representation of the course of infection after the introduction of a PI in a virus free and seronegative herd. This example herd is composed of 4 groups of animals: calves, heifers 1, heifers 2 and cows. Verticals lines represent individual bovine and the length of the line the time spend in each category. As time passes, animals will move to other groups. Grey dotted arrows represent movements between groups through time. Once a PI animal is introduced to the herd the infection spreads in all groups. The transmission thought different groups within a herd is linked to the herd structure and how animals from different groups are separated.**

#### *3.2.2.2 Herd structure and management that influence the course of infection*

Herd husbandry can affect infection transmission between groups of animals. Herd characteristics will be explained here: type, structure, contact structure and management.

##### 3.2.2.2.1 Type of herd

Only breeding beef and dairy herds through generation of PI calves will be considered in this work because they will be the main source of infection for other farms. Farms that specialise in fattening cattle will not be studied because the animals they sell are sent directly to slaughter and not to other farms. However, fattening units within beef and dairy farms can act as sources of infection for breeding units on those farms, and will therefore be included as risk factor.

##### 3.2.2.2.2 Herd contact structure

In cattle herds, animals live in separate groups. For example, in dairy herds, there are usually groups of calves, heifers and lactating cows. Animals in the same group have higher probabilities of contact than animals in separate groups. Furthermore, within a herd, the different groups can have more or less contact. For example, in beef herds, calves stay with their dam until weaning which can happen at up to 9 months, whereas in dairy herds, calves and dams are quickly separated. This results in PI calves being in close contact with the breeding herd for much longer in beef than in dairy herds. Within a herd, the different groups can be kept apart in different barns or on different pastures. The separation between groups can be quite different and herd specific.

Herd level structure associated with a risk of introduction, within-herd transmission or persistence of infection of BVDV:

- Size of the herd (or number of cows as a proxy) (Graham *et al.*, 2013)
- The age at which calves are separated from their dam. In breeding herds, calf stay with dams until weaning, meaning that if the calf is a PI it can transmit infection to other dams during the risk period of early pregnancy. In dairy herds, calves are separate from dams at birth meaning that transmission can only occurs between calves.
- The age at first calving: in dairy herds, the age at first calving is usually 24 months while it may be up to 36 months in beef herds. This implies that there are at least 2 groups of heifers in dairy herds and 3 in beef herds.
- The replacement rate determines the proportion of female calves born on the farm that are kept to replace breeding cows. The lower the replacement rate is, the higher the probability that a present PI calf is sold rather than kept as replacement.
- The proportion of time spent at pasture for the different groups determines the probability of being in contact with animals from neighbouring herds. Inversely proportional to the time spent indoors.
- Number of neighbouring herds (Graham *et al.*, 2016)
- Within herd biosecurity : how many barns for the different groups

##### 3.2.2.2.3 Herd Management

Some farm management practices are of major importance in the dynamics of BVD in these herds.

- Calving distribution can be seasonal, that means that all calving will occur in a short period (3 months). This is associated with most pregnant cows being in the window of susceptibility

for the formation of a PI calf at the same time. If an infectious animal is in contact with these cows during the window of susceptibility, this can result in a high number of PI calves. However, seasonal calving allows the identification and removed of pre-breeding cattle. Conversely, extending calving and breeding means that a PI born at any time of the year may have the opportunity to contact a pregnant animal in early gestation.

- The number of cattle purchased by the herd can lead to multiple re(-introduction) of BVDV in the herd
- Within herd biosecurity and hygiene measures can limit the spread of infection.
- Location: separation of calves from pregnant animals

#### 3.2.2.2.4 Vaccination

Vaccination can modify the course of infection within a herd as it can reduce or prevent in utero infections and limit the production of PIs. Cattle vaccination will also impact monitoring options for the presence of BVD infection in herds. With vaccination, animals may produce antibodies and all screening which is based, for example, on surveillance of antibodies in cows through bulk milk tank testing, may not be possible for a certain time period depending on the vaccine used and original immune status of the herd. Moreover, farm-level information about vaccination against BVDV is not readily available.

#### 3.2.2.2.5 Differences between dairy and beef herds

Table I shows differences between dairy and beef herds linked to risk factors for spread of BVDV that have been described above.

Table I: Example of the main differences between dairy and beef herds in France

| Risk factor | Dairy herds | Beef herds |
|---|---|---|
| **Age at which calves are separated from their dam** | At birth | After weaning |
| **Age at first calving** | 24 months<br>(at least 2 groups of heifers) | 36 months<br>(at least 3 groups of heifers) |
| **Replacement rate** | ~25-40% | ~20-30% |
| **Proportion of time spent at pasture** | Depend on :<br>Herd size<br>Region | Often<br>Can be Seasonal |
| **Vaccination** | Variable | Often |
| **Seasonal calving** | Often none | Often |

## 3.3 Available observations at herd level

At the herd level there are three types of available observations. First, risk factor information that can explain possible causes of infection wihtin the herd. Then, results from biological tests will inform about the consequences of this infection. Finally, factor that can lead to delayed detection

### 3.3.1 Risk factors associated with herd characteristics

Herds can be described by several characteristics that can be involved in BVDV dynamics. These characteristics can inform about the contact structure within the herd and with neighbouring herds and can be linked to the risk of introduction or the risk of transmission within the herd.

#### 3.3.1.1 Risk factors for BVDV introduction

- Number of neighbouring cattle herds with common boundaries, size of the common boundaries
- Number of cattle purchased and number of purchased animals that are pregnant?
- Time spent at pasture
- Biosecurity measures for professional and visitors (i.e. farmers, veterinary, AI technicians, traders)
- Herd size

#### 3.3.1.2 Risk factors for BVDV transmission within a herd

- Surface area (km²):
    - Building
    - Pasturing area : a large surface of pasturing area can increase the number of potential neighbours
- Number of animals
- Density of animals (km²)
- Calving distribution
- Contact structure within the herd (individual within herd biosecurity)
- Age at which calves are separated from their dam

### 3.3.2 Results from biological tests

Herd diagnosis can be conducted at the level of the individual or a group of animals. It can involve testing of samples individually (refer to part 2.3.Available information at individual animal level to observe the system) or in pools. Depending on context and territory, programmes can have different aims and lead to different screening strategies. Furthermore herd type can impact the strategy used. The main difference between beef and dairy herds is the use of bulk tank milk (BTM) to monitor BVDV infection. Diagnostic strategies at the herd level involve antibody detection and virus detection.

#### 3.3.2.1 Detection on Bulk milk test

##### 3.3.2.1.1 Ab ELISA

Monitoring BTM can detect seroconversion of a herd with Ab-ELISA. The level of BVDV antibodies in milk can even be correlated semi-quantitatively to the prevalence of seropositive animals in the dairy cows (F Beaudeau *et al.*, 2001; Eiras *et al.*, 2012b). Depending on the test and the assessment objective, the interpretation and the threshold value of tests can be quite different. Ab tests are widely used for herd diagnosis. Aims are either: (i) provide on-going evidence of freedom thought repeat negative tests or (ii) detect introduction of infection.

Ab levels in BTM can give an indication on the prevalence of sero-positive cows in the dairy herd (F Beaudeau *et al.*, 2001; Eiras *et al.*, 2012a) and variation in Ab levels can indicate a new infection of the herd (F Beaudeau *et al.*, 2001). One rare limitation to this test is if a PI contributes to the BTM : antigens can neutralize antibodies and can cause a negative BTM Ab ELISA (Sandvik, Larsen and

Nyberg, 2001) but this is not a major risk as PI often die or are removed before adulthood and requires a specific ratio of PI/seropositive cows. Serial testing allows the observation of the evolution of the herd status over time. Figure 9 show an example of infection of a herd and evolution of Ab-ELISA on BTM results.

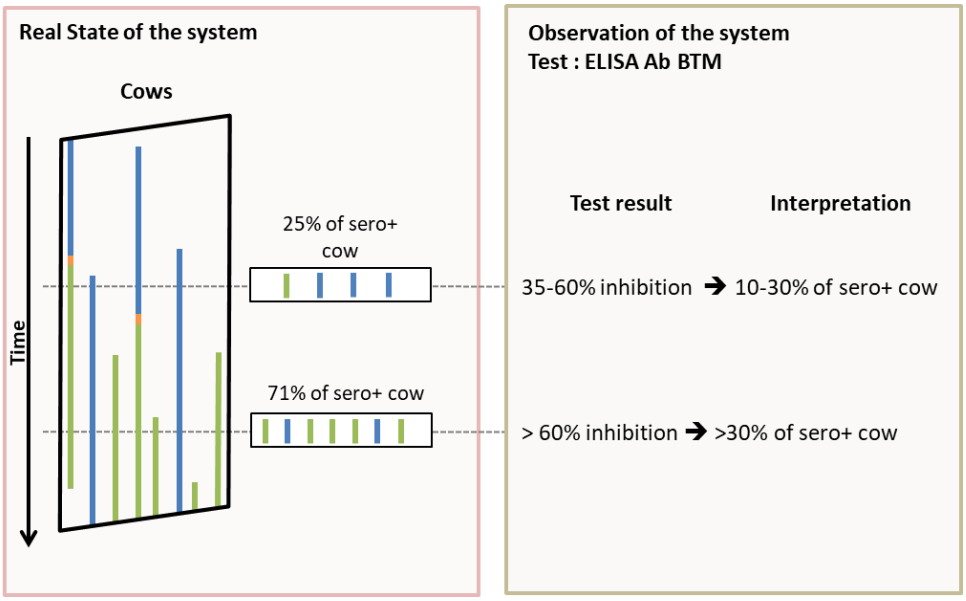



**Figure 6: Example of Ab-ELISA test on BTM in cows. This diagram on the left is from the figure 5 and represent the cows compartment in a herd each vertical line represent one animal which can be susceptible (in blue), transiently infected (in orange) and immune (in green) bovines. This diagram represents the link between the real state of the system (meaning the proportion of immune animals) and available observation at herd level based on an ELISA BVDp80 kit (example of Brittany,(F Beaudeau et al., 2001)).**

However, as this test relies on detection of antibodies to BVDV, its value is reduced in herds which apply vaccination against BVDV.

#### 3.3.2.1.2 PCR

PCR can also be used on BTM (Muñoz-Zanzi *et al.*, 2000). In practice RT-PCR on BTM is very sensitive as it has been proven that this test can detect 1 PI in a herd of 132 cows (2 in a herd of 800) (Drew, Yapp and Paton, 1999; Renshaw, Ray and Dubovi, 2000; Hill, Reichel and Tisdall, 2010).

#### 3.3.2.1.3 Spot test detection

##### *3.3.2.1.3.1 Pooled milk*

Ab-ELISA on pooled milk can be applied to cows that provide milk but that do not contribute to BTM, or can occasionally be applied to beef herds. This test can also be used to screen a specific age group (i.e. early lactation) as a negative result provides evidence of freedom even if BTM is positive. This test has the same limitations as the BTM testing.

##### *3.3.2.1.3.2 Pooled serum sample*

Ab-ELISA on pooled serum samples is usually applied to young stock and non-breeding beef cattle. This test is useful to predict presence or absences of PI in a dairy herd where BTM or first lactation tests are positive. Young stock will become Ab negative after the decrease of maternal antibodies,

after/at around 6 months of age. Testing those animals will give crucial information on the current situation within the herd. This variation highlights the importance of selection of animals for testing. In fact, the selection of animals tested is fundamental. Recently purchased animals have to be excluded from the test group and each separate group must be tested.

A PCR can also be applied on a pooled serum sample (Muñoz-Zanzi *et al.*, 2000). In such samples, PCR may be able to detect any individual infected up to a pooled sample of 50 individuals (Smith *et al.*, 2008; Yan *et al.*, 2011).

#### 3.3.2.1.4 Vaccinated herd

Vaccination can limit detection of current infections when using tests based on Ab detection. A solution is to use unvaccinated sentinels and test them for Ab detection. Pillars and Grooms (Pillars and Grooms, 2002) have shown that serological testing of unvaccinated heifers within a vaccinated herd can be used to detect the presence of PI in a herd with a sensitivity and specificity of respectively 66% and 100%. Ab titers can also be useful to distinguish vaccinated herds with and without the presence of a PI animal. Houe *et al.*, 1995, show that the screening of 5 young stocks can distinguish vaccinated herd with or without PI. The probability to find at least 3 of (in) 5 animals with higher titer in a herd where killed-virus vaccine was used and in the absence of PI, was P<0.01, while it was P>0.99 in a similar herd in the presence of a PI.

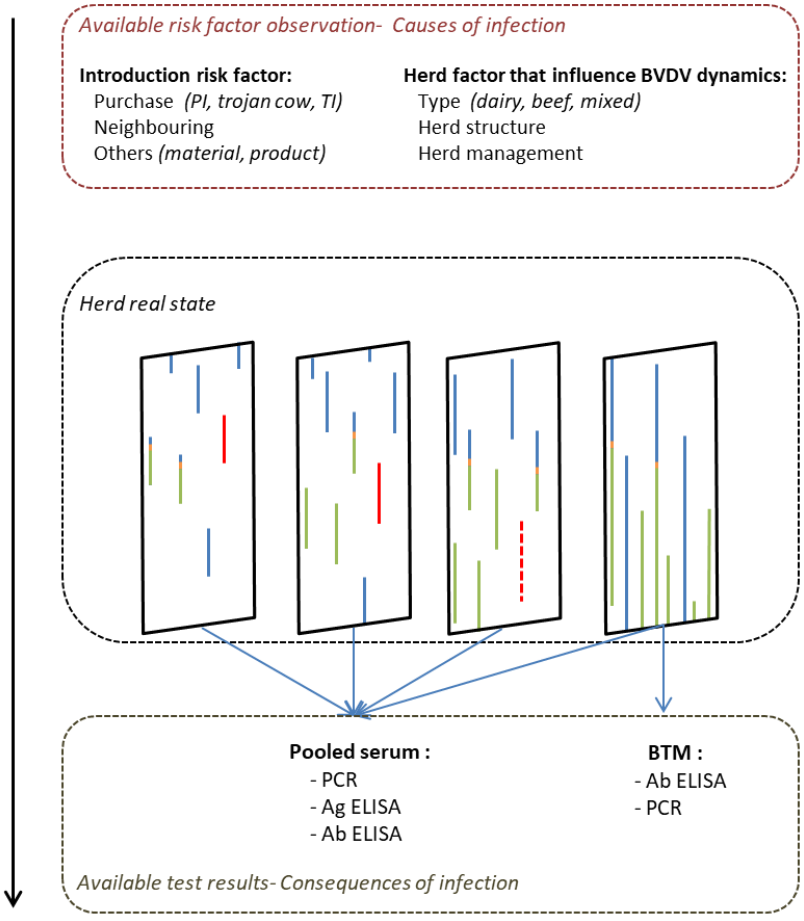



**Figure 7: Available information at herd level: causes and consequences of infection by BVDV within a herd.**

### 3.3.3 Delayed detection of BVDV in a herd

The time interval between the introduction of the infection and its detection can vary. As this interval increases the probability that infection spreads within the herd and between herds increases. The lengthening of this interval is what we call delayed detection.

Delayed detection can be associated with the design of the testing procedure such as test frequencies; animals tested and test performance (test characteristics).

#### 3.3.3.1.1 Surveillance programme design

Surveillance programme design will determine how the presence of infection within the herd is detected. In fact, if the test is carried out just after introduction of a PI in a herd, depending on the test used, the result can be negative as the infection will take some time to spread in the entire herd over time and for individual animals to develop detectable levels of Ab. Serial testing allows herd monitoring and an increased frequency of testing can limit delayed detection. The length of time between two screening tests impacts the risk of delayed detection. Keeping the time low will reduce the risk of delayed detection.

#### 3.3.3.1.2 Test performance

Test performance can contribute to delayed detection when sensitivity is not 100%. In general, test sensitivity is quite high for BVDV diagnostics, which should limit the impact of test performance to cause delayed detection. However, it worth noting that applications of these tests may result in a reduced sensitivity and specificity because of human errors during the diagnosis process (sampling, labelling, laboratory errors …).

# 4 Conceptual model for BVD at the territory level

Territory level can be either a region or a country depending on the way that BVDV eradication/control is managed. A territory is defined as an area where herds follow the same control measure (programme) and where information is gathered together. A territory has one BVD programme which can be based on different components (i.e. different component for dairy and beef herds).

Within the consortium, the BVDV control programmes for the Netherlands, Sweden, Ireland and Scotland, are applied on country level. For Germany and France, because of variability of types of herds, territory level will not correspond to country. For Germany it will be the Federal States and for France a region, or even a department.

## 4.1 Epidemiological state at territory level

There are at least three different BVD states at territory level.

### 4.1.1 Infection free and seronegative territory

A infection-free territory is defined as a territory composed of seronegative herds that are currently not infected by BVDV and where all cattle are susceptible.

### 4.1.2 Territory with infected herds

An infected territory is defined as a territory with at least one infected herd(s) meaning that the infection is present or spreading within the territory. In this defined territory, herds can be naïve and infection free, currently infected or seropositive (some or all animals). The proportion of herds in each state depends on the prevalence of BVDV infection and the control measure in place (endemic territory versus on-going eradication programme). Over time and depending on the contact between herds within the territory and the actions taken to trace and eradicate infected animals these proportions can change.

### 4.1.3 Post-eradication territory: Infection free and seropositive territory

A post-eradication territory does not have any infected herds within but can be composed of seropositive and seronegative herds.

## 4.2 Risk of introduction of BVDV into a territory

### 4.2.1 Cattle movement

As for herd level, cattle movement through purchase and market outside of a territory can be sources of (re-)introduction of BVDV into a territory. As PI animals are the main source of (re-) introduction of BVDV, purchasing young animals or pregnant dams (with a chance of being a Trojan cow) is particularly risky. More information about that risk can be found under herd level risk factors (section 3.2.1.1)

### 4.2.2 Infection prevalence in neighbouring territories

Infection prevalence in neighbouring territories can also be a risk factor for introduction of BVD within territories, when cattle are moved to/through or grazed in the neighbouring territory.

### 4.2.3 Wildlife (reservoir)

BVDV have been reported for over 40 different species, including domestic and wildlife species (Nelson *et al.*, 2016). As in domestic cattle, BVDV can induce persistent infections in 8 other species: white-tailed deer, mule deer, eland, mousedeer, mountain goats, alpacas, sheep, and domestic swine (Terpstra and Wensvoort, 1997; Scherer *et al.*, 2001; Duncan *et al.*, 2008; Nelson *et al.*, 2008; Bachofen *et al.*, 2013). Sources of infection for non-bovine species can be a spillover from cattle population by sharing environment or through direct contact (Nelson *et al.*, 2016). Despite this, infection through wildlife is not considered a major cause of introduction.

## 4.3 Within territory dynamics of BVDV

### 4.3.1 Territory representation

A territory is defined by an area where herds within the area follow the same programme against BVDV. Important territory characteristics that can vary from one territory to another and influence BVDV dynamics:

- Proportion of beef and dairy herds : as practises differ between beef and dairy herds
- Cattle density, herd density and degree of fragmentation of farms may influence the contact structure and potential contact between herds within the territory. Intensity of contact between herds can influence the transmission between herds once BVDV have been introduced.
- Purchase: the proportion of herds that purchase at least one animal and the total number of purchases can also influence transmission of BVDV once it has been introduced in the territory. If the proportion is high the transmission between herds is likely to be high.
- Infection prevalence within territory

## 4.4 Available observations at territory level

### 4.4.1 Territory structure information

Available information linked with territory structure can be listed as:

- Number of herd within the territory (number of herds)
- Density of herds within the territory (number of herds/km²)
- Surface area of the territory (km²)
- Proportion of dairy and beef herds (%)
- Infection prevalence of neighbouring territories
- Number of cattle purchased from outside of the territory and their source
- Participation in market/trade shows either inside or outside of the territory with participant from everywhere
- Information linked to wildlife: if available (qualitative/quantitative data) as the role of wildlife in BVDV dynamics is not considered significant.

### 4.4.2 Territory BVDV programme (surveillance/eradication)

Information at territory level will be derived from aggregation of observations at herd level. Seroprevalence of BVDV at territory level can be available. Programme information will also be available and will help to estimate the situation of the territory. First, the programme can be defined as compulsory or voluntary. In voluntary programmes not all herds within the territory are likely to

be involved in the programme. In this case information for herds that are in the territory but outside the programme can be missing. Vaccination campaign features like type and name of the vaccines used within territory may be also available.

Tests used in the programme will also be available. Type (sample and test) and performance (sensitivity and specificity) parameters will be provided. Other parameters involved in programme features such as time in between tests or group of cattle tested is referred at herds' level.

### 4.4.3 Delayed detection of BVDV in a territory

Delayed detection of BVDV within a territory can be linked to the efficiency of the surveillance/control programme. Depending on features of the programme, BVDV can spread within territory without being detected for a certain time period

#### 4.4.3.1.1 Voluntary versus compulsory programme

Surveillance programme can be compulsory, meaning that all herds are participating, or voluntary meaning that only some herds of the territory are participating. If BVDV transmission occurs in non-participating herds in voluntary programmes, detection of BVD in participating herds can take more time than in a territory with a compulsory programme because it will take more time to find the herd at the source of infection. We can imagine that higher percentages of participation in the programme result in a lower risk of delayed detection.

#### 4.4.3.1.2 Programme design

Characteristics of a programme may influence delayed detection. Performance of the chosen test (sensitivity and specificity), length of time between the test and group of animal tested can lead to misclassification. It case of false negative that can lead to delayed detection.

# 5 Conclusion

This work is a first step towards estimating probabilities of freedom from infection and their associated uncertainties from information that is diverse and heterogeneous. The conceptual models presented here map various pieces of information onto BVDV infection systems,which represent infection by the virus at animal, herd and territory level. The next step will be to translate these conceptual models into statistical models. From a statistical point of view, the challenge of translating heterogeneous inputs at different scales into uniform output has some specific characteristics that are important for the choice of the method to use.

The first feature of the challenge is its structure. This can be illustrated using the example of the probability for a calf of being born PI (Figure 9). The dam of such a calf, while susceptible to infection before her pregnancy, will have been infected by the BVDV during the window of susceptibility of her pregnancy (30-120 days of gestation). Therefore, she would have tested negative for an antibody test before the pregnancy and positive after. Omitting interference with the colostrum, the calf would multiply the virus and would not produce antibodies against it. This calf would therefore test positive to antigen ELISA or PCR tests and negative to antibody ELISA tests. With this example, we see that the information (test results), can be mapped onto the probability for a calf of being a PI using a conceptual representation of the infection epidemiology. Given the calf status, test results can differ depending on test characteristics as measured by sensitivity and specificity, but the underlying representation of how they are connected will not.

The second feature of our problem is the heterogeneity in the data. Under some control programmes, BTM antibody results are measured at regular intervals allowing the detection of herd seroconversion and thereby the probability of PI calf births. Other control programmes look for virus antigens or RNA in new-born calves. Going back to Figure 9, we would have information on either the dam statuses before and after pregnancy or on the calf status after birth. This heterogeneity in input can be a difficulty when our aim is to estimate a probability that is independent from the available data.

For statistical modelling, we will turn to Bayesian methods. A Bayesian representation of our problem will allow us to address these two features. In Bayesian statistics, models can be represented using directed acyclic graphs (DAG). Figure 9 is a DAG. Each box in the DAG (called a node) is connected to one or more other boxes with arrows (called edges). Each node can contain either observed data (test result) or unobserved/missing data (calf infection status, dam infection status, no test results (if no test has been performed)). The DAG describes the relationships between these nodes. Bayesian models go further by assigning statistical distributions to nodes and by providing mathematical descriptions of the relationships between nodes. For example, the node *PI calf* in Figure contains the calf's infection status. The calf is either PI or not PI. Being in one of two mutually exclusive categories is modelled using a Bernoulli distribution, which has a parameter $p$ which is the probability for this calf of being PI. This probability $p$ can be made dependent on the dam's status. In turn, the calf's status influences the probabilities of testing positive to specific tests, which will depend on test sensitivities and specificities. Bayesian models allow chaining these relationships within a single model.

The next step of our work will therefore consist in translating our conceptual models into Bayesian statistical models and then to parameterise these models so that they estimate a probability of freedom from infection regardless of the heterogeneity in input.

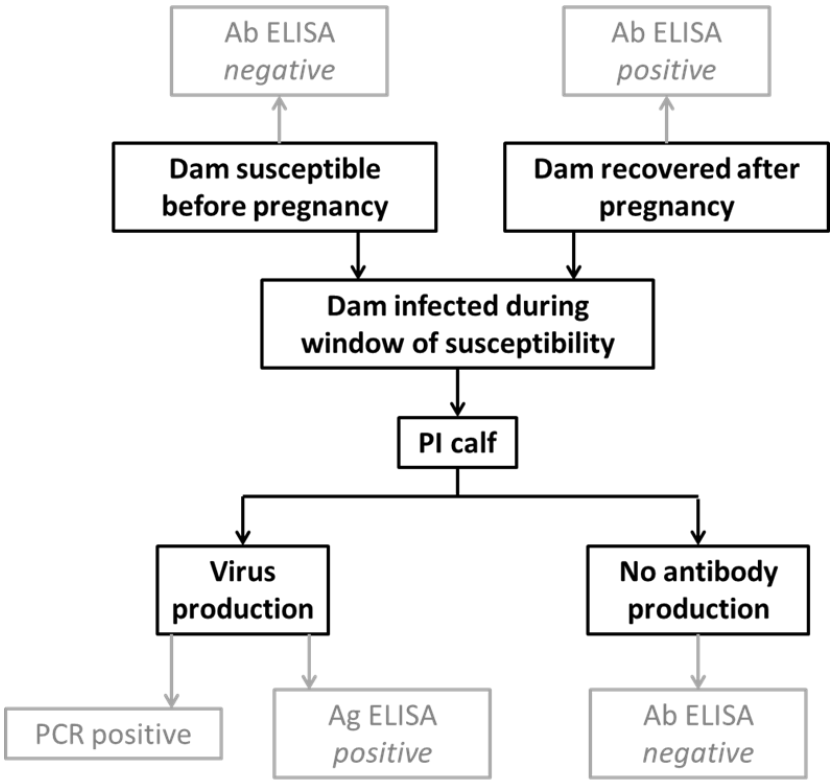



**Figure 9: Representation in a DAG (Directed Acyclic graph) of the causes and consequences of a calf being born PI (black) and examples of associated test results (grey), omitting interference with maternally derived antibodies. Test results can differ depending on test characteristics.**

# 6 Bibliography


Bachofen, C. *et al.* (2013) 'Persistent infections after natural transmission of bovine viral diarrhoea virus from cattle to goats and among goats', *Veterinary Research*, 44(1), pp. 1–10. doi: 10.1186/1297-9716-44-32.

Baker, J. C. (1995) 'The Clinical Manifestations of Bovine Viral Diarrhea Infection', *Veterinary Clinics of North America: Food Animal Practice*. Elsevier Masson SAS, 11(3), pp. 425–445. doi: 10.1016/S0749-0720(15)30460-6.

Beaudeau, F. *et al.* (2001) 'Assessing the within-herd prevalence of cows antibody-positive to bovine viral diarrhoea virus with a blocking ELISA on bulk tank milk.', *The Veterinary record*, 149(8), pp. 236–40. doi: 10.1136/vr.149.8.236.

Beaudeau, F. *et al.* (2001) 'Evaluation of a blocking ELISA for the detection of bovine viral diarrhoea virus (BVDV) antibodies in serum and milk', *Veterinary Microbiology*, 80(4), pp. 329–337. doi: 10.1016/S0378-1135(01)00322-4.

Bhudevi, B. and Weinstock, D. (2001) 'Fluorogenic RT-PCR assay (TaqMan) for detection and classification of bovine viral diarrhea virus', *Veterinary Microbiology*, 83(1), pp. 1–10. doi: 10.1016/S0378-1135(01)00390-X.

Bhudevi, B. and Weinstock, D. (2003) 'Detection of bovine viral diarrhea virus in formalin fixed paraffin embedded tissue sections by real time RT-PCR (Taqman)', *Journal of Virological Methods*, 109(1), pp. 25–30. doi: 10.1016/S0166-0934(03)00040-5.

Brownlie, J. *et al.* (1987) 'Pathogenesis and Epidemiology of Bovine Virus Diarrhoea Virus Infection of Cattle.', *Annales de Recherches Vétérinaires*, 18, pp. 157–166. Available at: https://hal.inria.fr/file/index/docid/901705/filename/hal-00901705.pdf.

Brownlie, J. *et al.* (1998) 'Maternal recognition of foetal infection with bovine virus diarrhoea virus (BVDV) - The bovine pestivirus', *Clinical and Diagnostic Virology*, 10(2–3), pp. 141–150. doi: 10.1016/S0928-0197(98)00030-0.

Brownlie, J., Clarke, M. C. and Howard, C. J. (1984) 'Experimental production of fatal mucosal disease in cattle Brownlie et al 1984', *The Veterinary record*, 114(22), pp. 535–536.

Chamorro, M. F. *et al.* (2015) 'Efficacy of multivalent, modified- live virus (MLV) vaccines administered to early weaned beef calves subsequently challenged with virulent Bovine viral diarrhea virus type 2', *BMC Veterinary Research*, 11(1), pp. 1–9. doi: 10.1186/s12917-015-0342-8.

Cho, H. J. *et al.* (1991) 'Sensitivity and specificity of an enzyme-linked immunosorbent assay for the detection of bovine viral diarrhea virus antibody in cattle', *Can J Vet Res*, 55, pp. 56–59.

Drew, T. W., Yapp, F. and Paton, D. J. (1999) 'The detection of bovine viral diarrhoea virus in bulk milk samples by the use of a single-tube RT-PCR', *Veterinary Microbiology*, 64(2–3), pp. 145–154. doi: 10.1016/S0378-1135(98)00266-1.

Duncan, C. *et al.* (2008) 'Histopathologic and immunohistochemical findings in two white-tailed deer fawns persistently infected with Bovine viral diarrhea virus', *J Vet Diagn Invest*, 20(3), pp. 289–296. doi: 20/3/289 [pii].

Eiras, C. *et al.* (2012a) '*Bovine viral diarrhea virus*', *Journal of Veterinary Diagnostic Investigation*, 24(3), pp. 549–553. doi: 10.1177/1040638712440984.

Eiras, C. *et al.* (2012b) 'Bovine viral diarrhea virus: Correlation between herd seroprevalence and bulk tank milk antibody levels using 4 commercial immunoassays', *Journal of Veterinary Diagnostic Investigation*, 24(3), pp. 549–553. doi: 10.1177/1040638712440984.

Frey, H. R. *et al.* (2002) 'Foetal protection against bovine virus diarrhoea virus after two-step vaccination', *Journal of Veterinary Medicine, Series B*, 49(10), pp. 489–493. doi: 10.1046/j.1439-0450.2002.00599.x.

Fulton, R. W. *et al.* (2004) 'Maternally derived humoral immunity to bovine viral diarrhea virus (BVDV) 1a, BVDV1b, BVDV2, bovine herpesvirus-1, parainfluenza-3 virus bovine respiratory syncytial virus, Mannheimia haemolytica and Pasteurella multocida in beef calves, antibody decline ', *Vaccine*. Elsevier, 22(5–6), pp. 643–649. doi: 10.1016/j.vaccine.2003.08.033.

Fulton, R. W. (2013) 'Host response to bovine viral diarrhea virus and interactions with infectious agents in the feedlot and breeding herd', *Biologicals*. Elsevier Ltd, 41(1), pp. 31–38. doi: 10.1016/j.biologicals.2012.07.009.

Fux, R. and Wolf, G. (2012) 'Transient elimination of circulating bovine viral diarrhoea virus by colostral antibodies in persistently infected calves: A pitfall for BVDV-eradication programs?', *Veterinary Microbiology*. Elsevier B.V., 161(1–2), pp. 13–19. doi: 10.1016/j.vetmic.2012.07.001.

Givens, M. D. and Marley, M. S. (2013) 'Immunology of chronic BVDV infections', *Biologicals*. Elsevier Ltd, 41(1), pp. 26–30. doi: 10.1016/j.biologicals.2012.06.003.

Graham, D. A. *et al.* (2013) 'Herd-level factors associated with the presence of bovine viral diarrhoea virus in herds participating in the voluntary phase of the Irish national eradication programme', *Preventive Veterinary Medicine*. Elsevier B.V., 112(1–2), pp. 99–108. doi: 10.1016/j.prevetmed.2013.07.011.

Graham, D. A. *et al.* (2016) 'Quantifying the risk of spread of bovine viral diarrhoea virus (BVDV) between contiguous herds in Ireland', *Preventive Veterinary Medicine*. Elsevier B.V., 126, pp. 30–38. doi: 10.1016/j.prevetmed.2016.01.017.

Gunn, H. M. (1993) 'Role of fomites and flies in the transmission of bovine viral diarrhoea virus', *Veterinary Record*, 132(June 5), pp. 584–585.

Hanon, J. B. *et al.* (2014) 'Distinction between persistent and transient infection in a bovine viral diarrhoea (bvd) control programme: Appropriate interpretation of real-time RT-PCR and Antigen-ELISA test results', *Transboundary and Emerging Diseases*, 61(2), pp. 156–162. doi: 10.1111/tbed.12011.

Hertig, C. *et al.* (1991) 'Detection of bovine viral diarrhea virus, using degenerate oligonucleotide primers and the polymerase chain reaction', *Veterinary Microbiology*, 29, pp. 65–76. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1656820.

Hill, F. I., Reichel, M. P. and Tisdall, D. J. (2010) 'Use of molecular and milk production information for the cost-effective diagnosis of bovine viral diarrhoea infection in New Zealand dairy cattle', *Veterinary Microbiology*. Elsevier B.V., 142(1–2), pp. 87–89. doi: 10.1016/j.vetmic.2009.09.047.

Houe, H. (1993) 'Survivorship of animals persistently infected with bovine virus diarrhoea virus (BVDV)', *Preventive Veterinary Medicine*, 15(4), pp. 275–283. doi: 10.1016/0167-5877(93)90099-F.

Houe, H. *et al.* (1995) 'Application of Antibody-Titers against Bovine Viral Diarrhea Virus (Bvdv) as a

Measure to Detect Herds with Cattle Persistently Infected with Bvdv', *Journal of Veterinary Diagnostic InvestigationAlso part of Houe's thesis 1996*, 7(3), pp. 327–332.

Houe, H. (1995) 'Epidemiology of bovine viral diarrhea virus.', *The Veterinary clinics of North America. Food animal practice*. Elsevier Masson SAS, 11(3), pp. 521–47. doi: 10.1016/S0749-0720(15)30465-5.

Houe, H. (1999) 'Epidemiological features and economical importance of bovine virus diarrhoea virus (BVDV) infections', *Veterinary Microbiology*, 64(2–3), pp. 89–107. doi: 10.1016/S0378-1135(98)00262-4.

Hult, L. and Lindberg, A. (2005) 'Experiences from BVDV control in Sweden', *Preventive Veterinary Medicine*, 72(1–2), pp. 143–148. doi: 10.1016/j.prevetmed.2005.04.005.

Kim, S. G. and Dubovi, E. J. (2003) 'A novel simple one-step single-tube RT-duplex PCR method with an internal control for detection of bovine viral diarrhoea virus in bulk milk, blood, and follicular fluid samples', *Biologicals*, 31(2), pp. 103–106. doi: 10.1016/S1045-1056(03)00023-X.

Kliučinskas, R. *et al.* (2008) 'Detection of bovine viral diarrhoea virus in saliva samples', *Bulletin of the Veterinary Institute in Pulawy*, 52(1), pp. 31–37.

Kramps, J. A. *et al.* (1999) 'A simple, rapid and reliable enzyme-linked immunosorbent assay for the detection of bovine virus diarrhoea virus (BVDV) specific antibodies in cattle serum, plasma and bulk milk', *Veterinary Microbiology*, 64(2–3), pp. 135–144. doi: 10.1016/S0378-1135(98)00265-X.

Letellier, C. *et al.* (1999) 'Detection and genotyping of bovine diarrhea virus by reverse transcription-polymerase chain amplification of the 5' untranslated region', *Veterinary Microbiology*, 64(2–3), pp. 155–167. doi: 10.1016/S0378-1135(98)00267-3.

Lindberg, A. L. . and Alenius, S. (1999) 'Principles for eradication of bovine viral diarrhoea virus (BVDV) infections in cattle populations', *Veterinary Microbiology*, 64(2), pp. 197–222. doi: 10.1016/S0378-1135(98)00270-3.

Meyer, G. *et al.* (2012) 'Fetal protection against bovine viral diarrhoea type 1 virus infection after one administration of a live-attenuated vaccine', *Veterinary Journal*. Elsevier Ltd, 192(2), pp. 242–245. doi: 10.1016/j.tvjl.2011.05.011.

Meyling, A., Houe, H. and Jensen, a M. (1990) 'Epidemiology of bovine virus diarrhoea virus.', *Revue scientifique et technique (International Office of Epizootics)*, 9(1), pp. 75–93.

Meyling, A. and Mikél Jensen, A. (1988) 'Transmission of bovine virus diarrhoea virus (BVDV) by artificial insemination (AI) with semen from a persistently-infected bull', *Veterinary Microbiology*, 17(2), pp. 97–105. doi: 10.1016/0378-1135(88)90001-6.

Moen, A., Sol, J. and Sampimon, O. (2005) 'Indication of transmission of BVDV in the absence of persistently infected (PI) animals', *Preventive Veterinary Medicine*, 72(1–2), pp. 93–98. doi: 10.1016/j.prevetmed.2005.08.014.

Moennig, V. and Liess, B. (1995) 'Pathogenesis of intrauterine infections with bovine viral diarrhea virus.', *The Veterinary clinics of North America. Food animal practice*. Elsevier Masson SAS, 11(3), pp. 477–487. doi: 10.1016/S0749-0720(15)30462-X.

Moerman, A. *et al.* (1994) 'Clinical consequences of a bovine virus diarrhoea virus infection in a dairy herd: a longitudinal study.', *The Veterinary quarterly*, 16(2), pp. 115–119. doi: 10.1080/01652176.1994.9694430.

Muñoz-Zanzi, C. a *et al.* (2000) 'Pooled-sample testing as a herd-screening tool for detection of bovine viral diarrhea virus persistently infected cattle.', *Journal of veterinary diagnostic investigation : official publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, 12(3), pp. 195–203.

Nelson, D. D. *et al.* (2008) 'Evidence for Persistent *Bovine Viral Diarrhea Virus* Infection in a Captive Mountain Goat ( *Oreamnos Americanus* )', *Journal of Veterinary Diagnostic Investigation*, 20(6), pp. 752–759. doi: 10.1177/104063870802000606.

Nelson, D. D. *et al.* (2016) 'Persistent bovine viral diarrhea virus infection in domestic and wild small ruminants and camelids including the mountain goat (Oreamnos americanus)', *Frontiers in Microbiology*, 6(JAN), pp. 1–7. doi: 10.3389/fmicb.2015.01415.

Nettleton, P. F. (1990) 'Pestivirus infections in ruminants other than cattle.', *Revue scientifique et technique (International Office of Epizootics)*, 9(1), pp. 131–50. doi: 1990-002.

Nettleton, P. F. and Entrican, G. (1995) 'Ruminant Pestivirus', *British Veterinary Journal*, 151, pp. 615–642.

Nickell, J. S. *et al.* (2009) 'Onset and duration of transient infections among antibody- Diverse beef calves exposed to a bovine viral diarrhea virus persistently infected calf', *International Journal of Applied Research in Veterinary Medicine*, 9(1–2), pp. 29–39.

Niskanen, R. *et al.* (1996) 'Primarily BVDV infected calves as transmitters of the infection', *Proc. XIX World Buiatrics Congress, Edinburgh, Scotland, 8-12 july*, pp. 593–595.

Niskanen, R. and Lindberg, A. (2003) 'Transmission of bovine viral diarrhoea virus by unhygienic vaccination procedures, ambient air, and from contaminated pens', *Veterinary Journal*, 165(2), pp. 125–130. doi: 10.1016/S1090-0233(02)00161-2.

Niskanen, R., Lindberg, A. and Tråvén, M. (2002) 'Failure to spread bovine virus diarrhoea virus infection from primarily infected calves Despite Concurrent Infection with Bovine Coronavirus', *Veterinary Journal*, 163(3), pp. 251–259. doi: 10.1053/tvjl.2001.0657.

Patel, J. R. *et al.* (2002) 'Prevention of transplacental infection of bovine foetus by bovine viral diarrhoea virus through vaccination', *Archives of Virology*, 147(12), pp. 2453–2463. doi: 10.1007/s00705-002-0878-3.

Paton, D. J. *et al.* (1995) 'Identification of herd-specific bovine viral diarrhoea virus isolates from infected cattle and sheep', *Veterinary Microbiology; 43 (4) p283-294*, 43(94), pp. 283–294.

Pedrera, M. *et al.* (2012) 'Quantification and determination of spread mechanisms of bovine viral diarrhoea virus in blood and tissues from colostrum-deprived calves during an experimental acute infection induced by a non-cytopathic genotype 1 strain', *Transboundary and Emerging Diseases*, 59(5), pp. 377–384. doi: 10.1111/j.1865-1682.2011.01281.x.

Pillars, R. B. and Grooms, D. L. (2002) 'Serologic evaluation of five unvaccinated heifers to detect herds that have cattle persistently infected with bovine viral diarrhea virus', *American Journal of Veterinary Research*, 63(4), pp. 499–505. doi: 10.2460/ajvr.2002.63.499.

Renshaw, R. W., Ray, R. and Dubovi, E. J. (2000) 'Comparison of Virus Isolation and Reverse Transcription Polymerase Chain Reaction Assay for Detection of Bovine Viral Diarrhea Virus in Bulk Milk Tank Samples', *Journal of Veterinary Diagnostic Investigation*, 12(2), pp. 184–186. doi: 10.1177/104063870001200219.

Ridpath, J. F. and Bolin, S. R. (1995) 'Delayed Onset Postvaccinal Mucosal Disease as a Result of Genetic Recombination between Genotype 1 and Genotype 2 BVDV', *Virology*, pp. 259–262. doi: 10.1006/viro.1995.1480.

Rikula, U. *et al.* (2008) 'Transmission of bovine viral diarrhoea virus through the semen of acutely infected bulls under field conditions', *Veterinary Record*, 192, pp. 79–82.

Sandvik, T., Larsen, I. L. and Nyberg, O. (2001) 'Influence of milk from cows persistently infected with BVD virus on bulk milk antibody levels', *Vet Rec.*, 148(Fig 4), pp. 82–84. Available at: http://veterinaryrecord.bvapublications.com.

Sarrazin, S. *et al.* (2014) 'Virulence comparison and quantification of horizontal bovine viral diarrhoea virus transmission following experimental infection in calves', *Veterinary Journal*. Elsevier Ltd, 202(2), pp. 244–249. doi: 10.1016/j.tvjl.2014.07.010.

Scherer, C. F. C. *et al.* (2001) 'Experimental infection of pregnant ewes with bovine viral diarrhea virus type-2 (BVDV-2): Effects on the pregnancy and fetus', *Veterinary Microbiology*, 79(4), pp. 285–299. doi: 10.1016/S0378-1135(00)00357-6.

Smith, R. L. *et al.* (2008) 'Sensitivity of Polymerase Chain Reaction for Detection of Bovine Viral Diarrhea Virus in Pooled Serum Samples and Use of Pooled Polymerase Chain Reaction to Determine Prevalence of Bovine Viral Diarrhea Virus in Auction Market Cattle', *Journal of Veterinary Diagnostic Investigation*, 20(2008), pp. 75–78. doi: 10.1177/104063870802000115.

Synge, B. *et al.* (1999) 'The control of bovine virus diarrhoea virus in Shetland.', *Veterinary microbiology*, 64(2–3), pp. 223–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10028174.

Terpstra, C. and Wensvoort, G. (1997) 'A congenital persistent infection of bovine virus diarrhoea virus in PIGS: Clinical, virological and immunological observations', *Veterinary Quarterly*, 19(3), pp. 97–101. doi: 10.1080/01652176.1997.9694750.

Victora, C. G. *et al.* (1997) 'The role of conceptual frameworks in epidemiological analysis: a hierarchical approach', *International journal of epidemiology*, 26(1), pp. 224–227. doi: 10.1093/ije/26.1.224.

Voges, H. *et al.* (1998) 'Persistent bovine pestivirus infection localized in the testes of an immuno-competent, non-viraemic bull', *Veterinary Microbiology*, 61(3), pp. 165–175. doi: 10.1016/S0378-1135(98)00177-1.

Voges, H. *et al.* (2006) 'Direct adverse effects of persistent BVDv infection in dairy heifers – a retrospective case control study', *VetScript*, (September), pp. 4–6.

Walz, P. *et al.* (2010) 'Control of bovine viral diarrhea virus in ruminants', *Journal of Veterinary Internal Medicine*, 24(3), pp. 476–486. doi: 10.1017/CBO9781107415324.004.

Yan, L. *et al.* (2011) 'Combination of reverse transcription real-time polymerase chain reaction and antigen capture enzyme-linked immunosorbent assay for the detection of animals persistently infected with Bovine viral diarrhea virus', 25, pp. 16–25.

# 7 Annex

**ANNEX I:** Risk factors for introduction of BVDV at herd level

**Table I:** Risk factor for introduction of BVDV and their need for control. Reproduced from (Lindberg and Alenius, 1999).

| Risk | Perceived need for control | Plausible ways through which BVDV is introduced into a non-infected herd | Comments |
|---|---|---|---|
| **Livestock trade** | Yes, imperative | Purchase of :<br>1) A PI animal.<br>2) A dam carrying a PI calf.<br>3) A seronegative animal in early pregnancy, infected during trade.<br>4) Other animal which has attained transient infection during trade and transmits virus to newly pregnant non immune animals in the destination herd. | (a) Effect on disease spread by PIs in the market will be multiplied if contacts with seronegative animals in early pregnancy can occur.<br>(b) Prevalence of dams carrying PIs likely to be higher than prevalence of PI animals. The latter has been estimated to 1-2% in endemic situation (Houe, 1995).<br>(c) Transiently infected animals are regarded as low impact transmitters (Niskanen *et al.*, 1996). |
| **Exhibitions** | Yes | 1) Seronegative animal in early pregnancy becomes infected at the exhibition.<br>2) (An animal attains a transient infection and succeeds in infecting newly-pregnant non-immune animals after returning home.) | (a) PIs present at exhibitions will constitute a severe risk for farmers bringing seronegative animals in early pregnancy.<br>(b) Transiently infected animals are regarded as low impact transmitters. |
| **Animal contacts on pasture or over fences** | Yes | 1) Seronegative animals in early pregnancy become infected on pasture<br>2) (Some other animal attains a transient infection and subsequently transmits the infection to others, newly-pregnant non-immune animals in the herds.) | (a) Not controlling for release of PIs on common pastures will constitute a severe risk for farmers pasturing seronegative animals in early pregnancy.<br>(b) PI carrying dams may spread disease if they abort or calve on pasture.<br>(c) From a disease control point-of-view, and in terms of herd incidence, over-fence contacts will be less important than common pasturing. |
| **Live vaccines** | In the context of BVDV control, the use of live BVDV vaccines should be banned until proven safe. | At least one susceptible animal in early pregnancy becomes infected due to usage of live vaccine contaminated with non-cytopathogenic BVDV strains in the production process, or disease emerge as a result of recombination between vaccine and field strains (Ridpath and Bolin, 1995, Desport el al., 1997). | Risk of introducing strains new to the cattle population in question. |

| Risk | Perceived need for control | Plausible ways through which BVDV is introduced into a non-infected herd | Comments |
|---|---|---|---|
| **Semen and embryos** | Yes | At least one susceptible animal in early pregnancy becomes infected by other dams transiently infected due to AI with semen from PI bull or transiently infected bull, or persistent foetal infection develops in dam receiving AI with semen form PI bull or transiently infected bull. | Risk of introducing new strains to the cattle population in question.<br>A case has been reported with a seropositive bull constantly shedding virus in semen, in the absence of general persistent infection (Voges *et al.*, 1998). Although this phenomenon is probably of low frequency occurrence, it should be noted that such bulls could only be detected by testing semen. |
| **Visitors, including vets, AI technicians and herdsmen in the replacement system** | Unlikely to be of major importance and impact, but preventive measures are appropriates in scheme rules. | At least one susceptible animal in early pregnancy becomes infected due to contact with inadequately cleaned and/or disinfected clothes, boots, and instruments and similar. | Risk for transmission will depend upon :<br>- Interval time between visit in infected/non-infected herd (prevalence of infection in the area)<br>- Types of vehicles (faeces, clothes instruments (Gunn, 1993), contaminated injectable) and amount of virus transmitted (Houe, 1999)<br>- Pregnancy and immune statues of in-contact animal(s) in the herd |
| **On farm collection of slaughter animals or brokered calves by professional transportation staff** | Preventive measures are appropriate in scheme regulation. | At least one susceptible animal in early pregnancy becomes infected due to virus transfer by :<br>- Transportation staff<br>- Farmer entering transportation vehicle<br>Risk for airborne transmission of virus from transportation vehicles parked close to stable entrances or air intakes has not been investigated | Risk of successful transmission will depend upon :<br>- Number of infected animals in the vehicle, and type of infection (PI/transient)<br>- Time interval between visit in infected/non-infected herd<br>- Degree of handling at pick-up or delivery, i.e. degree of contact between transportation staff and cattle in the herd and/or between farmer and cattle in the vehicle<br>- Pregnancy and immune status of in-contact animals in the herd. |
| **Other species (sheep, goats, swine, deer, elks)** | Preventive measures for sheep are appropriate in scheme regulation. | At least one susceptible animal in early pregnancy becomes infected due to contact with a persistently infected sheep/goat/pig/deer/elk. | No evidence exists that wild ungulates, swine or goats has transmitted the infection to cattle, even though interspecies transmission is possible (Nettleton, 1990). Strains proven to be involved in transmission from sheep to cattle have been of bovine origin(Paton *et al.*, 1995). BVD control was not compromised by sheep when implemented on the Shetland Islands (Synge *et al.*, 1999). |
| **Vectors (ticks, mosquitos, flies)** | No, at least not in the temperate climate zones. | At least one susceptible animal in early pregnancy becomes infected due to contact with virus-carrying vector. | Insects, such as biting flies have been shown to be capable of carry BVDV under experimental conditions (Tarry et al., 1991). Vector-borne transmission has never been described under natural |

*Table II: Types of contacts that can act as routes for transmission of BVDV infection between herds (from Lindberg and Houe 2005)*

| Characteristic of contact type | | | $\beta_A$ | $\beta_B$ | $\kappa$ |
|---|---|---|---|---|---|
| Source of infectivity | Length of contact[a] | Type of recipient[b] | Probability of transmission $A$[c] | Probability of transmission $B\|A$[c] | Number of potential infectious contacts per time unit[d] |
| PI animals | Permanent | | | 1 | Low |
| | Transient | Not early pregnancy | Close to 1 | Negligible | Low–very high |
| | | Early pregnancy | Close to 1 | Close to 1 | |
| Dams pregnant with PI foetuses | Permanent | | | Close to 1 | Low |
| Transiently infected animals | Permanent | Not early pregnancy | Very low | Negligible | Low–moderate |
| | | Early pregnancy | Very low | Close to 1 | |
| | Transient | Not early pregnancy | Negligible | Negligible | Low–very high |
| | | Early pregnancy | Negligible | Close to 1 | |
| Contaminated biologicals | n.a. | Not early pregnancy | Very low–very high, dependent on dose and way of administration! | Negligible | Low–extreme |
| | | Early pregnancy | | Close to 1 | |
| Contaminated equipment, persons etc | n.a. | Not early pregnancy | | Negligible | Low–moderate |
| | | Early pregnancy | | Close to 1 | |

[a] Permanent: more long-term introduction of the infectious animal into the herd; Transient: the infectious animal is in contact with other herds for a limited period of time, e.g. during an exhibition, an auction or a common pasture.

[b] Denotes whether the infectious animal is in contact with a susceptible animal in early pregnancy in the recipient herd for the duration of its infectious period, or not.

[c] $A$, the probability that the contact results in a transient infection in *the recipient(s)*; $B$, the probability that the contact results in a persistent infection in *the herd*. $B$ is conditional on $A$ for all scenarios except for permanent contacts with PI animals or PI carriers.

[d] Refers to the (potential) frequency of contacts between infected and a non-infected herds, by the contact type in question. Permanent contacts translates into intensity of trade, and transient contacts into intensity of contacts over fences, on common pastures, at auctions, exhibitions, etc.

# Annexe 2 : STOC free : WP1 and WP2, Deliverable 2. Guidelines for the identification of sources of the data available to quantify the confidence of freedom from infection, with an application to BVDV, 07/2018

M. Mercat, A. van Roon, I. Santman, G. van Schaik, M. Nielen, L. van Duijn, S. More, D. Graham, J. Frössling, A. Lindberg, J. Gethmann, C. Sauter-Louis, G. Gunn, C. Gomes, M. Henry, C. Fourichon, A. Madouasse.

**STOC free: WP1 and WP2, Deliverable 2**

# Guidelines for the identification of sources of the data available to quantify the confidence of freedom from infection, with an application to BVDV

**M. Mercat[1], A. van Roon[2], I. Santman[2,3], G. van Schaik[2,3], M. Nielen[2], L. van Duijn[3], S. More[4], D. Graham[5], J. Frössling[6], A. Lindberg[6], J. Gethmann[7], C. Sauter-Louis[7], G. Gunn[8], C. Gomes[8], M. Henry[8], C. Fourichon[1], A. Madouasse[1]**

[1] BIOEPAR, INRA, Oniris, Université Bretagne Loire, 44307, Nantes, France, [2]Urecht University, the Netherlands, [3]GD Animal Health, the Netherlands, [4]University College Dublin, Ireland, [5]Animal Health Ireland, [6]Swedish National Veterinary Institute, Sweden, [7]FLI, Germany, [8]SRUC, Scotland

July 2018

GA/EFSA/AFSCO/2016/01-03



**This study was awarded a grant by EFSA and was co-financed by public organisations in the countries participating in the study.**

# TABLE OF CONTENTS

# INTRODUCTION

The second deliverable of WP1 (D1.2) "Guidelines for the identification and sources of data" is developed in close collaboration with WP2. This deliverable is linked to the conceptual model representing the course and dynamics of infection at different levels (D1.1) and the questionnaire which captures data about aspects of control programmes that influence the confidence of freedom (D2.1 and D2.2). In this deliverable, the data needed for calculation of the confidence of freedom within the STOC free framework are listed. The data of interest include information issued from monitoring the infection (e.g. programme output and test characteristics) and information on factors that could influence the probability for a given entity to be free from infection (e.g. contact structure, infection pressure and presence of risk factors for introduction or delayed detection).

This deliverable consists of:

1. a table for collecting all data that is possibly important for calculation of the confidence of freedom within the STOC free framework 2.
    a. Template
    b. Filled in for the Netherlands
    c. Filled in for France
2. a table that describes all data that is possibly important for calculation of the confidence of freedom within the STOC free framework,
3. a table that gives an overview of BVDV diagnostic tests that are used in Europe with associated test characteristics and
4. a table that lists risk factors for introduction and delayed detection of BVDV ordered on importance by the six countries within the consortium.

The aim of the first table that lists important data for calculation of confidence of freedom is not to collect the data itself but to indicate on the territory level whether quantitative or qualitative data are available for each variable of interest, the sources of the data and the strengths and limitations of the data.

The overview of diagnostic tests for BVDV and their characteristics is a first inventory. Currently, the consortium is working on a systematic review about risk factors for introduction and/or delayed detection which also includes papers about test performance. If the systematic review does not provide sufficient data on the test characteristics, the consortium will explore other options to complete the overview presented in this deliverable.

The third table presents the risk factors for introduction or delayed detection that were deemed most relevant by the partners collaborating in the STOC free consortium. The table

distinguishes risk factors on animal, herd and territory level and was stratified towards a disease free or endemic situation. In the systematic review that is currently conducted, risk factors with their risk estimates will be identified, which will be used to complete the current risk factor overview.

This deliverable gives a comprehensive overview of all relevant and potentially available data for calculating the confidence of freedom. The tables will guide the further development of both the statistical model (STOC free MODEL) and the data collection tool (STOC free DATA).

## GUIDELINES FOR FILLING IN THE TABLE FOR IDENTIFICATION AND SOURCES OF DATA

In table 1, all relevant variables are listed followed by a definition of the information requested and the type of data. The column "importance of data", indicates the expected importance of the data at this point in the project. This expected importance may be updated during development of the statistical model. In the next column, it is asked whether exact quantitative data are available for each individual variable. This information is requested for all cattle. Then this information is further tailored to, dairy and non-dairy and subsequently a relevant subset of non-dairy: beef breeding. If the country has no exact quantitative data (e.g. the distribution of the parameters) available for the variable, it should be indicated whether they can provide a qualitative estimation (e.g. estimation by expert opinion). Thereafter, the owner of the data and the organisation with access to the data should be specified. Then there is a column about the strengths and limitations of the data. Here countries can indicate the quality of their data and what the limitations are, for example national coverage of the data as a strength and the lack of recent data as a limitation. In the comments column, all additional information on the data can be provided.

The table has been tested for clarity and user-friendliness by three countries within the consortium (i.e. NL, FR, SE), this helped to further develop the table into the current final version. This Table together with the table about the test characteristics and the information and estimations for the risk factors will be used to guide the further development of STOC free model. The information that is relevant input for STOC free model will be included in STOC free DATA to gather the necessary quantitative data.

## DATA IDENTIFICATION TABLE

For each of the parameters that were defined in the data information table the following information is requested to evaluate the potential for inclusion as input in STOC free model:

- Is there quantitative information available (No/Yes), for all cattle and stratified to dairy, non-dairy and beef breeding
- If no quantitative data is available, can a qualitative estimation be provided (No/Yes), for all cattle and stratified to dairy, non-dairy and beef breeding
- The owner of the data
- The organisation with access to the data
- Strengths and limitations of the data
- Comments about the data

The parameters that were included are provided in the table.

# I. TEMPLATE OF DATA IDENTIFICATION TABLE

## 1. Demographics

| | Variable | Definition | Type of data | Importance of data |
|---|---|---|---|---|
| Demographics (For the most recent full calendar year) | Number of cattle | Only cattle older than 1 year | Number of individual animals | ++ |
| | Number of cattle herds | Total number of cattle herds | Number of herds | ++ |
| | Average number of cattle per herd | Only cattle older than 1 year | Distribution [mean, median, SD, 5 and 95 percentiles] | +++ |
| | Number of births in the territory | Within the past 12 months in the territory | Number of individual births | +++ |
| | Average number of births per herd | Within the past 12 months per herd | Distribution [mean, median, SD, 5 and 95 percentiles] | +++ |
| | Calving pattern | Precentage of all calvings by month within the past 12 months | Distribution [mean, median, SD, 5 and 95 percentiles] | ++ |
| | Cattle density | The number of cattle per km2 | Distribution [mean, median, SD, 5 and 95 percentiles] | +++ |
| | Percentage of dairy cattle herds that have also beef cattle on the same location | All dairy herds that also have a type of beef cattle such as veal calf, suckler cattle etc. | Percentage of herds | ++ |
| | Number of farmed goat and/or sheep herds | Commercial goat and sheep herds | Number of herds | + |
| | Percentage of cattle herds that also have goat and/or sheep on the same location | Cattle herds with goat and sheep on the same location | Percentage of herds | + |
| | Percentage of cattle herds that could possibly have contact with wild ruminants | | Percentage of herds | + |

## 2. Control programme

| | | Variable | Definition | Type of data | Importance of data |
|---|---|---|---|---|---|
| Control programme | Previous year | Percentage of eligible cattle herds that participate in the control programme | Percentage of eligible herds that participate in the control programme at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of animals tested | Percentage of cattle tested for BVD in the territory , during the year | Percentage of individual animals | +++ |
| | | Number of herds that identified one or more PI's. | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of herds | +++ |
| | | Number of PI's identified in the territory | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of individual animals | +++ |
| | | Age at which PI animals were culled | Age at which PI animals were culled during the year | Distribution [mean, median, SD, 5 and 95 percentiles] of age at which PI animals were culled | +++ |
| | | Percentage of free cattle herds | Percentage of cattle herds participating in the CP that have any free status according to the control programme, at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of free cattle herds that had a breakdown | Percentage of herds participating in the CP that have a free status at the beginning of the year and that during that year had a breakdown. Breakdown: an antibody or virus positive test while the herd was free before, during the year | Percentage of herds | +++ |
| | -1 * | Percentage of eligible cattle herds that participate in the control programme | Percentage of eligible herds that participate in the control programme at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of animals tested | Percentage of cattle tested for BVD in the territory , during the year | Percentage of individual animals | +++ |
| | | Number of herds that identified one or more PI's. | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of herds | +++ |
| | | Number of PI's identified in the territory | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of individual animals | +++ |
| | | Age at which PI animals were culled | Age at which PI animals were culled during the year | Distribution [mean, median, SD, 5 and 95 percentiles] of age at which PI animals were culled | +++ |
| | | Percentage of free cattle herds | Percentage of cattle herds participating in the CP that have any free status according to the control programme, at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of free cattle herds that had a breakdown | Percentage of herds participating in the CP that have a free status at the beginning of the year and that during that year had a breakdown. Breakdown: an antibody or virus positive test while the herd was free before, during the year | Percentage of herds | +++ |

| | | Variable | Definition | Type of data | Importance of data |
|---|---|---|---|---|---|
| Control programme | -2 ˚ | Percentage of eligible cattle herds that participate in the control programme | Percentage of eligible herds that participate in the control programme at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of animals tested | Percentage of cattle tested for BVD in the territory , during the year | Percentage of individual animals | +++ |
| | | Number of herds that identified one or more PI's. | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of herds | +++ |
| | | Number of PI's identified in the territory | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of individual animals | +++ |
| | | Age at which PI animals were culled | Age at which PI animals were culled during the year | Distribution [mean, median, SD, 5 and 95 percentiles] of age at which PI animals were culled | +++ |
| | | Percentage of free cattle herds | Percentage of cattle herds participating in the CP that have any free status according to the control programme, at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of free cattle herds that had a breakdown | Percentage of herds participating in the CP that have a free status at the beginning of the year and that during that year had a breakdown. Breakdown: an antibody or virus positive test while the herd was free before, during the year | Percentage of herds | +++ |
| | -3 ˚ | Percentage of eligible cattle herds that participate in the control programme | Percentage of eligible herds that participate in the control programme at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of animals tested | Percentage of cattle tested for BVD in the territory , during the year | Percentage of individual animals | +++ |
| | | Number of herds that identified one or more PI's. | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of herds | +++ |
| | | Number of PI's identified in the territory | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of individual animals | +++ |
| | | Age at which PI animals were culled | Age at which PI animals were culled during the year | Distribution [mean, median, SD, 5 and 95 percentiles] of age at which PI animals were culled | +++ |
| | | Percentage of free cattle herds | Percentage of cattle herds participating in the CP that have any free status according to the control programme, at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of free cattle herds that had a breakdown | Percentage of herds participating in the CP that have a free status at the beginning of the year and that during that year had a breakdown. Breakdown: an antibody or virus positive test while the herd was free before, during the year | Percentage of herds | +++ |
| | -4 ˚ | Percentage of eligible cattle herds that participate in the control programme | Percentage of eligible herds that participate in the control programme at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of animals tested | Percentage of cattle tested for BVD in the territory , during the year | Percentage of individual animals | +++ |
| | | Number of herds that identified one or more PI's. | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of herds | +++ |
| | | Number of PI's identified in the territory | A PI is an animal that was positive to BVDV at the initial test and did not have a negative re-test, during the year | Number of individual animals | +++ |
| | | Age at which PI animals were culled | Age at which PI animals were culled during the year | Distribution [mean, median, SD, 5 and 95 percentiles] of age at which PI animals were culled | +++ |
| | | Percentage of free cattle herds | Percentage of cattle herds participating in the CP that have any free status according to the control programme, at the beginning of the year | Percentage of herds | +++ |
| | | Percentage of free cattle herds that had a breakdown | Percentage of herds participating in the CP that have a free status at the beginning of the year and that during that year had a breakdown. Breakdown: an antibody or virus positive test while the herd was free before, during the year | Percentage of herds | +++ |

## 3. Management

| | | Variable | Definition | Type of data | Importance of data |
|---|---|---|---|---|---|
| Management (For the most recent full calendar year) | Purchase | Percentage of herds that purchased cattle | Percentage of cattle herds that purchased one or more cattle, within or from outside the territory | Percentage of herds | +++ |
| | | Percentage of herds that purchased all animals within the territory | | Percentage of herds | ++ |
| | | Percentage of herds that purchased at least one animal from markets/traders | | Percentage of herds | ++ |
| | | Number of cattle that was purchased | Total number of cattle (all age categories) that was purchased | Number of individual animals | +++ |
| | | Percentage of cattle that was purchased within the territory | | Percentage of individual animals | ++ |
| | | Percentage of cattle that was purchased from markets/traders | | Percentage of individual animals | ++ |
| | | Number of purchase moments in the territory | Purchase moment : a purchase event on a specific day to one specific herd from another herd | Distribution [mean, median, SD, 5 and 95 percentiles] of times purchased cattle is introduced in a herd | +++ |
| | | Average number of cattle purchased at each purchase moment | | | +++ |
| | | Territories where most cattle was purchased from | Percentage of cattle per territory from the five territories where most cattle were purchased from | Percentage of cattle per territory | +++ |
| | | Percentage of purchased animals that were a calf at the moment of purchase | Calf: an animal in its first year | Percentage of individual animals | +++ |
| | | Percentage of purchased animals that were pregnant at the moment of purchase | | Percentage of individual animals | +++ |
| | | Percentage of herds that use quarantine for their purchased animals that have not been tested before arrival in the herd | | Percentage of herds | + |

| | | Variable | Definition | Type of data | Importance of data |
|---|---|---|---|---|---|
| Management (For the most recent full calendar year) | Grazing | Percentage of cattle herds practicing zero grazing | Zero grazing: no grazing during the whole year | Percentage of herds | +++ |
| | | Percentage of cattle herds involved in communal grazing | Communal grazing: grazing animals from different cattle herds together | Percentage of herds | + |
| | | Percentage of cattle farms that are fragmented | Fragmented farm: a farm where two or more geographically separated tracts of lands are operated | Percentage of herds | + |
| | | Number of neighbours at pasture per herd | Neighbours at pasture: pasture where cattle from different herds can have nose to nose contact | Distribution [mean, median, SD, 5 and 95 percentiles] number of neighbours | ++ |
| | | Percentage of herds where calves possibly have nose to nose contact with pregnant cattle on pasture | A calf is cattle up to 1 year old. | Percentage of herds | ++ |
| | Breeding | Percentage of herds that apply natural breeding | Percentage of herds that breed. All herds that used at least once natural breeding during the previous year | Percentage of herds | + |
| | | Percentage of herds that use artificial insemination | Percentage of herds that breed. All herds that used at least once artificial insemination during the previous year | Percentage of herds | + |
| | Cattle shows | Percentage of herds that have animals attending shows | | Percentage of herds | + |
| | Vaccination | Percentage of herds that vaccinate cattle against BVD | | Percentage of herds | ++ |
| | Housing | Percentage of herds that house calves separately from pregnant cattle | Percentage of herds that breed and that house calves separately from pregnant cattle. | Percentage of herds | + |
| | | Percentage of herds where calves possibly have nose to nose contact with pregnant cattle in the barn | A calf is cattle up to 1 year old. | Percentage of herds | + |
| | Biosecurity | Percentage of herds that share transport vehicles with other cattle herds | | Percentage of herds | + |
| | | Percentage of herds that share equipment with other cattle herds | | Percentage of herds | + |
| | | Percentage of herds that provide clothing for visitors | | Percentage of herds | + |

## II. DATA IDENTIFICATION TABLE FILLED IN FOR THE NETHERLANDS

**Territory : Netherlands**
*A territory is defined as a geographical area in which herds participate in the same control programme. The information provided below should be specific for this territory*
**Date of filling in : 25/06/2018**
**Period for which the data is available: 5 years (2017 and before). In 2018 a new programme started.**
*Preferably the most recent full calendar year up to five years back*
**Please specify how you would define non-dairy and beef-breeding: Non-dairy: beef breeding + beef non breeding (Farms keeping bulls for bull meat production and veal) Beef breeding: suckler**
*These categories are included in the table below*

### 1. Demographics

| | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Demographics (For the most recent full calendar year) | Number of cattle | ++ | Yes | Yes | Yes | Yes | | | | | RVO and GD | GD | Strength: Census data | Valid for all demographics-variables. |
| | Number of cattle herds | ++ | Yes | Yes | Yes | Yes | | | | | RVO and GD | GD | Limitation: Data not available from herds that refuse to participate in the | Valid for all demographics-variables. |
| | Average number of cattle per herd | +++ | Yes | Yes | Yes | Yes | | | | | RVO and GD | GD | | |
| | Number of births in the territory | +++ | Yes | Yes | Yes | Yes | | | | | RVO and GD | GD | Limitation: No access to data of stillborn twin calves, so numbers are very | |
| | Average number of births per herd | +++ | Yes | Yes | Yes | Yes | | | | | RVO and GD | GD | Limitation: No access to data of stillborn twin calves, so numbers are very | |
| | Calving pattern | ++ | Yes | Yes | Yes | Yes | | | | | RVO and GD | GD | | |
| | Cattle density | +++ | Yes | Yes | Yes | Yes | | | | | RVO and GD | GD | | |
| | Percentage of dairy cattle herds that have also beef cattle on the same location | ++ | | | | | | Yes | | | | | | Estimation based on former research |
| | Number of farmed goat and/or sheep herds | + | Yes | | | | | | | | RVO | All | | |
| | Percentage of cattle herds that also have goat and/or sheep on the same location | + | | | | | Yes | No | No | No | | | | Estimation based on former research |
| | Percentage of cattle herds that could possibly have contact with wild ruminants | + | | | | | Yes | No | No | No | | | | |

## 2. Control programme

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Control programme | Previous year | Percentage of eligible cattle herds that participate in the control programme | +++ | Yes | Yes | Yes | Yes | | | | | GD/ZuivelNL | GD | | Approval for access of the data should be requested for use |
| | | Percentage of animals tested | +++ | Yes | Yes | Yes | Yes | | | | | GD/ZuivelNL | GD | | Approval for access of the data should be requested for use |
| | | Number of herds that identified one or more PI's. | +++ | Yes | Yes | Yes | Yes | | | | | GD/ZuivelNL | GD | | Approval for access of the data should be requested for use |
| | | Number of PI's identified in the territory | +++ | Yes | Yes | Yes | Yes | | | | | GD/ZuivelNL | GD | | Approval for access of the data should be requested for use |
| | | Age at which PI animals were culled | +++ | Yes | Yes | Yes | Yes | | | | | GD/ZuivelNL | GD | | Approval for access of the data should be requested for use |
| | | Percentage of free cattle herds | +++ | Yes | Yes | Yes | Yes | | | | | GD/ZuivelNL | GD | | Approval for access of the data should be requested for use |
| | | Percentage of free cattle herds that had a breakdown | +++ | Yes | Yes | Yes | Yes | | | | | GD/ZuivelNL | GD | | Approval for access of the data should be requested for use |
| | -1 * | Percentage of eligible cattle herds that participate in the control programme | +++ | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of animals tested | +++ | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of herds that identified one or more PI's. | +++ | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of PI's identified in the territory | +++ | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Age at which PI animals were culled | +++ | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds | +++ | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds that had a breakdown | +++ | Yes | Yes | Yes | Yes | | | | | GD | GD | | |

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Control programme | -2 - | Percentage of eligible cattle herds that participate in the control programme | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of animals tested | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of herds that identified one or more PI's. | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of PI's identified in the territory | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Age at which PI animals were culled | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds that had a breakdown | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | -3 - | Percentage of eligible cattle herds that participate in the control programme | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of animals tested | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of herds that identified one or more PI's. | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of PI's identified in the territory | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Age at which PI animals were culled | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds that had a breakdown | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | -4 - | Percentage of eligible cattle herds that participate in the control programme | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of animals tested | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of herds that identified one or more PI's. | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Number of PI's identified in the territory | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Age at which PI animals were culled | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |
| | | Percentage of free cattle herds that had a breakdown | *** | Yes | Yes | Yes | Yes | | | | | GD | GD | | |

## 3. Management

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Management (For the most recent full calendar year) | Purchase | Percentage of herds that purchased cattle | +++ | Yes | Yes | Yes | Yes | | | | | RVO | GD | | |
| | | Percentage of herds that purchased all animals within the territory | ++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Percentage of herds that purchased at least one animal from markets/traders | ++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Number of cattle that was purchased | +++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Percentage of cattle that was purchased within the territory | ++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Percentage of cattle that was purchased from markets/traders | ++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Number of purchase moments in the territory | +++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Average number of cattle purchased at each purchase moment | +++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Territories where most cattle was purchased from | +++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Percentage of purchased animals that were a calf at the moment of purchase | +++ | Yes | Yes | Yes | Yes | | | | | | | | |
| | | Percentage of purchased animals that were pregnant at the moment of purchase | +++ | Yes | Yes | Yes | Yes | | | | | | | | Indirect by evaluating whether they gave birth to a calf within 9 months after purchase. |
| | | Percentage of herds that use quarantine for their purchased animals that have not been tested before arrival in the herd | + | | | | | No | No | No | No | | | | |

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Management (For the most recent full calendar year) | Grazing | Percentage of cattle herds practicing zero grazing | +++ | | | | | Yes | Yes | Yes | Yes | Dairy cooperations | Available to public | | Information based on annual sustainability report |
| | | Percentage of cattle herds involved in communal grazing | + | | | | | No | Yes | No | No | RVO | GD | | Is very rare |
| | | Percentage of cattle farms that are fragmented | + | | | | | No | No | No | No | | | | No information |
| | | Number of neighbours at pasture per herd | ++ | Yes | Yes | Yes | Yes | | | | | | | | Indirect through number of herds within a radius of 500 meters |
| | | Percentage of herds where calves possibly have nose to nose contact with pregnant cattle on pasture | ++ | | | | | No | No | No | No | | | | No information |
| | Breeding | Percentage of herds that apply natural breeding | + | | Yes | | | Yes | | Yes | Yes | CRV | CRV | Limitation: Only data available from 75% of dairy farms | Approval should be requested. |
| | | Percentage of herds that use artificial insemination | + | | Yes | | | Yes | | Yes | Yes | CRV | CRV | Limitation: Only data available from 75% of dairy farms | Approval should be requested. |
| | Cattle shows | Percentage of herds that have animals attending shows | + | | | | | Yes | No | No | No | RVO | GD | | Only certified shows |
| | Vaccination | Percentage of herds that vaccinate cattle against BVD | ++ | | | | | No | Yes | No | No | GD | GD | | |
| | Housing | Percentage of herds that house calves separately from pregnant cattle | + | | | | | No | No | No | No | | | | |
| | | Percentage of herds where calves possibly have nose to nose contact with pregnant cattle in the barn | + | | | | | No | No | No | No | | | | |
| | Biosecurity | Percentage of herds that share transport vehicles with other cattle herds | + | | | | | No | No | No | No | | | | |
| | | Percentage of herds that share equipment with other cattle herds | + | | | | | No | No | No | No | | | | |
| | | Percentage of herds that provide clothing for visitors | + | | | | | No | No | No | No | | | | |

# I. DATA IDENTIFICATION TABLE FILLED IN FOR BRITANNY (FRANCE)

**Territory : Britanny (France)**
*A territory is defined as a geographical area in which herds participate in the same control programme. The information provided below should be specific for this territory*
**Date of filling in : 7/07/2018**
**Period for which the data is available: 2017 and before**
*Preferably the most recent full calendar year up to five years back*
**Please specify how you would define non-dairy and beef-breeding: Non-dairy: beef breeding + beef non breeding (Farms keeping bulls for bull meat production and veal) Beef breeding: suckler**
*These categories are included in the table below*

## 1. Demographics

| | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Demographics (For the most recent full calendar year) | Number of cattle | ++ | yes | yes | yes | yes | | | | | State | GDS | | Definition of beef/dairy herds depends on cow breeds present on the farm |
| | Number of cattle herds | ++ | yes | yes | yes | yes | | | | | State | | | |
| | Average number of cattle per herd | +++ | yes | yes | yes | yes | | | | | State | | | |
| | Number of births in the territory | +++ | yes | yes | yes | yes | | | | | State | | | |
| | Average number of births per herd | +++ | yes | yes | yes | yes | | | | | State | | | |
| | Calving pattern | ++ | yes | yes | yes | yes | | | | | State | | | |
| | Cattle density | +++ | yes | yes | yes | yes | | | | | State | | | |
| | Percentage of dairy cattle herds that have also beef cattle on the same location | ++ | | yes | | | | | | | State | | | |
| | Number of farmed goat and/or sheep herds | + | no | | | | no | | | | State | | | |
| | Percentage of cattle herds that also have goat and/or sheep on the same location | + | no | no | no | no | no | no | no | no | State | | | |
| | Percentage of cattle herds that could possibly have contact with wild ruminants | + | no | no | no | no | no | no | no | no | State | | | |

## 2. Control programme

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Control programme | Previous year | Percentage of eligible cattle herds that participate in the control programme | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Percentage of animals tested | *** | no | yes | no | no | | | | | GDS | | | |
| | | Number of herds that identified one or more PI's. | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Number of PI's identified in the territory | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Age at which PI animals were culled | *** | | | | | | | | | GDS | | | |
| | | Percentage of free cattle herds | *** | no | yes | no | no | yes | | yes | yes | GDS | | | |
| | | Percentage of free cattle herds that had a breakdown | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | -1 * | Percentage of eligible cattle herds that participate in the control programme | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Percentage of animals tested | *** | no | yes | no | no | | | | | GDS | | | |
| | | Number of herds that identified one or more PI's. | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Number of PI's identified in the territory | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Age at which PI animals were culled | *** | | | | | | | | | GDS | | | |
| | | Percentage of free cattle herds | *** | no | yes | no | no | yes | | yes | yes | GDS | | | |
| | | Percentage of free cattle herds that had a breakdown | *** | yes | yes | yes | yes | | | | | GDS | | | |

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Control programme | -2 - | Percentage of eligible cattle herds that participate in the control programme | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Percentage of animals tested | *** | no | yes | no | no | | | | | GDS | | | |
| | | Number of herds that identified one or more PI's. | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Number of PI's identified in the territory | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Age at which PI animals were culled | *** | | | | | | | | | GDS | | | |
| | | Percentage of free cattle herds | *** | no | yes | no | no | yes | | yes | yes | GDS | | | |
| | | Percentage of free cattle herds that had a breakdown | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | -3 - | Percentage of eligible cattle herds that participate in the control programme | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Percentage of animals tested | *** | no | yes | no | no | | | | | GDS | | | |
| | | Number of herds that identified one or more PI's. | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Number of PI's identified in the territory | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Age at which PI animals were culled | *** | | | | | | | | | GDS | | | |
| | | Percentage of free cattle herds | *** | no | yes | no | no | yes | | yes | yes | GDS | | | |
| | | Percentage of free cattle herds that had a breakdown | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | -4 - | Percentage of eligible cattle herds that participate in the control programme | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Percentage of animals tested | *** | no | yes | no | no | | | | | GDS | | | |
| | | Number of herds that identified one or more PI's. | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Number of PI's identified in the territory | *** | yes | yes | yes | yes | | | | | GDS | | | |
| | | Age at which PI animals were culled | *** | | | | | | | | | GDS | | | |
| | | Percentage of free cattle herds | *** | no | yes | no | no | yes | | yes | yes | GDS | | | |
| | | Percentage of free cattle herds that had a breakdown | *** | yes | yes | yes | yes | | | | | GDS | | | |

## 3. Management

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Management (For the most recent full calendar year) | Purchase | Percentage of herds that purchased cattle | *** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Percentage of herds that purchased all animals within the territory | ** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Percentage of herds that purchased at least one animal from markets/traders | ** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Number of cattle that was purchased | *** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Percentage of cattle that was purchased within the territory | ** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Percentage of cattle that was purchased from markets/traders | ** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Number of purchase moments in the territory | *** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Average number of cattle purchased at each purchase moment | *** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Territories where most cattle was purchased from | *** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Percentage of purchased animals that were a calf at the moment of purchase | *** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Percentage of purchased animals that were pregnant at the moment of purchase | *** | yes | yes | yes | yes | | | | | State | GDS | | |
| | | Percentage of herds that use quarantine for their purchased animals that have not been tested before arrival in the herd | * | yes | yes | yes | yes | | | | | State | GDS | | |

| | | Variable | Importance of data | Quantitative (Yes/No) | | | | If no quantitative data are available : Qualitative (Yes/No) | | | | Owner of the data | Organisation with access to the data | Strengths and limitations of the data | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | All cattle (dairy + non-dairy) | Dairy | Non Dairy | Beef breeding | | | | |
| Management (For the most recent full calendar year) | Grazing | Percentage of cattle herds practicing zero grazing | +++ | no | no | no | no | no | no | no | no | | | | |
| | | Percentage of cattle herds involved in communal grazing | + | no | no | no | no | yes (0) | yes (0) | yes (0) | yes (0) | | | | |
| | | Percentage of cattle farms that are fragmented | + | no | no | no | no | no | no | no | no | | | | |
| | | Number of neighbours at pasture per herd | ++ | no | no | no | no | no | no | no | no | | | | |
| | | Percentage of herds where calves possibly have nose to nose contact with pregnant cattle on pasture | ++ | no | no | no | no | no | no | no | no | | | | |
| | Breeding | Percentage of herds that apply natural breeding | + | | | | | yes | yes | yes | yes | | | | |
| | | Percentage of herds that use artificial insemination | + | | | | | yes | yes | yes | yes | France génétique élevage | | | There exists published summaries |
| | Cattle shows | Percentage of herds that have animals attending shows | + | no | no | no | no | no | no | no | no | | | | |
| | Vaccination | Percentage of herds that vaccinate cattle against BVD | ++ | no | no | no | no | yes | yes | yes | yes | GDS | | | |
| | Housing | Percentage of herds that house calves separately from pregnant cattle | + | no | no | no | no | no | no | no | no | | | | |
| | | Percentage of herds where calves possibly have nose to nose contact with pregnant cattle in the barn | + | no | no | no | no | no | no | no | no | | | | |
| | Biosecurity | Percentage of herds that share transport vehicles with other cattle herds | + | no | no | no | no | no | no | no | no | | | | |
| | | Percentage of herds that share equipment with other cattle herds | + | no | no | no | no | no | no | no | no | | | | |
| | | Percentage of herds that provide clothing for visitors | + | no | no | no | no | no | no | no | no | | | | |

## OVERVIEW OF TEST CHARACTERISTICS FOR BVDV DIAGNOSTIC TESTING

| Antibody ELISA' s | Producer | Serum sample | | Technical (lab)/Diagnostic (field) | Gold standard | Reference | Milk sample | | Bulk milk/individual samples | Technical (lab)/Diagnostic (field) | Gold standard | Reference | Tissue sample | | Technical (lab)/Diagnostic (field) | Gold standard | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Se | Sp | | | | Se | Sp | | | | | Se | Sp | | | |
| BVDV (antibody) | BIO-X DIAGNOSTICS | | | | | | | | | | | | | | | | |
| BVDV (antibody) competition | BIO-X DIAGNOSTICS | | | | | | | | | | | | | | | | |
| HerdChek BVDV Antibody (Bovine Viral Diarrhea Virus) | IDEXX | 96,30% | 99,50% | Technical (lab) | | Hashemi Tabar et al., 2010 | | | | | | | | | | | |
| POURQUIER® ELISA BVD Ab | IDEXX | | | | | | 95% (93.2-96.8%) | 97,7% (96.5-98.9%) | Individual | Diagnostic (field) | Virus Neutralisation Test | Beaudeau et al., 2001a | | | | | |
| INGEZIM BVD COMPAC | INGENASA | | | | | | | | | | | | | | | | |
| PrioCHECK BVDV ab | Prionics | 98% (96-99%) | 99% (98-100%) | Diagnostic (field) | Virus neutralisation test | Kramps et al., 1999 | 65% (50-77%) | 100% (97-100%) | Individual | Diagnostic (field) | PrioCHECK on serum | Kramps et al., 1999 | | | | | |
| PrioCHECK BVDV ab Plus | Prionics | | | | | | | | | | | | | | | | |
| PrioCHECK BVDV ab Focus | Prionics | | | | | | | | | | | | | | | | |
| SVANOVIR® BVDV-Ab | SVANOVA Biotech AB | 98.2% | 100% | Technical (lab) | Virus neutralisation test | Svanova Biotech Ab, 2009 | 97.4 (95.2-99.0)<br>97.1 (95.2-98.5)<br>96.7 (93.4-99.6) | 98.7 (97.7-99.5)<br>97.8 (96.7-98.8)<br>98.4 (96.8-99.8) | Bulk milk | Diagnostic (field) | Virus isolation | Lindberg, 2000 | | | | | |
| CIVtest bovis BVD/BD p80 | HIPRA | | | | | | | | | | | | | | | | |
| SERELISA® BVD p80 Ab Mono Blocking | SYNBIOTICS Europe | | | | | | | | | | | | | | | | |
| BVD p80 Antibody competition | ID Vet | | | | | | | | | | | | | | | | |
| LSIVET BVD/BD p80 BLOCKING | Laboratoire Service International (LSI) | 96.9% (95.6-98.3%) | 97.8% (96.7-99.0%) | Diagnostic (field) | Virus Neutralisation Test | Beaudeau et al., 2001b | 96.9% (95.6-98.3%) | 97.3% (96-98.6%) | Individual | Diagnostic (field) | Virus Neutralisation Test | Beaudeau et al., 2001b | | | | | |

* PrioCheck (prionics) is the same as NS3 ELISA of CEDI Diagnostics

| Antigen ELISA's | Producer | Serum sample | | Technical (lab)/Diagnostic (field) | Gold standard | Reference | Milk sample | | Bulk milk/individual samples | Technical (lab)/Diagnostic (field) | Gold standard | Reference | Tissue sample | | Technical (lab)/Diagnostic (field) | Gold standard | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Se | Sp | | | | Se | Sp | | | | | Se | Sp | | | |
| BVDV (antigen) (on leucocytes) | BIO-X DIAGNOSTICS | | | | | | | | | | | | | | | | |
| Pulmotest BVDV (antigen) | BIO-X DIAGNOSTICS | | | | | | | | | | | | | | | | |
| \ | IDEXX | 100,00% | 100,00% | Diagnostic (field) | IHC | Hilbe et al., 2007 | | | | | | | | | | | |
| HerdChek BVDV Antigen Leukocytes (Bovine Viral Diarrhea Virus) | IDEXX | | | | | | | | | | | | | | | | |
| HerdChek BVDV Ag/Serum Plus (Bovine Viral Diarrhea Virus) | IDEXX | 99% / 100%* | 99.5% | Diagnostic (field) | PCR (* SerELISA BVD/MD Ag Mono-Indirect ) | Mars et al., 2005 | | | | | | | | | | | |
| HerdChek BVDV Antigen (Bovine Viral Diarrhea Virus) | IDEXX | | | | | | | | | | | | | | | | |
| INGEZIM BVD DAS | INGENASA | | | | | | | | | | | | | | | | |
| PrioCHECK BVDV ag | Prionics | | | | | | | | | | | | | | | | |
| SerELISA BVD/MD Ag Mono-Indirect | SYNBIOTICS Europe | 97% | 99% | Diagnostic (field) | Virus isolation | Brinkhof et al., 1996 | | | | | | | | | | | |

| PCR'S | Producer | Serum sample | | Technical (lab)/Diagnostic (field) | Gold standard | Reference | Milk sample | | Bulk milk/individual samples | Technical (lab)/Diagnostic (field) | Gold standard | Reference | Tissue sample | | Technical (lab)/Diagnostic (field) | Gold standard | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Se | Sp | | | | Se | Sp | | | | | Se | Sp | | | |
| Real time PCR / adjusted bij AHS | GD (Animal Health Service NL) in house test | | | | | | | | | | | | | | | | |
| ADIAVET BVD/MD | bioMérieux Deutschland | | | | | | | | | | | | | | | | |
| realtime PCR (virellaBVDV 2.0 real time RT-PCR Kit FLI-B 637) | Gerbion | | | | | | | | | | | | | | | | |
| realPCR BVDV RNA test | IDEXX | | | | | | | | | | | | | | | | |
| real BVDV | Ingenetix | | | | | | | | | | | | | | | | |
| LSI VetMAX BVDV 4ALL | Life Technologies | | | | | | | | | | | | | | | | |
| BVDV RT-PCR / virotype BVD RT-PCR kit | Qiagen | | | | | | | | | | | | | | | | |
| BoVir-SL BVDV realtime RT-PCR kit | Quidel | | | | | | | | | | | | | | | | |
| LSIVet BVDIL | Life Technologies | | | | | | | | | | | | | | | | |

# OVERVIEW OF THE RISK FACTORS FOR BVDV IN FREE AND NON-FREE TERRITORIES

## I. Risk factors at territory and herd level for BVD non-free territories.

| | | Dairy | Beef |
|---|---|---|---|
| a. Introduction | Territory 1 | Import/trade (TI animals and trojan cows) | Import/trade (TI animals and trojan cows) |
| | Territory 2 | Cattle farm density | Cattle farm density |
| | Herd 1 | Purchase/introduction of cattle (pregnant/trojan cows, cattle with unknown status, PI animals, TI animals ) | Purchase/introduction of cattle (pregnant/trojan cows, cattle with unknown status, PI animals, TI animals ) |
| | Herd 2 | Contact with neighbouring cattle | Contact with neighbouring cattle |
| | Herd 3 | Indirect contact with cattle in other herds through personnel/professional visitors, vehicles, fomites | Indirect contact with cattle in other herds through personnel/professional visitors, vehicles, fomites |
| | Herd 4 | Presence of beef cattle (fattening unit) on farm (animals not tested for BVD) | Natural breeding with a purchased bull |
| | Herd 5 | Location (underlying prevalence, advisory services, community attitudes etc) | Location (underlying prevalence, advisory services, community attitudes etc) |
| | Herd 6 | Inadequate quarantine for introduced or returning animals (e.g. unsold) | Inadequate quarantine for introduced or returning animals (e.g. unsold) |
| | Animal 1 | Age | Age |
| b. Delayed detection | Territory 1 | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) |
| | Territory 2 | Voluntary control programme | Voluntary control programme |
| | Territory 3 | Percentage of farms participating in the BVD control programme in case of a voluntary programme | Percentage of farms participating in the BVD control programme in case of a voluntary programme |
| | Territory 4 | Farmers demotivation on testing male calves (little economic value) | No BVD control in fattening farms |
| | Herd 1 | Delayed detection because introduction did not take place in the target group that is screened for BVD/nature of the disease | Delayed detection because introduction did not take place in the target group that is screened for BVD/nature of the disease |
| | Herd 2 | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) |
| | Herd 3 | Under reporting of clinical signs, abortions | Under reporting of clinical signs, abortions |
| | Herd 4 | Introduction of pregnant cows (delay between arrival and testing) | Introduction of pregnant cows (delay between arrival and testing) |
| | Herd 5 | Seasonal calving pattern | Seasonal calving pattern |
| | Herd 6 | | No BVD control in for example fattening farms |
| | Herd 7 | False negative test result | False negative test result |
| | Herd 8 | Extended interval between birth and testing | Extended interval between birth and testing |
| | Animal 1 | Age (the interpretation of test results can be influenced by the age of the animal) | Age (the interpretation of test results can be influenced by the age of the animal) |

## II. Risk factors at territory, herd and animal level for BVD-free territories.

| | | Dairy | Beef |
|---|---|---|---|
| a. Introduction | Territory 1 | Import of cattle (pregnant/trojan cows, cattle with unknown status, PI animals, TI animals ) | Import of cattle (pregnant/trojan cows, cattle with unknown status, PI animals, TI animals ) |
| | Herd 1 | Import of cattle (pregnant/trojan cows, cattle with unknown status, PI animals, TI animals ) | Import of cattle (pregnant/trojan cows, cattle with unknown status, PI animals, TI animals ) |
| | Herd 2 | Inadequate quarantine for imported animals | Inadequate quarantine for imported animals |
| b. Delayed detection | Territory 1 | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) |
| | Territory 2 | Voluntary control programme | Voluntary control programme |
| | Territory 3 | Percentage of farms participating in the BVD control programme in case of a voluntary programme | Percentage of farms participating in the BVD control programme in case of a voluntary programme |
| | Territory 4 | | No BVD control in fattening farms |
| | Herd 1 | Delayed detection because introduction did not take place in the target group that is screened for BVD/nature of the disease | Delayed detection because introduction did not take place in the target group that is screened for BVD/nature of the disease |
| | Herd 2 | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) | Farmer non-compliance with testing requirements (delayed tagging, submission of samples) |
| | Herd 3 | Introduction of pregnant cows (delay between arrival and testing) | Introduction of pregnant cows (delay between arrival and testing) |
| | Herd 4 | False negative test result | False negative test result |
| | Herd 7 | | No BVD control in fattening farms |

# BIBLIOGRAPHY


Beaudeau, F., Belloc, C., Seegers, H., Assie, S., Pourquier, P., & Joly, A. (2001a). Informative Value of an Indirect Enzyme-Linked Immunosorbent Assay (ELISA) for the Detection of Bovine Viral Diarrhoea Virus (BVDV) Antibodies in Milk. Zoonoses and Public Health, 48(9), 705-712.

BEAUDEAU, F., BELLOC, C., SEEGERS, H., ASSIE, S., SELLAL, E. & JOLY, A. (2001b ) Evaluation of a blocking virus (BVDV) antibodies in serum and milk. Veterinary Microbiology ELISA for the detection of bovine viral diarrhoea 80, 329-337

Brinkhof, J., Zimmer, G., Westenbrink, F., 1996. Comparative study on four enzyme-linked immunosorbent assays and a cocultivation assay for the detection of antigens associated with the bovine virus diarrhoea virus in persistently infected cattle. Vet. Microbiol. 50, 1–6l.

Hashemi Tabar, G.R., Haghparast, A., Naseri, Z., 2011. Prevalence of bovine viral diarrhoea virus antibodies and antigen among the aborted cows in industrial dairy cattle herds in Mashhad area of Iran. Archives of Razi Institute 66, 17–23.

Hilbe, M., Stalder, H., Peterhans, E., Haessig, M., Nussbaumer, M., Egli, C., & Ehrensperger, F. (2007). Comparison of five diagnostic methods for detecting bovine viral diarrhea virus infection in calves. Journal of Veterinary Diagnostic Investigation, 19(1), 28-34.

Kuijk, H., et al., Monitoring of BVDV in a vaccinated herd by testing milk for antibodies to NS3 protein. Vet Rec, 2008. 163(16): p. 482-4.

Lindberg, A., 2000. Validation of an antibody ELISA used on bulk milk to screen dairy herds for bovine viral diarrhoea virus infection. In Proceedings of the 9th International Symposium on Veterinary Epidemiology and Economics.

Makoschey, B., Sonnemans, D., Muñoz Bielsa, J., Franken, P.,Mars, M. H., Santos, L. & Álvarez, M. (2007) Evaluation of the induction of NS3 specific BVDV antibodies using a commercial inactivated BVDV vaccine in immunization and challenge trials. Vaccine 25, 6140-6145

Mars, M.H. and C. Van Maanen, Diagnostic assays applied in BVDV control in The Netherlands. Prev Vet Med, 2005. 72(1-2): p. 43-8; discussion 215-9.

Sayers, R.G., et al., Impact of three inactivated bovine viral diarrhoea virus vaccines on bulk milk p80 (NS3) ELISA test results in dairy herds. Vet J, 2015. 205(1): p. 56-61.

Svanova Biotech Ab, 2009. Product info SVANOVIR BVDV-Ab. <http://www.svanova.com/> (accessed 12-2-2018).

# Annexe 3 : Quantification of risk factors for bovine viral diarrhea virus in cattle herds : A systematic search and meta-analysis of observational studies, 2020.

A. M. van Roon, M. Mercat, G. van Schaik, M. Nielen, D. A. Graham, S. J. More, M. Guelbenzu-Gonzalo, C. Fourichon, A. Madouasse, and I. M. G. A. Santman-Berends.

J. Dairy Sci. 103:9446–9463
https://doi.org/10.3168/jds.2020-18193


# Quantification of risk factors for bovine viral diarrhea virus in cattle herds: A systematic search and meta-analysis of observational studies

**A. M. van Roon,[1]* M. Mercat,[2] G. van Schaik,[1,3]† M. Nielen,[1] D. A. Graham,[4] S. J. More,[5] M. Guelbenzu-Gonzalo,[4] C. Fourichon,[2] A. Madouasse,[2] and I. M. G. A. Santman-Berends[1,3]**
[1]Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, 3508, TD Utrecht, the Netherlands
[2]INRAE, Oniris, BIOEPAR, 44300, Nantes, France
[3]Royal GD, 7400 AA, Deventer, the Netherlands
[4]Animal Health Ireland, Carrick on Shannon, Co. Leitrim N41 WN27, Ireland
[5]Centre for Veterinary Epidemiology and Risk Analysis, UCD School of Veterinary Medicine, University College Dublin, Belfield, Dublin D04 W6F6, Ireland

## ABSTRACT

Bovine viral diarrhea virus (BVDV) is endemic in many parts of the world, and multiple countries have implemented surveillance activities for disease control or eradication. In such control programs, the disease-free status can be compromised by factors that pose risks for introduction or persistence of the virus. The aim of the present study was to gain a comprehensive overview of possible risk factors for BVDV infection in cattle herds in Europe and to assess their importance. Papers that considered risk factors for BVDV infection in cattle were identified through a systematic search. Further selection of papers eligible for quantitative analysis was performed using a predefined checklist, including (1) appropriate region (i.e., studies performed in Europe), (2) representativeness of the study population, (3) quality of statistical analysis, and (4) availability of sufficient quantitative data. In total, 18 observational studies were selected. Data were analyzed by a random-effects meta-analysis to obtain pooled estimates of the odds of BVDV infection. Meta-analyses were performed on 6 risk factors: herd type, herd size, participation in shows or markets, introduction of cattle, grazing, and contact with other cattle herds on pasture. Significant higher odds were found for dairy herds (odds ratio, OR = 1.63, 95% confidence interval, CI: 1.06–2.50) compared with beef herds, for larger herds (OR = 1.04 for every 10 extra animals in the herd, 95% CI: 1.02–1.06), for herds that participate in shows or markets (OR = 1.45, 95% CI: 1.10–1.91), for herds that introduced cattle into the herd (OR = 1.41, 95% CI: 1.18–1.69), and for herds that share pasture or have direct contact with cattle of other herds at pasture (OR = 1.32, 95% CI: 1.07–1.63). These pooled values must be interpreted with care, as there was a high level of heterogeneity between studies. However, they do give an indication of the importance of the most frequently studied risk factors and can therefore assist in the development, evaluation, and optimization of BVDV control programs.
**Key words:** risk factor, bovine viral diarrhea virus, review, meta-analysis, Europe

## INTRODUCTION

Bovine viral diarrhea (**BVD**) virus (**BVDV**) is a pestivirus belonging to the *Flaviviridae* family (Olafson and Rickard, 1947). It is one of the most common viral diseases in cattle and endemic in many parts of the world (Scharnböck et al., 2018). Bovine viral diarrhea virus is mainly spread by persistently infected (**PI**) cattle, which were infected in utero between 40 and 120 d of gestation and shed large amounts of virus into the environment after birth (McClurkin et al., 1984). Bovine viral diarrhea virus can be transmitted directly through nose-to-nose contact between cattle or indirectly through contaminated materials (Tråvén et al., 1991; Niskanen and Lindberg, 2003). Infections with BVDV can lead to respiratory and reproductive issues, causing major economic losses (Houe, 2003). Many European countries have implemented BVDV control or eradication programs, and some have already successfully eradicated the virus or reached a herd-level prevalence below 1.5% (Sweden, Norway, Finland, Denmark, Germany, Austria, Switzerland, and Ireland; Nuotio et al., 1999; Bitsch et al., 2000; Hult and Lindberg, 2005; Rikula et al., 2005; Rossmanith et al., 2010; Presi et al., 2011; Norström et al., 2014; Foddai et al., 2016; AHI, 2019). Within those control programs, animals, herds,

Received January 13, 2020.
Accepted May 20, 2020.
*Corresponding author: a.m.vanroon@uu.nl
†On behalf of the STOC free consortium: Jörn Gethmann, Carola Sauter-Louis, Jenny Frössling, Estelle Ågren, George Gunn, Madeleine Henry, and Jude Eze.

regions, or the country are ascribed a BVDV-free status that is subsequently monitored.

The probability that a herd categorized as free within a control program is truly free of infection will be influenced by risk factors for introduction of the virus (i.e., the probability that the virus is introduced or reintroduced into the herd between test moments) and factors that cause delayed detection of the virus after introduction or reintroduction (i.e., the probability that the virus had been introduced but not yet detected). The effectiveness of surveillance relies on an understanding of these risk factors. Delayed detection of the virus can be associated with herd management, control program design (e.g., test population, test frequency, sample size, test validity), and test performance. Risk factors for introduction depend on the contact structure between herds, such as purchase or contact with cattle from neighboring herds. The introduction of purchased animals is a well-known risk factor. However, an overview of the magnitude of the risk, and of country-level differences, is lacking.

Risk factors for introduction and delayed detection of BVDV are not easily studied in isolation due to the difficulty of determining exactly when the virus is introduced into a herd. Risk factors for the presence of infection are more often reported (e.g., Graham et al., 2013; Byrne et al., 2017; Amelung et al., 2018) and could serve as a proxy for introduction and delayed detection. In this study, we have conducted a systematic literature search, seeking to gain a comprehensive overview of possible risk factors for the presence of BVDV infection in cattle herds in Europe. We aimed to assess the importance of the most frequently studied risk factors and, depending on study quality and the availability of quantitative data, to perform meta-analyses to obtain pooled values. This information is critical for the development, evaluation, and optimization of BVDV control programs. Control program managers can list and prioritize risk factors in their country based on the pooled values or choose the results from countries most comparable with their situation.

## MATERIALS AND METHODS

This systematic review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) with the PRISMA 2009 Checklist (Supplemental File S1, https://doi.org/10.3168/jds.2020-18193).

### Search Strategy

Three databases (PubMed, CAB Abstracts, and Scopus) were interrogated using the search terms defined below. The final complete data search in all 3 databases was performed on September 21, 2018. An additional search was performed after the full-text screening and before data analysis on July 15, 2019. This additional search was performed only in PubMed because Scopus and CAB Abstracts do not allow selection for specific publication dates, only per year.

The research questions include 4 key aspects: BVDV, risk factors, introduction, and delayed detection. The BVDV search terms included the following: BVD, BVDV, bovine viral diarrh(o)ea, bovine viral diarrh(o)ea, and bovine viral diarrh(o)ea virus. Risk factor search terms included the following: risk factor, purchase, import, trade, market, grazing, nose-to-nose contact, direct contact, over the fence contact, density, contact structure, herd, herd size, seasonal calving, calving pattern, housing system, management, biosecurity, vaccination, artificial insemination, embryo transfer, PI, persistent infection, and persistently infected. Introduction search terms included the following (where * indicates a wildcard): introduction, pathway, epidemio*, incidence, prevalence, and contamin*. Finally, delayed detection search terms included the following: diagnostic test, persist*, delayed detection, test strategy, test scheme, test performance, test characteristics, sensitivity, control program*, eradication program*, surveillance, false negative, free, freedom, transmission, and spread. The full electronic search strategy is included in Supplemental File S2 (https://doi.org/10.3168/jds.2020-18193).

### Study Selection

Studies published in peer-reviewed journals with full text available were considered. They reported either risk factors for introduction of BVDV in cattle herds or risk factors for the presence of BVDV from which risk factors for introduction could be inferred. During the initial screening, studies were also included from which risk factors for delayed detection could be inferred (e.g., studies reporting test characteristics). In a later stage, it was decided to focus on risk factors for introduction and presence of BVDV to narrow down the search. Only studies with a cross-sectional, cohort, case-control, or randomized controlled trial study design were considered. Languages that were accepted were English, Dutch, French, Spanish, and German. Studies published since 1980 were included to focus on modern farm management systems.

The search in PubMed, CAB Abstracts, and Scopus was carried out by one researcher (AvR). The researcher imported all references into the online systematic review management tool Covidence (Veritas Health Innovation, Melbourne, VIC, Australia). In Covidence, duplicates were deleted automatically or following a

manual review. Two researchers (AvR and MM) both went independently through the following consecutive phases of the review: (1) screening titles and abstracts based on the inclusion criteria described above, and (2) reviewing full-text articles based on the inclusion criteria described above. After these review steps, conflicting opinions on papers were discussed with the other co-authors to reach consensus on inclusion or exclusion.

All full-text studies that were selected based on the inclusion criteria were further assessed for their appropriateness for meta-analyses by one researcher (AvR). This was done using the approach presented in Table 1. This checklist consists of 4 questions regarding internal validity (how well is the study conducted?) and external validity (generalizability). As no generic tool is available for appraisal of observational studies for meta-analysis (Sanderson et al., 2007), we created our own checklist with relevant checkpoints based on our own observations and in alignment with the methods used in previous studies (National Institutes of Health, 2014; Downes et al., 2016).

On several occasions, multiple studies were described in a single paper (so-called split studies)—for example, if a risk factor study was performed on different outcome variables (e.g., antibody or virus) or different types of cattle (e.g., beef or dairy) or if more than one final risk factor model was developed. We decided to include both split studies where beef and dairy herds were analyzed separately because these risk factor analyses were performed on different populations (e.g., Gates et al., 2013, 2014). When studies concluded with more than one final model, the model indicated by the authors as best describing the data was included. If no choice was made between the different final models, we selected the model that took into account the full data set. Risk factor analyses performed on subsets of the data were excluded.

### Data Collection

Data were extracted from all selected studies using an Excel (Microsoft Corp., Redmond, WA) form that was prepared in advance. Data were extracted by one researcher (AvR) and checked by the other researcher (MM). A pilot test of the Excel form was conducted by these 2 researchers working together on 3 selected papers to increase uniformity in extracting the data.

For each selected study, detailed data were extracted regarding study type, location, size of the study population, diagnostic tests used, risk factors studied in univariable and multivariable analysis, the effect size (odds ratio, **OR**; relative risk), confidence intervals, and the statistical analysis that was performed.

### Meta-Analysis

All risk factors from the studies that were selected for quantitative analysis were listed and combined into groups of similar risk factors. Per group, OR reported in at least 2 independent studies were analyzed by a random-effects meta-analysis to obtain pooled estimates of the odds of BVDV infection. In some cases, variables first had to be restructured to be able to include them in the meta-analysis. For example, this was the case with introduction of cattle where we wanted to combine variables with "yes introduction" versus "no introduction" with categorical variables where different numbers of introduced cattle were compared with zero introduction. In this case, we first performed a within-study fixed-effects meta-analysis on the different categories of this variable to obtain a summary estimate across all categories. This summary estimate could subsequently be included in the overall meta-analysis for introduction of cattle.

**Table 1.** Checklist study appraisal for quantitative analysis

| Item | Not appropriate for meta-analysis | Appropriate for meta-analysis |
|---|---|---|
| External validity | | |
| 1. Is the cattle production system comparable with the European situation? | Studies were performed outside of Europe. | Studies were performed in Europe. |
| 2. Are the selected animals or herds representative of the target population (commercial cattle herds in Europe)? | No, with high possibility of selection bias. Animals or herds are selected purposively. | Yes, with low or medium possibility of selection bias. Animals or herds are selected randomly or in a way that represents the target population. |
| Internal validity | | |
| 3. Was the unit of interest appropriate for a herd-level risk factor study? | Animal-level data were used without correction for within-herd correlation. | Herd-level data or animal-level data that were corrected for clustering were used. |
| 4. Are quantitative data available? | No, there are only descriptive studies, or some quantitative data but no odds ratios or data from which odds ratios could be derived. | Yes, there are quantitative data (odds ratios or data to derive odds ratios) of univariable or multivariable analysis. |

A random-effects approach is considered the default method in meta-analysis of observational studies (Mueller et al., 2018). This approach accounts for the fact that the study effect estimates are not drawn from a single population, which would be the case when using a fixed-effects approach (Harrer et al., 2019). The random-effects models were fitted in a 2-step approach. First, between-study variance, represented by the distribution of the true study effects, was estimated with the DerSimonian-Laird approach. Then, weights were assigned to all included studies based on the inverse of the variance as in general the population size between observational studies is not equal and pooled OR were estimated (Viechtbauer, 2010). In this process, the OR and their 95% confidence intervals (**CI**) as reported in the individual studies were log-transformed; therefore, due to rounding errors, the 95% CI in our results might differ slightly from the data reported in the individual studies. Preferably, adjusted OR that resulted from multivariable analysis were used. When no multivariable results were available, crude OR that resulted from univariable analysis were included. If no OR were available but frequencies were reported, OR were calculated. In each forest plot, the univariable results were marked. Also, subanalyses were performed in which univariable and multivariable results were analyzed separately.

Heterogeneity between studies was studied by the $I^2$ statistic. The $I^2$ statistic shows what proportion of the variance is due to heterogeneity in true effects rather than sampling error (Borenstein et al., 2017). To identify studies with the greatest influence on the results, an influential case analysis was performed with cut-off values proposed by Viechtbauer and Cheung (2010). The studies indicated as outliers were marked in each forest plot. The change in the summary estimates and $I^2$ statistic when retaining or removing outliers was of minor importance. Publication bias could not be properly assessed due to the low number of studies included in our meta-analyses ($n < 10$; Higgins et al., 2019). Funnel plots were checked for asymmetry, with some indication of publication bias, but these plots are not reported as it was not possible to determine whether this was by chance or real asymmetry due to the low number of studies. Meta-analyses were performed using R statistical software (R Core Team, 2019) and the metafor package (Viechtbauer, 2010).

## RESULTS

### Literature Search

The original searches revealed 12,028 papers, of which ultimately 259 papers were full-text screened and narrowed down to 51 papers (Figure 1). Based on Table 1, all 51 papers were screened for their appropriateness for quantitative analyses (Supplemental File S3, https://doi.org/10.3168/jds.2020-18193). Eventually, 18 papers (20 studies) were selected for inclusion in the meta-analysis (Table 2).

### Overview of Risk Factors

All risk factors that were studied in the final 18 papers were grouped into 6 risk factor categories: (1) herd and animal characteristics, (2) cattle movement, (3) reproduction, (4) neighborhood risk, (5) farm management and biosecurity, and (6) diagnostic testing and control programs.

### Description of Risk Factors

***Herd and Animal Characteristics.*** Herd and animal characteristics that were studied included milk yield, sex, age, infection with other pathogens, mortality, region, herd type, and herd size. Of all herd characteristics, variables describing herd size, herd type, and region were included most frequently (Table 3).

No further analysis could be performed on milk yield and sex as for both there was only 1 study with quantitative data. Age was included as a categorical variable in 2 studies (Mainar-Jaime et al., 2001; Hanon et al., 2018), both with higher OR for the presence of BVD antibodies in higher age classes. However, the age categories within those 2 studies were not comparable and therefore were unsuitable for meta-analysis. Infection with other pathogens associated with BVD infection was considered in 3 studies but could not be compared because different pathogens were studied (i.e., *Neospora caninum*, bovine herpes virus-1, and bovine tuberculosis). Mortality was considered in 5 studies, but as this was regarded more an outcome than a risk factor for BVDV, it was not included in the meta-analysis. Finally, region was not included in the meta-analysis even though this was one of the most studied risk factors within the herd and animal characteristics group. Because different regions were included in different studies, comparison of the risk estimates between regions was impossible. Nevertheless, most studies found significant differences between regions, which makes this an important risk factor to consider. Meta-analysis was performed on herd type and herd size.

***Cattle Movement.*** Movement characteristics that were studied included introduction of cattle, cattle shows or markets, and other movements (e.g., sale and exchange of calves). Of all cattle movement risk factors, variables describing introduction of cattle into a herd

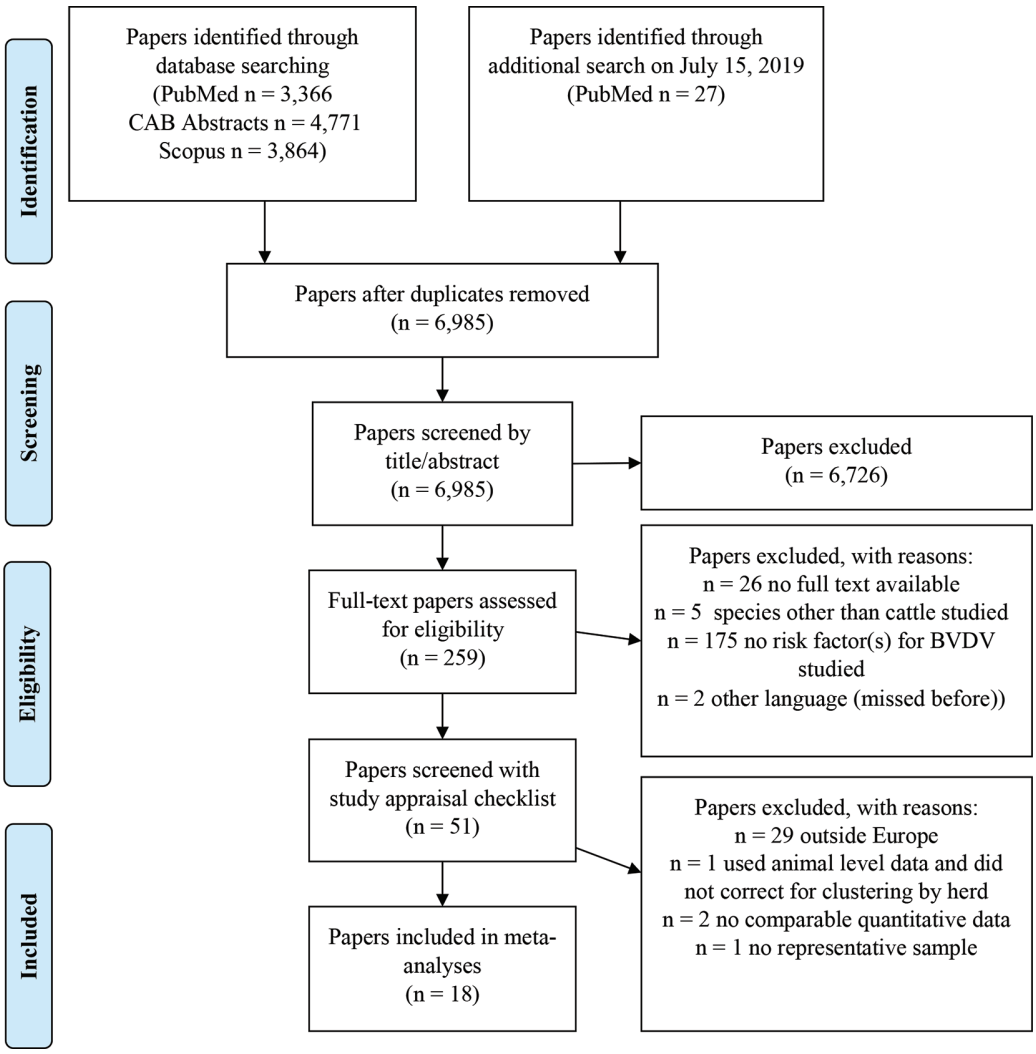



**Figure 1.** Flow diagram showing the total number of papers identified and excluded per stage of the selection process. At the eligibility stage, we decided to exclude papers that were initially selected for delayed detection. The 175 "no risk factor(s) for BVDV studied" papers were about BVDV test characteristics. BVDV = bovine viral diarrhea virus.

were included most frequently (Table 4). We considered studies on introduction of cattle into a herd and on purchase, where the latter assumes monetary transfer, which is not necessarily the case with introduction. In this paper, we use "introduction," which also covers purchase.

Other types of cattle movements were studied by Valle et al. (1999) and Amelung et al. (2018). Valle et al. (1999) looked at "other animal traffic," combining mainly exchange of calves and sharing of cattle housing with other farmers during summer. They found a very high OR of 28.60 (95% CI: 3.23–252.22). Amelung et al. (2018) studied sale of cattle, which was not comparable with the cattle movement studied in Valle et al. (1999). Meta-analysis was performed on cattle shows or markets and introduction of cattle.

***Reproduction.*** Reproduction variables that were studied included AI versus use of bulls and calving pattern (Table 5). The number of studies was too small or the definition of the variables varied too much between studies to enable a meta-analysis to be conducted.

Variables regarding AI or the use of bulls were included only in univariable analyses. In Amelung et al. (2018), higher but nonsignificant OR were found for BVD infection in herds with AI (OR = 1.28, 95% CI: 0.96–1.71) compared with herds without AI but also in herds with a bull for insemination (OR = 1.17, 95% CI: 0.93–1.48) compared with herds without a bull.

**Table 2.** Studies selected for the meta-analyses

| Study ID[1] | Study | Country | Study design | Unit of interest in risk factor analysis | Type of cattle studied | Outcome measure | Diagnostic test[2] |
|---|---|---|---|---|---|---|---|
| 2 | Amelung et al. (2018) | Germany | Cross-sectional | 2,542 herds | Combination | Virus | ELISA on ear notch followed by PCR on ear notch |
| 5B[3] | Barrett et al. (2018) | Ireland | Cross-sectional | 139 herds | Beef | Virus | On ear notch |
| 7 | Bishop et al. (2010) | Wales | Cross-sectional | 36 herds | Dairy | Antibodies | ELISA on BTM |
| 9B[3] | Byrne et al. (2017) | Northern Ireland | Cross-sectional | 2,827 herds | Combination | Virus | PCR on ear notch |
| 11 | Charoenlarp et al. (2018) | Northern Ireland | Cross-sectional | 17,186 herds | Combination | Virus | ELISA, PCR, or both on ear notch |
| 15 | Ersbøll et al. (2010) | Denmark | Cohort | 7,921 herds | Dairy | Virus | ELISA on BTM and blood |
| 18A[4] | Gates et al. (2014) | Scotland | Cross-sectional | 255 herds | Beef | Antibodies | ELISA on blood |
| 18B[4] | Gates et al. (2014) | Scotland | Cross-sectional | 189 herds | Dairy | Antibodies | ELISA on blood |
| 19A[4] | Gates et al. (2013) | Scotland | Case-control | 249 herds (65 cases and 184 controls) | Beef | Antibodies | ELISA on blood |
| 19B[4] | Gates et al. (2013) | Scotland | Case-control | 185 herds (119 cases and 66 controls) | Dairy | Antibodies | ELISA on blood |
| 20A[3] | Graham et al. (2013) | Ireland | Cross-sectional | 3,894 herds | Combination | Virus | ELISA or PCR on ear notch |
| 21A[3] | Graham et al. (2016) | Ireland | Cross-sectional | 58,479 herds | Combination | Virus | Unknown |
| 22 | Hanon et al. (2018) | Belgium | Cross-sectional | 51 herds and 3,017 cattle | Combination | Antibodies | Different ELISA and VNT on blood and milk |
| 24A[3] | Houe et al. (1995a,b) | Denmark | Cross-sectional | 19 herds | Dairy | Virus | Virus isolation and virus neutralization on blood |
| 30 | Mainar-Jaime et al. (2001) | Spain | Cross-sectional | 529 cattle | Dairy | Antibodies | ELISA on blood |
| 31 | Martinez-Ibeas et al. (2015) | Republic of Ireland | Cross-sectional | 305 herds | Dairy | Antibodies | ELISA on BTM and blood |
| 35 | Presi et al. (2011) | Switzerland | Cross-sectional | 33,188 herds | Combination | Virus | ELISA or PCR on ear notch |
| 40A[3] | Sarrazin et al. (2013) | Belgium | Cross-sectional | 664 herds | Combination | Antibodies and virus | ELISA on blood |
| 49A[3] | Valle et al. (1999) | Norway | Case-control | 314 herds (162 cases and 152 controls) | Dairy | Antibodies | BTM screening and pooled milk sample followed by ELISA on blood |
| 50 | Williams and Winden (2014) | United Kingdom | Cross-sectional | 1,088 herds | Dairy | Antibodies | ELISA on BTM |

[1]Study ID were assigned to the 51 papers that were selected in the second-last selection step (Supplemental File S3, https://doi.org/10.3168/jds.2020-18193).

[2]BTM = bulk tank milk; VNT = virus neutralization test.

[3]These rows represent one of 2 or 3 studies presented in a single paper. Each of these studies was chosen for inclusion in further analyses because they either present the best final model or were performed on the full data set. Excluded split studies can be found in Supplemental File S3.

[4]These rows represent one of 2 split studies presented in a single paper. Each of these studies had been conducted on different herds (beef or dairy) and has been analyzed separately.

**Table 3.** Overview of the number of risk factor studies (out of the selected 18 papers on 20 studies) that included herd and animal characteristics and the availability of quantitative data

| Factor | No. of studies | No. of variables | Study ID[1] | No. of variables with quantitative data |
|---|---|---|---|---|
| Milk yield | 2 | 2 | 2, 30 | 1 |
| Sex | 1 | 1 | 22 | 1 |
| Age | 2 | 2 | 22, 30 | 2 |
| Infection with other pathogens | 3 | 6 | 5B, 9B, 30 | 3 |
| Mortality | 5 | 7 | 5B, 9B, 20A, 30, 35 | 5 |
| Region | 8 | 8 | 2, 9B, 11, 15, 20A, 21A, 30, 31 | 7 |
| Herd type[2] | 9 | 11 | 2, 9B, 11, 20A, 21A, 22, 30, 35, 40A | 9 |
| Herd size[2] | 14 | 20 | 2, 5B, 9B, 11, 15, 20A, 21A, 22, 24A, 30, 31, 35, 40A, 50 | 13 |

[1]Study ID were assigned to the 51 papers that were selected in the second-last selection step (see Supplemental File S3, https://doi.org/10.3168/jds.2020-18193).

[2]Included in the meta-analysis.

**Table 4.** Overview of the number of risk factor studies (out of the selected 18 papers on 20 studies) that included cattle movement variables and the availability of quantitative data

| Factor | No. of studies | No. of variables | Study ID[1] | No. of variables with quantitative data |
|---|---|---|---|---|
| Other movement | 2 | 2 | 2, 35 | 2 |
| Cattle shows or markets[2] | 5 | 5 | 2, 19A, 19B, 22, 35 | 5 |
| Introduction of cattle[2] | 17 | 62 | 2, 5B, 7, 9B, 18A, 18B, 19A, 19B, 20A, 21A, 22, 24A, 30, 31, 35, 49A, 50 | 48 |

[1]Study ID were assigned to the 51 papers that were selected in the second-last selection step (see Supplemental File S3, https://doi.org/10.3168/jds.2020-18193).

[2]Included in the meta-analysis.

**Table 5.** Overview of the number of risk factor studies (out of the selected 18 papers on 20 studies) that included reproduction variables and the availability of quantitative data

| Factor | No. of studies | No. of variables | Study ID[1] | No. of variables with quantitative data |
|---|---|---|---|---|
| Calving pattern | 1 | 1 | 50 | 1 |
| AI/use of bulls | 3 | 4 | 2, 7, 50 | 3 |

[1]Study ID were assigned to the 51 papers that were selected in the second-last selection step (see Supplemental File S3, https://doi.org/10.3168/jds.2020-18193).

**Table 6.** Overview of the number of risk factor studies (out of the selected 18 papers on 20 studies) that included neighborhood variables and the availability of quantitative data

| Factor | No. of studies | No. of variables | Study ID[1] | No. of variables with quantitative data |
|---|---|---|---|---|
| Farm fragmentation | 1 | 1 | 20A | 1 |
| Environment | 1 | 4 | 11 | 4 |
| Cattle density | 6 | 9 | 11, 15, 19A, 19B, 21A, 30 | 7 |
| BVD[2]-positive neighbor herds | 3 | 11 | 11, 15, 21A | 8 |
| Contact with other animal species | 5 | 10 | 2, 19A, 19B, 20A, 49A | 8 |
| Pasture[3] | 8 | 20 | 2, 11, 19A, 19B, 22, 24A, 35, 49A | 14 |

[1]Study ID were assigned to the 51 papers that were selected in the second-last selection step (Supplemental File S3, https://doi.org/10.3168/jds.2020-18193).

[2]Bovine viral diarrhea.

[3]Included in the meta-analysis.

**Table 7.** Overview of the number of risk factor studies (out of the selected 18 papers on 20 studies) that included farm management and biosecurity variables and the availability of quantitative data

| Factor | No. of studies | No. of variables | Study ID[1] | No. of variables with quantitative data |
|---|---|---|---|---|
| Other biosecurity | 2 | 2 | 19A, 19B | 2 |
| Hygiene | 2 | 2 | 19A, 19B | 2 |
| Quarantine | 3 | 3 | 7, 19A, 19B | 2 |
| Mixed beef and dairy farm | 3 | 3 | 19A, 19B, 49A | 2 |
| Vaccination | 3 | 4 | 22, 31, 40A | 2 |
| Housing | 2 | 4 | 2, 22 | 4 |
| Shared equipment | 3 | 5 | 19A, 19B, 49A | 4 |
| People on farm | 2 | 8 | 19A, 19B | 8 |

[1]Study ID were assigned to the 51 papers that were selected in the second-last selection step (Supplemental File S3, https://doi.org/10.3168/jds.2020-18193).

Williams and Winden (2014) compared herds with a bull present on the farm with herds with AI only and found that herds with a bull present on the farm had higher but nonsignificant odds of infection with BVD (OR = 1.16, 95% CI: 0.90–1.49). Calving pattern was found only once in a univariable risk factor analysis and showed higher odds of infection (OR = 1.80, 95% CI: 1.22–2.67) in herds with year-round calving compared with seasonal calving (Williams and Winden, 2014).

***Neighborhood Risk.*** Variables related to neighborhood risk included farm fragmentation, environment, cattle density, BVD-positive neighbor herds, contact with other animal species, and pasture. Of all neighborhood risk factors, variables describing cattle density, contact with other animal species, and pasture were included most frequently (Table 6).

Farm fragmentation (number of individual noncontiguous parcels of land associated with the herd) and environment (i.e., natural grassland, forest) were both studied only once; therefore, no meta-analysis could be performed. Cattle density and BVD-positive neighbor herds were studied more frequently but in such different ways that meta-analysis was not possible. Both variables describe in different ways the distance to (positive) neighboring herds or the number of (positive) neighboring herds contiguous to the farm or in a 5- or 10-km radius. They are continuous or categorical. Most studies showed higher odds of BVD infection when the distance to (positive) neighbors is shorter, when there are more (positive or unknown status) neighbors close by, or when BVD-positive animals are retained for a longer period. One study found that seropositivity increased with a larger distance (in km) to the closest dairy farm (Mainar-Jaime et al., 2001). Variables regarding contact with other animal species included the presence of, contact with, close proximity of, or grazing with sheep, pigs, deer, or wildlife. No meta-analysis could be performed on contact with other animal species.

***Farm Management and Biosecurity.*** Variables included were quarantine, vaccination, mixed beef and dairy farm, type of housing, shared equipment, people on farm, and other biosecurity. None of these variables were suitable for meta-analysis because of noncomparable definitions or the low number of studies in which these factors were studied (Table 7).

Most farm management and biosecurity variables were studied by Gates et al. (2013). They studied the relative influence of cattle movements, local spread, and biosecurity on BVDV seropositivity. The variables we included in the farm management and biosecurity group were not exactly identical to the classification of biosecurity variables in the study of Gates et al. (2013), but especially for beef herds, cattle movement had the greatest influence on BVDV seropositivity. Also, in the other studies included in Table 7, most biosecurity variables were nonsignificant.

***Diagnostics Testing and Control Programs.*** Multiple papers studied variables related to diagnostic testing and control programs that we grouped into BVDV testing, farmer behavior, control program, and other (Table 8). However, either the number of studies was too small or the definition of these variables varied too much between studies to enable a meta-analysis to be conducted.

Within the diagnostic testing and control programs group, BVDV testing was studied most. Examples of variables studied are the total number of BVDV tests undertaken and detection of PI animals in the past. One study (Amelung et al., 2018) found that herds that participated in a control program has slightly higher odds (OR = 1.28, 95% CI: 1.01–1.64) for BVDV infection in univariable analysis than herds that do not participate. One of the studies looking at farmer behavior showed that the age of farmers was associated with the BVD status. Herds of farmers younger than 40 yr were more often infected than herds of farmers between 50 and 60 yr.

**Table 8.** Overview of the number of risk factor studies (out of the selected 18 papers on 20 studies) that included diagnostic testing and control program variables and the availability of quantitative data

| Factor | No. of studies | No. of variables | Study ID[1] | No. of variables with quantitative data |
|---|---|---|---|---|
| Other | 3 | 3 | 11, 40A, 21A | 2 |
| Farmer (behavior) | 2 | 4 | 2, 49A | 3 |
| Control program | 3 | 3 | 2, 11, 22 | 3 |
| BVDV[2] testing | 7 | 8 | 9B, 19A, 19B, 20A, 30, 31, 40A | 6 |

[1]Study ID were assigned to the 51 papers that were selected in the second-last selection step (Supplemental File S3, https://doi.org/10.3168/jds.2020-18193).
[2]Bovine viral diarrhea virus.

## Meta-Analyses

***Herd and Animal Characteristics.*** Herd type was studied frequently and was always included as a categorical variable (i.e., dairy, beef, mixed, beef breeding; Supplemental File S4, section 4.1.1, https://doi.org/10.3168/jds.2020-18193). A meta-analysis was conducted on the 6 studies that compared dairy versus beef herds (reference category; Supplemental File S4, section 4.1.2). We found a combined effect estimate of

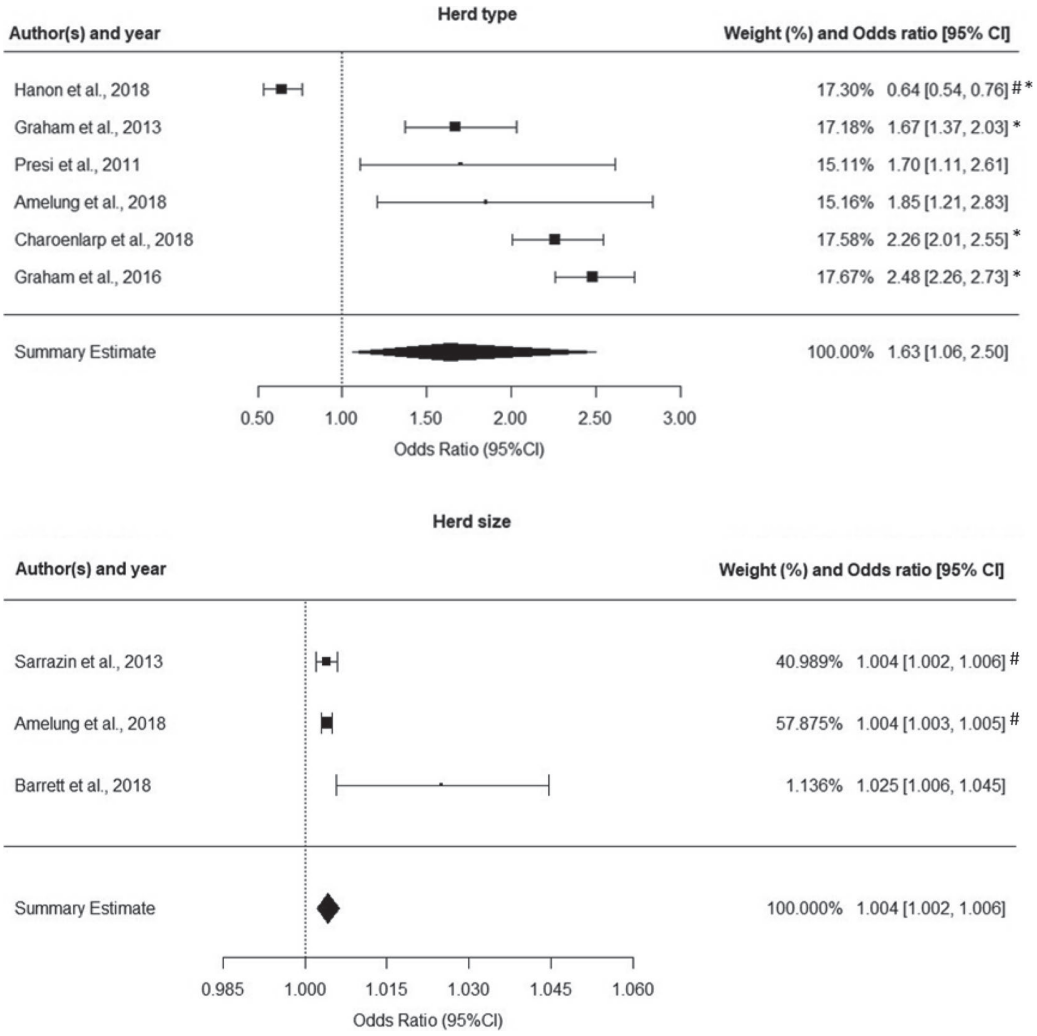



**Figure 2.** Forest plot of the effect of herd type with beef herds as reference category (upper plot) and herd size per additional animal in the herd (lower plot) on bovine viral diarrhea virus infection. *Univariable result; #study indicated as outlier in the influential case analysis.

1.63 higher odds (95% CI: 1.06–2.50) of BVDV infection in dairy herds compared with beef herds (Figure 2). The heterogeneity between studies ($I^2$) was 97.30% (95% CI: 91.87–99.47).

Herd size was studied frequently and was always included as an either categorical or continuous variable (Supplemental File S4, section 4.1.1). However, very few variables were comparable; therefore, meta-analysis was conducted on the 4 studies with OR per additional cow (Supplemental File S4, section 4.1.2). Other variables showing the log number of cows or different herd size categories were not included because they were not comparable.

For every extra animal in the herd, we found a combined effect estimate of 1.004 higher odds (95% CI: 1.002–1.006) of BVDV infection (Figure 2). For every 10 extra animals in the herd, this would be 1.04 higher odds of BVDV infection (95% CI: 1.02–1.06). The results of Presi et al. (2011) could not be included in the pooled estimate because weights are assigned to all factors based on the inverse of the variance and these results had a variance of zero. The heterogeneity between studies ($I^2$) was 55.96% (95% CI: 0.00–99.98).

***Cattle Movement.*** In all studies, participation in cattle shows or markets was included as a yes–no variable (Supplemental File S4, section 4.2.1) and therefore they could all be included in meta-analysis (Supplemental File S4, section 4.2.2). We found a combined effect estimate of 1.45 higher odds (95% CI: 1.10–1.91) of BVDV infection in herds that participated in shows

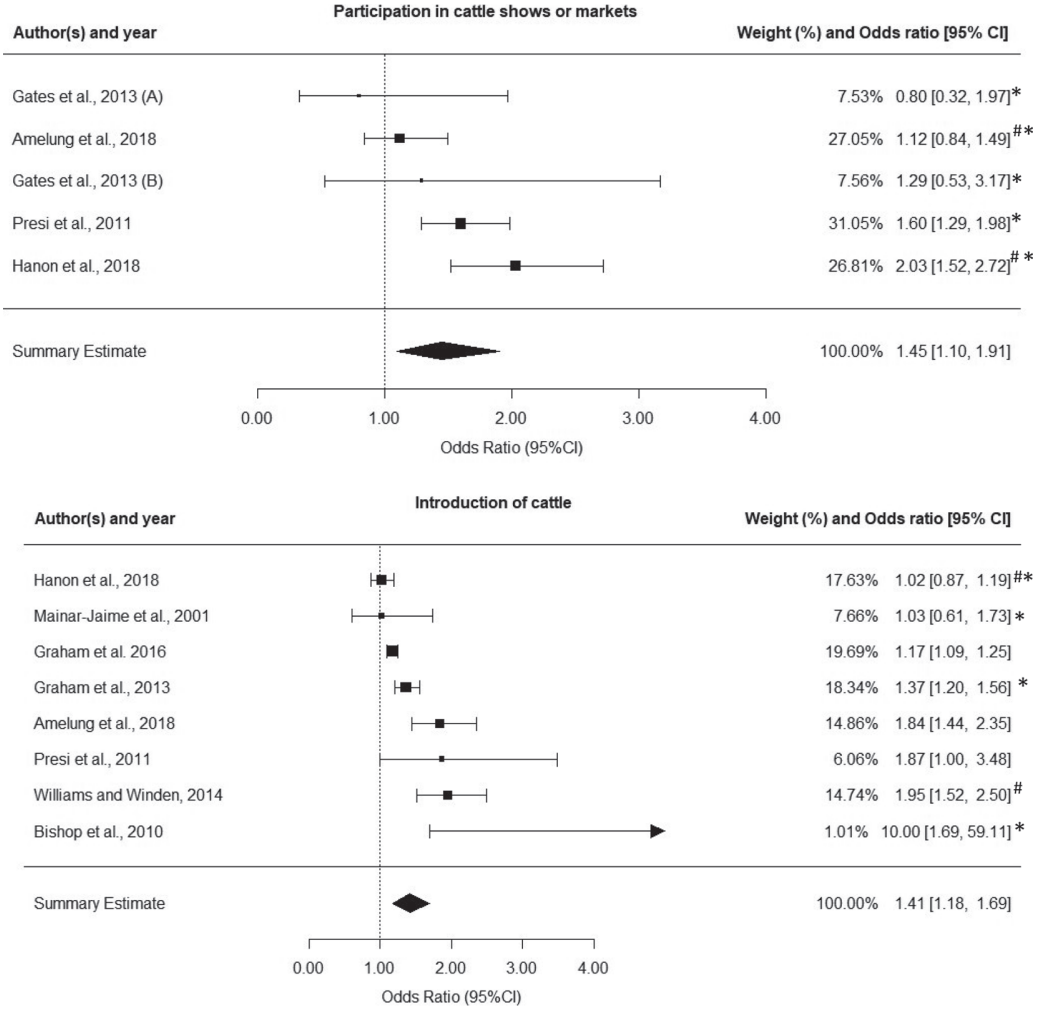



**Figure 3.** Forest plot of the effect of participation in shows or markets (upper plot) and introduction of cattle (lower plot) on bovine viral diarrhea virus infection. Gates et al., 2013 (A) and (B), refers to substudies, as indicated in Table 2. *Univariable result; #study indicated as outlier in the influential case analysis.

or markets compared with herds that did not (Figure 3). The heterogeneity between studies ($I^2$) was 61.70% (95% CI: 0.00–96.60).

Introduction of cattle was the most often studied movement variable but was not easily compared between studies because of the many different ways in which introduction of cattle was coded (i.e., introduction yes−no, source of introduced animals, continuous variables, and introduction of different types of cattle). We decided to focus further meta-analysis on introduction yes−no because these variables were most comparable (Supplemental File S4, section 4.2.2). In 2 studies (Graham et al., 2013, 2016), a sub-meta-analysis was first performed to obtain pooled estimates comparable with the estimates of the yes−no variables (Supplemental File S5, https://doi.org/10.3168/jds.2020-18193). We found a combined effect estimate of 1.41 higher odds (95% CI: 1.18–1.69) of BVDV infection in herds that introduce cattle into the herd compared with herds that do not (Figure 3). The heterogeneity between studies ($I^2$) was 82.98% (95% CI: 71.48–99.47).

***Neighborhood Risk.*** Pasturing of cattle was the most often studied neighborhood risk variable. Variables described whether cattle had access to pasture, the possibility of contact with cattle from other herds at pasture, and shared pasture (Supplemental File S4, section 4.3.1). First studies were compared that looked at the presence versus absence of pasture (Supplemental File S4, section 4.3.2) followed by contact between cattle on pasture (Supplemental File S4, section 4.3.2).

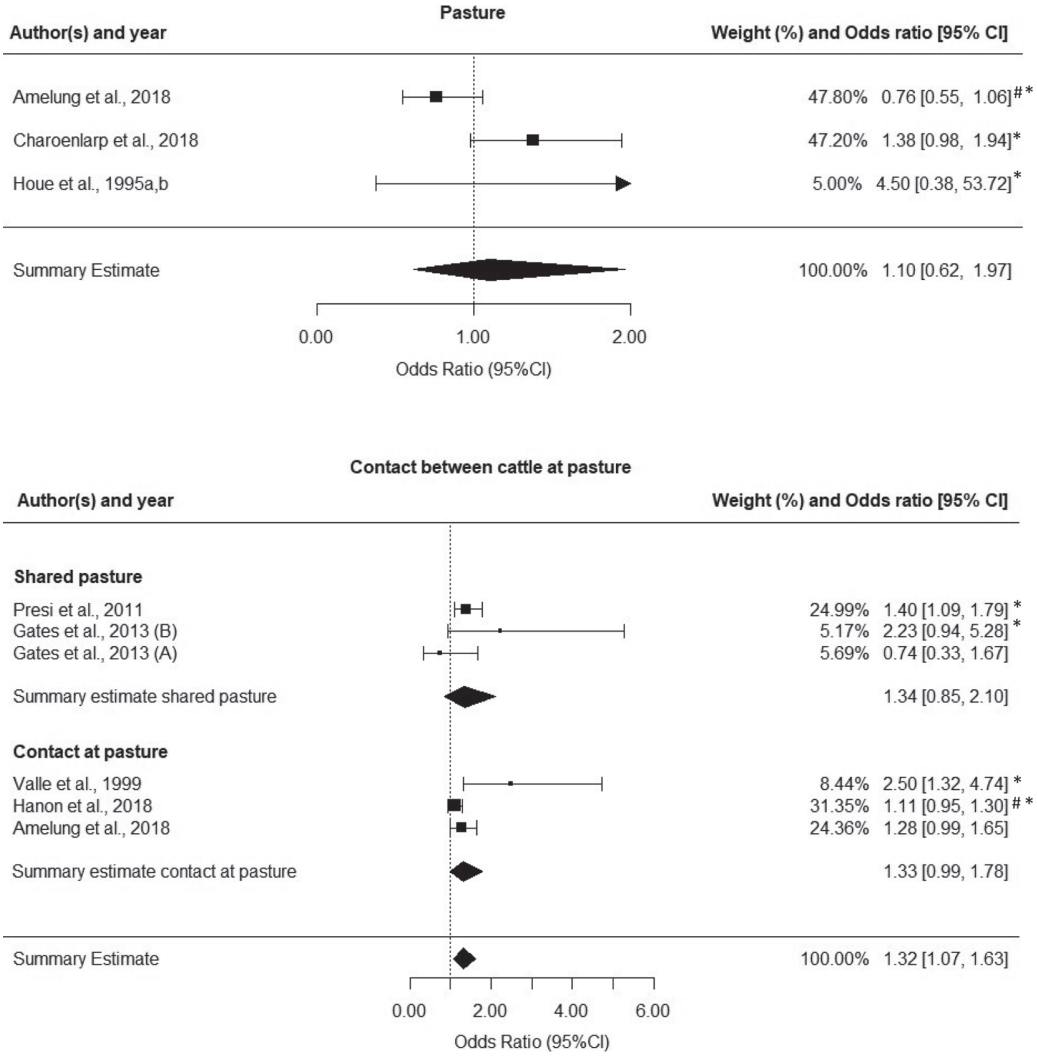



**Figure 4.** Forest plot of the effect of herds grazing (upper plot) and contact between cattle on pasture by either shared pasture or over-the-fence contact (lower plot) on bovine viral diarrhea virus infection. Gates et al., 2013 (A) and (B), refers to substudies, as indicated in Table 2. *Univariable result; #study indicated as outlier in the influential case analysis.

We found a nonsignificant combined effect estimate of 1.10 higher odds (95% CI: 0.62–1.97) of BVDV infection in herds that graze their cattle compared with herds that do not (Figure 4). The heterogeneity between studies ($I^2$) was 73.30% (95% CI: 0.83–99.80). Studies on contact between cattle at pasture were divided into shared pasture and the possibility of contact with cattle from other herds at pasture (e.g., contact over the fence) but were also analyzed together (Figure 4).

For both shared pasture and contact at pasture, we found nonsignificant odds of BVDV infection: 1.34 (95% CI: 0.85–2.10) and 1.33 (95% CI: 0.99–1.78), respectively (Figure 4). However, we found an overall significant combined effect estimate of 1.32 higher odds (95% CI: 1.07–1.63) of BVDV infection in herds where contact between cattle at pasture is possible either because different herds share pasture or because of contact between herds in contiguous pastures (Figure 4). The heterogeneity between studies ($I^2$) was 53.90% (95% CI: 0.00–97.70).

## DISCUSSION

By conducting this systematic literature search we have gained a comprehensive overview of potential risk factors for the presence of BVD in cattle herds. We decided to focus on studies performed in Europe in an attempt to reduce heterogeneity between results caused by different cattle production systems on different continents. However, the results could be generalized to areas outside Europe where there are similar cattle production systems (e.g., areas in the United States). The 18 European publications that were included in this study showed a wide range of potential risk factors that were grouped into 6 categories with similar characteristics: (1) herd and animal characteristics, (2) cattle movement, (3) reproduction, (4) neighborhood risk, (5) farm management and biosecurity, and (6) diagnostic testing and control programs. Although there was a lot of variation in risk factors between studies, we performed several meta-analyses and obtained pooled estimates for several frequently found risk factors.

Two herd characteristics that were frequently studied were herd size and herd type. Most studies found that larger herds were associated with higher odds of BVD infection. Only Hanon et al. (2018) found the highest seroprevalence in the smallest herds (<100 cattle). They did find a higher seroprevalence in farms with a higher number of stables (>3). The pooled estimate in our meta-analysis showed a significantly higher risk of infection per extra 10 animals in the herd (OR = 1.04, 95% CI: 1.02–1.06). This could be explained by the tendency for larger herds to have a decreased probability of self-clearance of infection and to be more likely to contain a higher number of pregnant cattle and purchased cattle, increasing the risk of introduction of PI into the herd (Lindberg and Houe, 2005; Sarrazin et al., 2013; Barrett et al., 2018). In our meta-analysis, dairy herds were also found to be at higher risk of infection than beef herds (OR = 1.63, 95% CI: 1.06–2.50). It has been suggested that this is related to the higher number of contacts between cattle and people and traffic on dairy farms compared with beef farms (Amelung et al., 2018).

Movement of cattle is considered one of the most important risk factors for BVD infection, especially purchase (Courcoul and Ezanno, 2010; Gates et al., 2013; Qi et al., 2019). Our meta-analysis showed higher odds (OR = 1.41, 95% CI: 1.18–1.69) for herds that introduced cattle into the herd in the previous year compared with herds that did not. However, Gates et al. (2014) illustrated that not all purchased cattle pose the same risk. They found that purchase of pregnant heifers and open cows with a calf at foot are associated with a higher risk of BVDV infection in beef herds, with OR of 2.18 (95% CI: 1.17–4.08) and 2.09 (95% CI: 1.13–3.88), respectively. The number of cattle introduced was also studied several times, generally showing increasing odds with increasing numbers of introduced cattle (Gates et al., 2013; Graham et al., 2013, 2016; Byrne at al., 2017). It was, however, suggested that the number of cattle introduced is related to herd size (Graham et al., 2016; Byrne et al., 2017), indicating the importance of correcting for herd size when studying purchase. A different way to study the risk of introduction is to look at the number of source herds. Gates et al. (2013) found a significant association between BVDV infection and a larger number of source herds in dairy herds (OR = 4.42 in units of 10 farms, 95% CI: 1.86–10.00) and beef herds (OR = 10.60 in units of 10 farms, 95% CI: 3.91–31.00). However, there was strong correlation between the number of cattle introduced and the number of source herds (Gates et al., 2013).

Another risk factor related to cattle movement that was studied frequently is participation in shows or markets. Our pooled estimate shows significant higher odds of infection for herds that visit cattle shows or markets (OR = 1.45, 95% CI: 1.10–1.91) compared with herds that do not. This could be explained by the possibility that cattle come in contact with BVDV-infected cattle at the show or market and infect the herd upon returning or because of infection during transport.

No meta-analysis could be performed on any of the reproduction variables because of the low number of comparable studies. However, concerns have been raised about transmission of BVDV by AI (Gard et al.,

2007; Rikula et al., 2008). This may be prevented by regular testing of bulls at AI centers and testing of imported semen (Eaglesome and Garcia, 1997; Wentink et al., 2000; Lindberg et al., 2006). Also, the within-herd calving pattern could not be compared between studies, but Williams and Winden (2014) found an increased likelihood of BVDV presence with year-round calving compared with seasonal calving. They indicated that this could be related to the fact that with year-round calving there are almost always pregnant cows present within the susceptible window for BVDV infection of the fetus. When developing or optimizing BVD control programs, calving pattern could be an important factor to consider. In block calving systems, tissue tag testing of newborn calves provides the opportunity to identify and remove the majority of PI calves before the breeding season commences, reducing the risk of establishing more PI calves to be born the following season. In year-round calving systems, spot testing could be a cost-effective option to monitor new infections (Tratalos et al., 2017).

Bovine viral diarrhea can easily spread between herds direct contact is possible between cattle (Tråvén et al., 1991). Therefore, grazing is considered a risk factor for BVD as nose-to-nose contact between cattle of different herds may occur. However, our pooled estimate did not show significant odds (OR = 1.10, 95% CI: 0.62–1.97) for BVD infection for herds that graze compared with herds that do not. When results that indicated shared pasture were separated from results that indicated whether contact between cattle at pasture could occur (e.g., over-the-fence contact), our pooled estimates were nonsignificant, but when taken altogether and thus increasing statistical power, we found a significant effect indicating that contact between cattle at pasture had a higher odds of BVD infection (OR = 1.32, 95% CI: 1.07–1.63). The risk of grazing is likely influenced by many factors, such as cattle density and the prevalence of BVDV in the area (Houe et al., 1995a), regulations around communal grazing (Rossmanith et al., 2005), the number of cattle and herds sharing pasture (Presi et al., 2011), and the number of neighbors.

In the current study, no meta-analysis was performed on any of the farm management and biosecurity variables due to the low number of studies and the differing ways in which biosecurity was measured. It was unexpected that most studies did not find a significant association between biosecurity measures and BVDV infection because biosecurity is considered an important aspect of BVDV control (Moennig et al., 2005; Lindberg et al., 2006). Gates et al. (2013) suggested that this could be related to the design of questionnaires (e.g., questionnaires that primarily use closed yes–no questions, which forces farmers to choose one of the options even if neither is completely true). Farmers could also give socially desirable answers because they fear possible consequences. Farmer behavior is another factor for which there were not enough quantitative data for meta-analysis. This lack of quantitative data does not necessarily mean that farmer behavior and biosecurity are not important factors for BVD, but they are more often studied qualitatively, which made it impossible to include them in the meta-analysis. Qualitative research into farmer behavior and biosecurity related to BVD stresses the importance of addressing farmer attitudes toward BVD control (Heffernan et al., 2016; Azbel-Jackson et al., 2018). A meta-analysis on epidemiological and mitigation measures that influence production losses in cattle due to BVDV has been reported (Pinior et al., 2019). These authors found that vaccination and biosecurity had a positive influence on the annual BVDV production losses per animal. We agree that farmers' attitudes toward BVD control and biosecurity-related measures are important and influence the effect of the risk factors we found in this paper. For example, when a new cow is kept in quarantine and tested for BVD before its introduction in the herd, the risk of introduction will be lower compared with new cows that are directly introduced in the herd. Therefore, we recommend further study of the quantitative association between BVD control and biosecurity and farmer behavior.

No meta-analysis could be performed on any of the diagnostic testing and control program variables because of both the small number of studies and the large variation between variables. One study found slightly higher odds for presence of BVDV when participating in control programs in univariable analysis (Amelung et al., 2018), which could probably be explained by the assumption that farms with BVDV problems are more likely to participate in a control program. Another interesting result was that herds of farmers younger than 40 yr were more often infected than herds of older farmers (Valle et al., 1999). According to Valle et al. (1999), this is probably due to different attitudes and management practices of younger farmers, such as not asking for health certificates when purchasing animals. This would be an interesting factor to consider in future quantitative studies about BVDV infection and farmer behavior.

In our meta-analyses, several pooled estimates were significant. However, the results could be biased because most studies looked at the presence of BVDV and not introduction of the virus. With presence of infection, it is unknown when the actual infection happened, which complicates finding direct associations between infection and risk factors. However, this would prob-

ably be less influential when considering risk factors that do not change much over time, such as whether herds graze at pasture, herd type, and herd size. When studying the introduction of BVDV, it is possible that there is a delay between introduction and detection. For example, a PI calf introduced on a farm that monitors by bulk milk testing is unlikely to be promptly detected unless individual animal testing is also conducted on newly imported animals on the farm. Such situations complicate efforts to identify direct associations between infection and risk factors. Therefore, we think that the presence of BVDV is a reasonable proxy for introduction of the virus. In addition, the presence of risk factors does not often change as they are part of regular farm management.

Another complicating factor in comparing different studies was the way in which herds were categorized as infected or not infected (e.g., based on antibodies or virus) using different sample types, tests, and strategies to confirm the infection status. These differences could be considered by performing a formal assessment of risk of bias. However, because we already had a low number of studies per meta-analysis, we did not want to exclude any more studies and decided to include only the most important internal and external validity checkpoints (Table 1). Also, not all information was available in each publication for a proper bias risk assessment.

For several risk factors, it was not appropriate to perform a meta-analysis given that there were not enough comparable studies with sufficient quantitative data. For the risk factors with sufficient data, the meta-analyses indicated high levels of heterogeneity. This was expected as all papers included in our meta-analyses were observational studies with different objectives, study designs, and context. For that reason, performing meta-analysis on observational studies and obtaining pooled estimates have been extensively debated (Egger et al., 1998; Blettner et al., 1999; Ioannidis et al., 2008). However, the number of published meta-analyses on observational data has substantially increased, and the need for guidelines for performing meta-analysis on observational data is emphasized (Mueller et al., 2018; Dekkers et al., 2019). In the current study, we decided to perform meta-analyses on observational studies to provide an overview of available quantitative data, including a weighted average estimate. In this subject area, quantitative risk factor information is available only from observational studies. A key principle underpinning this study is the potential for countries without local knowledge of risk factors for BVDV to learn from those countries where data are available. In our view, weighted average estimates have the potential to be more helpful to readers while being cognizant of heterogeneity between studies rather than being solely a listing of all available quantitative results.

In our study, we tried to control for heterogeneity and bias as much as possible through the checklist of study appraisal for quantitative analysis (Table 1) and by very carefully choosing the factors that could be compared. The $I^2$ statistics still showed a very high level of heterogeneity for all factors, but it is known to be not very accurate when only a small number of studies ($n < 20$) are available (Huedo-Medina et al., 2006). Also, the very wide 95% CI of the $I^2$ statistic we observed show the degree of uncertainty about the heterogeneity estimations. The influential case analyses showed that the $I^2$ estimate was often lower when removing outliers from the meta-analyses; however, CI remained wide. Given this result, and because $I^2$ is unreliable when few studies are available, we elected to retain the outliers but to show the summary estimates and $I^2$ of each meta-analysis when excluding the outliers (Supplemental File S6, https://doi.org/10.3168/jds.2020-18193).

To maximize the amount of quantitative data, we decided to include both univariable and multivariable OR in our analyses. Therefore, in 3 of the 6 meta-analyses we combined univariable and multivariable results. The rationale behind this is that in different studies the multivariable OR were adjusted for different factors and referred to different reference situations and are therefore not necessarily more comparable than unadjusted univariable results. On the other hand, univariable OR can under- or overestimate the strength of association. As there is not yet a uniform approach regarding the use of univariable and multivariable results in meta-analysis, adjusted and unadjusted OR often are combined (Liu et al., 2017). As we decided to combine adjusted and unadjusted OR, we have performed subanalyses in which we compared the results when including only the univariable results or the multivariable results. In most cases we observed only minor differences. In the meta-analyses on herd type and introduction of cattle, we did see a substantial decrease in heterogeneity ($I^2$). However, keeping in mind that the $I^2$ statistic becomes increasingly unreliable when even fewer studies are included and because the summary estimates did not change that much, we decided to combine univariable and multivariable results. The results of the subnalyses are reported in Supplemental File S7 (https://doi.org/10.3168/jds.2020-18193). We also selected different observational study designs to maximize the number of studies in our meta-analyses. Therefore, in 2 of the 6 meta-analyses (participation in cattle shows and markets and contact between cattle at pasture) we combined cross-sectional studies with

case-control studies. In the scientific literature, there is disagreement about whether different study designs can be combined (Mueller et al., 2018). The influential case analysis was conducted to determine whether the case-control studies (only 3 out of 20 studies) were indicated as outliers, which they were not. Consequently, leaving them out would not make much difference, and therefore we decided to retain both study designs. We note that these 2 study designs are differing types of observational studies and use OR as outcome.

All studies included in our meta-analyses used OR to show the strength of association between risk factors and BVD infection. It should be kept in mind that these OR are based on a certain reference population and are therefore sensitive to how the reference category is defined. For this reason, it can be questioned whether OR are the right means to compare studies. It would have been better to obtain probabilities of infection and risk factor occurrence. However, given that these were often not reported and the fact that OR do provide a rough risk estimate, it was decided to conduct the meta-analysis on OR. This should be considered when interpreting the results of this study.

## CONCLUSIONS

In this study, we found a wide range of potential risk factors and performed meta-analyses on 6 risk factors for BVDV: herd size, herd type, participation in shows or markets, introduction of cattle, pasture, and contact at pasture. We did not find any unexpected risk factors, and the pooled estimates can help guide advice to farmers and assist in the development, evaluation, and optimization of BVD control programs. The results of the meta-analyses must be interpreted with care due to a high level of study heterogeneity but can assist in the development, evaluation, and optimization of BVD control programs. They can also be used as input for BVDV modeling studies in herds that are comparable with the European cattle production systems. It was challenging to combine estimates of different studies due to heterogeneity between studies (e.g., study design, data analysis, data reporting), showing the need for more standardized methodologies in risk factor studies.

## ACKNOWLEDGMENTS

This work was carried out with the financial support of the Dutch Ministry of Agriculture, Nature and Food Quality (the Hague, the Netherlands) and is part of the STOC free project that was awarded a grant by the European Food Safety Authority (EFSA, Parma, Italy) and was co-financed by public organizations in the countries participating in the study. The authors have not stated any conflicts of interest.

## REFERENCES

AHI (Animal Health Ireland). 2019. Programme results. Accessed Sep. 17, 2019. http://animalhealthireland.ie/?page_id=229.

Almeida, L. L., I. C. S. Miranda, H. E. Hein, W. S. Neto, E. F. Costa, F. S. Marks, C. R. Rodenbusch, C. W. Canal, and L. G. Corbellini. 2013. Herd-level risk factors for bovine viral diarrhea virus infection in dairy herds from Southern Brazil. Res. Vet. Sci. 95:901–907. https://doi.org/10.1016/j.rvsc.2013.08.009.

Amelung, S., M. Hartmann, L. Haas, and L. Kreienbrock. 2018. Factors associated with the bovine viral diarrhoea (BVD) status in cattle herds in Northwest Germany. Vet. Microbiol. 216:212–217. https://doi.org/10.1016/j.vetmic.2018.01.018.

Aragaw, K., B. Sibhat, G. Ayelet, E. Skjerve, E. Z. Gebremedhin, and K. Asmare. 2018. Seroprevalence and factors associated with bovine viral diarrhea virus (BVDV) infection in dairy cattle in three milksheds in Ethiopia. Trop. Anim. Health Prod. 50:1821–1827. https://doi.org/10.1007/s11250-018-1624-5.

Aye, Y. M., M. Aung, W. O. Kyaw, T. Naing, and S. P. Po. 2017. Prevalence and associated factors with bovine viral diarrhoea virus antibodies in the bulk tank milk of small scale dairy herds in central Myanmar. Adv. Anim. Vet. Sci. 5:316–323. https://doi.org/10.17582/journal.aavs/2017/5.8.316.323.

Azbel-Jackson, L., C. Heffernan, G. Gunn, and J. Brownlie. 2018. Exploring the role of voluntary disease schemes on UK farmer bio-security behaviours: Findings from the Norfolk-Suffolk Bovine Viral Diarrhoea control scheme. PLoS One 13:e0179877. https://doi.org/10.1371/journal.pone.0179877.

Barrett, D., M. Parr, J. Fagan, A. Johnson, J. Tratalos, F. Lively, M. Diskin, and D. Kenny. 2018. Prevalence of bovine viral diarrhoea virus (BVDV), bovine herpes virus 1 (BHV 1), leptospirosis and neosporosis, and associated risk factors in 161 Irish beef herds. BMC Vet. Res. 14:8. https://doi.org/10.1186/s12917-017-1324-9.

Bedeković, T., N. Lemo, L. Barbić, Ž. Cvetnić, I. Lojkić, M. Benić, Ž. Čač, M. Lojkić, and J. Madić. 2013. Influence of category, herd size, grazing and management on epidemiology of bovine viral diarrhoea in dairy herds. Acta Vet. Brno 82:125–130. https://doi.org/10.2754/avb201382020125.

Bishop, H., J. Erkelens, and S. Van Winden. 2010. Indications of a relationship between buying-in policy and infectious diseases on dairy farms in Wales. Vet. Rec. 167:644–647. https://doi.org/10.1136/vr.c5256.

Bitsch, V., K. E. Hansen, and L. Rønsholt. 2000. Experiences from the Danish programme for eradication of bovine virus diarrhoea (BVD) 1994–1998 with special reference to legislation and causes of infection. Vet. Microbiol. 77:137–143. https://doi.org/10.1016/s0378-1135(00)00270-4.

Blettner, M., W. Sauerbrei, B. Schlehofer, T. Scheuchenpflug, and C. Friedenreich. 1999. Traditional reviews, meta-analyses and pooled analyses in epidemiology. Int. J. Epidemiol. 28:1–9. https://doi.org/10.1093/ije/28.1.1.

Borenstein, M., J. P. Higgins, L. V. Hedges, and H. R. Rothstein. 2017. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Res. Synth. Methods 8:5–18. https://doi.org/10.1002/jrsm.1230.

Buitrago Horta, E. R., C. Jiménez Escobar, and J. L. Zambrano Varón. 2018. Identification of factors associated with exposure to bovine viral diarrhea virus (BVDV) in dairy-herd calves in the Bogotá savanna. Rev. Med. Vet. 36:63–73. https://doi.org/10.19052/mv.5172.

Byrne, A. W., M. Guelbenzu-Gonzalo, S. A. J. Strain, S. McBride, J. Graham, A. Lahuerta-Marin, R. Harwood, D. A. Graham, and S. McDowell. 2017. Assessment of concurrent infection with bovine viral diarrhoea virus (BVDV) and *Mycobacterium bovis*: A herd-level risk factor analysis from Northern Ireland. Prev. Vet. Med. 141:38–47. https://doi.org/10.1016/j.prevetmed.2017.04.007.

Carbonero, A., A. Maldonado, A. Perea, I. García-Bocanegra, C. Borge, A. Torralbo, A. Arenas-Montes, and A. Arenas-Casas. 2011. Risk factors against bovine respiratory disease in suckling calves from Argentina. Arch. Zootec. 60:41–51. https://doi.org/10.21071/az.v60i229.4687.

Charoenlarp, W., K. Frankena, S. A. J. Strain, M. Guelbenzu-Gonzalo, J. Graham, and A. W. Byrne. 2018. Spatial and risk factor analysis of bovine viral diarrhoea (BVD) virus after the first-year compulsory phase of BVD eradication programme in Northern Ireland. Prev. Vet. Med. 157:34–43. https://doi.org/10.1016/j.prevetmed.2018.05.011.

Chi, J., J. A. Van Leeuwen, A. Weersink, and G. P. Keefe. 2002. Management factors related to seroprevalences to bovine viral-diarrhoea virus, bovine-leukosis virus, *Mycobacterium avium* subspecies *paratuberculosis*, and *Neospora caninum* in dairy herds in the Canadian Maritimes. Prev. Vet. Med. 55:57–68. https://doi.org/10.1016/s0167-5877(02)00067-3.

Courcoul, A., and P. Ezanno. 2010. Modelling the spread of bovine viral diarrhoea virus (BVDV) in a managed metapopulation of cattle herds. Vet. Microbiol. 142:119–128. https://doi.org/10.1016/j.vetmic.2009.09.052.

Cuttance, W. G., and E. L. Cuttance. 2014. Analysis of individual farm investigations into bovine viral diarrhoea in beef herds in the North Island of New Zealand. N. Z. Vet. J. 62:338–342. https://doi.org/10.1080/00480169.2014.928925.

Dekkers, O. M., J. P. Vandenbroucke, M. Cevallos, A. G. Renehan, D. G. Altman, and M. Egger. 2019. COSMOS-E: Guidance on conducting systematic reviews and meta-analyses of observational studies of etiology. PLoS Med. 16:e1002742. https://doi.org/10.1371/journal.pmed.1002742.

Downes, M. J., M. L. Brennan, H. C. Williams, and R. S. Dean. 2016. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open 6:e011458. https://doi.org/10.1136/bmjopen-2016-011458.

Eaglesome, M. D., and M. M. Garcia. 1997. Disease risks to animal health from artificial insemination with bovine semen. Rev. Sci. Tech. Off. Int. Epiz. 16:215–225. https://doi.org/10.20506/rst.16.1.1017.

Egger, M., M. Schneider, and G. D. Smith. 1998. Spurious precision? Meta-analysis of observational studies. BMJ 316:140–144. https://doi.org/10.1136/bmj.316.7125.140.

Erfani, A. M., M. Bakhshesh, M. H. Fallah, and M. Hashemi. 2019. Seroprevalence and risk factors associated with bovine viral diarrhea virus and bovine herpes virus-1 in Zanjan Province, Iran. Trop. Anim. Health Prod. 51:313–319. https://doi.org/10.1007/s11250-018-1687-3.

Ersbøll, A. K., B. K. Ersbøll, H. Houe, L. Alban, and A. M. Kjeldsen. 2010. Spatial modelling of the between-herd infection dynamics of bovine virus diarrhoea virus (BVDV) in dairy herds in Denmark. Prev. Vet. Med. 97:83–89. https://doi.org/10.1016/j.prevetmed.2010.08.004.

Fernandes, L. G., A. H. de Campos Nogueira, E. De Stefano, E. M. Pituco, C. P. Ribeiro, C. J. Alves, T. S. Oliveira, I. J. Clementino, and S. S. de Azevedo. 2016. Herd-level prevalence and risk factors for bovine viral diarrhea virus infection in cattle in the State of Paraíba, Northeastern Brazil. Trop. Anim. Health Prod. 48:157–165. https://doi.org/10.1007/s11250-015-0937-x.

Foddai, A., A. Stockmarr, and A. Boklund. 2016. Evaluation of temporal surveillance system sensitivity and freedom from bovine viral diarrhea in Danish dairy herds using scenario tree modelling. BMC Vet. Res. 12:118. https://doi.org/10.1186/s12917-016-0744-2.

Gädicke, P., T. Junod, J. López-Martin, R. Ortega, and G. Monti. 2016. Enfermedades abortigénicas en lecherías de la Provincia de Nuble: Prevalencia y análisis espacial. Arch. Med. Vet. 48:18–26. https://doi.org/10.4067/S0301-732X2016000100003.

Gard, J. A., M. D. Givens, and D. A. Stringfellow. 2007. Bovine viral diarrhea virus (BVDV): Epidemiologic concerns relative to semen and embryos. Theriogenology 68:434–442. https://doi.org/10.1016/j.theriogenology.2007.04.003.

Gates, M. C., R. W. Humphry, G. J. Gunn, and M. E. Woolhouse. 2014. Not all cows are epidemiologically equal: Quantifying the risks of bovine viral diarrhoea virus (BVDV) transmission through cattle movements. Vet. Res. 45:110. https://doi.org/10.1186/s13567-014-0110-y.

Gates, M. C., M. E. J. Woolhouse, G. J. Gunn, and R. W. Humphry. 2013. Relative associations of cattle movements, local spread, and biosecurity with bovine viral diarrhoea virus (BVDV) seropositivity in beef and dairy herds. Prev. Vet. Med. 112:285–295. https://doi.org/10.1016/j.prevetmed.2013.07.017.

Graham, D. A., T. A. Clegg, M. Lynch, and S. J. More. 2013. Herd-level factors associated with the presence of bovine viral diarrhoea virus in herds participating in the voluntary phase of the Irish national eradication programme. Prev. Vet. Med. 112:99–108. https://doi.org/10.1016/j.prevetmed.2013.07.011.

Graham, D. A., T. A. Clegg, H. H. Thulke, P. O'sullivan, G. McGrath, and S. J. More. 2016. Quantifying the risk of spread of bovine viral diarrhoea virus (BVDV) between contiguous herds in Ireland. Prev. Vet. Med. 126:30–38. https://doi.org/10.1016/j.prevetmed.2016.01.017.

Hanon, J. B., M. De Baere, C. de la Ferté, S. Roelandt, G. Guillot, Y. Van der Stede, and B. Cay. 2018. Serological monitoring on milk and serum samples in a BVD eradication program: A field study in Belgium showing antibody ELISA performances and epidemiological aspects. Prev. Vet. Med. 160:136–144. https://doi.org/10.1016/j.prevetmed.2018.07.008.

Harrer, M., P. Cuijpers, T. A. Furukawa, and D. D. Ebert. 2019. Doing Meta-Analysis in R: A Hands-on Guide. Accessed Nov. 20, 2019. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/.

Heffernan, C., L. Azbel-Jackson, J. Brownlie, and G. Gunn. 2016. Farmer attitudes and livestock disease: Exploring citizenship behaviour and peer monitoring across two BVD control schemes in the UK. PLoS One 11:e0152295. https://doi.org/10.1371/journal.pone.0152295.

Herrera-Yunga, V., J. Labada, F. Castillo, A. Torres, G. Escudero-Sanchez, M. Capa-Morocho, and R. Abad-Guaman. 2018. Prevalence of antibodies and risk factors to bovine viral diarrhea in non-vaccinated dairy cattle from southern Ecuador. Trop. Subtrop. Agroecosystems 21:11–18.

Higgins, J. P. T., J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, and V. A. Welch. 2019. Cochrane Handbook for Systematic Reviews of Interventions. Version 6.0. Accessed July 7, 2019. www.training.cochrane.org/handbook.

Houe, H. 2003. Economic impact of BVDV infection in dairies. Biologicals 31:137–143. https://doi.org/10.1016/s1045-1056(03)00030-7.

Houe, H., J. C. Baker, R. K. Maes, J. W. Lloyd, and C. Enevoldsen. 1995a. Comparison of the prevalence and incidence of infection with bovine virus diarrhoea virus (BVDV) in Denmark and Michigan and association with possible risk factors. Acta Vet. Scand. 36:521–531.

Houe, H., J. C. Baker, R. K. Maes, P. L. Ruegg, and J. W. Lloyd. 1995b. Application of antibody titers against bovine viral diarrhea virus (BVDV) as a measure to detect herds with cattle persistently infected with BVDV. J. Vet. Diagn. Invest. 7:327–332. https://doi.org/10.1177/104063879500700305.

Huedo-Medina, T. B., J. Sánchez-Meca, F. Marín-Martínez, and J. Botella. 2006. Assessing heterogeneity in meta-analysis: Q statistic or $I^2$ index? Psychol. Methods 11:193–206. https://doi.org/10.1037/1082-989X.11.2.193.

Hult, L., and A. Lindberg. 2005. Experiences from BVDV control in Sweden. Prev. Vet. Med. 72:143–148. https://doi.org/10.1016/j.prevetmed.2005.04.005.

Humphry, R. W., F. Brülisauer, I. J. McKendrick, P. F. Nettleton, and G. J. Gunn. 2012. Prevalence of antibodies to bovine viral diarrhoea virus in bulk tank milk and associated risk factors in Scottish dairy herds. Vet. Rec. 171:445. https://doi.org/10.1136/vr.100542.

Ioannidis, J. P., N. A. Patsopoulos, and H. R. Rothstein. 2008. Reasons or excuses for avoiding meta-analysis in forest plots. BMJ 336:1413–1415. https://doi.org/10.1136/bmj.a117.

Kadohira, M., and M. Tajima. 2010. A case control study of bovine viral diarrhea virus (BVDV) persistent infection (PI) in Betsukai,

Hokkaido, Japan. J. Vet. Med. Sci. 72:635–638. https://doi.org/10.1292/jvms.09-0305.
Kumar, S. K., K. M. Palanivel, K. Sukumar, B. S. M. Ronald, G. Selvaraju, and G. Ponnudurai. 2018. Herd-level risk factors for bovine viral diarrhea infection in cattle of Tamil Nadu. Trop. Anim. Health Prod. 50:793–799. https://doi.org/10.1007/s11250-017-1497-z.
Lindberg, A., J. Brownlie, G. Gunn, H. Houe, V. Moening, H. W. Saatkamp, T. Sandvik, and P. S. Valle. 2006. The control of bovine viral diarrhoea virus in Europe: Today and in the future. Rev. Sci. Tech. Off. Int. Epiz. 25:961–979.
Lindberg, A., and H. Houe. 2005. Characteristics in the epidemiology of bovine viral diarrhea virus (BVDV) of relevance to control. Prev. Vet. Med. 72:55–73. https://doi.org/10.1016/j.prevetmed.2005.07.018.
Liu, T., X. Nie, Z. Wu, Y. Zhang, G. Feng, S. Cai, Y. Lv, and X. Peng. 2017. Can statistic adjustment of OR minimize the potential confounding bias for meta-analysis of case-control study? A secondary data analysis. BMC Med. Res. Methodol. 17:179. https://doi.org/10.1186/s12874-017-0454-x.
Luzzago, C., M. Frigerio, R. Piccinini, V. Daprà, and A. Zecconi. 2008. A scoring system for risk assessment of the introduction and spread of bovine viral diarrhoea virus in dairy herds in Northern Italy. Vet. J. 177:236–241. https://doi.org/10.1016/j.tvjl.2007.04.017.
Machado, G., R. M. F. Egocheaga, H. E. Hein, I. C. S. Miranda, W. S. Neto, L. L. Almeida, C. W. Canal, M. C. Stein, and L. G. Corbellini. 2016. Bovine Viral Diarrhoea Virus (BVDV) in dairy cattle: A matched case–control study. Transbound. Emerg. Dis. 63:e1–e13. https://doi.org/10.1111/tbed.12219.
Mainar-Jaime, R. C., B. Berzal-Herranz, P. Arias, and F. A. Rojo-Vázquez. 2001. Epidemiological pattern and risk factors associated with bovine viral-diarrhoea virus (BVDV) infection in a non-vaccinated dairy-cattle population from the Asturias region of Spain. Prev. Vet. Med. 52:63–73. https://doi.org/10.1016/s0167-5877(01)00239-2.
Martinez-Ibeas, A. M., C. Power, J. McClure, and R. G. Sayers. 2015. Prevalence of BoHV-1 seropositive and BVD virus positive bulls on Irish dairy farms and associations between bull purchase and herd status. Ir. Vet. J. 68:28. https://doi.org/10.1186/s13620-015-0059-9.
McClurkin, A. W., E. T. Littledike, R. C. Cutlip, G. H. Frank, M. F. Coria, and S. R. Bolin. 1984. Production of cattle immunotolerant to bovine viral diarrhea virus. Can. J. Comp. Med. 48:156.
Milián-Suazo, F., R. Hernández-Ortíz, L. Hernández-Andrade, A. Alvarado-Islas, E. Díaz-Aparicio, F. Mejía-Estrada, E. G. Palomares-Reséndiz, I. Bárcenas Reyes, and H. Zendejas-Martínez. 2016. Seroprevalence and risk factors for reproductive diseases in dairy cattle in Mexico. J. Vet. Med. Anim. Health 8:89–98. https://doi.org/10.5897/JVMAH2016.0483.
Mockeliūniene, V., A. Šalomskas, R. Mockeliūnas, and S. Petkevičius. 2004. Prevalence and epidemiological features of bovine viral diarrhoea virus infection in Lithuania. Vet. Microbiol. 99:51–57. https://doi.org/10.1016/j.vetmic.2003.11.008.
Moennig, V., H. Houe, and A. Lindberg. 2005. BVD control in Europe: Current status and perspectives. Anim. Health Res. Rev. 6:63–74. https://doi.org/10.1079/ahr2005102.
Moher, D., A. Liberati, J. Tetzlaff, and D. G. Altman. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. PLoS Med. 6:e1000097. https://doi.org/10.1371/journal.pmed.1000097.
Mueller, M., M. D'Addario, M. Egger, M. Cevallos, O. Dekkers, C. Mugglin, and P. Scott. 2018. Methods to systematically review and meta-analyse observational studies: A systematic scoping review of recommendations. BMC Med. Res. Methodol. 18:44. https://doi.org/10.1186/s12874-018-0495-9.
National Institutes of Health. 2014. Study quality assessment tools. Accessed Sep. 17, 2019. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools.
Nava, L., P. Bracamonte, D. Hidalgo, and L. D. G. Escobar. 2013. Seroprevalence of bovine viral diarrhea in milk producing herds at Barinas State, Venezuela. Rev. Soc. Venez. Microbiol. 33:162–168.
Niskanen, R., and A. Lindberg. 2003. Transmission of bovine viral diarrhoea virus by unhygienic vaccination procedures, ambient air, and from contaminated pens. Vet. J. 165:125–130. https://doi.org/10.1016/s1090-0233(02)00161-2.
Norström, M., M. E. Jonsson, J. Åkerstedt, A. C. Whist, A. B. Kristoffersen, S. Sviland, P. Hopp, and H. Wahlström. 2014. Estimation of the probability of freedom from bovine virus diarrhoea virus in Norway using scenario tree modelling. Prev. Vet. Med. 116:37–46. https://doi.org/10.1016/j.prevetmed.2014.06.012.
Nuotio, L., M. Juvonen, E. Neuvonen, L. Sihvonen, and J. Husu-Kallio. 1999. Prevalence and geographic distribution of bovine viral diarrhoea (BVD) infection in Finland 1993–1997. Vet. Microbiol. 64:231–235. https://doi.org/10.1016/s0378-1135(98)00272-7.
Olafson, P., and C. G. Rickard. 1947. Further observations on the virus diarrhea (new transmissible disease) of cattle. Cornell Vet. 37:104–106.
Pinior, B., S. Garcia, J. J. Minviel, and D. Raboisson. 2019. Epidemiological factors and mitigation measures influencing production losses in cattle due to bovine viral diarrhoea virus infection: A meta-analysis. Transbound. Emerg. Dis. 66:2426–2439. https://doi.org/10.1111/tbed.13300.
Presi, P., R. Struchen, T. Knight-Jones, S. Scholl, and D. Heim. 2011. Bovine viral diarrhea (BVD) eradication in Switzerland—Experiences of the first two years. Prev. Vet. Med. 99:112–121. https://doi.org/10.1016/j.prevetmed.2011.01.012.
Qi, L., G. Beaunée, S. Arnoux, B. L. Dutta, A. Joly, E. Vergu, and P. Ezanno. 2019. Neighbourhood contacts and trade movements drive the regional spread of bovine viral diarrhoea virus (BVDV). Vet. Res. 50:30. https://doi.org/10.1186/s13567-019-0647-x.
Quevedo, D. C., B. B. Benavides, G. Cárdenas, and C. Herrera. 2011. Seroprevalence and risk factors associated to BHV-1 and DVBV in dairy herds in Pasto, Colombia, in 2011. Rev. Lasallista Investig. 8:61–68.
R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Ramírez Vásquez, N. F., D. Villar Argaiz, J. A. Fernández Silva, J. Londoño Pino, J. J. Chaparro Gutiérrez, and M. E. Olivera Ángel. 2016. Seroprevalence and risk factors of several bovine viral diseases in dairy farms of San Pedro de los Milagros, Antioquia, Colombia. CES Med. Vet. Zootec. 11:15–25. https://doi.org/10.21615/cesmvz.11.1.2.
Rêgo, M. J. P., A. F. B. Batista Filho, P. R. F. de Oliveira, J. de Melo Borges, C. A. B. de França, C. P. Ribeiro, E. M. Pituco, and J. W. Pinheiro Junior. 2016. Epidemiological analysis of infection by the bovine viral diarrhea virus on family farms in Brazil. Semin. Cienc. Agrar. 37:4119–4129. https://doi.org/10.5433/1679-0359.2016v37n6p4119.
Rikula, U., L. Nuotio, T. Aaltonen, and O. Ruoho. 2005. Bovine viral diarrhoea virus control in Finland 1998–2004. Prev. Vet. Med. 72:139–142. https://doi.org/10.1016/j.prevetmed.2005.08.010.
Rikula, U., L. Nuotio, U. I. Laamanen, and L. Sihvonen. 2008. Transmission of bovine viral diarrhoea virus through the semen of acutely infected bulls under field conditions. Vet. Rec. 162:79–82. https://doi.org/10.1136/vr.162.3.79.
Rossmanith, W., M. Deinhofer, R. Janacek, R. Trampler, and E. Wilhelm. 2010. Voluntary and compulsory eradication of bovine viral diarrhoea virus in Lower Austria. Vet. Microbiol. 142:143–149. https://doi.org/10.1016/j.vetmic.2009.09.055.
Rossmanith, W., R. Janacek, and E. Wilhelm. 2005. Control of BVDV-infection on common grassland—The key for successful BVDV-eradication in Lower Austria. Prev. Vet. Med. 72:133–137. https://doi.org/10.1016/j.prevetmed.2005.05.012.
Saa, L. R., A. Perea, I. García-Bocanegra, A. J. Arenas, D. V. Jara, R. Ramos, and A. Carbonero. 2012. Seroprevalence and risk factors associated with bovine viral diarrhea virus (BVDV) infection in non-vaccinated dairy and dual purpose cattle herds in Ecuador.

Trop. Anim. Health Prod. 44:645–649. https://doi.org/10.1007/s11250-011-9948-4.
Sanderson, S., I. D. Tatt, and J. Higgins. 2007. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: A systematic review and annotated bibliography. Int. J. Epidemiol. 36:666–676. https://doi.org/10.1093/ije/dym018.
Sarrazin, S., A. Veldhuis, E. Méroc, I. Vangeel, J. Laureyns, J. Dewulf, A. B. Caij, S. Piepers, J. Hooyberghs, S. Ribbens, and Y. Van Der Stede. 2013. Serological and virological BVDV prevalence and risk factor analysis for herds to be BVDV seropositive in Belgian cattle herds. Prev. Vet. Med. 108:28–37. https://doi.org/10.1016/j.prevetmed.2012.07.005.
Scharnböck, B., F.-F. Roch, V. Richter, C. Funke, C. L. Firth, W. Obritzhauser, W. Baumgartner, A. Käsbohrer, and B. Pinior. 2018. A meta-analysis of bovine viral diarrhoea virus (BVDV) prevalences in the global cattle population. Sci. Rep. 8:14420. https://doi.org/10.1038/s41598-018-32831-2.
Scott, H. M., O. Sorensen, J. T. Wu, E. Y. Chow, K. Manninen, and J. A. VanLeeuwen. 2006. Seroprevalence of *Mycobacterium avium* subspecies *paratuberculosis*, *Neospora caninum*, Bovine leukemia virus, and Bovine viral diarrhea virus infection among dairy cattle and herds in Alberta and agroecological risk factors associated with seropositivity. Can. Vet. J. 47:981.
Segura-Correa, J. C., J. L. Solorio-Rivera, and L. G. Sánchez-Gil. 2010. Seroconversion to bovine viral diarrhoea virus and infectious bovine rhinotracheitis virus in dairy herds of Michoacan, Mexico. Trop. Anim. Health Prod. 42:233–238. https://doi.org/10.1007/s11250-009-9411-y.
Segura-Correa, J. C., C. C. Zapata-Campos, J. O. Jasso-Obregón, J. Martinez-Burnes, and R. López-Zavala. 2016. Seroprevalence and risk factors associated with bovine herpesvirus 1 and bovine viral diarrhea virus in North-Eastern Mexico. Open Vet. J. 6:143–149. https://doi.org/10.4314/ovj.v6i2.12.
Selim, A. M., M. M. Elhaig, S. A. Moawed, and E. El-Nahas. 2018. Modeling the potential risk factors of bovine viral diarrhea prevalence in Egypt using univariable and multivariable logistic regression analyses. Vet. World 11:259–267. https://doi.org/10.14202/vetworld.2018.259-267.
Solis-Calderon, J. J., V. M. Segura-Correa, and J. C. Segura-Correa. 2005. Bovine viral diarrhoea virus in beef cattle herds of Yucatan, Mexico: Seroprevalence and risk factors. Prev. Vet. Med. 72:253–262. https://doi.org/10.1016/j.prevetmed.2005.06.004.
Sun, W. W., Q. F. Meng, W. Cong, X. F. Shan, C. F. Wang, and A. D. Qian. 2015. Herd-level prevalence and associated risk factors for *Toxoplasma gondii*, *Neospora caninum*, *Chlamydia abortus* and bovine viral diarrhoea virus in commercial dairy and beef cattle in eastern, northern and northeastern China. Parasitol. Res. 114:4211–4218. https://doi.org/10.1007/s00436-015-4655-0.
Talafha, A. Q., S. M. Hirche, M. M. Ababneh, A. M. Al-Majali, and M. M. Ababneh. 2009. Prevalence and risk factors associated with bovine viral diarrhea virus infection in dairy herds in Jordan. Trop. Anim. Health Prod. 41:499–506. https://doi.org/10.1007/s11250-008-9214-6.
Taylor, L. F., P. F. Black, D. J. Pitt, A. R. Mackenzie, S. J. Johnson, and B. J. Rodwell. 2006. A seroepidemiological study of bovine pestivirus in Queensland beef and dairy herds conducted in 1994/95. Aust. Vet. J. 84:163–168. https://doi.org/10.1111/j.1751-0813.2006.tb12771.x.
Tratalos, J. A., D. A. Graham, and S. J. More. 2017. Patterns of calving and young stock movement in Ireland and their implications for BVD serosurveillance. Prev. Vet. Med. 142:30–38. https://doi.org/10.1016/j.prevetmed.2017.04.005.
Tråvén, M., S. Alenius, C. Fossum, and B. Larsson. 1991. Primary bovine viral diarrhoea virus infection in calves following direct contact with a persistently viraemic calf. J. Vet. Med. Series B 38:453–462. https://doi.org/10.1111/j.1439-0450.1991.tb00895.x.
Valle, P. S., S. W. Martin, R. Tremblay, and K. Bateman. 1999. Factors associated with being a bovine-virus diarrhoea (BVD) seropositive dairy herd in the Møre and Romsdal County of Norway. Prev. Vet. Med. 40:165–177. https://doi.org/10.1016/s0167-5877(99)00030-6.
Viechtbauer, W. 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36:1–48. https://doi.org/10.18637/jss.v036.i03.
Viechtbauer, W., and M. W. L. Cheung. 2010. Outlier and influence diagnostics for meta-analysis. Res. Synth. Methods 1:112–125. https://doi.org/10.1002/jrsm.11.
Wentink, G. H., K. Frankena, J. C. Bosch, J. E. D. Vandehoek, and T. Van Den Berg. 2000. Prevention of disease transmission by semen in cattle. Livest. Prod. Sci. 62:207–220. https://doi.org/10.1016/S0301-6226(99)00158-X.
Williams, D., and S. V. Winden. 2014. Risk factors associated with high bulk milk antibody levels to common pathogens in UK dairies. Vet. Rec. 174:580. https://doi.org/10.1136/vr.102049.
Wittum, T. E., D. M. Grotelueschen, K. V. Brock, W. G. Kvasnicka, J. G. Floyd, C. L. Kelling, and K. G. Odde. 2001. Persistent bovine viral diarrhoea virus infection in US beef herds. Prev. Vet. Med. 49:83–94. https://doi.org/10.1016/s0167-5877(01)00181-7.

# Annexe 4 : STOC free : WP1, Deliverable 2.3. Description of the STOC free model, 04/2019.

A. Madouasse, M. Mercat, A. van Roon, I. Santman, M. Nielen, G. van Schaik, S. More, D.Graham, M. Guelbenzu, J. Frössling, E. Ågren, A. Lindberg, J. Gethmann, C. Sauter-Louis, G.Gunn, C. Gomes, C. Fourichon.

# STOC free: WP1, Deliverable 2.3

## Description of the STOC free model

**A. Madouasse[1], M. Mercat[1], A. van Roon[2], I. Santman[2,3], M. Nielen[2], G. van Schaik[2,3], S. More[4], D. Graham[5], M. Guelbenzu[5], J. Frössling[6], E. Ågren[6], A. Lindberg[6], J. Gethmann[7], C. Sauter-Louis[7], G. Gunn[8], C. Gomes[8], C. Fourichon[1]**

[1] BIOEPAR, INRA, Oniris, France, [2]Utrecht University, the Netherlands, [3]GD Animal Health, the Netherlands, [4]University College Dublin, Ireland, [5]Animal Health Ireland, [6]Swedish National Veterinary Institute, Sweden, [7]FLI, Germany, [8]SRUC, Scotland

April 2019

GA/EFSA/AFSCO/2016/01-03

**This study was awarded a grant by EFSA and was co-financed by public organisations in the countries participating in the study.**

**Abstract**

The aim of this document is to describe a statistical framework called the STOC free model for the estimation of a probability of infection that can incorporate the information available from different control programmes against cattle diseases. These control programmes rely on a surveillance programme for the definition of a status regarding infection (e.g. herd free from infection). It is assumed that, as part of these surveillance programmes, tests are carried out at regular time intervals. These tests are imperfect and defined by a sensitivity and a specificity. Knowledge about disease dynamics and risk factors of infection can also bring information for the estimation. Two situations are considered. When data from control programmes are available, statuses regarding infection as well as test characteristics, disease dynamics parameters and strengths of association between risk factors of infection can be estimated. A Bayesian model allowing to perform these estimations is described. When parameter estimation is not possible or wanted, statuses can be predicted assuming known values or distributions for all parameters. In the document, modelling hypotheses and a statistical model are described. Computer code to run the model is provided. The document ends with a description of the further areas that will be investigated on the modelling part of the STOC free project.

# Contents

# 1 About this file

This file mixes text and R code. It was generated using R and LaTeXfrom within RStudio[1]. The R code appears within grey boxes as this one:

```
print("This is R code")

## [1] "This is R code"
```

In order to generate this file from source, the corresponding `.Rnw` must be compiled with RStudio. The R code can be extracted from the `.Rnw` file with the `purl()` function from the `knitr` package.

The required LaTeXpackages are listed at the top of the `.Rnw` file.

In order to perform Bayesian inference, the following programmes need to be installed:

- `JAGS`: performs Bayesian inference using Markov Chain Monte Carlo methods such as Gibbs sampling[2].
- `Stan`: performs Bayesian inference using Hamiltonian Monte Carlo[3]. It is newer than JAGS and should perform better in most situations. However, the code may be harder to read.

The following R packages need to be installed.

- `rjags`: allows calling JAGS from R
- `rstan`: allows calling Stan from R

At various points in the document, we need to estimate the `a` and `b` parameters of a Beta distribution from its mean and variance. The code below allows doing this.

```
estBetaParams <- function(mu, var) {
  alpha <- ((1 - mu) / var - 1 / mu) * mu ^ 2
  beta <- alpha * (1 / mu - 1)
  return(params = list(alpha = alpha, beta = beta))
}
estBetaParams(.2, .001)

## $alpha
## [1] 31.8
##
## $beta
## [1] 127.2
```

[1] See RStudio website.
[2] See instructions for downloading and installing JAGS here
[3] See instructions for downloading and installing Stan here

In order to get the parameters of a distribution with a mean of 0.7 and a variance of 0.01, enter the following code:

```
estBetaParams(0.7, 0.01)

## $alpha
## [1] 14
##
## $beta
## [1] 6
```

The shape of the distribution can be checked by typing:

```
curve(dbeta(x, 14, 6))
```

Finally, in order to allow many people to work in parallel on the STOC free model, we use a version control software called `git`[4]. Our Git repository is available online on a website called SourceSup. It can be cloned onto your computer by entering in a command line interface:

```
git clone https://git.renater.fr/stocfree-model.git/
```

[4]See the Git Wikipedia page for more information: https://en.wikipedia.org/wiki/Git

# 2 Introduction

## 2.1 Context and objectives

Control programmes against non-regulated cattle diseases usually rely on a surveillance component in order to detect infected animals or herds. For the same disease, the surveillance programmes implemented in different territories are usually different. As a consequence, what is considered as a herd or an animal free from infection is different and difficult to compare between territories. There is a need for output based surveillance methods that allow to estimate probabilities of freedom from infection that are comparable regardless of the surveillance programme implemented (Cameron, 2012).

The aim of this document is to describe a statistical framework for the estimation of a probability of infection that can incorporate all the available information generated by different control programmes for the estimation, including context, information on the control programme, test results and risk factors. We focus here on surveillance programmes that operate at the herd level, but the concepts presented would apply to any level of investigation. We will therefore refer to the herd level as a unit.

From the surveillance activities carried out, the objective is to estimate a probability of infection as well as the uncertainty associated with this probability. From this probability of infection it is possible to calculate a probability of freedom from infection, which is 1 - probability of infection. However, this would obscure the description of the statistical models which usually model a probability of infection, especially when taking risk factors into account. This probability of infection will primarily be quantified from the results of biological tests performed at regular intervals, to which information on infection dynamics as well as on the presence of risk factors will be added in a second stage.

## 2.2 Type of surveillance programmes considered

We consider surveillance programmes that collect infection related information at regular time intervals, initially with infection by the BVDV in cattle in mind. The information collected usually consists of biological test results. For example, for BVDV, such programmes can consist in performing ELISA tests on bulk tank milk several times a year in dairy herds or performing such ELISA tests on pools of blood samples collected in different categories of young animals. The general organisation of such programmes is presented in Figure 1.

## 2.3 Choice of a family of methods

Previous work on output based surveillance has considered the scenario tree methodology (More *et al.*, 2009). With this method, surveillance is split in different components. Each component has its own sensitivity for the detection of infection, but the specificity of all components is considered to be equal to one, i.e. there are no false positives. This is because units positive to a test are

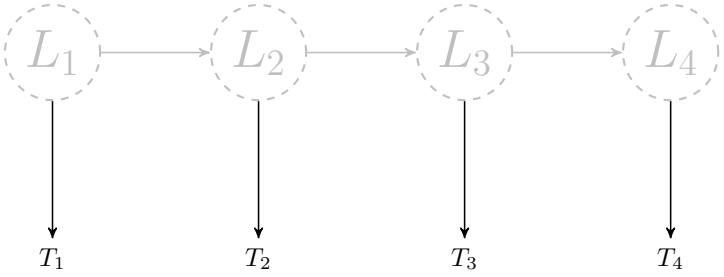



Figure 1: General organisation of the surveillance programmes considered. **L**: true (Latent) status regarding infection for the unit; **T**: biological test. The true status is monitored over time using biological tests.

re-tested with another test and either confirmed infected or found uninfected. The method returns a posterior probability of freedom from infection, given a sequence of negative test results for all the surveillance components for which there are available data. What is meant by freedom from infection is, *given a sequence of negative test results, the probability that the prevalence of infection in the unit is, with a predetermined level of confidence, below a chosen prevalence* called the **design prevalence**. This freedom from infection is a negative predictive value. For example, freedom from infection could mean that there is 95% chance that the prevalence in the unit of interest is below 2%. This method has been extensively used in animal disease surveillance at different levels such as herd (Meyer *et al.*, 2019) or country (Frössling *et al.*, 2009).

Another way of looking at this problem is to consider test results as imperfect measures of a true (latent) state regarding infection using latent class models. With Bayesian latent class models, it is possible to incorporate into the estimation of a probability of infection, previous knowledge on test sensitivities and specificities and disease dynamics (Nusinovici *et al.*, 2015; Courcoul *et al.*, 2010) or other variables of interest.

There are several differences between these 2 families of approaches. First, while scenario trees return a probability of being free from infection after having observed a certain number of negative test results, Bayesian latent class models return a probability of infection for different combinations of test results that can be positive or negative. Another major difference is that with scenario trees, the parameters for the simulations (or their distributions) must be determined from the literature or from expert knowledge. On the other hand with Bayesian models, given reasonable priors, the models can get more precise parameter estimates from the data. This could be valuable for control programmes against diseases that are still present in some territories. Such programmes often generate large amount of data that could be used to estimate parameters such as test sensitivities and specificities and probabilities of infection associated with risk factors.

In the remainder of this document, we chose to focus on a Bayesian model for

the estimation of a probability of infection from various pieces of information. Such a model can be used to estimate probabilities of (freedom from) infection, but could also provide parameter estimates that can later be used in scenario trees. In the following section, we provide a description of this model.

## 2.4 Basic description of the STOC free model

The STOC free model must be able to predict/ estimate a probability of infection and the associated uncertainty from all the available data. The hypotheses of the STOC free model are that:

- Infection modifies some biological parameters which can be measured using biological tests
- Biological tests sometimes provide inaccurate results by being negative in infected units (lack of sensitivity) or by being positive in uninfected units (lack of specificity)
- Information on infection dynamics (incidence / cure rates) is often available or can be estimated and can be incorporated into the model
- Information on risk factors acting on infection dynamics, such as risk factors of infection, is often available or can be estimated and can be incorporated in the model

The following concepts are useful for the design of such a model:

**Cause versus consequence.** There is conceptual difference between some of these types of information. The outcome of a biological test is a consequence of an (absence of) infection. The probability of a test being positive increases with time after the infection. The presence of some risk factors is associated with an increase in the probability of infection. In this case, the risk factor occurs before the infection.

**Structure of the model.** Our model is a structured representation of the associations between an event of interest (the infection) and causes and consequences of this event. Risk factors cause the infection which in turn determines test results.

**Data versus parameter.** The model will take various pieces of data as input and return a probability of infection as output. It provides a structure specifying how the different pieces of information are related. For each relationship encoded in the model structure, a strength of association must be either provided or estimated. This is what we call a parameter. For example, the amount by which the presence of a risk factor increases the probability of new infection must be provided or estimated. Parameters can be values or distributions.

| | | Infection | |
|---|---|---|---|
| | | + | - |
| Test | + | TRUE POSITIVE | FALSE POSITIVE |
| | - | FALSE NEGATIVE | TRUE NEGATIVE |

Table 1: Test result as a function of true status regarding infection.

**Predicting versus learning.** Once the structure of the model has been specified, the model parameters can either be provided by the user or estimated from data or both. When all parameter values are provided, a probability of infection can be predicted. When some of the parameter values are not known, they can be estimated from data. Learning parameter values from a statistical model and data is called statistical inference. There are various families of methods that allow performing statistical inference. In the STOC free model, we will use Bayesian inference which allows incorporating prior knowledge.

**Bayesian inference for learning and prediction.** When some of the parameters are unknown or known with too much uncertainty, Bayesian inference allows combining prior knowledge and data to produce better parameter estimates and to predict quantities of interest. This will be explained below.

# 3 Estimation of a probability of infection using biological tests

## 3.1 Tests characteristics

Biological tests are rarely able to perfectly discriminate infected from non infected units. These tests can make two types of errors. Infected units can test negative (false negative) or non infected units can test positive (false positive). Table 1 presents the labels associated to different test outcomes as a function of the true status regarding infection.

The ability of a test to correctly identify infected and uninfected units is measured by 2 characteristics.

**The sensitivity** is the probability for an infected unit ($I^+$) to test positive ($T^+$). This can be written as:

$$Se = p(T^+|I^+)$$

From Table 1, it can be calculated as:

$$Se = \frac{\text{TRUE POSITIVE}}{\text{TRUE POSITIVE} + \text{FALSE NEGATIVE}}$$

**The specificity** is the probability for a non-infected unit to test negative. This can be written as:

$$Sp = p(T^-|I^-)$$

From Table 1, it can be calculated as:

$$Sp = \frac{\text{TRUE NEGATIVE}}{\text{FALSE POSITIVE} + \text{TRUE NEGATIVE}}$$

**Interpretation of test sensitivity and specificity at different levels.** Imperfect sensitivities and specificities can originate from different levels. They can be inherent to a test which returns negative results in some of the positive samples or which return positive results in some of the negative samples. But the test charactersitics at the sample level are different from these same characteristics at the herd level. Usually, in a herd, not all animals are infected at the same time. The rarer the infection, the more animals must be sampled in order to get at least one positive test. Furthermore, things such as inadequate conservation of samples, errors in labelling the samples can be associated with wrong test results, even though the test is very good. Therefore, estimates of sensitivity and specificity are only valid at the level for which they are known.

## 3.2 Estimation of a probability of infection when test characteristics and disease prevalence are known

Sensitivity and specificity quantify the probabilities of each test result ($T^+$ / $T^-$) given a known status regarding infection ($I^+$ / $I^-$). However, when estimating a probability of infection, it is the test result which is known and it is the infection status that needs to be determined. Therefore, what needs to be quantified is $p(I^+|T+)$ (positive predictive value) as well as $p(I^+|T^-)$ (1 - negative predictive value). Table 2, shows how the sensitivity ($Se$), specificity ($Sp$) and prevalence of infection ($\pi$) are related. From this table it can be seen that the probability of being infected when the test is positive is:

$$p(I^+|T^+) = \frac{Se\pi}{p(T^+)}$$

$$p(I^+|T^+) = \frac{Se\pi}{Se\pi + (1 - Sp)(1 - \pi)}$$

and that the probability of being infected given a negative test is:

$$p(I^+|T^-) = \frac{(1 - Se)\pi}{p(T-)}$$

$$p(I^+|T^-) = \frac{(1 - Se)\pi}{(1 - Se)\pi + Sp(1 - \pi)}$$

As is apparent from this description, the estimation of a probability of infection from a test result requires an estimate of the infection prevalence. The

| | | Infection | |
|---|---|---|---|
| | | + | - |
| Test | + | $Se\pi$ | $(1-Sp)(1-\pi)$ |
| | - | $(1-Se)\pi$ | $Sp(1-\pi)$ |
| | | $\pi$ | $1-\pi$ |

Table 2: Relation of test results, true status regarding infection, sensitivity, specificity and prevalence

R function below allows predicting a probability of infection from the following pieces of information:

- `testRes`: test result. Positive = 1; negative = 0
- `Se`: sensitivity
- `Sp`: specificity
- `prev`: infection prevalence

```
probInf_F <- function(testRes, Se, Sp, prev){

  pInf <- testRes *
            (Se * prev) / (Se * prev + (1 - Sp) * (1 - prev)) +
          (1 - testRes) *
            (1 - Se) * prev / ((1 - Se) * prev + Sp * (1 - prev))
  pInf

}
```

The code below shows how to predict a probability of infection when infection prevalence is 0.05, test sensitivity is 0.8 and test specificity is 0.9.

```
probInf_F(testRes = 1, Se = .8, Sp = .9, prev = .05)

## [1] 0.2962963
```

Figure 2 shows the probability of being infected with a positive test result (red curve) and a negative test result (blue curve) as a function of infection prevalence in the population tested, for a test with a sensitivity of 0.8 and a specificity of 0.95. As infection prevalence increases, the probability of being infected increases to 1, even when the test result is negative. On the opposite, the probability of infection decreases when infection prevalence decreases, even when the test is positive.

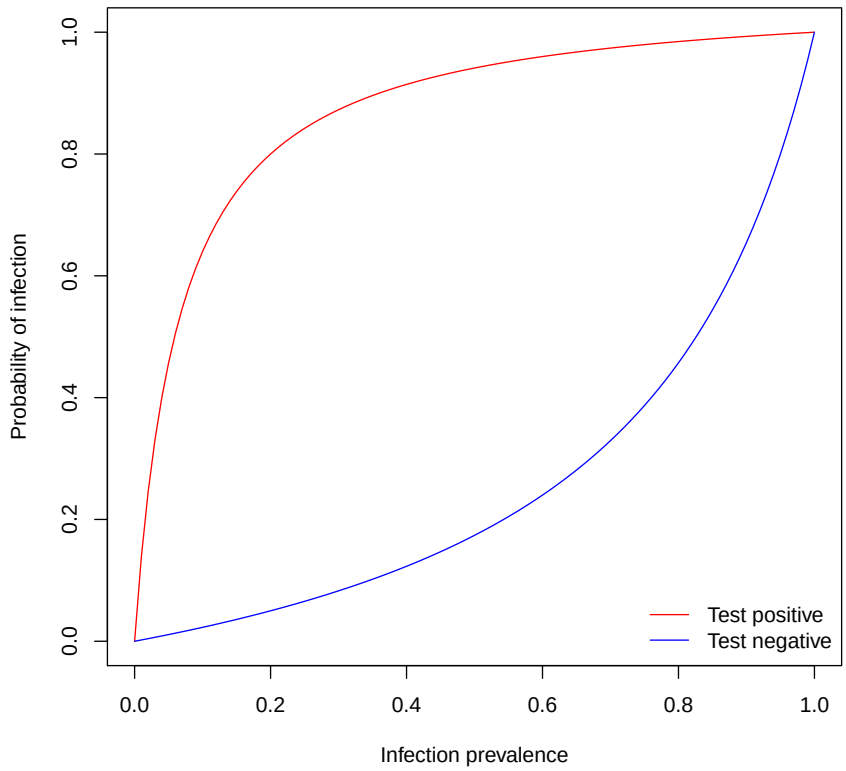



Figure 2: Probability of being infected as a function of infection prevalence in the population when the test result is positive (red curve) or negative (blue curve) for a test with a sensitivity of 0.8 and a specificity of 0.95.

## 3.3 Estimation of a probability of infection from imperfect information using Bayes' rule

### 3.3.1 Bayes' rule

As discussed above, the interpretation of a test result requires an estimate of the infection prevalence. Before obtaining a test result, and without any additional information, the probability of infection is equal to the prevalence. Obtaining a test results allows to update this probability. This updating is carried out using Bayes rules:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

If we label infected and non infected units as $I^+$ and $I^-$ respectively and test positive and test negative units as $T^+$ and $T^-$ respectively, the formula for the probability of being infected given a positive test result can be written as:

$$p(I^+|T^+) = \frac{p(T^+|I^+)p(I^+)}{p(T^+)}$$

In this formula $p(T^+)$ is the probability of testing positive. Test positive units gather true positives and false positives. Therefore, the formula can be re-written as:

$$p(I^+|T^+) = \frac{p(T^+|I^+)p(I^+)}{p(T^+|I^+)p(I^+) + p(T^+|I^-)(1 - p(I^+))}$$

It can be noted that $p(I^+)$ is the probability of infection in the population (i.e. the prevalence $\pi$), $p(T^+|I^+)$ is the sensitivity and $p(T^+|I^-)$ is 1 - specificity. Therefore, the formula can be re-written as:

$$p(I^+|T^+) = \frac{Se\pi}{Se\pi + (1 - Sp)(1 - \pi)}$$

The probability of infection before having obtained a test result is the infection prevalence. We call it a prior probability. We call the probability of being infected after having obtained a test result the posterior probability of infection.

### 3.3.2 Prior and posterior distributions

In the formula for the posterior probability of infection, not all the parameters are known precisely. It is rare to know that infection prevalence is, for example, exactly 0.17. Instead, rough estimates are usually known. The prevalence can be known to be of around 0.2 with the certainty that it is greater than 5% and less than 30%. This prior knowledge can be described with a distribution. There exists a vast number of statistical distributions that describe different types of data (discrete, continuous, counts ...). The Beta distribution describes

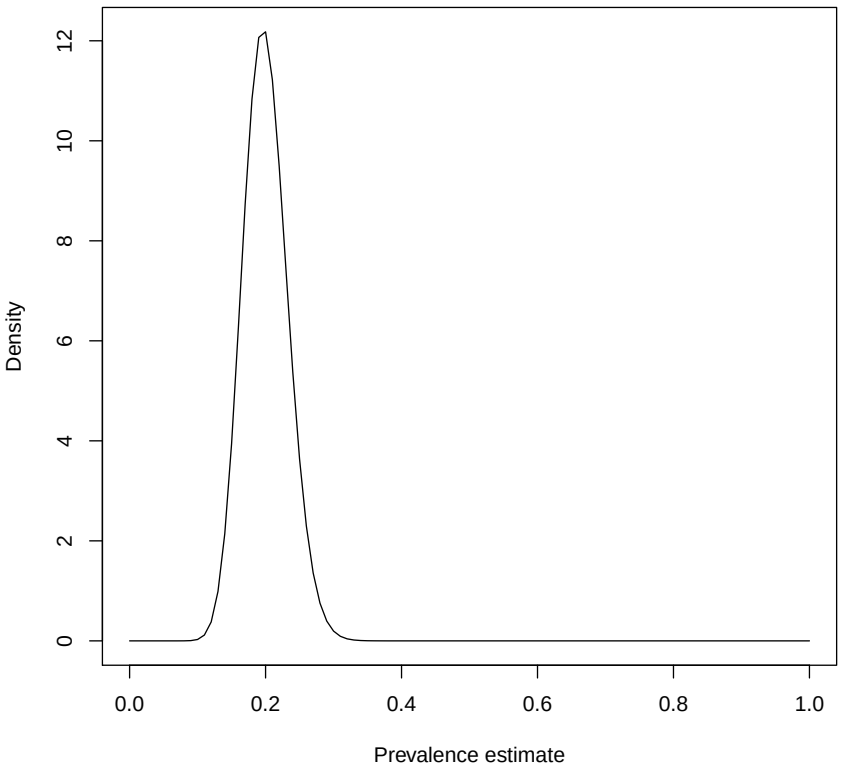



Figure 3: Prior for the prevalence of infection represented with a Beta distribution with parameters 30 and 120.

continuous values between 0 and 1[5]. Figure 3 represents the probability density function for a Beta distribution with parameters 30 and 120[6]. As can be seen, the most likely value is around 0.2 with all the values concentrated between 0.1 and 0.3. Figure 4 represents the probabilities of being infected before a test results is available (dashed grey) and after a positive (red) and negative (blue) test. The curves show the amount by which the probability of infection is increased after having observed a positive test result and the amount by which it is decreased after having observed a negative test result.

[5]See the Wikipedia page on the Beta ditribution: https://en.wikipedia.org/wiki/Beta_distribution

[6]In R, this distribution can be plotted using the following line of code: curve(dbeta(x, 30, 120)) as explained in Section 1.

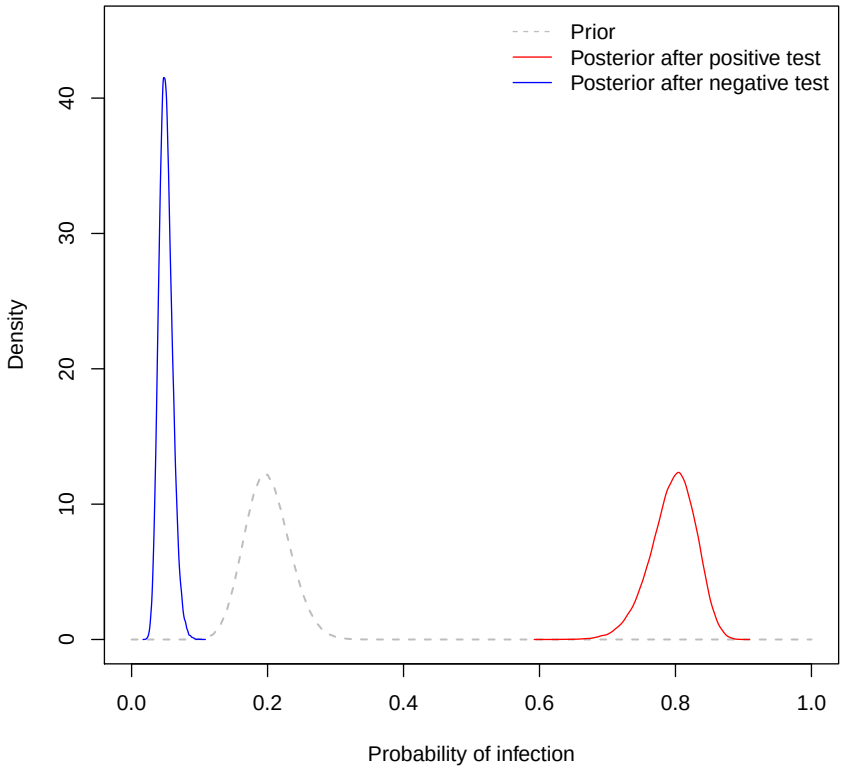



Figure 4: Distributions of the probability of infection: before testing (prior), after a positive test and after a negative test.

### 3.3.3 Bayesian inference

In Bayesian inference, Bayes' rule is used to estimate model parameters.

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

In this equation, $\theta$ are the model parameters and $y$ are the data. $p(\theta|y)$ is the posterior distribtion for the parameters, $p(\theta)$ are the prior distributions for all $\theta$, $p(y|\theta)$ is the probability of observing the data given a set of parameter values, which is called the likelihood. In essence, in Bayesian inference, the data are used to update the prior probabilities through the likelihood function. $p(y)$ is called a normalising constant. This constant cannot be estimated analytically which has prevented the development of Bayesian inference until the early XXI century. Bayesian parameter estimation is now carried out using a set of techniques called Markov chain Monte Carlo (MCMC). The first readily usable computer programme that allowed performing Bayesian inference was called BUGS for *Bayesian inference Using Gibbs Sampling*. Gibbs sampling is a type of MCMC algorithm. With the BUGS language, it is straightforward to specify a model for the data and priors for the parameters. The programme works by drawing values from the full posterior distribution for all parameters. Below, we will perform Bayesian inference using a programme called JAGS (Just Another Gibbs Sampler), which uses the BUGS language, with slight modifications.

### 3.3.4 Bayesian inference using JAGS

As an example, let's say we carried out a study to estimate the prevalence of infection by the BVDV last year, with a test that has a sensitivity of 0.8 and a specificity of 0.9. We found a prevalence of 0.2. This year, we repeat the study with the same test. We test 100 herds and get 10 positive results. We want to update our estimate for infection prevalence with these data.

A test result is either positive (=1) or negative (=0). Our data can therefore be modelled with a Bernoulli distribution.

$$T_i \sim Bernoulli(p_i)$$

where $T_i$ is a variable for test result (0 or 1) for unit $i$ and $p_i$ is the proportion of positive tests. $p_i$ is modelled as:

$$p_i = Se\pi + (1 - Sp)(1 - \pi)$$

The value of Se and Sp are set to 0.8 and 0.95 respectively. We assume a Beta(30, 120) prior for $\pi$.

$$Se = 0.8$$

$$Sp = 0.95$$

$$\pi \sim Beta(30, 120)$$

The model is written in the JAGS language, from within R using the `cat()` function and is stored as a text file called `modelFile.txt`.

```
cat("model{

  for(i in 1:100){

  testRes[i] ~ dbern(p[i])

  p[i] <- Se * prev + (1 - Sp) * (1 - prev)

  }

  prev ~ dbeta(30, 120)
  Se <- .8
  Sp <- .95

}", file =  "modelFile.txt")
```

The prevalence will be estimated from a vector of 10 positive test results and 90 negative test results.

```
dataJAGS <- list(testRes = c(rep(1, 10), rep(0, 90)))
```

The `rjags` package is loaded. This package allows calling JAGS from R.

```
library(rjags)

## Loading required package:  coda
## Linked to JAGS 4.3.0
## Loaded modules:  basemod,bugs
```

The model is compiled using the `jags.model()` function. We just need to specify the location of the model file, the data and the number of chains to be run in parallel.

```
modelJAGS <- jags.model(
               file = "modelFile.txt",
               data = dataJAGS,
               n.chains = 3)

## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 100
```

```
##    Unobserved stochastic nodes: 1
##    Total graph size: 111
##
## Initializing model
```

Samples are drawn from the posterior distribution of `prev` using the `coda.samples()` function. From the model, 1000 samples from the posterior distribution of the `prev` parameter are kept for later analysis.

```
samplesJAGS <- coda.samples(modelJAGS,
                            variable.names = "prev",
                            n.iter = 1000)
```

Figure 5 shows the model results. We have run 3 simulations (chains) in parallel. These 3 chains have converged to the same posterior distribution. The graph on the right hand side of the figure shows the density for the 3000 samples from the posterior distribution of `prev`.

The model results can be summarised as shown below. As is visible, the mean of the posterior distribution of `prev` is somewhere between the proportion of positive samples and the mean of its prior distribution. Adding more observations would have shifted the posterior towards the proportion observed in the data. Having less data would give more weight to the prior distribution.

```
summary(samplesJAGS)

##
## Iterations = 1001:2000
## Thinning interval = 1
## Number of chains = 3
## Sample size per chain = 1000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean             SD       Naive SE Time-series SE
##      0.1651916      0.0256125      0.0004676      0.0006300
##
## 2. Quantiles for each variable:
##
##   2.5%    25%    50%    75%  97.5%
## 0.1175 0.1471 0.1642 0.1820 0.2180
```

Keeping the same data, the model is run with different priors: we assume that the test never returns false positives (Sp = 1) and that the sensitivity is lower than previously thought with a Beta(15, 5) prior.

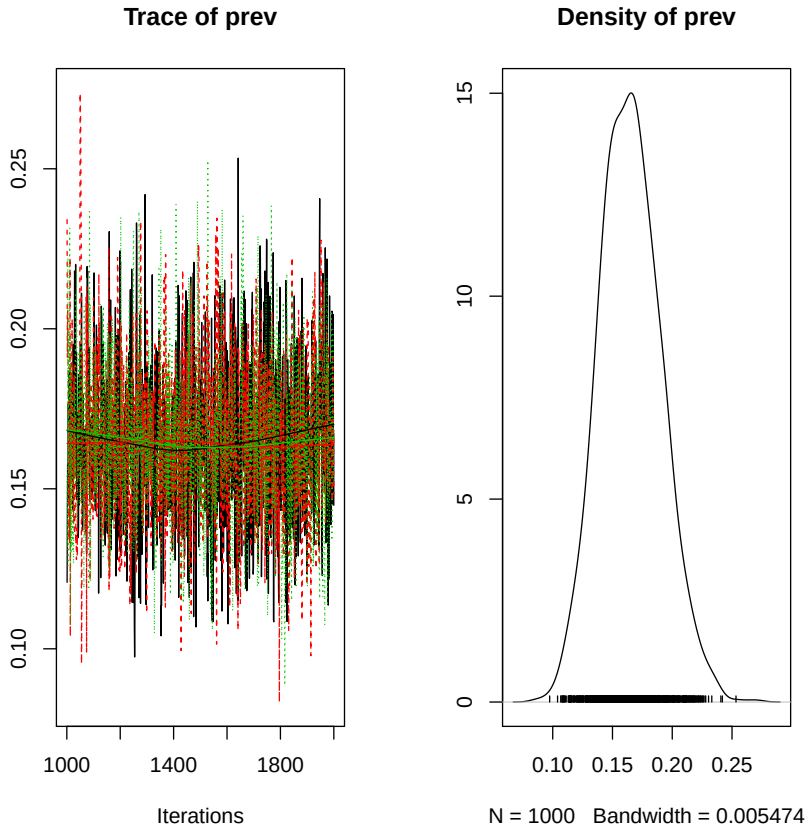



Figure 5: Results of the JAGS model for the estimation of infection prevalence.

```
cat("model{

  for(i in 1:100){

  testRes[i] ~ dbern(p[i])

  p[i] <- Se * prev + (1 - Sp) * (1 - prev)

  }

  prev ~ dbeta(30, 120)
  Se ~ dbeta(15, 15)
  Sp <- 1

}", file =  "modelFile1.txt")

modelJAGS1 <- jags.model(
              file = "modelFile1.txt",
              data = dataJAGS,
              n.chains = 3)

## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 100
##    Unobserved stochastic nodes: 2
##    Total graph size: 111
##
## Initializing model

samplesJAGS1 <- coda.samples(modelJAGS1,
                             variable.names = c("prev", "Se"),
                             n.iter = 1000)
```

In this case, the prevalence estimate is higher than in the previous example (See Figure 6).

Lastly, the same Beta(1, 1) prior is put on `Se`, `Sp` and `prev`. This prior is uniform on the interval 0, 1. It can be called uninformative since does not put any constraint on parameter values.

```
cat("model{

  for(i in 1:100){

  testRes[i] ~ dbern(p[i])
```

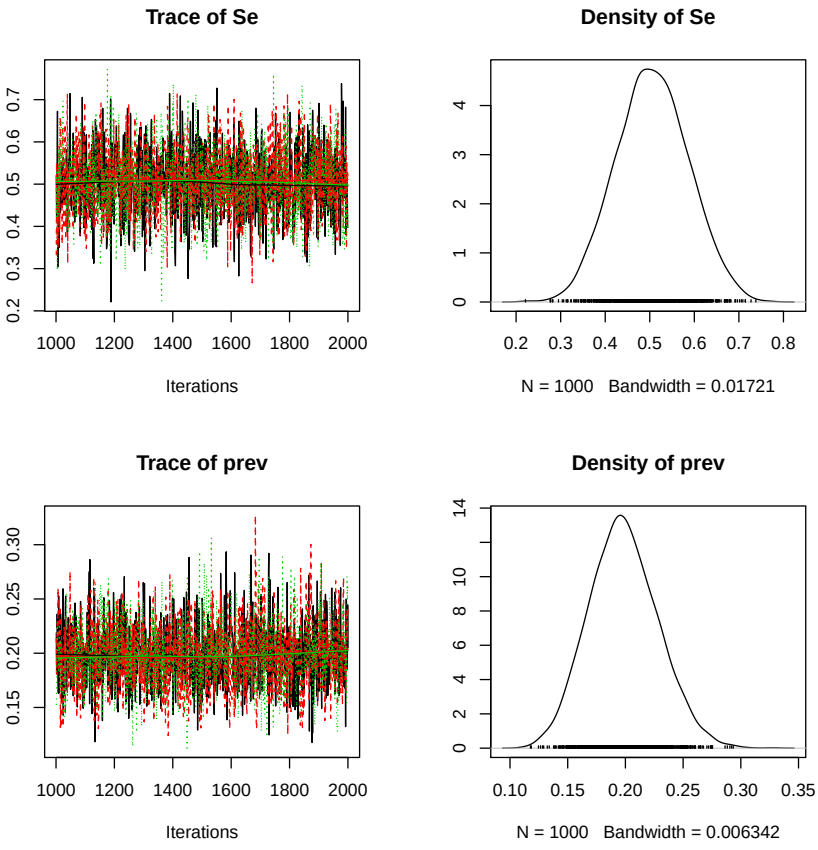



Figure 6: Results of JAGS model 1 for the estimation of infection prevalence.

```
  p[i] <- Se * prev + (1 - Sp) * (1 - prev)

  }

  prev ~ dbeta(1, 1)
  Se ~ dbeta(1, 1)
  Sp ~ dbeta(1, 1)

}", file =  "modelFile2.txt")

modelJAGS2 <- jags.model(
               file = "modelFile2.txt",
               data = dataJAGS,
               n.chains = 3)









samplesJAGS2 <- coda.samples(modelJAGS2,
                             variable.names = c("prev", "Se", "Sp"),
                             n.iter = 1000)
```

The problem with this model is that there are many parameter values compatible with the prior that could have produced the data. As can be seen in Figure 7, the 3 chains are often far apart: they do not converge to the same posterior distributions. The parameters are not identifiable. The reasons for running several chains is to be able to check that several independent simulations converge to the same posterior distribution.

# 4 Contribution of the knowledge about infection dynamics on the estimation of a probability of infection

The STOC free model needs to be able to incorporate information other than test results in order to estimate a probability of infection. In this section, we consider information on infection dynamics as well as on risk factors.

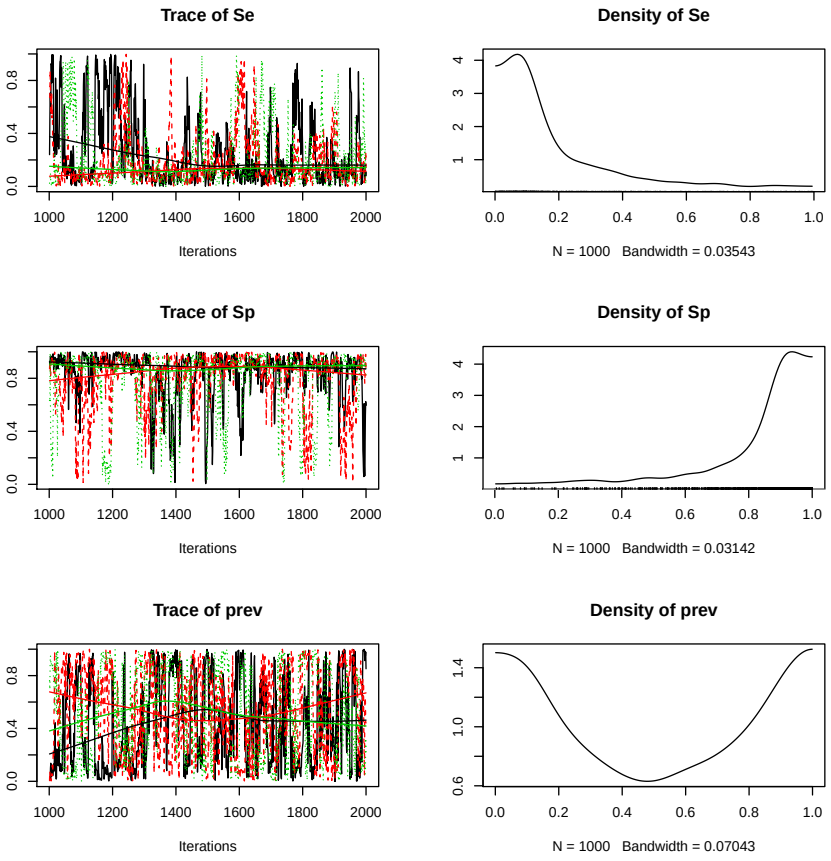



Figure 7: Results of JAGS model 2 for the estimation of infection prevalence.

## 4.1 Infection dynamics

For simplicity, we assume that the latent infection status can be known perfectly ($Se = 1$ ; $Sp = 1$). We also assume that there are only 2 possible states regarding infection: infected and not infected. Considering time as discrete, being infected at time $t$ can be represented as the outcome of 2 possible events:

- a new infection in units that were not infected at $t-1$
- the absence of elimination of infection in units that were infected at $t-1$

We use $\tau_1$ to denote a new infection between $t-1$ and $t$ and $\tau_2$ to denote the non elimination of an infection between $t-1$ and $t$. This is represented in Figure 8.

Most of the time, some prior knowledge on infection dynamics is available. For example, in cattle herds, infections by *Mycobacterium avium paratuberculosis* is notoriously difficult to eliminate. Therefore, between 2 consecutive months, the probability for infected herds of becoming uninfected must be extremely low. In Bayesian inference, this can be incorporated in the model as a prior distribution.

Furthermore, the interrelations between new infections, infection elimination and prevalence can be exploited. We call equilibrium the situation in which infection prevalence is stable in a population. At equilibrium, for the prevalence to remain stable, the number of new infections must compensate the number of infection eliminations. If we call $\pi$ the infection prevalence, $\tau_1$ the probability of new infection and $\tau_2$ the probability of not eliminating the infection, this implies that:

$$\tau_1(1-\pi) = \tau_2\pi$$

This means that knowing any 2 of the 3 parameters allows knowing the third one. For example, at equilibrium, the prevalence is:

$$\pi = \frac{\tau_1}{\tau_1 + \tau_2}$$

## 4.2 Modelling the probability of new infection

The focus of STOC free is on detecting new infections as early as possible after their occurrence. In this context, being able to incoporate risk factors of introduction of infection is of interest. In order to do this, we specify a model for $\tau_1$. In Figure 8, this is represented as a single risk factor $x$ acting on $\tau_1$. The association between $\tau_1$ and $x$ is modelled with a logistic regression:

$$ln(\frac{\tau_1}{1-\tau_1}) = \theta_1 + \theta_2 x$$

Here, for simplicity, we only consider one risk factor. Eventually, as many risk factors as necessary can be included in the model.

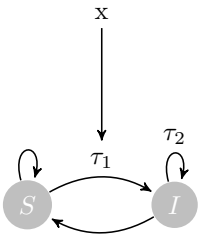



Figure 8: Representation of infection dynamics. Two latent states are modelled: $S$ for susceptible and $I$ for infected. $\tau_1$ is the probability of moving from $S$ to $I$ between $t-1$ and $t$ or probability of new infection. $\tau_2$ is the probability of not eliminating the infection between $t-1$ and $t$, i.e. remaining $I$. Risk factor $X$ acts on the probability of new infection $\tau_1$.

# 5 The STOC free model: a framework for the estimation of a probability of infection from heterogeneous data

What we call the STOC free model is in fact a set of computer programmes that have the same structure and parameters for the prediction of a probability of infection from various pieces of information. Section 5.3 describes an R function that uses information on the status at the previous test as well on infection dynamics (possibly including risk factors) to predict a probability of infection. Section 5.6 and 5.7 presents the same model coded in 2 different Bayesian inference programmes: JAGS and Stan. We are still exploring the pros and cons of each one.

## 5.1 Modelling hypotheses

The models combine information on regular test results as well as the explicit modelling of infection dynamics. These dynamics are described by a probability of new infection and a probability of infection elimination between consecutive tests. The dynamics parameters can be modelled as a function of the presence/absence of risk factors through logistic regression. Below is a summary of the modelling hypotheses:

- Biolgical tests are imperfect measures of an unobserved latent state. A test characteristics are described by its sensitivity and a specificity
- Two latent states of interest are considered: infected ($L = 1$) and uninfected ($L = 0$)
- For unit $i$, the probability of being in latent state $L$ at time $t$ only depends on the latent state of unit $i$ at time $t-1$ (Markovian property)

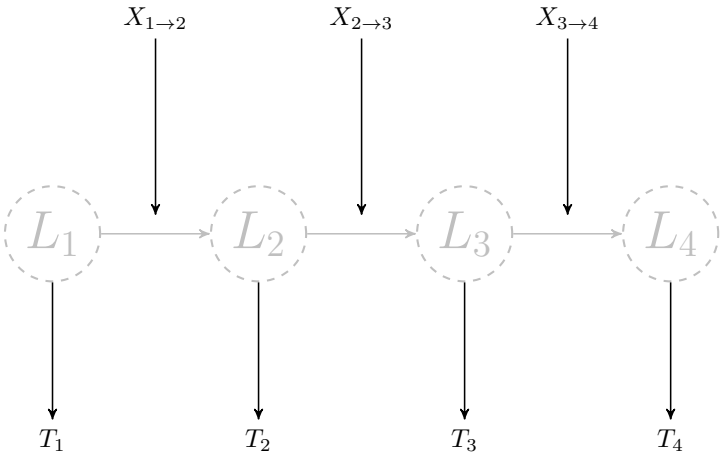



Figure 9: Working hypotheses for the modelling of latent states in unit $i$ over 4 consecutive time points. The true state $L_{it}$ is imperfectly measured with tests $T_t$. Risk factor $X_{t-1\to t}$ acts on the transition between latent states between consecutive time points.

- The probability of becoming infected between $t-1$ and $t$ can be modelled as a function of one or several risk factors using logistic regression.

These modelling hypotheses are represented in Figure 9 and the notation used is presented in Figure 10. The fact that the latent state at time $t$ only depends on the latent state at time $t-1$ is called the Markovian property. Models that describe an imperfectly measured latent state with a Markovian dynamics are called hidden Markov models (HMMs). There are several methods that allow estimating the parameters of such models (for an introduction to HMMs see Zucchini *et al.* (2017)). Several examples have been written in Stan[7].

## 5.2 Model equations

The STOC free model can be represented by the following set of equations. Note that in the text $\tau_1$ and $\tau_2$ have been changed to $\tau^1$ and $\tau^2$ to make the equations easier to read.

**Latent state.** It can take 2 values: 0 = uninfected; 1 = infected. For a given unit $i$ at a given time $t$, it follows a Bernoulli distribution:

$$L_{it} \sim Bernoulli(\pi_{it})$$

The distribution has a parameter $\pi_{it}$ which represents the probability of being infected.

[7]See this page of the Stan manual or this Tutorial on Hidden Markov Models using Stan.

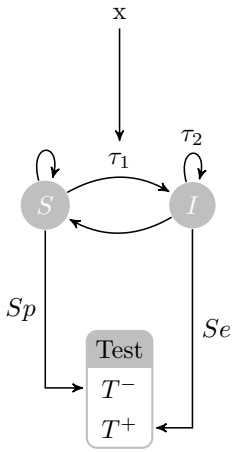



Figure 10: Notation used in the STOC free model. Two latent states are modelled: $S$ for susceptible and $I$ for infected. $\tau_1$ is the probability of moving from $S$ to $I$ between $t-1$ and $t$ or probability of new infection. $\tau^2$ is the probability of not eliminating the infection between $t-1$ and $t$, i.e. remaining $I$. Risk factor $X$ acts on the probability of new infection $\tau^1$. Units in the $S$ state test negative with probability $Sp$ and test positive with probability $1-Sp$. Units in the $I$ state test positive with probability $Se$ and test negative with probability $1-Se$.

**Infection dynamics.** The probability of being infected at $t$ depends on the status at $t-1$. For units that were uninfected at $t-1$ (i.e. $L_{i(t-1)} = 0$), a probability of new infection ($\tau^1$) is modelled. For units that were infected at $t-1$ (i.e. $L_{i(t-1)} = 1$), a probability of remaining infected ($\tau^2$) is modelled.

$$\pi_{it} = (1 - p_{i(t-1)})\tau^1 + p_{i(t-1)}\tau^2$$

**Probability of new infection.** In STOC free, the probability of new infection is modelled as a function of the presence/absence of risk factors $X$ using a logistic regression.

$$ln(\frac{\tau^1_{it}}{1-\tau^1_{it}}) = \theta_1 + \theta_2 X_{i(t-1)}$$

**Test results** ($T$) can be positive or negative. Each result follows a Bernoulli distribution with a probability $p$ of being positive.

$$T_{it} \sim Bernoulli(p_{it})$$

$p_{it}$ is modelled as depending on the latent status through sensitivity and specificity.

$$p_{it} = SeL_{it} + (1 - Sp)(1 - L_{it})$$

## 5.3 The STOC free model in R: an R function to predict a probability of infection when all information is available

Below is an R function that predicts a probability of infection when information on the previous status, $\tau_1$ and $\tau_2$ is available. The function can take the following arguments:

- `L_previous`: previous status (0 or 1) or infection prevalence (any value between 0 and 1.)
- `tau2`: for units infected at the previous test, probability of not having eliminated the infection. Must be between 0 and 1.
- `tau1`: for units not infected at the previous test, probability of having acquired the infection. Must be between 0 and 1. When supplied, the arguments below must not be provided.
- `rf`: risk factor presence. 0 = absence; 1 = presence
- `theta1`: intercept of the logistic regression modelling the probability of new infection as a function of the presence/absence of the risk factor. It is the odds of infection in units not exposed to the risk factor.

- `theta2`: coefficient of the logistic regression modelling the probability of new infection as a function of the presence/absence of the risk factor. It is the natural logarithm of the odds ratio associated with the presence of the risk factor.

```
predProbInf_F <- function(L_previous, tau1, tau2,
                          rf, theta1, theta2){

  if(missing(tau1)){

    logit_tau1 <- theta1 + theta2 * rf
    tau1 <- exp(logit_tau1) / (1 + exp(logit_tau1))

  } else {

    cat("Value provided for tau1.
Values for rf, theta1 and theta2 will be ignored.\n\n")
  }

  L_status <- L_previous * tau2 +
              (1- L_previous) * tau1

  cat("Predicted probability of infection:", L_status)

}
```

The function is tested with all arguments supplied.

```
predProbInf_F(L_previous = 0,
              tau1 = .1,
              tau2 = .5,
              rf = 1,
              theta1 = -1,
              theta2 = 1)

## Value provided for tau1.
## Values for rf, theta1 and theta2 will be ignored.
##
## Predicted probability of infection: 0.1
```

The function is tested with tau1 not supplied.

```
predProbInf_F(L_previous = 0,
              tau2 = .5,
              rf = 1,
```

```
              theta1 = -1,
              theta2 = 1)

## Predicted probability of infection: 0.5
```

The function is tested with all tau1 supplied and arguments associated with the risk factor not supplied.

```
predProbInf_F(L_previous = 0,
              tau1 = .1,
              tau2 = .5)

## Value provided for tau1.
## Values for rf, theta1 and theta2 will be ignored.
##
## Predicted probability of infection: 0.1
```

## 5.4 Toy dataset for inference with JAGS and Stan

In this section, we provide a toy dataset in order for the reader to be able to test the JAGS and Stan codes. There are only 3 herds with 4 times each.

The dataset contains the following columns:

- `herdID`: herd id
- `test_t`: time
- `status`: true (latent) status. Infected = 1; uninfected = 0
- `testRes`: test result. Positive = 1; negative = 0
- `rf`: risk factor. Present = 1; absent = 0

```
herdInf

##    herdID test_t status testRes rf
## 1       1      1      1       1  0
## 2       1      2      1       1  0
## 3       1      3      1       1  1
## 4       1      4      1       1  0
## 5       2      1      0       0  1
## 6       2      2      0       0  0
## 7       2      3      1       1  0
## 8       2      4      1       1  1
## 9       3      1      1       1  1
## 10      3      2      1       1  0
## 11      3      3      1       1  1
## 12      3      4      0       1  1
```

The latent status is used for building the dataset and checking results but is assumed to be unknown in the models.

## 5.5 Parameters for the prior distribution used for inference with JAGS and Stan

The same prior will be used by both programmes. It has to be noted that for the normal distribution, JAGS uses the precision which is the inverse of the variance while Stan use the standard deviation.

In JAGS:

$$y \sim Normal(\mu, \frac{1}{\sigma^2})$$

In Stan:

$$y \sim Normal(\mu, \sigma)$$

The list of priors that will be used:

```
Se_beta_a <- 1000
Se_beta_b <- 200
Sp_beta_a <- 450
Sp_beta_b <- 25
theta1_mean <- 0
theta1_sd <- 10
theta2_mean <- 0
theta2_sd <- 10
tau2_beta_a <- 80
tau2_beta_b <- 8
pi1_beta_a <- 1
pi1_beta_b <- 1
```

## 5.6 The STOC free model in JAGS: Bayesian inference for learning parameter values and predicting a probability of infection.

Below is the JAGS code for the STOC free model. The difference with the R function described above is that the model will learn from the data and prior. As explained above, the model works by drawing samples from the posterior distributions of all the parameters. It combines data, likelihood and prior information. Before embarking upon serious analysis using this programme, readers are encouraged to read a bit more about Bayesian inference, especially on the subject of model convergence.

The data used by the model are the variables named:

- `testRes`: test result. Positive = 1; negative = 0

- `rf`: risk factor. Present = 1; absent = 0

A few other variables are supplied as data: `nHerds` is the number of herds (3 here) and `N` is the total number of observations (12 here).

The quantity of interest in the model is the variable `status`. It is not supplied as data, although it is present in the toy dataset. The JAGS model will consider it as a vector of parameters and impute a value for each status at each iteration of the algorithm. The proportion of iterations where the imputed status is infected will be taken as the probability that the unit is infected. `status` is modelled as following a Bernoulli distribution with proportion `pi`. In each unit, the first and later statuses are modelled differently. For the first status, we have no previous information on neither the previous status nor on the presence of the risk factor. In this case, we assume that the probability of being infected is the overall prevalence `pi`. From the second test onwards, the probability of being infected depends on both the previous status as well as on the presence of a risk factor.

```
cat("model{

  for(h in 1:nHerds){

    pi[indI[h]] ~ dbeta(pi1_beta_a, pi1_beta_b)
    status[indI[h]] ~ dbern(pi[indI[h]])

   for(t in indJ[h]:indF[h]){

    logit(tau1[t]) <- theta[1] + theta[2] * rf[t-1]

    pi[t] <- (1 - status[t-1]) * tau1[t] + status[t-1] * tau2
    status[t] ~ dbern(pi[t])

      } #t

    } #h

  for(i in 1:N){

    pTestPos[i] <- status[i] * Se + (1 - status[i]) * (1 - Sp)
    testRes[i] ~ dbern(pTestPos[i])

  }

  ## Priors

  ## Priors for sensitivities and specificities
    Se ~ dbeta(Se_beta_a, Se_beta_b)
```

```
    Sp ~ dbeta(Sp_beta_a, Sp_beta_b)

  ## Probability of new infection
    theta[1] ~ dnorm(theta1_mean, theta1_tau)
    theta[2] ~ dnorm(theta2_mean, theta2_tau)

  ## Probability of not elimintaing the infection
    tau2 ~ dbeta(tau2_beta_a, tau2_beta_b)

    }", file =  "modelFile.txt")
```

Formatting the data for JAGS. The parameters for the prior distributions are also supplied as data.

```
dataJAGS <- list(
    nHerds = max(herdInf$herdID),
    N = nrow(herdInf),
    indI = which(herdInf$test_t == 1),
    indJ = which(herdInf$test_t == 2),
    indF = which(herdInf$test_t == max(herdInf$test_t)),
    testRes = herdInf$testRes,
    rf = herdInf$rf,
    Se_beta_a = Se_beta_a,
    Se_beta_b = Se_beta_b,
    Sp_beta_a = Sp_beta_a,
    Sp_beta_b = Sp_beta_b,
    theta1_mean = theta1_mean,
    theta1_tau = 1 / theta1_sd^2,
    theta2_mean = theta2_mean,
    theta2_tau = 1 / theta2_sd^2,
    tau2_beta_a = tau2_beta_a,
    tau2_beta_b = tau2_beta_b,
    pi1_beta_a = pi1_beta_a,
    pi1_beta_b = pi1_beta_b
  )
```

The model is compiled. Three chains will be run in parallel.

```
modelJAGS <- jags.model( file = "modelFile.txt",
                         data = dataJAGS,
                         n.chains = 3)

## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
```

```
##    Observed stochastic nodes: 12
##    Unobserved stochastic nodes: 20
##    Total graph size: 150
##
## Initializing model
```

Here is a list of the parameters for which we want to keep the samples from the posterior distribution for later analysis. Some care is needed here. At each iteration, one value per parameter per chain is saved. As the number of iterations increases, the programme can take a lot of RAM.

```
savedParam <- c("Se", "Sp", "status")
```

Sampling from the posterior distribution. We run the model for 100 iterations only. The aim here is to make sure that the model runs, not to perform inference.

```
samplesJAGS <- coda.samples(modelJAGS,
                            variable.names = savedParam,
                            n.iter = 100)
```

## 5.7 The STOC free model in Stan

Stan is a computer programme that performs Bayesian inference using Hamiltonian Monte Carlo[8]. Unlike JAGS, Stan cannot sample from the posterior distributions of discrete parameters such as the latent status wich is of interest to us in this work. Therefore, in Stan, the latent status is modelled as a mixture of 2 Bernoulli distributions which makes the code a bit harder to read.

Below is the model code. The code has 4 blocks:

- `data`: the variables and their type must be declared.
- `parameters`: list of the variables that will be modelled
- `transformed parameters`: calculations performed on the data and parameters
- `model`: sampling statements

In the code below, the same bits as in the JAGS code can be found, mostly in the `transformed parameters` and `model` blocks.

The `rstan` library needs to be loaded

[8]To get an intuition of the difference between Hamiltonian Monte Carlo and older methods see here. This method of estimation presents many advantages.

```
library(rstan)
```

Below are the data that will be used by Stan.

```
dataStan <- list(
    nHerds = max(herdInf$herdID),
    N = nrow(herdInf),
    indI = which(herdInf$test_t == 1),
    indJ = which(herdInf$test_t == 2),
    indF = which(herdInf$test_t == max(herdInf$test_t)),
    testRes = herdInf$testRes,
    x = herdInf$rf,
    Se_beta_a = Se_beta_a,
    Se_beta_b = Se_beta_b,
    Sp_beta_a = Sp_beta_a,
    Sp_beta_b = Sp_beta_b,
    theta1_mean = theta1_mean,
    theta1_sd = theta1_sd,
    theta2_mean = theta2_mean,
    theta2_sd = theta2_sd,
    tau2_beta_a = tau2_beta_a,
    tau2_beta_b = tau2_beta_b,
    pi1_beta_a = pi1_beta_a,
    pi1_beta_b = pi1_beta_b
  )
```

The model is compiled.

```
StanM <- stan_model(model_code = STOCfreeM_StanCode)
```

Samples from the parameter posterior distributions. In this case, the model status is not modelled explicitely but its probability `ppi` is.

```
samplesStan <- sampling(StanM,
     data = dataStan,
     chains = 3,
     iter = 100,
     pars = c("Se", "Sp", "ppi"))
```

# 6 Further work

The model is currently tested on simulated data as well as on data generated by a BVD control programme which has been running for several years in France. Eventually, the model will be tested on programmes from other countries. The questions that are explored are:

**Importance of prior distributions.** The Bayesian model has to estimate parameters for test characteristics as well as for associations between risk factors and infection. If the priors are too vague, there may be problems of indentifiability: there would be different parameter values/ distributions compatible with the observed data (see section 3.3.4). Avoiding this type of problem will require setting priors that are not too vague on some of the parameters in order to constrain the estimation to plausible values. Which parameters need to be constrained and by how much needs to be investigated. The ability of the model to recover known parameter values given different priors can be explored using simulated data.

**Incorporation of test results collected at different levels.** Different testing protocols are used for BVD surveillance. The most frequently seen are regular bulk tank milk testing, tests on pools of samples from different age groups (spot tests) or tissue tag testing on all newborn calves. While bulk tank milk ELISA gives one piece of information for the entire lactating herd, tissue tag testing on all newborn calves provides animal level information. The sensitivities and specificities that are associated with the tests used are usually only valid at the sample level, but hardly apply at the group or herd level (see section 3.1). There are in fact two issues here. First, the rarer the infection in a given group, the more individuals have to be sampled to be sure to include at least one infected. This will be associated with a decrease in sensitivity compared to the sample level. Secondly, within a herd, the infection may have different probabilities to reach different age groups. For the BVDV, infection will mostly spread by the birth of PI calves which will act as sources of infection for the rest of the herd. The infection may reach the lactating herd some time after the birth or the introduction of a PI calf. Therefore, testing bulk tank milk could be associated with a delayed detection. Both issues need to be further explored.

**Identification of risk factors to include in the model.** The risk factors of interest are the risk factors of disease introduction in a herd or in a territory. The two major classes of risk factors will be through animal purchases and through neigbourhood contacts. But the way this is modelled still needs to be specified. For example, the probability of introduction through neighbourhood depends on the infection prevalence in the neighbourhood as well as on the probability of contact with the animals from the neighbourhood. In turn, this probability of contact depends on the categories of animals that graze, the amount of time spent at pasture ... In order to progress on this question, a systematic review of the risk factors of BVDV introduction is currently underway within the STOC free project.

**Changes in the association between risk factors and probability of infection over time.** As a result of programme success in reducing infection prevalence in a territory, the association between risk factors and probability

of infection may evolve over time. In practice, the probabilities of purchasing an infected animal and of introducing the infection through contact with neighbouring herds will decrease as infection prevalence decreases. Therefore, the coefficients of the logistic regression used in the model may have to be allowed to vary over time. One possibility that would mimic surveillance would be ta carry out sequential evaluations with estimates from the previous round of testing used as priors for the current estimation. This is explored with data from a Frecnh BVD control programme.

**Combining inference and prediction.** Most control programmes against BVD based on bulk milk testing test herds once or twice a year. In this situation, the statuses are only updated when a new test result is available. This is also on these occasions that the model parameters can be re-estimated with the Bayesian model. However, risk factors such as animal purchases can be available at a much higher frequency. In this case, a probability of infection can be predicted from knowledge on risk factor occurrence even though there is no test result. Predictions can be performed using the function presented in section 5.3 or scenario tree modelling. In theory, it would even be possible to ask for a new test when, given knwoledge on risk factors, the probability of infection is above some predefined level. The way to combine inference using test results, when they are available, and prediction at more frequent time intervals needs to be explored at a later stage.

# References


Cameron, AR. 2012. The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease. *Preventive veterinary medicine*, **105**(4), 280–286.

Courcoul, A, Vergu, E, Denis, J-B, & Beaudeau, F. 2010. Spread of Q fever within dairy cattle herds: key parameters inferred using a Bayesian approach. *Proceedings of the Royal Society B: Biological Sciences*, **277**(1695), 2857–2865.

Frössling, J, Ågren, ECC, Eliasson-Selling, L, & Lewerin, SS. 2009. Probability of freedom from disease after the first detection and eradication of PRRS in Sweden: scenario-tree modelling of the surveillance system. *Preventive veterinary medicine*, **91**(2-4), 137–145.

Meyer, A, McAloon, CG, Tratalos, JA, More, SJ, Citer, LR, Graham, DA, & Sergeant, ESG. 2019. Modeling of alternative testing strategies to demonstrate freedom from Mycobacterium avium ssp. paratuberculosis infection in test-negative dairy herds in the Republic of Ireland. *Journal of dairy science*, **102**(3), 2427–2442.

More, S J, Cameron, AR, Greiner, M, Clifton-Hadley, R S, Rodeia, SC, Bakker, D, Salman, Mo D, Sharp, J M, De Massis, F, Aranaz, A, *et al.* 2009. Defining output-based standards to achieve and maintain tuberculosis freedom in farmed deer, with reference to member states of the European Union. *Preventive veterinary medicine*, **90**(3-4), 254–267.

Nusinovici, S, Madouasse, A, Hoch, T, Guatteo, R, & Beaudeau, F. 2015. Evaluation of two PCR tests for Coxiella burnetii detection in dairy cattle farms using latent class analysis. *PloS one*, **10**(12), e0144608.

Zucchini, W., MacDonald, I.L., & Langrock, R. 2017. *Hidden Markov Models for Time Series: An Introduction Using R, Second Edition.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press.

# Annexe 5 : Supplementary Information Chapter 4

## 5.1. Table of tests results depending on the simulated herd status

TABLE 6.1 – Probability of test results depending on the simulated herd status. $\hat{Se}$ and $\hat{Sp}$ refer to estimated herd level test sensitivity and specificity. $p\left(\tilde{S}^{+*}_{h,t}\right)$ is the predicted probability of being herd status positive depending on previous predicted status $\left(\hat{S}^{+}_{h,t-1}\right)$.

| | | Herd status at $t$ | |
|---|---|---|---|
| | | + | - |
| Test | + | $\hat{Se}.p\left(\tilde{S}^{+*}_{h,t}\right)$ | $\left(1-\hat{Sp}\right).\left(1-p\left(1-\tilde{S}^{+*}_{h,t}\right)\right)$ |
| | - | $\left(1-\hat{Se}\right).p\left(\tilde{S}^{+*}_{h,t}\right)$ | $\hat{Sp}.\left(1-p\left(1-\tilde{S}^{+*}_{h,t}\right)\right)$ |

## 5.2. Priors distribution for STOC free model parameters

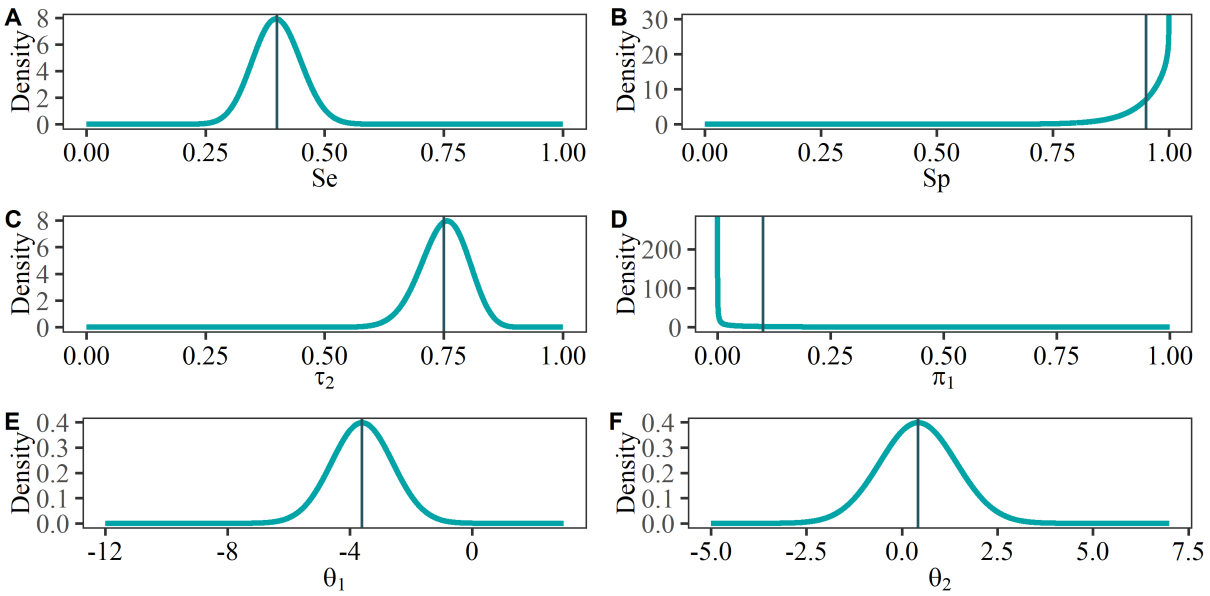



FIGURE 6.1 – Priors distribution for STOC free model parameters. (A) $Se$ (test sensitivity), (B) $Sp$ (test specificity), (C) $\pi_1$ (herd probability of being infected at time step 1), (D) $\tau_2$ (probability to remain infected), (E) $\theta_1$ (logistic regression parameters) and (F) $\theta_2$ (logistic regression parameters) for one scenario, as an example. Solid line represents the chosen parameter value for simulation.

## 5.3. Posterior distribution, simulated population value and the 95% interval of the distribution

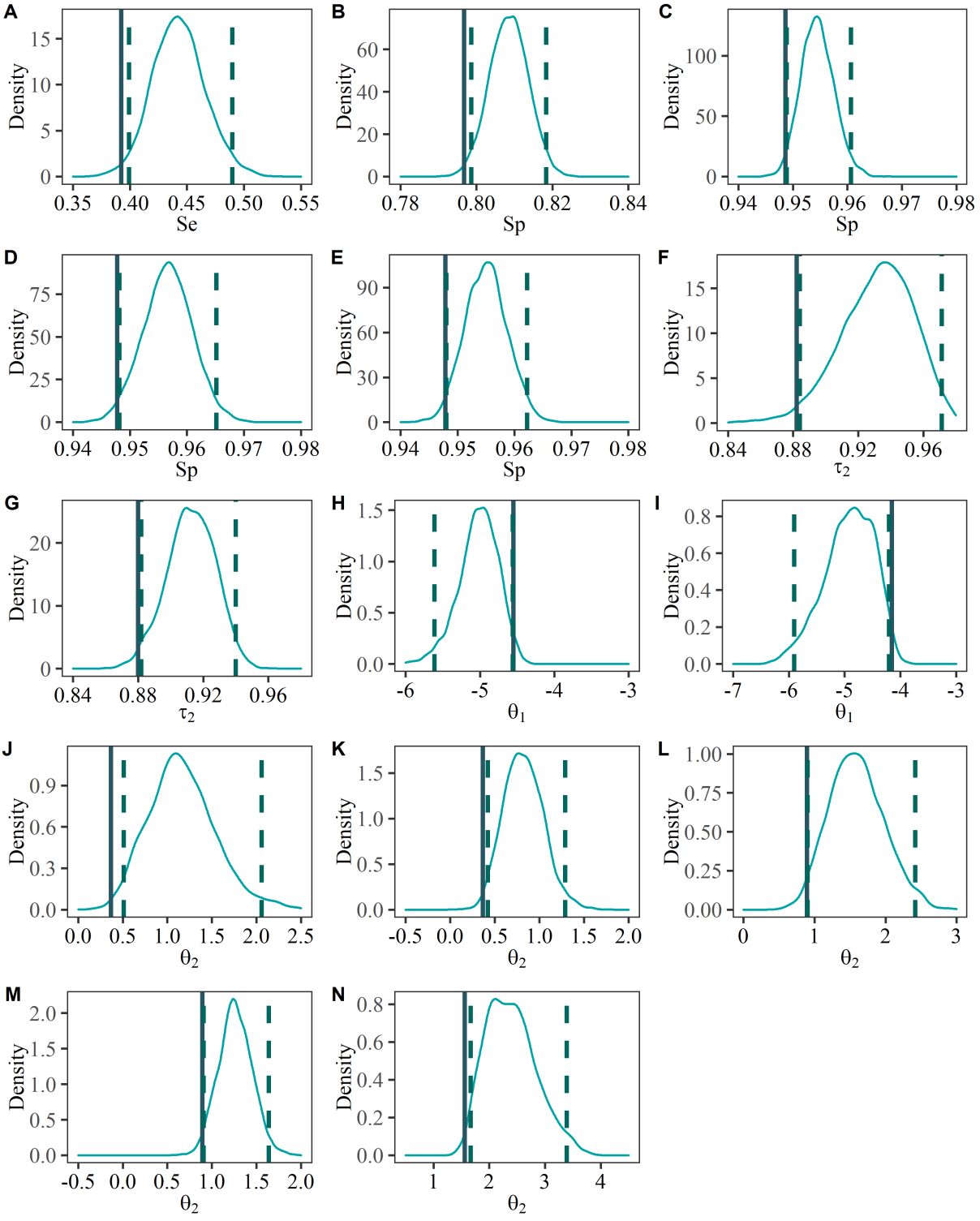



FIGURE 6.2 – Posterior distribution (in blue), simulated population value (in solid line) and the 95% interval of the distribution for the 13 scenarios for which at least one parameter was outside the 95% interval. For only one scenario, parameters simulated value were outside the 95% interval (I and P).

## 5.4. Distribution of cut-off value using Youden index

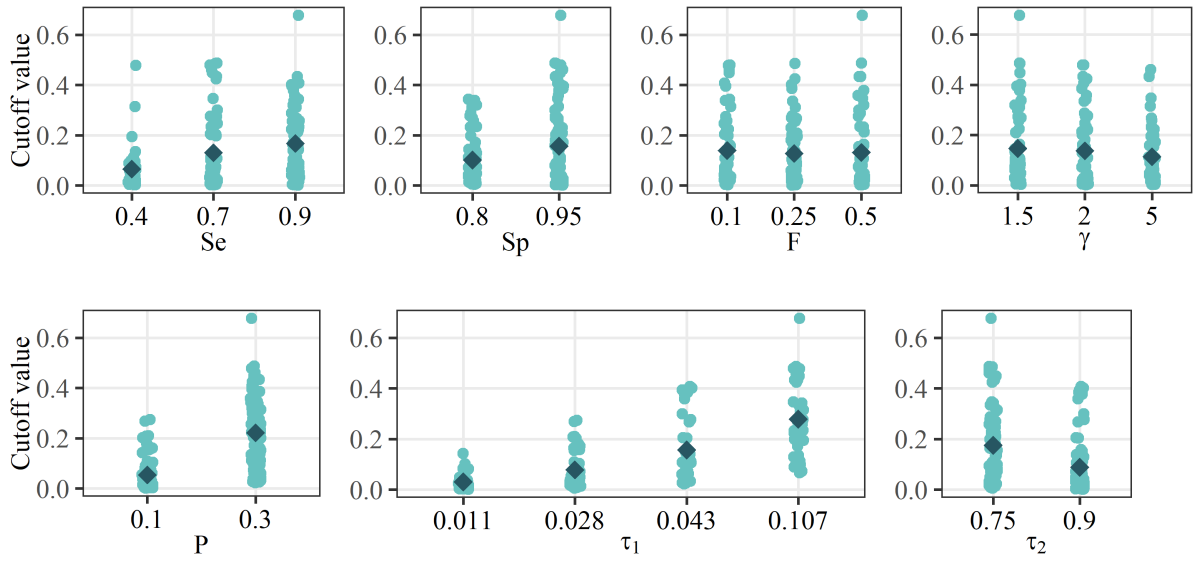



FIGURE 6.3 – Distribution of cut-off value using Youden index for all scenarios depending on simulated parameter values. The seven simulated parameters are represented : $Se$ (test sensitivity), $Sp$ (test specificity), $F$ (frequency of the risk factor), $\gamma$ (relative risk associated with the risk factor), $P$ (prevalence), $\tau_1$ (probability of being newly infected) and $\tau_2$ (probability of remaining infected). Dark blue diamond represents the mean of each distribution.

## 5.5. Distribution of cut-off value using NewI cost index

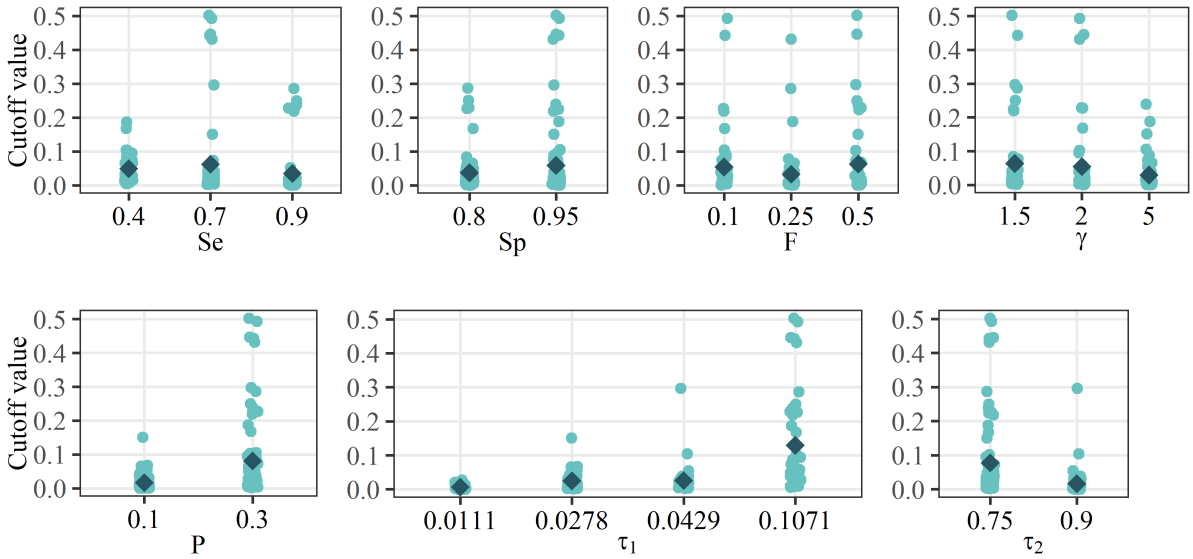



FIGURE 6.4 – Distribution of cut-off value using NewI cost index for all scenario depending on simulated parameter values The seven simulated parameters are represented : $Se$ (test sensitivity), $Sp$ (test specificity), $F$ (frequency of the risk factor), $\gamma$ (relative risk associated with the risk factor), $P$ (prevalence), $\tau_1$ (probability of being newly infected) and $\tau_2$ (probability of remaining infected). Dark blue diamond represents the mean of each distribution.

# Liste des productions scientifiques

## Publications dans des revues à comité de lecture

### En cours de révision

Soumis à *Preventive veterinary medecine*

— **Mercat, M.**, van Roon, A.M., Santman-Berends, I. van Schaik, G., Nielen, M., Graham, D., More, S. J., Guelbenzu-Gonzalo, M., Fourichon, C. and Madouasse, A. 2021. Capacity of a Bayesian model to detect infected herds using disease dynamics and risk factor information from surveillance programmes : A simulation study.

Soumis à *PCI Animal Science*

— Madouasse, A.,**Mercat, M.**, van Roon, A., Graham, D., Guelbenzu, M., Berends, I.S., van Schaik, G., Nielen, M., Frössling, J., Âgren, E., Humphry, R., Eze, J., Gunn, G., Henry, M., Gethmann, J., More, S. and Fourichon, C., 2020. A modelling framework for the prediction of the herd-level probability of infection from longitudinal data. bioRxiv 1–36. `https://doi.org/10.1101/2020.07.10.197426`

### Articles Publiés

— van Roon, A.M. , **Mercat, M.**, van Schaik, G., Nielen, M., Graham, D., More, S. J., Guelbenzu-Gonzalo, M., Fourichon, C., Madouasse, A. and Santman, I.M.G.A., 2020. Quantification of risk factors for bovine viral diarrhea virus in cattle herds : A systematic search and meta-analysis of observational studies. Quantification of risk factors for bovine viral diarrhea virus in cattle herds : A systematic search and meta-analysis of observational studies. J. Dairy Sci. 103, 9446–9463.

— van Roon, A.M., Rapaliute, E., Koleci, X., Muñoz, V., **Mercat, M.**, Faverjon, C., Santman-Berends, I.M.G.A., Nielen, M., More, S.J., Graham, D., Guelbenzu-Gonzalo, M., Fourichon, C., Madouasse, A. and van Schaik, G., 2021. Key Learnings During the Development of a Generic Data Collection Tool to Support Assessment of Freedom of Infection in Cattle Herds. Frontiers in veterinary science. 8, 395.
https://doi.org/10.3168/jds.2020-18193

— van Roon, A.M., Santman-Berends, I.M.G.A., Graham, D., More, S.J., Nielen, M., van Duijn, L., **Mercat, M.**, Fourichon, C., Madouasse, A., Gethmann, J., Sauter-Louis, C., Frössling, J., Lindberg, A., Correia-Gomes, C., Gunn, G.J., Henry, M.K. and van Schaik, G., 2020b. A description and qualitative comparison of the elements of heterogeneous bovine viral diarrhea control programs that influence confidence of freedom. J. Dairy Sci. 103, 4654–4671.
https://doi.org/10.3168/jds.2019-16915

— van Roon, A. M., Santman-Berends, I. M. G. A., Graham, D., More, S. J., Nielen, M., Madouasse, A., **Mercat, M.**, Fourichon, C. Gethmann, J., Frössling, J., Lindberg, A., Correia-Gomes, C., Gunn, G. J., Sauter-Louis, C., Henry, M. K., van Duijn, L. and van Schaik, G.,. 2019. STOC Free : An innovative framework to compare probability of freedom from infection in heterogeneous control programmes. Front. Vet. Sci. 6 :133.
https://doi.org/10.3389/fvets.2019.00133

## Rapports techniques

— **Mercat, M.**, van Roon, A., Santman, I., van Schaik, G., Nielen, M., van Duijn, L., More, S. J., Graham, D., Frossling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., Fourichon, C., & Madouasse, A. 2018. *STOC free : WP1, Deliverable 1.1Guidelines for the design of conceptual models representing the infectious process at different levels, from animal to region, with an application to BVD. techreport*. EFSA. 32p.

— **Mercat, M.**, van Roon, A., Santman, I., van Schaik, G., Nielen, M., van Duijn, L., More, S. J., Graham, D., Frossling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., Fourichon, C., & Madouasse, A. 2018. *STOC*

*free : WP1 and WP2, Deliverable 1.2. Guidelines for the identification of sources of the data available to quantify the confidence of freedom from infection, with an application to BVDV*. techreport. EFSA. 26p.

— Madouasse, A, **Mercat, M.**, van Roon, A., Santman, I., van Schaik, G., Nielen, M., van Duijn, L., More, S. J., Graham, D., Frossling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., & Fourichon, C. 2018. *STOC free : WP1 1.3Description of the STOC free model*. techreport. EFSA. 39p.

## Congrès internationaux

### Présentations orales

— van Roon, A., Santman I., Graham, D., More, S., Nielen1, M., van Duijn, L., **Mercat, M.**, Fourichon, C., Madouasse, A., Gethmann, J., Sauter-Louis, C., Frössling, J., Lindberg, A., Gomes, C., Gunn, G., Henry, M., van Schaik, G. Qualitative comparison of BVDV control programmes in Europe to substantiate freedom from infection. ISVEE. November, 12-16$^th$, Chiang Mai, Thailand.

— Santman, I., Madouasse, A., **Mercat, M.**, Fourichon, C., More, S. J., Graham, D., Frössling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., van Roon, A., Nielen, M., van Duijn, L., van Schaik, G. An innovative surveillance analysis tool for outcome-based comparison of freedom from infection in heterogeneous control program. November, 12-16$^th$, Chiang Mai, Thailand.

### Présentations de posters

— **Mercat, M.**, van Roon, A., Santman, I., Nielen, M., van Duijn, L., van Schaik, G., More, S. J., Graham, D., Frössling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., Fourichon, C., Madouasse, A. Using heterogeneous information for the estimation of the probability of freedom from infection with BVD virus : A conceptual model mapping information onto infection biology. SVEPM. March 21-23$^th$ 2019, Tallin, Estonia.

— van Roon, A., Santman, I., Graham, D., More, S., Nielen, M., van Duijn, L., **Mercat, M.**, Fourichon, C., Madouasse, A., Gethmann, J., Sauter-Louis, C., Frössling, J., Lindberg, A., Gomes, C., Gunn, G., Henry, M., A. van Schaik, G. Chal-

lenges when comparing disease control programmes. SVEPM. March 27-29$^th$ 2019, Utrecht, Netherlands.

— Santman, I., Madouasse, A., **Mercat, M.**, Fourichon, C., More, S. J., Graham, D., Frössling, J., Lindberg, A., Gethmann, J., Sauter-Louis, C., Gunn, G., Gomes, C., Henry, M., van Roon, A., Nielen, M., van Duijn, L., van Schaik, G. Development of a Surveillance analysis Tool for Outcome-based Comparison of the confidence of freedom generated by control programmes. SVEPM. March 21-23$^th$ 2019, Tallin, Estonia.

DOCTORAT ECOLOGIE
BRETAGNE GEOSCIENCES
LOIRE AGRONOMIE ALIMENTATION

Oniris
École Nationale Vétérinaire, Agroalimentaire et de l'Alimentation
Nantes Atlantique
Santé et alimentation au cœur de la vie

**Titre :** Conception et évaluation d'une méthode d'estimation d'une probabilité d'infection d'un troupeau à partir de données hétérogènes : contribution au développement d'une surveillance épidémiologique basée sur la comparabilité des résultats

**Mot clés :** épidémiologie, bovins, surveillance *output-based*, absence d'infection, programme de maîtrise, données longitudinales

**Résumé :** A l'échelle de territoires, les programmes collectifs de maîtrise des maladies infectieuses non règlementées des bovins présentent de multiples bénéfices. Ils créent aussi des difficultés dans les échanges entre territoires car leurs définitions du statut « indemne d'infection » diffèrent. Estimer une probabilité (d'absence) d'infection pour chaque troupeau, calculée indépendamment des modalités de surveillance, permettrait de sécuriser le commerce d'animaux entre territoires. Ce type d'estimation pourrait servir à une surveillance dite output-based, basée sur un résultat à atteindre et non sur les moyens mis en œuvre. Les objectifs de cette thèse étaient de contribuer à l'élaboration puis d'évaluer une méthode d'estimation de probabilité d'infection à l'échelle du troupeau, à partir de données de surveillance hétérogènes. En prenant l'exemple de l'infection par le virus de la diarrhée virale bovine, les informations pertinentes et disponibles ont été identifiées et organisées. Le modèle développé est un modèle de Markov caché estimant une probabilité d'infection à l'échelle du troupeau à partir de résultats de test et de facteurs de risque d'infection. Ses performances ont été évaluées sur des données simulées, représentant une diversité de dynamiques d'infection et de programmes de maîtrise. Il ressort de l'évaluation que la valeur ajoutée du modèle est d'autant plus importante que la sensibilité du test de diagnostic est faible. La valeur ajoutée des facteurs de risque semble limitée. L'utilisation de ce modèle requiert des développements supplémentaires pour la classification des troupeaux en indemne / infecté à partir des probabilités d'infection prédites.

**Title:** Design and evaluation of a method for the estimation of a herd-level probability of infection from heterogeneous data : contribution to the development of output-based surveillance

**Keywords:** epidemiology, cattle, output based surveillance, freedom from infection, disease control programme longitudinal data

**Abstract:** On a territorial scale, programmes to control non-regulated infectious diseases of cattle have multiple benefits. They also create difficulties in exchanges between territories because the definitions of 'infection-free' status differ between programmes. Estimating a probability (of absence) of infection for each herd calculated independently of the surveillance modalities would make it possible to secure trade in animals between territories. This type of estimate could be used for output-based surveillance, a type of surveillance based on a result to be achieved and not on the means implemented. The objectives of this thesis work were to contribute to the development and evaluation of a method for estimating infection probabilities at the herd level, based on heterogeneous surveillance data. Using the example of bovine viral diarrhea virus infection, relevant and available information was identified and organised. The model developed is a hidden Markov model estimating a probability of infection at the herd level from repeated test results and risk factors for infection. Its performance was evaluated on simulated data representing a variety of infection dynamics and control programmes. The evaluation showed that the added value of the model was greater the lower the sensitivity of the diagnostic test. The added value of the risk factors was moderate in the range of situations evaluated. The use of this model requires further development for the classification of herds as free/infected based on predicted infection probabilities.