



HAL
open science

Utilisation des analyses en mélanges pour l'évaluation et le suivi du statut sanitaire des troupeaux : application à la paratuberculose des ovins

Yoann Mathevon

► To cite this version:

Yoann Mathevon. Utilisation des analyses en mélanges pour l'évaluation et le suivi du statut sanitaire des troupeaux : application à la paratuberculose des ovins. Médecine vétérinaire et santé animale. Université Paul Sabatier - Toulouse III, 2018. Français. NNT : 2018TOU30044 . tel-02089274

HAL Id: tel-02089274

<https://theses.hal.science/tel-02089274>

Submitted on 3 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *12 avril 2018 (12/04/2018)* par :

Yoann MATHEVON

**Utilisation des analyses en mélanges pour l'évaluation et le
suivi du statut sanitaire des troupeaux :
application à la paratuberculose des ovins**

JURY

MARIE LAURE
DELIGNETTE-MULLER Professeur des Universités Présidente du jury

VIRGINIE THIBAUT Docteur Membre du jury

École doctorale et spécialité :

SEVAB : Pathologie, Toxicologie, Génétique et Nutrition

Unité de Recherche :

UMR Interactions Hôtes Agents Pathogènes 1225 INRA-ENVT

Directeur(s) de Thèse :

Pr. Gilles FOUCRAS et Dr. Fabien CORBIERE

Rapporteurs :

Pr. Raphaël GUATTEO et Pr. Emmanuelle GILLOT-FROMONT

Remerciements

En tout premier lieu je souhaite remercier chaleureusement mes directeurs de thèse. Merci à M. Gilles Foucras pour m'avoir fait découvrir les subtilités des analyses de laboratoire ainsi que pour l'aide apportée tout au long de ces trois ans de thèse. Merci à M. Fabien Corbière, qui ne m'a pas oublié après ma thèse vétérinaire, ce qui a donné lieu à cette coopération composée de prise de tête sur des lignes de codes et de canalisation de mon travail lorsque l'entêtement prenait le dessus. Un grand merci à vous deux.

Mes remerciements s'adressent à Mme. Emmanuelle Gillot-Fromont et M. Raphael Guatteo pour avoir accepté d'être les rapporteurs de ma thèse. Ils m'ont fait l'honneur d'être aussi parmi le jury pour juger mon travail. Je voudrais remercier de même Mme. Marie Laure Derlignette Muller pour avoir accepté d'être la présidente du jury, et aussi Mme Virginie Thibault pour avoir accepté d'être une examinatrice dans ce jury. J'apprécie leurs contributions pour rendre ma présentation plus intéressante.

Je souhaite remercier les éleveurs et le personnel du groupe Ovilot, sans qui cette thèse n'aurait pas eu lieu et en particulier M. Remy Falguières pour la coordination lors de la réalisation des prélèvements.

Bien évidemment je souhaite remercier mon équipe d'accueil EPIDEC, de l'UMR IHAP de l'INRA. Pour le support mutuel en stats, pour les opportunités d'enseignement que vous m'avez confiées et pour la bonne ambiance du bureau des "jeunes". En espérant travailler/collaborer à nouveau avec vous un jour!

Toujours sur le site de l'ENVT, je tiens à remercier les personnes avec qui j'ai pu travailler/échanger : les doctorants de la virologie (Ellias et Guillaume), Christian, Guillaume et Sarah pour les manip en immunologie ainsi que tous les autres doctorants du site pour les séminaires plus ou moins mensuels.

De manière plus personnelle je tiens à remercier les habitants de Yalma pour ces trois ans au bord du Touch. Les amis de Lyon et Toulouse, maintenant disséminés aux quatre coins de la France et du monde, cela devient difficile de vous voir sur une base régulière mais c'est à chaque fois des retrouvailles de qualités.

Du côté de la famille, bien évidemment mes parents, qui ont cru en moi depuis que je veux être vétérinaire (ce qui remonte à ma prime enfance) et qui ont continué même une fois cette étape atteinte. A mon frère Kevin et Clothilde qui ont eu la bonne idée de migrer à Toulouse nous permettant de se voir bien plus souvent qu'avant, et ça c'était cool! A ma soeur Tiphaine et Baptiste et bien sur Camille ma nièce. Je suis un tonton un peu loin (et ça ne va pas s'arranger dans les 6 prochains mois) mais tu sais que je ferais baby-sitter dès que je le pourrais tellement tu as fait un beau boulot avec cette petite! A mes grands-parents, qui sont tellement fiers de moi, cela me touche énormément. Je repasserais vous voir à chaque fois que mes tribulations me ramènent dans la Loire.

La famille encore et toujours avec mon parrain (Yvan) et ma marraine (Marie-Cécile), mes presque-cousins de Haute-Loire (les Magat).

A mamie Chouchou...

Et enfin, cette thèse est dédiée de A à Z à Sarah. Pour notre histoire et pour nos projets, que notre futur soit aussi brillant que notre passé.

Table des matières

Table des matières	i
Liste des figures	iii
Liste des tableaux	v
Liste des sigles et des abréviations	vii
1 <i>Mycobacterium avium</i> ssp <i>paratuberculosis</i> et paratuberculose : éléments de bibliographie	1
1 <i>Mycobacterium avium</i> ssp <i>paratuberculosis</i> , agent de la paratuberculose . . .	1
1.1 Taxonomie et biologie	2
1.2 Paratuberculose : éléments de pathogénie	7
2 Épidémiologie	11
2.1 Épidémiologie descriptive	11
2.2 Épidémiologie analytique	14
2.3 Importance de la paratuberculose	19
3 Tests pour le diagnostic individuel de la paratuberculose	24
3.1 Méthodes directes	25
3.2 Méthodes indirectes	33
4 Performances diagnostiques des tests biologiques	39
4.1 Indicateurs des performances des tests	39
4.2 Évaluation des performances absolues en l'absence de test parfait . . .	43
4.3 Performances diagnostiques des tests à l'échelle individuelle	49
4.4 Diagnostic à l'échelle du troupeau	53
5 Prévention et maîtrise	64
5.1 Mesures préventives	64
5.2 Mesures de maîtrise	65
5.3 Sélection génétique	76
6 Conclusions du chapitre	79
2 Contexte et objectifs des travaux de thèse	83
1 Paratuberculose et race ovine Causse du Lot	83
2 Objectifs de la thèse	85
3 Évaluation de la sensibilité et de la spécificité de deux trousse de sérologie ELISA et d'une qPCR sur fèces pour le dépistage de la paratuberculose à l'échelle individuelle	89
1 Article 1 : estimation par une approche bayésienne à classes latentes	90
2 Article 2 : estimation par une approche fréquentiste	112
3 Conclusions du chapitre	150

4	Stratégies de dépistage à l'échelle collective	153
1	Article 3 : évaluation des performances en mélanges d'échantillons	154
2	Compléments : analyses de mélanges d'échantillons	178
2.1	Analyse par qPCR de mélanges de fèces	178
2.2	Analyse par sérologie ELISA de mélanges de sérums	181
3	Conclusions du chapitre	186
5	Maîtrise de la paratuberculose par la vaccination	187
1	Effets de la vaccination vis-à-vis de la paratuberculose en élevages ovins al- laitants : premiers résultats en France	187
2	Suivi de la paratuberculose en contexte vacciné	192
3	Conclusions du chapitre	195
6	Discussion générale et perspectives	197
1	Estimation des performances diagnostiques à l'échelle individuelle	198
1.1	Méthodes d'estimation et principaux résultats	198
1.2	Résultats comparés entre troupes ELISA et qPCR sur fèces	202
2	Estimation des performances relatives à l'échelle des mélanges	206
2.1	Performances relatives des troupes ELISA appliquées à des mélanges de sérums	207
2.2	Performances relatives de la qPCR dans des mélanges de fèces	209
3	Définition et suivi du statut sanitaire sur la base de résultats de mélanges d'échantillons	211
4	Application au contexte des troupeaux vaccinés	214
	Bibliographie	221
	Annexes	253
I	Liste des communications	253
II	High throughput qPCR analysis : Application to the detection of Mycobac- terium Paratuberculosis in sheep	257
III	Vaccination against Paratuberculosis in sheep : methodological developments and evaluation of efficacy	259

Liste des figures

1	Représentation de la diversité génétique au sein du complexe <i>Mycobacterium avium</i>	3
2	Schéma évolutif à partir du séquençage par Shotgun optical mapping	4
3	Représentation schématique de l'évolution de l'infection par <i>Map</i> après ingestion d'après Arsenault et al. (2014)	8
4	Représentation schématique de la réponse immunitaire dite "classique" avec l'évolution de la paratuberculose	11
5	Carte de la répartition mondiale de la paratuberculose au premier semestre 2016, (OIE, 2017)	13
6	Carte de France de l'estimation de la prévalence de la paratuberculose ovine à l'échelle des départements	15
7	Statut de 20 ovins exposés à <i>Map</i> au cours d'un suivi de deux ans et demi (de Silva et al., 2018)	22
8	Courbe d'amplification caractéristique de qPCR avec les différentes définitions du cycle de sortie.	29
9	Description d'un essai phage basé sur l'amplification des phages (Merwe et al., 2014)	32
10	Variation d'expression de gènes codant pour différentes cytokines en relation avec différentes formes de la maladie chez les ovins (Smeed et al., 2007)	37
11	Représentation des quatre statuts selon la classification proposée par Nielsen and Toft (2008)	51
12	Représentation de la variation de la sensibilité à l'échelle d'un l'élevage Se_T en fonction du nombre d'animaux testés, de la prévalence réelle, et des performances intrinsèques du test utilisé à l'échelle individuelle.	55
13	Variation de la spécificité à l'échelle d'un élevage Sp_T , en fonction du nombre d'animaux testés et de la spécificité du test utilisé à l'échelle individuelle.	56
14	Représentation graphique des évaluations réalisées sur l'efficacité de la vaccination, regroupées par résultat en fonction de l'espèce animale et des effets étudiés (production, épidémiologie ou pathogénicité) (d'après Bastida and Juste (2011)).	72
15	Suivi sérologique et $IFN\gamma$ d'ovins vaccinées et non vaccinés dans trois fermes	74
16	Schéma de production et de croisement dit "en double étage" de la race Causse du Lot.	84
17	Schéma des relations entre les objectifs de la thèse.	87
18	Plan d'analyse pour l'analyse de la Se_D de la qPCR en mélange de fèces	179
19	Résultats des analyses de qPCR sur des mélanges de fèces	181

20	Résultats des dilutions sériées pour le kitID Screen Paratuberculosis Indirect®	183
21	Résultats des analyses en dilutions sériées pour la trousse IDEXX paratuberculosis screening®	184
22	Nombre de mélanges de fèces détectés par qPCR après un suivi annuel d'une durée de trois ans et de 50 animaux (10 mélanges de cinq échantillons) . . .	193
23	Nombre de mélanges de fèces détectés par qPCR après un suivi annuel de cinq ans sur 50 (10 mélanges de cinq échantillons) ou 100 animaux (10 mélanges de 10 échantillons)	194
24	Statut des mélanges de taille cinq et 10 analysés en ELISA en fonction de la prévalence d'infection	213
25	Relation entre séroprévalence apparente et prévalence apparente d'excrétion fécale dans 14 troupeaux ovins non vaccinés du Lot	215

Liste des tableaux

1	Synthèse des méthodes les plus utilisées pour diagnostic ou le dépistage de la paratuberculose	38
2	Table de contingence des résultats d'un test	40
3	Caractéristiques des vaccins contre la paratuberculose	71
4	Réduction de l'expression clinique, de l'excrétion ou des aspects lésionnels obtenus après vaccination contre la paratuberculose (d'après Bastida and Juste (2011)).	72
5	Résultats de l'analyse qPCR sur fèces des animaux choisis pour l'évaluation de la sensibilité de la qPCR ADIAVET™ paraTB Real Time (BioX) appliquée à des mélanges de fèces.	180
6	Classes de sérum en fonction de la valeur S/P	182
7	Résultats des analyses de mélanges de sérums de taille 5 (trousse ID Screen Paratuberculosis Indirect®, lot 800, IdVet)	185
8	Résultats des analyses de mélanges de sérums de taille 10 (trousse ID Screen Paratuberculosis Indirect®, lot 800, IdVet)	186
9	Résultats des analyses de mélanges de fèces en phase liquide par qPCR . . .	210

Liste des sigles et des abréviations

ACERSA : Association pour la certification en santé animale	LP : Low positive
ADN : Acide désoxiribo-nucléique	Map : <i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i>
AEEMA : Association pour l'étude de l'épidémiologie des maladies animales	OIE : Office international des épizooties
AGID : Agar gel immunodiffusion	OMS : Organisation mondiale de la santé
AIC : Aikake information criterion	OR : Odds ratio
AMM : Autorisation de mise sur le marché	PA : Prevalence apparente
ANMV : Agence nationale du médicament vétérinaire	PCI 95% : Intervalle de crédibilité de la distribution à <i>posteriori</i> à 95%
BIC : Bayesian information criterion	PCR : Réaction en chaîne par polymérase
Ct : Cycle threshold	pH : Potentiel hydrogène
CV : Coefficient de variation	PICSAR : Paratuberculosis control In Cattle and Sheep : innovative Approach based on biosecurity and genetic Resistance
ddl : Degrés de liberté	ppm : partie par million
DIC : Deviance information criterion	qPCR : Réaction en chaîne par polymérase quantitative
ELISA : Enzyme linked immunosorbent assay	SDRP : Syndrome dysgénésique et respiratoire du porc
EM : Expectation-maximization	Se_A : sensibilité analytique
FN : Faux négatif	Se_D : sensibilité diagnostique
FP : Faux positif	Se_T : Sensibilité diagnostique de troupeau
GDS : Groupement de défense sanitaire	SNGTV : Société nationale des groupements techniques vétérinaires
GISA : Gestion intégrée de la santé des animaux	SNP : Single-nucleotide polymorphism
HCl : Acide chloridrique	Sp_A : spécificité analytique
HN : High negative	Sp_D : spécificité diagnostique
HP : High positive	Sp_T : spécificité diagnostique de troupeau
IBR : Infectious bovine rhinotracheitis	S/P : Sample-to-positive
IC 95% : Intervalle de confiance à 95%	TLR : Toll-like receptor
ICP : International colloquium on paratuberculosis	TNF : Tumor necrosis factor
IDT : Intradermotuberculation	UE : Union européenne
IFAT : Immunofluorescence antibody test	UFC : Unité formant colonie
IFNγ : Interferon γ	VBJDCP : Voluntary bovine Johne's disease control program
IHAP : Interactions hôtes agents pathogènes	VN : Vrai négatif
IL : Interleukine	VP : Vrai positif
INRA : Institut national de la recherche agronomique	VPN : Valeur prédictive négative
LN : Low negative	VPP : Valeur prédictive positive

Chapitre 1

Mycobacterium avium ssp *paratuberculosis* et paratuberculose : éléments de bibliographie

Harmlessly passing your time in the grassland
away
Only dimly aware of a certain unease in the air

Roger Waters

Sheep; Animals

1 *Mycobacterium avium* ssp *paratuberculosis*, agent de la paratuberculose

La littérature scientifique relative à la paratuberculose est extrêmement vaste et s'enrichit tous les jours de nouvelles références (3759 articles référencés avec les mots clés [Paratuberculosis] ou [Johnes' disease] dans la base PudMed entre 1936 et le 31 décembre 2017, dont 1345 depuis 2010). Ce premier chapitre n'a pas l'ambition de faire une synthèse exhaustive des connaissances acquises à ce jour sur cette maladie, mais de fournir au lecteur quelques éléments de biologie, de physiopathologie et d'épidémiologie utiles à la compréhension de la question posée. Nous nous intéressons ensuite plus spécifique-

ment au diagnostic de la paratuberculose, tant à l'échelle individuelle que collective, et aux difficultés méthodologiques que représente l'évaluation des performances diagnostiques des tests biologiques disponibles.

1.1 Taxonomie et biologie

1.1.1 Historique

La paratuberculose a été décrite pour la première fois en 1895 par John et Frothingham comme une forme particulière de tuberculose chez une vache caractérisée par une entérite chronique avec une diarrhée intermittente. L'autopsie révèle un épaissement de la paroi muqueuse ainsi que des lésions avec des bacilles acido-résistants, justifiant le lien potentiel avec la tuberculose. En 1906, Bang propose le nom de paratuberculose après avoir remarqué que les animaux infectés ne réagissent pas aux antigènes de tuberculine mais à une préparation similaire à partir de bacille tuberculeux aviaire. Le premier isolement de l'agent pathogène est attribué à Twort qui, après avoir observé la croissance de colonies en périphérie d'une culture contaminée par *Mycobacterium phlei* parvient à cultiver le bacille responsable de la paratuberculose par ajout d'un extrait de *Mycobacterium phlei* au milieu de culture (Clarke, 1997 ; Chiodini, 1992 ; Twort and Ingram, 1912).

1.1.2 Taxonomie et enveloppe mycobactérienne

Mycobacterium avium subsp. *paratuberculosis* (*Map*) est une bactérie immobile aérobie classée Gram positive, bien que la coloration ne soit pas très efficace. Elle fait partie de la famille des Mycobacteriaceae, qui sont notamment caractérisées par leur enveloppe épaisse composée de trois couches (Manning and Collins, 2001). De l'intérieur vers l'extérieur, la première couche est la membrane plasmique similaire à celle des autres bactéries, puis d'une paroi riche en lipide et enfin d'une capsule. Cette enveloppe protège de l'action des antibiotiques, des acides, de la phagocytose et améliore la survie dans le milieu extérieur. Il en résulte également une résistance aux acides et alcools, rendant la coloration de Ziehl-Neelsen possible (Lefrançois, 2012). *Map* fait partie du complexe *Mycobacterium avium* regroupant les espèces *M. avium* et *M. intracellulare*.

1.1.3 Structure génétique

Le génome de *Map* est un acide désoxiribo-nucléique (ADN) circulaire de 4,8 millions de paires de bases codant pour environ 4300 à 4700 cadres de lecture ouverte (Li et al., 2005), ce qui représente environ 91,5% de son génome. L'espèce *M. avium* est divisée en quatre sous-espèces, la plus ancienne étant *M. avium hominissuis* à partir de laquelle ont évoluées d'une part *M. avium avium* et *M. avium sylvaticum*, plus proches phylogénétiquement, et d'autre part *M. avium paratuberculosis*, comme nous pouvons le voir dans la figure 1 (Turenne et al., 2008).

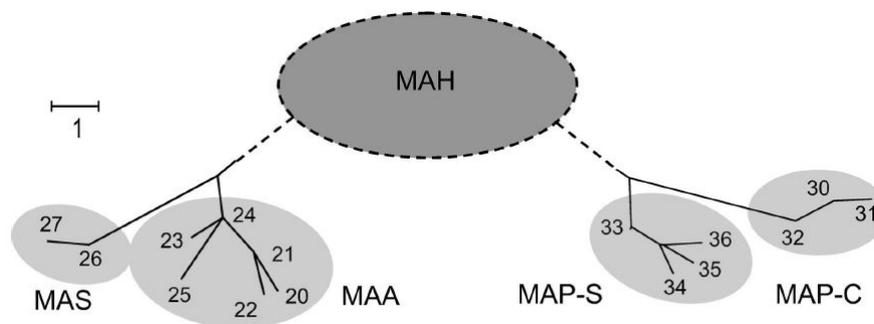


FIGURE 1 – Représentation de la diversité génétique au sein du complexe *Mycobacterium avium*.

A partir d'une analyse de plusieurs séquences d'ADN, Turenne montre la variabilité génétique de plusieurs souches du complexe *Mycobacterium avium*

(MAH) *M. avium hominissuis* (MAS) *M. avium sylvaticum* (MAA) *M. avium avium* (MAP-S) *Map* souche S (MAP-C) *Map* souche C

L'échelle représente une différence d'un nucléotide.

La comparaison du génome de *Map* avec celui de *Mycobacterium avium* et de *Mycobacterium bovis* a permis d'identifier la mutation du complexe de gènes responsable de la sécrétion de mycobactine, un complexe lipidique nécessaire à la culture de certaines mycobactéries inopérant chez *Map*. Le gène *mbtA*, initiant la production de mycobactine chez les mycobactéries, est plus court chez *Map* en comparaison de celui des deux autres mycobactéries étudiées (Li et al., 2005).

Bannantine et al. (2012) ont comparé par séquençage la souche K 10 de *Map*, infectant les bovins et la souche S397 isolée à partir d'ovins. Nous pouvons retenir deux informations principales : la présence d'environ 70 cadres de lecture ouverts supplémentaires chez S397 par rapport à K10, dont 30 coderaient pour des protéines pouvant expliquer une adaptation à l'hôte. Deuxièmement, par l'analyse des inversions de séquences entre

K10, S397 et *Mycobacterium avium* subsp *hominisuis*, les auteurs retracent l'évolution des différentes souches de *Map* comme nous pouvons le voir dans la figure 2 .

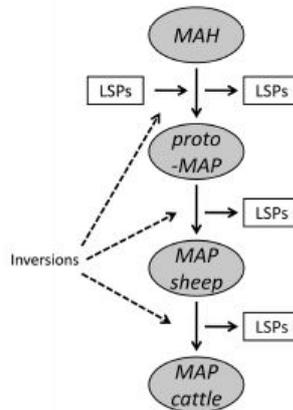


FIGURE 2 – Schéma évolutif à partir du séquençage par Shotgun optical mapping

A partir de souche C et S de *Map*, ainsi que d'une souche de *Mah*, Bannantine et al. (2012) ont analysé les différences de grandes séquences du génome (LSP) spécifiques de chaque souche ainsi que les phénomènes d'inversions du génome observés.

Le génotypage de souches de *Map* a mis en évidence plusieurs séquences d'insertions spécifiques. Li et al. (2005) ont identifié 19 séquences d'insertion répétées 58 fois sur l'ensemble du chromosome. Pour la souche K10 de *Map*, 17 copies d'*IS900* sont retrouvées, contre trois copies pour *ISMav2*.

1.1.4 Variabilité des souches

La sous-espèce paratuberculosis peut être séparée en deux types par hybridation/restriction d'ADN (Collins et al., 1990). Le type C (pour Cattle) a été isolé à partir de souches bovines en opposition au type S (pour Sheep) isolé à partir de souches ovines. Cette distinction était suspectée suite aux observations phénotypiques telles que la vitesse de croissance ainsi que par l'analyse de données épidémiologiques. Le type C se caractérise par une croissance plus rapide (trois à quatre semaines) et un spectre d'hôtes large (Pavlík et al., 1995). Le type S quant à lui à une croissance lente (16 à 20 semaines de culture) et un spectre d'hôtes autrefois considéré comme limité aux petits ruminants. Cependant, des études récentes ont montré une contamination croisée possible par une souche de type S entre ovins et bovins (Verdugo et al., 2014b). Fernández et al. (2014) ont inoculé l'équivalent de 2×10^9 bactéries de souches C et S par voie orale à des agneaux de un mois et

deuxièmes (12 animaux par groupes, ainsi que six témoins non inoculés). Dans chaque groupe, quatre animaux ont été euthanasiés 150 jours après l'infection et le reste des animaux a été sacrifié 390 jours post-infection. Une analyse histologique des tissus de l'intestin a été réalisée suite aux autopsies. Chez neuf des agneaux infectés (cinq pour la souche C, quatre pour la souche S), des lésions microscopiques caractéristiques de la paratuberculose ont été observées. Néanmoins la forme des lésions différait entre les deux groupes : les agneaux infectés par une souche C présentaient des lésions plus marquées 150 jours post-infection et avec une tendance à la rémission à 390 jours post-infection. Les résultats sont concordants avec ceux de l'étude de Verna et al. (2007).

L'infection des bovins par une souche de type S est possible (Stevenson, 2015), même si les études d'infection expérimentale requièrent généralement une plus grande dose d'agent infectieux (Taylor, 1953).

Un troisième type de souches a été décrit par Whittington et al. (2001) après isolement de *Map* provenant de bisons présentant des signes cliniques de paratuberculose. Ce troisième type se différencie par une mutation sur le fragment *IS1311*.

1.1.5 Résistance de *Map*

Comme de nombreuses mycobactéries, *Map* est résistante aux variations des paramètres physiques de son environnement. Les températures extrêmes de survie s'étendent de -70°C sur une période de trois mois (Khare et al., 2008) jusqu'à 72°C pendant 15 secondes, équivalent aux conditions de la pasteurisation du lait (Chiodini and Hermon-Taylor, 1993). Le temps de réduction décimale à 71°C est estimé entre 10 et 14 secondes (Sung and Collins, 1998). Bien qu'aérobie, plusieurs études montrent une tolérance de *Map* aux conditions anaérobies (Parrish et al., 2017), associée à des phénomènes de dormance (Whittington et al., 2004; 2005 ; Eppleston et al., 2014). La plage de pH faiblement acide dans laquelle la croissance de *Map* a été observée s'étend de pH = 5,4 à pH = 7, avec un optimum en dessous de 6,6 (Portaels and Pattyn, 1982).

Ces caractéristiques font de *Map* une bactérie résistante aux conditions extérieures, quel que soit le type C ou S. Whittington et al. (2004) montrent, à l'aide de la culture, une survie possible jusqu'à un an (55 semaines) de *Map* de type S dans des fèces au pâturage. L'humidité

dité ou l'acidité du sol ne semblaient pas influencer la durée de survie, alors que les fèces procurent une protection vis-à-vis des rayons du soleil. Des résultats similaires, avec un degré d'ombrage de 70%, ont été obtenus pour une souche C (Eppleston et al., 2014). La colonisation des végétaux à partir du sol (Whittington et al., 2004 ; Kaevska et al., 2014) a également été montrée avec une survie jusqu'à huit semaines dans les différentes parties du végétal (racine, tige et feuille).

Map peut survivre dans l'eau ainsi que dans les dépôts de sédiments lacustres et fluviaux. Pickup et al. (2005; 2006) ont cultivé *Map* à partir d'échantillons d'eau dans deux rivières écossaises, indiquant la contamination du milieu naturel et ont montré en parallèle une survie possible de *Map* jusqu'à 640 jours dans l'eau en conditions expérimentales. Durant une année, pour un total de 96 jours de collecte, 31 échantillons étaient positifs par réaction en chaîne par polymérase (PCR) (fragment *IS900*) et la culture de 18 de ces échantillons positifs a abouti à 12 isollements. Ces résultats sont en contradiction avec ceux de Whittington et al. (2004) qui ont obtenu des cultures positives de *Map* seulement jusqu'à 250 jours. Ces différences peuvent s'expliquer par le traitement initial du substrat, qui est stérilisé dans les études de Pickup et al., ainsi que par la source de *Map* utilisée pour la contamination. Whittington et al. (2004) ont contaminé l'eau prélevée à partir de fèces, alors que Pickup et al. avaient utilisé un extrait de culture.

Du fait de la présence d'une enveloppe associée à une paroi, *Map* est également résistante aux désinfectants chlorés, avec des différences observées entre trois souches. Kralik et al. (2014) montrent une action plus efficace après une exposition à une concentration élevée de chlore (100 partie par million (ppm)) pendant 30 minutes, répétée sur trois jours, entraînant une réduction de 90% à 95% du nombre de *Map* vivantes dans une solution de concentration initiale de 10^8 bactéries. La réduction observée lors de l'application d'un seul traitement de 30 minutes d'une concentration de chlore de 10 ppm est seulement de 10% à 30% selon la souche testée. Dans cette même étude, *Map* montre une sensibilité à l'acide peracétique à partir d'une concentration de 0,5% en une seule application de 30 minutes ou à partir de 0,05% en deux applications. Il est également rapporté que *Map* est sensible aux composés phénoliques et crésyliques (Collins, 2003).

Concernant la résistance aux antibiotiques évaluée en culture, *Map* est apparue sen-

sible aux macrolides (Krishnan et al., 2009) à partir d'un panel de 10 isolats principalement d'origine humaine (neuf isolats humains, un isolat bovin). Sept des dix isolats testés étaient également sensibles aux fluoroquinolones, les trois restant étant classés comme résistants. Zanetti et al. (2006) ont étudié la sensibilité de 12 isolats de *Map* (six isolats humains, cinq bovins et un isolat provenant de l'American Type Culture Collection) et ont observé une sensibilité de *Map* à la streptomycine, la rifabutine, ainsi qu'à la rifampicine, alors qu'elle est résistante à la dapsonne, classiquement utilisée pour le traitement de la lèpre chez l'homme. Les autres antibiotiques testés indiquent des résultats variés en fonction des isolats de *Map* sur lesquels ils sont appliqués. Nous retiendrons que ces résultats ont été obtenus *in vitro*, et ne sont pas totalement prédictifs d'une efficacité sur le traitement d'une infection par *Map* et concernent essentiellement le traitement de ces infections en médecine humaine. En effet, l'utilisation d'antibiotiques pour le traitement d'une maladie à potentiel zoonotique est déconseillée sur une espèce animale par l'Organisation Mondiale de la Santé (OMS) (OMS, 2017).

1.2 Paratuberculose : éléments de pathogénie

1.2.1 Infection par *Map*

Environ 30 minutes après ingestion, *Map* pénètre dans les plaques de Peyer, plus fréquemment dans la partie iléale de l'intestin, et préférentiellement à travers les cellules M (Momotani et al., 1988 ; Sweeney et al., 2006). Le mécanisme de pénétration dans les cellules M implique la présence de fibronectine, et est facilité si la bactérie est entourée d'une enveloppe de complément (Hostetter et al., 2005). *Map* peut également atteindre les cellules M de l'iléon par infection des glandes salivaires, puis transfert par la voie lymphatique et sanguine. Une fois présente dans la sous-muqueuse, *Map* est phagocytée par les macrophages, impliquant plusieurs familles de récepteurs qui sont à l'origine d'une activation cellulaire faible comparée à celle provoquée par d'autres mycobactéries. La survie de *Map* dans les macrophages se fait soit par blocage de la formation du phagolysosome soit par inhibition de son acidification. L'agent pathogène bloque également l'activation du macrophage, notamment par détournement de plusieurs voies des récepteurs Toll-like

(TLR) (Arsenault et al., 2014). Les macrophages infectés libèrent des cytokines induisant le recrutement d'autres macrophages, ainsi que celui des lymphocytes T qui vont former, en plus de la participation de fibroblastes, un granulome à l'intérieur duquel *Map* se multiplie. Ce granulome constitue la lésion microscopique caractéristique des infections mycobactériennes. La figure 3 synthétise les étapes depuis l'entrée de l'agent pathogène jusqu'à l'établissement du premier granulome.

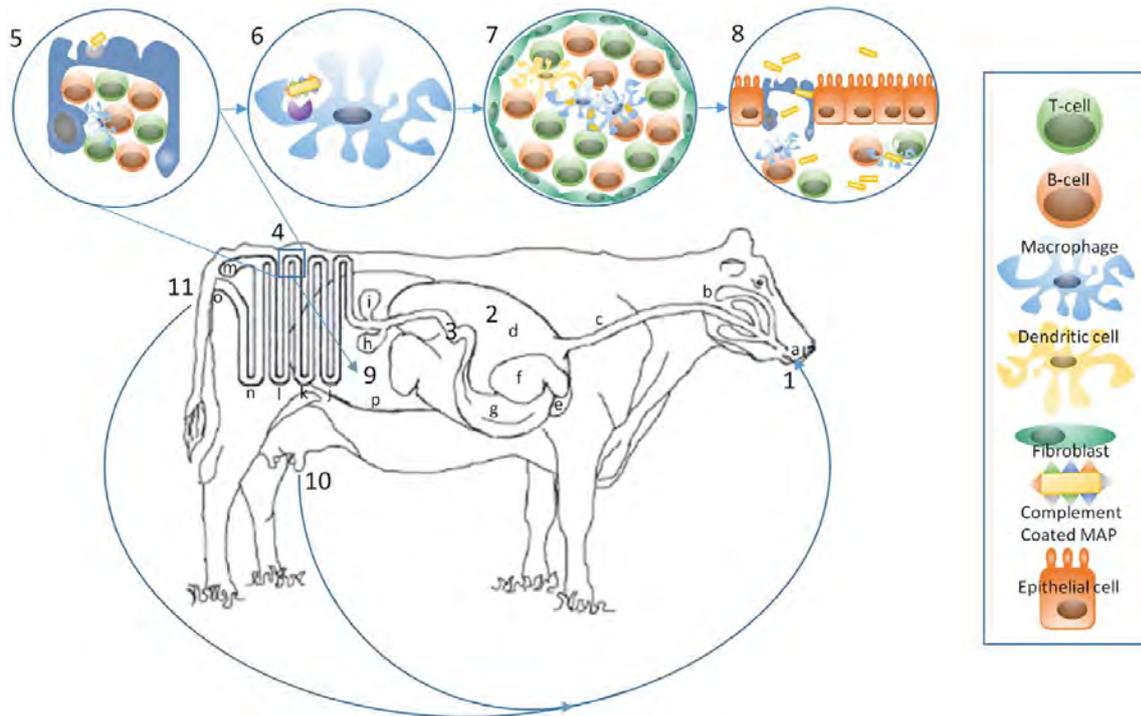


FIGURE 3 – Représentation schématique de l'évolution de l'infection par *Map* après ingestion d'après Arsenault et al. (2014)

(1) *Map* est ingérée et est entraînée dans le tractus intestinal (2) Dans le rumen, *Map* est activée et opsonisée avant d'entrer dans le tube digestif inférieur (3). Dans l'iléon (4) *Map* est captée par les cellules M des plaques de Peyer (5) puis traverse l'épithélium et est prise en charge par les macrophages, reconnaissant *Map* enveloppée dans du complément (6). Les macrophages infectés forment un granulome hébergeant *Map* (7). Pendant les phases actives de paratuberculose les *Map* sont excrétées dans la lumière de l'intestin (8) et peuvent coloniser le foetus (9), les glandes mammaires (10) et être excrétées dans l'environnement (11).

(a) Bouche (b) Glande salivaire (c) Œsophage (d) Rumen (e) Reticulum (f) Omasum (g) Abomasum (h) Vésicule biliaire (i) Pancréas (j) Duodenum (k) Jejunum (l) Ileum (m) Caecum (n) Gros intestin (o) Anus (p) Utérus

1.2.2 Établissement de l'infection et développement de la maladie

Map entame sa réplication dans le macrophage jusqu'à la lyse de ce dernier au bout de quelques semaines, en fonction de la quantité initiale de bactéries phagocytées par le

macrophage. Les bactéries ainsi libérées se retrouvent dans le liquide interstitiel et ont plusieurs devenir possibles :

- Excrétion dans la lumière de l'intestin;
- Captation par d'autres macrophages.

Les macrophages circulant ayant capté des mycobactéries vont soit engendrer de nouvelles lésions à proximité du premier granulome, soit rejoindre le nœud lymphatique drainant, permettant alors une dissémination de *Map* dans l'organisme. Deux formes de lésions ont été décrites chez les ovins par Clarke and Little (1996). Les lésions multibacillaires présentent de très nombreuses *Map* dans des macrophages agrégés en amas, distendent la lamina propria et s'infiltrent dans la sous-muqueuse. A l'inverse, les lésions paucibacillaires se caractérisent par la présence de seulement quelques *Map* dans les macrophages, des granulomes non encapsulés, plus petits et une infiltration moindre de la lamina propria et de la sous-muqueuse. Classiquement, les lésions paucibacillaires sont présentes durant la phase précoce de la maladie et évolueront vers une généralisation des lésions multibacillaires, chaque granulome pouvant évoluer indifféremment vers l'un ou l'autre des types lésionnels.

Ces lésions, originellement décrites chez les ovins, sont aussi présentes chez les bovins (Koets et al., 2015) et les caprins (Catton, 2002 ; Kruze et al., 2006). Elles provoquent une érosion des villosités épithéliales, parfois associée à de la nécrose épithéliale. Au bout de plusieurs mois ou années (en fonction de la dose infectieuse, de l'âge et de l'espèce), une multiplication des granulomes est observée dans la muqueuse intestinale, et de manière moins systématique, dans d'autres organes (foie, reins, poumon). L'avancée des lésions fait l'objet de nombreux systèmes de classification, initialement créés pour les ovins (Pérez et al., 1996 ; Clarke, 1997). Ces lésions envahissent progressivement l'intestin, à l'origine d'un épaissement de la paroi et d'un syndrome de malabsorption. Ceci se traduit par la présence de diarrhée chez l'animal atteint, plus fréquente chez les bovins que chez les ovins et les caprins, alors que l'appétit est conservé. Ce signe, initialement intermittent, devient constant et ne rétrocede pas à l'utilisation des antibiotiques usuels. L'animal en phase terminale maigrit jusqu'à devenir une non-valeur économique avant de succomber de manière certaine à l'évolution de la maladie.

1.2.3 Dynamique de la réponse immunitaire

Classiquement, la dynamique de la réponse de l'animal à une infection par *Map* est décrite comme étant associée à une immunité de type cellulaire précoce, suivie d'un changement pour une immunité à dominante humorale avec l'avancée de la maladie (Gillan et al., 2010 ; Clarke, 1997). L'immunité protectrice est de type Th1 ; elle est mise en place par la production d'interleukine (IL) 12 par les macrophages/cellules dendritiques infectés, induisant un recrutement de lymphocytes T qui vont produire de l'IL-2 orientant la différenciation des lymphocytes T en CD4+ producteurs d'interferon γ (IFN γ). Le type 2 de la réponse CD4 est caractérisé par la production de cytokines différentes dont l'IL-4, IL-5, IL-6 et IL-10 (Stabel, 2000; 2006). Ce profil de réaction immunitaire est remis en question ; plusieurs études montrent une simultanété de la réponse Th1 et Th2 durant la phase précoce de la maladie, aussi bien chez les ovins (Begg et al., 2011) que chez les bovins (Stabel and Robbe-Austerman, 2011). Chez les ovins, Begg et al. (2011) ont mis en évidence 3 profils types de réactions immunitaires en mesurant les niveaux d'IFN γ et d'anticorps :

- "classique", avec une augmentation précoce d'IFN γ suivi d'un effondrement conjointement à l'augmentation de la production d'anticorps ;
- "simultané", caractérisé par une augmentation précoce des niveaux d'IFN γ et d'anticorps ;
- et "cellulaire", décrit par une production d'IFN γ sans réponse anticorps.

Nous pouvons synthétiser la réponse dite "classique" de l'hôte avec la figure 4.

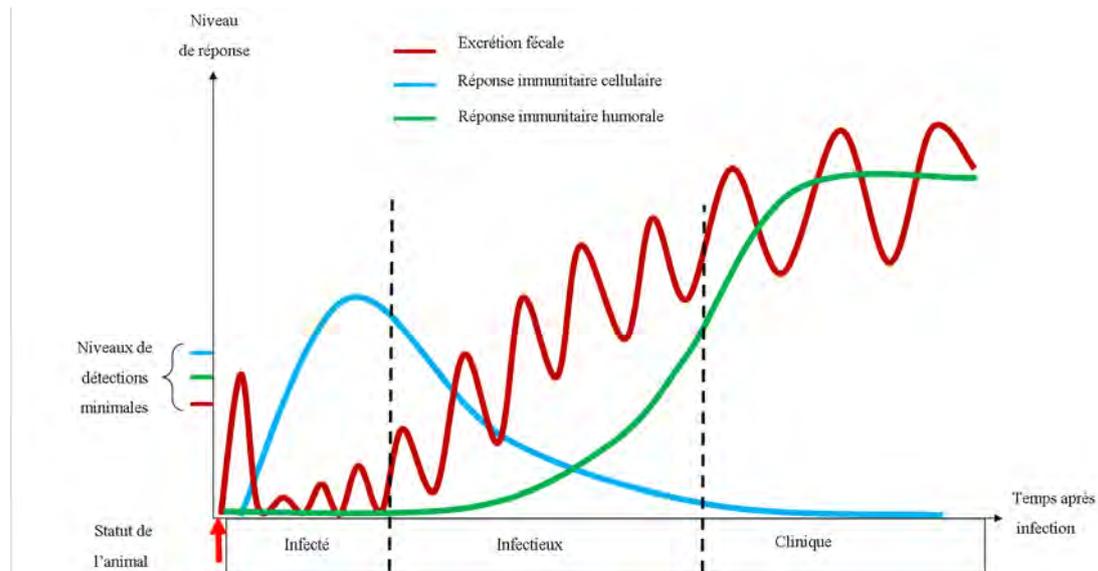


FIGURE 4 – Représentation schématique de la réponse immunitaire dite "classique" avec l'évolution de la paratuberculose

A partir de l'ingestion de *Map* (représentée par la flèche rouge) les différentes réponses de l'organisme sont représentées schématiquement.

En marron, l'excrétion fécale se caractérise par une phase de transport passif détectable rapidement après infection puis une excrétion qui n'est détectable que de façon intermittente, avant de devenir permanente. En bleu est représentée la réponse immunitaire cellulaire qui classiquement augmente en début d'infection avant de s'effondrer avec la progression de la maladie.

En vert est représenté le niveau de réponse immunitaire humorale qui, au début de l'infection est faible et indétectable et augmente au fur et à mesure de l'évolution vers la maladie.

Les niveaux minimums de détection sont l'équivalent de la sensibilité analytique (Se_a) des tests utilisés pour la détection des différentes réponses de l'organisme.

2 Épidémiologie

2.1 Épidémiologie descriptive

2.1.1 Animaux cibles

De nombreuses espèces animales peuvent être infectées par *Map*, mais les cas cliniques sont principalement observés chez les ruminants domestiques : bovins, buffles, ovins et caprins, cerf élaphe. Des espèces de la faune sauvages sont également sensibles à la paratuberculose, des cas sont décrits chez des carnivores sauvages (ratons laveurs, opossums, coyotes (Anderson et al., 2007)) et des oiseaux (étourneau, pie) (Waddell et al., 2016) ainsi que chez des ruminants et pseudo-ruminants (Carta et al., 2013).

2.1.2 Animaux réservoirs

Tous les animaux chez lesquels la maladie a été décrite n'ont pas la même importance vis-a-vis de l'épidémiologie de la paratuberculose. La notion de réservoir telle que décrite par l'Association pour l'Étude de l'Épidémiologie des Maladies Animales (AEEMA) permet de distinguer les espèces jouant un rôle prépondérant dans la dissémination de la bactérie, de celles pouvant être considérées comme des culs-de-sac épidémiologiques. Une espèce réservoir est une entité assurant la conservation d'un agent pathogène biologique, considéré en tant qu'espèce, et sa fourniture aux sujets réceptifs (AEEMA, 2017). Utilisant cette définition, nous retenons deux points nécessaires pour faire d'une espèce un réservoir de la paratuberculose :

- Assurer la conservation : un réservoir doit pouvoir être infecté par *Map*
- Assurer sa fourniture aux sujets réceptifs : un réservoir doit pouvoir excréter *Map*, entraînant un contact potentiel avec un animal réceptif à la maladie.

A l'échelle d'un élevage ou d'un territoire, la connaissance des espèces réservoirs de la paratuberculose est indispensable pour entreprendre une lutte contre la maladie. Parmi les espèces réservoirs, nous retrouvons les ruminants domestiques (bovins, ovins, caprins, voire cervidés d'élevages et buffles), et on dénombre cinq espèces réservoirs dans la faune sauvage (quatre espèces de cervidés et le lapin sauvage) (Judge et al., 2007 ; Maio et al., 2011 ; Shaughnessy et al., 2013 ; Carta et al., 2013)

2.1.3 Répartition spatiale

La paratuberculose a une répartition mondiale comme le montre la figure 5 qui recense les cas déclarés chez les animaux domestiques et sauvages à l'Office International des Épizooties (OIE) durant le premier semestre 2016 (OIE, 2017). Comme nous le verrons au paragraphe 4.3, les tests diagnostiques utilisés pour le dépistage de la paratuberculose ont des performances variables selon la cible épidémiologique, mais sont globalement médiocres en début d'infection. Les études d'évaluation de la prévalence sont donc souvent biaisées et difficilement comparables, mais elles permettent d'approcher la situation épidémiologique de la région concernée.

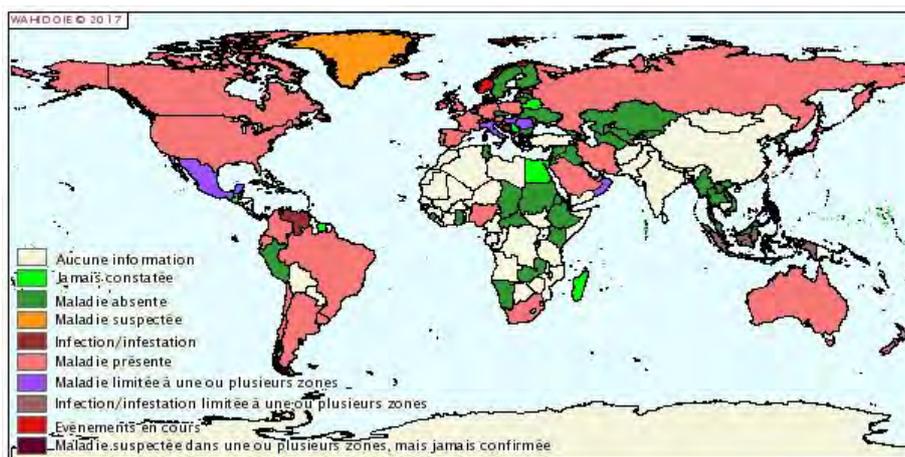


FIGURE 5 – Carte de la répartition mondiale de la paratuberculose au premier semestre 2016, (OIE, 2017)

En se basant sur les déclarations envoyées à l’OIE durant le premier semestre 2016, les pays en rose sont ceux où la paratuberculose est présente, en opposition aux pays en vert dans lesquels la paratuberculose est absente. On constate que la maladie est présente sur chaque continent et on peut supposer un effet de sous-déclaration au vu des résultats retrouvés dans la littérature.

2.1.4 Prévalence à l’échelle mondiale

Plusieurs revues de la littérature synthétisent les études de prévalence par grandes régions du globe. La paratuberculose est décrite en Afrique avec des prévalences variant entre la simple description de cas jusqu’à une prévalence de 56% dans les cheptels bovins au Soudan. La paratuberculose est retrouvée chez les bovins, les ovins les caprins ainsi que chez des antilopes (Okuni, 2013). En Amérique du sud, comme en Afrique, la paratuberculose est décrite dans tous les pays et chez toutes les espèces de ruminants domestiques. Les prévalences observées sont variables avec des extrêmes compris entre 18,7% au Costa Rica et 100% au Venezuela à l’échelle des élevages bovins et entre 2,3% à la Grenade et 44% au Mexique à l’échelle individuelle pour les ovins (Fernández-Silva et al., 2014). En Europe, de nombreuses études sont disponibles et ont fait l’objet d’une revue (Nielsen and Toft, 2009), déjà ancienne mais corrigeant les prévalences apparentes observées en tenant compte de la sensibilité diagnostique (Se_D) et de la spécificité diagnostique (Sp_D) imparfaites des tests utilisés. Les auteurs concluent à une prévalence à l’échelle des troupeaux, supérieure à 50% chez les bovins et supérieure à 20% chez les caprins et les ovins. De plus, aucune des études incluses dans cette revue ne permet d’assurer l’absence de la paratuberculose dans quelque pays que ce soit en Europe. L’Australie et la Nouvelle-

Zélande sont également touchées par la paratuberculose. Les prévalences de troupeaux observées varient de 42% pour les bovins à 76% pour les ovins en Nouvelle-Zélande, avec des différences significatives entre les deux îles (Verdugo et al., 2014a). En Australie, la prévalence des troupeaux ovins atteints de la paratuberculose varie entre 0,2% et 33% en fonction de la région (Sergeant and Baldock, 2002).

2.1.5 Prévalence en France

Les données sur la prévalence de la paratuberculose ovine en France sont parcelaires et peu fiables. Nous retrouvons notamment une approximation de la répartition dans la figure 6 (Petit, 2006) qui présente les résultats d'une enquête envoyée aux Groupements de Défense Sanitaire (GDS) français. Le taux de réponse à l'enquête était de 100% et les résultats sont indicateurs d'une méconnaissance de la répartition réelle de la maladie. L'absence de dépistage systématique et organisé, ainsi que l'utilisation quasi-systématique des tests sérologiques (qui ont une Se_D médiocre comme nous le verrons, au paragraphe 4.3) permettent raisonnablement de suspecter non seulement une répartition plus large en France mais également des niveaux de prévalence plus élevés.

Des données plus solides et plus récentes sont disponibles sur la prévalence dans l'espèce caprine en France à partir des résultats de sérologie ELISA obtenus sur l'ensemble des caprins adultes de 105 élevages répartis sur tout le territoire français selon un schéma d'échantillonnage stratifié (Mercier et al., 2010). La prévalence réelle des troupeaux infectés, corrigée en prenant en compte les performances imparfaites du test diagnostique utilisé, a été estimée à 63% avec un intervalle de confiance à 95% (IC 95%) entre 41 et 84% avec une prévalence réelle intra-élevage variant de 1 à 33%.

2.2 Épidémiologie analytique

2.2.1 Sources de *Map*

En élevage, les sources de *Map* sont majoritairement les bovins et les ovins cliniquement atteints de paratuberculose. Viennent ensuite les animaux infectés de manière asymptomatique qui ont une excrétion intermittente et en quantité moindre.

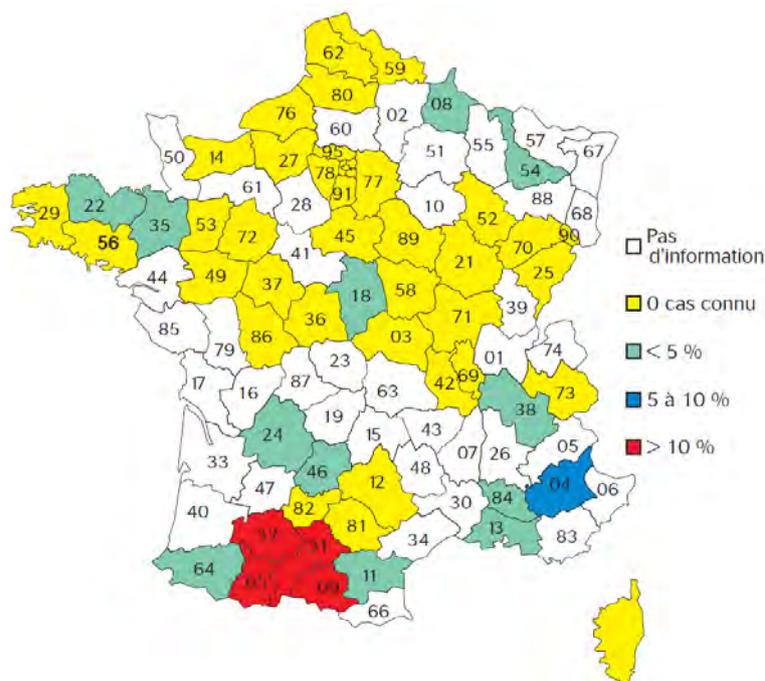


FIGURE 6 – Carte de France de l’estimation de la prévalence de la paratuberculose ovine à l’échelle des départements

La prévalence a été estimée à l’échelle départementale en fonction des cas traités par les GDS. Les départements en blanc sont ceux où la maladie n’a pas été mise en évidence et non pas une absence de réponse à l’enquête.

2.2.2 Matières infectieuses

Les fèces sont la principale voie d’excrétion de *Map* dans l’environnement. Du fait des caractéristiques de la mycobactérie et notamment sa résistance dans le milieu extérieur comme nous l’avons vu au paragraphe 1.1.5, la contamination de l’environnement peut persister plusieurs mois, voire des années dans un élevage. La quantité de *Map* excrétée par cette voie va augmenter avec l’évolution de la maladie, de $2,4 \times 10^6$ unité formant colonie (UFC) par grammes de fèces et par jour pour les formes paucibacillaires, jusqu’à 10^{11} UFC/g et par jour pour les formes multibacillaires (Whittington et al., 2000b). L’excrétion de *Map* peut être intermittente ou continue. En analysant les résultats de coproculture individuelle de 2800 bovins répartis dans 21 élevages suivis durant une période allant de trois ans et demi à 20 ans ainsi que 20 bovins infectés expérimentalement, Mitchell et al. (2015) définissent plusieurs profils d’excrétion fécale et d’évolution de la paratuberculose. Parmi les bovins infectés naturellement les auteurs ont observés que les excréteurs continus ont une probabilité deux fois plus élevée de devenir des forts excréteurs comparé aux

excréteurs intermittents.

Il est à noter qu'un phénomène de transport passif est possible, des animaux non infectés peuvent excréter la bactérie jusqu'à sept jours après son ingestion (Sweeney et al., 1992a ; Mortier et al., 2014 ; Schwarz et al., 2017). Par sécurité, les recommandations pour la conduite d'infections expérimentales considèrent une période de 15 jours durant laquelle l'excrétion passive est possible (Hines et al., 2007), passé ce laps de temps un animal excréteur peut être considéré comme multipliant la mycobactérie.

De manière similaire à l'excrétion de *Map* dans les fèces, une excrétion dans le **colostrum** et le **lait** est possible. Elle est plus fréquente aux stades avancés de la maladie. Sweeney et al. (1992b) ont mis en évidence une corrélation entre le niveau d'excrétion fécale et mammaire. Soixante-dix sept vaches ayant une culture fécale positive (37 fortes excrétrices, neuf intermédiaires et 31 faibles excrétrices) ont été testées par culture à partir du lait. Les résultats montrent un pourcentage de cultures positives variant de 3% pour le groupe des faibles excrétrices à 19% pour les fortes excrétrices. L'importance épidémiologique de cette voie d'excrétion apparaît toutefois moindre que la voie fécale car les quantités de *Map* disséminées par cette voie sont plus faibles et la contamination de l'environnement moins probable. L'excrétion dans le lait a également été démontrée chez les ovins et les caprins (Nebbia et al., 2006 ; Verin et al., 2016). Les échantillons de lait de 29 animaux (20 brebis et neuf chèvres) provenant d'élevages infectés ont été testés par Nebbia et al. (2006) en utilisant une réaction en chaîne par polymérase quantitative (qPCR) ciblant le fragment *IS900*. L'ADN de *Map* a été isolé sur 13 de ces animaux (six caprins, sept ovins), les animaux séropositifs au début de l'étude avait un risque plus élevé d'excréter *Map* (60%) par rapport aux animaux séronégatifs (28%).

La **semence** peut également être contaminée par *Map*, mais dans une moindre mesure que les fèces et les sécrétions mammaires. Néanmoins *Map* a pu être détecté par culture et/ou par qPCR chez des taureaux reproducteurs asymptomatiques qui excrétaient *Map* dans les fèces (Ayele et al., 2004 ; Münster et al., 2013), ainsi que chez des béliers en phase clinique (Eppleston and Whittington, 2001).

2.2.3 Voies de transmission intra-élevage

La transmission peut être séparée en trois principales voies que nous pouvons définir et hiérarchiser de la façon suivante :

- Horizontale : transmission entre deux individus sans relation avec la période péri-partum.
- Pseudo-verticale : transmission entre deux individus pendant la période néonatale.
- Verticale : transmission pendant la gestation

Transmission horizontale. C'est la voie la plus commune de transmission intra-élevage. Un animal asymptomatique pouvant excréter jusqu'à $8 * 10^{10}$ *Map* par jour (Whittington et al., 2000b), la contamination de l'environnement d'élevage (notamment la litière), de l'eau de boisson, de la mamelle, et du matériel d'élevage sont autant de sources d'infection potentielles. La dose infectieuse minimale est mal connue, néanmoins Sweeney et al. (2006) ont réussi à reproduire la maladie chez des ovins avec une dose de $1,5 * 10^6$ unités formant colonie (UFC) distribuée à 20 jours de vie.

Transmission pseudo-verticale. Comme vu au paragraphe 1.2.1, une infection par *Map* est associée à une phase de bactériémie. Une excrétion de l'agent pathogène dans la mamelle et dans le lait est donc possible et a été démontrée chez les bovins (Sweeney et al., 1992b) et chez les petits ruminants (Nebbia et al., 2006). Cette excrétion survient aussi bien chez les animaux en phase clinique qu'asymptomatiques. De plus les traitements thermiques n'empêchent pas la survie de *Map* dans les poudres de lait (Grant et al., 2017) et dans le lait pasteurisé (Ellingson et al., 2005), à partir desquels elle peut être isolée.

Transmission verticale. Plusieurs études ont mis en évidence une colonisation du tractus génital chez les bovins et les petits ruminants (Lambeth et al., 2004 ; Whittington and Windsor, 2009), le risque d'infection transplacentaire semblant varier en fonction du statut de l'animal. Dans leur méta-analyse portant sur les bovins, Whittington and Windsor (2009) estiment ainsi un taux d'infection du fœtus de 9% si la vache est asymptomatique contre 39% si elle est en phase clinique. Les résultats de (Lambeth et al., 2004) vont dans

le même sens dans l'espèce ovine.

2.2.4 Facteurs de réceptivité

Au même titre que la dose infectieuse, l'âge des animaux lors de leur premier contact avec l'agent pathogène est un facteur primordial dans le déroulement de l'infection. Lors d'une étude longitudinale sur deux ans et demi, McGregor et al. (2012) ont infecté naturellement, grâce à un pâturage partagé avec des ovins excréteurs durant 100 jours, des groupes de 280 animaux : des agneaux (âge médian de six mois), des brebis antenaises (âge médian de deux ans et demi) et des brebis adultes (âge médian de cinq ans). Les agneaux exposés étaient plus souvent (odds ratio (OR) : 7,1; IC 95% [2,43 - 21,05]) et plus précocement excréteurs que les ovins adultes (6 mois après exposition pour les agneaux contre 11 mois pour les brebis adultes). Par ailleurs, les lésions histologiques étaient plus fréquentes (OR : 3,3; IC 95% [1,36 - 7,88]) et plus sévères selon la classification de Pérez et al. (1996) chez les agneaux.

Dans cette même étude, l'effet de la dose sur les mêmes paramètres a été étudié. Trois niveaux d'exposition (faible, moyen et fort) ont été définis en fonction du nombre d'ovins excréteurs/jour mis en contact sur les pâtures. Dans les conditions naturelles d'infection, une exposition à une grande quantité de *Map* a conduit plus fréquemment à une excrétion fécale (OR : 3,5; IC 95% [1,30 - 9,31]) et ce de manière plus précoce. Contrairement à l'âge, la dose d'exposition n'avait aucun effet significatif sur les lésions histopathologiques observées. A l'inverse, dans un essai expérimental sur des veaux, Mortier et al. (2013) ont observé des lésions histopathologiques plus sévères lors d'une infection avec une forte dose de *Map* ($5 * 10^9$ UFC) qu'avec un inoculum 100 fois plus petit ($5 * 10^7$ UFC).

2.2.5 Transmission entre élevages

Du fait de la longue période d'incubation et de la faible Se_D des tests pour la détection des animaux en phase subclinique (comme nous le verrons au paragraphe 4.3) la transmission de la paratuberculose entre élevages est principalement due à l'introduction d'un animal infecté dans un élevage indemne. À partir d'un modèle décrivant la répartition individuelle intra-troupeau de l'infection et en se basant sur les données de mouvements

d'achats de bovins laitiers, Beaunée et al. (2015) ont modélisé la dissémination de la paratuberculose à l'échelle de la Bretagne sur une période de neuf ans. Sans surprise, la probabilité d'acquérir l'infection augmente avec le nombre d'achats réalisés par an, et un élevage initialement indemne introduisant au moins trois bovins par an pendant neuf ans avait un risque supérieur à 50% d'être contaminé.

Map étant une bactérie résistante à de larges plages de pH, de température et de taux d'oxygène, sa survie a été démontrée, par culture ou par PCR, jusqu'à trois semaines dans les effluents liquides (Olsen et al., 1985) ou solides (Grewal et al., 2006 ; Fecteau et al., 2013). Les méthodes de traitement (compostage (Grewal et al., 2006), méthanisation (Slana et al., 2011)) semblent inefficaces pour assainir les effluents. De plus l'aérosolisation des mycobactéries a été mise en évidence (Rowe and Grant, 2006 ; Falkinham, 2003) ce qui, associé à la persistance de *Map* dans les effluents représente un risque potentiel de transmission entre élevages, la charge infectieuse étant cependant faible et la voie de contamination peu efficace.

2.3 Importance de la paratuberculose

Au delà de la mortalité accrue, la paratuberculose provoque une baisse de la production chez les animaux infectés. De plus l'implication de *Map* dans la maladie de Crohn a été suspectée dès les années 1910, soulevant une possible question de santé publique. Nous séparerons l'étude de l'importance de la paratuberculose selon les trois axes suivants :

- Importance économique;
- Importance en santé animale;
- Importance en santé publique.

2.3.1 Importance économique

L'inflammation chronique de l'intestin résultant de l'infection par *Map* est associée à un syndrome de malabsorption des nutriments, principalement des protéines. Il en ré-

sulte un déséquilibre du ratio énergie/protéines ayant un effet direct sur les capacités de production de tous les ruminants infectés. De manière classique, nous retrouvons une diminution de la production laitière et bouchère ainsi qu'un raccourcissement de la vie économique des animaux infectés.

McGregor et al. (2015) ont classé une cohorte de 83 agneaux, tous exposés à *Map* sur une pâture contaminée, en s'appuyant sur des résultats de culture tissulaire ainsi que la gradation des lésions histopathologiques observées après euthanasie. Le suivi du poids des animaux sur trois ans (16, 21, 26, 32 et 36 mois) indique une réduction significative du poids des animaux avec des lésions multibacillaires par rapport à ceux ne présentant ni lésions, ni isolement à partir des tissus. Cette baisse de poids est apparue dès 21 mois d'âge et est devenue clairement significative à partir de 32 mois. À la fin du suivi, les animaux avec une forme multibacillaire pesaient en moyenne 4,85 kg de moins que les animaux témoins. Dans cette même étude, la production de laine était également suivie, sans qu'une différence n'ait été observée entre les animaux sévèrement affectés et les autres. La perte de poids est également constatée chez les bovins dans des études rétrospectives aussi bien en production laitière (Kudahl and Nielsen, 2009) que bouchère (Bhattarai et al., 2013), en se basant sur les pesées d'abattoir et en définissant le statut infectieux des animaux à partir des résultats d'ELISA ou de coproculture.

L'effet d'une infection par *Map* sur la production de lait chez les bovins est sujet à controverse. Les nombreuses études disponibles rapportent soit une diminution de la quantité de lait produite soit un faible effet, voire une production plus élevée. La majorité de ces études sont construites en opposant un groupe "indemne" à un groupe "infecté" dont les définitions sont variables (animal ou cheptel positif à un test ELISA sur le lait, sur le sang, animal ou cheptel excréteur) (Lombard et al., 2005 ; Hoogendam et al., 2009 ; Gonda et al., 2007 ; Johnson et al., 2001). Cette différence de cible rend difficile la comparaison des effets rapportés par ces études, mais l'effet le plus probable serait une production plus élevée de lait pendant la première lactation avant une réduction de celle-ci avec l'avancée de la maladie. Il n'existe pas d'étude sur la production laitière ovine, mais nous pouvons supposer que les effets sont similaires à ceux observés chez les bovins.

Les performances de reproduction chez les ovins et les caprins ont notamment été étudiées par Kostoulas et al. (2006a) dans quatre élevages infectés par *Map*. Dans cette étude rétrospective, le statut de 173 ovins et 196 caprins ne présentant pas de symptômes de paratuberculose a été déterminé par sérologie ELISA et culture fécale, un animal étant considéré comme infecté dès lors qu'un des deux tests était positif. La réussite des performances de reproduction était codée par une variable binaire (un si l'animal avait donné naissance à une progéniture vivante l'année précédant les tests, zéro dans le cas opposé) et modélisée en fonction de variables explicatives comprenant la parité, le statut infectieux ainsi que l'élevage comme effet aléatoire. Les résultats obtenus étaient là aussi variables selon l'âge des brebis : les animaux infectés de parité inférieure à quatre avaient un OR de réussite à la reproduction de cinq (IC 95% : [1,5-13,3]) par rapport aux animaux non-infectés, alors que les animaux infectés de parité supérieure à six avaient un OR de 0,05 (IC 95% : [0,007-0,4]) par rapport aux animaux de même parité considérés comme non-infectés.

Deux hypothèses simples permettent d'expliquer les effets variables selon l'âge de l'infection par *Map*, que ce soit pour la production laitière ou les performances de reproduction. Un biais de classification est possible du fait de la définition imprécise d'animal asymptomatique infecté qui peut être excréteur ou non au moment du prélèvement servant à définir son statut infectieux. L'autre hypothèse serait une plus grande réceptivité à l'infection des animaux les plus performants qui, durant les premières années de vie, contiennent l'infection avant d'observer un effondrement de leurs performances avec l'avancée de la maladie.

2.3.2 Importance en santé animale

Comme nous l'avons vu précédemment, l'évolution d'une infection par *Map* est le plus souvent irrémédiable une fois l'apparition des premiers signes cliniques. Whittington et al. (2017) définissent un phénomène de résilience, difficile à mettre en évidence du fait de l'existence de l'excrétion passive qui peut faire conclure à tort à la présence

d'une infection. Pour explorer les mécanismes de cette résilience, de Silva et al. (2018) ont exposé 20 ovins qu'ils ont suivi durant deux ans et demi. Onze de ces animaux ont été qualifiés de résilients par les auteurs car ils ne présentaient pas de lésions histologiques ou d'excrétion fécale continue. La figure 7 montre la répartition des 20 animaux à la fin de l'essai et décrit les réponses de l'hôte (IFN γ et sérologie) associées à chaque catégorie.

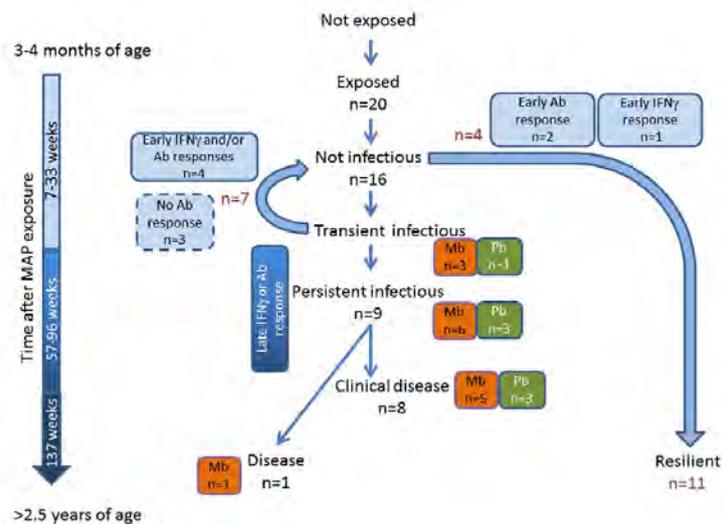


FIGURE 7 – Statut de 20 ovins exposés à *Map* au cours d'un suivi de deux ans et demi (de Silva et al., 2018)

Ab : Anticorps
 IFN γ : taux d'IFN γ
 Pb : animaux paucibacillaires
 Mb : animaux multibacillaires

Onze des animaux enrôlés dans cette étude ont été considérés comme résilients à la fin du suivi, sept ayant été détectés comme excréteurs à un moment donné du suivi. Les auteurs n'ont pas rapporté de différences dans les paramètres immunologiques entre ces deux groupes.

D'autre part, la résistance de *Map* aux antibiotiques rend leur utilisation sur des animaux atteints de paratuberculose peu efficace et coûteuse. De plus, la restriction de l'usage de certaines familles d'antibiotiques sur les animaux d'élevage impose une contrainte supplémentaire pour la maîtrise de l'infection par *Map*. En l'absence de traitement, une dégradation progressive de l'état de l'animal infecté est observée. La seule possibilité pour l'éleveur pour limiter les pertes économiques est alors la réforme précoce des animaux infectés, avant qu'ils ne deviennent des non-valeurs. Le diagnostic des animaux en phase

préclinique est donc un aspect essentiel de la maîtrise de l'infection en élevage. Il est rendu difficile par les piètres performances des tests diagnostiques comme nous le verrons au paragraphe 4.3.

En France, un vaccin inactivé (Gudair[®]) est disponible pour les ovins et les caprins depuis 2009 et depuis fin 2014 pour les bovins (Silirum[®]) (cf. chapitre 5). Leur utilisation est réglementée du fait de l'existence de réactions croisées lors du dépistage de la tuberculose (infection par *Mycobacterium bovis*) par intradermotuberculation (Garrido et al., 2013 ; Serrano et al., 2017).

2.3.3 Importance en santé publique

Le lien entre la paratuberculose et la maladie de Crohn a été suspecté dès 1913 par Dalziel (1989) suite à l'observation de trois cas cliniques humains présentant une inflammation de l'intestin. La participation possible de *Map* dans l'étiologie de la maladie de Crohn fait depuis l'objet d'une controverse scientifique sur le statut zoonotique de l'infection paratuberculeuse. En 1945, Taylor (1953) décrit la lésion caractéristique comme une inflammation chronique de l'intestin, associé à l'épaississement de la paroi aboutissant à une sténose de la partie de l'intestin touchée. Ce tableau lésionnel n'est pas forcément associé à la présence de granulomes et à l'époque, l'agent étiologique de la maladie de Crohn ne pouvait que faire l'objet de suppositions. Trois hypothèses principales ont été proposées par Chamberlin and Naser (2006) pour expliquer la maladie de Crohn, soulignant qu'il s'agit surtout d'une manière d'interpréter les données disponibles en mettant l'accent sur tel ou tel point de la maladie

- un processus auto-immun ;
- une immunodéficience de l'hôte ;
- une infection par *Map*.

L'hypothèse auto-immune considère la maladie de Crohn comme une réaction anormale du système immunitaire digestif à la flore commensale, cette réaction est facilitée par des brèches dans la barrière mucoale, pouvant éventuellement être causées par *Map*. L'hypothèse de l'immunodéficience de l'hôte est basée sur le fait qu'un déficit d'activité des cellules neutrophiles (causé par des facteurs environnementaux ou génétiques) fa-

cilite l'établissement des bactéries dans les macrophages, provoquant une réponse Th1 excessive à l'origine de symptômes similaires à ceux de la maladie de Crohn mais sans spécificité d'agent pathogène. La dernière hypothèse est simplement que l'agent pathogène responsable de la maladie de Crohn est *Map*. *Map* est plus souvent détectée chez les patients atteints de maladie de Crohn que chez les individus témoins. Cette détection peut être réalisée par mesure des anticorps dirigés contre *Map*, par PCR/qPCR, par culture ou supposée suite à un traitement antibiotique efficace contre *Map*, permettant une réduction des symptômes chez les patients affectés (Kuenstner et al., 2017). Ces éléments, même s'ils ne permettent pas de démontrer un lien de causalité strict entre *Map* et la maladie de Crohn vont dans le sens d'une corrélation. *Map* peut donc être vue comme au minimum un facteur de risque pour la maladie de Crohn. De plus, du fait de la résistance de *Map* dans l'environnement (paragraphe 1.1.5) ainsi que son excrétion dans le lait des ruminants (paragraphe 2.2.2), l'Homme est exposé à la mycobactérie (Pickup et al., 2005; 2006 ; Klanicova et al., 2012 ; Atreya et al., 2014 ; Grant et al., 2017). Si les caractéristiques observées pour la paratuberculose, notamment la sensibilité plus élevée des individus les plus jeunes, sont transposables entre espèces, alors *Map* a le potentiel pour devenir un enjeu de santé publique majeur dans les années à venir.

3 Tests biologiques pour le diagnostic individuel de la paratuberculose

Les tests diagnostiques utilisés pour le diagnostic individuel de la paratuberculose ou le dépistage de l'infection par *Map* peuvent être subdivisés en tests directs et indirects. Nous présenterons ici les principes sur lesquels reposent les tests les plus fréquemment utilisés, une synthèse sur leurs performances diagnostiques et sur les méthodes employées pour les évaluer étant décrite dans le paragraphe 4.

3.1 Méthodes directes

Les tests directs présentent l'avantage de mettre en évidence l'agent pathogène chez l'animal testé ce qui, bien que ne permettant pas d'établir un lien de causalité strict avec une infection, apporte une information plus précise que les tests indirects. Pour le diagnostic de la paratuberculose les tests directs disponibles sont

- la coloration de Ziehl-Neelsen (Weber et al., 2009);
- l'isolement en culture (Kostoulas et al., 2006b);
- la détection d'une séquence d'ADN spécifique de *Map* par PCR/ qPCR (Sting et al., 2014 ; Kawaji et al., 2011; 2007);
- le test d'amplification des phages (Foley-Thomas et al., 1995).

3.1.1 Coloration de Ziehl-Neelsen

Comme vu paragraphe 1.1.2, *Map* est un bacille acido-alcool résistant ce qui rend possible son identification par la coloration de Ziehl-Neelsen. À partir d'échantillons de lésions ou de fèces et après fixation sur une lame, la coloration se décompose en quatre étapes :

- une coloration par la fuschine qui pénètre à l'intérieur de la cellule
- une décoloration avec une solution acide, généralement HCl
- une décoloration avec une solution d'éthanol
- une recoloration du fond de lame avec une solution de méthylène

La recherche de *Map* par la coloration de Ziehl-Neelsen met en évidence la bactérie vivante et peut être utilisée ante-mortem à partir de fèces ou post-mortem sur des échantillons tissulaires (iléon notamment). Cependant cette technique n'est pas spécifique de *Map* mais des mycobactéries ; en général elle présente un faible caractère de spécificité la rendant difficilement utilisable pour le suivi de la paratuberculose en élevage.

3.1.2 Isolement en culture

Initialement réalisé à partir d'échantillons de tissus, l'isolement de *Map* à l'aide de la culture est désormais réalisé à partir d'échantillons de fèces pour le diagnostic ante-

mortem de l'infection. Un milieu de culture particulier est nécessaire puisque *Map* ne synthétise pas la mycobactine, nécessaire à l'acquisition de fer et donc à sa multiplication. Un des désavantages majeurs de la culture est le temps nécessaire avant de conclure à un résultat négatif, variant de quelques semaines à plusieurs mois pour les souches S de *Map*.

Après une phase de décontamination, nécessaire en raison de la forte contamination microbienne des échantillons utilisés, un milieu de culture enrichi en mycobactine est ensemencé. Classiquement le milieu utilisé est une gélose d'agar complétée avec du jaune d'œuf et de la mycobactine, appelé milieu de Herrold. L'échantillon est ensuite incubé à 37°C avec contrôle visuel à plusieurs reprises jusqu'à ce que les colonies deviennent visibles. Une culture est considérée comme négative en l'absence de développement de colonies de *Map*, généralement au bout de 16 à 20 semaines.

Le milieu de culture a été amélioré pour répondre aux spécificités de l'isolement des *Map* de type S, et leur croissance plus lente que celles des souches C. Whittington et al. (1999) ont évalué le milieu liquide BACTEC 12B pour la culture fécale et tissulaire. Après culture, la confirmation de la présence de *Map* nécessite une étape de révélation du développement des *Map* : coloration de Ziehl-Neelsen ou PCR/qPCR sur les colonies observées. En 2013, suite à l'arrêt de la commercialisation du BACTEC 12B, en raison des contraintes de radioprotection, un nouveau milieu liquide permettant la culture des souches S, a été proposé par Whittington et al. (2013). Ce milieu, le M7H9C, présente des performances de culture similaires au BACTEC 12B ; il requiert cependant une confirmation par PCR/qPCR lors de résultat considéré positif.

Un résultat de l'isolement peut être quantifié, en fonction du temps nécessaire à l'obtention d'une réponse positive, du nombre de colonies observées, ou de l'intensité de la réaction de qPCR.

3.1.3 La réaction en chaîne par polymérisation - PCR - et PCR quantitative (qPCR)

La PCR, développée en 1985 par Saiki et al. (1985), permet l'amplification d'un fragment d'ADN cible dans un échantillon en vue de sa détection. L'ADN de l'échantillon est tout d'abord purifié afin de retirer les inhibiteurs de la réaction. Après ajout de fragments d'ADN complémentaires (amorces) du fragment recherché et d'un mélange contenant

des nucléotides ainsi que des enzymes comme l'ADN polymérase, la PCR débute avec une phase de dénaturation permettant de séparer les deux brins d'ADN. Les amorces s'hybrident au fragment cible, ce qui permet la fixation de l'ADN polymérase et le début de la réplication. Deux ADN double brins sont obtenus après refroidissement avant le début du second cycle reprenant les mêmes étapes, la quantité d'ADN est donc doublée à chaque cycle. En fin de réaction (généralement entre 30 et 50 cycles) classiquement le produit obtenu était récupéré pour détection ou mesurer la quantité produite.

La technique de la PCR a subi une avancée notable en 1996 avec le développement de la qPCR (Heid et al., 1996) en combinaison avec la détection fluorimétrique de l'ADN amplifié. Il peut s'agir par exemple de l'ajout d'un fluorophore et d'un désactivateur (quencher) sur un fragment d'ADN spécifique. En l'absence de réplication, le quencher absorbe la fluorescence du fluorophore. Lors de réplication de l'ADN cible, le fluorophore et le quencher sont séparés par l'ADN polymérase qui dégrade le fragment d'ADN qui les lie, provoquant ainsi une émission de fluorescence qui est enregistrée à chaque cycle. Les résultats obtenus sont des courbes de fluorescence avec un point de mesure à chaque cycle de la qPCR, et se décomposent en trois phases :

- **la ligne de base**, correspondant aux premiers cycles de réplication durant laquelle la fluorescence est inférieure au seuil de détection;
- **la phase exponentielle**, correspondant à une phase durant laquelle la fluorescence augmente fortement (phase de réplication de l'ADN cible);
- **le plateau**, correspondant aux derniers cycles durant lesquels la fluorescence dépasse les limites de détection du capteur, ou la réaction est à saturation.

Le résultat d'une qPCR est appelé cycle de sortie (Cycle Threshold, Ct) et peut être calculé de plusieurs façons. La première méthode d'interprétation utilisée est la détermination par l'utilisateur d'un seuil de fluorescence; le Ct est alors calculé par la projection sur l'axe des abscisses de l'intersection entre la courbe de fluorescence et ce seuil. Le calcul du Ct par cette méthode de seuil est donc subjectif et ne tient pas compte de la forme de la courbe de fluorescence. Le Ct peut également être calculé en modélisant la courbe sigmoïde de fluorescence et en calculant un indicateur à partir de la fonction qui passe au mieux par tous les points de fluorescence mesurés (courbe d'ajustement).

Cette étape de modélisation de la fluorescence permet d'identifier les courbes ayant des formes aberrantes avant une interprétation objective du cycle de sortie. Nous retrouvons principalement 3 indicateurs :

- **CpD1** est le cycle au cours duquel la dérivée première de la courbe d'ajustement est maximale. Ce cycle est situé au milieu de la phase exponentielle;
- **CpD2** est le cycle au cours duquel la dérivée seconde de la courbe d'ajustement est maximale. Ce cycle est situé au tout début de la phase exponentielle (Luu-The et al., 2005);
- **Cy0** est défini par l'intersection entre l'axe des abscisses et la tangente à la courbe d'ajustement au point d'inflexion (Guescini et al., 2008).

le CpD2 fut le premier indicateur utilisé pour l'interprétation des résultats de qPCR avec pour objectif de rendre la qPCR compatible avec une approche haut débit et donc pour cela indépendante de l'utilisateur (Luu-The et al., 2005). Plus récemment, Guescini et al. (2008) ont proposé le Cy0 comme indicateur moins sensible à la présence d'inhibiteurs d'amplification. La figure 8 illustre les différents indicateurs calculés sur la courbe d'ajustement et montre les différences de résultats obtenus selon l'indicateur choisi. Néanmoins, les appareils d'amplification étant fournis avec un logiciel d'interprétation des données, la détermination du cycle de sortie se fait souvent par un algorithme de calcul qui n'est ni connu, ni contrôlable par l'utilisateur.

La qPCR permet de quantifier l'ADN de *Map* extrait à partir d'une matrice (sang, tissu, fèces, échantillon d'environnement, lait) que la mycobactérie soit morte ou vivante. L'amplification ne peut être réalisée qu'après une étape de traitement des échantillons et d'extraction de l'ADN présent qui conduit à la perte de cette information. Le volume de matrice utilisé pour l'analyse va déterminer la représentativité du protocole. Nous pouvons citer pour le diagnostic de la paratuberculose, l'exemple de la qPCR sur fèces qui peut être réalisée à partir d'une quantité de fèces variant de 0,5g à 10g selon les recommandations fabricants. *Map* est distribuée en amas de bactéries dans les fèces et non de façon uniforme (Christopher-Hennings et al., 2003). Une plus grande quantité de fèces initiale augmente donc la probabilité d'inclure au sein de l'échantillon un ou plusieurs amas et donc la Se_D de la qPCR. D'autre part, la présence d'inhibiteurs de la qPCR dans la ma-

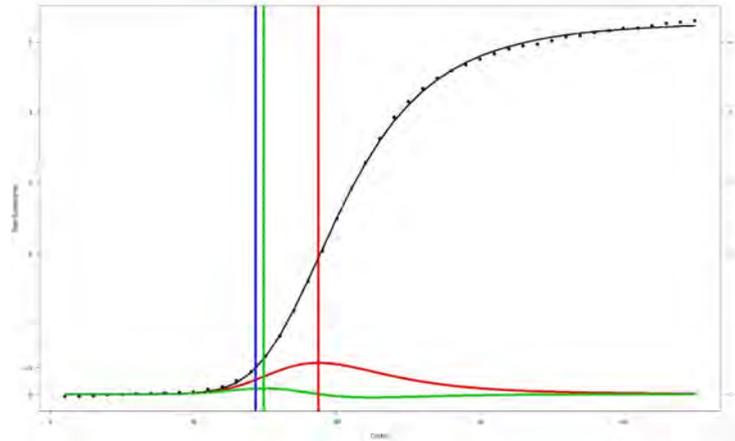


FIGURE 8 – Courbe d’amplification caractéristique de qPCR avec les différentes définitions du cycle de sortie.

La modélisation de la fluorescence est représentée en noir à partir des points enregistrés à chaque cycle. La dérivée première est en rouge, le CpD1 (ligne verticale rouge) est à 18,72 cycles. La dérivée seconde est en vert, le CpD2 (ligne verticale verte) est à 14,9 cycles. Le Cy0 est en bleu à 14,35 cycles

trice initiale va diminuer la sensibilité analytique (Se_A) de l’analyse (La Se_A est la plus petite variation biologique détectable par un test biologique cf.paragraphe 4.1.2). Acharya et al. (2017) ont étudié l’influence de la présence d’inhibiteurs dans des échantillons d’ADN extraits de fèces. En amplifiant 265 extraits ainsi que leur dilution au 1/5 dans un tampon neutre (tampon phosphate salin), les auteurs ont constaté la présence d’une inhibition de la qPCR pour 59 d’entre eux (20%). Parmi ces 59 échantillons, 42 (89%) avaient des résultats discordants entre la qPCR réalisée sur les extraits purs et les extraits dilués. La dilution dans un tampon réduisant la concentration d’inhibiteurs dans l’échantillon a permis l’amplification alors qu’elle n’était pas possible pour les échantillons purs. Les 17 échantillons restant avaient des résultats positifs dans les deux cas, l’inhibition provoquant cependant une diminution de la quantité d’ADN estimée dans le prélèvement initial.

L’étape d’extraction d’ADN doit par ailleurs être prise en compte lors de l’évaluation de la Se_A de la qPCR. Kawaji et al. (2007) ont analysé la Se_A d’une qPCR ciblant le gène *IS900* en amplifiant directement de l’ADN plasmidique contenant ce gène et d’autre part en extrayant de l’ADN de *Map* contenu dans 0,5g de fèces ovines. L’amplification à partir d’ADN plasmidique avait une Se_A telle qu’elle permettait de détecter une seule copie d’*IS900*. L’étape d’extraction d’ADN diminuait clairement cette Se_A , le protocole complet

permettant uniquement de détecter de façon certaine la présence de 10 organismes de *Map* (soit entre 150 et 200 copies d'*IS900*) par gramme de fèces traité. De plus, Timms et al. (2015) ; Fock-Chow-Tho et al. (2016) ont montré des variations de Se_A de plusieurs trousseaux d'amplification en fonction de la méthode d'extraction employée. Il convient donc, lors de l'évaluation des performances diagnostiques d'une qPCR, de tenir compte de l'ensemble de la chaîne d'analyse, y compris de la méthode d'extraction employée.

La quantification de l'ADN présent dans l'échantillon initial peut être réalisée relativement à une séquence de référence ou de façon absolue en quantifiant le fragment cible après réalisation d'une gamme de dilution (Tse and Capeau, 2003). Toute stratégie de quantification nécessite de faire l'hypothèse que les conditions de qPCR et notamment l'efficacité d'amplification sont les mêmes pour l'échantillon à quantifier et le fragment de référence ou la gamme de dilution. Pour les travaux réalisés dans le cadre de cette thèse, nous avons privilégié une approche absolue à partir d'une gamme de dilution externe, en se basant sur la répétabilité des cycles de sortie des contrôles positifs, nous assurant une efficacité similaire et rendant la comparaison possible entre les tests. Une gamme de dilution de sept points d'ADN de *Map* purifié a été analysée en triplicata pour construire une droite de régression en échelle logarithmique de base 10, reliant le Ct et la concentration initiale selon la formule suivante :

$$[°] = 10^{-a * cT + b} \quad (1)$$

avec $[°]$ qui est la concentration initiale de l'échantillon, et $-a$ et b qui sont respectivement la pente et l'ordonnée à l'origine de la droite de régression.

En 1989, plusieurs auteurs ont identifié le fragment *IS900* comme un gène spécifique de *Map* (Green et al., 1989 ; Collins et al., 1989). Ce gène, répété entre 15 et 20 fois dans le génome de *Map*, est un fragment de 1451 paires de bases contenant 66% de guanine et de cytosine. La présence de nombreuses copies d'*IS900* au sein du génome augmente la Se_A des PCR basées sur l'amplification de ce fragment. Des études complémentaires sur la PCR *IS900* ont montré l'existence de fragments d'ADN avec un degré d'homologie élevé chez d'autres mycobactéries, réduisant potentiellement la spécificité de cette réac-

tion (Cousins et al., 1999 ; Englund et al., 2002). Ces résultats ont conduit à l'élaboration d'amorces plus spécifiques ou à la définition d'autres cibles : *F57*, présent en une copie (Poupart et al., 1993), *HSPX* présent en une copie (Ellingson et al., 1998) ou encore *IS-mav2*, présent en au moins trois copies (Strommenger et al., 2001). Le plus faible nombre de copies de ces gènes dans le génome de *Map* rend cependant la Se_A de ces PCR moins élevée.

En France, cinq fabricants commercialisent six trousse de qPCR pour la détection de *Map* :

- Thermofisher LSI VetMax™ M. paratuberculosis advanced real time PCR
- Thermofisher LSI Vetmax™ Triplex M. paratuberculosis
- IDvet ID Gene™ Paratuberculosis duplex
- Idexx Real PCR MAP DNA mix
- Qiagen bactotype MAP PCR
- BioXdiagnostics Adiavet™ ParaTB real time

La plupart des trousse ciblent le gène *IS900* à l'exception du kit LSI Vetmax™ Triplex M. paratuberculosis qui mesure en parallèle les gènes *IS900* et *F57*.

3.1.4 Phage assay ou amplification de mycophages

Jones and David (1971) ; David et al. (1980b;a) ont posé les bases de cet essai par amplification de bactériophages, notant que le bactériophage D29 infecte *M. smegmatis*, une mycobactérie à croissance rapide et *M. tuberculosis*. Le protocole utilisé comprend quatre étapes :

- Mise en contact du phage D29 avec les mycobactéries potentiellement présentes qu'il infecte ;
- Ajout des bactériophages libres, puis ajout de mycobactéries libérant les phages ;
- Apport de bactéries hôtes à croissance rapide assurant une deuxième phase de multiplication des phages ;
- Observation des plages de lyse des bactéries qui ont répliqué le phage ou PCR ciblant le phage

La figure 9 (Merwe et al., 2014) décrit les étapes d'un essai par amplification des phages avec lecture des plages de lyse.

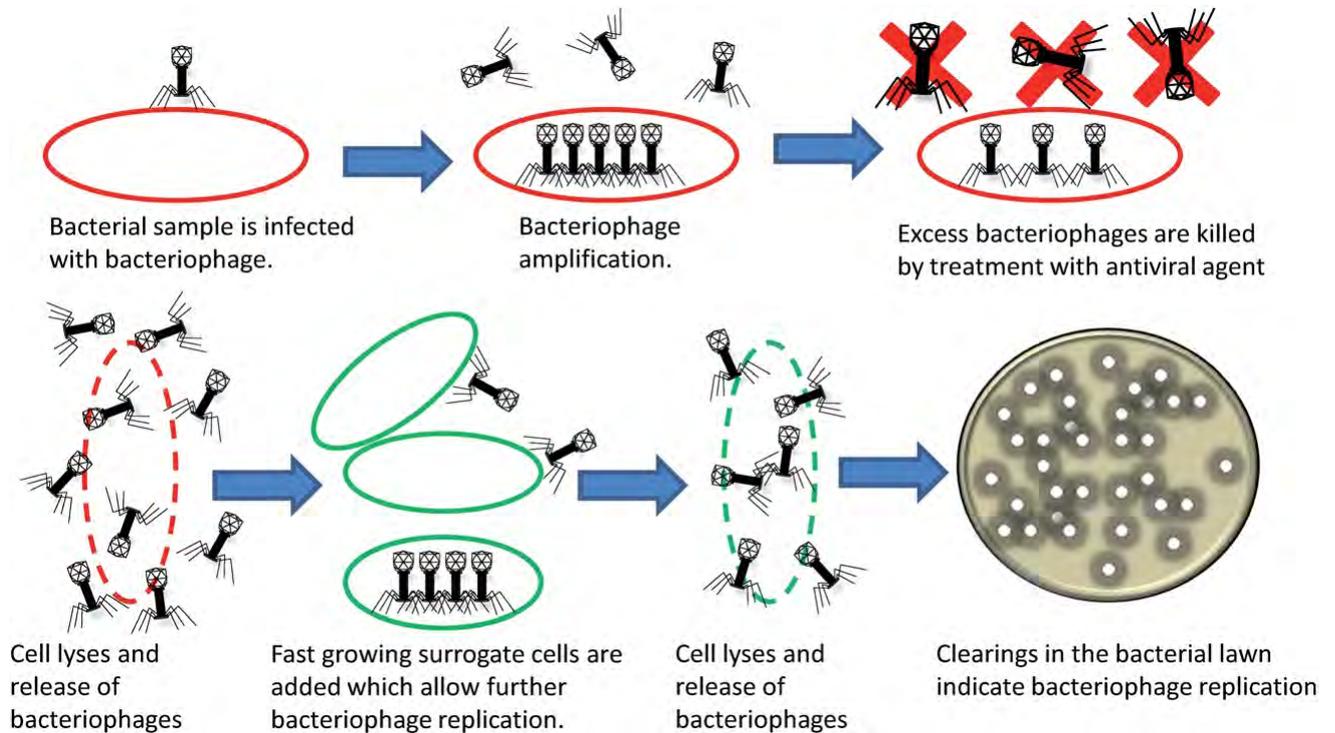


FIGURE 9 – Description d'un essai phage basé sur l'amplification des phages (Merwe et al., 2014)

Swift et al. (2013) ont proposé un protocole de récupération des cellules de *Map* par séparation magnétique avec des billes portant des peptides capables de lier *Map* de façon spécifique. Brièvement, un prélèvement sanguin est incubé avec des billes magnétiques couvertes de peptides. Les mycobactéries présentes dans le prélèvement vont se lier à ces billes grâce aux peptides et sont ainsi récupérées dans le surnageant et avant analyse par amplification de bactériophages comme présenté précédemment. Ce protocole a été testé pour la détection de *Map* circulant dans le sang de 19 bovins infectés expérimentalement (Swift et al., 2016). Sept d'entre eux étaient positifs avec ce test, aucun de ces animaux n'étant détecté par ELISA ou avec une culture fécale. A l'inverse, deux animaux négatifs pour le test basé sur les phages étaient identifiés par la culture fécale. Les méthodes basées sur l'amplification des phages présentent deux intérêts majeurs pour le diagnostic de la paratuberculose. Premièrement et contrairement à la qPCR, ils identifient les mycobactéries vivantes, la réplification des phages nécessitant une machinerie cellulaire fonctionnelle. Ils sont par ailleurs beaucoup plus rapides (de l'ordre de quelques heures en fonction des protocoles) que la culture qui ne peut être déclarée négative qu'après plu-

sieurs mois. Les méthodes basées sur l'amplification des phages ne sont cependant pas utilisées en routine à l'heure actuelle, et ont principalement été mises en oeuvre dans le cadre de projets de recherche pour mettre en évidence la présence de *Map* sur des matrices dans lesquelles les concentrations attendues étaient très faibles (sang (Swift et al., 2016), lait de consommation (Foddai and Grant, 2017 ; O'Brien et al.), poudre de lait infantile (Botsaris et al., 2016) et pour animaux d'élevage (Grant et al., 2017))

3.2 Méthodes indirectes

Les tests indirects révèlent la réponse de l'hôte à l'infection par *Map*, et ciblent soit l'immunité cellulaire, soit l'immunité anticorps (paragraphe 1.2.3). Un résultat positif à l'un de ces tests ne renseigne donc pas sur la capacité de contamination du milieu extérieur par l'animal mais permet d'identifier les animaux qui ont été au contact de *Map* et qui sont potentiellement infectés voire excréteurs.

3.2.1 Méthodes basées sur la détection de l'immunité cellulaire

Deux tests ciblant l'immunité cellulaire peuvent être utilisés pour le diagnostic de la paratuberculose en élevage : l'intradermotuberculation (IDT) et le dosage d'IFN γ après stimulation sanguine.

Intradermotuberculation. Ces tests consistent à injecter une solution antigénique par voie intradermique puis à mesurer l'augmentation du pli de peau au site de l'injection 72 heures après injection, révélant ainsi une hypersensibilité retardée de type IV. Cette famille de tests, développée originellement pour le diagnostic de la tuberculose humaine a ensuite été adaptée pour la recherche de la paratuberculose chez les bovins ainsi que chez les ovins (Robbe-Austerman, 2007). Dans le cadre du diagnostic de *Map*, l'allergène utilisé est la johnine (un dérivé de protéine purifié), obtenu après chauffage du milieu ayant servi à la culture de la mycobactérie. Tout comme pour le diagnostic de la tuberculose, l'injection de johnine provoque une période d'anergie de six semaines durant laquelle l'animal ne pourra être détecté par intradermotuberculation.

Stimulation antigénique et dosage d'IFN γ . Le dosage de l'IFN γ , consiste à stimuler un prélèvement sanguin à l'aide de dérivés de protéines purifiées. Les lymphocytes circulant, s'ils sont présents, reconnaissent les antigènes spécifiques de *Map* et sécrètent alors de l'IFN γ qui est ensuite dosé à l'aide d'un test ELISA (Rothel et al., 1990). La réalisation d'un test IFN γ présente certains inconvénients tels que un coût élevé, une valeur prédictive positive faible ainsi qu'un délai très court (moins d'une demi-journée) entre prélèvement et stimulation antigénique. Pour surmonter ce dernier problème, Jungersen et al. (2012) ont évalué la possibilité de réaliser un test IFN γ 24 heures après le prélèvement, en complétant ce dernier avec des interleukines (IL-2 et IL-12). L'ajout d'IL-2, bien qu'assurant la détection de l'IFN γ 24 heures après le prélèvement, dégradait fortement la spécificité contrairement au protocole avec complémentation en IL-12.

3.2.2 Tests basés sur la mesure de l'immunité anticorps

Ces tests sont basés sur la détection d'une réponse anticorps chez des animaux infectés par *Map*, qu'ils soient excréteurs ou non.

Test d'immunodiffusion en gélose (AGID). Le test AGID est réalisé par migration entre des puits percés dans une gélose d'agar d'une part d'antigènes de *Map* et d'autre part du sérum de l'animal à tester. Après une incubation de 24 à 48 heures, le test est considéré positif s'il y a apparition d'une ligne de précipitation entre le puits contenant l'antigène et le puits contenant le sérum à tester. Ce test n'est plus utilisé en routine.

Test de fixation du complément. Le test de fixation du complément se réalise à partir de sérum ayant une concentration standardisée de protéines du complément. Cette étape est réalisée par dénaturation des protéines du complément puis ajout d'une quantité standardisée de complément. Cela est possible car les protéines du complément sont moins résistantes à la chaleur que les anticorps. L'antigène est ensuite ajouté au sérum normalisé. Si des anticorps contre *Map* sont présents les protéines du complément vont être consommées. On ajoute ensuite un complexe immun formé de globules rouges et d'anticorps. Si les protéines du complément n'ont pas été consommées, elles provoquent la lyse des globules rouges donnant au sérum une coloration rouge. Si le sérum contient

des anticorps dirigés contre l'antigène recherché, il demeurera incolore. De même, ce test n'est plus utilisé en routine.

ELISA indirect. Les deux tests décrits précédemment ont progressivement été remplacés par des tests ELISA indirects. Le principe du test ELISA repose sur la formation et la détection de complexes immuns. Une première étape de préabsorption permet de fixer les anticorps dirigés contre d'autres mycobactéries et idéalement, de limiter les réactions croisées. Le sérum à tester est ensuite placé dans un puits au fond duquel sont fixés des antigènes de *Map*. Les anticorps du sérum se fixent sur les antigènes et le reste du sérum est éliminé par lavage. Un conjugué (composé à forte affinité pour les anticorps) couplé à une enzyme est ensuite ajouté dans le puits. Le reste du conjugué est éliminé par lavage et le substrat de l'enzyme est ajouté dans le puits. Si des anticorps sont présents dans le sérum testé, il y a alors, formation de complexes antigènes – anticorps – conjugué – enzyme. L'enzyme dégrade alors le substrat ce qui produit une réaction colorée dont l'intensité est proportionnelle à la quantité d'anticorps dirigés contre *Map* présents dans le sérum. Le résultat d'un test ELISA peut être quantifié par le calcul du ratio entre l'intensité de réaction de l'échantillon et celle du contrôle positif (Sample to positive ratio ; S/P). En l'absence d'anticorps dirigés contre *Map* dans l'échantillon testé le substrat n'est pas dégradé. Dans le cas de la paratuberculose, les tests ELISA peuvent être réalisés à partir d'un échantillon de sérum, de plasma, ou de lait.

En France, cinq kits ELISA indirects commerciaux sont actuellement disponibles :

- IDEXX Paratuberculosis Screening Ab Test
- ID Screen Paratuberculosis Indirect – IDVet
- SERELISA[®] M. ParaTB Ab Mono Indirect développé par Symbotics et désormais commercialisé par Zoetis
- LSIVet Ruminant Paratuberculosis Advanced Serum ELISA commercialisé par Thermo-Fisher
- PARACHEK[®] 2 (Prionics) commercialisé par Thermo-Fisher

Ils se différencient par les antigènes de *Map* utilisés dont la nature n'est pas communiquée par les fabricants.

Évaluation de la réponse immunitaire. L'évaluation des profils cytokiniques peut être considérée comme une extension du dosage de l'IFN γ . La possibilité d'identifier les animaux infectés par *Map* en mesurant les concentrations plasmatiques de différentes cytokines ou des niveaux d'expression des gènes associés a fait l'objet de plusieurs études. Lee et al. (2001) ont comparé les niveaux d'expression des gènes codant pour six cytokines différentes (IL-1 α , IL-1 β , IL-6, IFN γ , TNF α et β -Actin) dans l'iléon de sept bovins en phase clinique de la maladie et de huit bovins indemnes. Les résultats indiquent une expression augmentée des gènes codant pour l'IL-1 α , l'IL-1 β , l'IL-6, et l'IFN γ chez les bovins cliniquement affectés. Pour améliorer le potentiel diagnostique des profils cytokiniques, Coussens et al. (2004) ont comparé les niveaux d'expression, avec ou sans stimulation préalable par *Map*, des gènes codant pour 10 cytokines dans les cellules sanguines mononuclées, l'iléon et les noeuds lymphatiques mésentériques chez quatre bovins indemnes, six infectés, et quatre en phase clinique. Dans les cellules sanguines mononuclées, des différences significatives de niveaux d'expression des gènes codant pour l'IFN γ , le TGF- β , TNF- α , l'IL-1 α , l'IL-4, l'IL-6, l'IL-8, and l'IL-12p35 ont pu être mises en évidence entre les bovins indemnes et les bovins infectés et/ou en phase clinique et ceci sans stimulation préalable par *Map*. Après stimulation le gène codant pour IL-10 avait aussi des niveaux d'expression supérieurs chez les animaux infectés.

Peu d'études sont disponibles sur les niveaux d'expression mesurés à partir de prélèvements sanguins chez les ovins de plusieurs cytokines. Une augmentation significative et durable des concentrations en IL-10 a été mise en évidence entre des ovins inoculés par voie orale et le groupe témoins, dès quatre mois post-infection (de Silva et al., 2011).

Les études menées chez les ovins se sont concentrées sur les différentes formes lésionnelles de la maladie (paucibacillaire ou multibacillaire). Smeed et al. (2007) ont étudié le profil cytokinique à partir de prélèvements d'iléon de 10 ovins avec la forme multibacillaires, 10 pour le type paucibacillaires et 10 animaux sans signes cliniques. Les profils d'expression de gènes entre les groupes d'animaux sont apparus différents, comme le montre la figure 10. Des différences du niveau d'expression des gènes codant pour TNF- α sont aussi retrouvées à plusieurs reprises chez les ovins (Alzuherri et al., 1996 ; Rossi et al., 2009). En revanche, l'IL-1 α , cytokine identifiée à plusieurs reprises avec une expression

différente chez les bovins infectés ou non infectés ne semble pas retrouvée chez les ovins.

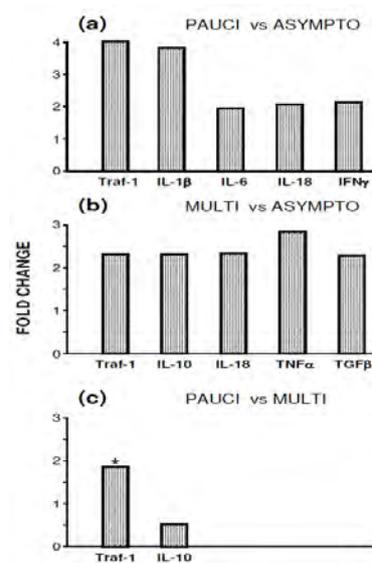


FIGURE 10 – Variation d’expression de gènes codant pour différentes cytokines en relation avec différentes formes de la maladie chez les ovins (Smeed et al., 2007)

(a) paucibacillaires *versus* asymptomatiques; (b) multibacillaires *versus* asymptomatiques; (c) paucibacillaires *versus* multibacillaires. Les résultats sont exprimés en facteur de multiplication du nombre moyen de copies des gènes évalués par PCR dans le premier groupe par rapport au second. Tous les résultats présentés sont significatifs ($p \leq 0,05$)

* retrait d’une valeur aberrante.

L’évaluation des profils cytokiniques pourrait ainsi servir de support à un test de dépistage précoce de l’infection par *Map*. En évaluant le type de réponse immunitaire présent chez l’animal, le pronostic d’évolution (résolution, installation de l’infection) pourrait par ailleurs en être déduit (de Silva et al., 2013 ; Khare et al., 2016). A l’heure actuelle ce test n’est pas applicable en pratique, pour des raisons de coût, de lourdeur de mise en oeuvre, du manque de valeurs de référence et d’absence d’évaluation des performances diagnostiques associées.

Tableau 1 – Synthèse des méthodes les plus utilisées pour diagnostic ou le dépistage de la paratuberculose

	Test	Durée d'analyse	Sensibilité analytique			
			sang	lait	fèces	tissu
Directes	Ziehl-Nielsen	jour	ND	ND	-	-
	Culture	mois	-	+	++	++
	PCR	jour	-	+	+++	+++
	phage assay	heures	+++	+++	ND	ND
Indirectes	IDT	jours	ND	ND	ND	ND
	ELISA	heure/jour	+	+	ND	ND
	IFN	heure	+++	ND	ND	ND

Sensibilité analytique relative : ND : tests non utilisable sur cette matrice, - : peu sensible, +++ : très sensible

4 Performances diagnostiques des tests biologiques

L'utilisation d'un examen complémentaire sur un animal, ou sur un groupe d'animaux, a pour objectif d'obtenir une information sur le statut sanitaire de cet animal ou de ce groupe; celle ci peut être de nature qualitative ou quantitative. En suivant la définition de Toma et al. (2010), deux objectifs peuvent être recherchés lors de l'utilisation d'un test :

- diagnostic : l'application d'un examen de laboratoire vise à confirmer ou infirmer la présence d'une maladie chez un sujet qui présente des troubles de la santé.
- dépistage : l'application d'un examen de laboratoire vise à la recherche systématique, dans une population, des individus (ou des groupes) atteints par un trouble de la santé donné, passé jusque-là inaperçu.

Les tests peuvent détecter l'agent pathogène et sont alors qualifiés de directs. La positivité d'un test direct renseigne donc sur la présence de l'agent pathogène chez l'animal contrairement aux tests indirects, qui permettent de révéler la présence de différents marqueurs induits chez l'hôte par l'infection. L'information obtenue par un test indirect ne renseigne donc pas nécessairement sur la présence de l'agent pathogène au moment du prélèvement. Le résultat d'un test, qu'il soit direct ou indirect, n'implique donc pas de façon certaine l'infection de l'animal par l'agent pathogène recherché au moment où le test est réalisé.

4.1 Indicateurs des performances des tests

Lors de l'utilisation d'un test diagnostique, il convient de définir précisément la cible recherchée (paragraphe 4.3). Dans le cadre de la paratuberculose, les cibles peuvent être i) les animaux cliniquement atteints (affectés), ii) les animaux excréant *Map* dans les fèces (infectieux) et , iii) les animaux infectés sans signe clinique ni excrétion. La réponse à un test parfait sera positive chez tous les animaux cibles et sera négative chez tous les autres. La plupart des tests n'étant pas parfaits, ils sont évalués en fonction de plusieurs indicateurs qui sont :

- **les indicateurs intrinsèques**, c'est-à-dire ceux qui ne dépendent que du test : c'est le cas de la **sensibilité diagnostique** définie comme la probabilité d'obtenir une ré-

ponse positive au test chez un animal correspondant à la cible recherchée, et de la *spécificité diagnostique* définie par la probabilité d'obtenir une réponse négative chez un animal ne correspondant pas à la cible recherchée;

- **les indicateurs extrinsèques**, c'est-à-dire qui dépendent aussi du contexte d'utilisation. C'est le cas de la *valeur prédictive positive* (VPP) définie comme la probabilité que l'animal corresponde à la cible recherchée sachant que la réponse au test est positive, et de la *valeur prédictive négative* (VPN) qui est la probabilité que l'animal ne corresponde pas à la cible recherchée sachant que la réponse au test est négative.

4.1.1 Statut de l'animal connu

Lorsque le statut réel de l'animal est connu, évaluer les performances diagnostiques d'un test consiste donc à calculer la Se_D et la Sp_D obtenues à partir de la table de contingence présentée dans le tableau 2.

Tableau 2 – Table de contingence des résultats d'un test

	Cible(C ⁺)	Non-cible (C ⁻)	
Test positif (T ⁺)	Vrais positifs (VP)	Faux positifs (FP)	VP + FP
Test négatif (T ⁻)	Faux négatifs (FN)	Vrais négatifs (VN)	FN + VN
	VP + FN	FP + VN	

La Se_D et la Sp_D d'un test sont calculées à partir du tableau 2 au moyen des équations 2 et 3

$$Se_D = P(T^+|C^+) = \frac{VP}{(VP + FN)} \quad (2)$$

$$Sp_D = P(T^-|C^-) = \frac{VN}{(FP + VN)} \quad (3)$$

Bien que qualifiées d'intrinsèques, Se_D et Sp_D vont varier en fonction de la prévalence, mais dans une moindre mesure cependant que les VPP et VPN. Brenner and Gefeller (1997) ont posé les bases mathématiques de ces variations en modélisant la variable continue de réponse du test diagnostique aboutissant à un résultat binaire. Les auteurs

concluent à une augmentation de la Se_D et à une diminution de la Sp_D lorsque la prévalence augmente et inversement lorsque elle diminue. Ces résultats théoriques ont été vérifiés par Leeflang et al. (2009) qui ont défini les facteurs influençant la prévalence, la Se_D et la Sp_D du test appliqué. Bien que ce travail ait été réalisé en médecine humaine, un des facteurs peut être transposé à la médecine vétérinaire : lorsque la prévalence dans une population change, la distribution et la sévérité des symptômes vont être modifiées, ce qui va faire varier les performances diagnostiques des tests. En conduisant une méta-analyse des estimateurs obtenus dans plusieurs méta-analyses estimant les performances diagnostiques de tests pour diverses maladies humaines, Leeflang et al. (2013) ont montré une corrélation inverse et significative entre la Sp_D et la prévalence, les variations de Se_D n'étant pas clairement associées aux variations de la prévalence.

Nous pouvons également calculer les indicateurs de performances extrinsèques avec les équations 4 et 5 :

$$VPP = P(C^+|T^+) = \frac{VP}{(VP + FP)} \quad (4)$$

$$VPN = P(C^-|T^-) = \frac{VN}{(VN + FN)} \quad (5)$$

En utilisant le théorème de Bayes, nous pouvons ré-écrire les équations 4 et 5 comme suit :

$$VPP = P(C^+|T^+) = \frac{P(T^+|C^+) * P(C^+)}{P(T^+)} = \frac{Se_D * prev}{(Se_D * prev) + (1 - Sp_D) * (1 - prev)} \quad (6)$$

$$VPN = P(C^-|T^-) = \frac{P(T^-|C^-) * P(C^-)}{P(T^-)} = \frac{Sp_D * (1 - prev)}{Sp_D * (1 - prev) + (1 - Se_D) * prev} \quad (7)$$

Les équations 6 et 7 montrent que les indicateurs de performances extrinsèques d'un test varient avec la prévalence de la cible recherchée dans la population testée ainsi qu'avec la Se_D et la Sp_D du test. Ce sont les indicateurs les plus utiles pour la prise de décision puisqu'ils renseignent directement sur le statut de l'animal testé en fonction de la réponse du test.

4.1.2 Indicateurs relatifs, indicateurs analytiques

Comme nous l'avons vu dans le tableau 2, l'évaluation des performances d'un test nécessite de connaître le statut des animaux testés. S'il existe un test parfait (i.e. dont la Se_D et la Sp_D sont égales à un), ce classement est possible et le calcul des indicateurs est trivial. En l'absence de test parfait, il est possible d'évaluer un test par rapport à un autre test dont les performances sont connues en utilisant le même principe que celui permettant d'écrire les équations 2 et 3. Les indicateurs obtenus dans ce cadre seront qualifiés de relatifs et nécessitent l'hypothèse que le test évalué a des performances moindres ou égales à celle du test imparfait de référence. Nous parlerons par opposition d'indicateurs absolus lorsque ces derniers sont calculés par rapport à un test parfait ou faisant appel à des méthodes statistiques permettant de construire une référence (cf. paragraphe 4.2.1).

Lors de l'utilisation de tests biologiques, la *sensibilité analytique* (Se_A) est définie comme la plus petite variation observable d'un caractère biologique et la *spécificité analytique* (Sp_A) comme la capacité d'un test à identifier uniquement et seulement la variation du caractère biologique évalué (Saah and Hoover, 1997). Se_A et Se_D ne décrivent pas la même probabilité, et une forte Se_A ne garantit pas une Se_D élevée. La qPCR sur fèces utilisée dans le cas de la paratuberculose illustre très bien la différence entre ces deux notions. En effet la qPCR, capable de détecter la présence d'une très faible quantité de *Map* (de l'ordre de quelques équivalents génomes par gramme de fèces) est considérée comme ayant une forte Se_A (Möbius et al., 2008). Pour autant, la physiopathologie de la paratuberculose avec notamment une longue période d'incubation et l'existence d'une excrétion intermittente en début d'évolution réduit la Se_D de l'infection par *Map* des animaux asymptomatique (Mitchell et al., 2015).

Pour plus de clarté, les termes sensibilité et spécificité devraient être employés suivis du qualificatif *analytique* ou *diagnostique* afin d'éviter les confusions.

4.2 Évaluation des performances absolues en l'absence de test parfait

4.2.1 Méthodes disponibles

Les méthodes d'évaluation des performances des tests en l'absence de test de référence parfait, suivant les recommandations de Rutjes et al. (2007) consistent à :

- Corriger les erreurs d'un test de référence imparfait;
- Construire une nouvelle référence;
- Valider les résultats d'un test imparfait par apport d'informations extérieures, le plus souvent en s'appuyant sur les signes cliniques.

La correction d'un test de référence imparfait nécessite de connaître les erreurs faites par ce dernier et d'ajuster les résultats du test de référence et du test que l'on cherche à évaluer. Cela est possible en faisant l'hypothèse, d'une part que les résultats des deux tests sont indépendants mais également que les erreurs sont indépendantes. Ces deux hypothèses, si elles ne se vérifient pas, vont conduire à une estimation biaisée des performances du test évalué. Les modèles à classes latentes peuvent être vus comme une évolution de cette méthode de correction du test de référence.

La construction d'une nouvelle référence peut se faire selon les quatre méthodes suivantes, qui peuvent éventuellement être combinées :

- utilisation d'un second test afin de résoudre les résultats discordants entre le test évalué et celui de référence;
- utilisation systématique de deux tests de référence, et utilisation d'un algorithme de tri permettant de déduire le statut de chaque individu;
- recours à un panel d'experts pour classer les individus;
- recours aux modèles à classes latentes (détaillés dans le paragraphe 4.2.2)

L'utilisation d'un second test pour résoudre les discordances est limitant puisque cela peut entraîner des erreurs de classement des individus dont les résultats ne sont pas discordants. De plus, si les résultats des différents tests sont corrélés, les biais d'évaluation seront plus grands. L'application de deux tests de référence en parallèle requiert une définition précise de l'état cible et de s'assurer que ces deux tests sont indépendants et qu'ils

ciblent le même processus biologique. Le classement des individus par un panel d'experts reste subjectif et peut conduire à des classements discordants, notamment si la définition de l'état cible n'est pas stricte.

4.2.2 Modèles à classes latentes

Principe. Les méthodes basées sur les classes latentes, initialement proposées par Hui and Walter (1980), sont de plus en plus utilisées pour l'évaluation des performances diagnostiques des tests, que ce soit en médecine humaine ou vétérinaire (Collins and Huynh, 2014). De façon schématique, la classe latente correspond à une séparation des animaux selon un critère qui n'est pas atteignable par l'observateur (non observable). Les réponses à un test diagnostique seront modélisées avec des paramètres différents selon chaque classe latente dans le but d'approcher au mieux la distribution observée des résultats dans la population d'étude. Dans la cas de l'évaluation des performances diagnostiques des tests biologiques, deux classes latentes sont généralement définies : une classe latente correspond à la cible recherchée, l'autre inclus les animaux ne correspondant pas à la cible (par exemple malades/non malades, infectés/non infectés). L'approche par des modèles à classes latentes repose sur l'analyse des résultats de plusieurs tests sur de nombreux individus permettant de construire une nouvelle référence et sur une évaluation simultanée des performances de tous les tests.

Conditionnellement à la classe latente, et en prenant l'exemple de deux tests binaires conditionnellement indépendants (T_1) et T_2) nous pouvons classer les individus selon quatre classes en fonction des résultats obtenus à chaque test pour parvenir à une description des différentes combinaisons possibles d'évènements. Ainsi pour les individus appartenant à la classe latente C^+ :

$$\left\{ \begin{array}{l} P(T_1^+ \cap T_2^+ | C^+) = P(T_1^+ | C^+) * P(T_2^+ | C^+) \\ P(T_1^+ \cap T_2^- | C^+) = P(T_1^+ | C^+) * P(T_2^- | C^+) \\ P(T_1^- \cap T_2^+ | C^+) = P(T_1^- | C^+) * P(T_2^+ | C^+) \\ P(T_1^- \cap T_2^- | C^+) = P(T_1^- | C^+) * P(T_2^- | C^+) \end{array} \right. \quad (8)$$

Le même système d'équations peut être établi pour les animaux de la classe latente

C⁻.

Toujours sous l'hypothèse d'indépendance des tests conditionnellement à la classe latente, nous pouvons donc écrire :

$$\begin{aligned}
 P(T_1^+ \cap T_2^+) &= P(T_1^+ \cap T_2^+ | C^+) * P(C^+) + P(T_1^+ \cap T_2^+ | C^-) * P(C^-) \\
 &= P(T_1^+ | C^+) * P(T_2^+ | C^+) * P(C^+) + P(T_1^+ | C^-) * P(T_2^+ | C^-) * (1 - P(C^+)) \\
 &= Se_{D1} * Se_{D2} * prev + (1 - Sp_{D1}) * (1 - Sp_{D2}) * (1 - prev)
 \end{aligned} \tag{9}$$

En appliquant cette décomposition au quatre types d'évènements observables, nous obtenons le système d'équations suivant, permettant d'explicitier les indicateurs recherchés :

$$\left\{ \begin{array}{l}
 P(T_1^+ \cap T_2^+) = Se_{D1} * Se_{D2} * prev + (1 - Sp_{D1}) * (1 - Sp_{D2}) * (1 - prev) \\
 P(T_1^+ \cap T_2^-) = Se_{D1} * (1 - Se_{D2}) * prev + (1 - Sp_{D1}) * Sp_{D2} * (1 - prev) \\
 P(T_1^- \cap T_2^+) = (1 - Se_{D1}) * Se_{D2} * prev + Sp_{D1} * (1 - Sp_{D2}) * (1 - prev) \\
 P(T_1^- \cap T_2^-) = (1 - Se_{D1}) * (1 - Se_{D2}) * prev + Sp_{D1} * Sp_{D2} * (1 - prev)
 \end{array} \right. \tag{10}$$

auquel nous pouvons ajouter l'équation suivante :

$$P(T_1^+ \cap T_2^+) + P(T_1^+ \cap T_2^-) + P(T_1^- \cap T_2^+) + P(T_1^- \cap T_2^-) = 1 \tag{11}$$

Le modèle doit alors être résolu afin de déterminer les paramètres à estimer en fonction des données observées. Dans le cas présenté, nous avons donc cinq paramètres à estimer (Se_{D1} , Se_{D2} , Sp_{D1} , Sp_{D2} et $prev$), alors que la structure des données n'autorise que trois degrés de liberté (ddl). En effet en fixant trois des quatre probabilités dans l'équation (11), la dernière est automatiquement déduite.

Un modèle peut être qualifié d'identifiable (ou inversible) s'il existe une unique série de paramètres aboutissant à la répartition observée des données. Jones et al. (2010) séparent les modèles en trois classes en fonction de leur identifiabilité :

- identifiable : le modèle est inversible sur tout l'espace des paramètres;
- localement identifiable : le modèle est inversible dans une partie de l'espace des

paramètres ;

- non identifiable : il n'existe pas d'espace, même réduit, dans lequel le modèle est inversible.

Comme montré par de nombreux auteurs (Goodman, 1974 ; Allman et al., 2009 ; Elmore et al., 2005), un nombre de ddl supérieur au nombre de paramètres à estimer est une condition nécessaire, mais non suffisante, assurant l'identifiabilité au moins locale d'un système d'équations. De plus, les modèles à classes latentes sont par nature non-identifiables du fait du phénomène de "label swapping" (Allman et al., 2009) (nommé "label-switching" par Jones et al. (2010)). Les classes latentes étant cachées aux yeux de l'observateur, une inversion des étiquettes entre population cible et population saine produira des estimateurs biaisés qui permettent d'obtenir les données observées. Le problème du "label swapping" peut être contourné, dans le cas des études d'évaluation des performances diagnostiques des tests, par recours au contexte, qui oriente les résultats en faisant l'hypothèse que les tests sont informatifs (c'est-à-dire en posant $Se_D + Sp_D > 1$).

Afin d'augmenter le nombre de degrés de liberté, il est possible d'appliquer plus de tests à la population, ou d'appliquer ces tests à un plus grand nombre de populations ayant des prévalences différentes. En effet, dans le cas d'une seule population et avec trois tests l'équation (11) devient :

$$\begin{aligned}
 &P(T_1^+ \cap T_2^+ \cap T_3^+) + P(T_1^+ \cap T_2^+ \cap T_3^-) + P(T_1^+ \cap T_2^- \cap T_3^+) + P(T_1^+ \cap T_2^- \cap T_3^-) + \\
 &P(T_1^- \cap T_2^+ \cap T_3^+) + P(T_1^- \cap T_2^+ \cap T_3^-) + P(T_1^- \cap T_2^- \cap T_3^+) + P(T_1^- \cap T_2^- \cap T_3^-) = 1 \quad (12)
 \end{aligned}$$

Ce modèle autorise sept ddl pour sept paramètres à estimer (deux paramètres par test et la prévalence dans la population). Quand deux tests sont employés dans deux populations de prévalences différentes, et en appliquant les équations 10 et 11 dans les deux populations, nous obtenons six paramètres à estimer (deux paramètres par test et les deux prévalences) ainsi que six ddl. Nous pouvons généraliser le nombre de ddl et de paramètres

(ndp) pour les modèles à t tests et p populations comme suit :

$$\begin{cases} ddl = (2^t - 1) * p \\ ndp = 2^t + p \end{cases} \quad (13)$$

Les modèles présentés jusqu'alors font l'hypothèse que les tests appliqués aux populations sont conditionnellement indépendants. Deux événements A et B sont considérés comme indépendants si :

$$P(A \cap B) = P(A) * P(B) \quad (14)$$

La notion d'indépendance conditionnelle se définit avec un troisième événement (C) comme suit :

$$P(A \cap B | C) = P(A | C) * P(B | C) \quad (15)$$

Dans le cadre des modèles à classes latentes, les événements A et B sont les résultats des deux tests et l'évènement C est la classe latente d'appartenance. De manière générale, l'indépendance conditionnelle signifie que sachant la classe latente d'appartenance (par exemple en sachant qu'un animal est infecté), le résultat à un test n'apporte pas d'informations sur le résultat des autres tests. Les tests diagnostiques peuvent donc être conditionnellement indépendants par rapport à une des classes latentes tout en étant dépendants lorsqu'il s'agit de l'autre classe latente.

Dans les faits, l'hypothèse forte d'indépendance conditionnelle des tests est rarement rencontrée, et cela même si les tests ont pour cible des phénomènes biologiques différents. Plusieurs auteurs (Vacek, 1985 ; Qu et al., 1996 ; Dendukuri and Joseph, 2001 ; Georgiadis et al., 2003) ont proposé des améliorations de ce modèle pour prendre en compte les corrélations possibles entre les réponses aux tests utilisés. Nous retrouvons deux principales approches pour écrire la dépendance dans un modèle à classes latentes. En considérant deux tests dont les résultats sur les animaux cibles sont dépendants, Dendukuri and Joseph (2001) proposent de modéliser explicitement la covariance entre le test i et le test j en ajoutant un terme (μ_{ij}) par rapport au calcul de la probabilité en cas d'indépendance, conduisant aux équations suivantes, pour la Se_D (les mêmes équations peuvent

être écrites pour la Sp_D) :

$$\begin{aligned} P(T_i^+ \cap T_i^+ | C^+) &= Se_{Di} * Se_{Dj} + \mu_{ij} \\ P(T_i^+ \cap T_j^- | C^+) &= Se_{Di} * (1 - Se_{Dj}) - \mu_{ij} \end{aligned} \quad (16)$$

Si les deux tests fournissent la même réponse, la covariance μ_{ij} est ajoutée; dans l'autre cas elle est retranchée.

Qu et al. (1996) et Georgiadis et al. (2003) définissent quant à eux la covariance ρ_{ij} entre le test i et le test j comme suit dans le cas de la Se_D (les mêmes relations valent pour la Sp_D) :

$$\rho_{ij} = \frac{P(T_i^+ \cap T_j^+ | C^+) - Se_{Di} * Se_{Dj}}{\sqrt{Se_{Di}(1 - Se_{Di})Se_{Dj}(1 - Se_{Dj})}} \quad (17)$$

L'équation précédente peut être ré-écrite comme suit :

$$P(T_i^+ \cap T_j^+ | C^+) = Se_{Di} * Se_{Dj} + \rho_{ij} * \sqrt{Se_{Di}(1 - Se_{Di})Se_{Dj}(1 - Se_{Dj})} \quad (18)$$

soit :

$$\mu_{ij} = \rho_{ij} * \sqrt{Se_{Di}(1 - Se_{Di})Se_{Dj}(1 - Se_{Dj})} \quad (19)$$

L'équation (19) permet de mieux comprendre la différence de paramétrisation entre les deux approches permettant de modéliser la covariance. La paramétrisation par μ est plus directe, facilitant la définition des bornes de l'intervalle des valeurs possibles. A l'inverse ρ ne dépend pas des paramètres (Se_D ou Sp_D) des tests corrélés, permettant d'obtenir des estimateurs de covariance comparables entre études pour lesquelles les estimateurs de Se_D ou Sp_D sont différents.

Le modèle complet, dit aussi modèle saturé, est celui qui tient compte d'une corrélation entre tous les paramètres de Se_D ainsi qu'entre tous les paramètres de Sp_D des tests. Dans le cas de l'évaluation conjointe de j tests, cette paramétrisation ajoute le nombre de paramètres suivants à estimer au modèle :

$$\sum_{i=2}^j \binom{i}{j} \quad (20)$$

Dans le cas de l'évaluation conjointe de trois tests cela représente quatre paramètres pour la Se_D et quatre pour la Sp_D .

Standardisation des rapports d'évaluation des performances diagnostiques des tests.

Pour rationaliser les protocoles d'évaluation des tests et rendre comparables les estimateurs obtenus, des standards ont émergé tout d'abord en médecine humaine (Bossuyt et al., 2003), puis en médecine vétérinaire, en prenant la paratuberculose comme exemple d'étude (Gardner et al., 2011). Cette standardisation des protocoles d'évaluation des performances diagnostiques des tests évite la production d'estimateurs biaisés (Lijmer et al., 1999). Une mise à jour spécifiquement dédiée à l'approche par modèles à classes latentes a été récemment proposée (Kostoulas et al., 2017).

La paratuberculose est emblématique des situations où le statut réel des individus est inconnu. En raison de la physiopathologie complexe de l'infection par *Map*, aucun test diagnostique ne peut en effet être considéré comme un test de référence parfait, qui assurerait la détection de toutes les cibles d'intérêt avec une Se_D et une Sp_D parfaites. C'est pourquoi, les performances diagnostiques des tests appliqués pour le diagnostic/dépistage de la paratuberculose ont fait l'objet de nombreuses études, difficilement comparables tant d'un point de vue des méthodes employées que des résultats obtenus.

4.3 Performances diagnostiques des tests à l'échelle individuelle

Nous chercherons dans un premier temps à répertorier les études d'évaluation des performances diagnostiques de tests réalisées chez les ovins, mais certains résultats obtenus dans l'espèce bovine sont aussi présentés, la littérature étant beaucoup plus fournie pour cette espèce.

Les performances des tests pour le diagnostic ou le dépistage de la paratuberculose sont extrêmement variables. Elles dépendent de l'espèce étudiée, de la trousse et/ou de la technique employées, du modèle statistique utilisé, et également de la cible définie par les auteurs. Nous nous intéresserons ici aux examens de laboratoire ante-mortem qui sont de loin les plus couramment utilisés pour le dépistage de la paratuberculose (ELISA, qPCR et

isolement par culture fécale). Les performances des tests seront comparées en fonction du statut des animaux selon la classification proposée par Nielsen and Toft (2008). Ces auteurs décrivent quatre statuts possibles vis-à-vis de la paratuberculose :

- Animal affecté : qui exprime des signes cliniques ;
- Animal infectieux : qui excrète *Map* dans les fèces ;
- Animal infecté : dont des tissus (en particulier digestifs) contiennent *Map* ;
- Animal sain : qui n'appartient à aucun des trois statuts précédents.

Les quatre statuts, présentés à la figure 11, ne sont pas exclusifs, un animal infectieux n'étant pas nécessairement infecté par *Map* (cas des animaux excréteurs passifs). Le statut d'animal au stade clinique de la maladie pourrait être une cible plus convenable en terme d'homogénéité de statut des animaux, contrairement au statut d'animal infectieux. Cependant, étant donné la longue période d'incubation durant laquelle l'animal peut être infectieux, l'identification des seuls animaux en phase clinique présente moins d'intérêt ; ils sont par ailleurs en plus petit nombre dans un élevage infecté.

4.3.1 Sérologie

Dans leur revue sur les performances diagnostiques des tests pour le diagnostic de la paratuberculose, Nielsen and Toft (2008) rapportent une Se_D de l'ELISA sérologie à partir du sérum pour la détection des ovins infectés variant de 16 à 45% dans cinq études. La Se_D augmente entre 36 et 85% lorsque les tests sont utilisés pour détecter les ovins affectés. La Sp_D varie selon les études incluses dans cette revue de 95% à 99%. Angelidou et al. (2016) ont évalué, par une approche bayésienne, les performances de la sérologie ELISA pour le diagnostic des ovins infectés par la paratuberculose dans un troupeau ovin laitier en Grèce. Les auteurs ont estimé la Se_D en fonction du stade de lactation, obtenant des moyennes estimées de Se_D entre 46% (intervalle de crédibilité à 95% (PCI 95%) : [27 - 67]) pendant la phase colostrale, et 49% (PCI 95% : [32 - 74]) en fin de lactation. La Sp_D aux mêmes stades de lactation a été estimée à 94% (PCI 95% : [88 - 97]). Robbe-Austerman et al. (2006) ont également évalué les performances diagnostiques de l'ELISA par une approche bayésienne et ont obtenu des valeurs estimées plus faibles que celles présentées précédemment (moyenne : 8,3% ; IC 95% : [6,0 - 10,4]). La Sp_D était estimée dans cette

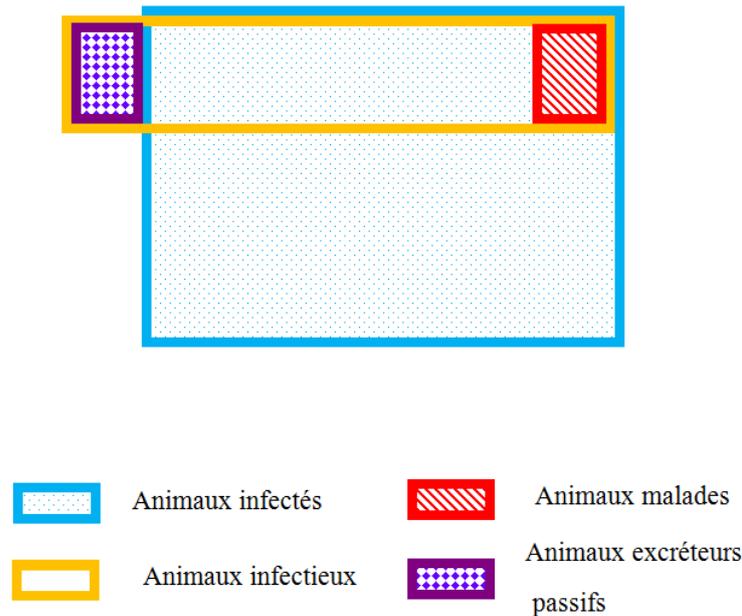


FIGURE 11 – Représentation des quatre statuts selon la classification proposée par Nielsen and Toft (2008)

Le schéma montre la relation entre les quatre statuts vis-à-vis de la paratuberculose. Il est important de noter que les animaux affectés sont aussi infectieux et infectés. Les animaux infectieux sont généralement infectés, à l'exception des animaux excréteurs passifs.

La part de chaque groupe n'est représentée qu'à titre indicatif et il n'existe pas à notre connaissance d'étude sur la part relative de ces statuts à l'échelle d'un élevage.

étude à 99,3% (IC 95% : [98,4 - 99,9]). Les variations observées entre les études peuvent s'expliquer par des différences de méthodes statistiques, de trousse ELISA utilisées ainsi possiblement que de la souche de *Map* circulant dans les populations étudiées.

4.3.2 Isolement en culture fécale

Deux études récentes traitent des performances diagnostiques de la **culture fécale** chez les ovins infectés. Kostoulas et al. (2006c) ont évalué en parallèle la culture fécale sur milieu solide HEYM et la sérologie ELISA dans un modèle bayésien à classes latentes. Les tests étaient considérés comme conditionnellement indépendants dans le modèle choisi par les auteurs. La Se_D de la culture fécale a été estimée à 16% (PCI 95% : [2 - 48]) avec une Sp_D de 97% (PCI 95% : [95 - 99]). Les performances de la culture fécale ont également été évaluées par Bauman et al. (2016). Dans le cadre d'une analyse bayésienne, les auteurs ont évalué en parallèle la culture fécale et une qPCR ciblant le gène *hsp-X* sur fèces pour le diagnostic de la paratuberculose sur 397 ovins infectés. Le modèle dans lequel les tests

étaient considérés comme conditionnellement indépendants a été retenu, et deux estimations des performances ont été produites : la première en respectant un temps d'incubation de 49 jours pour la culture sur milieu liquide BACTEC, la seconde en allongeant le temps d'incubation à 240 jours. Les valeurs estimées de Se_D étaient de 5,8% (PCI 95% : [2,3 - 12,4]) pour l'incubation courte et de 19,0% (PCI 95% : [11,9 - 28,9]) pour l'incubation de 240 jours. La Sp_D variait peu entre les deux protocoles d'incubation : 98,8% (PCI 95% : [97,8 - 99,4]) et 98,9% (PCI 95% : [97,8 - 99,5]).

4.3.3 qPCR

La PCR sur fèces pour la détection de *Map* chez les ovins a tout d'abord été évaluée par rapport à d'autres méthodes. Kawaji et al. (2011) ont comparé les performances d'une qPCR ayant pour cible le fragment *IS900* de *Map* à celles de la culture tissulaire et de la culture fécale pour la détection d'animaux infectés expérimentalement. Trente-huit ovins inoculés et 20 témoins ont été évalués à l'aide de la culture fécale et de la qPCR, au bout de quatre, huit et 13 mois après inoculation, une autopsie avec analyse histopathologique étant réalisée à la fin du suivi. De manière similaire à ce qui a été observé pour la sérologie ELISA, les auteurs ont constaté une augmentation de la Se_D relative avec l'avancée des lésions histologiques, de 80% de détection en l'absence de lésions (10 animaux) jusqu'à 100% de détection dans le cas d'animaux avec des lésions multibacillaires (10 animaux). Les performances de la qPCR pour la détection des animaux infectés qui ne présentaient pas encore de lésions (80% de détection) étaient supérieures à celles de la culture fécale (0%) ou tissulaire (10%). La culture fécale n'a permis de détecter que 33% des animaux ayant des lésions paucibacillaires, contre 83% pour la culture tissulaire et la qPCR. La Se_D globale, tous types lésionnels confondus, calculée à partir des données obtenues chez 38 animaux inoculés était de 87% pour la qPCR réalisée sur les fèces.

À notre connaissance, il n'existe qu'une seule étude, présentée précédemment, ou Bauman et al. (2016) évaluent les performances de la qPCR sur fèces pour la détection de la paratuberculose chez les ovins de manière absolue. La Se_D estimée était significativement plus élevée que celle de la culture fécale suivant un protocole court (42,6%; PCI 95% : [8,8 - 63,3]). La différence devenait non significative en comparaison avec la culture

fécale en suivant le protocole long (35,3%; PCI 95% : [24,7 - 49,8]). La Sp_D de la qPCR était évaluée entre 97,8% (PCI 95% : [93,0 - 99,9]) et 97,4% (PCI 95% : [93,5 - 99,7]).

Au bilan, les sensibilités diagnostiques des tests les plus couramment utilisés sont variables, en fonction des antigènes cibles des trousse de sérologie ELISA, du milieu de culture, du gène cible de la qPCR, mais aussi de la souche de *Map* circulant et du stade de l'infection (asymptomatique/clinique) chez les animaux échantillonnés. Toutefois nous pouvons conclure de l'ensemble des études disponibles que la Se_D de la sérologie est globalement faible (dans tous les cas inférieure à 50%, voire 30%) pour la détection des animaux en phase subclinique, elle est plus élevée pour les animaux en phase clinique. Les spécificités diagnostiques de la culture et de la qPCR sur fèces sont proches, et leurs valeurs estimées sont homogènes entre les études, entre 97% et 99%. Pour l'ELISA, la Sp_D estimée varie de 94% à 99% en fonction des études. Ainsi aucun des tests biologiques couramment utilisés ne peut être considéré comme parfait, ni en terme de Sp_D , ni encore moins en terme de Se_D , ce qui rend le diagnostic de l'infection par *Map* incertain, et ce d'autant plus si l'on cherche à détecter des animaux à la phase subclinique de l'infection.

4.4 Diagnostic à l'échelle du troupeau

Le diagnostic de la paratuberculose à l'échelle collective peut s'effectuer selon deux approches : l'une repose sur la somme de diagnostics individuels réalisés sur tout ou partie des animaux de l'élevage, l'autre s'attache à définir un statut global de l'élevage. Dans les élevages bovins français, l'approche individuelle est encore privilégiée à l'heure actuelle, les plans de maîtrise reposant souvent sur une stratégie de tests exhaustifs et de réforme anticipée des bovins considérés comme positifs. Par ailleurs le référentiel de garantie de cheptel en matière de paratuberculose bovine, établi par l'association pour la certification en santé animale (ACERSA) en 2005 n'est basé que sur des résultats individuels. De plus la forte valeur marchande des bovins par rapport aux ovins augmente la rentabilité économique relative des tests appliqués. A l'inverse, dans les troupeaux ovins, voire caprins, l'approche collective est la plus souvent adoptée. En effet la faible valeur économique individuelle et les effectifs élevés rendent inenvisageable l'application exhaustive de tests diagnostiques individuels. Dans les élevages de petits ruminants, les

plans de maîtrise, quand ils existent, sont principalement orientés vers la mise en oeuvre de mesures zootechniques et le recours à la vaccination.

L'approche collective comporte deux objectifs principaux : le premier s'attache à détecter la présence de la maladie ou de l'infection dans le troupeau, le second cherche à estimer la prévalence intra-troupeau.

4.4.1 Performances d'un test diagnostique à l'échelle du troupeau

Les performances d'un test diagnostique peuvent être calculées à l'échelle d'un élevage de manière similaire à leur évaluation au niveau individuel. En se plaçant dans le cadre de la détection d'élevages infectés, nous considérerons que l'obtention d'une seule réponse positive à un test suffit à considérer l'élevage comme positif. Les performances du test à l'échelle de l'élevage peuvent être calculées en connaissant les performances individuelles des tests, de la prévalence de la cible recherchée et du nombre de tests réalisés. La **sensibilité diagnostique de troupeau** (Se_T) est définie comme la probabilité d'obtenir au moins un test positif, sachant que l'élevage est infecté. En considérant un troupeau d'effectif N avec une prévalence notée $prev$ et en appliquant de manière exhaustive un test de Se_D Se et de Sp_D Sp à l'échelle individuelle, nous pouvons calculer la probabilité que tous les tests soient négatifs ($P(T^-)$) de la manière suivante :

$$P(T^-) = (1 - Se)^{N*prev} + Sp^{N*(1-prev)} \quad (21)$$

Nous pouvons donc en déduire Se_T :

$$Se_T = 1 - (P(T^-)) \quad (22)$$

Un des objectifs principaux de l'approche collective étant la réduction des coûts analytiques, le dépistage exhaustif n'est que rarement réalisé en pratique. Il faut donc tenir compte du taux de sondage ainsi que de la prévalence apparente (PA) dans le groupe considéré. La prévalence apparente est la proportion d'animaux réagissant positivement au test et dépend à la fois de la prévalence et des performances intrinsèques du test ap-

pliqué :

$$PA = Se * prev + (1 - Sp)(1 - prev) \quad (23)$$

La probabilité de n'obtenir que des résultats négatifs en testant t animaux est égale à

$$P(T^-) = \frac{\binom{(1-PA)*N}{t}}{\binom{N}{t}} = \frac{((1-PA) * N)! * (N! * t! * (N - t)!)}{t! * ((1-PA) * N - t)!} \quad (24)$$

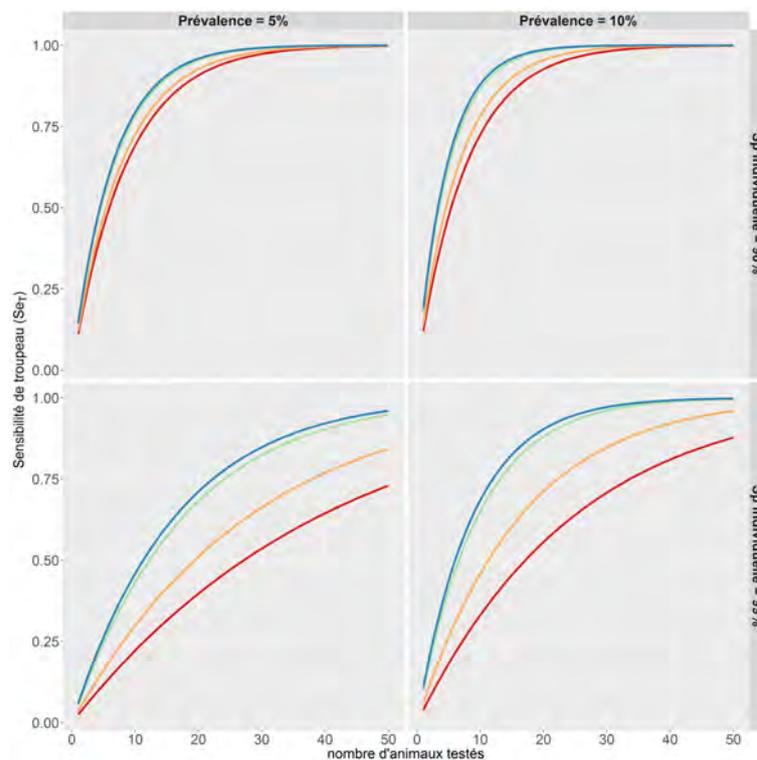


FIGURE 12 – Représentation de la variation de la sensibilité à l'échelle d'un élevage Se_T en fonction du nombre d'animaux testés, de la prévalence réelle, et des performances intrinsèques du test utilisé à l'échelle individuelle.

Un élevage de 500 animaux est simulé avec deux prévalences réelles différentes (0.05 dans le panel de gauche et 0.10 à droite). La Se_T est représentée en fonction du nombre d'animaux testés (entre 1 et 50) ainsi que des performances intrinsèques (Se_D et Sp_D) du test utilisé à l'échelle individuelle. Les résultats sont présentés pour quatre valeurs de Se_D (0.30 en rouge, 0.50 en orange, 0.90 en vert et 0.99 en bleu) et deux valeurs de Sp_D : 0.90 (en haut) et 0.99 (en bas).

La figure 12 présente les variations de Se_T en fonction du nombre d'animaux testés dans un groupe de 500 animaux, de la prévalence réelle, et des performances intrinsèques du test utilisé. Pour un nombre fixé d'animaux testés, nous observons que la Se_T augmente avec la prévalence réelle et avec la Se_D individuelle. La Sp_D individuelle du test utilisé influence également fortement la Se_T , une diminution de celle-ci augmentant le

nombre de résultats faussement positifs et donc également la Se_T .

De la même manière, nous pouvons calculer la **spécificité troupeau** (Sp_T) qui correspond à la probabilité de n'obtenir que des tests négatifs dans un élevage indemne. Nous obtenons la même équation que précédemment, le seul changement étant la prévalence apparente qui devient :

$$PA = (1 - Sp) \quad (25)$$

La figure 13 présente les variations de la Sp_T en fonction du nombre de tests individuels appliqués et de la Sp_D individuelle du test utilisé. La Sp_T diminue avec le nombre de tests réalisés.

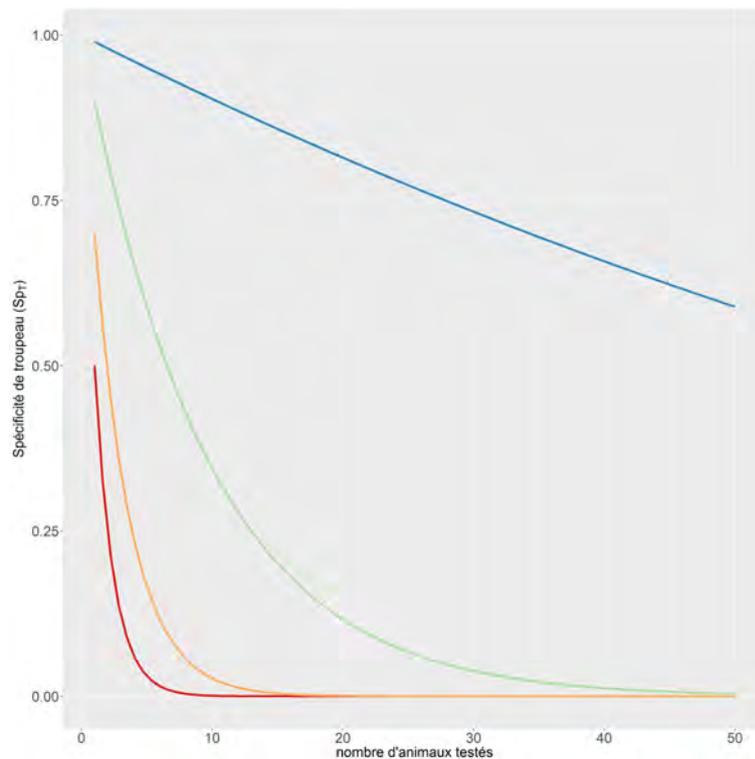


FIGURE 13 – Variation de la spécificité à l'échelle d'un élevage Sp_T , en fonction du nombre d'animaux testés et de la spécificité du test utilisé à l'échelle individuelle.

Dans un élevage de 500 animaux indemnes, la Sp_T est représentée en fonction du nombre d'animaux testés (entre 1 et 50) ainsi que de la Sp_D individuelle du test utilisé (0.50 en rouge, 0.70 en orange, 0.90 en vert et 0.99 en bleu).

Initialement basée sur des analyses individuelles, l'approche troupeau peut également être réalisée à partir de prélèvements environnementaux ou par des analyses en mélanges de prélèvements. Ces deux méthodes permettent une réduction notable des coûts analytiques.

4.4.2 Analyses de mélanges d'échantillons

Les analyses des mélanges d'échantillons, ou analyses de mélanges, ont tout d'abord été utilisées dès les années 1940 chez l'homme pour la détection des individus atteints de maladies à faibles prévalences. Dorfman (1943) proposait pour la première fois l'analyse de mélanges d'échantillons d'urine afin de réduire les coûts d'analyses dans le cadre de la détection de la syphilis chez les soldats américains pendant la seconde guerre mondiale. Seuls les mélanges positifs faisaient l'objet d'une investigation complémentaire à l'échelle individuelle, les mélanges négatifs étant considérés comme composés uniquement de sérums individuels négatifs.

Les analyses de mélanges permettent également d'approcher la prévalence d'une maladie dans une population, augmentant le nombre d'individus testés pour le même coût, le coût de prélèvement étant généralement largement inférieur à celui de l'analyse biologique. Cette approche a été appliquée en médecine humaine pour étudier la prévalence de circulation du virus HIV dans la population (Tu et al., 1995).

Les analyses en mélanges ont largement été utilisées en médecine vétérinaire, que ce soit pour l'identification des animaux atteints par une maladie à faible prévalence intra-troupeau (comme c'est le cas pour le dépistage de l'infection par le virus herpès bovin de type I, virus de la rhinotrachéite infectieuse bovine (IBR) dans le cadre de la prophylaxie obligatoire en France) ou pour privilégier une approche de troupeau. C'est notamment le cas dans les élevages ovins du fait de la faible valeur économique des animaux et la grande taille des troupeaux comparativement aux cheptels bovins.

Selon le test utilisé, les performances diagnostiques des analyses de mélanges peuvent être similaires à celles observées à l'échelle individuelle pour des tailles de mélange faibles (Litvak et al., 1994). L'intensité du phénomène enregistré par le test dans les échantillons individuels va également modifier les performances diagnostiques des analyses de mélanges.

Culture de mélanges de fèces. Pour le dépistage de la paratuberculose, les analyses de mélanges sont principalement réalisées avec la culture fécale sur des mélanges de taille cinq à 10 pour les bovins et jusqu'à 25 ou 50 pour les ovins et les caprins (Whittington

et al., 2000a ; Sergeant et al., 2002 ; Kalis et al., 2004 ; Eamens et al., 2007 ; van Schaik et al., 2007 ; Aly et al., 2010 ; Dhand et al., 2010a ; Raizman et al., 2011). Whittington et al. (2000a) ont montré une différence de sensibilité de détection des mélanges selon le niveau d'excrétion des ovins composant ces mélanges. Pour 20 ovins avec la forme paucibacillaire de la maladie (animaux faibles excréteurs) et 20 ovins avec la forme multibacillaire (forts excréteurs), les auteurs ont réalisé des mélanges avec des fèces d'ovins sains de taille croissante entre 10, 50 et 100 (1g de fèces par animal). Après culture sur milieu BACTEC 12B, la sensibilité de détection pour les mélanges contenant les fèces d'un animal avec la forme paucibacillaire pour une taille de mélange de 50, était de 50% (IC 95% : [27 - 73]) et de 100% (IC 95% : [83 - 100]) pour les mélanges contenant les fèces d'un ovin avec la forme multibacillaire.

qPCR sur mélanges de fèces. Les performances diagnostiques de la qPCR sur mélanges de fèces ont été évaluées par Mita et al. (2015) à partir de fèces de deux bovins forts excréteurs. Trois tailles de mélanges ont été étudiées (cinq, 10 et 50) et les fèces des bovins ont été diluées pour simuler des fèces provenant de bovins plus faiblement excréteurs. Plusieurs trousse d'extraction ont été employées avant l'étape de qPCR ciblant le fragment *IS900*. Les performances diagnostiques obtenues sont variables : les variations de la Se_D sont les mêmes que celles observées pour la culture fécale, les animaux fortement excréteurs et les mélanges de petite taille rendant la détection plus facile.

qPCR sur lait de grand mélange ou lait de tank. Les données relatives à l'utilisation de la qPCR sur le lait de tank chez les bovins rapportent un seuil de détection de 10 à 100 UFC/mL (Pillai and Jayarao, 2002). L'étude de Wilson et al. (2010) montre que la méthode qPCR (basée sur la détection d'*IS900*) et la trousse commerciale ELISA Pourquier sur des prélèvements de lait de tank (deux prélèvements à un mois d'intervalle) ont une concordance globale de 76%, mais elle se réduit à 53% pour les troupeaux infectés. Herthnek et al. (2008) considèrent eux que l'utilisation de la qPCR sur le lait serait plus utile pour contrôler la présence de *Map* dans le lait, dans le but d'éviter l'exposition des consommateurs, que pour évaluer la prévalence des troupeaux infectés.

Sérologie ELISA sur mélanges de sérums. A notre connaissance, aucune étude n'a été publiée sur l'utilisation de la sérologie ELISA sur des mélange de sérums pour le dépistage de la paratuberculose. Cette approche a pourtant été appliquée à d'autres maladies (Visna-Maedi chez les ovins (Brinkhof et al., 2007), syndrome dysgénésique et respiratoire du porc (SDRP) chez le porc (Rovira et al., 2008) par exemple). Comme pour la qPCR ou la culture fécale, la Se_D diminue avec la taille du mélange parfois dès une tailles de mélanges petite (quatre à huit échantillons) (Brinkhof et al., 2007). Dans les deux études citées, un abaissement du seuil à partir duquel la positivité de l'ELISA est considérée a été appliqué en fonction de la taille des mélanges.

ELISA sur lait de grand mélange ou lait de tank. L'analyse du lait de tank peut être vue comme une étape ultime d'exploitation de grands mélanges. Nielsen et al. (2000), ont évalué la Se_T et la Sp_T d'un ELISA (M. paratuberculosis souche 18, Allied Monitors, Fayette, Missouri, USA) adapté pour l'évaluation des échantillons de lait de tank provenant de 45 troupeaux bovins danois. La culture fécale servait de test de référence. La Sp_T de ce test a été évaluée sur 96 échantillons de lait de tank provenant d'élevages norvégiens considérés comme exempts de paratuberculose au moment de l'étude. Au seuil de l'ELISA statistiquement optimal, la Se_T était de 97,1% et la Sp_T de 83,3%. Les auteurs concluent que les performances techniques de cet ELISA ne sont pas suffisantes pour en faire un outil de surveillance efficace car de faibles variations de densité optique peuvent faire changer le classement de l'échantillon.

La Se_T et la Sp_T d'une sérologie ELISA (Pourquier) ont été évaluées dans une étude (Delbrouck, 2008) menée sur 383 échantillons de lait de tank. La méthode de référence était la sérologie individuelle et les troupeaux étaient classés en fonction de la séroprévalence apparente. En utilisant le seuil recommandé par le fabricant de l'ELISA (valeur S/P de 30%), la performance de l'ELISA sur lait de tank était plutôt faible, correspondant à une Se_T de 24% et une Sp_T de 99%. Cependant à un seuil de décision revu à 12,5% pour le lait de tank, et une séroprévalence intra-troupeau de 3%, la Se_T et la Sp_T étaient respectivement de 85%, et 96%.

Une autre étude porte sur l'analyse à l'aide de deux troussees ELISA (Idexx Paratuber-

culosis Screening Ab Test et Mycobacterium paratuberculosis test kit Parachek 2, Prionics AG) de 203 échantillons de lait de tank en comparaison de la séroprévalence apparente au sein des élevages par la sérologie ELISA (Pesqueira et al., 2017). La Se_T du test sur le lait de tank dans des troupeaux fortement infectés (séroprévalence apparente $\geq 10\%$) était respectivement de 85,7% et 71,4% pour les 2 troussees évaluées. La Sp_T pour détecter des troupeaux sans animal séropositif variait de 70,5% à 53%. Ainsi, selon les auteurs, bien que les résultats du test étaient relativement bien corrélés à la prévalence intra-troupeau, la faible Sp_T obtenue limitait fortement l'utilité pratique de l'ELISA sur lait de tank.

4.4.3 Analyse d'échantillons d'environnement

L'excrétion fécale et la grande résistance de *Map* dans l'environnement rend possible l'utilisation d'échantillons de matières fécales prélevées au sol, de pèdisacs appliqués au sol, voire de lisier pour la détection directe de l'agent pathogène, par culture ou qPCR. Cette approche s'appuyant sur l'échantillonnage de lieux de passage des animaux pour obtenir une bonne représentativité de l'élevage, elle est principalement utilisée dans les élevages bovins laitiers (Lombard et al., 2006 ; Berghaus et al., 2006b ; Tavorntpanich et al., 2008 ; Pillars et al., 2009 ; Aly et al., 2009 ; Guatteo et al., 2012 ; Moravkova et al., 2012 ; Eisenberg et al., 2013 ; Donat et al., 2015 ; Arango-Sabogal et al., 2016 ; Hahn et al., 2017), un exemple étant retrouvé dans les élevages caprins (Frederic, 2016).

Les performances diagnostiques des analyses d'échantillons environnementaux sont difficilement comparables entre études du fait de la diversité des lieux et des méthodes de prélèvement appliquées.

La culture à partir d'échantillons environnementaux n'apparaît pas très sensible pour détecter les troupeaux à faible prévalence d'excrétion. Ainsi, Pillars et al. (2009) ont montré que lorsque la prévalence d'excrétion intra-troupeau était supérieure à 2%, *Map* était identifié par la culture d'échantillons réalisée dans l'aire d'exercice des animaux et/ou de fumier dans seulement 75% des cas. Quand la prévalence intra-troupeau était inférieure à 2%, aucune culture n'était positive. Ainsi, la proportion de résultats positifs sur échantillons d'environnement était significativement corrélée au nombre d'individus séropositifs dans l'élevage suggérant que les prélèvements environnementaux puissent indiquer

une estimation qualitative de la prévalence intra-troupeau.

Sur la base de simulations, Tavornpanich et al. (2008) ont estimé la Se_T pour différentes méthodes (ELISA, coproculture individuelle, en mélange et sur prélèvements environnementaux). Parmi les méthodes de tests évalués dont la Sp_T serait de 100%, la culture sur prélèvements environnementaux était la méthode la plus efficace du point de vue économique, quelle que soit la prévalence intra-troupeau (5%, 16% ou 35%). La Se_T permise par la culture de quatre échantillons environnementaux était respectivement de 99% et 100% pour des prévalences intra-troupeaux de 16% et 35%. En revanche, même en augmentant le nombre de prélèvements environnementaux jusqu'à 10, la Se_T était seulement de 38% dans un contexte de faible prévalence (5%). Les auteurs soulignaient que la Se_T est fortement associée à la prévalence intra-troupeau, au niveau d'excrétion dans les fèces par les vaches infectées, et au nombre d'échantillons testés. Ainsi, dans cette étude, les auteurs ont fait l'hypothèse que chaque prélèvement environnemental était la contribution équivalente, des vaches infectées et de celles qui ne l'étaient pas partageant le même environnement. Cependant il est possible que les vaches au stade clinique de la paratuberculose (avec de la diarrhée) ou excréant massivement des bactéries *Map* dès la phase pré-clinique contaminent l'environnement de façon plus fréquente ou plus étendue, à l'origine d'une Se_T sur prélèvements environnementaux plus grande.

Wolf et al. (2015) se sont intéressés à l'influence de la zone de prélèvement dans l'exploitation, de la taille du troupeau, et de la saison à laquelle étaient réalisés les prélèvements. Une meilleure sensibilité de détection était obtenue lorsque les prélèvements étaient faits dans des zones où les fèces de nombreuses vaches s'accumulent et se mélangent. (Berghaus et al., 2006a) observent que les échantillons prélevés autour des abreuvoirs donnent significativement plus de résultats positifs (15/23; 65%) que ceux prélevés derrière les cornadis ou à la sortie de la salle de traite (respectivement 8/22; 36% et 9/23; 39%). Kruze et al. (2013) identifient quant à eux la zone d'attente et la zone de stockage du fumier comme les deux sites les plus fréquemment positifs à la culture environnementale (respectivement 35% et 33% des prélèvements positifs).

Plus récemment, Hahn et al. (2017) ont comparé les performances diagnostiques de deux trousse de qPCR ciblant le gène *IS900* (Adiavet[®] Paratb Real time kit (Adiagene) et

VetMAX[®] MAP Real-Time PCR screening kit (Life Technologies)), et de la culture (milieu HEYM) pour la détection de *Map* sur deux matrices environnementales : des pédisacs ainsi qu'un échantillon de lisier. Soixante-dix sept élevages bovins laitiers ont été visités (19 indemnes de paratuberculose et 58 ayant des prévalences d'infection variables), les sensibilités de détection des deux trousseaux n'étaient pas statistiquement différentes et variaient entre 61,40% et 67,92% en fonction de la matrice analysée. La Sp_T observée durant cette étude était imparfaite pour les deux qPCR en analysant le lisier (kit A : 94,12%; IC 95% : [71,31 – 99,85] et kit B : 88,24%; IC 95% : [63,56 – 98,54]) et était légèrement supérieure lors de l'analyse des pédisacs (kit A : 100,00%; IC 95% : [80,49 – 100,00] et kit B : 88,24%; IC 95% : [72,71 – 99,86]). Les performances diagnostique de la culture n'ont pas été évaluées à proprement parler par les auteurs; ils relatent cependant un plus grand nombre d'échantillons positifs avec la culture qu'avec la qPCR, sans que cette augmentation soit significativement différente.

Les prélèvements issus de l'environnement semblent donc de bons prédicteurs du statut du troupeau et ce, à des coûts moindres. Ils ne nécessitent en outre aucune intervention sur les animaux. Cette approche permet de discriminer les élevages infectés contenant des animaux excréteurs des élevages indemnes ou éventuellement infectés mais où l'excrétion ou la proportion d'animaux excréteurs est faible. Cependant, la Se_T de ces tests reste fortement dépendante de la prévalence intra-troupeau. Ainsi, la culture ou la qPCR sur prélèvements environnementaux ont été considérés comme méthodes d'analyse dans le programme Voluntary Bovine Johne's Disease Control Program (VBJDCP), mise en place aux Etats-Unis pour identifier les troupeaux à faible risque d'infection par *Map*. Les stratégies approuvées pour le dépistage initial de chaque troupeau, nécessaire à la détermination préalable de leur statut, incluent désormais trois modalités alternatives (VBJDCP, 2010) : i) sérologie ELISA sur 30 vaches de plus de trois ans, sélectionnées aléatoirement, ii) culture ou qPCR de six échantillons environnementaux de trois emplacements ou iii) culture de 10 mélanges de fèces de cinq vaches de plus de trois ans.

Enfin, nous citerons l'étude de Eisenberg et al. (2012) dans laquelle *Map* a été mise en évidence dans les poussières accumulées et les filtres à air (extracteurs) par qPCR *IS900*,

suggérant que ces échantillons pourraient aussi être utilisés pour le diagnostic de troupeau.

4.4.4 Du statut de troupeau à la prévalence intra-troupeau

Les approches par analyse de mélanges ou analyse d'échantillons environnementaux ont initialement été utilisées pour déterminer la présence ou non d'une infection par *Map* dans l'élevage. Cependant, l'estimation de la prévalence individuelle à partir de ces données pourrait aussi être intéressante.

Cowling et al. (1999) proposent plusieurs approches statistiques permettant d'estimer la prévalence à partir des résultats des analyses de mélanges. Ces développements ont été appliqués par Toribio and Sergeant (2007) au cas de la paratuberculose chez les ovins en comparant les prévalences estimées par les différentes méthodes proposées par Cowling et al. (1999) à la prévalence individuelle à l'aide de la culture fécale dans trois élevages. Certaines de ces méthodes tenaient compte du caractère imparfait des tests diagnostiques. Par ailleurs les auteurs proposaient une approche bayésienne basée sur une information *a priori* des performances diagnostiques des tests et de la prévalence. Sous l'hypothèse d'un test parfait, les intervalles de confiance des estimateurs de la prévalence individuelle contenaient à chaque fois la prévalence observée. Sous l'hypothèse d'un test imparfait, les estimateurs de la prévalence étaient généralement supérieurs aux prévalences déterminées par culture fécale. Enfin le modèle bayésien est apparu très sensible à l'information *a priori* portée sur la prévalence.

Les méthodes statistiques, notamment bayésiennes ont été améliorées, de sorte qu'elles peuvent désormais prendre en compte des tailles de mélanges différentes et le caractère imparfait des tests diagnostiques (Dhand et al., 2010a). Ces développements ont fait l'objet de la production d'un module d'analyse pour le logiciel R (RSurveillance disponible à l'adresse <https://cran.r-project.org/web/packages/RSurveillance/RSurveillance.pdf>) et sont accessibles par un calculateur en ligne sur le site <http://epitools.ausvet.com.au>.

5 Prévention et maîtrise

Il est classique de distinguer les mesures de prévention à déployer dans les élevages indemnes de paratuberculose dans le but de prévenir l'introduction de la maladie dans l'élevage, des mesures de maîtrise utiles à réduire la prévalence, voire éradiquer, si tenté que cela soit possible, l'infection dans les élevages déjà touchés par la paratuberculose.

5.1 Mesures préventives

Dans les élevages considérés indemnes de paratuberculose, les mesures de prévention visent à limiter les risques liés à l'introduction d'animaux infectés et les contacts entre troupeaux via des programmes de certification ou de garantie.

Nielsen and Toft (2011) ont analysé les pratiques relatives à l'achat et à l'introduction de bovins dans 1081 élevages inscrits dans le programme de contrôle national au Pays-Bas. A l'aide d'un modèle tenant compte de la durée d'adhésion au plan de contrôle, ils ont montré qu'un fort taux de bovins nouvellement introduits (>15% de la taille de l'élevage par an) constituait un risque accru d'augmenter la prévalence intra-élevage, sur la base d'une évaluation par ELISA sur le lait, comparé aux élevages ayant un taux de renouvellement par achat inférieur à 15%. Plus les mesures bénéfiques étaient appliquées longtemps plus la diminution de prévalence était grande. Pour faciliter la maîtrise des achats de bovins, en prenant en compte la faible sensibilité de dépistage de la paratuberculose, des plans de certification ont été mis en place. Ils reposent généralement sur le classement des élevages en fonction de leur prévalence estimée ou apparente, les élevages les moins touchés ne pouvant pas acheter des animaux auprès des élevages les plus touchés. De tels plans de certifications ont été mis en place en Australie, au Canada, et au Pays-Bas (Benedictus et al., 2000).

En France, un référentiel technique de garantie des cheptels en matière de paratuberculose bovine a été proposé en 2005 par l'ACERSA. Compte tenu de la faible Se_D des tests utilisables, aucune garantie absolue ne peut être délivrée, mais le niveau de garantie proposé permet une réduction du risque par rapport à un cheptel tout venant. Cette garantie est obtenue après deux contrôles négatifs successifs espacés dans le temps (neuf mois

minimum à 30 mois maximum) obtenus sur l'ensemble des bovins reproducteurs de plus de deux ans. Ces contrôles sont basés sur l'utilisation à l'échelle individuelle de la sérologie ELISA et/ou de la culture/qPCR fécales. Un dépistage tous les ans (de neuf à 15 mois après le second contrôle d'acquisition de la garantie) puis tous les deux ans (entre 21 et 27 mois) à l'aide des mêmes tests est ensuite nécessaire au maintien de la garantie. En cas de résultat positif à la sérologie ELISA sur un ou au maximum 2% de l'effectif testé, une confirmation est effectuée par qPCR ou culture fécale, permettant le maintien ou non de la garantie selon l'issue de ces examens de confirmation. Dans un élevage sous garantie, les bovins introduits sont soumis, après avoir atteint l'âge de 18 mois, à deux tests de dépistage entre neuf et 15 mois d'intervalle s'ils proviennent d'un élevage non garanti. Pour les animaux provenant d'un cheptel sous garantie, aucun contrôle à l'introduction n'est nécessaire.

Dans la région Limousin, plus de 850 cheptels bénéficient de cette garantie ACERSA Paratuberculose. Elle entre dans le cahier des charges des élevages sélectionneurs en race bovine Limousine. A l'échelle de la France, ce dispositif reste cependant limité, et concerne le plus souvent les élevages vendeurs d'animaux de très haute valeur génétique ou de mâles reproducteurs destinés aux stations de testage. Force est de constater que la plupart des éleveurs de bovins laitiers ou allaitants, régulièrement vendeurs d'animaux reproducteurs (génisses, vaches en lactation), ne connaissent pas le statut de leur troupeau vis-à-vis de l'infection par *Map* et contribuent donc à disséminer l'infection.

5.2 Mesures de maîtrise

En plus des mesures de prévention décrites au paragraphe précédent, des mesures de maîtrise supplémentaires peuvent être mises en place dans les élevages infectés par la paratuberculose.

5.2.1 Mesures zootechniques

L'objectif des mesures zootechniques va être de diminuer la probabilité de transmission à l'intérieur du cheptel et de réduire la contamination de l'environnement par *Map*. Tenant compte du fait que la réceptivité à l'agent de la paratuberculose décroît avec l'âge

(McGregor et al., 2015), ces mesures s'appliqueront pour limiter l'exposition des jeunes à des matières infectieuses (fèces, colostrum et dans une moindre mesure le lait). Nielsen et al. (2008) ont étudié les pratiques d'élevage autour de l'alimentation colostrale et lactée en les mettant en relation avec le statut sérologique de 93994 bovins laitiers, obtenu par ELISA sur lait individuel. A partir de 808 questionnaires rapportant les conduites d'élevage de troupeaux bovins laitiers inscrits dans un plan de contrôle national au Pays-Bas, les auteurs ont construit un modèle linéaire généralisé identifiant les pratiques à risque. Les facteurs qui augmentent le risque sont :

- l'alimentation colostrale ou lactée avec mélange de la production de plusieurs vaches;
- l'adoption des veaux par une autre vache

L'alimentation du veau avec un mélange de colostrum (OR : 1,243; IC95% : [1,089 - 1,418]) ou de lait (OR : 1,137; IC 95% : [1,000 - 1,292]) de plusieurs vaches constituent des pratiques à risque d'obtenir un résultat positif à un test ELISA sur le lait. Bien que significatif ou proche de l'être, la puissance d'association reste mineure concernant ces deux pratiques, ce qui peut s'expliquer par une excrétion faible dans les sécrétions mammaires, comparée à l'excrétion fécale. L'adoption des veaux par d'autres vaches a également été identifiée comme un facteur de risque (OR : 2,012; IC 95% : [1,370 - 2,956]), mais cette pratique n'a été recensée que pour 934 bovins (0,994% de l'effectif total) et peut refléter une contamination par voie lactée ou fécale, le veau adopté étant en contact proche avec une vache potentiellement infectieuse.

La séparation stricte des jeunes et des adultes est souvent recommandée afin de rompre la transmission pseudo-verticale de la paratuberculose. Marcé et al. (2011) ont modélisé l'évolution de la prévalence de la paratuberculose pendant 25 ans après introduction d'un animal infectieux. Les paramètres diminuant significativement la prévalence étaient le taux d'exposition des jeunes (rythme de nettoyage des boxes, contact réduit avec les adultes), ainsi que la séparation stricte des classes d'âge extrêmes. Norton et al. (2009) ont vérifié l'association de ce dernier paramètre avec la prévalence à l'aide d'un questionnaire envoyé à des éleveurs. Dans cette étude, si les veaux étaient mélangés avec les adultes à un âge plus tardif, la probabilité d'être classé parmi les élevages ayant une prévalence élevée était réduite.

Collins and Morgan (1992) ont exploré à l'aide de simulations différentes stratégies de maîtrise (identification et réforme des animaux infectés, séparation précoce des veaux, ou les deux associés) en élevage bovin laitier. Les résultats sont en faveur d'une meilleure efficacité des mesures lorsqu'elles sont appliquées en parallèle. Une étude en élevages, menée par les mêmes auteurs, a permis de vérifier les résultats de ces simulations (Collins et al., 2010). Neuf élevages laitiers ont été suivis durant 6 ans, le troupeau reproducteur étant testé par sérologie ELISA une fois par lactation. Les éleveurs devaient réformer les animaux fortement séropositifs et respecter certaines règles de gestion zootechnique :

- La séparation des boxes de vêlage en deux secteurs (en fonction des résultats ELISA) ;
- La séparation des veaux dans les deux heures suivant la naissance et ce jusqu'au sevrage ;
- la distribution de colostrum provenant uniquement de vaches négatives à la sérologie ELISA ;
- La distribution de lait pasteurisé ou de lait en poudre uniquement ;

Après six ans d'application de ces pratiques d'élevages, huit des neuf élevages ayant participé à l'étude jusqu'à son terme ont montré une diminution significative de la prévalence apparente, la prévalence cumulée entre les neuf élevages diminuant de 11,6% à 5,6% à la fin du suivi. L'étude était ainsi construite qu'il était impossible de hiérarchiser l'importance des mesures mises en place par les éleveurs. De plus, certaines des mesures préconisées ne semblent pas efficaces pour prévenir complètement l'exposition des veaux à *Map*, des mycobactéries vivantes pouvant être retrouvées dans les laitiers après thermisation (Grant et al., 2017).

Les mesures visant la séparation précoce et durable des jeunes animaux, bien que contraignantes, sont envisageables dans les troupeaux bovins laitiers, mais ne peuvent pas être mises en oeuvre dans les troupeaux allaitants et les élevages ovins, dans lesquels les jeunes animaux restent en contact prolongé avec leur mère.

5.2.2 Plans de maîtrise en France

En France, un groupe de travail associant GDS France et la société nationale des groupements techniques vétérinaires (SNGTV) a été constitué en 1999 afin d'élaborer un programme national pour la maîtrise de la maladie en matière de paratuberculose. Sont présentés ici les éléments fondamentaux de la mise en place d'un plan d'assainissement dans un élevage vis-à-vis de la paratuberculose.

Tout animal suspect doit être immédiatement isolé de façon à limiter les risques de contamination d'autres animaux et du milieu. La confirmation de l'existence d'un foyer de paratuberculose clinique nécessite d'avoir recours à l'outil analytique (qPCR sur fèces, sérologie ELISA). Le plan de lutte dans les élevages à foyer confirmé de paratuberculose clinique s'appuie sur deux catégories de mesures fondamentales d'importance identique :

- la détection et la réforme précoces des animaux excréteurs et de leur dernier descendant;
- la maîtrise sanitaire des risques de contamination au sein de l'effectif.

Les objectifs de **la détection et de la réforme précoces des bovins excréteurs** sont d'une part, de limiter la contamination du milieu et d'abaisser le risque de contamination des autres bovins, et d'autre part de détecter les animaux qui sont le plus susceptibles de développer une paratuberculose clinique. La fréquence de dépistage recommandée est une fois par an. Les animaux de moins de 24 mois sont exclus du dépistage, sauf dans les cheptels où la prévalence est élevée (taux d'excréteurs grand) ou si des cas cliniques sont observés sur de jeunes animaux (avant le premier vêlage). Dans ce cas les bovins de 12 à 24 mois, au moins la première année, seront également testés. Les animaux positifs doivent être isolés des autres animaux, en particulier des jeunes, et réformés dans un délai si possible court (moins de 6 mois). De même le dernier veau né de la vache concernée (ou toute la descendance s'il s'agit d'un cas clinique avéré) sera réformé. En cas d'effectif trop grand, cela implique une conduite isolée et une gestion très suivie du lot des animaux positifs. Deux techniques d'analyse peuvent être utilisées : la qPCR et la sérologie ELISA. L'importance de l'atteinte clinique et du nombre d'animaux positifs, la proximité veau – vache (élevage allaitant) orientent vers le choix de la qPCR étant donné son caractère

plus informatif sur la contamination du milieu extérieur. Les bovins avec une diarrhée chronique sont isolés et font l'objet d'un prélèvement de fèces et de sang pour recherche individuelle.

La maîtrise sanitaire des risques de contamination au sein de l'effectif représente un pôle d'action tout aussi important. Il vise à réduire les risques de contact entre les jeunes animaux et les matières infectieuses. Les mesures suivantes sont préconisées :

- **Conduite d'élevage des veaux** : le local de vêlage doit être propre, régulièrement désinfecté, suffisamment paillé et ne doit en aucun cas servir de local d'infirmierie. En élevage laitier, la séparation du veau de sa mère dès la naissance en veillant à une prise de colostrum adaptée est fortement conseillée. En élevage allaitant, une attention particulière doit être portée à l'hygiène au quotidien des bâtiments d'élevage qui abritent les couples mères - veaux dans les premières semaines de vie ;
- **Maîtrise des effluents d'élevage** : le stockage des effluents et leur écoulement doivent être parfaitement maîtrisés au sein de l'élevage. Le plan d'épandage doit prendre en compte le fait de ne pas épandre (ou seulement après stockage de six mois minimum) des déjections sur des parcelles accueillant des bovins de moins d'un an.
- **Désinfection** : le nettoyage et la désinfection réguliers des matériels et des bâtiments d'élevage doivent être réalisés, notamment ceux accueillant les animaux jusqu'à l'âge d'un an.
- **Environnement** : les points d'eau doivent être entretenus de façon optimale. Un amendement calcique des sols acides peut être envisagé afin de rééquilibrer le pH et ainsi limiter la persistance de *Map* dans le sol.
- **Alimentation** : l'équilibre alimentaire en énergie, protéines, minéraux et oligo-éléments de l'ensemble des catégories d'animaux est conseillé.
- **Maîtrise du parasitisme** : le parasitisme en tant que facteur éventuellement favorisant l'expression clinique chez les animaux infectés par *Map* doit être maîtrisé.

La **sortie du plan** se fait généralement lorsqu'il n'y a plus aucun cas clinique depuis trois ans dans l'élevage, qu'aucun animal positif n'a été réformé depuis au moins deux ans et après obtention de deux séries de contrôles négatifs sur les bovins de plus de 24

mois.

Malgré ces mesures drastiques, la réussite de tels plans de maîtrise est rarement atteinte. En effet, la longue période d'incubation, associée à la mauvaise Se_D de la sérologie pour le dépistage des phases précoces de l'infection, rend l'identification des animaux infectés hasardeuse. Ce type de plan de maîtrise a démontré une efficacité limitée dans les élevages bovins infectés et engendre des abandons fréquents, en raison de son inefficacité à endiguer les nouvelles infections, associée à des charges économiques et de travail trop grande pour les éleveurs (malgré la prise en charge fréquente d'une partie des coûts par les associations d'éleveurs).

Dans les élevages de petits ruminants, la taille des troupeaux empêche, pour des raisons financières et logistiques, le déploiement de telles mesures. L'efficacité de ces plans de maîtrise ne peut donc être jugée objectivement pour la filière ovine mais nous pouvons supposer que les résultats seront similaires à ceux observés pour la filière bovine.

5.2.3 Vaccination

La vaccination contre *Map* a été utilisée en 1926 pour la première fois et plusieurs vaccins ont depuis vu le jour. En France, le seul vaccin à avoir obtenu une autorisation de mise sur le marché (AMM) était un vaccin vivant (Neoparasec®). Il n'est plus commercialisé depuis 2001 notamment pour éviter une réaction croisée possible lors du diagnostic de la tuberculose par intradermo-tuberculation. Actuellement, seuls trois vaccins inactivés sont commercialisés dans le monde (tableau 3). Seul le vaccin Gudair® commercialisé en Espagne est actuellement disponible pour les ovins. Depuis 2009, l'Agence Nationale du Médicament Vétérinaire (ANMV) autorise l'importation de ce vaccin, à condition que les vétérinaires sanitaires en fassent la demande écrite et que celle-ci soit autorisée par l'ANMV. Pour obtenir l'autorisation, il est nécessaire de prouver l'existence de la paratuberculose par des analyses de laboratoire dans les cheptels souhaitant utiliser le vaccin. L'utilisation de ce vaccin provenant d'un état membre de l'UE respecte la cascade de prescription hors AMM.

Tableau 3 – Caractéristiques des vaccins contre la paratuberculose

Nom déposé	Mycopar®	Gudair®	Silirum®
laboratoire	Solvay Animal Health	CZ veterinaria SA	CZ veterinaria SA
Nature du vaccin	Inactivé	Inactivé	Inactivé
Souche vaccinale	<i>M. avium avium St 18</i>	<i>M. avium paratuberculosis 316F</i>	<i>M. avium paratuberculosis 316F</i>
Adjuvant	Huile minérale	Huile minérale	Huile minérale
Espèce	Bovins	Ovins, Caprins	Bovins
Année de début d'utilisation	1972	1994	2014

Résultats de la vaccination. Parmi les références publiées, l'efficacité des vaccins contre la paratuberculose a été démontrée à plusieurs reprises dans les trois espèces cibles (bovins, ovins et caprins).

Bastida and Juste (2011) ont répertorié dans une méta-analyse, toutes les études relatives à l'efficacité de la vaccination pour le contrôle de la paratuberculose. Au total, 118 études, basées sur l'utilisation des différents vaccins dont Néoparasec® et Gudair®, ont été analysées (dont cinq réalisées en France). Les expériences ont été classées selon l'espèce (bovine, ovine ou caprine) et les modalités d'évaluation de l'efficacité du vaccin (effets sur la clinique, sur l'excrétion ou sur les lésions macroscopiques et/ou microscopiques). Les études rapportées sont soit des comparaisons groupe vacciné/groupe non vacciné (témoin), soit des comparaisons temporelles avant/après vaccination. La majorité des études montrent des résultats favorables de la vaccination, quel que soit l'effet étudié (figure 14).

Quarante-neuf études ont été réalisées dans l'espèce ovine ; elles sont décrites dans 21 publications. Parmi elles, 14 études ont montré une réduction du nombre de cas cliniques ou du taux de mortalité, 15 études ont montré une réduction de l'isolement de *Map* dans les fèces ou les tissus, et enfin 17 études ont montré une réduction des lésions (observées après abattage).

Les résultats (en moyenne) obtenus pour les trois espèces de ruminants et pour les trois catégories d'effets sont proposés au tableau 4. Chaque résultat est un pourcentage de réduction calculé par le ratio : (fréquence lot témoin- fréquence lot vacciné) / fréquence lot témoin. Ce calcul est pondéré par la taille des effectifs impliqués dans les différentes

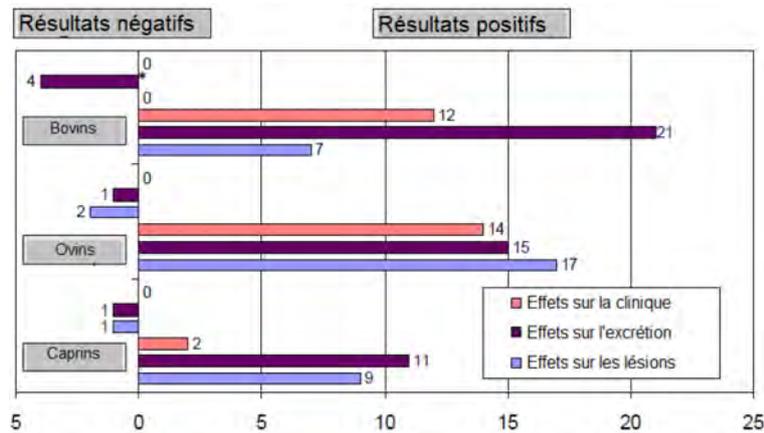


FIGURE 14 – Représentation graphique des évaluations réalisées sur l'efficacité de la vaccination, regroupées par résultat en fonction de l'espèce animale et des effets étudiés (production, épidémiologie ou pathogénicité) (d'après Bastida and Juste (2011)).

* Une expérience n'a montré ni résultat négatif ni résultat positif.

Les expériences avec des résultats négatifs sont représentées par les barres horizontales vers la gauche du graphique. Les expériences avec des résultats positifs sont représentées par les barres vers la droite du graphique. Les nombres adjacents aux barres correspondent aux nombres d'expériences publiées.

études.

Tableau 4 – Réduction de l'expression clinique, de l'excrétion ou des aspects lésionnels obtenus après vaccination contre la paratuberculose (d'après Bastida and Juste (2011)).

Type d'effets	Réduction du nombre de cas cliniques	Réduction du nombre d'excréteurs ou de la charge bactérienne excrétée	Réduction de l'intensité des lésions
Bovins	96,0% [35 - 100]*	72,6% [-36 - 99]	57,5 [17 - 58]
Ovins	67,5% [14 - 91]	76,4% [-10 - 93]	89,7% [-3 - 100]
Caprins	45,1% [34 - 82]	79,3% [-57 - 95]	94,8% [-17 - 97]

* Pourcentage moyen [minimum, maximum].

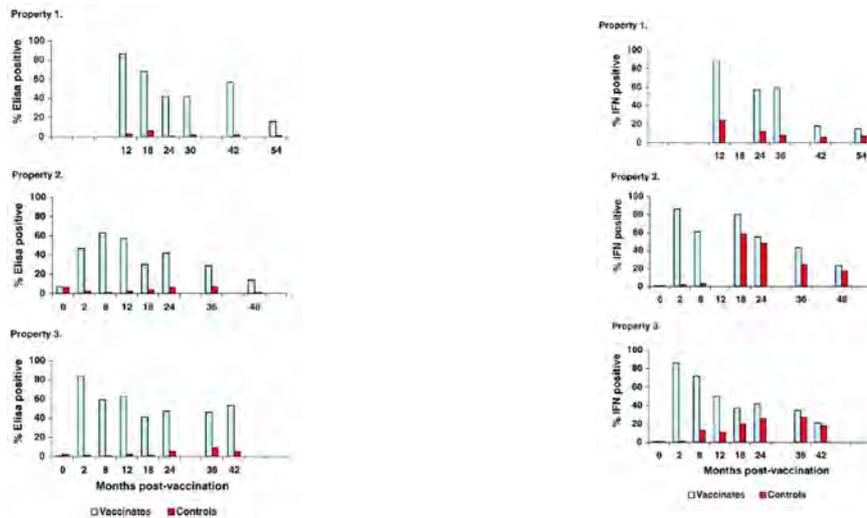
Efficacité du vaccin Gudair®. Le vaccin Gudair® est le vaccin destiné aux ovins et aux caprins; Il contient la souche inactivée 316F de *Map* et est combiné avec une nouvelle génération d'adjuvant. Il permet une immunisation active des ovins et des caprins afin de prévenir et réduire les signes cliniques, les lésions et la mortalité dus à la paratuberculose. Il réduit également l'excrétion fécale de *Map* chez les animaux vaccinés. Les animaux peuvent être vaccinés à partir de deux à trois semaines d'âge, à l'aide d'une seule dose de 1 mL injectée par voie sous cutanée.

La méta-analyse réalisée par Bastida and Juste (2011) ne répertorie que sept publications qui décrivent les effets du vaccin Gudair[®] chez les ovins. Le tableau 4 recense ces publications en mentionnant le nom du premier auteur, le pays et l'année de chaque publication, ainsi que les effets du vaccin qui ont été étudiés.

Les études d'évaluation de l'efficacité du vaccin Gudair[®] chez les ovins ont fait l'objet d'un nombre limité d'études de longue durée en élevages. Les seules études basées sur des grands effectifs, réalisées en élevages commerciaux, sont le fait de l'équipe de Windsor et Whittington de la Faculté de Sciences Vétérinaires de l'Université de Sydney en Australie.

Dès son autorisation de mise sur le marché en avril 2002 en Australie, le vaccin Gudair[®] a fait l'objet d'une étude visant à évaluer son efficacité vis-à-vis de la réduction de l'excrétion fécale dans des troupeaux commerciaux (Dhand et al., 2013). Cette étude portait sur 37 troupeaux fermés de plus de 500 brebis Mérinos où une vaccination des agneaux avant l'âge de six mois avait été réalisée pendant au moins cinq ans. La prévalence d'infection avant la vaccination était estimée en se basant sur des résultats de sérologie (méthode AGID) et/ou la culture de mélanges de fèces. La prévalence d'excrétion fécale après vaccination a été évaluée par culture de sept mélanges de 50 échantillons de fèces (soit environ une crotte par animal) par élevage, par la méthode radiométrique BACTEC 12B, suivie d'une confirmation des mélanges positifs par PCR *IS900*. Les estimateurs de prévalence d'excrétion pré et post-vaccinal ont été obtenus par une approche bayésienne tenant compte de la sensibilité et de la spécificité imparfaite des tests biologiques employés. La prévalence médiane pré-vaccinale a été estimée à 2,72% (PCI 95% : [1,40 - 6,86]). Après au moins cinq ans de vaccination, une diminution significative de la prévalence a été mise en évidence (médiane de 0,72%; PCI 95% : [0,39 - 1,27]). Cependant une persistance de l'excrétion fécale était retrouvée dans 30 des 37 élevages étudiés. Cette étude a par ailleurs permis d'identifier certains facteurs à l'origine des variations de l'efficacité vaccinale. Les troupeaux vaccinés ayant accès à des parcelles fréquentées par des animaux non vaccinés avaient une prévalence plus élevée. De même, l'introduction d'animaux non vaccinés réduisait également l'efficacité vaccinale au sein des troupeaux vaccinés.

Plus récemment, les mêmes auteurs ont rapporté les résultats d'une étude qui a duré dix ans dans 12 troupeaux à des niveaux de prévalence variables et ayant utilisés la vac-



(a) Proportion des animaux positifs au test IFN γ après vaccination (b) Proportion des animaux positifs au test ELISA après vaccination

FIGURE 15 – Suivi sérologique et IFN γ d'ovins vaccinés et non vaccinés dans trois fermes

cination des agneaux depuis plus de 10 ans (Dhand et al., 2016). L'excrétion fécale a été mesurée tous les deux ans à l'aide de la culture (milieu BACTEC 12B puis M7H9C) de 7 à 14 mélanges de fèces de 25 ou 50 animaux par classes d'âge (3, 4, 5 et 6 ans). Les agneaux étaient vaccinés entre l'âge de un à quatre mois. Les résultats obtenus confirment la réduction de la prévalence individuelle moyenne d'excrétion fécale (de 7,64% à 0,12%) rapportée dans l'étude précédente. De même, des mélanges positifs ont été retrouvés après 10 ans de vaccination dans trois des huit élevages encore présents à la fin de l'étude, correspondant à des prévalences individuelles d'excrétion de 0,25 à 0,47%.

La réponse immunitaire suite à la vaccination a été mesurée dans une étude de Reddacliff et al. (2006). Les figures 15a and 15b montrent les proportions de brebis chez lesquelles la production d'IFN γ a été détectée et avec une réponse mesurée par sérologie ELISA dans les groupes témoins et les groupes vaccinés.

La stimulation par la vaccination des deux types de la réponse immunitaire est mise en évidence par la proportion élevée d'animaux positifs pour les deux tests chez les vaccinés par rapport aux témoins non-vaccinés et par l'augmentation rapide du nombre de réponses positives chez les vaccinés deux mois après l'injection vaccinale (dans les fermes 2 et 3). Parmi les animaux témoins, des réponses positives à l'IFN γ ont été détectées huit à 12 mois après le début de la vaccination; elles sont certainement dues à l'exposition à *Map* dans l'environnement. La proportion de réponses positives chez les ovins vaccinés

est maximale au premier test après la vaccination. Les proportions d'animaux positifs restent plus élevées chez les vaccinés que chez les non-vaccinés quel que soit le moment du prélèvement. Au 18^{ème} mois après vaccination dans la ferme deux et au 24^{ème} mois dans la ferme 3, le pourcentage d'animaux vaccinés qui ont une réponse positive à la mesure de l'IFN γ augmente, ce qui peut indiquer que les animaux vaccinés ont été exposés à une pression infectieuse plus grande et que leur système immunitaire y répond.

Les réponses sérologiques sont similaires dans les fermes 1 et 3, avec des niveaux de séroconversion élevés chez les vaccinés qui diminuent ensuite lentement jusqu'à environ 18 mois après l'injection et les anticorps diminuent rapidement par la suite. Dans la ferme 2, les anticorps contre *Map* sont présents chez 6% des agneaux au moment de la vaccination, et peu d'animaux ont séroconverti au premier test après vaccination par rapport à ceux des autres fermes. Le pourcentage le plus élevé d'animaux positifs dans la ferme 2 s'élève à 63% au huitième mois alors qu'il atteint 80% dans les deux autres fermes. Au 24^{ème} mois post-vaccination, le pourcentage d'animaux positifs augmente à nouveau dans la ferme 2 ce qui indique une possible restimulation de la réponse immunitaire face à des niveaux d'infection élevés. Parmi les animaux non vaccinés, une séroconversion est observée vers 12 mois dans toutes les fermes, et la proportion de résultats positifs en sérologie ELISA est la plus élevée entre le 18^{ème} et le 36^{ème} mois post-vaccination. Ces séroconversions reflètent probablement l'évolution progressive vers la maladie.

Limites de la vaccination et effets indésirables. La vaccination n'empêche pas l'infection, tout comme elle ne prévient pas totalement l'apparition de la maladie. Dans plusieurs études portant sur l'efficacité vaccinale, des animaux vaccinés fortement excréteurs ont été détectés (Muñoz et al., 2005 ; Reddacliff et al., 2006). Ainsi, les animaux vaccinés doivent toujours être considérés comme potentiellement contaminants, même si le risque qu'ils représentent est significativement réduit par rapport à des animaux de statut inconnu (AFSSA, 2009).

Lors du dépistage des animaux infectés, les tests sérologiques ne permettent pas de distinguer les animaux vaccinés des animaux infectés. En effet, la vaccination engendre une forte réponse humorale pouvant être confondue avec celle apparue chez un animal

infecté. Dans la littérature, peu d'informations étaient disponibles chez les ovins jusqu'à récemment. Les résultats de l'étude d'Hilbink and West (1990) indiquent que la vaccination avec un vaccin vivant atténué (Neoparasec[®]) est susceptible de provoquer une réponse sérologique positive durable au test de fixation du complément et au test de diffusion en gélose. La réponse positive au test de fixation du complément est forte quel que soit le délai entre la vaccination et le test (jusqu'à 31 mois), alors qu'elle est plutôt hétérogène pour le test de diffusion en gélose (mais pouvant donner lieu à plus de 50% de résultats positifs 16 à 31 mois après vaccination). Ainsi la vaccination peut interférer avec les tests sérologiques de diagnostic de la paratuberculose.

Les animaux vaccinés peuvent réagir positivement à l'IDT à la tuberculine bovine. Toutefois, la réaction à la tuberculine aviaire est plus intense que celle à la tuberculine bovine chez les animaux infectés par *Map*. Ainsi, le recours à l'IDT comparative peut permettre de différencier un animal vacciné contre la paratuberculose et d'un animal tuberculeux. Cette problématique touche les bovins et les caprins, mais beaucoup moins les ovins, qui sont plus résistants à l'infection par *M. bovis*.

Enfin, il faut noter qu'il n'est pas rare d'observer une inflammation locale au site d'injection du vaccin, en particulier à cause des adjuvants huileux. Cette inflammation évolue en nodule sous-cutané fibreux qui persiste pendant plusieurs mois. Chez les petits ruminants, ce nodule peut parfois s'abcéder. Eppleston and Windsor (2007) ont réalisé une étude dont l'objectif était de déterminer la prévalence des lésions au point d'injection du vaccin Gudair[®] et l'impact économique de ces lésions à l'abattoir. La prévalence des lésions observées était de 18% pour les adultes et 65% pour les agneaux. Le parage réalisé pour retirer les lésions était négligeable, le surcoût du travail de retrait était nul et aucune carcasse n'a été déclassée. Les effets collatéraux sur la valeur marchande des carcasses du fait de la vaccination sont donc nuls à négligeables.

5.3 Sélection génétique

Le rôle du fond génétique de l'hôte dans la résistance ou la résilience à l'infection par *Map* vient initialement de l'observation de l'incidence variable des cas cliniques ou des résultats positifs en sérologie ELISA entre les différentes races ovines et bovines dans les

élevages. Ces différences de réceptivité et de sensibilité n'ont cependant fait l'objet que de très peu de travaux de confirmation par inoculation expérimentale. Dans une étude récente, (Begg et al., 2017) des différences significatives de l'incidence des cas cliniques et des pourcentages d'infection ont été mises en évidence entre des animaux de races Merinos, F1 Suffolk-Merinos, Border Leicester et Poll Dorset après inoculation expérimentale par voie orale (trois doses de $2,74 * 10^9$ *Map* à un mois d'intervalle, souche S Telford 9.2), suggérant des différences de prédisposition liées à des facteurs génétiques.

L'estimation de l'héritabilité des phénotypes observés et des coefficients de variation génétique associés permet d'évaluer l'opportunité de mettre en place des schémas pour améliorer la résistance/résilience par la voie de la sélection génétique. Les estimateurs de l'héritabilité varient de 0,031 à 0,283 chez les bovins (Kirkpatrick and Shook, 2011), suggérant, pour les plus élevés, la possibilité d'utiliser la sélection génétique pour une plus grande résistance à l'infection par *Map*. Chez les ovins, les résultats sont en nombres très limités, avec des héritabilités rapportées de $0,07 \pm 0,14$ pour la race Romneys et $0,18 \pm 0,11$ pour la race Merinos (Hickey et al., 2003).

Par ailleurs la recherche de gènes candidats à la résistance/résilience pour une maladie donnée a été permise par les mesures d'association entre un phénotype observé et des marqueurs génomiques dans l'espèce concernée. Ces études sont basées sur des approches génétiques (microsatellites, single-nucleotide polymorphism (SNP)) et analyses statistiques diverses (Analyses d'association à l'échelle du génome (GWAS) sur un marqueur SNP à la fois, analyses bayésiennes multi-marqueurs (Zare et al., 2014), ou plus récemment GSEA-SNP (Kiser et al., 2017)). Plus de 50 marqueurs ont été identifiés sur les chromosomes 1, 3, 5, 6, 7, 8, 9, 11, 12, 13, 16, 21, 23, et 25 chez les bovins (Minozzi et al., 2012), indiquant un phénomène polygénique et une part variable mais globalement faible de chaque gène dans la résistance/résilience à l'infection par *Map*.

Certains gènes candidats ont cependant été proposés à plusieurs reprises et dans des études indépendantes, comme étant associés à une réduction de la sensibilité à l'infection par *Map*. Ainsi, pour l'espèce ovine, les gènes codant pour les récepteurs TLR 1, 2 et 4; tous trois associés à la réponse immunitaire de l'hôte, le gène *SLC11A1*, influençant

sur la réponse précoce des macrophages à une infection bactérienne intracellulaire, ainsi qu'un ou plusieurs gènes du complexe majeur d'histocompatibilité, associé à la présentation des antigènes, ont été identifiés (Kirkpatrick and Shook, 2011 ; Purdie et al., 2011).

Les différences observées entre études peuvent avoir plusieurs sources, notamment des évaluations conduites dans des races ou des pays différents, des puissances statistiques variables selon les effectifs inclus ou des différences d'analyses statistiques. Cependant le facteur limitant principal est celui de la qualité des phénotypes recueillis pour les analyses d'association, et en définitive la définition des cas et des témoins. En l'absence de test diagnostique parfait pour le dépistage de l'infection par *Map*, la définition du phénotype repose sur les résultats combinés de un ou plusieurs tests (majoritairement sérologie ELISA, parfois associée à la culture fécale ou la culture tissulaire) plutôt que sur le statut infectieux avéré vis-a-vis de la paratuberculose. Des erreurs de classement sont donc très probables, avec principalement des animaux témoins considérés à tort comme non infectés (FN). Par ailleurs même si un ajustement ou un appariement sur l'âge est effectué, il est impossible de s'assurer que l'exposition à *Map* des animaux évalués est comparable en termes à la fois de précocité, de durée, de fréquence et d'intensité, et ce d'autant plus que les cas et les témoins sont parfois sélectionnés dans des élevages différents. Par ailleurs, dans la majorité des études, le statut des animaux est défini de manière transversale, à un temps donné, sans vérification que les animaux considérés comme non infectés à un moment donné le restent effectivement.

Sur les bases du modèle stochastique de diffusion intra-troupeau de la paratuberculose en élevage bovin laitier de (Marcé et al., 2011) et (Beaunée et al., 2015), les effets d'un schéma de sélection ciblant l'amélioration de différents caractères phénotypiques ont été quantifiés chez les bovins (Ben Romdhane et al., 2017). Les 14 caractéristiques phénotypiques évaluées étaient liées à la sensibilité à l'infection par *Map* en fonction de l'âge, la dose infectieuse minimale requise pour infecter un animal, la durée des différentes phases asymptomatiques et cliniques, et le niveau d'excrétion et les risques de transmission verticale associés. Les quatre caractères conduisant à l'amélioration la plus significative des critères épidémiologiques (prévalence d'animaux infectés et d'animaux malades,

incidence cumulée, persistance de la maladie dans l'élevage) étaient la diminution de la sensibilité avec l'âge, la réceptivité à une dose infectieuse donnée, la durée de la phase d'incubation avant le début d'une forte excrétion et la quantité de *Map* excrétée par les bovins forts excréteurs. Une sélection génétique permettant d'améliorer conjointement ces quatre caractères permettrait de réduire l'incidence cumulée au bout de 25 ans (moins de 25 nouvelles infections) assurant une maîtrise de l'infection au sein des élevages simulés. Les deux premières caractéristiques ne peuvent cependant pas faire l'objet, à l'heure actuelle, d'une sélection génétique, contrairement aux deux suivantes pour lesquelles des marqueurs génétiques ont été identifiés (Zare et al., 2014 ; Kirkpatrick et al., 2011). Dans ces études, les auteurs ont estimé l'héritabilité du caractère non-excréteur (sur la base de la culture fécale) entre 0.10 et 0.15. Il est donc théoriquement envisageable d'organiser une sélection génétique sur ce critère, même si les résultats rapportés par Ben Romdhane et al. (2017) montrent un bénéfice d'une sélection basée sur ce critère au bout d'une durée de 25 ans.

6 Conclusions du chapitre

Les éléments de bibliographie présentés mettent en évidence les limites des méthodes actuelles pour l'évaluation et le suivi du statut sanitaire vis-à-vis de la paratuberculose dans les élevages ovins. Par ailleurs cette revue des connaissances actuelles permet d'identifier un certain nombre de questions de recherche auxquelles il convient de répondre afin d'améliorer ces méthodes ou de les adapter à des contextes différents de ceux rencontrés dans la littérature.

La connaissance de la prévalence d'infection, à l'échelle des troupeaux où intra-troupeau, est nécessaire afin de rationaliser les mesures de maîtrise et de prévention. Du fait des faibles performances diagnostiques des tests ante-mortem individuels les estimations de ces prévalences sont peu précises et incertaines. Dans le contexte français, la paratuberculose n'étant pas une maladie réglementée, les études de prévalence sont basées sur le volontariat et sont à l'initiative de laboratoires de recherche (Mercier et al., 2010), des Groupements de Défense Sanitaire (GDS) ou des Organismes de Sélection (cas de l'OS Ro-

mane, ou Causse de Lot en France). Ces études sont peu nombreuses, notamment dans la filière ovine, et les protocoles associés sont très variables (type de test, nombre et âge des animaux prélevés, étude transversale ou suivi longitudinal) ce qui limite les comparaisons entre études. Le coût des évaluations basées sur le diagnostic individuel représente également un frein à la meilleure connaissance de la prévalence intra-troupeau du fait du faible nombre d'animaux généralement testés. Cette limitation pourrait être levée par l'utilisation d'analyse de mélanges d'échantillons, permettant une réduction substantielle des coûts. Dès lors le statut sanitaire du troupeau, et non plus le statut individuel des animaux, deviendrait la cible d'intérêt. Cette approche répond pleinement aux besoins de la filière ovine, dans laquelle la gestion de la maladie est réalisée à l'échelle collective et non individuelle.

L'utilisation des analyses en mélanges pour définir et suivre le statut sanitaire des troupeaux nécessite d'en connaître les performances diagnostiques (sensibilité et spécificité). Ces performances sont variables selon la cible diagnostique, c'est à dire le statut infectieux qui est recherché (infecté, infectieux, malade), la nature du test utilisé (culture fécale, PCR, sérologie...) et des techniques ou des trousseaux commerciales qui y sont associées. Les estimateurs de ces performances fournis par la littérature sont donc variables et difficilement comparables ou extrapolables. Une estimation théorique des performances des analyses de mélanges à l'échelle des troupeaux ovins français serait possible en se basant sur ces estimateurs, mais pourrait être fortement biaisée si les conditions ou les méthodes prévalant en France étaient éloignées de celles rapportées dans les études publiées. Par ailleurs, pour certaines techniques, ces estimateurs ne sont pas disponibles chez les ovins (qPCR sur mélanges de fèces par exemple) ou sont inexistantes (sérologie de mélange par exemple dans le cadre de la paratuberculose). Pour ces raisons, une estimation des performances diagnostiques des analyses de mélanges d'échantillons dans notre contexte d'étude nous semblait nécessaire.

Par ailleurs, la plupart études relatives aux analyses de mélanges ont eu pour objectif d'évaluer leur performances diagnostiques *relatives*, c'est à dire en la comparant aux ré-

sultats des analyses individuelles. A titre d'exemple Whittington et al. (2000a) ou Dhand et al. (2010b) ont estimé la probabilité qu'un mélange contenant un échantillon de fèces confirmé positif, dilué à un facteur 5 à 50 voire 100 dans des fèces négatifs et cultivé en milieu BACTEC 12B ou Middlebrook 7H10 modifié soit retrouvé positif selon que l'échantillon positif soit issu d'une brebis pauci- ou multibacillaire. Cette probabilité est donc relative à l'analyse à l'échelle individuelle et ne caractérise pas la capacité de l'analyse du mélange à détecter un *animal infecté* ou *infectieux*, mais seulement un *échantillon* de fèces contenant *Map*. Le lien entre ces deux informations est contenu dans la probabilité qu'un animal infecté ou infectieux puisse fournir un échantillon positif, ou, autrement dit, dans les performances des analyses individuelles pour le diagnostic des individus infectés ou infectieux. Les estimateurs relatifs à ces performances fournis par la littérature sont aussi très variables, pour les raisons exposées précédemment. De ce fait une évaluation, à l'échelle individuelle, des performances diagnostiques des tests utilisés dans notre travail, était nécessaire pour correspondre au mieux au contexte des élevages investigués.

Enfin, les effets de la vaccination au Gudair[®] sur la réponse sérologique et l'excrétion fécale de *Map* des ovins ont principalement été documentés dans le contexte d'élevage Australien. Les différences observées entre les protocoles vaccinaux australiens et français ne permettent pas d'extrapoler de manière univoque les résultats publiés par Reddacliff et al. (2006). De ce fait une étude spécifique au contexte de l'élevage ovin allaitant français apparaissait nécessaire et justifiée afin d'évaluer les conséquences de la mise en place d'une vaccination sur l'évaluation du statut sanitaire des troupeaux vaccinés par l'analyse de mélanges d'échantillons.

Les questions de recherche soulevées ci-dessous ont permis de définir les objectifs spécifiques de ce travail de thèse, qui sont présentés de manière détaillée dans le chapitre suivant.

Chapitre 2

Contexte et objectifs des travaux de thèse

- Voilà : la lune est-elle une banane et dans ce cas, où se trouve l'épluchure?
- Ça n'est pas une question scientifique...
- Ben qu'est-ce qu'il te faut?!!!

F'MURR

Le génie des alpages

1 Paratuberculose et race ovine Causse du Lot

La filière ovine de production de viande dans le département du Lot est organisée autour d'un schéma de sélection (OVILOT) porté notamment par un groupement d'éleveurs (CAPEL). Dans un objectif d'amélioration des performances de la race Causse du Lot, 72 éleveurs sont inscrits dans un schéma de sélection pyramidal à double étage (figure 16), reposant sur un petit nombre d'élevages sélectionneurs fournissant aux élevages utilisateurs et multiplicateurs des agnelles de race pure Causse du Lot. Ce flux d'animaux provenant des élevages sélectionneurs vers les élevages utilisateurs et multiplicateurs a favorisé la diffusion des maladies infectieuses dans toute la filière, et notamment celle de la paratuberculose. Un suivi annuel du statut des troupeaux sélectionneurs vis-à-vis de la paratuberculose a été mis en place, sur la base des résultats conjoints de la sérologie ELISA sur des groupes de 50 brebis adultes, tous âges confondus, et des résultats d'autopsie (avec confirmation par histologie, puis plus récemment par qPCR sur l'iléon) d'au

moins deux ovins avec un mauvais état corporel et d'éventuelles suspicions cliniques. Ce dispositif de surveillance a été amélioré en 2008, en partenariat avec l'UMR INRA-ENVT Interactions Hôtes - Agents Pathogènes (IHAP), par la définition de plans d'échantillonnage de sérologie ciblant une seule classe d'âge (brebis âgées de trois ans), permettant la comparaison des résultats entre années successives (Corbière, communication personnelle) et un suivi longitudinal plus efficace. Pour des raisons de coût et de logistique, ce suivi repose jusqu'à aujourd'hui, sur la sérologie ELISA appliquée à des lots de 50 brebis, les animaux prélevés servant aussi pour le dépistage sérologique réglementaire de la brucellose et de la Visna-Maedi.

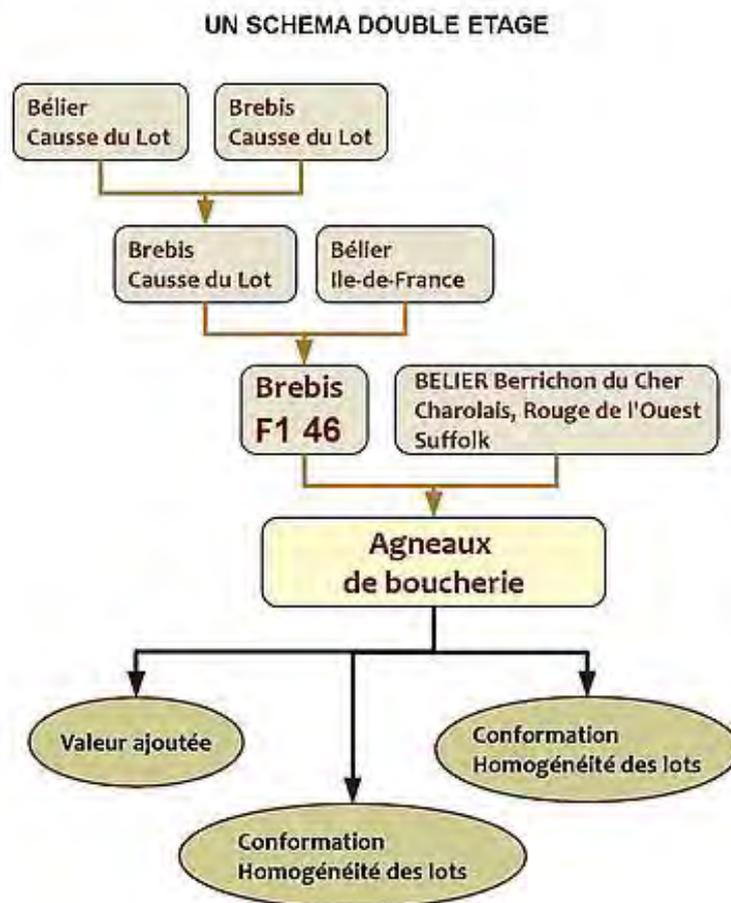


FIGURE 16 – Schéma de production et de croisement dit "en double étage" de la race Causse du Lot.

Le premier étage est en charge de la production d'agnelles et de brebis Causse du Lot qui sont vendues aux élevages de l'étage multiplicateur qui assure le croisement des brebis Causse du Lot avec des béliers Île-de-France. L'engraissement est ensuite assuré par le dernier étage du schéma.

Dans de nombreux élevages sélectionneurs, mais aussi utilisateurs de brebis F1, les

conséquences économiques de la paratuberculose (mortalité accrue, réformes précoces) ont été marquées. En 2009, malgré les contraintes, la commercialisation du vaccin Gudair® (sous autorisation temporaire d'utilisation), a permis de mettre en place un programme de vaccination basé sur le volontariat dans les élevages sélectionneurs les plus affectés, avec une prise en charge financière partielle par l'organisation de producteurs Ovilot. Une réduction de l'incidence des cas cliniques a rapidement été observée dans la plupart des élevages participants, mais le niveau résiduel d'infection est demeuré inconnu dans ces troupeaux.

Suite à la mise en place de la vaccination, les attentes de Capel/Ovilot ont évolué. Une meilleure détection des troupeaux infectés, sans augmentation excessive des coûts analytiques, était nécessaire afin de mieux sécuriser les mouvements d'animaux entre les élevages. Par ailleurs, dans le cadre du suivi sérologique annuel, de nombreux résultats positifs étaient observés dans les élevages vaccinés, soulevant la question de la pertinence de ce protocole de suivi et de la possibilité d'utiliser d'autres méthodes diagnostiques. La qPCR sur fèces apparaissait comme une bonne alternative, mais son coût rendait son utilisation à l'échelle individuelle prohibitive sur de nombreux animaux. Par ailleurs, en 2014, soit après quatre années de protocole vaccinal ininterrompu, la question de l'arrêt de la vaccination a été soulevée. La vaccination n'empêchant ni l'infection ni la présence d'animaux excréteurs au sein des troupeaux vaccinés, aucune réponse ne pouvait être apportée sans l'obtention d'estimateurs de la prévalence de l'excrétion fécale dans ces troupeaux. C'est de ce partenariat et de ces attentes qu'ont émergé les questions de recherche abordées dans ce mémoire. Ces travaux ont été rendus possibles par le soutien financier de l'INRA (Projet PICSAR du Métaprogramme GISA) et de la Région Occitanie (anciennement Midi-Pyrénées, projet PAROVLOT).

2 Objectifs de la thèse

L'objectif général de ce travail de doctorat était de réaliser une évaluation des performances diagnostiques de la sérologie ELISA et de la qPCR sur fèces appliquées à des mélanges d'échantillons pour l'évaluation et le suivi du statut sanitaire des troupeaux ovins

vis-à-vis de la paratuberculose. Le choix des tests biologiques s'est porté sur la sérologie ELISA et la qPCR sur fèces en raison de leur utilisation fréquente dans les élevages, de leurs coûts accessibles et de la rapidité d'obtention des résultats (notamment en comparaison de l'isolement en culture). Par ailleurs la plupart des laboratoires départementaux d'analyses (dont celui du département de Lot qui réalise les analyses pour Capel / Ovilot) propose ces deux tests diagnostiques, assurant ainsi le transfert des connaissances produites à l'issue de ce travail de thèse.

Plus spécifiquement, les objectifs suivants ont été définis :

1. **Évaluation des performances diagnostiques de la sérologie ELISA et de la qPCR sur fèces à l'échelle individuelle.** Compte tenu des estimateurs très variables recensés dans les publications, dépendants du contexte d'élevage, des cibles épidémiologiques définies et spécifiques des outils diagnostiques utilisés, il n'était pas envisageable d'extrapoler les résultats disponibles. Par ailleurs, la Se_D de la qPCR sur fèces n'avait jamais été évaluée jusqu'alors chez les ovins en phase subclinique et très partiellement chez les bovins. Il est apparu donc essentiel de réaliser cette évaluation dans le contexte des élevages ovins du Lot.
2. **Performances diagnostiques relatives de la sérologie ELISA et de la qPCR sur fèces appliquées à des mélanges d'échantillons.** La grande taille des troupeaux ovins et la faible valeur économique individuelle rend le diagnostic individuel exhaustif envisageable à grande échelle. Par ailleurs, une information individuelle n'est généralement pas utile en production ovine, où la gestion collective prime. L'analyse de mélanges d'échantillons permet d'augmenter le nombre d'individus testés, tout en conservant un coût de réalisation raisonnable. Le deuxième objectif de ce travail était donc d'explorer les conséquences de la réalisation de mélanges d'échantillons sur les performances diagnostiques de l'ELISA et de la qPCR.
3. **Sensibilité et spécificité à l'échelle des élevages de différentes stratégies de dépistage de la paratuberculose basées sur la sérologie ELISA et la qPCR sur fèces appliquées à des mélanges d'échantillons.** En s'appuyant sur les résultats des deux premières parties du travail de thèse, l'objectif était de proposer des plans de dépistage associés à des performances diagnostiques et économiques adaptées aux

élevages ovins allaitants du Lot.

4. Effets de la vaccination sur la réponse sérologique et l'excrétion fécale de *Map* en élevages ovins allaitants et conséquences pour le suivi des troupeaux vaccinés.

Ce dernier axe de travail avait pour objectif d'évaluer les prévalences et les niveaux d'excrétion fécale et de la réponse sérologique dans les troupeaux vaccinés et de les comparer aux indicateurs obtenus dans des troupeaux infectés non vaccinés. Par ailleurs l'obtention de ces résultats permettait d'évaluer les modalités de suivi du statut infectieux des troupeaux par l'utilisation de la sérologie ELISA et de la qPCR sur fèces appliquées à des mélanges d'échantillons dans un contexte de vaccination.

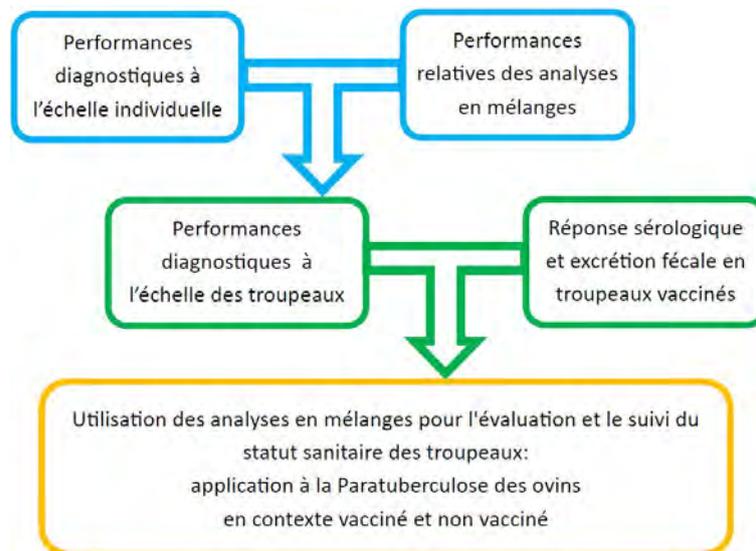


FIGURE 17 – Schéma des relations entre les objectifs de la thèse.

Chapitre 3

Évaluation de la sensibilité et de la spécificité de deux trousse de sérologie ELISA et d'une qPCR sur fèces pour le dépistage de la paratuberculose à l'échelle individuelle

L'intuition est l'art, spécifique à l'esprit humain, d'élaborer une réponse correcte à partir de données qui, par elles-mêmes, sont incomplètes, voire trompeuses.

Isaac Asimov

L'aube de Fondation

L'évaluation des performances des analyses de mélanges à l'échelle des troupeaux pour déterminer les prévalences d'infection nécessite une bonne connaissance des performances diagnostiques à l'échelle individuelle. Dans ce chapitre nous présentons notre travail portant sur l'estimation des performances diagnostiques de deux trousse de sérologie ELISA et d'une qPCR sur fèces pour le dépistage de la paratuberculose ovine à l'échelle individuelle.

L'analyse des courbes de fluorescence produites lors des réactions de qPCR a fait l'objet d'un travail spécifique. L'objectif était de s'affranchir des algorithmes commerciaux fournis avec les appareils d'amplification/détection qPCR. L'approche retenue, basée sur la modélisation des courbes de fluorescence par l'utilisation du package `qpcR` (Ritz and Spiess, 2008) pour le logiciel R (Team, 2017), a permis d'obtenir une modélisation indépendante de l'appareil utilisé. Le Ct produit en résultat de la qPCR était de ce fait plus homogène. Ces travaux ont fait l'objet d'une communication sous forme de poster présenté lors du 13^{ème} *Congrès International sur la Paratuberculose (ICP)* en 2016 et disponible à l'annexe II.

En l'absence de test parfait, des modèles à classes latentes ont été utilisés, selon deux approches statistiques, l'une bayésienne, l'autre fréquentiste. La première approche a fait l'objet d'un article publié en 2017 dans la revue *BMC Veterinary Research* et présenté ci-dessous. La deuxième approche est le support d'un article en cours de préparation.

1 Article 1 : estimation par une approche bayésienne à classes latentes

Résumé de l'article. L'objectif de cette étude était d'évaluer les performances diagnostiques de trois tests pour le dépistage de la paratuberculose chez les ovins infectés subcliniques, deux trousse ELISA appliquées aux échantillons de sérum (ID Screen Paratuberculosis Indirect[®], lot 602 et IDEXX paratuberculosis screening[®], lot 5074) et une trousse de qPCR appliquée aux fèces (ADIAVET[™] paraTB Real Time, BioX). Les prélèvements ont été réalisés dans le cadre d'une étude transversale sur 1197 ovins âgés de deux à trois ans, sans symptôme de paratuberculose, et issus de 14 troupeaux fermés de race Causse du Lot, où la paratuberculose était enzootique, membres à l'organisme de sélection OVILOT. Afin d'augmenter la puissance statistique, les élevages ont été regroupés en quatre classes de prévalence croissante et l'évaluation de la Se_D et de la Sp_D de ces trois tests a fait l'objet de plusieurs modèles bayésiens pour quatre populations et trois tests. Ces différents modèles ont permis d'évaluer différentes hypothèses de dépendance entre les tests. Par

ailleurs, ces modèles ne prenant en compte que les données binaires, les résultats douteux en sérologie ont été considérés comme positifs ou négatifs dans des analyses différentes. Enfin deux seuils de positivité ($Ct \leq 40$ ou $Ct \leq 42$) ont été évalués pour la qPCR sur fèces. Dans tous les cas, le modèle le plus parcimonieux associé au meilleur ajustement aux données, sélectionné sur le critère d'information bayésienne (DIC), était celui faisant l'hypothèse d'une dépendance entre les sensibilités diagnostiques des 2 trousses, et de dépendance deux à deux entre les spécificités diagnostique des trois tests.

Selon ce modèle, les sensibilités diagnostiques des 2 trousses ELISA étaient proches et évaluées à 17,4% (PCI 95% : [10,6 – 25,9]) pour la trousse IDEXX paratuberculosis screening[®], et 17,9% (PCI 95% : [11,4 – 25,6]) pour la trousse ID Screen Paratuberculosis Indirect[®], avec des spécificités estimées respectivement à 94,0% (PCI 95% : [92,2 – 95,7]) et 94,8% (PCI95% : [93,1 – 96,3]). La qPCR était significativement plus sensible (47,5%; PCI 95% : [29,3 – 69,9]) et plus spécifique (99,0%; PCI 95% : [97,9 – 99,9]) que les deux trousses ELISA. Les modèles ne tenant pas compte des corrélations entre la qPCR et les deux trousses ELISA ont conduit à des estimateurs de Se_D et de Sp_D légèrement supérieurs. Enfin le modèle faisant l'hypothèse d'une indépendance totale entre les trois tests aboutissait au moins bon ajustement aux données et était associé à des estimateurs de performances diagnostiques peu crédibles (sensibilités et spécificités diagnostiques des trousses ELISA très supérieures à celles de la qPCR). Ces résultats plaident pour une utilisation de la qPCR pour le dépistage de la paratuberculose chez les ovins infectés subcliniques, malgré une Se_D médiocre. Par ailleurs ce premier article illustre les conséquences des hypothèses faites sur la structure de corrélation entre tests sur les estimateurs obtenus par les modèles à classes latentes.

RESEARCH ARTICLE

Open Access



Estimation of the sensitivity and specificity of two serum ELISAs and one fecal qPCR for diagnosis of paratuberculosis in sub-clinically infected young-adult French sheep using latent class Bayesian modeling

Yoann Mathevon¹, Gilles Foucras¹, Rémy Falguières² and Fabien Corbiere^{1*} 

Abstract

Background: The objective was to evaluate the diagnostic accuracy of two serum ELISAs and one quantitative PCR on feces for the diagnosis of paratuberculosis in sub-clinically infected young-adult sheep. A cross-sectional study was performed to collect 1197 individual blood and fecal samples from 2- to 3-year-old sub-clinically infected ewes in 14 closed meat sheep flocks in France. Fecal excretion was determined using qPCR based on IS900 sequence detection, and serology was performed on serum samples using two commercial ELISAs. Data were analyzed in a 3-test multiple-population Bayesian latent class model accounting for potential dependence between the three tests fitted in OpenBUGS. Separate analyses were performed according to whether doubtful ELISA results were handled as positive or negative and based on two thresholds for fecal qPCR ($Ct \leq 42$ or $Ct \leq 40$).

Results: The best fit to the data was provided by accounting for a pairwise dependence between the two ELISAs on sensitivity and pairwise dependence between the three tests on specificity. Under this model, the estimated ELISA sensitivities were 17.4% (95% PCI: 10.6 – 25.9) and 17.9% (95% PCI: 11.4 – 25.6), with estimated specificities of 94.8% (95% PCI: 93.1 – 96.3) and 94.0% (95% PCI: 92.2 – 95.7). Fecal qPCR demonstrated significantly higher sensitivity (47.5%; 95% PCI: 29.3 – 69.9) and specificity (99.0%; 95% PCI: 97.9 – 99.9) than the ELISAs. Assumptions regarding doubtful ELISA results and qPCR thresholds had only a slight impact on test accuracy estimates. Models not accounting for pairwise dependence between ELISA and fecal qPCR results yielded higher sensitivity and specificity estimates but always provided a worse fit to the data.

Conclusions: Although the overall sensitivity of serum ELISAs and fecal qPCR remains low, the higher diagnostic performances of fecal qPCR make it more suitable for paratuberculosis diagnosis in sub-clinically infected sheep. Our results also illustrate that all dependence structures should be investigated when evaluating diagnostic test accuracy and selection based on a rigorous statistical approach.

Keywords: Paratuberculosis, Sheep, Elisa, Fecal quantitative PCR, Sensitivity, Specificity, Bayesian latent class model

* Correspondence: f.corbiere@envt.fr

¹IHAP, Ecole Nationale Vétérinaire de Toulouse, 23 chemin des Capelles, F-31076 Toulouse Cedex 3, France

Full list of author information is available at the end of the article



© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Background

Surveillance and control of paratuberculosis are largely hampered by the lack of sensitivity of available diagnostic tests, especially for the detection of sub-clinically infected (i.e., clinically healthy) animals. Historically, the evaluation of diagnostic test accuracy for the diagnosis of paratuberculosis has been based on cases confirmed by histopathological examination, fecal or tissue culture or repeated fecal culture for the detection of *Mycobacterium avium* subsp. *paratuberculosis* (MAP), the causative agent of paratuberculosis. However, due to the long and complex physiopathology of the disease, these cases do not include all latent cases of infection, generally leading to biased estimates of sensibility of diagnostic tests [1, 2].

In the last few decades, however, special attention has been given to the evaluation of diagnostic test accuracy in sub-clinically infected animals. Because of the unknown true disease status of the study subjects, due to the absence of a perfect reference test, latent class models have been increasingly used. These non-gold standard methods were first introduced by Hui and Walter (1980) [3] for 2 conditionally independent tests and two populations and were further extended to take account of conditional dependence between tests [2–6]. Bayesian modeling has been extensively developed to tackle non-identifiability issues that might arise in such models by incorporating prior knowledge of test performances [7, 8].

When erroneously assumed, the assumption of conditional independence between tests can seriously bias parameter estimations [9, 10]. Conditional dependence has been taken into account in most cases when evaluating two or more tests based on the same biological process (i.e., two fecal culture methods or two serological tests) for the diagnosis of paratuberculosis in cattle or in small ruminants [11–14]. Conversely, the a priori assumption of conditional independence between tests based on the identification of MAP (i.e., fecal culture, Ziehl-Neelsen stained fecal smear or fecal PCR) and those targeting the immune response (i.e., serum ELISA or AGID) has often been made [12, 15] but has been explicitly evaluated in only a few studies [16, 17].

One other assumption underlying latent class models is that the accuracy of tests is constant across all populations, or in other words, that the various infection stages among different populations are homogeneously distributed [2]. This assumption may, however, be difficult to stick to in practice, especially when sampling without controlling for factors that influence test accuracy [18]. For paratuberculosis, the increasing test sensitivity with the course of infection at the individual level would advocate for an age-specific evaluation of test accuracy [19]. This may lead to wide confidence or credible intervals of accuracy estimates, especially when prevalence is low and sample size is limited, as shown by simulation studies [20] and experience in

field studies [16, 15]. Furthermore, a diagnostic test's sensitivity may also vary between species, age and possibly MAP strains [2], while its specificity may be influenced by the presence of environmental mycobacterial [21].

Fecal quantitative PCR (qPCR) has been widely developed in the last two decades as an alternative to fecal culture for the detection of animals. It is less time consuming, especially for the detection of S-strains (sheep strains) of MAP that grow slowly in vitro compared to C-strains (cattle strains) [22]. There is also growing evidence that fecal qPCR might be at least as sensitive as, or even more sensitive than, fecal culture [23, 24]. However, its analytical sensitivity depends on several factors, including sample quality, DNA extraction methods, DNA target and qPCR systems [25, 26]. Furthermore, from an epidemiological point of view, Bayesian latent class models have seldom been applied to evaluate the diagnostic accuracy of fecal PCR [12], and estimates for sheep are scarce [14].

In this context, it would be unwise to simply extrapolate already published estimates of diagnostic test accuracy to any situation without utmost caution. In this study, we used a latent class approach in a Bayesian framework to estimate the diagnostic accuracy of two serum ELISAs and one fecal qPCR for the diagnosis of paratuberculosis in sub-clinically infected young-adult meat sheep, focusing on a narrow age range. Special attention was paid to the possibility of conditional dependence between tests under evaluation.

Methods

Flock and animal selection

Fourteen meat flocks with a size ranging from 290 to 1400 adult ewes (median 610) were selected for the study. They all belonged to the same breeders' association located in the Lot administrative region of France. Inclusion criteria were (i) Causse du Lot purebred closed flocks with no introduction of replacement ewes for at least 4 years, (ii) history of positive serological results and/or of clinical cases of paratuberculosis, and (iii) no history of vaccination against paratuberculosis. Sampling was performed from March 2014 to March 2015, avoiding the month before and after lambing as well as the month after artificial insemination or mating. Although it has been shown that the sensitivity of serological testing may be enhanced in early and late lactation in cattle [15, 21], this sampling scheme was applied to fulfill breeders' requests to reduce animal stress. Only 2- to 3-year-old ewes were included, using their eartag as an indicator of their birth cohort. Individual ages at sampling were calculated based on birth date available from the French System National d'Information Génétique (SNIG) database. Ewes showing obvious clinical signs of paratuberculosis, if any, were excluded because the target population was sub-clinically infected animals. If no feces could be retrieved intra-rectum at the

time of sampling, animals were excluded and the next one fulfilling the inclusion criteria was substituted. Depending on flock size, the target sample size ranged between 60 and 150 ewes per flock.

Sample collection and handling

A handful of feces was sampled from the rectum of selected animals using single-use gloves without lubricant and was placed in an individually identified sterile plastic bag for transportation. In parallel, a five-milliliter blood sample was also collected from the jugular vein in vacuum tubes without anticoagulant (Vacutainer® System). Feces and blood samples were frozen at -20°C prior to analysis. Animal handling was performed in compliance with the European Commission Directive 2010/63/EU. All farmers gave written consent for their animals to be used in this study.

Laboratory testing

Serological tests

Two commercial ELISA tests were applied to serum samples using an overnight incubation protocol following the manufacturer's instructions: ELISA A (ID Screen Paratuberculosis Indirect®, batch 602, IDVet, Montpellier, France) and ELISA B (IDEXX paratuberculosis screening® kit, batch 5074, IDEXX, Montpellier, France). Negative and positive controls provided by the manufacturers were included on each ELISA plate, and manufacturer's guidelines were strictly followed for interpretation of sample to positive (S/P) ratio results: for ELISA A serum, samples with S/P values $<60\%$, between 60 and 70%, and $\geq 70\%$ were considered negative, doubtful, and positive for MAP antibodies, respectively. For ELISA B, the negative and positive thresholds were 45% and 55%, respectively.

Fecal real-time PCR

First, fecal samples underwent a concentration procedure using the ADIAFILTER system (BioX, Rochefort, Belgium) following the manufacturer's instructions. Ten grams of feces were rehydrated overnight in 70 mL of bidistilled sterile water. The top 10 mL of the supernatant were then filtered and centrifuged using the ADIAFILTER® disposal. Pellets were then resuspended in 500 μL of bi-distilled water and mixed with 300 mg of 150-250 μm silica beads (Silibeads, Sigmund Lindner, Warmensteinach, Germany) for 30 s at 6800 rpm three times in a bead beater (Precellys 24®, Bertin Technologies, Montigny-le-Bretonneux, France). A magnetic bead-based DNA extraction was performed on a Kingfisher Flex® magnetic particle processor (Thermo Fisher Scientific, Courtaboeuf, France) following the NucleoMag 96 tissue protocol (Macherey-Nagel, Hoerdts, France), with addition of an extraction control (ADIAVET™ PARATB REAL TIME, BioX, Rochefort, Belgium) in each plate well. Samples were subjected to qPCR (ADIAVET™ PARATB REAL TIME, BioX,

Rochefort, Belgium), following the manufacturer's instructions. Each sample was also tested for amplification of the internal control. Bi-distilled water and synthetic IS900 DNA provided in the amplification kit were used as negative and positive controls, respectively. Forty-five amplification cycles were performed on a LightCycler 480 (Roche Life Science, Meylan, France), and fluorescent signals were recorded in two channels, with FAM detecting IS900 and VIC detecting the extraction control. Due to the overlapping spectra of the two dyes, a color compensation step was applied. Raw fluorescence data were obtained from the LightCycler 480 and modeled using the qpcR package [27] in R software [28]. Cycle thresholds were determined using second derivative maximum (CpD2). According to the manufacturer's recommendations, samples that reached fluorescence with a cycle count (Ct) below 40 were considered positive. A higher threshold ($\text{Ct} \leq 42$) was also considered. Indeed, careful examination of late fluorescence curves indicated that they were associated with low but unambiguously positive results up to 42 Ct, while non-specific amplification results could not be ruled out beyond this threshold.

All tests were performed blind for other test outcomes.

Target conditions

The purpose of this evaluation was to provide an accurate appraisal of sensitivity and specificity of two ELISAs and one fecal qPCR for the diagnosis of paratuberculosis in sub-clinically infected 2- to 3-year-old ewes. The target condition for this evaluation was MAP-infected animals that shed enough bacteria in their feces to potentially test positive on fecal PCR at the time of sampling, that mounted an antibody response towards MAP that could be detected by ELISA, or both. Following the Nielsen and Toft (2008) definition [29], this target condition included both infected and infectious animals but probably only few affected ones, as ewes showing obvious clinical signs of paratuberculosis were excluded on farms. Note that animals passively shedding MAP in their feces [30, 31] as a result of heavy environmental contamination were also included in our target conditions.

Statistical analysis

Separate analyses were performed for the four scenarios according to whether doubtful ELISA results were handled as positive or negative and on the choice of the positive cut-off for fecal qPCR ($\text{Ct} \leq 42$ or $\text{Ct} \leq 40$). Based on previous serological results, history of paratuberculosis clinical cases and judgment of practicing veterinarians and technicians supervising the flocks, flocks were grouped into 4 sub-populations according to the within-flock suspected prevalence of infection: very low (3 flocks, 287 sampled ewes), low (5 flocks, 299

sampled ewes), moderate to high (6 flocks, 447 sampled ewes) and very high (2 flocks, 164 samples ewes).

Model definition

We applied multiple populations Bayesian Latent Class models [32, 33] to estimate the diagnostic accuracy of the two ELISAs and the fecal qPCR in the absence of gold standard.

The models were defined following the approach by Dendikuri and Joseph (2001) [4] that uses a multinomial distribution to model the frequency of the 8 observed combinations of test outcomes. The simplest model assumes conditional independence between tests (i.e., given the true disease state of a sample, the outcome of one test does not have any influence on the probability of a positive or negative outcome in a second test). Under this assumption, the probability of a combination of test outcomes in a given population only depends on the true prevalence within this population and the sensitivities and specificities of diagnostic tests, which are assumed constant across all populations [3]. If T_i^+ denotes the event of a positive outcome for test i , $i = 1, \dots, 3$, Se_i and Sp_i denote the sensitivity and specificity of test i , respectively, and π_j , the true prevalence in a given population j , $j = 1 \dots 4$, then the probability of all three test being positive on a sample in this population is given by

$$P(T_1^+, T_2^+, T_3^+) = \pi_j Se_1 Se_2 Se_3 + (1 - \pi_j)(1 - Sp_1) \times (1 - Sp_2)(1 - Sp_3)$$

The probability of other combinations of test outcomes can be easily derived analogously. The assumption of conditional independence between tests may, however, not hold in practice and should be challenged against models allowing for the conditional dependence between tests [2]. We considered the approach proposed by Dendikuri and Joseph (2001) [4], where pairwise dependence of sensitivities and specificities of tests are explicitly modeled by covariance terms ($Covse$ and $Covsp$). In the fully dependent case, the probability of all three tests being positive on a sample within population j is then given by

$$P(T_1^+, T_2^+, T_3^+) = \pi_j (Se_1 Se_2 Se_3 + Covse_{23} Se_1 + Covse_{13} Se_2 + Covse_{12} Se_3 + Covse_{123}) + (1 - \pi_j) ((1 - Sp_1)(1 - Sp_2)(1 - Sp_3) + Covsp_{23}(1 - Sp_1) + Covsp_{13}(1 - Sp_2) + Covsp_{12}(1 - Sp_3) - Covsp_{123})$$

Starting from the fully saturated model below, covariance terms were removed one-by-one following a step-wise backward selection procedure using the Deviance Information Criterion (DIC) as the selection criterion

[34]. The DIC evaluates the model fit while penalizing the number of parameters, and it is generally accepted that models with smaller DIC are better supported by the data.

Comparing diagnostic test accuracies

The Bayesian posterior probability of difference (PPD) in sensitivity and specificity between tests was estimated using the Boolean step function in OpenBUGS [12, 16]. If $PPD < 0.05$ or > 0.95 , we concluded that the sensitivities (or specificities) of two compared tests were significantly different.

Serial and parallel testing

The accuracy of serial and parallel testing for the combinations of one ELISA and fecal qPCR was finally evaluated. For two conditionally dependent tests, namely, Test 1 and Test 2, the sensitivity (Se_{ser}) and specificity (Sp_{ser}) of serial testing are given by

$$Se_{ser} = Se_1 Se_2 + CovSe_{12}$$

$$Sp_{ser} = 1 - ((1 - Sp_1)(1 - Sp_2) + CovSp_{12}),$$

where $CovSe_{12}$ and $CovSp_{12}$ denote the covariance terms for the pairwise dependence of sensitivities and specificities, respectively.

Sensitivity (Se_{par}) and specificity (Sp_{par}) of parallel testing were given by

$$Se_{par} = 1 - ((1 - Se_1)(1 - Se_2) + CovSe_{12})$$

$$Sp_{par} = Sp_1 Sp_2 + CovSp_{12}$$

Prior distributions

Uniform distributions in the range from 0 to 1 were used as priors for sensitivity and prevalence model parameters. Based on previous published estimates in sheep [16, 35–37], the specificity of ELISAs and fecal qPCR was set at 0.95, with 95% certainty to be greater than 0.80. The corresponding Beta distribution Beta (21.20, 2.06) was generated using the epi.betabuster function embedded in the epiR package in R software [38] and was used as prior distribution for all specificity parameters.

Constraints were defined for covariance terms so that each of the 8 probabilities of combinations of test outcomes was between 0 and 1 [4], and uniform distributions between the lower and upper constraint bounds were used as non-informative priors.

Implementation

Computations were performed with OpenBUGS [39] embedded in R software using the R2OpenBUGS package [40]. Posterior estimates for test sensitivity and specificity

were generated using the Markov Chain Monte Carlo (MCMC) sampling method and the Gibbs algorithm. Three simulation chains of 200,000 iterations were run with different starting values, with the first 10,000 iterations discarded as the burn-in period. The chains were then thinned, taking every tenth sample to reduce autocorrelation among the samples. The convergence of the chains following the initial burn-in period was assessed visually by examining the traces, histories, Monte Carlo errors and the Gelman-Rubin diagnostic plots [41, 42]. The posterior distribution of each parameter was summarized using the mean and the 95% posterior credible interval (95% PCI). Analysis and graphing of the MCMC output were conducted using the coda package in R [43].

The aggregated data sets supporting the results of this article and the R2OpenBUGS code used are provided as additional files (Additional files 1 and 2).

Sensitivity analysis and model assumption checking

To assess the influence of prior information on the estimates of model parameters, poorly informative uniform distributions in the range of 0.5 to 1 were also considered for specificities. These truncated distributions were chosen to avoid convergence issues of single MCMC chains due to label switching [44].

To verify the assumption of constant test accuracy across all populations, we first excluded each of the 4 populations and subsequently each of the 14 flocks, one at a time, and re-ran all investigated models.

Results

Complete tests results were available for 1197 animals fulfilling the inclusion criteria, with a median sample size per flock of 89 (minimum 59, maximum 147). The median age at sampling was 2.5 years (lower quartile 2.3, upper quartile 2.7).

Test results

The cross-tabulated counts of the dichotomous outcome of the three tests are given in Table 1 for the 1197 sampled animals when assuming a fecal qPCR positive threshold of $Ct \leq 42$. The proportion of concordant test results was greater between the two ELISAs (1137/1197 = 95%) than between fecal qPCR and ELISA A (1047/1137 = 87%) or ELISA B (1051/1197 = 88%). Both ELISAs yielded fewer positive test results ($n = 85$ for ELISA A, $n = 93$ for ELISA B) than fecal qPCR ($n = 105$).

Doubtful results were few for both ELISAs tests and significantly fewer for ELISA A ($n = 8$) compared to ELISA B ($n = 23$, Fisher's Exact test $p = 0.0109$). Setting the positive cut-off at $Ct \leq 42$ for fecal qPCR, rather than $Ct \leq 40$, yielded 32 more positive samples.

Model selection

Doubtful ELISA results and moving the positive cut-off from 40 to 42 for fecal qPCR had no influence on model selection. Based on DIC, the best fitting model (model 1) was the one with a pairwise dependence between ELISA A and ELISA B on sensitivity and pairwise dependence between the three tests on specificity (Table 2). This model always outperformed the one assuming a conditional independence between fecal qPCR and both ELISAs on sensitivity and specificity (model 2). The difference in the DIC of model 1 and model 2 was always greater than 12.5, suggesting that including covariance terms between the fecal qPCR and both ELISAs provides a better fit to the data, although this was only significant for specificity. As expected, the assumption of conditionally independent ELISAs was not supported by the data, as shown by the high DIC values of model 3 (Table 2).

Estimated accuracy of diagnostic tests

The posterior distributions for sensitivity and specificity of the three tests and prevalence are summarized in

Table 1 Cross-classified positive (+) and negative (–) results of two serum ELISAs and one fecal PCR in sub-populations 1 to 4 for sub-clinically infected 2- to 3-year-old French Causse du Lot sheep

ELISA A	ELISA B	Fecal qPCR	Subpop. 1	Subpop. 2	Subpop. 3	Subpop. 4	Total
+	+	+	0	2	5(4)	8 (5)	15 (11)
+	+	–	9 (7)	11 (7)	14 (14)	10 (6)	44 (34)
+	–	+	0	1	2	2	5 (5)
+	–	–	6 (7)	5 (8)	9 (6)	1 (6)	21 (27)
–	+	+	1	3	6 (4)	1	11 (9)
–	+	–	9 (5)	6 (4)	7 (6)	1	23 (16)
–	–	+	0	4(3)	37 (17)	33 (28)	74 (48)
–	–	–	262 (267)	267 (271)	367 (394)	108 (115)	1004 (1047)
	Total		287	299	447	164	1197

Doubtful ELISA and fecal qPCR $Ct \leq 42$ results were treated as positive. When doubtful ELISA and fecal qPCR $40 \leq Ct \leq 42$ results were treated as negative, counts are shown in brackets

Table 2 Bayesian Deviance Information Criterion (DIC) for model 1 to 3 under different scenarios

Doubtful ELISA results	Fecal qPCR positive cut-off	Deviance Information Criterion		
		Model 1 ^a	Model 2 ^b	Model 3 ^c
Positive	Ct ≤ 42	120.4	134.2	227.9
Positive	Ct ≤ 40	114.3	130.1	197.5
Negative	Ct ≤ 42	115.6	128.2	222.1
Negative	Ct ≤ 40	114.3	127.5	194.5

^aModels with pairwise dependence between ELISA A and ELISA B for sensitivity and pairwise dependence between the three tests for specificity

^bModels with pairwise dependence between ELISA A and ELISA B for sensitivity and specificity and assuming conditional independence of fecal qPCR

^cModels assuming fully conditional independence between the three tests

Table 3 in form-of-point estimates (mean) and 95% Bayesian posterior density credible intervals (95% PCI). For comparison purposes, the results from model 2 and model 3 are also shown. The estimated sensitivity and specificity were similar for ELISA A and ELISA B (Se ≈ 17%, PPD = 0.121; Sp ≈ 95%, PPD = 0.401) (Table 3, model 1). The fecal qPCR was found to be more sensitive (47.5%) and specific (99.0%) than ELISA tests, with PPD > 0.999 and posterior 95% credible interval excluding zero. Under the assumption of complete independence between the fecal qPCR and both ELISA tests (model 2), higher estimated sensitivities were obtained, especially for

fecal qPCR (56.3%), without substantial changes for estimated specificities. The fully conditional independent model (model 3) yielded unrealistic significantly higher estimated sensitivity and specificity for ELISA A (Se = 70.0%, Sp = 98.7%) and ELISA B (Se = 80.0%, Sp = 98.9%) than for fecal qPCR (Se = 31.3%, Sp = 93.2%).

From model 1, ELISA A and ELISA B appeared positively correlated for sensitivity and specificity (Covse median of 0.108 and 95% PCI between 0.068 and 0.153; Covsp median 0.029 and 95% PCI between 0.018 and 0.033). No evidence of correlation was found between ELISAs and fecal qPCR for sensitivity. In model 1, covariance terms for specificity between the fecal qPCR and ELISA A (Covsp median 0.001 and 95% PCI between 0.0009 and 0.00529) and ELISA B (Covsp median 0.00472 and 95% PCI between 0.00029 and 0.01179) were very small, although significantly different from 0.

No substantial differences in estimated sensitivity and specificity were observed when analyzing the three other datasets (Table 4). Treating doubtful ELISA results as negative mostly induced a slightly lower estimated sensitivity of ELISA B (14.7%), which was expected from the larger number of doubtful results obtained with this test compared to ELISA A. Similarly, changing the positive cut-off for fecal qPCR from Ct ≤ 42 to Ct ≤ 40 yielded a slightly lower estimated sensitivity for fecal qPCR

Table 3 Mean and 95% posterior credible intervals (PCI) for the sensitivity (Se) and specificity (Sp) of two serum ELISAs and on fecal qPCR and true prevalence (Ps) of MAP in sub-populations 1 to 4

		Mean (95% PCI)		
		Model 1 ^a	Model 2 ^b	Model 3 ^c
DIC		120.4	134.2	227.9
ELISA A	Se	17.9 (11.4; 25.6)	19.7 (13.1; 27.3)	70.0 (56.0; 85.6)
	Sp	94.8 (93.1; 96.3)	95.0 (93.4; 96.5)	98.7 (97.4; 99.9)
ELISA B	Se	17.4 (10.6; 25.9)	23.2 (15.8; 31.5)	80.0 (63.7; 95.3)
	Sp	94.0 (92.2; 95.7)	94.8 (93.2; 96.4)	98.9 (97.5; 99.9)
Fecal qPCR	Se	47.5 (29.3; 69.9)	56.3 (36.9; 77.3)	31.3 (21.8; 41.9)
	Sp	99.0 (97.9; 99.7)	99.4 (98.4; 99.9)	93.2 (91.6; 94.7)
Sub-pop. 1	P1	0.01 (0.0; 0.03)	1.2 (0.1; 3.8)	6.6 (3.1; 11.0)
Sub-pop. 2	P2	5.4 (1.6; 11.5)	6.2 (2.5; 11.9)	8.1 (4.7; 12.3)
Sub-pop. 3	P3	21.7 (13.1; 34.5)	19.2 (12.1; 29.4)	8.4 (5.2; 12.3)
Sub-pop. 4	P4	57.3 (35.4; 88.8)	47.2 (31.0; 70.4)	15.5 (9.7; 22.6)
Δ _{ELISA A - ELISA B}	Se	0.5 (-5.6; 5.9)	-3.5 (-9.9; 2.5)	-10.0 (-28.8; 9.3)
	Sp	0.8 (-0.8; 2.4)	0.2 (-1.2; 1.5)	-0.2 (-1.9; 1.5)
Δ _{Fecal PCR - ELISA A}	Se	29.6 (13.7; 50.9)	36.7 (18.6; 57.9)	-38.8 (-56.4; -22.4)
	Sp	4.2 (2.5; 6.0)	4.4 (2.6; 6.2)	-5.5 (-7.4; -3.7)
Δ _{Fecal PCR - ELISA B}	Se	30.1 (14.5; 50.8)	33.1 (16.0; 53.4)	-48.8 (-66.4; -30.8)
	Sp	5.0 (3.4; 6.7)	4.6 (2.8; 6.4)	-5.7 (-7.7; -3.8)

Doubtful ELISA and fecal qPCR Ct ≤ 42 results were treated as positive

^aModels with pairwise dependence between ELISA A and ELISA B for sensitivity and pairwise dependence between the three tests for specificity

^bModels with pairwise dependence between ELISA A and ELISA B for sensitivity and specificity and assuming conditional independence of fecal qPCR

^cModels assuming fully conditional independence between the three tests

Table 4 Mean and 95% posterior credible intervals (PCI) for the sensitivity (Se) and specificity (Sp) of two serum ELISAs and one fecal qPCR, depending on different scenarios

Dataset		Mean (95% PCI)	
		Se	Sp
Doubtful ELISA results as negative, fecal qPCR cut-off Ct ≤ 42	ELISA A	17.3 (11.0; 24.8)	95.3 (93.7; 96.7)
	ELISA B	14.7 (8.5; 22.4)	95.6 (94.0; 97.0)
	Fecal qPCR	50.9 (31.3; 73.3)	99.2 (98.2; 99.8)
Doubtful ELISA results as positive, fecal qPCR cut-off Ct ≤ 40	ELISA A	21.0 (12.8; 30.8)	94.8 (93.1; 96.3)
	ELISA B	20.0 (11.9; 30.2)	94.0 (92.2; 95.6)
	Fecal qPCR	40.7 (23.5; 63.2)	99.0 (98.0; 99.7)
Doubtful ELISA results as negative, fecal qPCR cut-off Ct ≤ 40	ELISA A	21.2 (12.9; 31.0)	95.3 (93.7; 96.7)
	ELISA B	17.4 (9.5; 27.5)	95.5 (94.0; 96.9)
	Fecal qPCR	45.8 (26.2; 69.3)	99.2 (98.2; 99.8)

Estimates were obtained with model 1

(40.7%) and slightly higher estimated sensitivity for ELISA A (21.0%) and ELISA B (20.0%). In any case, the estimated specificity of the three tests remained mostly unchanged.

Serial and parallel testing

Serial and parallel testing were evaluated for model 1 (Table 5). For both ELISA and fecal qPCR combinations, serial testing was associated with a slight increase in specificity but a strong drop in sensitivity to below 9%. The use of ELISA and fecal qPCR in parallel testing led to an increased estimated sensitivity compared to fecal qPCR alone, though at the price of a loss of specificity.

Sensitivity analysis and model assumption checking

The use of poorly informative prior distributions for specificities and resampling subpopulations or flocks did not yield any substantial change of the parameter estimates. This suggests a very weak influence of prior distributions on estimation and that the assumption of constant sensitivities and specificities was not unreasonable. Furthermore, model selection based on DIC remained unchanged, strengthening our findings regarding the conditional dependence between test results.

Table 5 Mean and 95% posterior credible intervals (PCI) for the sensitivity (Se) and specificity (Sp) of serial and parallel testing using one serum ELISA and the fecal PCR

		Mean (95% PCI)	
		ELISA A – Fecal qPCR	ELISA B – Fecal qPCR
Serial testing	Se _{ser}	8.7 (3.8; 15.8)	8.5 (3.5; 16.0)
	Sp _{ser}	99.8 (99.3 – 100.0)	99.3 (98.5; 99.9)
Parallel testing	Se _{par}	56.6 (38.8; 76.1)	55.8 (37.4; 75.9)
	Sp _{par}	93.8 (91.8; 95.6)	93.1 (90.7; 95.1)

Estimates were obtained with model 1, with doubtful ELISA results and fecal qPCR Ct ≤ 42 treated as positive

Discussion

We used a Bayesian latent class approach to estimate the diagnostic accuracy of two serum ELISAs and one fecal qPCR for the detection of 2- to 3-year-old subclinically infected sheep. This evaluation follows the standards for the reporting of diagnosis accuracy for paratuberculosis [1] that were recently extended to Bayesian latent class models [2].

Latent class models are highly sensitive to assumptions made regarding the conditional dependence between tests [10]. We found that treating all three tests as conditionally independent (model 3) led to biased results, with strongly overestimated sensitivities for both ELISAs. This finding is supported by the high DIC value obtained for this model and was already emphasized by simulation studies [20]. In the same way, we found that the assumption of conditional independence between fecal qPCR and ELISAs (model 2) was not supported by the data. Although the conditional independence between fecal culture and ELISA may hold [16, 17], to our knowledge, there is no available study evaluating the conditional dependence between fecal qPCR and ELISA. Indeed, the a priori assumption of conditional independence is made in most cases but not formally tested [12, 15]. In our study, covariance terms between fecal qPCR and ELISAs were only significant in the specificity part of the model and were considerably less than the one found between the two ELISAs. However, based on DIC values, models that accounted for this dependence were unambiguously favored and led to estimates that were moderately lower than those obtained under the conditional independence assumption. These findings may or may not apply to evaluations of other commercial ELISAs and PCRs, depending on the antigens used and gene targets, respectively. In some instances, moreover, the dependence between tests may be of minimal importance, especially if the individual estimates (i.e., specificity) are close to one [2]. However, our results suggest that, when possible, models accounting for all dependence of sensitivities and specificities should be evaluated first and possibly simplified based on a rigorous selection process. Complete saturated models may, however, not be identifiable (i.e., with the number of parameters greater than the degrees of freedom permitted by the data), allowing only restrained covariance structures to be evaluated [44].

One other assumption underlying latent class models is that the various infection stages among the different populations are homogeneously distributed [2]. Our study was based on animals belonging to a narrow age range (2 to 3 years), which, to our point of view, offers several advantages. First, it might have lessened the selection biases related to a non-homogenous sampling across the different infection stages among populations, since an age representative sample of animals might be

difficult to achieve in practice. In large herds/flocks where only partial sampling is often applied due to cost constraints, focusing on specific age cohorts may also allow for an easier and more robust comparison of prevalence estimates between herds/flocks. Finally, at the herd/flock level, a narrow age range may facilitate year-over-year comparison of results. The drawback of such an approach is that our results may be strongly linked to our study population and should be carefully extrapolated to other situations.

As both ELISA and fecal qPCR provide a continuous range of result values, the classification of samples as positive or negative results in a loss of information [45] and in inconclusive test results (in our case, doubtful ELISA results and characteristic amplification curves with $Ct > 40$ for fecal qPCR according to the manufacturer's recommended positive threshold). As they may have a strong influence on accuracy estimates [46], inconclusive results were classified either as positive or negative in separate analyses following standards for reporting of diagnosis accuracy studies. However, because there were only a few, doubtful ELISA results did not cause any considerable differences in the summaries of test performances. In the same way, choosing a $Ct \leq 42$ rather than $Ct \leq 40$ threshold for the fecal qPCR did not lead to a dramatic change in sensitivity estimates. These changes were of the order of magnitude as those observed between statistical models 1 and 2.

Point estimates of both ELISA sensitivities obtained in our study (14 - 21%) are similar to or slightly lower than those obtained in other studies for the detection of subclinically infected sheep reviewed in Nielsen and Toft (2008) [29]. ELISA B was recently applied on serum and milk in Greek dairy sheep and yielded higher sensitivity estimates (46-49%) [47]. The reasons for these discrepancies are not known but could be related to the age structure of study samples, breed differences or possibly regional MAP strain variations. Conversely, our specificity estimates (94-96%) were in concordance with those found in already mentioned studies in sheep [16, 35-37] and support the idea that ELISA is far from being perfectly specific.

Fecal qPCR has the potential to be a rapid and sensitive method of MAP diagnosis, especially in sheep in which fecal cultures performed poorly. We found that fecal qPCR had higher diagnostic accuracy than ELISA, with sensitivity estimates close to those obtained by Baumann et al. [14] in sheep when using the $Ct \leq 40$ cut-off for positive results. Moving the cut-off up to $Ct \leq 42$ was associated with slightly enhanced sensitivity estimates with almost no change in specificity estimates. While the specificity of fecal qPCR was very high, it was not absolute at the $Ct \leq 42$ or $Ct \leq 40$ cut-off. An even more conservative value (i.e., $Ct \leq 38$) was also evaluated

without improvement of specificity estimates (results not shown). Although the specificity of the IS900 target for the detection of MAP is of concern, as other mycobacteria with IS900-like sequences have been described [48], considerable improvements have been made in PCR probe and primer designs in recent years [49, 50], and this hypothesis is currently unlikely. However, other targets exclusive to MAP, such as the *hspX* gene [51], have shown non-perfect specificity for the detection of infectious animals when evaluated in Bayesian latent class models [14]. Rather, this might reflect the potential of pass through of orally ingested organisms by uninfected animals [30, 31] or the small yet existent possibility of cross-contamination of samples during collection or laboratory processing. The multi-copy presence of the IS900 target in the MAP genome (14-18 copies) might conversely provide higher analytical sensitivity compared to some specific alternative targets (*f57*, ISMAP02, *hspX*) that are only present in six or fewer copies [52, 53]. Moreover, 10.0 g of feces were processed for the qPCR detection, lowering the possibility of missing MAP aggregates [54]. Nevertheless, as stated in our results, the epidemiological sensitivity of fecal qPCR, even based on the IS900 target, remains low in 2- to 3-year-old subclinically infected sheep (40-50%). This might reflect the low number of infected animals that shed MAP in their feces within this age cohort, or that intermittent shedding prevented their detection at the time of sampling, or both.

Our specificity estimates for ELISAs and qPCR are based on data collected in flocks suspected or known to be infected by MAP. Therefore, they may not reflect those that would have been obtained in truly paratuberculosis-free flocks, in which they could be expected to be higher [55]. However, the large-scale application of an imperfectly specific test (even with specificity as high as 99.5%) is questionable for detection purposes, as it would lead to numerous false positive results in paratuberculosis-free flocks that would require further investigation. Conversely, this lack of specificity may have fewer adverse impacts on infected flock monitoring programs, as the positive predictive value of tests will be higher, and no confirmatory testing will generally be requested [56].

Finally, the estimated sensitivity of fecal qPCR had wide credible intervals. In latent class model analysis, reasons responsible for such findings are low true values of diagnostic test accuracy, low true prevalence, small sample size, small difference in prevalence between sub-populations, lack of global identifiability of the model, or parameter estimates close to 0.5 [20, 44, 57]. Although a large number of sheep were sampled, the estimated true prevalence was rather low in two out of four sub-populations (0.8% and 5.4%, respectively), and therefore, the sensitivity estimates were based on a limited number of positive results. This is

also illustrated by the very narrow intervals for fecal qPCR sensitivity estimates provided by multiplying the original data by ten (11,970 animals) (results not shown).

The serial use of fecal qPCR for the confirmation of ELISA-positive individuals allows for an almost perfect specificity, especially for ELISA A (99.8%). Serial testing was, however, associated with a very low global sensitivity, meaning that the true infectious status of an ELISA-positive individual that would be subsequently tested as qPCR-negative in feces would remain uncertain. The interferon-gamma release assay provides a positive response earlier in the course of the disease than fecal culture [58] and would therefore be advised in such cases. However, this assay also requires careful interpretation, as it cannot distinguish between infected and exposed animals [59]. As shown in Table 5, the diagnostic accuracy at the individual level could be enhanced by the use of serum ELISA and fecal qPCR in parallel testing. This reflects the fact that fecal shedding of MAP and the humoral response are poorly correlated and that parallel testing might target different individuals. This is also stated by the non-significant covariance terms for sensitivity between fecal qPCR and serum ELISAs in our Bayesian latent class model and is in accordance with experimental infection results indicating that some persistently shedding sheep may develop clinical disease in the absence of an antibody response [60]. The use of tests in combination, however, substantially adds to the cost of control, which may or may not be acceptable to sheep owners. Moreover, the higher cost of individual fecal qPCR (approximately 35 euros or 39 USD) compared to serum ELISA (approximately 6 euros or 7 USD) limits its use at a large scale in France.

Conclusions

An accurate appraisal of diagnostic test accuracy is of critical importance for a better evaluation of paratuberculosis control programs. In this study, we showed that the assumption of conditional independence between fecal qPCR and serum ELISA was not supported by the data and that accounting for this dependence provided slightly different accuracy estimates. Fecal qPCR demonstrated a higher sensitivity and specificity than serum ELISA, but the overall sensitivity of both diagnostic approaches remains low in 2- to 3-year-old sub-clinically infected animals. These findings advocate for more frequent testing of animals in a longitudinal follow-up scenario. Studies are in progress to evaluate the consequence of these estimated diagnostic test accuracy for surveillance programs at the flock level.

Additional files

Additional file 1: R code for Bayesian Latent Class models. Bayesian Latent Class Models (model 1 to 3) that were used in this study. (DOCX 25 kb)

Additional file 2: Aggregated data set. Aggregated diagnostic test results for the 4 sub-populations, given whether doubtful ELISA results were handled as positive or negative and based on two thresholds for fecal qPCR. (DOCX 15 kb)

Abbreviations

CovSe: covariance term on sensitivity; CovSp: covariance term on specificity; CpD2: second derivative maximum; Ct: cycle threshold; DIC: deviance information criterion; MAP: *Mycobacterium avium* spp. paratuberculosis; MCMC: Markov Chain Monte Carlo; PCI: Bayesian posterior credible interval; PPD: Bayesian posterior probability of difference; qPCR: quantitative polymerase chain reaction; S/P: sample-to-positive ratio; Se: sensitivity; Se_{par}: sensitivity of parallel testing; Se_{ser}: sensitivity of serial testing; SNIG: Systeme National d'Information Génétique; Sp: specificity; Sp_{par}: specificity of parallel testing; Sp_{ser}: specificity of serial testing

Acknowledgements

The authors thank all farmers, farm technicians and veterinary practitioners and students for their involvement in and dedication to this study.

Funding

This study was funded by the INRA GISA Metaprogram PICSAR and Region Midi-Pyrénées (PAROVLOT program). The lead author received a PhD grant from both funding bodies. The funding bodies did not have any direct role in the study design or sample collection.

Availability of data and materials

Aggregated data are presented within the manuscript (Table 1) and in Additional file 2. Individual data analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

YM, FC and GF developed and designed the study and participated in the animal sampling. YM carried out the laboratory analysis. YM and FC carried out the data analysis. YM, FC and GF drafted the initial manuscript. RF provided additional assistance in coordination of the study and animal sampling. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Animal studies were compliant with all applicable provisions established by the European Commission Directive 2010/63/UE. All animals used in this study were handled in strict accordance with good clinical practices and all efforts were made to minimize suffering. All animal owners gave written consent for their animals to be used in this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹IHAP, Ecole Nationale Vétérinaire de Toulouse, 23 chemin des Capelles, F-31076 Toulouse Cedex 3, France. ²CAPEL Ovilot, 237 avenue Pierre Semard, 46000 Cahors, France.

Received: 14 April 2017 Accepted: 19 July 2017

Published online: 03 August 2017

References

- Gardner IA, Nielsen SS, Whittington RJ, Collins MT, Bakker D, Harris B, et al. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. *Prev Vet Med.* 2011;101:18–34.
- Kostoulas P, Nielsen SS, Branscum AJ, Johnson WO, Dendukuri N, Dhand NK, et al. STARD-BLCM: standards for the reporting of diagnostic accuracy studies that use Bayesian latent class models. *Prev Vet Med.* 2017;138:37–47.

3. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics*. 1980;36:167–71.
4. Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*. 2001;57:158–67.
5. Georgiadis MP, Johnson WO, Gardner IA, Singh R. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *J R Stat Soc Series B Stat Methodol*. 2003;52:63–76.
6. Berkvens D, Speybroeck N, Praet N, Adel A, Lesaffre E. Estimating disease prevalence in a Bayesian framework using probabilistic constraints. *Epidemiology*. 2006;17:145–53.
7. Johnson WO, Gastwirth JL, Pearson LM. Screening without a "gold standard": the Hui-Walter paradigm revisited. *Am J Epidemiol*. 2001;153:921–4.
8. Branscum AJ, Gardner IA, Johnson WO. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Prev Vet Med*. 2005;68:145–63.
9. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics*. 1985;41:959–68.
10. Gardner IA, Stryhn H, Lind P, Collins MT. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Prev Vet Med*. 2000;45:107–22.
11. Kostoulas P, Leontides L, Billinis C, Florou M. Application of a semi-dependent latent model in the Bayesian estimation of the sensitivity and specificity of two faecal culture methods for diagnosis of paratuberculosis in sub-clinically infected Greek dairy sheep and goats. *Prev Vet Med*. 2006;76:121–34.
12. Wells SJ, Collins MT, Faaborg KS, Wees C, Tavoranpanich S, Petrini KR, et al. Evaluation of a rapid fecal PCR test for detection of *Mycobacterium Avium* subsp. *paratuberculosis* in dairy cattle. *Clin Vaccine Immunol*. 2006;13:1125–30.
13. Angelidou E, Kostoulas P, Leontides L. Bayesian validation of a serum and milk ELISA for antibodies against *Mycobacterium Avium* subspecies *paratuberculosis* in Greek dairy goats across lactation. *J Dairy Sci*. 2014;97:819–28.
14. Bauman CA, Jones-Bittton A, Jansen J, Kelton D, Menzies P. Evaluation of fecal culture and fecal RT-PCR to detect *Mycobacterium Avium* Ssp. *paratuberculosis* fecal shedding in dairy goats and dairy sheep using latent class Bayesian modeling. *BMC Vet Res*. 2016;12:212.
15. Norton S, Johnson WO, Jones G, Heuer C. Evaluation of diagnostic tests for Johne's disease (*Mycobacterium Avium* subspecies *paratuberculosis*) in New Zealand dairy cows. *J Vet Diagn Investig*. 2010;22:341–51.
16. Kostoulas P, Leontides L, Enoe C, Billinis C, Florou M, Sofia M. Bayesian estimation of sensitivity and specificity of serum ELISA and faecal culture for diagnosis of paratuberculosis in Greek dairy sheep and goats. *Prev Vet Med*. 2006;76:56–73.
17. Weber MF, Verhoeff J, van Schaik G, van Maanen C. Evaluation of Ziehl-Neelsen stained faecal smear and ELISA as tools for surveillance of clinical paratuberculosis in cattle in the Netherlands. *Prev Vet Med*. 2009;92:256–66.
18. Bermingham ML, Handel IG, Glass EJ, Woolliams JA, de Clare Bronsvooort BM, McBride SH, et al. Hui and Walter's latent-class model extended to estimate diagnostic test properties from surveillance data: a latent model for latent data. *Sci Rep*. 2015;5:11861.
19. Nielsen SS, Toft N. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. *J Dairy Sci*. 2006;89:569–79.
20. Toft N, Jorgensen E, Hojsgaard S. Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. *Prev Vet Med*. 2005;68:19–33.
21. Nielsen SS, Gronbaek C, Agger JF, Houe H. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. *Prev Vet Med*. 2002;53:191–204.
22. Stevenson K. Genetic diversity of *Mycobacterium Avium* subspecies *paratuberculosis* and the influence of strain type on infection and pathogenesis: a review. *Vet Res*. 2015;46:64.
23. Kawaji S, Begg DJ, Plain KM, Whittington RJ. A longitudinal study to evaluate the diagnostic potential of a direct faecal quantitative PCR test for Johne's disease in sheep. *Vet Microbiol*. 2011;148:35–44.
24. Laurin EL, Chaffer M, McClure JT, McKenna SL, Keefe GP. The association of detection method, season, and lactation stage on identification of fecal shedding in *Mycobacterium Avium* Ssp. *paratuberculosis* infectious dairy cows. *J Dairy Sci*. 2015;98:211–20.
25. Timms VJ, Mitchell HM, Neilan BA. Optimisation of DNA extraction and validation of PCR assays to detect *Mycobacterium Avium* subsp. *paratuberculosis*. *J Microbiol Methods*. 2015;112:99–103.
26. Fock-Chow-Tho D, Topp E, Ibeagha-Awemu EA, Bissonnette N. Comparison of commercial DNA extraction kits and quantitative PCR systems for better sensitivity in detecting the causative agent of paratuberculosis in dairy cow fecal samples. *J Dairy Sci*. 2017;100:572–81.
27. Spiess AN. qpcR: Modelling and analysis of real-time PCR data. R package version 1.4-0. [https://cran.r-project.org/web/packages/qpcR/index.html]. 2014. Accessed 31 July 2016.
28. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [https://www.R-project.org/]. Accessed 31 Jul 2016.
29. Nielsen SS, Toft N. Ante mortem diagnosis of paratuberculosis: a review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. *Vet Microbiol*. 2008;129:217–35.
30. Sweeney RW, Whitlock RH, Hamir AN, Rosenberger AE, Herr SA. Isolation of *Mycobacterium paratuberculosis* after oral inoculation in uninfected cattle. *Am J Vet Res*. 1992;53:1312–4.
31. Moloney BJ, Whittington RJ. Cross species transmission of ovine Johne's disease from sheep to cattle: an estimate of prevalence in exposed susceptible cattle. *Aust Vet J*. 2008;86:117–23.
32. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol*. 1995;141:263–72.
33. Enoe C, Georgiadis MP, Johnson WO. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med*. 2000;45:61–81.
34. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol*. 2002;64:583–639.
35. Hope A, Kluver P, Jones S, Condron R. Sensitivity and specificity of two serological tests for the detection of ovine paratuberculosis. *Aust Vet J*. 2000;78:850–6.
36. Sergeant ES, Marshall DJ, Eamens GJ, Kearns C, Whittington RJ. Evaluation of an absorbed ELISA and an agar-gel immuno-diffusion test for ovine paratuberculosis in sheep in Australia. *Prev Vet Med*. 2003;61:235–48.
37. Gumber S, Eamens G, Whittington RJ. Evaluation of a Pourquier ELISA kit in relation to agar gel immunodiffusion (AGID) test for assessment of the humoral immune response in sheep and goats with and without *Mycobacterium paratuberculosis* infection. *Vet Microbiol*. 2006;115:91–101.
38. Stevenson M. EpiR: Tools for the analysis of epidemiological data (Version 0.9-79) [https://cran.r-project.org/web/packages/epiR/index.html]. 2016. Accessed 31 July 2016.
39. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Stat Med*. 2009;28:3049–67.
40. Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. *J Stat Softw*. 2005;12:1–16.
41. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat*. 1998;7:434–55.
42. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Stat Sci*. 1992;4:57–72.
43. Plummer M, Best N, Cowles K, Vines K. Coda: convergence diagnosis and output analysis for MCMC. *R News*. 2006;6(1):7–11.
44. Jones G, Johnson WO, Hanson TE, Christensen R. Identifiability of models for multiple diagnostic testing in the absence of a gold standard. *Biometrics*. 2010;66:855–63.
45. Nielsen SS, Toft N, Jorgensen E, Bibby BM. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. *Prev Vet Med*. 2007;81:290–305.
46. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. *BMJ*. 2013;346:f2778.
47. Angelidou E, Kostoulas P, Leontides L. Bayesian estimation of sensitivity and specificity of a commercial serum/milk ELISA against the *Mycobacterium Avium* subsp. *Paratuberculosis* (MAP) antibody response for each lactation stage in Greek dairy sheep. *Prev Vet Med*. 2016;124:102–5.
48. Englund S, Bolske G, Johansson KE. An IS900-like sequence found in a *Mycobacterium* sp. other than *Mycobacterium Avium* subsp. *paratuberculosis*. *FEMS Microbiol Lett*. 2002;209:267–71.
49. Kawaji S, Taylor DL, Mori Y, Whittington RJ. Detection of *Mycobacterium Avium* subsp. *paratuberculosis* in ovine faeces by direct quantitative PCR has similar or greater sensitivity compared to radiometric culture. *Vet Microbiol*. 2007;125:36–48.
50. Plain KM, Marsh IB, Waldron AM, Galea F, Whittington AM, Saunders VF, et al. High-throughput direct fecal PCR assay for detection of *Mycobacterium Avium* subsp. *paratuberculosis* in sheep and cattle. *J Clin Microbiol*. 2014;52:745–57.
51. Ellingson JL, Bolin CA, Stabel JR. Identification of a gene unique to *Mycobacterium Avium* subspecies *paratuberculosis* and application to diagnosis of paratuberculosis. *Mol Cell Probes*. 1998;12:133–42.

52. Li L, Bannantine JP, Zhang Q, Amonsin A, May BJ, Alt D, et al. The complete genome sequence of *Mycobacterium Avium* subspecies paratuberculosis. *Proc Natl Acad Sci U S A*. 2005;102:12344–9.
53. Mobius P, Hotzel H, Rassbach A, Kohler H. Comparison of 13 single-round and nested PCR assays targeting IS900, ISMav2, f57 and locus 255 for detection of *Mycobacterium Avium* subsp. paratuberculosis. *Vet Microbiol*. 2008;126:324–33.
54. Christopher-Hennings J, Dammen MA, Weeks SR, Epperson WB, Singh SN, Steinlicht GL, et al. Comparison of two DNA extractions and nested PCR, real-time PCR, a new commercial PCR assay, and bacterial culture for detection of *Mycobacterium Avium* subsp. paratuberculosis in bovine feces. *J Vet Diagn Investig*. 2003;15:87–93.
55. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med*. 1997;16:981–91.
56. Thrusfield M. *Veterinary epidemiology*. Third ed. USA: Wiley-Blackwell; 2007. p. 305–30.
57. Georgiadis MP, Johnson WO, Gardner IA. Sample size determination for estimation of the accuracy of two conditionally independent tests in the absence of a gold standard. *Prev Vet Med*. 2005;71:1–10.
58. Huda A, Jungersen G, Lind P. Longitudinal study of interferon-gamma, serum antibody and milk antibody responses in cattle infected with *Mycobacterium Avium* subsp. paratuberculosis. *Vet Microbiol*. 2004;104:43–53.
59. Mortier RA, Barkema HW, De Buck J. Susceptibility to and diagnosis of *Mycobacterium Avium* subspecies paratuberculosis infection in dairy calves: a review. *Prev Vet Med*. 2015;121:189–98.
60. Stewart DJ, Vaughan JA, Stiles PL, Noske PJ, Tizard ML, Prowse SJ, et al. A long-term study in merino sheep experimentally infected with *Mycobacterium Avium* subsp. paratuberculosis: clinical disease, faecal culture and immunological studies. *Vet Microbiol*. 2004;104:165–78.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



Additional file 1 for Mathevon Y, Foucras G, Falguières R and Corbiere F.” Estimation of the sensitivity and specificity of two serum ELISAs and one fecal qPCR for diagnosis of paratuberculosis in sub-clinically infected young adult French sheep using latent class Bayesian modeling”. BMC Vet Research

R code for model 1

#model definition

#AP[k,j] = multinomial probability for the kth population and the jth combination of the tests outcome

model1 <- function() {

 for (k in 1:4){
 pop[k, 1:8] ~ dmulti(AP[k,1:8], n[k])

#multinomial distribution for the infected animals

#a = CovSe12

 prob_se[k,1] <- SE[1] * SE[2] * SE[3] + a * SE[3]
 prob_se[k,2] <- SE[1] * SE[2] * (1 - SE[3]) + a * (1 - SE[3])
 prob_se[k,3] <- SE[1] * (1 - SE[2]) * SE[3] - a * SE[3]
 prob_se[k,4] <- SE[1] * (1 - SE[2]) * (1 - SE[3]) - a * (1 - SE[3])
 prob_se[k,5] <- (1 - SE[1]) * SE[2] * SE[3] - a * SE[3]
 prob_se[k,6] <- (1 - SE[1]) * SE[2] * (1 - SE[3]) - a * (1 - SE[3])
 prob_se[k,7] <- (1 - SE[1]) * (1 - SE[2]) * SE[3] + a * SE[3]
 prob_se[k,8] <- (1 - SE[1]) * (1 - SE[2]) * (1 - SE[3]) + a * (1 - SE[3])

#multinomial distribution for the non infected animals

#b[1] = CovSp12

#b[2] = CovSp13

#b[3] = CovSp23

 prob_sp[k,8] <- SP[1] * SP[2] * SP[3] + b[1] * SP[3] + b[2] * SP[2] + b[3] * SP[1]
 prob_sp[k,7] <- SP[1] * SP[2] * (1 - SP[3]) + b[1] * (1 - SP[3]) - b[2] * SP[2] - b[3] * SP[1]
 prob_sp[k,6] <- SP[1] * (1 - SP[2]) * SP[3] - b[1] * SP[3] + b[2] * (1 - SP[2]) - b[3] * SP[1]
 prob_sp[k,5] <- SP[1] * (1 - SP[2]) * (1 - SP[3]) - b[1] * (1 - SP[3]) - b[2] * (1 - SP[2]) + b[3] * SP[1]
 prob_sp[k,4] <- (1 - SP[1]) * SP[2] * SP[3] - b[1] * SP[3] - b[2] * SP[2] + b[3] * (1 - SP[1])
 prob_sp[k,3] <- (1 - SP[1]) * SP[2] * (1 - SP[3]) - b[1] * (1 - SP[3]) + b[2] * SP[2] - b[3] * (1 - SP[1])
 prob_sp[k,2] <- (1 - SP[1]) * (1 - SP[2]) * SP[3] + b[1] * SP[3] - b[2] * (1 - SP[2]) - b[3] * (1 - SP[1])
 prob_sp[k,1] <- (1 - SP[1]) * (1 - SP[2]) * (1 - SP[3]) + b[1] * (1 - SP[3]) + b[2] * (1 - SP[2]) + b[3] * (1 - SP[1])

#Computation of multinomial probability for the kth population and the jth combination of the tests outcome

#prev[k] = prevalence of the kth population

#control procedure for multinomial probabilities

 for (i in 1:8) {

 AP[k,i] <- prev[k] * prob_se[k,i] + (1 - prev[k]) * prob_sp[k,i]
 constraint3[k,i] <- step(prob_se[k,i])
 constraint4[k,i] <- step(prob_se[k,i] - 1)
 constraint5[k,i] <- step(prob_sp[k,i])

```
  constraint6[k,i] <- step(prob_sp[k,i] - 1)
}
n[k]<- sum(pop[k,1:8])
}
```

#significance levels of covariance terms using the boolean “step” function

```
for (r in 1:3){
  p.b[r] <- step(b[r])
}
p.a<- step(a)
```

#significance of sensitivities specificities and prevalences differences using “step” function

```
SE12<-step(SE[1]-SE[2])
SE13<-step(SE[3]-SE[1])
SE23<-step(SE[3]-SE[2])
SP12<-step(SP[1]-SP[2])
SP13<-step(SP[3]-SP[1])
SP23<-step(SP[3]-SP[2])
prev12<-step(prev[2]-prev[1])
prev23<-step(prev[3]-prev[2])
prev34<-step(prev[4]-prev[3])
```

#computation of sensitivity and specificity of parallel and serial testing**# serial and parallel strategy using Elisa A and Elisa B**

```
se.ser[1]<-SE[1]*SE[2]+a
sp.ser[1]<- 1-((1-SP[1])*(1-SP[2])+b[1])
sp.par[1]<-SP[1]*SP[2]+b[1]
se.par[1]<- 1-((1-SE[1])*(1-SE[2])+a)
```

serial and parallel strategy using Elisa A and qPCR

```
se.ser[2]<-SE[1]*SE[3]
sp.ser[2]<- 1-((1-SP[1])*(1-SP[3])+b[2])
sp.par[2]<-SP[1]*SP[3]+b[2]
se.par[2]<- 1-((1-SE[1])*(1-SE[3]))
```

serial and parallel strategy using Elisa B and qPCR

```
se.ser[3]<-SE[2]*SE[3]
sp.ser[3]<- 1-((1-SP[2])*(1-SP[3])+b[3])
sp.par[3]<-SP[2]*SP[3] +b[3]
se.par[3]<- 1-((1-SE[2])*(1-SE[3]))
```

priors

```
for (r in 1:4){
  prev[r] ~ dbeta(1,1)
}
```

```
SE[1] ~ dbeta(1,1)
SE[2] ~ dbeta(1,1)
SE[3] ~ dbeta(1,1)
```

```
SP[1] ~ dbeta(21.20, 2.06)
SP[2] ~ dbeta(21.20, 2.06)
SP[3] ~ dbeta(21.20, 2.06)
```

#constraints applied to covariance terms

```
se.l <- max(-(1-SE[2])*(1-SE[1]),-SE[2]*SE[1])
se.u <- min(SE[2]*(1-SE[1]),SE[1]*(1-SE[2]))
a ~ dunif(se.l,se.u)
```

```
sp.l1 <- max(((1-SP[1]*SP[2]*(1-SP[3])+b[2]*SP[2]+b[3]*SP[1])/(1-SP[3])),
  ((1+(1-SP[1])*SP[2]*(1-SP[3])+b[2]*SP[2]-b[3]*(1-SP[1]))/(-(1-SP[3])) ) )
sp.u1 <- min(((1-SP[1]*SP[2]*(1-SP[3])+b[2]*SP[2]+b[3]*SP[1])/(1-SP[3])),
  (-(1-SP[1])*SP[2]*(1-SP[3])-b[2]*SP[2]+b[3]*(1-SP[1]))/(-(1-SP[3]))))
b[1] ~ dunif(sp.l1,sp.u1)
```

```
sp.l2 <- max(((1-SP[1]*SP[2]*(1-SP[3]))- b[1]*(1-SP[3])+ b[3]*SP[1] )/(-SP[2])),
  ((-(1-SP[1])*SP[2]*(1-SP[3])+b[1]*(1-SP[3])+b[3]*(1-SP[1]))/(SP[2])))
sp.u2 <- min(((1-SP[1]*SP[2]*(1-SP[3]))- b[1]*(1-SP[3])+ b[3]*SP[1] )/(-SP[2])),
  (1-(1-SP[1])*SP[2]*(1-SP[3])+b[1]*(1-SP[3])+b[3]*(1-SP[1]))/(SP[2]))
b[2] ~ dunif(sp.l2,sp.u2)
```

```
sp.l3 <- max(((1-SP[1]*SP[2]*(1-SP[3]))- b[1]*(1-SP[3])+ b[2]*SP[2] )/(-SP[1])),
  (1+(1-SP[1])*SP[2]*(1-SP[3])+b[2]*SP[2]-b[1]*(1-SP[3]))/(-(1-SP[1])))
sp.u3 <- min(((1-SP[1]*SP[2]*(1-SP[3]))- b[1]*(1-SP[3])+ b[2]*SP[2] )/(-SP[1])),
  ((-(1-SP[1])*SP[2]*(1-SP[3])-b[2]*SP[2]+b[1]*(1-SP[3]))/(-(1-SP[1]))))
b[3] ~ dunif(sp.l3,sp.u3)
}
```

R code for model 2

#AP[k,j] = multinomial probability for the kth population and the jth combination of the tests outcome

```
model2 <- function() {
```

```
  for (k in 1:4){
```

```
    pop[k, 1:8] ~ dmulti(AP[k,1:8], n[k])
```

#multinomial distribution for the infected animals

#a = CovSe12

```
  prob_se[k,1] <- SE[1] * SE[2] * SE[3] + a * SE[3]
```

```
  prob_se[k,2] <- SE[1] * SE[2] * (1 - SE[3]) + a * (1 - SE[3])
```

```
  prob_se[k,3] <- SE[1] * (1 - SE[2]) * SE[3] - a * SE[3]
```

```
  prob_se[k,4] <- SE[1] * (1 - SE[2]) * (1 - SE[3]) - a * (1 - SE[3])
```

```
  prob_se[k,5] <- (1 - SE[1]) * SE[2] * SE[3] - a * SE[3]
```

```
  prob_se[k,6] <- (1 - SE[1]) * SE[2] * (1 - SE[3]) - a * (1 - SE[3])
```

```
  prob_se[k,7] <- (1 - SE[1]) * (1 - SE[2]) * SE[3] + a * SE[3]
```

```
  prob_se[k,8] <- (1 - SE[1]) * (1 - SE[2]) * (1 - SE[3]) + a * (1 - SE[3])
```

#multinomial distribution for the non infected animals

#b = CovSp12

```
  prob_sp[k,8] <- SP[1] * SP[2] * SP[3] + b * SP[3]
```

```
  prob_sp[k,7] <- SP[1] * SP[2] * (1 - SP[3]) + b * (1 - SP[3])
```

```
  prob_sp[k,6] <- SP[1] * (1 - SP[2]) * SP[3] - b * SP[3]
```

```
  prob_sp[k,5] <- SP[1] * (1 - SP[2]) * (1 - SP[3]) - b * (1 - SP[3])
```

```
  prob_sp[k,4] <- (1 - SP[1]) * SP[2] * SP[3] - b * SP[3]
```

```
  prob_sp[k,3] <- (1 - SP[1]) * SP[2] * (1 - SP[3]) - b * (1 - SP[3])
```

```
  prob_sp[k,2] <- (1 - SP[1]) * (1 - SP[2]) * SP[3] + b * SP[3]
```

```
  prob_sp[k,1] <- (1 - SP[1]) * (1 - SP[2]) * (1 - SP[3]) + b * (1 - SP[3])
```

#Computation of multinomial probability for the kth population and the jth combination of the tests outcome

#prev[k] = prevalence of the kth population

#control procedure for multinomial probabilities

```
  for (i in 1:8) {
```

```
    AP[k,i] <- prev[k] * prob_se[k,i] + (1 - prev[k]) * prob_sp[k,i]
```

```
    constraint3[k,i] <- step(prob_se[k,i])
```

```
    constraint4[k,i] <- step(prob_se[k,i] - 1)
```

```
    constraint5[k,i] <- step(prob_sp[k,i])
```

```
    constraint6[k,i] <- step(prob_sp[k,i] - 1)
```

```
  }
```

```
  n[k] <- sum(pop[k,1:8])
```

```
}
```

#significance levels of covariance terms using the boolean "step" function

```
p.a <- step(a)
p.b <- step(b)
```

#significance of sensitivities specificities and prevalences differences using "step" function

```
SE12<-step(SE[1]-SE[2])
SE13<-step(SE[3]-SE[1])
SE23<-step(SE[3]-SE[2])
SP12<-step(SP[2]-SP[1])
SP13<-step(SP[3]-SP[1])
SP23<-step(SP[3]-SP[2])
```

#computation of sensitivity and specificity of parallel and serial testing

serial and parallel strategy using Elisa A and Elisa B

```
se.ser[1]<-SE[1]*SE[2]+a
sp.ser[1]<- 1-((1-SP[1])*(1-SP[2])+b)
sp.par[1]<-SP[1]*SP[2]+b
se.par[1]<- 1-((1-SE[1])*(1-SE[2])+a)
```

serial and parallel strategy using Elisa A and qPCR

```
se.ser[2]<-SE[1]*SE[3]
sp.ser[2]<- 1-((1-SP[1])*(1-SP[3]))
sp.par[2]<-SP[1]*SP[3]
se.par[2]<- 1-((1-SE[1])*(1-SE[3]))
```

serial and parallel strategy using Elisa B and qPCR

```
se.ser[3]<-SE[2]*SE[3]
sp.ser[3]<- 1-((1-SP[2])*(1-SP[3]))
sp.par[3]<-SP[2]*SP[3]
se.par[3]<- 1-((1-SE[2])*(1-SE[3]))
```

#priors

```
for (r in 1:4){
  prev[r] ~ dbeta(1,1)
}
SE[1] ~ dbeta(1,1)
SE[2] ~ dbeta(1,1)
SE[3] ~ dbeta(1,1)
```

```
SP[1] ~ dbeta(21.20, 2.06)
SP[2] ~ dbeta(21.20, 2.06)
SP[3] ~ dbeta(21.20, 2.06)
```

#constraints applied to covariance terms

```
se.l <- max(-(1-SE[2])*(1-SE[1]),-SE[2]*SE[1])
se.u <- min(SE[2]*(1-SE[1]),SE[1]*(1-SE[2]))
a ~ dunif(se.l,se.u)
```

```
sp.l <- max(-(1-SP[2])*(1-SP[1]),-SP[2]*SP[1])
sp.u <- min(SP[2]*(1-SP[1]),SP[1]*(1-SP[2]))
b ~ dunif(sp.l,sp.u)
}
```

R code for model 3

#AP[k,j] = multinomial probability for the kth population and the jth combination of the tests outcome

```
model3 <- function() {
```

```
  for (k in 1:4){
```

```
    pop[k, 1:8] ~ dmulti(AP[k,1:8], n[k])
```

#multinomial distribution for the infected animals

```
    prob_se[k,1] <- SE[1] * SE[2] * SE[3]
```

```
    prob_se[k,2] <- SE[1] * SE[2] * (1 - SE[3])
```

```
    prob_se[k,3] <- SE[1] * (1 - SE[2]) * SE[3]
```

```
    prob_se[k,4] <- SE[1] * (1 - SE[2]) * (1 - SE[3])
```

```
    prob_se[k,5] <- (1 - SE[1]) * SE[2] * SE[3]
```

```
    prob_se[k,6] <- (1 - SE[1]) * SE[2] * (1 - SE[3])
```

```
    prob_se[k,7] <- (1 - SE[1]) * (1 - SE[2]) * SE[3]
```

```
    prob_se[k,8] <- (1 - SE[1]) * (1 - SE[2]) * (1 - SE[3])
```

#multinomial distribution for the non infected animals

```
    prob_sp[k,8] <- SP[1] * SP[2] * SP[3]
```

```
    prob_sp[k,7] <- SP[1] * SP[2] * (1 - SP[3])
```

```
    prob_sp[k,6] <- SP[1] * (1 - SP[2]) * SP[3]
```

```
    prob_sp[k,5] <- SP[1] * (1 - SP[2]) * (1 - SP[3])
```

```
    prob_sp[k,4] <- (1 - SP[1]) * SP[2] * SP[3]
```

```
    prob_sp[k,3] <- (1 - SP[1]) * SP[2] * (1 - SP[3])
```

```
    prob_sp[k,2] <- (1 - SP[1]) * (1 - SP[2]) * SP[3]
```

```
    prob_sp[k,1] <- (1 - SP[1]) * (1 - SP[2]) * (1 - SP[3])
```

#Computation of multinomial probability for the kth population and the jth combination of the tests outcome

#prev[k] = prevalence of the kth population

#control procedure for multinomial probabilities

```
  for (i in 1:8) {
```

```
    AP[k,i] <- prev[k] * prob_se[k,i] + (1 - prev[k]) * prob_sp[k,i]
```

```
    constraint3[k,i] <- step(prob_se[k,i])
```

```
    constraint4[k,i] <- step(prob_se[k,i] - 1)
```

```
    constraint5[k,i] <- step(prob_sp[k,i])
```

```
    constraint6[k,i] <- step(prob_sp[k,i] - 1)
```

```
  }
```

```
  n[k] <- sum(pop[k,1:8])
```

```
}
```

#significance of sensitivities specificities and prevalences differences using "step" function

```
SE12 <- step(SE[2] - SE[1])
```

```
SE13 <- step(SE[1] - SE[3])
```

```
SE23 <- step(SE[2] - SE[3])
```

```
SP12 <- step(SP[2] - SP[1])
```

```
SP13<-step(SP[1]-SP[3])
```

```
SP23<-step(SP[2]-SP[3])
```

```
##computation of sensitivity and specificity of parallel and serial testing
```

```
# serial and parallel testing strategy using Elisa A and Elisa B
```

```
se.ser[1]<-SE[1]*SE[2]
```

```
sp.ser[1]<- 1-((1-SP[1])*(1-SP[2]))
```

```
sp.par[1]<-SP[1]*SP[2]
```

```
se.par[1]<- 1-((1-SE[1])*(1-SE[2]))
```

```
# serial and parallel testing strategy using Elisa A and qPCR
```

```
se.ser[2]<-SE[1]*SE[3]
```

```
sp.ser[2]<- 1-((1-SP[1])*(1-SP[3]))
```

```
sp.par[2]<-SP[1]*SP[3]
```

```
se.par[2]<- 1-((1-SE[1])*(1-SE[3]))
```

```
# serial and parallel testing strategy using Elisa B and qPCR
```

```
se.ser[3]<-SE[2]*SE[3]
```

```
sp.ser[3]<- 1-((1-SP[2])*(1-SP[3]))
```

```
sp.par[3]<-SP[2]*SP[3]
```

```
se.par[3]<- 1-((1-SE[2])*(1-SE[3]))
```

```
# priors
```

```
for (r in 1:4){
```

```
  prev[r] ~ dbeta(1,1)
```

```
}
```

```
SE[1] ~ dbeta(1,1)
```

```
SE[2] ~ dbeta(1,1)
```

```
SE[3] ~ dbeta(1,1)
```

```
SP[1] ~ dbeta(21.20, 2.06)
```

```
SP[2] ~ dbeta(21.20, 2.06)
```

```
SP[3] ~ dbeta(21.20, 2.06)
```

```
}
```

Additional file 2 for Mathevon Y, Foucras G, Falguières R and Corbiere F.” Estimation of the sensitivity and specificity of two serum ELISAs and one fecal qPCR for diagnosis of paratuberculosis in sub-clinically infected young adult French sheep using latent class Bayesian modeling”. BMC Vet Research.

Aggregated data set (4 sub-populations models)

Order of appearance of test results: ELISA A, ELISA B, fecal qPCR

Example: 111 = number of animal with both ELISA A, ELISA B and fecal qPCR results positive

1 line for each subpopulation

\$ Doubtful ELISA results were handled as positive, positive qPCR cut-off Ct ≤ 42

	111	110	101	100	011	010	001	000
1	0	9	0	6	1	9	0	262
2	2	11	1	5	3	6	4	267
3	5	14	2	9	6	7	37	367
4	8	10	2	1	1	1	33	108

\$ Doubtful ELISA results were handled as positive, positive qPCR cut-off Ct ≤ 40

	111	110	101	100	011	010	001	000
1	0	9	0	6	1	9	0	262
2	2	11	1	5	3	6	3	268
3	4	15	2	9	4	9	17	387
4	7	11	1	2	1	1	27	114

\$ Doubtful ELISA results were handled as negative, positive qPCR cut-off Ct ≤ 42

	111	110	101	100	011	010	001	000
1	0	7	0	7	1	5	0	267
2	2	7	1	8	3	4	4	270
3	5	13	2	6	4	6	39	372
4	6	5	3	5	1	1	34	109

\$ Doubtful ELISA results were handled as negative, positive qPCR cut-off Ct ≤ 40

	111	110	101	100	011	010	001	000
1	0	7	0	7	1	5	0	267
2	2	7	1	8	3	4	3	271
3	4	14	2	6	4	6	17	394
4	5	6	2	6	1	1	28	115

2 Article 2 : estimation par une approche fréquentiste à classes latentes

Résumé de l'article. Nous proposons une approche fréquentiste en classes latentes d'évaluation des performances diagnostiques des k tests ($k \geq 2$) dans n populations ($n \geq 2$) en l'absence de test de référence parfait. La fonction d'estimation, nommée Freak et programmée sous le logiciel R, est basée sur l'algorithme Expectation-Maximisation (EM) et repose sur l'hypothèse de performances constantes des tests entre les différentes populations. Pour des études évaluant deux ou trois tests, elle permet de comparer les modèles basés sur des structures différentes de dépendance conditionnelle par un test de rapport de vraisemblance pour les modèles emboîtés ou selon le critère d'Akaike (AIC) pour les modèles non emboîtés. Pour quatre tests et plus, compte tenu de la multiplicité des structures de covariance possibles entre tests (4096 modèles possibles pour quatre tests en prenant uniquement en compte les covariances 2 à 2), seuls sont programmés dans la version actuelle les modèles faisant l'hypothèse d'indépendance conditionnelle entre tous les tests ou à l'inverse ne faisant aucune hypothèse *a priori* (modèle saturé). Un des avantages de notre approche est qu'elle ne nécessite pas de spécifier de contraintes de distribution sur les paramètres de covariance entre tests, contraintes difficiles à calculer lorsque plus de deux tests sont évalués. Notre méthode permet par ailleurs de savoir rapidement si le modèle évalué est identifiable ou non, et dans ce dernier cas, l'espace des solutions possibles au maximum de vraisemblance est fourni. Pour les modèles identifiables, les IC 95% des estimateurs au maximum de vraisemblance sont calculés par bootstrap paramétrique.

Nous avons appliqué notre méthode d'estimation à deux jeux de données ayant fait l'objet d'analyses en classes latentes par des approches bayésiennes, le premier concernant l'évaluation d'un test sérologique ELISA en comparaison d'un test IFAT pour la détection des bovins séropositifs vis-à-vis de *Neospora caninum* (Frössling et al., 2003), le second étant celui analysé dans notre premier article (Mathevon et al., 2017). Sur la base du maximum de vraisemblance ou de l'AIC, la sélection des modèles associés au meilleur ajustement aux données est en accord avec les résultats obtenus par les approches bayé-

siennes basées sur le critère de déviance (DIC). Cependant, Les estimateurs du maximum de vraisemblance obtenus par notre approche fréquentiste, tout en restant proches des estimateurs bayésiens, différaient parfois. De même les intervalles de confiance obtenus par bootstrap étaient globalement plus larges, en raison de la sensibilité de cette approche aux données très déséquilibrées. Notre estimation fréquentiste n'apporte pas de solution définitive aux problèmes de non-identifiabilité, et est moins adaptable que les méthodes bayésiennes dans sa version actuelle. Elle offre cependant au chercheur non rompu à la programmation des méthodes bayésiennes un outil simple d'évaluation des performances diagnostiques des tests en l'absence de test de référence parfait. Par ailleurs des améliorations, notamment par la prise en compte de variables explicatives des résultats aux tests diagnostiques sont possibles et seront envisagées.

Freak : a frequentist approach for diagnosis accuracy studies

Yoann Mathevon¹, Didier Concordet², Fabien Corbière^{1*}

¹ UMR INRA-ENVT 1225 IHAP, Ecole Nationale Vétérinaire de Toulouse, 23 chemin des Capelles, F-31076, Toulouse Cedex 3, France

² UMR INRA-ENVT Toxalim, Ecole Nationale Vétérinaire de Toulouse, 23 chemin des Capelles, F-31076, Toulouse Cedex 3, France

* Corresponding author : f.corbiere@envt.fr

Abstract

We propose a frequentist latent class approach to evaluate the diagnostic accuracy of tests in the absence of a perfect gold standard. Based on an EM algorithm, this estimation procedure allows for comparing several models based on different correlation structures between test results at the same time, from the simplest fully conditionally independent to the fully saturated one. The estimation function, called Freak, is written in R language and also allows detecting non identifiable models in a very simple way, and in such cases, gives all equally plausible parameters estimates obtained at the maximum likelihood. For identifiable models, parametric bootstrap confidence intervals are proposed. The proposed approach was applied to two working examples from the literature and yielded accuracy estimates close to those obtained by Bayesian latent class modeling.

Keywords: EM algorithm, frequentist models, latent class models, sensitivity, specificity, diagnostic accuracy

Estimating diagnostic test performances (sensitivity, specificity) and disease prevalence in the absence of results from a perfect gold standard is of major interest in epidemiology. It is a challenging task because statistical models are not identifiable when no assumption is made about the correlation structure between the test results. Thus an infinite number of model parameter values can lead to the observed data. Various bayesian and frequentist approaches have been proposed to overcome this difficulty [14]. The model is then made identifiable either by modeling the tests dependence by a common latent random variable (probit models [13, 15], and/or by assuming specific correlation structures between the test results [6, 3, 8]. These approaches are therefore based on prior information and strong model assumptions, respectively. Consequently, no method is currently available for analyzing data in the absence of hypotheses or priors. Without any prior information or strong assumption, we propose a method that provides the entire maximum likelihood set of solutions for non-identifiable models, which is reduced to a unique solution for identifiable ones. This approach is written in R language and also allows medical researchers not familiar with mathematical programming to compare models based on different assumptions regarding the correlation structures between test results.

1 Background and notations

1.1 Notations

To present our method and the notation used we will rely on the example of 2 imperfect binary tests (T_1 and T_2). Assuming that there are no missing results, we can describe a population of size n according to tests results as follows:

- n^{++} : animals testing positive for both tests (T_1^+ & T_2^+)
- n^{+-} : animals testing positive for test 1 (T_1^+) and negative for test 2 (T_2^-)
- n^{-+} : animals testing negative for test 1 (T_1^-) and positive for test 2 (T_2^+)
- n^{--} : animals testing negative for both tests (T_1^- & T_2^-)
- with $n = n^{++} + n^{+-} + n^{-+} + n^{--}$

In a latent class analysis, a non observable target is used to divide the population between two classes which we will call healthy and diseased for a better understanding. Given the unknown prevalence (p) there are $n_1 = n * p$ diseased individuals and $n_0 = n * (1 - p)$ healthy ones, and the probability distribution function for the test outcomes is defined as described in Table 1. Two sample spaces are now defined, \mathcal{X} containing complete data (i.e. multinomial parameters and prevalence) and \mathcal{Y} containing the observed data (i.e. distribution of the n individuals).

Table 1: Contingency table for 2 binary tests

Diseased	T_2^+	T_2^-	
T_1^+	$n_1 * q_1^{++}$	$n_1 * q_1^{+-}$	$n_1 * q_1^{+ \cdot}$
T_1^-	$n_1 * q_1^{-+}$	$n_1 * q_1^{--}$	$n_1 * q_1^{- \cdot}$
	$n_1 * q_1^{+ \cdot}$	$n_1 * q_1^{- \cdot}$	n_1
Healthy	T_2^+	T_2^-	
T_1^+	$n_0 * q_0^{++}$	$n_0 * q_0^{+-}$	$n_0 * q_0^{+ \cdot}$
T_1^-	$n_0 * q_0^{-+}$	$n_0 * q_0^{--}$	$n_0 * q_0^{- \cdot}$
	$n_0 * q_0^{+ \cdot}$	$n_0 * q_0^{- \cdot}$	n_0

As D^+ and D^- is a finite partition of \mathcal{X} , $P(T_1^+) = P(T_1^+|D^+) * P(D^+) + P(T_1^+|D^-) * P(D^-)$. Other probabilities are easily derived using the same formula. Diagnostic performances for test i , (i.e sensitivity, Se_i and specificity Sp_i , $i = 1, \dots, 2$) can be computed from the marginal distribution over the 2 sub-populations :

$$\begin{cases} Se_1 &= P(T_1^+|D^+) = q_1^{++} + q_1^{+-} = q_1^{+ \cdot} \\ Sp_1 &= P(T_1^-|D^-) = q_0^{--} + q_0^{-+} = q_0^{- \cdot} \\ Se_2 &= P(T_2^+|D^+) = q_1^{++} + q_1^{-+} = q_1^{+ \cdot} \\ Sp_2 &= P(T_2^-|D^-) = q_0^{--} + q_0^{+-} = q_0^{- \cdot} \end{cases} \quad (1)$$

1.2 E-M algorithm

The best estimates in these parametric models such as the one described in table 1 can be obtained by maximising the likelihood. A convenient way to estimate the parameters in these models is to apply the Expectation-Maximization (EM) algorithm [5]. The EM algorithm increases the likelihood function at each iteration step and converges on a maximum of this likelihood. Unfortunately, this maximum is not necessarily the global maximum. The EM algorithm adapted to this class of models is described in the appendix. The models we are dealing with in this article can be identifiable (they have a single likelihood maximum) or non-identifiable (they have a likelihood with a continuum of global maxima). In both cases they probably have a number of local maxima too. So, using an EM algorithm will guarantee failure. A deterministic EM algorithm can be interpreted as a function which provides a single maximum (global and local) from a chosen initial parameter value (p, q_0, q_1) . For a given maximum, there exists a set of initial values of the EM algorithm. This set of initial values is usually named the "basin of attraction" for the corresponding maximum. The set of basins of attraction forms a partition of the parametric space. Two basins of attraction are disjoint and the union of all these basins equals the parametric space. If we were able to choose an initial value for the EM algorithm in each basin of attraction, we would get all the maxima (local and global) of the likelihood. An iterative method that allows exploration of the set of parameter values maximising the likelihood, is to:

- choose a starting point for the EM algorithm in one basin of attraction
- run the EM algorithm
- compare the obtained maximum with those already collected, and finally
- collect the final parameters if the absolute maximum is reached.

Because the set of basins of attraction is not known, we choose to draw it uniformly within the parameter space. Starting values for the prevalences p_i are therefore drawn from independent uniform distributions on $(0; 1)$ while the starting values for q_0 and q_1 are drawn from a $2 * (N - 1)$ simplex, with N the number of tests, allowing to sample every point in a $(N - 1)$ -dimensions space with the sum of each vector equals to 1. In short, we randomly draw the starting values of EM and keep only those final values for which the absolute likelihood maximum is reached.

1.3 Correlation structures and model selection

Without further parametrization, the EM algorithm will compute the correlations between tests on diseased animals as well as on healthy ones. Two events are independent if their joint probability equals the product of each probabilities Here independence is defined *conditionally* on the latent class D^+ or D^- , leading to the following system of equations:

$$\begin{cases} P(T_1^+ \cap T_2^+ | D^+) = P(T_1^+ | D^+) * P(T_2^+ | D^+) = S e_1 * S e_2 \\ P(T_1^- \cap T_2^- | D^-) = P(T_1^- | D^-) * P(T_2^- | D^-) = S p_1 * S p_2 \end{cases} \quad (2)$$

Using the notations defined earlier, pairwise correlations between tests can be defined as previously proposed by Dendukuri et al. [6].

$$\begin{cases} q_1^{++} - Se_1 * Se_2 = CovSe_{12} \\ q_0^{--} - Sp_1 * Sp_2 = CovSp_{12} \end{cases} \quad (3)$$

As the fully saturated model might not be identifiable, or not be the model leading to the best fit to the data, correlation structure can be forced around the EM algorithm. After drawing initial values, the vectors q_0 and q_1 are bent to have the same correlation structure without modifying sensitivity or specificity parameter values. After the maximisation step but prior to the assessment of convergence, q_0 and q_1 are modified again. This can be seen as exploring only the part of the parametric space that corresponds to the desired correlation structure. In the current version of our R function, correlation structures for 2-test and 3-test models are implemented, from the fully independent to the fully dependant (saturated) one, with the possibility of specifying that a test is independent from the 2 others. For more than 3 tests, only the fully dependant and independent models are currently implemented.

Extending notations defined earlier for 3 tests, and assuming that T_1 is independent from T_2 and T_3 conditionally on both healthy and diseased latent class, the following system of equations is obtained:

$$\begin{cases} P(T_1^+ \cap T_2^+ \cap T_3^+ | D^+) = P(T_1^+ | D^+) * P(T_2^+ \cap T_3^+ | D^+) = q_1^{++} * q_1^{++} \\ P(T_1^- \cap T_2^- \cap T_3^- | D^-) = P(T_1^- | D^-) * P(T_2^- \cap T_3^- | D^-) = q_0^{--} * q_0^{--} \end{cases} \quad (4)$$

The total number of correlation structures is represented in fig. 1 for 3 tests. Three pairwise dependencies are possible and the total number of models considering only healthy or diseased individual is equal to:

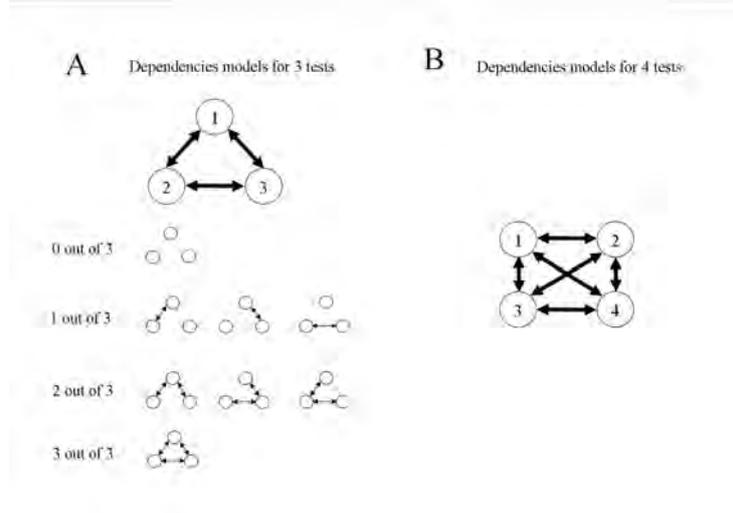
$$\sum_{i=0}^3 \binom{i}{3} = 8 \quad (5)$$

As models with 2 pairwise dependencies between 3 tests are difficult to interpret at a biological level, we choose to consider only models based either on 1 pairwise dependence, 3 dependent tests or 3 independent tests, reducing the numbers of possibilities to 5 for both healthy and diseased latent classes, and overall to 25 models. When considering 4 tests, the possible links between tests are shown in fig. 1. There are 6 possibles links when considering only pairwise dependencies and the total numbers of correlation structures for each latent class (healthy or diseased) equals 64, leading to 4096 models. The number of models has to be reduced and they cannot be quickly evaluated exhaustively. One approach to reduced the number of models could be to consider only 2 or less pairwise dependencies, leading to 22 models for each latent class, with a total numbers of models of 484.

The fit to the data provided by different models can be compared using several indicators, including the likelihood ratio test for nested models. The Aikake information criterion (AIC) and its corrected version taking into account small samples are also computed to allow comparing non nested models.

1.4 Identifiability

Latent class models, because they are a mixture of unobserved distributions, are never identifiable because of the "label-switching" or "label swapping" phenomenon [2, 10]. Switching the label for the latent class will result in a symmetric sets of solutions to the estimation problem. To overcome this issue, some authors suggested imposing $Se + Sp > 1$ for all diagnostic



Tests are represented by circles and pairwise dependencies by arrows.

A) With 3 tests, 8 models are possible on sensitivity or specificity. Choosing 2 out of 3 pairwise dependencies implies that 1 test is correlated with the 2 others while these 2 are independent. This situation does not seem credible and is not currently implemented in our function.

B) With 4 tests the total number of possible correlation structure based on pairwise dependencies is 64 for each latent class.

Figure 1: Correlation structures for 3-test and 4-test models.

tests [9, 10, 16]. By doing so, it is assumed that all tests provide information about the disease status of individuals. Moreover, imposing $Se + Sp > 1$ on all tests also leads to the selection of prevalence values.

However, even if a single set of solutions has been selected in response to label switching, identifiability cannot be guaranteed. In fact, when q_0 and q_1 are free to vary within the whole simplex (that is, when no specific structure of conditional correlation is imposed between tests), the model still has an infinite number of solutions whatever the number of populations and tests. There always exists a continuum of solutions that live in a space with at least 2 dimensions. The identifiability determination is basically a problem of finding a unique solution of an equation system. The numbers of independent equations is equal to the number of degrees of freedom (df) in the data and the parameters (r) to be estimated are the unknowns of this system. If there are more parameters to be estimated than distinct equations ($r > df$), then there are no unique set of answers and the system is non-identifiable.

If $df \geq r$, the system may be solved, at least partially and the model is locally identifiable once the label switching is dealt with. In the situation of 2 populations (with distinct prevalences p_1 and p_2) and 2 conditionally independent tests, the parameters to be estimated are the 2 prevalences, the 2 Se and the 2 Sp . For one population, df is given by the 2x2 contingency table where if 3 numbers are set, the last one will be deduced from the total. If the 2 tests are not conditionally independent, 2 more parameters have to be estimated i.e. the covariance terms Cov_{Se} and Cov_{Sp} between the 2 tests on Se and Sp , respectively. Equation 6 is the

generalisation of the 2-test 2-population example to m populations and k tests:

$$\begin{cases} df = m * (2^k - 1) \\ r = 2 * k + \binom{2}{k} + m \end{cases} \quad (6)$$

It is possible to obtain identifiability for some specific correlation structures between tests. As a special case, models assuming the independence of tests conditionally to the actual disease status of individuals, as proposed by Hui and Walter [9] are identifiable and have been widely studied. This assumption may sometimes hold, but does not seem to be biologically reasonable especially when different diagnostic tests are based on the same biological process. When erroneously assumed, the conditional independence between tests assumption can seriously bias parameter estimations [17, 8, 1, 12]. For some intricate correlation structures between tests, models may be identifiable [10, 6]. However, the *a priori* determination of identifiability properties for such models is rather technical (formal computation of the Jacobian of the transformation described in Equation 1 and study of its zeros)[10]. As formal identifiability is not needed to evaluate the diagnostic performances of tests, in our function a model was assumed to be identifiable if all the parameter estimates from EM runs converged with a maximal difference of 10^{-4} for each elements of the q_0 and q_1 vectors, after discarding the 5% extreme values.

1.5 Computation of parameter estimate confidence intervals

For identifiable models only confidence interval (CI) for parameter estimates are based on parametric bootstrap. Briefly, 10000 samples are simulated following 2 multinomial distributions, one for each parameter vector q_0 and q_1 , and 10 EM estimations, based on 10 different starting points, are conducted for each sample. Only maximum likelihood vectors of solutions are kept for each estimation run and 95% confidence intervals are built based on 2.5% and 97.5% percentiles of the estimate distributions over the 10000 iterations.

2 Application to real data

Two datasets were evaluated using the proposed frequentist approach, both drawn from already published articles deemed to evaluate the diagnostic accuracy of diagnostic tests using latent class Bayesian modelling. Model selection and test evaluation was performed on these two datasets following the steps presented in fig. 2. One thousand initial values were drawn in the N-simplex and each correlation structure was evaluated. If a model was judged identifiable, parametric bootstrap confidence for parameter estimates were computed based on 10000 resampling.

The first example was presented by Frösling et al [7]. The aim of the study was to evaluate the performances of serum ELISA and an indirect-fluorescence antibody test (IFAT) for the detection of *Neospora caninum* in five cattle herds grouped in 2 populations based on previous seroprevalence results. The contingency table 2 shows the observed results for the 2 tests from 244 sampled animals.

2.1 Evaluation in the 2-test 2-population design: *Neospora caninum* example

The maximum likelihood and corresponding Aikake Information Criterion (AIC) and Corrected AIC (AICc) estimated using our proposed method, are shown in table 3 for the 4 possible

Table 2: Contingency table for IFAT and ELISA in 5 herds grouped in 2 populations

Test		Populations	
IFAT	ELISA	Low prevalence	High prevalence
+	+	8	55
+	-	0	1
-	+	7	22
-	-	115	36

Data from Frösling et al (2003)

correlation structures using the *Neospora caninum* example. As expected the fully saturated model (model 2) is not identifiable. Whatever the criterion used (maximum likelihood, AIC, AICc), the best fitting model was the one where the 2 tests are assumed conditionally independent on diseased animals (i.e. on sensitivity) and conditionally dependent on healthy ones (i.e. on specificity) (model 1). Interestingly, model 4, which assumes that the 2 tests are assumed conditionally dependent on sensitivity but independent on specificity, was not identifiable. This result highlights the impossibility of assessing identifiability solely based on degrees of freedom and numbers of parameters to be estimated.

Table 3: Model selection step for all the correlation structure between 2 tests using data from Frösling et al [7]

	Se	Model		Maximum LL ¹	Selection		
		Sp	# parameters		AIC	AICc	Identifiability ²
1	IND ³	DEP ⁴	7	-179.369	362.738	344.071	yes
2	DEP	DEP	8	-179.369	364.738	364.167	no
3	IND	IND	6	-179.504	363.007	346.207	yes
4	DEP	IND	6	-179.504	365.007	346.341	no

¹ maximum likelihood at the end of the FREAK run

² Identifiability was assessed by calculating the range of each element of the q_0 and q_1 vector. Model was identifiable if the maximum range was less than 10^4 for all the run ending at the maximum likelihood.

³ tests were considered independent (on either Se or Sp) ⁴ tests were considered dependent (on either Se or Sp)

Two models were selected for further computation, the best fitting model (model 1) and the independent one (model 3), to be compared with Frösling et al. [7] (table 4). Results from the independent model are close to those presented by Frösling et al. [7] with non informative priors. Using this model, prevalence estimates are also similar to those calculated with the same dataset in an other study by Branscum et al. [4] who also took into account different correlation structures under a Bayesian framework. These authors also concluded that tests results were correlated on healthy animals. Our estimates from this model were the same than those proposed by Branscum et al. (2005) for both sensitivities and moderately lower for both specificities. Also, the confidence intervals computed using 1000 parametric bootstrap resampling were really close to the credible interval provided by the Bayesian posterior distribution. Fi-

nally, the fully dependent model was not identifiable and all the equally probable parameters space is provided (table 4), which should not be confounded with a confidence interval as each set of parameters resulted in the same maximum likelihood.

Table 4: Point estimates and confidence interval for the *Neospora caninum* dataset example [7]

Models		fully saturated	independent		dependency on Sp	
		PE ¹	PE	CI ²	PE	CI
IFAT	Se	[0.491 ; 0.730]	0.726	[0.611 ; 0.829]	0.726	[0.617 ; 0.824]
	Sp	[0.938 ; 1.000]	1	[0.992 ; 1.000]	0.938	[0.901 ; 0.984]
ELISA	Se	[0.675 ; 0.986]	0.981	[0.946 ; 1.000]	0.981	[0.870 ; 1.000]
	Sp	[0.885 ; 0.966]	0.966	[0.916 ; 1.000]	0.885	[0.792 ; 0.976]
CovSe ³		[-0.003 ; 0.151]	-	-	-	-
CovSp ⁴		[0.000 ; 0.054]	-	-	0.054	[0.014 ; 0.089]
Prevalence pop. 1		[0.000 ; 0.125]	0.085	[0.048 ; 0.121]	0.000	[0.000 ; 0.064]
Prevalence pop. 2		[0.643 ; 1.000]	0.676	[0.649 ; 0.707]	0.645	[0.576 ; 0.763]

¹ Point estimates.

² Parametric bootstrap confidence interval based on 100 resampling.

³ Covariance term on diseased animals.

⁴ Covariance term on healthy animals.

2.2 Evaluation in the 3-test multiple-population design: *Ovine paratuberculosis* example

We also applied our proposed frequentist estimation approach to data that we previously analyzed [11]. Based on results from 1197 ewes from 14 flocks, grouped in 4 populations of increasing apparent prevalence, the sensitivity and specificity of two serum ELISAs and one fecal qPCR were evaluated for diagnosis of paratuberculosis in sub-clinically infected sheep using a latent class Bayesian approach. Results for the 3 tests from the 1197 tested ewes are summarised in table 5.

Table 5: Cross-classified positive (+) and negative (-) results of ELISA A, ELISA B and fecal qPCR in four sub-populations of sheep [11].

Test			Populations			
ELISA A	ELISA B	Fecal QPCR	Pop. 1	Pop. 2	Pop. 3	Pop.4
+	+	+	0	2	5	8
+	+	-	9	11	14	10
+	-	+	0	1	2	2
+	-	-	6	5	9	1
-	+	+	1	3	6	4
-	+	-	9	6	7	1
-	-	+	0	4	37	32
-	-	-	262	267	367	108

The model selection step for the ovine paratuberculosis dataset is shown in Table 6 for 5 models based on different correlation structures. The fully saturated model (model 1) was not identifiable, despite the fact that the number of degrees of freedom (28) was greater than the number of parameters to be estimated (18). Similarly model 3 was not identifiable. Based on maximum likelihood, AIC or AICc, the best identifiable model was the one with pairwise dependence between ELISA on sensitivity and pairwise dependence for the 3 tests on specificity (model 2), which is in accordance with results obtained in our previous analysis using a Bayesian framework. Furthermore this model was not statistically different from the full saturated one (likelihood ratio test, $\chi_3^2 = 0.3925$). As with the *Neospora caninum* dataset, we found that 2 models with the same numbers of parameters (models 2 and 3) but based on different correlation structures can be identifiable or not.

Table 6: **Model selection step for the correlation structure between 3 tests using data from Mathevon et al (2017) [11]**

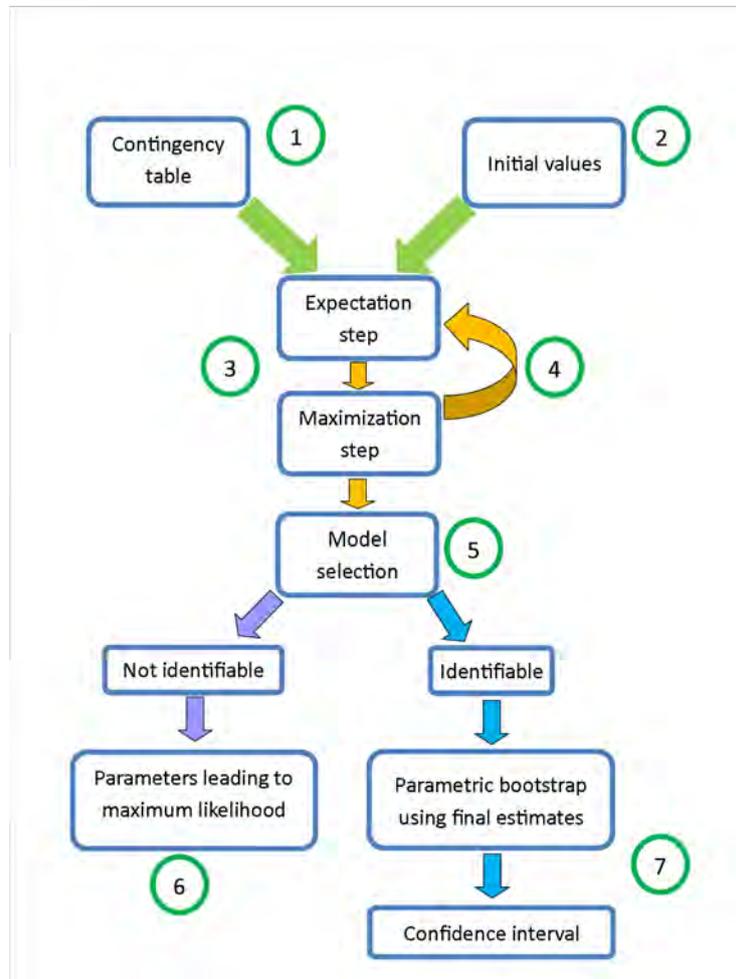
	Se ¹	Model		Maximum LL ²	Selection		Identifiability ³
		Sp ¹	# parameters		AIC	AICc	
1	DEP	DEP	18	-786.716	1609.433	1563.833	no
2	3	DEP	15	-788.213	1606.426	1566.426	yes
3	DEP	3	15	-791.629	1613.258	1573.258	no
4	3	3	12	-794.927	1613.855	1579.188	yes
5	IND	IND	10	-845.4257	1710.851	1679.423	yes

¹ correlation structure can be either

- DEP: the 3 tests are dependent
- IND: the 3 tests are independent
- 3: the 3rd test is independent from the 2 others

² Maximum log-likelihood at the end of the FREAK run

³ Identifiability was assessed by calculating the range of each element of the q_0 and q_1 vector. The model was assumed to be identifiable if the maximum range was less than 10^{-4} for all the runs leading to the maximum likelihood.



1) Using contingency table for each population 2) and initial values for the EM algorithm sampled from a N-simplex 3) there is a first expectation step followed by a maximization step. 4) After the maximization step, if convergence is not attained, another run of EM algorithm is performed 5) After all models were evaluated, a selection step using maximum likelihood or AIC is needed 6) If the best model is not identifiable all possible results are given 7) if the selected model is identifiable, a bootstrap based on parameters estimated is performed to calculate confidence interval

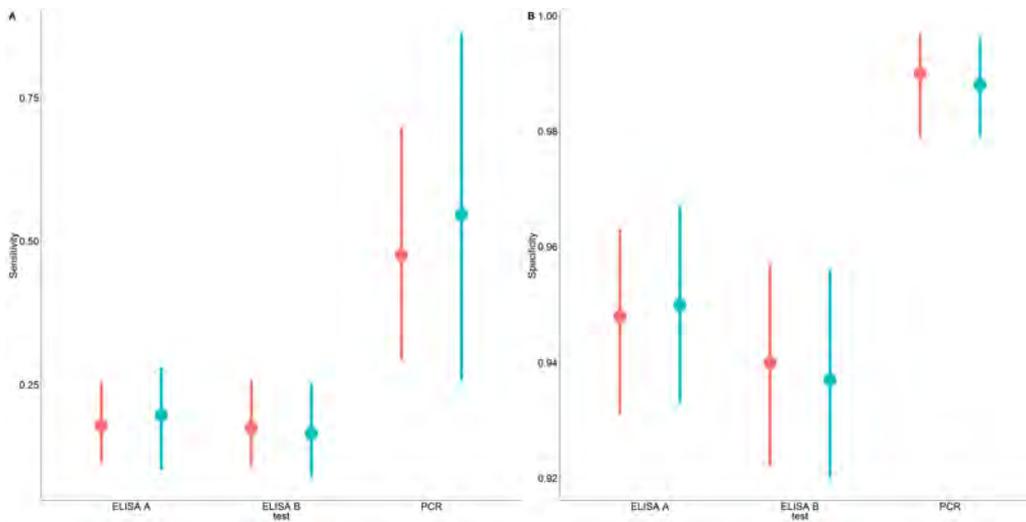
Figure 2: Freak function using EM algorithm

Table 7: Estimates for sensitivity and specificity of two serum ELISAs and one fecal qPCR using a frequentist approach (*ovine paratuberculosis* example).

Model	Se_1	Se_2	Se_3	Sp_1	Sp_2	Sp_3
1	[0.11-0.13]	[0.11-0.12]	[0.28-0.35]	[0.95-0.95]	[0.94-0.94]	[0.99-0.99]
2	0.19 (0.10-0.28)	0.16 (0.09-0.25)	0.55 (0.26-0.86)	0.95 (0.93-0.97)	0.94 (0.92-0.96)	0.99 (0.98-0.99)
3	[0.11-0.13]	[0.12-0.14]	[0.27-0.34]	0.95	0.94	1
4	0.20 (0.12-0.28)	0.25 (0.17-0.34)	0.85 (0.48-0.98)	0.94 (0.93-0.96)	0.94 (0.93-0.96)	1 (0.99-1.00)

For non-identifiable models (1 and 3) the range of equally possible solutions are given in square brackets; for identifiable models (2 and 4) point estimates are given with 95% confidence intervals in brackets

Point estimates for models 1 to 4 are shown in table 7. Although not strictly identifiable, model 1 and 3 led to very narrow parameter spaces (maximum difference less than 0.02), except for Se_3 estimates (maximum difference 0.7). Interestingly, model 3 led to non-identifiability for sensitivity estimates, but as Sp_3 was estimated to be 1, specificity estimates were mathematically conditionally independent. Comparison between Bayesian and frequentist sensitivity and specificity estimates from model 2 are depicted in fig. 3. The frequentist confidence intervals were slightly larger than the Bayesian credible intervals for sensitivity estimates, with a maximum difference in point estimates of 2% for serum ELISAs. The frequentist point estimate of the fecal qPCR sensitivity (55%) was however higher than the median of the posterior distribution provided by the Bayesian approach (47.5%).



Medians and 95% Bayesian posterior credible intervals are displayed in red. Frequentist point estimates and 95% bootstrap confidence intervals are displayed in blue.

Figure 3: Bayesian and frequentist sensitivity (left) and specificity (right) estimates (dots) and 95% posterior credible or confidence interval (bars) for the *ovine paratuberculosis* example.

3 Discussion

We propose a frequentist approach to evaluate the diagnostic accuracy of tests in the absence of a perfect gold standard. Based on two illustrative examples, we showed that estimates from our frequentist approach were close to those obtained under the Bayesian framework. In the current version of our function, confidence intervals are computed using parametric bootstrap resampling which may be computationally intensive and time consuming. The parametric bootstrap computation of confidence intervals is sensitive to very unbalanced data (i.e. with many 0 in cross-tabulated test results) possibly leading to non identifiability issues for some bootstrap samples and to slightly increased confidence intervals.

For up to 3 tests, our approach allows for the evaluation of all models based on different correlation structures between tests. For more than 3 tests, the current version of our function

only computes estimates for the fully saturated and independent models. Indeed the number of possible correlation structure grows exponentially with the number of tests (for instance 4096 for 4 tests), which rapidly becomes intractable. The only main assumption that is still required is that test performances are constant across multiple populations. One of the main benefits of our approach over the Bayesian modeling is that no constraints have to be defined for covariance terms, which may be difficult to specify when more than 2 two are evaluated [6]. Furthermore for users not familiar with Bayesian modeling and programming, the proposed function allows for a one-step evaluation of several models correlation structures and for ranking them according to classical statistical criterion (maximum likelihood, AIC or AICc). Using Bayesian modeling would require to program and run separately all models and compare them based, for instance, on the Deviance Information Criterion. Finally, while detecting non identifiable models is sometimes difficult under the Bayesian framework, as the only indicator may be that several simulation chains failed to converge to a stationary posterior distribution, our approach is based on a simple comparison between parameters estimates that yielded to the same (maximum) likelihood. While not better than or a substitute for Bayesian latent class modeling, our frequentist approach shall rather be regarded as a complementary approach, especially for medical researchers not familiar with programming. Some extensions to the present function are possible, for instance by taking into account the possibility that test results may depend on covariates. The estimation function is currently implemented in R software language and is provided as additional file.

Appendix A EM algorithm

Let us first choose notations in the general case where K tests are studied using samples coming from S populations.

For each individual, the observed results for the K tests are stored in a cell of a K -way contingency table containing 2^K cells. This table is obtained by crossing all the test results. To simplify the notations, we choose to denote 0 and 1 respectively as a negative and a positive test result. We shall first number these cells in order to store column-wise the results of each individual.

Each cell of a given contingency table is numbered with an integer $k \in \{1, \dots, 2^K\}$. The cell corresponding to the test results $(j_1, \dots, j_K) \in \{0,1\}^K$ is numbered as $k = 1 + j_1 2^0 + \dots + j_K 2^{K-1} \in \{1, \dots, 2^K\}$. Actually, the results of these K tests can be interpreted as the dyadic decomposition of $k-1$.

Now, let us build the 2^K vector of observation \mathbf{Z}_{si} for the i^{th} individual from the s^{th} population. All the components of this vector are null except the k^{th} component corresponding to the cell containing the tests results for this individual.

Actually, the observation Z_{si} can be written as

$$Z_{si} = X_{si} \mathbf{Z}_{si}^1 + (1 - X_{si}) \mathbf{Z}_{si}^0, i = 1, \dots, n_s$$

where the 2^K vectors \mathbf{Z}_{si}^1 and \mathbf{Z}_{si}^0 contain the results that would be observed if the i^{th} individual from the s^{th} population were diseased and healthy, respectively. The Bernoulli random variable X_{si} takes the value 1 with probability p_s for a diseased individual and equals 0 otherwise. If the test performances do not depend on the population and if all the individuals are independent, each of the unobserved vectors \mathbf{Z}_{si}^1 and \mathbf{Z}_{si}^0 follows a multinomial distribution and we have

$\mathbf{Z}_{si}^1 \sim_{iid} M(1, q1)$ and $\mathbf{Z}_{si}^0 \sim_{iid} M(1, q0)$.

The vectors $\mathbf{q1}$ and $\mathbf{q0}$ contain the probabilities to belong to the different cells of the K-way contingency table for the diseased and healthy individuals.

Written like this, the distribution of the observations \mathbf{Z}_{si} can be seen as a mixture of two multinomial distributions; the prevalences $\mathbf{p} = (p_1, \dots, p_S)$ being the weights of this mixture.

Now, we have enough notations to estimate the parameter $\theta = (\mathbf{p}, \mathbf{q0}, \mathbf{q1})$ of this model. The EM algorithm converges to a maximum (global or local) of the marginal likelihood. It iterates between an E-step and an M-step. The E-step computes, at each iteration t , the expectation of complete data log-likelihood given the observations and the current model parameter estimates obtained at iteration $(t-1)$.

The M-step finds new parameter estimates that maximize this expected log-likelihood.

At iteration t , we have θ^{t-1} and the expectation step has the following expression

$$Q(\theta, \theta^{t-1}) = \sum_{s=1}^S \sum_{i=1}^{n_s} E_{\theta^{t-1}} [\log L_{\theta}(Z_{si}, X_{si}) | Z_{si}]$$

where

$$L_{\theta}(Z_{si}, X_{si}) = P_{X_{si}} \prod_{l=1}^{2^K} q_{X_{si}l}^{Z_{sil}}$$

and $P_{X_{si}} = p_s, q_{X_{si}} = q1$ when $X_{si} = 1, P_{X_{si}} = 1 - p_s, q_{X_{si}} = q0$ when $X_{si} = 0$.

The M step consists of maximizing $Q(\theta, \theta^{t-1})$ with respect to θ to get

$$\theta^{t-1} = \underset{\theta}{\text{Arg sup}} Q(\theta, \theta^{t-1}).$$

Let us denote

$$\left\{ \begin{array}{l} p_s = p_s^{t-1}, q1_l = q1_l^{t-1}, q0_l = q0_l^{t-1}, \\ a_{si} = \frac{p_s \prod_{l=1}^{2^K} q1_l^{Z_{sil}}}{p_s \prod_{l=1}^{2^K} q1_l^{Z_{sil}} + (1 - p_s) \prod_{l=1}^{2^K} q0_l^{Z_{sil}}}, \\ b_l = \sum_{s=1}^S \sum_i^{n_s} Z_{sil} a_{si}, l = 1, \dots, 2^K, \\ c_l = \sum_{s=1}^S \sum_i^{n_s} Z_{sil} (1 - a_{si}), l = 1, \dots, 2^K, \end{array} \right.$$

using the Lagrange multiplier method, at iteration t of the EM algorithm we get $\theta^t =$

$(p^t, q0^t, q1^t)$ by computing:

$$\left\{ \begin{array}{l} p_s^t = \sum_{i=1}^{n_s} a_{si} / n_s, s = 1, \dots, S \\ q1_l^t = b_l / \sum_{m=1}^{2^K} b_m, l = 1, \dots, 2^K \\ q0_l^t = c_l / \sum_{m=1}^{2^K} c_m, l = 1, \dots, 2^K. \end{array} \right.$$

Appendix B R code of the freak function

Freak function

Description

Performs a frequentist test evaluation of 2 to 3 diagnostics tests applied on p populations.

Arguments

- np:** numeric: number of population in the study.
- count:** vector: number of results for each combination (as described in the article) of tests results in each population.
- nb_try:** numeric: number of starting values for the EM algorithm.
- nt:** numeric: number of tests under evaluation.
- model:** NULL for a screening (all models are evaluated)
dataframe: models to be evaluated are coded as follow:
- "DEP":** fully saturated
 - "IND":** fully independent
 - numeric:** indicates which test is independant from the 2 other;
- first column of the dataframe contain dependency structure for sensitivities, second one for specificities
- Spf:** numeric: Number of the test with a Sp fixed to 1
- full:** FALSE: freak function returns a synthetic data frame of results for identifiable models
TRUE: freak function returns a list with the full data available

Values

if full==FALSE, returns a dataframe of results for identifiable models

if Full==TRUE, returns a list with a sublist for each model evaluated containing all results (parameters obtained). Initials values and statistical hierarchisation paameters can be found in the 2 last sublists.

```
freak <- function(np, count = NULL, nb_try = 100, nt, model = NULL, Spf = NULL, full = F, ...)
{
  require(parallel)

  if (length(count) != np * 2^nt) {
    stop("Check your data, there are inconsistencies between number
         of population/nb of test/ count")
  }

  if (is.null(model) == F & is.data.frame(model) == F) {
    stop("models to be evaluated must be given in a data frame
         (first column Se, seconde column Sp) or left null for a full evaluation")
  }
  results <- list()
  log_like_dep = -1e+09
  lambda = 1
  pop <- rep(1:np, each = 2^nt)
  nb_tir <- nb_try
  Test <- as.vector(
    as.matrix(expand.grid(
      as.data.frame(matrix(rep(c(0, 1), time = nt), ncol = nt))
    )))
}
```

```

)
Test <- matrix(data = Test[length(Test):1], ncol = nt)
nn <- numeric(np)
q0 <- numeric(2^nt)
q1 <- numeric(2^nt)
p <- numeric(np)
cell <- rep(1:2^nt, time = np)

if (!is.null(count)) {
  comptage <- tapply(count, list(pop, cell), sum)
} else {
  comptage <- table(pop, cell)
}
q0[2:2^nt] = 0
q0[1] = sum(count[cell == 1])/sum(count)
q1[1:2^nt - 1] = 0
q1[2^nt] = sum(count[cell == 2^nt])/sum(count)
nn <- tapply(count, pop, sum)
p <- tapply(count[cell == 2^nt], pop[cell == 2^nt], sum)/nn
dd <- tire_simplexe(np = np, nt, nb_try = nb_try)

if (is.null(model) == T & nt == 2) {
  models <- expand.grid(c("DEP", "IND"), c("DEP", "IND"))
  models$nom <- apply(models, 1, function(x){
    paste(c("Se", x[1], "Sp", x[2]), collapse = "_")
  })
  modelsl <- split(models, seq(nrow(models)))
} else if (is.null(model) == T & nt == 3) {
  models <- expand.grid(c(1:nt, "DEP", "IND"), c(1:nt, "DEP", "IND"))
  models$nom <- apply(models, 1, function(x){
    paste(c("Se", x[1], "Sp", x[2]), collapse = "_")
  })
  modelsl <- split(models, seq(nrow(models)))
} else if (is.data.frame(model) == T) {
  models <- model
  models$nom <- apply(models, 1, function(x){
    paste(c("Se", x[1], "Sp", x[2]), collapse = "_")
  })
  modelsl <- split(models, seq(nrow(models)))
}

cl <- makeCluster(detectCores() - 1)
clusterExport(cl, c("EM_algo", "modifq",
  "nn", "np", "comptage",
  "nt", "Test", "modelsl",
  "dd", "Spf"), envir = environment())
ff <- parLapply(cl, modelsl, function(x) {
  as.matrix(t(apply(dd, 1, function(y) {
    EM_algo(n_pop = np,
      comptage = comptage,
      n_test = nt,
      pe = y[1:np],
      q0 = y[(np + 1):(np + 2^nt)],
      q1 = y[(np + 1 + 2^nt):(length(y))],

```

```

        nne = nn,
        MSe = as.character(x[[1]]),
        MSp = as.character(x[[2]]),
        Test,
        Spf = Spf)
    })))
})

names(ff) <- models[, 3]
clusterExport(cl, c("ff"), envir = environment())
Se <- parLapply(cl, ff, function(x) {
  t(apply(x, 1, function(y) {
    apply(t(Test), 1, function(z) {
      c(sum(y[(np + 1 + 2^nt):length(y)][z == 1]))
    })
  })
})

Sp <- parLapply(cl, ff, function(x) {
  t(apply(x, 1, function(y) {
    apply(t(Test), 1, function(z) {
      c(sum(y[(np + 1):(np + 2^nt)][z == 0]))
    })
  })
})

SeSp <- Map(cbind, Se, Sp)
clusterExport(cl, c("SeSp"), envir = environment())
SeSp2 <- parLapply(cl, SeSp, function(x) cbind(x[, 1] + x[, nt + 1], x))
SeSp3 <- Map(cbind, SeSp2, ff)
clusterExport(cl, c("log_vrais", "SeSp3"), envir = environment())

ffv <- parLapply(cl, SeSp3, function(y) {
  apply(y, 1, function(x) {
    log_vrais(comptage = comptage,
              q0 = x[(2 * nt + 2 + np):(2 * nt + 1 + 2^nt + np)],
              q1 = x[(2 * nt + 2 + 2^nt + np):(length(x))],
              p = x[(2 * nt + 2):(2 * nt + 1 + np)])
  })
})

SeSp4 <- Map(cbind, SeSp3, ffv)
clusterExport(cl, c("SeSp4"), envir = environment())
fit <- parLapply(cl, SeSp4, function(y) {
  t(apply(y, 1, function(x) {
    as.vector(t(diag(nn) %*% ((x[(2 * (nt + 1)):(2 * (nt) + 1 + np)] %o%
x[(2 * nt + np + 2^nt + 2):(2 * nt + np + 2 * (2^nt) + 1)]) +
(1 - x[(2 * (nt + 1)):(2 * (nt) + 1 + np)]) %o%
x[(2 * (nt) + 2 + np):(2 * (nt) + 1 + np + 2^nt)])))
  })
})

clusterExport(cl, c("covdep"), envir = environment())
covcov <- parLapply(cl, SeSp3, function(x) {

```

```

t(apply(x, 1, function(y) {
  covdep(Se = y[2:(1 + nt)],
        Sp = y[(nt + 2):(2 * nt + 1)],
        q0 = y[(2 * nt + 2 + np):(2 * nt + 1 + 2^nt + np)],
        q1 = y[(2 * nt + 2 + 2^nt + np):(length(y))],
        nt = nt)
}))
})

paramSe <- lapply(models1, function(x) {
  if (x$Var1 %in% c(1:3)) {
    x$paramSe <- (2 * nt - 2)
  } else {
    if (x$Var1 == "IND") {
      x$paramSe <- (nt)
    } else {
      if (x$Var1 == "DEP") {
        x$paramSe <- ((2^nt) - 1)
      }
    }
  }
})

paramSp <- lapply(models1, function(x) {
  if (x$Var2 %in% c(1:3)) {
    paramSp <- (2 * nt - 2)
  } else {
    if (x$Var2 == "IND") {
      paramSp <- (nt)
    } else {
      if (x$Var2 == "DEP") {
        paramSp <- ((2^nt) - 1)
      }
    }
  }
})

models12 <- Map(cbind, paramSp, paramSe, models1)
models13 <- lapply(models12, function(x) as.numeric(np + x[1] + x[2]))
SeSp5 <- Map(cbind, SeSp4,
            fit, covcov,
            param = models13, paramSp = paramSp,
            paramSe = paramSe, models1)

results <- parLapply(cl, SeSp5, function(x) {
  list(SeSp = x[, 1],
       Se = x[(x[, 1] > 1), 2:(nt + 1)],
       Sp = x[(x[, 1] > 1), (nt + 2):(2 * nt + 1)],
       p = x[(x[, 1] > 1), (2 * nt + 2):(2 * nt + 1 + np)],
       q0 = x[(x[, 1] > 1), (2 * nt + 2 + np):(2 * nt + 1 + np + 2^nt)],
       q1 = x[(x[, 1] > 1), (2 * nt + 2 + np + 2^nt):(2 * nt + 1 + np + 2 * (2^nt))],
       loglike = x[(x[, 1] > 1), (2 * nt + 2 + np + 2 * (2^nt))],
       fitted = x[(x[, 1] > 1),
                  (2 * nt + 3 + np + 2 * (2^nt)):(2 * nt + 2 + np + 2 * (2^nt) + (2^nt) * np)],

```

```

CovSe = x[(x[, 1] > 1),
(1 + (2 * nt + np + 2 * (2^nt) + 2 + (np * 2^nt))):
(1 + (2 * nt + np + 2 * (2^nt) + 2 + (np * 2^nt)) + choose(nt, 2) - 1)],
CovSp = x[(x[, 1] > 1),
(1 + (2 * nt + np + 2 * (2^nt) + 2 + (np * 2^nt)) + choose(nt, 2)):
(1 + (2 * nt + np + 2 * (2^nt) + 2 + (np * 2^nt)) + 2 * choose(nt, 2) - 1)],
param = x[1, (1 + (2 * nt + np + 2 * (2^nt) + 2 + (np * 2^nt)) + 2 * choose(nt, 2))],
maxLL = max(x[(x[, 1] > 1), (2 * nt + 2 + np + 2 * (2^nt))]),
AIC = -2 * (max(x[(x[, 1] > 1), (2 * nt + 2 + np + 2 * (2^nt))]))
+ 2 * x[1, (4 * nt + 3 + np + 2 * (2^nt) + (2^nt) * np)],
MSE = x[1, (1 + (2 * nt + np + 2 * (2^nt) + 2 + (np * 2^nt)) + 2 * choose(nt, 2) + 3)],
MSP = x[1, (1 + (2 * nt + np + 2 * (2^nt) + 2 + (np * 2^nt)) + 2 * choose(nt, 2) + 4)])
})

ident <- sapply((1:length(results)), function(y) {
  abs(max(apply(results[[y]]$Se, 2, function(x) max(x) - min(x)),
    apply(results[[y]]$Sp, 2, function(x) max(x) - min(x)),
    apply(results[[y]]$p, 2, function(x) max(x) - min(x))))
}) < 10^-4

modid <- results[ident]

syn <- data.frame(t(sapply(1:length(modid), function(x) {
  (c(apply(modid[[x]]$Se, 2, mean),
    apply(modid[[x]]$Sp, 2, mean),
    apply(modid[[x]]$CovSe, 2, mean),
    apply(modid[[x]]$CovSp, 2, mean),
    apply(modid[[x]]$p, 2, mean),
    modid[[x]]$maxLL,
    modid[[x]]$AIC)
})))

nomc <- c(apply(combn(1:nt, 2), 2, function(x) paste("T", x[1], "T", x[2], sep = "")))

colnames(syn) <- c(paste0("Se_", 1:nt),
  paste0("Sp_", 1:nt),
  paste0("Se_", nomc),
  paste0("Sp_", nomc),
  paste0("pop_", 1:np),
  "LL", "AIC")
rownames(syn) <- models$nom[ident]
syn <- syn[order(syn$LL, decreasing = T), ]

maxLL <- do.call(cbind, do.call(c, lapply(results, function(x) (LL = (x["maxLL"]))))))
modcomp <- lapply(results, function(x){
  cbind(maxLL = x["maxLL"], AIC = x["AIC"], param = x["param"])
})
modcomp <- do.call(rbind, modcomp)
rownames(modcomp) <- models$nom
modcomp <- cbind(unlist(modcomp[, 1]),
  unlist(modcomp[, 2]),
  unlist(modcomp[, 3]))
modcomp <- as.data.frame(modcomp)

```

```

names(modcomp) <- c("maxLL", "AIC", "param")
modcomp$AICc <- modcomp$AIC +
  ((2 * modcomp$param * (modcomp$param + 1))/(np - modcomp$param - 1))
modcomp <- modcomp[order((modcomp$AICc), decreasing = F), ]

pchi <- do.call(rbind, lapply(maxLL, function(x){
  pchisq((x) - maxLL, 2 * (2^nt - 1 - nt), lower.tail = FALSE)
}))
stopCluster(cl)
results$init <- list(tirage = dd, count = count)
results$test <- list(chi2 = pchi, AIC = modcomp)

if (full == F) {
  syn
} else results
}

```

Reproducible example

Data

Tests resultst for the 4 populations 3 tests study presented in the article

```

count<-c(0,9,0,6,1,9,0,262,
         2,11,1,5,3,6,4,267,
         5,14,2,9,6,7,37,367,
         8,10,2,1,1,1,33,108)

```

Call

Call of the freak function. The computation might take long.

```

dffreak<-freak(np=4,
               count=count,
               nb_try=100,
               model=NULL,
               nt=3,
               Spf=NULL,
               full=F)

```

Results

Results obtained, they may vary due to the relative small number of initial values. Discrepancies are resolved when sampling a higher number of initial values.

	Se_1	Se_2	Se_3	Sp_1	Sp_2	Sp_3
Se_3_Sp_DEP	0.197	0.165	0.546	0.950	0.937	0.988
Se_3_Sp_3	0.197	0.249	0.849	0.944	0.942	1.000
Se_3_Sp_1	0.203	0.177	0.281	0.981	0.961	0.988
Se_1_Sp_3	0.190	0.248	1.000	0.940	0.939	1.000
Se_2_Sp_3	0.190	0.248	1.000	0.940	0.939	1.000
Se_IND_Sp_3	0.190	0.248	1.000	0.940	0.939	1.000
Se_3_Sp_2	0.215	0.245	0.333	0.980	0.981	0.999
Se_3_Sp_IND	0.220	0.245	0.337	0.982	0.981	1.000

	Se_T1T2	Se_T1T3	Se_T2T3	Sp_T1T2	Sp_T1T3	Sp_T2T3
Se_3_Sp_DEP	0.130	0	0	0.027	0.001	0.010
Se_3_Sp_3	0.102	0	0	0.034	0.000	0.000
Se_3_Sp_1	0.136	0	0	0.000	0.000	0.012
Se_1_Sp_3	0.000	0	0	0.037	0.000	0.000
Se_2_Sp_3	0.000	0	0	0.037	0.000	0.000
Se_IND_Sp_3	0.000	0	0	0.037	0.000	0.000
Se_3_Sp_2	0.134	0	0	0.000	0.001	0.000
Se_3_Sp_IND	0.133	0	0	0.000	0.000	0.000

	pop_1	pop_2	pop_3	pop_4	LL	AIC
Se_3_Sp_DEP	0.000	0.040	0.176	0.488	-788.213	1606.426
Se_3_Sp_3	0.004	0.039	0.131	0.318	-794.927	1613.855
Se_3_Sp_1	0.074	0.147	0.320	0.785	-805.011	1634.022
Se_1_Sp_3	0.003	0.033	0.112	0.268	-809.783	1643.566
Se_2_Sp_3	0.003	0.033	0.112	0.268	-809.783	1643.566
Se_IND_Sp_3	0.003	0.033	0.112	0.268	-809.783	1641.566
Se_3_Sp_2	0.077	0.149	0.304	0.673	-811.675	1647.349
Se_3_Sp_IND	0.077	0.150	0.304	0.667	-811.731	1645.462

Appendix C R code of the additional needed functions

bootstrap.freak

Description

Bootstrap estimation of confidence interval for parameters estimates from an identifiable model.

Arguments

freaklist: list: a list obtained from the freak function for an identifiable model.
nb_simul: numeric : number of simulated dataset.
nb_pop: numeric: number of populations in the study.
nb_test: numeric: number of tests under evaluation.
nb_try: numeric: number of starting values for the EM algorithm.
count: vector: number of results for each combination (as described in the article) of tests results in each population.

Values

Returns a list containing parameters estimate for each run of the bootstrap.

```
bootstrap.freak<-function(freaklist,nb_simul,nb_pop,nb_test,nb_try,count)
{
  pop<-rep(1:nb_pop,each=2^nb_test)
  cell<-rep(1:2^nb_test,time=nb_pop)
  nn<-tapply(count,list(pop,cell),sum)
  nn<-tapply(count,pop,sum)

  Test<-as.vector(as.matrix(expand.grid(as.data.frame(
    matrix(rep(c(0,1),time=nb_test),ncol=nb_test))))))
  Test<-matrix(data=Test[length(Test):1],ncol=nb_test)

  qqp<-cbind(freaklist$q0[freaklist$loglike==max(freaklist$loglike),][1,],
             freaklist$q1[freaklist$loglike==max(freaklist$loglike),][1,],
             freaklist$p[freaklist$loglike==max(freaklist$loglike),][1,])
  if(is.vector(qqp)==F){
    qqp<-qqp[1,]
  } else {
    qqp<-qqp
  }
  ll<-apply(qqp,1,function(x) {
    count=simule_data(q0=x[1:(2^nb_test)],
                     q1=x[((2^nb_test))+1):(2*(2^nb_test))],
                     p=x[(2*(2^nb_test)+1):dim(qqp)[2]],
                     nb_simu=nb_simul,nn)}

  dd<-tire_simplexe(nb_pop=nb_pop,nb_test,nb_try= nb_try)
  cl<-makeCluster(detectCores() - 1)
  clusterExport(cl, c("estim","modifq",
                    "nn","nb_pop",
                    "nb_test","Test",
                    "dd","ll","freaklist"),
               envir=environment())
}
```

```

ff<-parLapply(cl,ll,function(x){
  lapply(x,function(y){
    list(estim=t(apply(dd,1,function(z) {
      estim(n_pop=nb_pop,
            comptage=y$count,
            n_test=nb_test,
            pe=z[1:nb_pop],
            q0=z[(nb_pop+1):(nb_pop+2^nb_test)],
            q1=z[(nb_pop+1+2^nb_test):(length(z))],
            nne=nn,
            MSe=freaklist$MSE,
            MSp=freaklist$MSP,
            Test,
            Spf=F)
      })),count=y$count)
    })
  })

ff2<-lapply(ff,function(x){
  lapply(x,function(y){
    cbind(y$estim,apply(y$estim,1,function(z){
      log_vrais(comptage=y$count,
                q0=z[(nb_pop+1):(nb_pop+2^nb_test)],
                q1=z[(nb_pop+1+2^nb_test):(length(z))],
                p=z[1:nb_pop])
      })))
    })
  })

ff3<-do.call(rbind,lapply(ff2,function(x) do.call(rbind,x)))
Se<-t(apply(ff3,1,function(x){
  apply (t(Test),1,function(y){
    c(sum(x[(nb_pop+1+2^nb_test):(length(x)-1)][y==1]))
  })
}))

Sp<-t(apply(ff3,1,function(x){
  apply((t(Test)),1,function(y){
    c(sum(x[(nb_pop+1):(nb_pop+2^nb_test)][y==0]))
  })
}))

p<-ff3[,1:nb_pop]
SeSp<-cbind(Se,Sp,Se[,1]+Sp[,1],ff3[, (2^(nb_test+1)+nb_pop+1)])
results<-list(SeSp_bt=SeSp,
              prev=p,
              init_bt=ll,
              misc_bt=ff3,
              MSe=freaklist$MSE,
              MSp=freaklist$MSP)

results
}

```

covdep

Description

Computes the calculation of covariances terms

Arguments

Se: sensitivities estimates.
Sp: specificities estimates.
q0: multinomial parameters for healthy animals.
q1: multinomial parameters for infected animals.
nt: number of tests.

Values

Returns a vector of estimated covariances terms for sensitivities and specificities.

```
covdep<- function(Se,Sp,q0,q1,nt)
{
  Test<-as.vector(as.matrix(expand.grid(as.data.frame(matrix(rep(c(0,1),time=nt),ncol=nt))))))
  Test<-matrix(data=Test[length(Test):1],ncol=nt)
  liste_combinaison<-(combn(1:nt,2))

  CovSe<-apply(apply(liste_combinaison,2,function(x){
    apply(Test,1,function(y) sum(y[x])==2)
  })*q1,2,sum)-apply(liste_combinaison,2,function(x){
    Se[x[1]]*Se[x[2]]
  })

  CovSp<-apply(apply(liste_combinaison,2,function(x){
    apply(Test,1,function(y) sum(y[x])==0)
  })*q0,2,sum)-apply(liste_combinaison,2,function(x) {
    Sp[x[1]]*Sp[x[2]]
  })

  nom<-c(apply(liste_combinaison,2,function(x) {
    paste("T",x[1],"T",x[2],sep="")
  }
  ))
  names(CovSe)<-nom
  names(CovSp)<-nom
  result<-c(CovSe,CovSp)

  return(result)
}
```

EM_algo

Description

Computes the Expectation-Maximisation algorithm.

Arguments

n_pop:	number of population.
comptage:	tests results.
n_test:	number of tests.
pe:	prevalence initial draw.
q0:	multinomial parameters for healthy animals initial draw.
q1:	multinomial parameters for diseased animals initial draw.
ne:	number of animal by population.
MSe:	covariance model for sensitivity.
MSp:	covariance model for specificity.
Test:	matrix of tests results.
Spf:	test with a fixed Sp of 1.

Values

Returns a vector of estimated parameters when the EM algorithm reach convergence.

```
EM_algo<-function(n_pop,comptage,n_test,pe,q0,q1,nne,MSe,MSp,Test,Spf)
{
  b <- numeric(2^n_test)
  c <- numeric(2^n_test)
  new_q0<- numeric(2^n_test)
  new_q1<- numeric(2^n_test)
  new_p<- numeric(n_pop)
  kk<-1
  converge <- F

  Test1<-as.vector(as.matrix(expand.grid(as.data.frame(
    matrix(rep(c(0,1),time=(n_test-1)),ncol=(n_test-1))))))
  Test1<-matrix(data=Test1[length(Test1):1],ncol=(n_test-1))

  if (is.numeric(Spf)==T){
    if (n_test==2){
      q0[Test[,Spf]==0]<-sapply(Test1, function(y){
        sum(q0[lapply(Test[,~Spf],function(x)
          paste0(x,collapse=""))==paste0(y,collapse="")])
      })
      q0[Test[,Spf]==1]<-0
    } else {
      q0[Test[,Spf]==0]<-apply(Test1,1, function(y){
        sum(q0[apply(Test[,~Spf],1,function(x)
          paste0(x,collapse=""))==paste0(y,collapse="")])
      })
      q0[Test[,Spf]==1]<-0
    }
  }
}
```

```

}
}

if (MSe %in% c(1:3)){
  q1<-modifq(q=q1,tt=as.numeric(MSe),Test)
} else {
  if (MSe=="IND") {
    Se<-(c(apply(t(Test),1,function(y) c(sum(q1[y==1])))))
    q1<-(apply(Se*t(Test)+(1-Se)*t(1-Test),2,prod))
  }
  else {
    if (MSe=="DEP") {
      q1<-q1
    }
  }
}

if (MSp %in% c(1:3)){
  q0<-modifq(q=q0,tt=as.numeric(MSp),Test)
} else {
  if (MSp=="IND") {
    Sp<-(c(apply(t(Test),1,function(y) c(sum(q0[y==0])))))
    q0<-(apply(Sp*t(1-Test)+(1-Sp)*t(Test),2,prod))
  }
  else {
    if (MSp=="DEP") {
      q0<-q0
    }
  }
}

while(!converge) {
  a<-((1.0-pe)%*%t(q0))/(pe%*%t(q1))
  a<-1.0/(1.0+a)
  a[a>1.0]=1.0
  a[a<0.0]=0.0
  a[q1==0.0]=0.0
  a[(a=="NaN")]=0.0
  aa=1-a
  b<-apply(a*comptage,2,sum)
  c<-apply(aa*comptage,2,sum)
  new_q1<-b/sum(b)
  new_q0<-c/sum(c)

  new_p<-apply(a*comptage,1,sum)/nne

  if (is.numeric(Spf)==T){
    if (nt==2){
      new_q0[Test[,Spf]==0]<-sapply(Test1, function(y){
        sum(new_q0[lapply(Test[, -Spf], function(x)
          paste0(x,collapse=""))]==paste0(y,collapse=")"])
      })
      new_q0[Test[,Spf]==1]<-0
    } else {

```

```

new_q0[Test[,Spf]==0]<-apply(Test1,1, function(y){
  sum(new_q0[apply(Test[, -Spf], 1, function(x)
    paste0(x, collapse=""))]==paste0(y, collapse="")])
})
new_q0[Test[,Spf]==1]<-0
}
}
  if (MSe %in% c(1:3)){
new_q1<-modifq(q=new_q1,tt=as.numeric(MSe),Test)
} else {
  if (MSe=="IND") {
    Se<-(c(apply(t(Test), 1, function(y) c(sum(new_q1[y==1])))))
    new_q1<-(apply(Se*t(Test)+(1-Se)*t(1-Test), 2, prod))
  } else {
    if (MSe=="DEP") {
      new_q1<-new_q1
    }
  }
}
}
if (MSp %in% c(1:3)){
new_q0<-modifq(q=new_q0,tt=as.numeric(MSp),Test)
} else {
  if (MSp=="IND") {
    Sp<-(c(apply(t(Test), 1, function(y) c(sum(new_q0[y==0])))))
    new_q0<-(apply(Sp*t(1-Test)+(1-Sp)*t(Test), 2, prod))
  } else {
    if (MSp=="DEP") {
      new_q0<-new_q0
    }
  }
}
}
converge=((max(abs(pe-new_p))<1e-8)&(max(abs(q0-new_q0))<1e-8)&(max(abs(q1-new_q1))<1e-8))
pe<-new_p
q0<-new_q0
q1<-new_q1
if (is.na(converge)|(kk>50000))
{
  converge=T
  pe=rep(-999,np)
  q0=rep(-999,2^nt)
  q1=rep(-999,2^nt)
}
kk<-kk+1
}
results<-c(pe,q0,q1)
results
}

```

log_vrais

Description

Computes log-likelihood calculation.

Arguments

- comptage:** vector: number of results for each combination (as described in the article) of tests results in each population.
- q0:** vector: multinomial parameters estimated for healthy animals.
- q1:** vector: multinomial parameters estimated for infected animals.
- p:** vector: prevalence estimates for each population.

Values

Returns a matrix containing initial values for prevalence, and parameters for the 2 distributions (healthy and diseased).

```
log_vrais<-function(comptage,q0,q1,p)
{
  u=log(p%%t(q1)+(1-p)%%t(q0))
  u[(u=="-Inf")]=-1000
  return (sum(u*comptage))
}
```

modifq

Description

Modifies distributions parameters when estimating a 3-tests-model with one of the test independent from the 2 others.

Arguments

- q** : vector: multinomial parameters.
- tt** : test independent from the others (from 1 to 3).
- Test** : matrix of tests results.

Values

Returns a modified vector of parameters to force a correlation structure.

```
modifq<-function(q,tt,Test)
{
  q_1<-apply(apply(Test[,-tt],1,function(x){
    apply(Test[,-tt],1,function(y) identical(y,x)==T)}
  )*q,2,sum)
  qq1<-Test[,tt]*sum(q[Test[,tt]==1])+(1-Test[,tt])*sum(q[Test[,tt]==0])
  q1i2<-(qq1*q_1)
  q1i2
}
```

simule_data

Description

Simulates data based on estimated parameters from a freaklist.

Arguments

q0: vector: multinomial parameters for healthy animals.
q1: vector: multinomial parameters for infected animals.
nn: vector: number of animal in each population
nb_simu: numeric: number of bootstrap run

Values

returns a vector of tests results to perform an EM_algorithm procedure

```
simule_data<- function(q0,q1,p,nn,nb_simu=1000)
{
  l1<-list()
  length(l1)<-nb_simu
  nnp<-rbind(nn,p)
  l12<-lapply(l1,function(x){
    x<-list(count=t(apply(nnp,2,function(y){
      x=(rmultinom(1,round(y[1]*y[2]),prob=q1)+rmultinom(1,(y[1]-round(y[1]*y[2])),prob=q0))
    })),q0=q0,q1=q1,p=p,nn=nn)
  })
  l12
}
```

tire_simplexe

Description

Randomly samples values in a N-simplex and in uniform distribution for prevalences.

Arguments

np: number of population in the study.
nt: number of tests under evaluation.
nb_try: number of initial value to sample.

Values

Returns a matrix containing initial values for prevalence, and parameters for the 2 distributions (healthy and diseased).

```

tire_simplexe<-function(np,nt,nb_try)
{
  p<-matrix(nrow=nb_try,ncol=np,runif(nb_try*np))
  q0=matrix(nrow=nb_try,ncol=2^nt)
  q1=matrix(nrow=nb_try,ncol=2^nt)

  qq0<-matrix(nrow=nb_try,ncol=2^nt-1,runif(nb_try*(2^nt-1)))
  qq0<-t(apply(t(qq0),2, sort))
  qq1<-matrix(nrow=nb_try,ncol=2^nt-1,runif(nb_try*(2^nt-1)))
  qq1<-t(apply(t(qq1),2, sort))

  sous<-matrix(nrow=2,ncol=2^nt-1,byrow=F,c(1,1,1,rep(2:(2^nt-2),each=2),2^nt-1))
  q0<-abs(qq0[,sous[1,]]-qq0[,sous[2,]])
  q0<-cbind(qq0[,1],q0[,-1],1-qq0[,2^nt-1])
  q1<-abs(qq1[,sous[1,]]-qq1[,sous[2,]])
  q1<-cbind(qq1[,1],q1[,-1],1-qq1[,2^nt-1])

  ll<-list(p=p,q0=q0,q1=q1)
  cbind(ll$p,ll$q0,ll$q1)
}

```

- [1] Albert, P. S. and Dodd, L. E. (2004). A Cautionary Note on the Robustness of Latent Class Models for Estimating Diagnostic Error without a Gold Standard. *Biometrics*, 60(2):427–435.
- [2] Allman, E. S., Matias, C., and Rhodes, J. A. (2009). Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics*, 37(6A):3099–3132. arXiv: 0809.5032.
- [3] Berkvens, D., Speybroeck, N., Praet, N., Adel, A., and Lesaffre, E. (2006). Estimating disease prevalence in a Bayesian framework using probabilistic constraints. *Epidemiology (Cambridge, Mass.)*, 17(2):145–153.
- [4] Branscum, A. J., Gardner, I. A., and Johnson, W. O. (2005). Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Preventive Veterinary Medicine*, 68(2–4):145–163.
- [5] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1976). Maximum Likelihood from Incomplete Data via the EM Algorithm.
- [6] Dendukuri, N. and Joseph, L. (2001). Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*, 57(1):158–167.
- [7] Frössling, J., Bonnett, B., Lindberg, A., and Björkman, C. (2003). Validation of a *Neospora caninum* iscom ELISA without a gold standard. *Preventive Veterinary Medicine*, 57(3):141–153.
- [8] Georgiadis, M. P., Johnson, W. O., Gardner, I. A., and Singh, R. (2003). Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(1):63–76.
- [9] Hui, S. L. and Walter, S. D. (1980). Estimating the Error Rates of Diagnostic Tests. *Biometrics*, 36(1):167.
- [10] Jones, G., Johnson, W. O., Hanson, T. E., and Christensen, R. (2010). Identifiability of Models for Multiple Diagnostic Testing in the Absence of a Gold Standard. *Biometrics*, 66(3):855–863.
- [11] Mathevon, Y., Foucras, G., Falguières, R., and Corbiere, F. (2017). Estimation of the sensitivity and specificity of two serum ELISAs and one fecal qPCR for diagnosis of paratuberculosis in sub-clinically infected young-adult French sheep using latent class Bayesian modeling. *BMC Veterinary Research*, 13:230.
- [12] Pepe, M. S. and Janes, H. (2007). Insights into latent class analysis of diagnostic test performance. *Biostatistics*, 8(2):474–484.
- [13] Qu, Y., Tan, M., and Kutner, M. H. (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostics tests. *Biometrics*, 52(3):797–810.
- [14] Rutjes, A. W. S., Reitsma, J. B., Coomarasamy, A., Khan, K. S., and Bossuyt, P. M. M. (2007). Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technology Assessment (Winchester, England)*, 11(50):iii, ix–51.

- [15] Shih, J. H. and Albert, P. S. (1999). Latent model for correlated binary data with diagnostic error. *Biometrics*, 55(4):1232–1235.
- [16] Toft, N., Innocent, G. T., Gettinby, G., and Reid, S. W. (2007). Assessing the convergence of Markov Chain Monte Carlo methods: An example from evaluation of diagnostic tests in absence of a gold standard. *Preventive Veterinary Medicine*, 79(2-4):244–256.
- [17] Vacek, P. M. (1985). The Effect of Conditional Dependence on the Evaluation of Diagnostic Tests. *Biometrics*, 41(4):959–968.

3 Conclusions du chapitre

Concernant les méthodes statistiques d'évaluation des performances diagnostiques des tests en l'absence de test de référence parfait :

- L'évaluation des performances diagnostiques des tests, pour produire des résultats comparables, doit être basée sur une définition précise de la cible recherchée;
- Les modèles basés sur différentes structures de corrélation entre tests doivent être évalués et sélectionnés selon des critères statistiques objectifs (vraisemblance, AIC, BIC/DIC...);
- La méthode d'estimation fréquentiste en classes latentes que nous proposons fournit une alternative pour l'utilisateur non rompu à l'estimation bayésienne en permettant d'évaluer rapidement l'identifiabilité du modèle ainsi qu'une hiérarchisation des modèles basés sur différentes structures de corrélation entre tests.

Concernant l'évaluation des performances diagnostiques des tests de dépistage de la paratuberculose à l'échelle individuelle :

- La Se_D à l'échelle individuelle des trois tests étudiés (deux trousse de sérologie ELISA et une trousse de qPCR sur fèces) est mauvaise à médiocre pour le dépistage des ovins infectés au stade subclinique de la maladie;
- Pour la sérologie ELISA, les résultats obtenus sont globalement en accord avec ceux décrits dans la littérature pour les animaux infectés subcliniques, avec une Sp_D loin d'être parfaite;
- Nos résultats concernant la qPCR sont, à notre connaissance, les premiers présentés selon les recommandations faites pour la standardisation des rapports d'évaluation des performances diagnostiques des tests vis-à-vis de la paratuberculose faisant appel à des modèles à classes latentes.

	ELISA A		ELISA B		fqPCR	
	Moy ¹	IC ²	Moy	IC	Moy	IC
Se	18%	[11%-25%]	17%	[10%-25%]	47%	[29%-70%]
Sp	95%	[93%-96%]	94%	[92%-96%]	99%	[98%-100%]

ELISA A : ID Screen Paratuberculosis Indirect[®], lot 602, IdVet; ELISA B : IDEXX paratuberculosis screening[®], lot 5074; fqPCR : ADIAVET[™] paraTB Real Time, BioX; ¹ : moyenne du paramètre estimé; ² : IC 95%/PCI 95% du paramètre estimé.

Chapitre 4

Stratégies de dépistage à l'échelle collective

Bien sûr, certains de leurs animaux étaient eux aussi des contrefaçons électroniques. Mais il n'était bien sûr jamais allé y mettre le nez, pas plus que ses voisins n'étaient venus voir de près la nature véritable de son mouton.

Philip K. Dick

Les androïdes rêvent-ils de moutons électriques?

Le chapitre précédent nous a permis d'obtenir une évaluation des performances diagnostiques des tests appliqués à l'échelle individuelle pour le diagnostic de la paratuberculose chez des ovins infectés à la phase subclinique. L'évaluation des performances diagnostiques des analyses de mélanges d'échantillons pour le dépistage et le suivi de la paratuberculose à l'échelle collective a été réalisée en deux temps. La première partie est une étude expérimentale basée, aussi bien pour la qPCR sur fèces que pour les deux trousseaux de sérologie ELISA, sur des dilutions d'échantillons de statut connu dans une matrice négative. L'objectif de cette première étape était de définir les performances diagnostiques relatives des analyses de mélanges permettant dans un deuxième temps de réaliser une évaluation, par simulations, de différentes stratégies de dépistage.

Les résultats obtenus font l'objet d'un article en cours de finalisation. Nous fournis-

sons aux lecteurs des informations complémentaires sur les expérimentations réalisées après la présentation de cet article.

1 Article 3 : évaluation de la sensibilité et de la spécificité de la sérologie ELISA et de la qPCR sur fèces appliquées à des mélanges d'échantillons pour le dépistage à l'échelle des troupeaux de la paratuberculose ovine en France.

Résumé de l'article. En combinant les résultats de l'évaluation des performances diagnostiques à l'échelle individuelle de la sérologie ELISA et de la qPCR sur fèces, vus au chapitre 3, et ceux obtenus sur des analyses de mélanges, une étude par simulations a été réalisée pour évaluer les performances diagnostiques de différentes stratégies de dépistage de la paratuberculose à l'échelle des troupeaux.

L'analyse de mélanges expérimentaux de fèces a permis d'estimer la Se_D relative de la qPCR ADIAVET™ paraTB Real Time (BioX). Des mélanges de 5 (195 mélanges) et 10 échantillons (228 mélanges) ont été mimés par dilution de 3 ou 10 grammes de fèces positives dans des fèces négatives. Les fèces positives étaient soit des échantillons ayant un Ct très précoce en qPCR individuelle ($Ct \leq 30$; 6 animaux), soit des échantillons ayant un Ct très tardif (Ct compris entre 32 à 38; 9 animaux). Les échantillons ayant un Ct initial précoce ($Ct \leq 30$) demeuraient tous détectables en mélanges de taille 5 ou 10, quelle que soit la prise d'échantillon utilisée pour réaliser les mélanges (3 ou 10 grammes), suggérant une Se_D relative parfaite pour les fèces issus d'animaux forts excréteurs. A l'inverse seuls 87% à 90% des échantillons issus d'animaux faibles excréteurs ($Ct \text{ initial} \geq 30$) étaient détectés en mélanges de taille 5, et 63% à 70% en mélanges de taille 10, sans influence notable de la prise d'échantillon utilisée pour réaliser les mélanges.

L'analyse par sérologie ELISA (trousse IDEXX Paratuberculosis Screening®) de 362 mélanges de 5 ou 10 sérums de valeurs S/P individuelles variables a permis de définir de nouveaux seuils de décision (respectivement $S/P > 25\%$ et $S/P > 15\%$ pour les mélanges de 5 et 10 échantillons) associés à une Sp_D parfaite quelle que soit la composition des mé-

langes. A l'inverse, aux seuils sélectionnés, la Se_D relative dépendait du nombre d'échantillons individuels positifs composant les mélanges et de leurs valeurs S/P initiales, avec des sensibilités diagnostiques variant d'environ 62% (mélanges contenant un seul échantillon faiblement positif) à 100% (mélanges contenant au moins un échantillon fortement positif ou deux échantillons faiblement positifs).

Cet article présente par ailleurs les distributions des résultats quantitatifs en sérologie ELISA (valeurs S/P) et qPCR sur fèces (Ct) observées chez les 1197 animaux prélevés dans 14 élevages infectés dans le cadre de l'étude transversale présentée au chapitre 3, et servant de base de modélisation du présent article. Les résultats indiquent que les animaux fortement excréteurs ($Ct \leq 30$) sont très minoritaires (moins de 6%) parmi les animaux détectés par qPCR. De même, les animaux séropositifs ayant des valeurs S/P élevées ($>90\%$) sont peu nombreux (14% des animaux non-négatifs). Ces distributions ont aussi été obtenues pour deux élevages indemnes de paratuberculose, permettant de modéliser l'utilisation des analyses de mélanges d'échantillons dans des troupeaux sains.

Les études par simulations ont permis de tenir compte de l'ensemble des paramètres de variation (performances diagnostiques à l'échelle individuelle et des mélanges, distributions des résultats quantitatifs individuels) et de l'incertitude associée à chacun d'entre eux. Les simulations ont été réalisées pour des prévalences d'infection variant de 1 à 30%, pour un nombre d'animaux prélevés compris entre 50 et la totalité de l'effectif présent (soit 300 animaux). Pour chaque combinaison de prévalence d'infection, 1000 élevages ont été simulés, chacun d'entre eux étant soumis 1000 fois à chaque stratégie de dépistage. Des dépistages réalisés en troupeaux indemnes ont aussi été simulés afin d'évaluer la Sp_T des différentes stratégies envisagées. Un élevage était considéré comme détecté dès lors qu'au moins un mélange donnait un résultat positif. Le nombre de mélanges positifs était par ailleurs enregistré pour chaque élevage et chaque stratégie de dépistage simulée.

Concernant les paramètres de modélisation des performances intrinsèques à l'échelle individuelle de la sérologie ELISA et de la qPCR sur fèces, deux scénarios ont été envisagés. Le premier reposait sur les estimateurs obtenus dans le premier article publié dans ce travail de thèse (voir chapitre 3), notamment caractérisés par une Sp_D non parfaite des tests dans les troupeaux infectés. Dans une seconde approche, des valeurs plus optimistes

de Sp_D ont été fixées, semblables à celles obtenues dans les troupeaux non infectés (Sp_D de la qPCR : 100,0%; Sp_D de la sérologie ELISA : 99,0%). La comparaison de ces deux approches a permis d'évaluer l'impact de la Sp_D non parfaite à l'échelle individuelle sur la Sp_D et la Se_D à l'échelle des troupeaux.

Les résultats des simulations suggèrent que l'utilisation de la sérologie ELISA sur des mélanges de sérums dans le cadre d'un dépistage de la paratuberculose conduirait à des résultats faussement positifs dans une trop grande proportion d'élevages indemnes. Ce défaut de spécificité est très largement lié à une spécificité non parfaite à l'échelle individuelle. La sérologie ELISA appliquée à des mélanges de sérums n'est donc conseillée que dans le cadre d'une suspicion de prévalence d'infection élevée pour laquelle le nombre de mélanges positifs apportera une réponse sans ambiguïté. L'utilisation de la qPCR sur des mélanges de fèces semble plus prometteuse pour identifier de manière plus spécifique les troupeaux infectés, mais nécessiterait un nombre élevé d'animaux prélevés (plus de 100) pour permettre la détection dans plus de 90% des cas, des élevages à prévalence d'infection inférieure à 5%.

Flock sensitivity and specificity of pooled fecal qPCR and pooled serum ELISA for screening ovine paratuberculosis

Yoann Mathevon¹, Gilles Foucras¹, Fabien Corbiere^{1*},

1 UMR INRA ENVT 1225 IHAP, Ecole Nationale Vétérinaire de Toulouse, 23 chemin des Capelles, F-31076 Toulouse Cedex 3, France

* Corresponding author E-mail: f.corbiere@envt.fr

Abstract

The aim of our study was to evaluate the flock sensitivity and specificity of fecal qPCR and serum ELISA using pooled samples for screening paratuberculosis in french sheep. Using individual feces with low (n=6) or high (n=9) qPCR Ct values from ewes sampled in 14 infected flocks, 195 pools of size 5 and 210 pools of size 10 were mimicked by diluting 3 or 10g of individual materials in negative feces and analysed using a commercial *IS900* qPCR kit. The relative sensitivity and specificity of pooled serum analysis was evaluated based on the analysis of 181 different pools of size 5 or 10, composed of individual serum samples of various S/P values. Results showed that for sample pool of size 5 or 10, individual fecal samples with low Ct values were invariably detected. Conversely fecal samples with high Ct values were associated with a lower detection rate in both pools of size 5 (87% to 90%) and 10 (63% à 70%). Regarding pooled serum ELISA analysis, the detection rates were highly dependant on the pool composition. Lowering the decision threshold to 25% and 15% for samples of size 5 and 10 respectively, allowed for still variable but efficient detection sensitivity and was associated with perfect specificity. Finally, based on these results and the distribution of individual test results in 14 infected and 2 paratuberculosis-free flocks, we conducted a simulation study in order to evaluate sensitivity and specificity of various detection strategies at flock levels under various assumptions regarding the diagnostic accuracy of tests at the individual levels. The use of pooled serum ELISA led to very high false positive detection rate in paratuberculosis free flocks and prevents its further use in that context. Conversely the use of pool fecal qPCR led to much more relevant detection strategies. However our results showed that a substantial number of individual sheep should be sampled (more than 100) to allow the detection of true infection prevalence lower than 5%.

Introduction

Paratuberculosis is the chronic infectious disease affecting the digestive tract of ruminants, caused by *Mycobacterium avium* subsp. *paratuberculosis* (*Map*). Correct classification and estimation of the prevalence in infected herds/flocks are important to provide guidance for control programs, including vaccination and/or restrictions on animal movements. Flock or herd-testing strategies based on individual milk or serum ELISA and fecal culture have been assessed in several studies but appear expensive (1 ; 2 ; 3). Furthermore, the lack of sensitivity of individual tests, especially for the detection of sub-clinically infected animals (4 ; 5 ; 6 ; 7), lead to inaccurate negative predictive value regarding the individual infectious status. In sheep production, the main objective of paratuberculosis surveillance programs is usually to prove freedom of disease or to estimate prevalence rather than to identify individual infectious animals.

Pool sample testing (PST), proposed initially for screening syphilis in US soldiers during World War II (8), has become increasingly popular to monitor paratuberculosis in herds/flocks as it allows substantial savings in laboratory costs (9 ; 10). For detection of *Map*, PST was originally applied to faecal culture by physically combining and mixing a number of individual fecal samples (11 ; 9 ; 12) and has been extended to

polymerase chain reaction (qPCR) testing (13 ; 14). In parallel bulk tank milk ELISA (15) or qPCR (16 ; 17) and environmental sampling, including manure pit (18), has been proposed, mostly in dairy cattle, as alternative testing approaches and can be regarded as extreme pooling strategies (19). However, whatever the approach, combining and mixing individual samples may lead to a drop in the concentration of *Map* or antibodies toward *Map* to a level that cannot be detected (20), making the sensitivity of the pooled-sample approach lower than the approach based on individual testing. Therefore when evaluating the diagnostic accuracy of pooled-sample based herd/flock-testing, the influence of the dilution effect on pool sensitivity is an essential pre-requisite. Furthermore depending on the surveillance purposes, the best testing strategy (i.e. number and pool size per herd/flock) may differ and should be evaluated. Finally interpretation of results are also challenging to estimate accurately within-herd prevalence, which is the main (only) indicator needed to trigger control measures by farmers. The analytical sensitivity of pooled-sample approach based on *Map* detection may vary according to sample quality, pooling and mixing methods, culture media (21), DNA extraction methods, DNA target and qPCR systems (22 ; 23 ; 24). Similarly, the accuracy of bulk tank milk antibody detection may depend on the ELISA kit used or to the applied decision threshold (15 ; 25). Finally it is unwise to simply extrapolate already published estimates to any other method. Most of all, the intensity levels of individual samples composing the pool (i.e. number of *Map* bacteria or antibody titers) may have a strong influence on the pool result (11 ; 26).

To our knowledge, the detection of antibody response toward *Map* based on pooled serum samples has not been published. This approach has however already been evaluated for other sheep or porcine diseases (27 ; 28) and showed that decision thresholds should be re-evaluated for different pool sizes to allow a satisfactory detection rate.

In this context, we aimed at evaluating the flock sensitivity and specificity of pooled fecal qPCR and pooled serum ELISA for screening ovine paratuberculosis. Based on the evaluation of the dilution effect on pool fecal or serum sensitivity and the distribution of individual results in infected and non-infected flocks, we conducted a simulation study at the flock levels in order to evaluate various detection strategies.

Materials and methods

Flocks, animals and biological samples

The diagnostic test results from a cross-sectional study were previously published by the study investigators (29). Briefly, 1197 individual serum and fecal samples were collected from 2- to 3-year-old sheep without clinical signs of paratuberculosis in 14 purebred Causse du Lot closed meat sheep flocks where paratuberculosis infection was enzootic. Fecal excretion was determined using qPCR based on *IS900* sequence detection, and serology was performed using two commercial ELISAs. Beyond the binary (positive/negative) test results that were analysed in this study, we also evaluated the distribution of the quantitative responses for serum ELISA (i.e sample to positive ratio, S/P) and fecal qPCR (i.e. cycle threshold, Ct) across the sampled population. For the particular aims of the present study, blood and fecal samples were also collected from 387 ewes in three flocks known to be free from paratuberculosis (closed purebred Lacaune flocks) using the same sampling methods.

Pooled-sample analysis

We hypothesized that pooled-sample analysis was sensitive to the values of individual samples that composed the pool, i.e. that the effect of dilution was less detrimental for pooled-sample sensitivity when individual samples gave strong positive results. For practical reasons, only a limited number of serum and fecal samples could be included in the experiment, and simplifications were made in order to assess the effect of dilution together with the assessment of uncertainty. The continuous test results were therefore discretized in different ordered response categories. Note that for both serum ELISA and fecal qPCR, these response categories did not rely on established biological knowledge but were rather driven by practical considerations related to the pooled-sample analysis and the design of the simulation study. For serum ELISA the decision threshold of $S/P > 45\%$ and $S/P > 55\%$ are defined by the manufacturer to distinguish between negative,

doubtful and positive results, respectively. Based on the first cut-off value, individual S/P results were arbitrarily separated in four response categories as follows : S/P < 22.5% : low negative (LN) ; [22.5% - 45.0%[: high negative (HN) ; [45.0% - 90%[: low positive (LP) and S/P ≥ 90% : high positive (HP). For fecal qPCR, two response categories were used, based on the distribution of Ct values of the samples recovered in the 14 investigated flocks infected with paratuberculosis. We arbitrarily classified samples with Ct ≤ 30 as “highly contaminated” (HC) with *Map* and those with the Ct > 30 as “lowly contaminated” (LC).

Pooled serum ELISA

Pools of size 5 and 10 were manually constituted in duplicate using 20 μl of each individual serum sample. Samples with individual S/P values below 22.5% were used as negative diluent. Several types of pools were created, based on the number of individual samples from the different response categories that were pooled together. To evaluate the relative sensitivity of pooled serum ELISA (PSe_{ELISA}), pools containing either one “high positive (HP)”, 1 “low positive (LP)” and 2 “low positive” individual sample(s) were investigated. To evaluate the relative specificity of pooled serum ELISA analysis (PSP_{ELISA}), we constructed pools containing either only “low negative (LN)” and 1 up to 3 “high negative (HN)” individual sample(s). Using 13 HP, 37 LP, 41 HN, and 156 LN unique individual samples, 181 pools were created for each pool size.

Pooled fecal qPCR

Feces from 6 individual ewes with Ct ≤ 30 (range 20.42 – 28.12) and 9 with Ct > 30 (range 34.8 – 41.7) from the investigated flocks were used to build the pools. Confirmed qPCR-negative fecal samples collected from uninfected animals in a research flock known to be free of paratuberculosis (closed purebred Lacaune flock) were used for creation of pools. Before pooling, positive feces from each individual ewe were mixed thoroughly. Three grams of each positive fecal sample were mixed manually in a sterile plastic bag until complete homogenization with either 12 or 27 grams of negative feces. The same was also done using 10 grams of positive feces and either 40 or 90 grams of negative diluent. Thus, the pools contained positive and negative feces in a ratio of 1:4 or 1:9 which is equivalent to creating pools of size 5 and 10. A total of 195 5-samples and 210 10-samples pools were created to evaluate the relative sensitivity of pooled fecal qPCR (PSe_{qPCR}). In addition, 40 pools containing only negative feces were constructed to evaluate the specificity of pooled fecal PCR (PSP_{qPCR}).

Laboratory testing

Serum ELISA

The same commercial serum ELISA (IDEXX paratuberculosis screening[®] kit, batch 5074, IDEXX, Montpellier, France) as in (29) was applied to individual or pooled serum samples using an overnight incubation protocol following the manufacturer’s instructions. Negative and positive controls provided by the manufacturer were included on each ELISA plate. After blocking, absorbance was recorded at 450 nm and 650 nm and sample to positive (S/P) ratio results were calculated for computation. Coefficients of variation from duplicate pools of serum were measured, and samples with values greater than 10% were regarded as indicative of incorrect manipulations during the pooling process and data exclude from the analysis.

Fecal qPCR

The same qPCR protocol targeting the IS900 DNA fragment of *Map* as in (29) was applied to individual and pool fecal samples, using 10 grams as working material, as previously described (29). Briefly samples underwent a concentration procedure using the ADIAFILTER system (BioX, Rochefort, Belgium) following the manufacturer’s instructions. DNA extraction was then performed on a Kingfisher Flex[®] magnetic particle processor (Thermo Fisher Scientific, Courtaboeuf, France) following the NucleoMag 96 tissue protocol (Macherey-Nagel, Hoerdt, France), with addition of an extraction control from the qPCR kit (ADIAVET[™] PARATB REAL TIME, BioX, Rochefort, Belgium) in each sample. Forty-five amplification cycles were performed on a LightCycler 480 (RocheLife Science, Meylan, France), and fluorescent signals

were recorded in two channels, corresponding to FAM for *IS900* and VIC for the extraction control, respectively. Due to the overlapping spectra of the two dyes, a color compensation step was applied. Raw fluorescence was analyzed using the qpcR package in R software and cycle threshold (Ct) were determined using second derivative maximum (CpD2). Samples with Ct below 42 were considered as positive.

Relative accuracy of pooled-sample analysis

The relative sensitivities of pooled serum ELISA (PS_{eELISA}) and pooled fecal qPCR (PS_{eqPCR}), were defined as the proportion of true positive pools (i.e. that actually contained at least one positive individual sample) that yielded a positive result. The relative specificities PS_{pELISA} and PS_{pPCR} were defined as the proportion of true negative pools (i.e. that actually contained only negative individual samples) that yielded a negative result. For serum ELISA, when determining whether pools are test positive or negative, it can be beneficial to establish alternative cut-offs that are lower than those recommended by the manufacturer. Indeed pooling method dilutes test-positive individual samples, causing a lower signal that requires a lower cut-off to be regarded as positive. However choosing very low cut-offs may lead to false positive results in true negative pools, especially when several HN samples are pooled together. Therefore ELISA (PS_{eELISA}) and PS_{pELISA} estimates were computed and compared for different decision thresholds ranging from $S/P \geq 10\%$ to $S/P \geq 45\%$. Note that here sensitivity and specificity estimates are relative to the individual analysis, as we evaluated the proportion of known positive and negative individual samples that were correctly recovered when mixed with negative diluent. These estimates are different from the absolute accuracy of pooled-sample analysis, which is its ability to correctly detect individual animals of a given infectious status (i.e non-infected, infected, infectious) (30). In fact due to the long and complex physiopathology of the disease, some infected or infectious individuals may not shed a detectable amount of *Map* in their feces or have mounted a detectable antibody response toward *Map* at the time of sampling. Conversely some truly uninfected animals may yield positive individual test results, for instance due to an imperfect analytical specificity or because the presence of passively shed MAP in their feces.

Performances of pooled-samples based detection at the flock level

A simulation study was carried out to estimate the flock-sensitivity (FPS_e) and flock-specificity (FPS_p) of pooled fecal qPCR and pooled serum ELISA in sheep in simulated flocks with a range of infection prevalence levels and under different sampling and pooling scenarios. The simulation program comprised 3 independent compartments, hereafter named the flock, individual and pool compartments. In the flock compartment the number of truly infected and truly uninfected sheep within a flock was determined, given the flock size and infection prevalence. The binary (positive/negative) serum ELISA and fecal qPCR results for each sheep within a flock were then modeled in the individual compartment, given its true infectious status and the sensitivity (IS_{eELISA} , IS_{eqPCR}) and specificity (IS_{pELISA} , IS_{pPCR}) of diagnostic tests applied at the individual levels. For infected flocks, these estimates were derived from (29) in which the diagnostic accuracy of the IDEXX serum ELISA and the ADIAVET fecal qPCR were estimated in sub-clinically infected sheep using Bayesian modeling. Interested readers are referred to (29) for further details about the estimating procedure. For paratuberculosis-free flocks, the accuracy of these diagnostic test at the individual level was computed based on results obtained in the present study. Given individual binary test results, each individual was then allocated to a serum ELISA (LN, HN, LP, HP) and fecal qPCR (LC, HC) response category, based on the probability distribution functions observed in the 14 flocks infected and the 3 paratuberculosis-free flocks investigated in the present study. Finally, random sampling of individual sheep, random creation of pools and pool analysis were modeled in the pool compartment of the simulation program. For a given serum pool (respectively fecal pool), the probability that it yielded a positive ELISA (respectively qPCR) result was modeled based on the relative diagnostic performances which were estimated in the experimental part of our study. For each combination of flock size, infection prevalence, number and size of pools, 1000 flocks were simulated, each being sampled over 1000 iterations. For each iteration, the number of pools that yielded a positive test result was recorded for each simulated flock and a flock was considered to be positive when a single pool was found positive. Flock sensitivities of pooled serum ELISA ($FPSe_{ELISA}$) and of pooled fecal qPCR ($FPSe_{qPCR}$) were computed as the proportion of infected flocks that

yield at least one positive pool and flock specificities ($FPSp_{ELISA}$ and $FPSp_{qPCR}$) as the proportion of paratuberculosis-free flocks that yielded no positive pool at all. Finally, the overall $FPSe_{ELISA}$, $FPSe_{qPCR}$, $FPSp_{ELISA}$ and $FPSp_{qPCR}$, number of positive pools and 95% confidence interval were calculated using the median value and the 2.5% and 97.5% percentiles of the estimate distributions over the 1000 iterations. The schematic representation of the simulated sampling approach is summarized in Fig. 1 for fecal qPCR and is provided as additional file for serum ELISA. Two flock sizes (100 and 300) were simulated with infection prevalence ranging from 0% (paratuberculosis free flocks) to 30%. The number of ewes sampled in each flock was 50 (5 pools of size 10 or 10 pools of size 5) and 100 (10 pools of size 10). We also investigated scenarios where all animals within flocks were sampled and submitted to pooled serum ELISA or pooled fecal qPCR (pools of size 10). Although this situation is unlikely to occur in France due to logistical and cost constraints, it allowed us estimating the lowest infection prevalence that could be detected based on pooled-samples analysis. Beta probability distributions were used to account for uncertainty about parameter estimates. For sensitivity of pooled-samples analysis, numbers of positive and total pools were used to construct α and β parameters of beta probability distributions:

$$\begin{cases} \alpha = \text{number of positive results} + 1 \\ \beta = \text{number of truly positive samples tested} - \text{number of positive results} + 1 \end{cases} \quad (1)$$

The same approach was used for pooled-sample specificities, based on the numbers of true-negative pools and observed negative results. These distributions closely reflect the probability density functions of expected values under a binomial likelihood (31). For pooled serum ELISA, at the selected decision thresholds (see results section), no difference was evidenced between negative pools containing either only LN or up to 3 HN individual serums and results were gathered together. Similarly, results from serum pools constructed with one HP or 2 LP individual sample(s) were merged, as detection rates were the same. Finally, the relative diagnostic performances were similar for pools of size 5 and 10 at the selected decision thresholds. Therefore the same beta distributions were used for the two pool sizes. For the same reasons, results from pooled fecal qPCR based on an individual fecal amount of 3 or 10 grams were merged. Although a limited number of true-negative fecal pools were tested, all yielded a negative results, the specificity of pooled fecal qPCR (PSp_{qPCR}) was assumed to be perfect. The possibility of cross-contamination was assumed to be minimal due to inclusion of negative controls and other quality control procedures. Input parameters are summarized in Table 1. For further information, Beta distributions that were used to model these parameters are provided as additional data. The specificity estimates for serum ELISA and fecal qPCR at the individual levels used for infected flocks in our simulation study may appear rather low. We therefore investigated another scenario in which $ISp_{ELISA} = 99\%$ and $ISp_{qPCR} = 100\%$. This amounts to reducing the number of false-positive individual results compared to scenarios based on values derived from (29), and allowed estimating their influence on simulation results at the flock levels. Finally, we also compared results from scenarios based on pooled-samples analysis to those based on individual testing, which served as the (probably imperfect) gold standard.

Fig 1. Schematic representation of the simulation sampling approach for fecal qPCR.

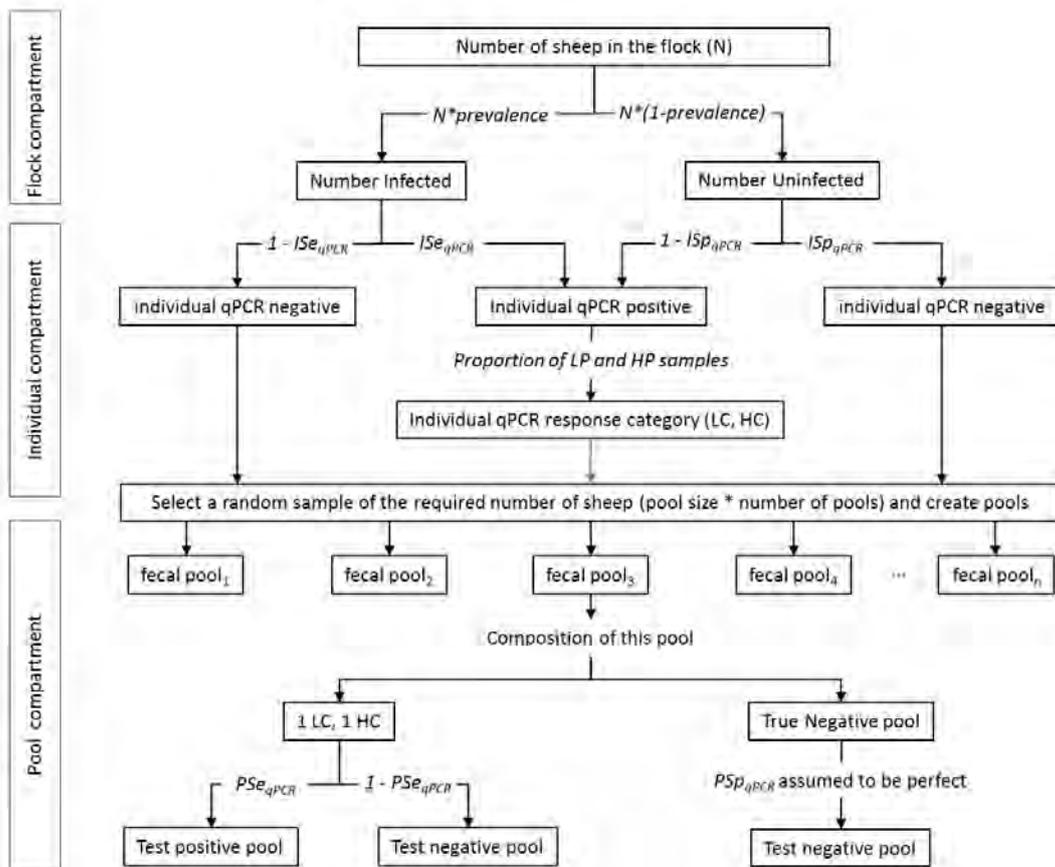


Table 1. Input parameters used in the simulation study

Parameters	Values	95% Confidence interval	Distribution
Flock size	100, 300		
Infection prevalence (%)	0, 1, 2.5, 5, 10, 15, 20, 30		
Individual serum ELISA			
ISeELISA (%), infected flock	17.4	10.8 - 26.6	Beta(16.3, 73.7)
ISpELISA (%), infected flock	94.0	91.8 - 95.6	Beta(583.8, 38.2)
ISpELISA (%), uninfected flock	99.0	96.6 - 99.7	Beta(210.0, 3.0)
Individual fecal qPCR			
ISe _q PCR (%), infected flock	47.5	26.2 - 69.9	Beta(8.8, 9.7)
ISp _q PCR (%), infected flock	99.0	97.6 - 99.6	Beta(487.4, 5.9)
ISp _q PCR (%), uninfected flock	100.0	99.0 - 100.0	Beta(388.0, 1.0)
% of LP serum samples	86.0	77.5 - 91.6	Beta(100.0, 7.0)
	100.0		
% of LC fecal samples	94.3	88.1 - 97.3	Beta(100.0, 7.0)
	100.0		
Number of sampled ewes	50, 100, all		
Number of pools	5, 10		
Pool size	5, 10		
Pooled serum ELISA			
PSeELISA (%), 1 LP	62.2	46.0 - 76.0	Beta(24.0, 15.0)
PSeELISA (%), ≥ 2 LP or ≥ 1 HP	100.0	91.8 - 99.9	Beta(43.0, 1.0)
PSpELISA (%)	100.0	96.5 - 100.0	Beta(103.0, 1.0)
Pooled fecal			
PSe _q PCR (%), 1 LC, pool size 5	89.0	79.8 - 94.3	Beta(66.0, 9.0)
PSe _q PCR (%), 1 LC, pool size 10	68.2	57.7 - 77.1	Beta(59.0, 28.0)
PSe _q PCR (%), 1 HC, pool size 5	99.2	95.5 - 99.8	Beta(122.0, 2.0)
PSe _q PCR (%), 1 HC, pool size 10	100.0	97.1 - 100.0	Beta(126.0, 1.0)
PSp _q PCR (%)	100.0		

Results

Individual test result distributions

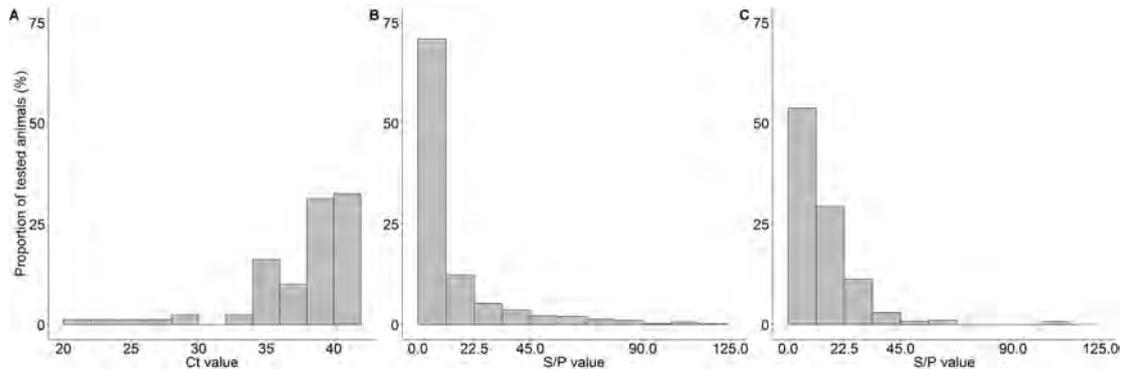
Individual serum ELISA results

The distributions of quantitative results for individual serum ELISA (S/P) and fecal qPCR (Ct) are shown in Fig. 2 for the 14 infected and the 3 paratuberculosis-free flocks. Overall, at the $S/P \geq 45\%$ decision threshold, 20 and 73 individual samples were considered as doubtful ($45 \geq SP < 55\%$) and positive ($SP \geq 55\%$) in infected flocks, respectively. Among them, 13 samples (14.0%) were classified as “High Positive” (HP, S/P value $\geq 90\%$). Among the 1104 negative results, 128 (11.6%) had S/P values between 22.5% and 45% and were regarded as “High Negative” (HN). At the flock level, the ELISA-based individual apparent prevalence ranged between 1.05% and 16.47% (median value 7.2%). HP samples were found in almost all infected flocks. Although the proportion of HN samples differed between flocks (range 4.6 – 26.8%) there was no evidence that it was influenced by the apparent prevalence within each flock. Finally, over the 211 individual serum ELISA results available in paratuberculosis-free flocks, 2 “Low Positive” (LP) samples were found in two different flocks while 34 (16.3%) negative results had HN values.

Individual fecal qPCR results

No positive result was found in the 387 ewes sampled from the three paratuberculosis free flocks. In infected flocks 105 samples (8.78%) yielded a positive qPCR response, with intra-flock individual apparent prevalence ranging from 0.0% to 29.1 % (first quartile 3.0%, third quartile 12.8%). Among those, only 6 (5.7%) with Ct < 30 were found in 5 different flocks and were classified as “Highly Contaminated” (HC).

Fig 2. Distribution of individual test results. distribution of individual Ct and S/P values for serum ELISA and fecal qPCR in 14 infected and 2 paratuberculosis-free flocks. A: fecal qPCR Ct ; B: ELISA S/P values in infected flocks; C: ELISA S/P values in paratuberculosis-free flocks.



Relative performances of pooled-sample analysis

Pooled serum ELISA

As expected, pooled serum ELISA was strongly influenced by both pool size and sample composition. Results are summarized in Table 2 and Table 3 as numbers of pools that yielded a positive response at different chosen S/P decision thresholds for pools of size 5 and 10, respectively. Retaining the $S/P \geq 45\%$ threshold that is applied to individual serum samples led to poor detection rates, especially when only one LP serum sample was incorporated (21.6% and 0% for pool size 5 and 10, respectively). Lowering the decision threshold improved the detection rate, but choosing very low cut-off values led to false positive results in true negative pools. For pools of size 5, the $S/P > 25\%$ threshold was associated with a perfect specificity for all investigated combinations of negative individual samples and with moderate to high

sensitivity estimates (62.2% for pools containing one LP sample, 100% for pools with 2 LN samples or 1 HP sample). Almost identical results were obtained when applying the $S/P > 15\%$ decision threshold for pools of size 10. The relative diagnostic performances associated to these thresholds were therefore selected for the simulation study and were hereafter assumed to be equal for both pool sizes.

Table 2. Pooled serum ELISA results for pools of size 5. Number of pools that were deemed to be positive given the S/P decision threshold

Pool composition*	# tested	# positive for different S/P (%) decision thresholds				
		20%	25%	30%	35%	45%
5 LN	30	0	0	0	0	0
1 HN	41	2	0	0	0	0
2 HN	21	4	0	0	0	0
3 HN	10	2	0	0	0	0
1 LP	37	28	23	21	19	8
2 LP	21	21	21	21	20	19
1 HP	21	21	21	20	20	18

* LN: $S/P < 22.5\%$; HN: $22.5\% \leq S/P < 45.0\%$; LP: $45 \leq S/P < 90\%$ and HP: $S/P \geq 90\%$.

Table 3. Pool serum ELISA results for pools of size 10.

Pool composition*	# tested	# positive for different S/P (%) decision thresholds				
		10%	15%	20%	25%	45%
10 LN	21	1	0	0	0	0
1 HN	38	7	0	0	0	0
2 HN	27	5	0	0	0	0
3 HN	16	3	0	0	0	0
1 LP	37	32	24	17	14	0
2 LP	21	21	21	20	20	8
1 HP	21	21	21	19	19	11

* LN: $S/P < 22.5\%$; HN: $22.5\% \leq S/P < 45.0\%$; LP: $45 \leq S/P < 90\%$ and HP: $S/P \geq 90\%$.

Pooled fecal qPCR

None of the 40 truly-negative fecal pools yielded a positive qPCR result, indicating that the analytical specificity of pool fecal qPCR was perfect ($PSP_{qPCR} = 100\%$). All except one of the 217 pools constructed by including a HC sample were detected by qPCR, whatever the pool size (5 or 10) or the amount of feces (3 or 10 grams) used to construct the pools (Table 4). Conversely, respectively 20/23 (87.0%) and 45/50 (90.0%) pools of size 5 containing one LC sample and constructed with 3 and 10 grams of positive feces were detected. Pools of size 10 containing one LC sample were even less frequently detected, with 63% to 70% of positive results (Table 4). Noteworthy detection rates were similar whatever the individual fecal amount (3 or 10 grams) used to construct the pools, indicating that pooling higher amounts of feces did not lead to substantial improvement of sensitivity.

Simulation studies

Results in paratuberculosis-free flocks (Fig. 3) are imperfect for both qPCR and ELISA. For each test, $FPSp$ was similar when sampling 50 animals. $FPSp_{qPCR}$ (respectively $FPSp_{ELISA}$) range between 92.00% (respectively 62.00%) when testing 5 pools of 10 samples and 90.00% (respectively 61.00%) when testing 10 pools of 5 samples. When more animals are sampled, $FPSp$ decrease for both tests (85.00% for $FPSp_{qPCR}$

Table 4. Pool fecal qPCR results

Pool composition*	Individual fecal amount (g)	Pool size	# positive / # tested	Pooled fecal qPCR Se (%)	Exact 95% CI (%)
1 HC	3	5	62 / 62	100.0	94.2 - 100.0
		10	65 / 65	100.0	94.4 - 100.0
	10	5	59 / 60	98.3	91.1 - 99.7
		10	60 / 60	100.0	94.0 - 100.0
1 LC	3	5	20 / 23	87.0	67.9 - 95.5
		10	17 / 27	63.0	44.2 - 78.5
	10	5	45 / 50	90.0	78.6 - 95.6
		10	41 / 58	70.7	58.0 - 80.8

* HC: sample with qPCR Ct<30; LC: sample with qPCR Ct>30.

and 41.00% for FPS_{DELISA} when testing 100 animals). FPSe increase with the number of animals sampled and the within-herd prevalence. In each sampling strategy, FPSe_{ELISA} was higher than FPSe_{qPCR} for low simulated within-herd prevalence. When infection prevalence is over 15%, FPSe_{qPCR} was superior or equal at FPSe_{ELISA}. Sampling 100 or more animals in high prevalence flocks (within-herd prevalence $\geq 15\%$) results in perfect FPSe for both tests. At least one pool was found positive in flocks with an infection prevalence over 20% (respectively 15%) when testing 10 pools of 5 samples (respectively 5 pools of 10 samples).

Considering flocks as infected if 2 pools were positive leads to a higher FPS_p and lower FPSe. In this case FPS_{p,qPCR} was of 99% when testing 50 animals, 98% when testing 100 animals. The minimal within flock prevalence detected in 95% of the case was with this cut-off of 30% when testing 10 pools of 5 animals. When testing 5 pools of 10 animals this minimal prevalence was over 30%.

Number of pools detected by qPCR are shown in Fig. 4. We did not find a clear threshold to discriminate precisely flocks according to within-herd prevalence. For the same number of animals sampled, testing a higher number of pools lead to more accurate estimation of within-herd prevalence. For the same number of pool tested, the estimation accuracy is higher when sampling more animals.

Under the higher IS_p scenario (Fig. 5), FPS_{p,qPCR} was perfect as both IS_{p,qPCR} and PS_{p,qPCR} were assumed perfect. FPS_{DELISA} did not increase when sampling the whole flock. There was an increase of 6 to 7% when sampling less animals. For both tests FPSe was lower in the high IS_p scenario for all within-herd prevalence and sampling strategy. As within-herd prevalence increase, the difference between FPSe in the 2 scenarios decrease.

Discussion

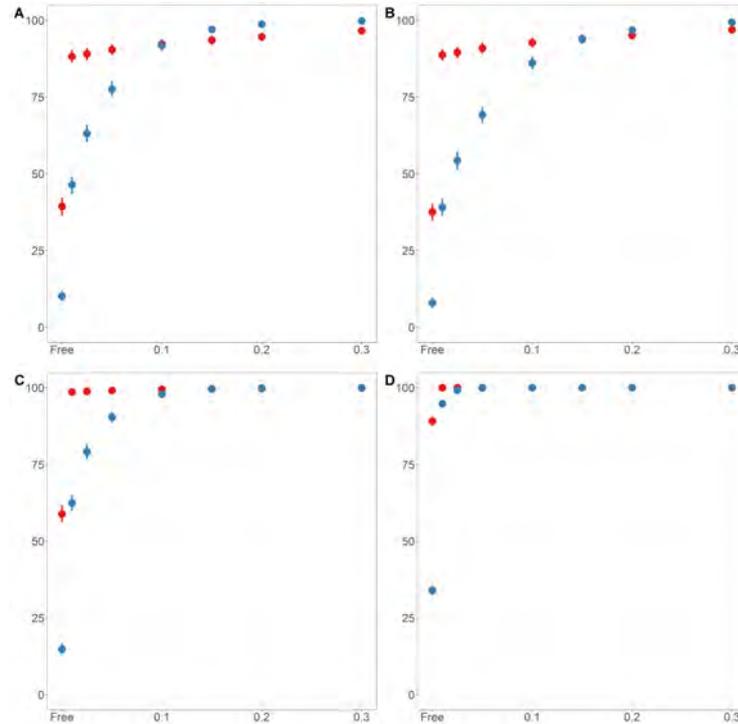
The observed proportion of HC samples was low in our study (5.7% of qPCR-positive animals). Compared to the multibacillary sheep proportion observed by Abbott et al. (32) or the proportion of high-shedding cows observed by (33). Even if these three cases definitions may be closely related (11) it can explain part of the differences observed between our results and those found in literature. In these 2 studies, HC samples proportion was calculated relatively to faecal culture positive animals and was based on an exhaustive sampling of the herd/flock. We chose to sample only 2 to 3 years old sub-clinical animals lowering by construction the HC samples prevalence. Although, qPCR is assumed to have a lowered detection threshold and a better diagnostic sensitivity than faecal culture (34) which can explain our results.

In this study, ISe and IS_p of both tests were drawn from beta distribution. We assume that the imperfect IS_{p,qPCR} observed in infected flocks was due to passive shedding. As none of the 237 samples from paratuberculosis-free flocks were positive to qPCR, the uncertainty of the IS_{p,qPCR} in paratuberculosis-free flocks results from the relatively small sampling size.

To evaluate pooled-sample ELISA performances, we lowered the cut-off value to compensate the dilution of positive samples. This approach has already been used by several authors (27 ; 28 ; 35) to improve

Fig 3. Proportion of detected flocks Proportion of flocks detected according to sampling strategy. Results show mean detection proportion and 95 confidence intervals for ELISA (red) and PCR (blue) varying with prevalence

A: 10 pools of 5 samples ; B: 5 pools of 10 samples; C: 10 pools of 10 samples; D 30 pools of 10 samples



sensitivity. As the lack of specificity is the main concern when screening a large population for a low-prevalence disease and as IS_{PELISA} estimate was poor in our study, we redefine S/P threshold to maximise PSe_{ELISA} and keeping PSP_{ELISA} equal to 1. Furthermore, the increase in PSe_{ELISA} when lowering the cut-off affected only pools containing one LP sample, PSe_{ELISA} for other pool compositions (2LP and 1HP) was already perfect.

PSe_{qPCR} was higher when pool contained HC samples and for smaller pool size. These results have been observed by several authors for *Map* using faecal culture (36 ; 26) or qPCR (22) (26) observed similar results using pooled faecal culture on pool of various size containing multibacillary or paucibacillary samples. In this study, PSe_{FC} was quasi-perfect up to a pool size of 50 samples without significant differences according to pool size. We also found a perfect PSe_{qPCR} when testing pools of 20 samples containing 1 HC sample (data not shown). We did not include this pool size in our simulation study as PSe_{qPCR} was lower and heterogeneous (from 13.3% to 40.0%) when testing pools of 20 samples containing one LC sample. As few HC samples were found in the 14 flocks, we focused on pool sizes maximising PSe_{PCR} .

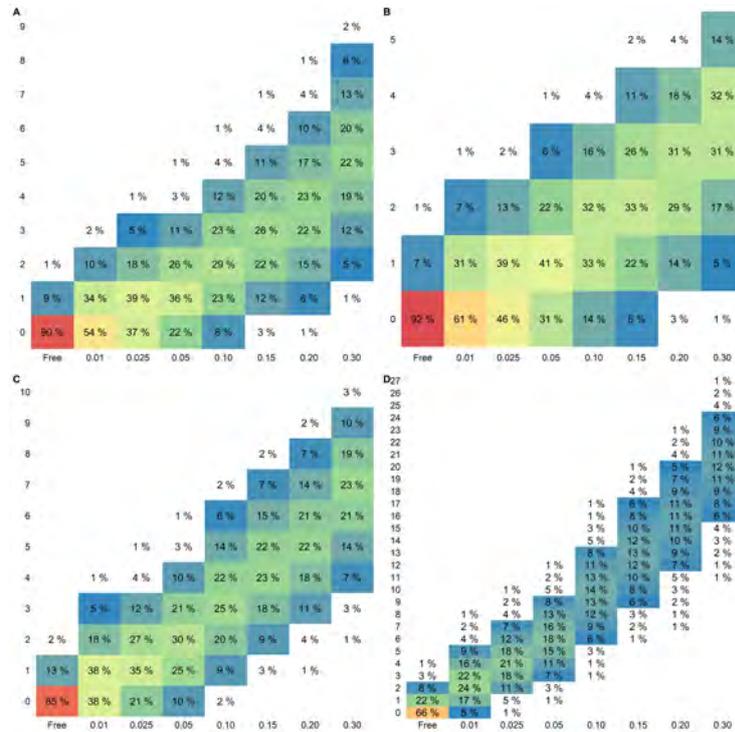
$FPSe_{qPCR}$ was higher in our study compared to the estimation obtained by Dhand et al. (2010) (26) when testing low prevalence flocks. We evaluate pooled-sample analysis to detect infection prevalence in opposition with shedding prevalence. As all infected animals are not shedding *Map*, infection prevalence must be higher than shedding prevalence. This could explain the better $FPSe_{qPCR}$ we observed in low prevalence flocks. There was less and less discrepancies between flock-level performances as the prevalence simulated was increasing. Whittington and McGregor(37) showed that lambs exposed to higher doses of *Map* are more likely to develop multibacillary lesions than those exposed to lower doses. The proportion of shedding animals among infected ones may increase with within-flock prevalence.

Some authors consider relative pool performances as perfect (38 ; 39) or independent from pool composition (40). If relative pool performances are considered perfect, $FPSe$ and $FPSp$ can be calculated

272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292

Fig 4. Number of pool detected by qPCR Colour scale represent the distribution of number of pools detected for a given prevalence. Percentages have to be read vertically and, for a given prevalence, gives the probability to have a number of pool detected

A: 10 pools of 5 samples ; B: 5 pools of 10 samples; C: 10 pools of 10 samples; D: 30 pools of 10 samples



knowing the within-flock prevalence (p), ISe and ISp , pool size (k) and number of pool tested (n) with the following equations:

$$\begin{aligned}
 FPS_e &= 1 - (((1 - ISe)^{k*p}) * (ISp^{k*(1-p)}))^n \\
 FPS_p &= ISp^{k*n}
 \end{aligned}
 \tag{2}$$

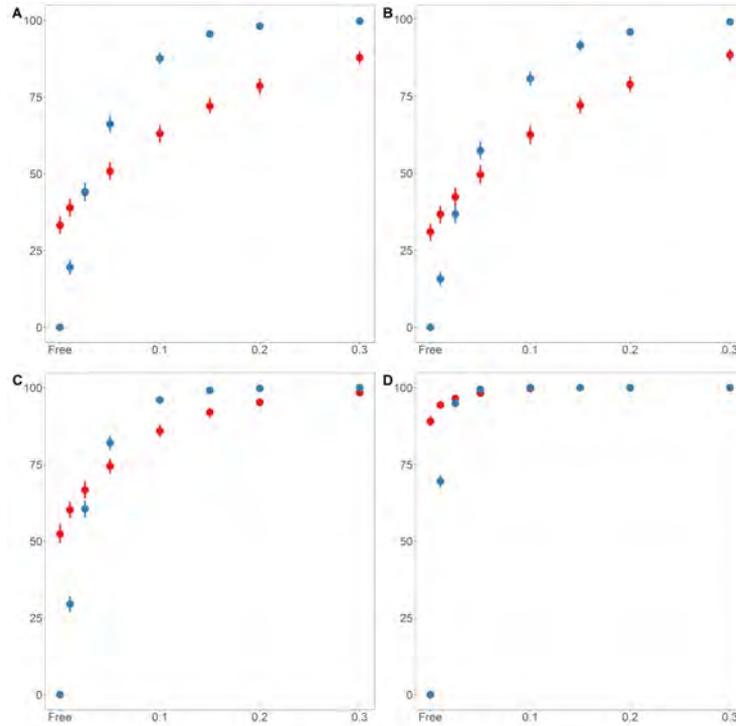
When these equations are applied with our parameters, theoretic $FPS_{e,qPCR}$ and $FPS_{e,ELISA}$ were equal to 1 when testing 50 samples in all infected flocks. Theoretic FPS_p is lower than FPS_p estimated in this study for both tests when sampling 50 animals. When ISp used for calculation was the upper bound of the 95% confidence interval this difference was smaller (theoretic $FPS_{p,qPCR} = 86.0\%$, theoretic $FPS_{p,ELISA} = 36.4\%$). Relative pooled performances should be evaluated before estimating flock-level performances. For both tests in our study the hypothesis of perfect pool performances was incorrect and led to biased flock-level performances.

Our within-flock prevalence estimations were in accordance with those calculated using the approach proposed by Speybroeck et al (39) implemented in the prevalence package (41) for the R (42). When testing 5 pools of 10 samples, we found slightly larger prevalence range when one or less pool was positive. One way to provide better within-flocks prevalence estimations may be to test pools of variable size as presented by Gu et al (43).

Flock-level performances in infected flock were worse when we fixed ISp_{qPCR} to 1. Imperfect ISp_{qPCR} , associated with a high PSe_{qPCR} , leads to falsely positive results that actually increase $FPS_{e,qPCR}$. Se and Sp are usually considered as intrinsic estimators of diagnostic performances, they will not vary with the prevalence of the population. This hypothesis is challenged by some authors (44 ; 45) and may not hold with paratuberculosis. Passive shedding will occur after a non productive ingestion of *Map* that can thereafter be

Fig 5. proportion of detected flocks under the higher ISp scenario Proportion of flocks detected according to sampling strategy. Results show mean detection proportion for ELISA (red) and PCR (blue) varying with prevalence

A: 10 pools of 5 samples; B: 5 pools of 10 samples ; C: 10 pools of 10 samples; D: 30 pools of 10 samples



detected in faeces and produce false positive results. This phenomenon is less likely to happen in paratuberculosis-free flocks as less *Map* are present in the environment. The exact diagnostic performances of pool testing are between the two scenarios evaluated, further studies will be needed to investigate the potential correlation between ISP_{qPCR} and within-flock prevalence.

Barkema et al. (46) pointed out the lack of reliable on-farm tests for paratuberculosis. Testing 10 pools of 5 animals by qPCR could be a sampling protocol fitted into the framework of a control program to monitor paratuberculosis infection in flocks. If testing occurs each year, results could be aggregated resulting in higher $FPS_{e_{qPCR}}$ over the years and a more precise estimation of the within-flock prevalence.

Acknowledgements

The authors thank all farmers, farm technicians and veterinary practitioners and students for their involvement in and dedication to this study.

Funding

This study was funded by the INRA GISA Metaprogram PICSAR and Region Midi-Pyrénées (PAROVLOT program). The lead author received a PhD grant from both funding bodies. The funding bodies did not have any direct role in the study design or sample collection.

Availability of data and materials 325

Individual data analyzed during the current study are available from the corresponding author on reasonable request. 326
327

Authors' contributions 328

YM, FC and GF developed and designed the study and participated in the animal sampling. YM carried out the laboratory analysis. YM and FC carried out the data analysis. YM, FC and GF drafted the initial manuscript. All authors read and approved the final manuscript. 329
330
331

Ethics approval and consent to participate 332

Animal studies were compliant with all applicable provisions established by the European Commission Directive 2010/63/UE. All animals used in this study were handled in strict accordance with good clinical practices and all efforts were made to minimize suffering. All animal owners gave written consent for their animals to be used in this study. 333
334
335
336

Consent for publication 337

Not applicable. 338

Competing interests 339

The authors declare that they have no competing interests. 340

1. Allworth MB, Kennedy DJ. Progress in national control and assurance programs for ovine Johne's disease in Australia. *Veterinary microbiology*. 2000;77(3):415–422. 341
342
2. Groenendaal H, Nielen M, Hesselink JW. Development of the Dutch Johne's disease control program supported by a simulation model. *Preventive Veterinary Medicine*. 2003;60(1):69–90. doi:10.1016/S0167-5877(03)00083-7. 343
344
345
3. Benedictus G, Verhoeff J, Schukken YH, Hesselink JW. Dutch paratuberculosis programme history, principles and development. *Veterinary microbiology*. 2000;77(3):399–413. 346
347
4. Kostoulas P, Leontides L, Billinis C, Florou M. Application of a semi-dependent latent model in the Bayesian estimation of the sensitivity and specificity of two faecal culture methods for diagnosis of paratuberculosis in sub-clinically infected Greek dairy sheep and goats. *Preventive Veterinary Medicine*. 2006;76(1-2):121–134. doi:10.1016/j.prevetmed.2006.04.008. 348
349
350
351
5. Kostoulas P, Leontides L, Enøe C, Billinis C, Florou M, Sofia M. Bayesian estimation of sensitivity and specificity of serum ELISA and faecal culture for diagnosis of paratuberculosis in Greek dairy sheep and goats. *Preventive Veterinary Medicine*. 2006;76(1-2):56–73. doi:10.1016/j.prevetmed.2006.04.006. 352
353
354
6. Acevedo-Whitehouse K, Vicente J, Gortazar C, Hofle U, Fernandez-De-Mera IG, Amos W. Genetic resistance to bovine tuberculosis in the Iberian wild boar. *Molecular Ecology*. 2005;14(10):3209–3217. doi:10.1111/j.1365-294X.2005.02656.x. 355
356
357
7. Nielsen SS, Toft N. Age-Specific Characteristics of ELISA and Fecal Culture for Purpose-Specific Testing for Paratuberculosis. *Journal of Dairy Science*. 2006;89(2):569–579. doi:10.3168/jds.S0022-0302(06)72120-8. 358
359
360

8. Dorfman R. The Detection of Defective Members of Large Populations. *The Annals of Mathematical Statistics*. 1943;14(4):436–440. doi:10.1214/aoms/1177731363. 361
362
9. Kalis CHJ, Hesselink JW, Barkema HW, Collins MT. Culture of Strategically Pooled Bovine Fecal Samples as a Method to Screen Herds for Paratuberculosis. *Journal of Veterinary Diagnostic Investigation*. 2000;12(6):547–551. doi:10.1177/104063870001200609. 363
364
365
10. Weber MF, Groenendaal H, van Roermund HJW, Nielen M. Simulation of alternatives for the Dutch Johne's disease certification-and-monitoring program. *Preventive Veterinary Medicine*. 2004;62(1):1–17. doi:10.1016/j.prevetmed.2003.11.006. 366
367
368
11. Whittington RJ, Fell S, Walker D, McAllister S, Marsh I, Sergeant E, et al. Use of Pooled Fecal Culture for Sensitive and Economic Detection of *Mycobacterium avium* subsp. paratuberculosis Infection in Flocks of Sheep. *Journal of clinical microbiology*. 2000;38(7):2550–2556. 369
370
371
12. Eamens GJ, Whittington RJ, Turner MJ, Austin SL, Fell SA, Marsh IB. Evaluation of radiometric faecal culture and direct PCR on pooled faeces for detection of *Mycobacterium avium* subsp. paratuberculosis in cattle. *Veterinary Microbiology*. 2007;125(1-2):22–35. doi:10.1016/j.vetmic.2007.04.043. 372
373
374
375
13. Eamens G, Walker D, Porter N, Fell S. Pooled faecal culture for the detection of *Mycobacterium avium* subsp paratuberculosis in goats. *Australian Veterinary Journal*. 2007;85(6):243–251. doi:10.1111/j.1751-0813.2007.00160.x. 376
377
378
14. Aly SS, Mangold BL, Whitlock RH, Sweeney RW, Anderson RJ, Jiang J, et al. Correlation between Herrold Egg Yolk Medium Culture and Real-Time Quantitative Polymerase Chain Reaction Results for *Mycobacterium Avium* Subspecies Paratuberculosis in Pooled Fecal and Environmental Samples. *Journal of Veterinary Diagnostic Investigation*. 2010;22(5):677–683. doi:10.1177/104063871002200501. 379
380
381
382
15. Nielsen SS, Thamsborg SM, Houe H, Bitsch V. Bulk-tank milk ELISA antibodies for estimating the prevalence of paratuberculosis in Danish dairy herds. *Preventive Veterinary Medicine*. 2000;44(1):1–7. doi:10.1016/S0167-5877(00)00098-2. 383
384
385
16. Stabel JR. Transitions in immune responses to *Mycobacterium paratuberculosis*. *Veterinary Microbiology*. 2000;77(3):465–473. doi:10.1016/S0378-1135(00)00331-X. 386
387
17. Slana I, Liapi M, Moravkova M, Kralova A, Pavlik I. *Mycobacterium avium* subsp. paratuberculosis in cow bulk tank milk in Cyprus detected by culture and quantitative IS900 and F57 real-time PCR. *Preventive Veterinary Medicine*. 2009;89(3):223–226. doi:10.1016/j.prevetmed.2009.02.020. 388
389
390
18. Raizman EA, Wells SJ, Godden SM, Bey RF, Oakes MJ, Bentley DC, et al. The Distribution of *Mycobacterium avium* ssp. paratuberculosis in the Environment Surrounding Minnesota Dairy Farms. *Journal of Dairy Science*. 2004;87(9):2959–2966. doi:10.3168/jds.S0022-0302(04)73427-X. 391
392
393
19. Whittington R, Marsh I, Taylor P, Marshall D, Taragel C, Reddacliff L. Isolation of *Mycobacterium avium* subsp paratuberculosis from environmental samples collected from farms before and after destocking sheep with paratuberculosis. *Australian Veterinary Journal*. 2003;81(9):559–563. doi:10.1111/j.1751-0813.2003.tb12887.x. 394
395
396
397
20. van Schaik G, R Rossiter C, Stehman S, J Shin S, Schukken Y. Longitudinal study to investigate variation in results of repeated ELISA and culture of fecal samples for *Mycobacterium avium* subsp paratuberculosis in commercial dairy herds. *American journal of veterinary research*. 2003;64:479–84. doi:10.2460/ajvr.2003.64.479. 398
399
400
401
21. Whittington RJ, Whittington AM, Waldron A, Begg DJ, de Silva K, Purdie AC, et al. Development and Validation of a Liquid Medium (M7H9C) for Routine Culture of *Mycobacterium avium* subsp. paratuberculosis To Replace Modified Bactec 12B Medium. *Journal of Clinical Microbiology*. 2013;51(12):3993–4000. doi:10.1128/JCM.01373-13. 402
403
404
405

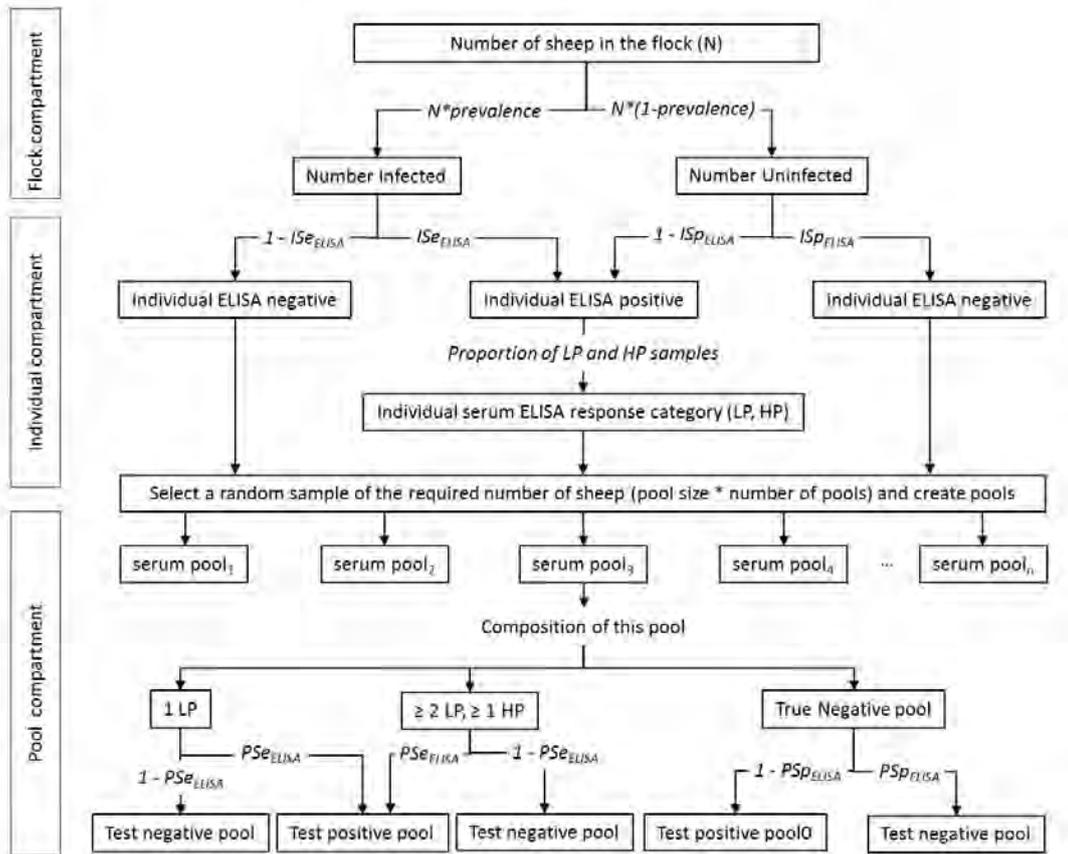
22. Mita A, Mori Y, Nakagawa T, Tasaki T, Utiyama K, Mori H. Comparison of fecal pooling methods and DNA extraction kits for the detection of *Mycobacterium avium* subspecies paratuberculosis. *MicrobiologyOpen*. 2015;5(1):134–142. doi:10.1002/mbo3.318. 406–408
23. Timms VJ, Mitchell HM, Neilan BA. Optimisation of DNA extraction and validation of PCR assays to detect *Mycobacterium avium* subsp. paratuberculosis. *Journal of Microbiological Methods*. 2015;112:99–103. doi:10.1016/j.mimet.2015.03.016. 409–411
24. Fock-Chow-Tho D, Topp E, Ibeagha-Awemu EA, Bissonnette N. Comparison of commercial DNA extraction kits and quantitative PCR systems for better sensitivity in detecting the causative agent of paratuberculosis in dairy cow fecal samples. *Journal of Dairy Science*. 2016;doi:10.3168/jds.2016-11384. 412–414
25. Pesqueira MN, Yus E, Factor C, Mato I, Sanjuán ML, Eiras C, et al. Short communication: Correlation between within-herd antibody-prevalence and bulk tank milk antibody levels to *Mycobacterium avium* ssp. paratuberculosis using 2 commercial immunoassays. *Journal of Dairy Science*. 2017;100(9):7544–7548. doi:10.3168/jds.2017-12706. 415–418
26. Dhand NK, Sergeant E, Toribio JALML, Whittington RJ. Estimation of sensitivity and flock-sensitivity of pooled faecal culture for *Mycobacterium avium* subsp. paratuberculosis in sheep. *Preventive Veterinary Medicine*. 2010;95(3-4):248–257. doi:10.1016/j.prevetmed.2010.03.013. 419–421
27. Brinkhof JMA, Houwers DJ, van Maanen C. Development of a sample pooling strategy for the serodiagnosis of small ruminant lentiviral infections using the ELITEST-MVV ELISA. *Small Ruminant Research*. 2007;70(2-3):194–199. doi:10.1016/j.smallrumres.2006.03.003. 422–424
28. Rovira A, Cano JP, Muñoz-Zanzi C. Feasibility of pooled-sample testing for the detection of porcine reproductive and respiratory syndrome virus antibodies on serum samples by ELISA. *Veterinary Microbiology*. 2008;130(1):60–68. doi:10.1016/j.vetmic.2007.12.016. 425–427
29. Mathevon Y, Foucras G, Falguières R, Corbiere F. Estimation of the sensitivity and specificity of two serum ELISAs and one fecal qPCR for diagnosis of paratuberculosis in sub-clinically infected young-adult French sheep using latent class Bayesian modeling. *BMC Veterinary Research*. 2017;13:230. doi:10.1186/s12917-017-1145-x. 428–431
30. Nielsen SS, Toft N. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon- γ assay and faecal culture techniques. *Veterinary Microbiology*. 2008;129(3-4):217–235. doi:10.1016/j.vetmic.2007.12.011. 432–434
31. McV Messam LL, Branscum AJ, Collins MT, Gardner IA. Frequentist and Bayesian approaches to prevalence estimation using examples from Johne's disease. *Animal Health Research Reviews*. 2008;9(01):1–23. doi:10.1017/S1466252307001314. 435–437
32. Abbott KA, Whittington RJ, McGregor H. Exposure factors leading to establishment of OJD infection and clinical disease: epidemiology of OJD-1. *Meat and Livestock Australia Limited*; 2004. 438–439
33. Tavoranpanich S, Muñoz-Zanzi CA, Wells SJ, Raizman EA, Carpenter TE, Johnson WO, et al. Simulation model for evaluation of testing strategies for detection of paratuberculosis in Midwestern US dairy herds. *Preventive Veterinary Medicine*. 2008;83(1):65–82. doi:10.1016/j.prevetmed.2007.06.010. 440–443
34. Soumya MP, Pillai RM, Antony PX, Mukhopadhyay HK, Rao VN. Comparison of faecal culture and IS900 PCR assay for the detection of *Mycobacterium avium* subsp. paratuberculosis in bovine faecal samples. *Veterinary Research Communications*. 2009;33(7):781–791. doi:10.1007/s11259-009-9226-3. 444–446
35. Græsboell K, Andresen LO, Halasa T, Toft N. Opportunities and challenges when pooling milk samples using ELISA. *Preventive Veterinary Medicine*. 2017;139:93–98. doi:10.1016/j.prevetmed.2016.08.001. 447–448

36. van Schaik G, Pradenas F M, Mella N A, Kruze V J. Diagnostic validity and costs of pooled fecal samples and individual blood or fecal samples to determine the cow- and herd-status for *Mycobacterium avium* subsp. *paratuberculosis*. *Preventive Veterinary Medicine*. 2007;82(1-2):159–165. doi:10.1016/j.prevetmed.2007.05.018. 449–452
37. Whittington RJ, McGregor H. Epidemiology of ovine Johne's disease 2 – pasture contamination level and age susceptibility; 2005. 1 74036 682 4. Available from: <https://ses.library.usyd.edu.au/bitstream/2123/950/1/OJD.028> 453–455
38. Boelaert F, Biront P, Soumare B, Dispas M, Vanopdenbosch E, Vermeersch JP, et al. Prevalence of bovine herpesvirus-1 in the Belgian cattle population. *Preventive Veterinary Medicine*. 2000;45(3):285–295. doi:10.1016/S0167-5877(00)00128-8. 456–458
39. Speybroeck N, Williams CJ, Lafia KB, Devleeschauwer B, Berkvens D. Estimating the prevalence of infections in vector populations using pools of samples. *Medical and Veterinary Entomology*. 2012;26(4):361–371. doi:10.1111/j.1365-2915.2012.01015.x. 459–461
40. Toribio JA, Sergeant E. A comparison of methods to estimate the prevalence of ovine Johne's infection from pooled faecal samples. *Australian Veterinary Journal*. 2007;85(8):317–324. doi:10.1111/j.1751-0813.2007.00188.x. 462–464
41. Devleeschauwer B, Torgerson P, Charlier J, Levecke B, Praet N, Roelandt S, et al.. prevalence: Tools for Prevalence Assessment Studies; 2015. Available from: <https://cran.r-project.org/web/packages/prevalence/index.html>. 465–467
42. Team RDC. R: A Language and Environment for Statistical Computing; 2017. Available from: <https://www.r-project.org/>. 468–469
43. Gu W, Lampman R, Novak RJ. Assessment of arbovirus vector infection rates using variable size pooling. *Medical and Veterinary Entomology*. 2004;18(2):200–204. doi:10.1111/j.0269-283X.2004.00482.x. 470–472
44. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Statistics in Medicine*. 1997;16(9):981–991. 473–474
45. Leeflang MMG, Bossuyt PMM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *Journal of Clinical Epidemiology*. 2009;62(1):5–12. doi:10.1016/j.jclinepi.2008.04.007. 475–477
46. Barkema H, Orsel K, Nielsen S, Koets A, Rutten VP, P Bannantine J, et al. Knowledge gaps that hamper prevention and control of *Mycobacterium avium* subspecies *paratuberculosis* infection. *Transboundary and Emerging Diseases*. 2017;doi:10.1111/tbed.12723. 478–480

Supporting information

481

Schematic representation of the simulation sampling approach for ELISA.

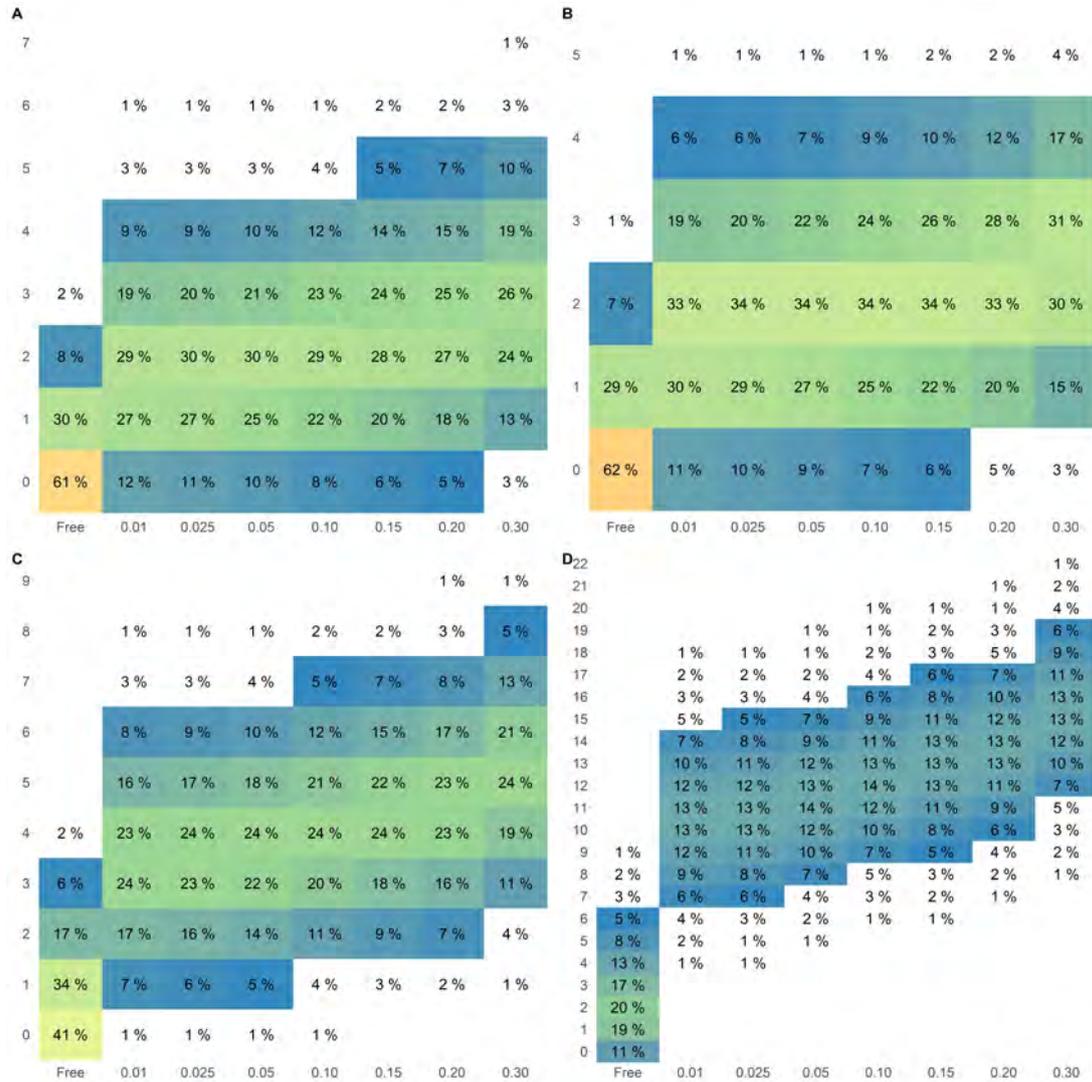


S1 Fig.

482

Number of pool detected by ELISA Colour scale represent the distribution of number of pools detected for a given prevalence. Percentages have to be read vertically and, for a given prevalence, gives the probability to have a number of pool detected

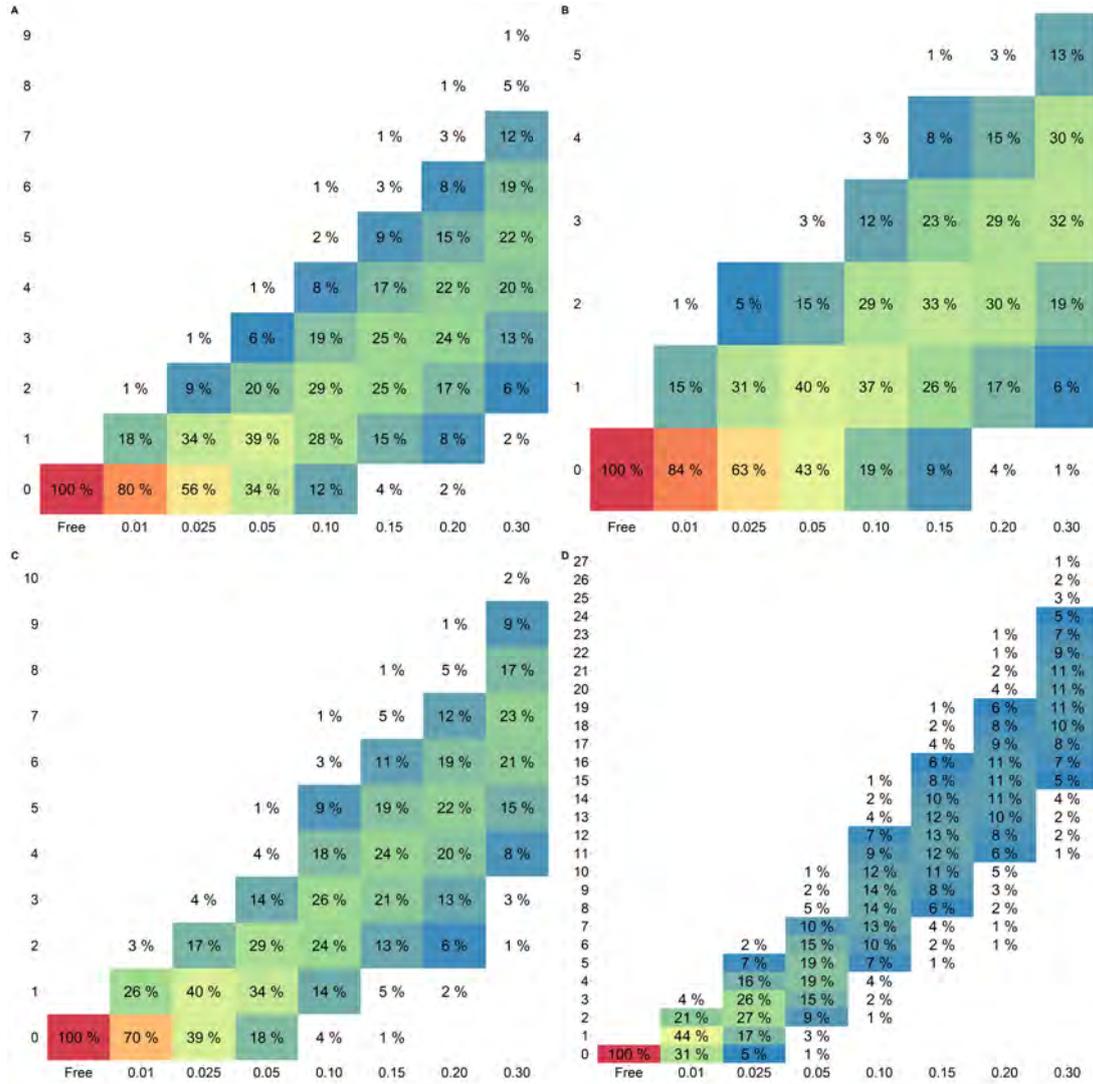
A: 10 pools of 5 samples ; B: 5 pools of 10 samples; C: 10 pools of 10 samples; D 30 pools of 10 samples



S2 Fig.

Number of pool detected by PCR under higher ISp scenario Colour scale represent the distribution of number of pools detected for a given prevalence. Percentages have to be read vertically and, for a given prevalence, gives the probability to have a number of pool detected

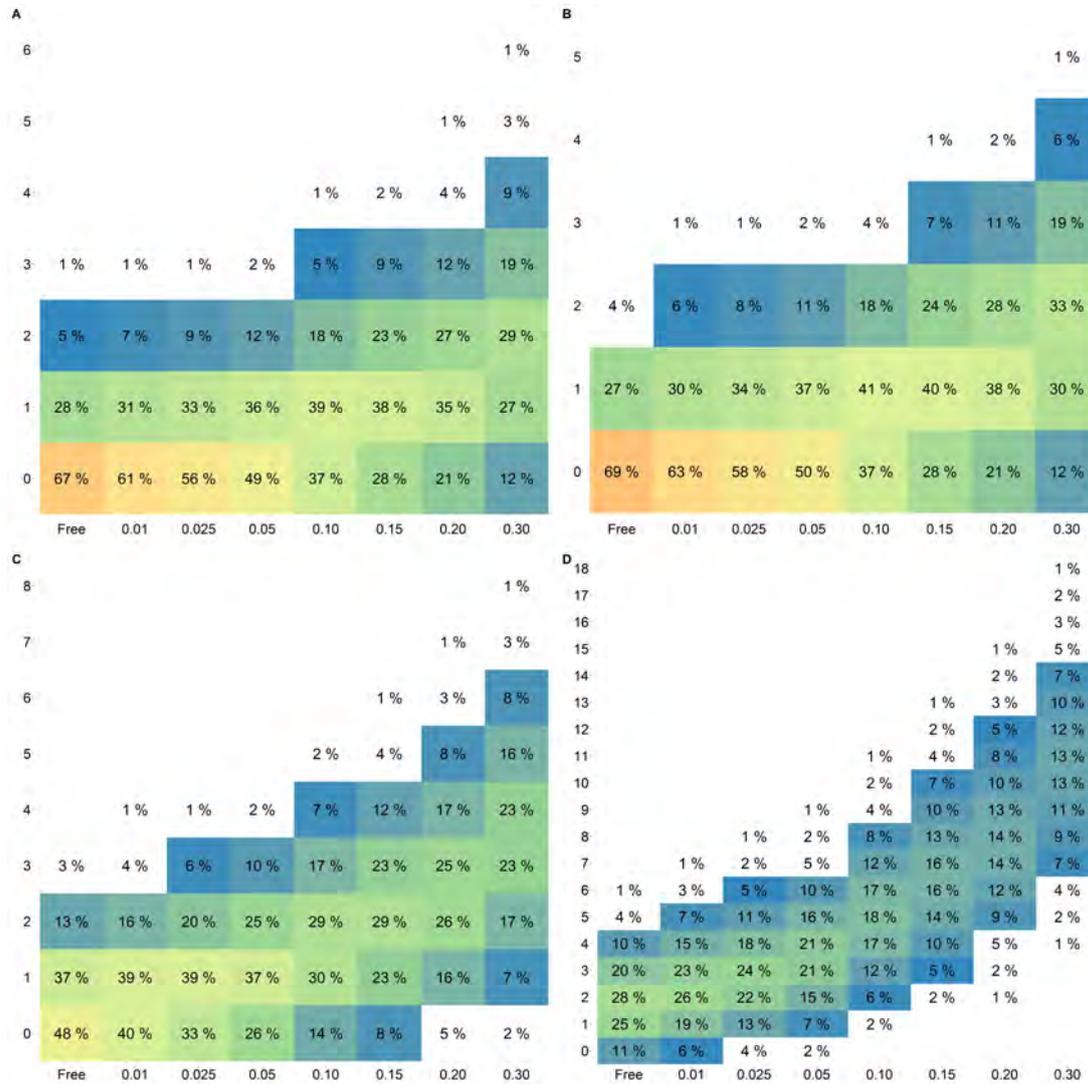
A: 10 pools of 5 samples ; B: 5 pools of 10 samples; C: 10 pools of 10 samples; D 30 pools of 10 samples



S3 Fig.

Number of pool detected by ELISA under higher ISp scenario Colour scale represent the distribution of number of pools detected for a given prevalence. Percentages have to be read vertically and, for a given prevalence, gives the probability to have a number of pool detected

A: 10 pools of 5 samples ; B: 5 pools of 10 samples; C: 10 pools of 10 samples; D: 30 pools of 10 samples



S4 Fig.

2 Compléments : analyses de mélanges d'échantillons

2.1 Analyse par qPCR de mélanges de fèces

2.1.1 Matériels et méthodes

Six ovins ont été sélectionnés sur la base de leurs résultats de qPCR individuelle (tableau 5). Ces animaux ont été suivis sur le site de l'ENVT, en cases individuelles, afin de réaliser des prélèvements de fèces durant deux semaines. Des sachets journaliers de fèces ont été constitués et identifiés pour chaque animal. Un sachet contenant plus de 300g de fèces a été sélectionné pour chaque animal et huit analyses de qPCR, quatre à partir d'une prise d'échantillon de trois grammes et quatre à partir d'une prise d'échantillon de 10 grammes, ont été réalisées. Le protocole de qPCR appliqué est celui de la trousse ADIAVETTM paraTB Real Time, BioX. Deux extractions d'ADN ont été réalisées pour chaque point et l'amplification a été réalisée deux fois, afin d'évaluer la variabilité inter- et intra-plaque.

Cette étape d'analyse individuelle nous a permis de connaître précisément le niveau d'excrétion individuel correspondant aux sachets de fèces sélectionnés avant d'entreprendre l'analyse de mélanges d'échantillons. Nous avons réalisé des dilutions de ces fèces positives en utilisant des fèces négatives provenant d'un élevage expérimental de brebis Lacaune indemne de paratuberculose (troupeau INRA de La Fage).

Le plan d'analyse est présenté à la figure 18. Chaque sachet a été testé en qPCR en fonction de deux prises initiales de fèces différentes (3 et 10 grammes) correspondant aux extremums fournis par le fabricant de la trousse de qPCR. Trois tailles de mélanges (5, 10 et 20) ont été évaluées. Chaque mélange a été réalisé à partir d'une prise initiale d'un sachet d'un des animaux infectés, complétée avec des fèces négatives et suivi d'une homogénéisation manuelle effectuée dans un sachet plastique pendant une à deux minutes.

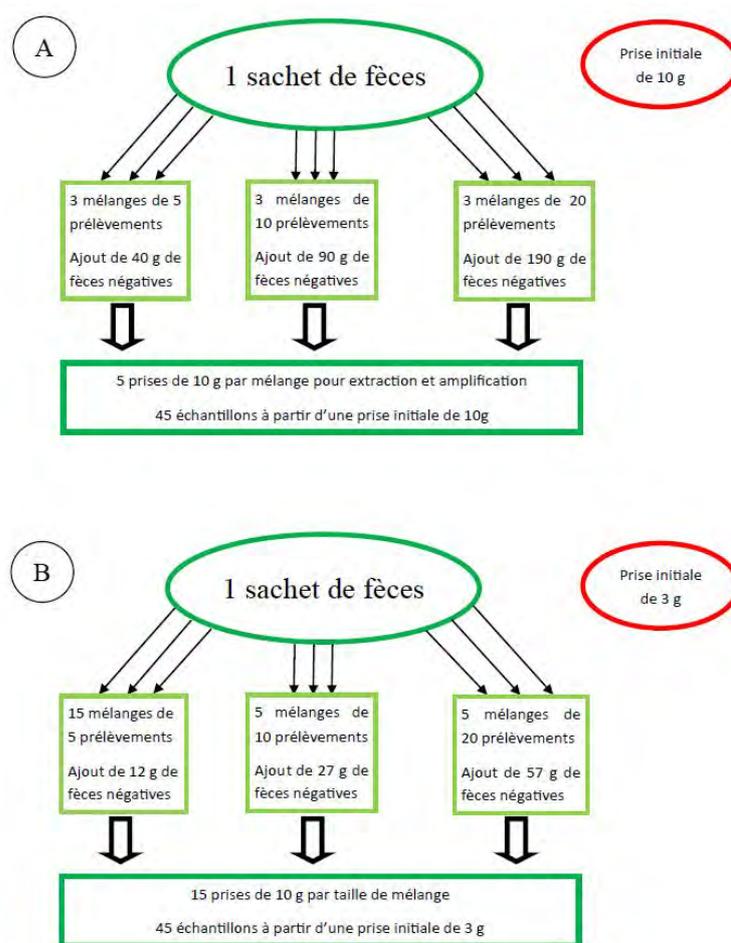


FIGURE 18 – Plan d'analyse pour l'analyse de la Se_D de la qPCR en mélange de fèces

Le plan d'analyse est construit en fonction de la prise initiale dans le sachet de fèces.

A) Neuf prises de 10 g ont été effectuées, trois permettant de faire des mélanges de cinq animaux par ajout de 40 g de fèces négatives, trois pour des mélanges de taille 10, par ajout de 90 g de fèces négatives et trois pour des mélanges de taille 20 par ajout de 210 g de fèces négatives. Chaque mélange (15 au total) a été extrait et amplifié cinq fois à partir d'une prise de 10 g.

B) Vingt-cinq prises de trois g ont été effectuées, 15 permettant de faire des mélanges de cinq animaux par ajout de 12 g de fèces négatives, cinq pour des mélanges de taille 10 par ajout de 27 g de fèces négatives et cinq pour des mélanges de taille 20 par ajout de 57 g de fèces négatives. Pour chaque taille de mélange, 15 extractions et amplifications ont été réalisées à partir de 10 g.

2.1.2 Résultats

Les résultats des analyses individuelles sont présentés dans la figure 19. Les variations inter-plaques ont été évaluées en calculant le coefficient de variation (CV) défini par le rapport entre l'écart type des valeurs de Ct et la moyenne des valeurs de Ct. Le CV a tout d'abord été calculé entre les deux plaques d'amplification à partir des échantillons individuels afin de s'assurer de la répétabilité de la méthode de qPCR (préparation des échantillons, extraction et amplification). Pour chaque animal ce coefficient a été comparé en

Tableau 5 – Résultats de l'analyse qPCR sur fèces des animaux choisis pour l'évaluation de la sensibilité de la qPCR ADIAVET™ paraTB Real Time (BioX) appliquée à des mélanges de fèces.

ID	Ct moyen	Ct minimum	Ct maximum
1	39,3	35,3	Négatif
2	35,8	34,1	39,6
3	23,0	20,5	25,0
4	20,0	19,2	22,2
5	20,9	18,0	23,2
6	26,7	26,2	27,7

Liste des animaux prélevés afin de déterminer la Se_D relative de la qPCR sur mélange de fèces.
Le Ct présenté est la moyenne des Ct des huit analyses individuelles réalisées avant les analyses en mélanges

fonction de la prise initiale de fèces.

Nous n'avons pas observé de différences du CV entre les prises de 3 g (minimum : 0,05%, moyenne : 4,11%, maximum :13,77%) et de 10 g (minimum : 0,18%, moyenne : 3,55%, maximum :13,73%) ni de relation entre la valeur du Ct et le CV. Ces résultats nous permettent de conclure à une bonne répétabilité de l'analyse qPCR. Le même travail a été effectué pour évaluer la variation de la valeur de Ct intra-plaque entre les différentes prises d'échantillons provenant d'un même sachet. A l'exception de deux points de qPCR, toutes les valeurs de CV étaient inférieures à 10%.

Les résultats observés dans la figure 19 sont en accord avec les différences de niveau d'excrétion individuel. La valeur de Ct théorique a été calculée en appliquant la relation définie par l'équation (1) présentée au chapitre 1. En utilisant le seuil de positivité de $Ct \leq 42$, tous les mélanges contenant des fèces d'animaux forts excréteurs étaient positifs, et un seul mélange était considéré comme négatif en abaissant le seuil de positivité de deux cycles ($Ct \leq 40$).

Pour les mélanges contenant un échantillon dont la valeur de Ct individuelle était élevée (brebis 1 ou 2), les résultats sont apparus plus variables et peuvent refléter une mauvaise homogénéisation des mélanges de fèces.

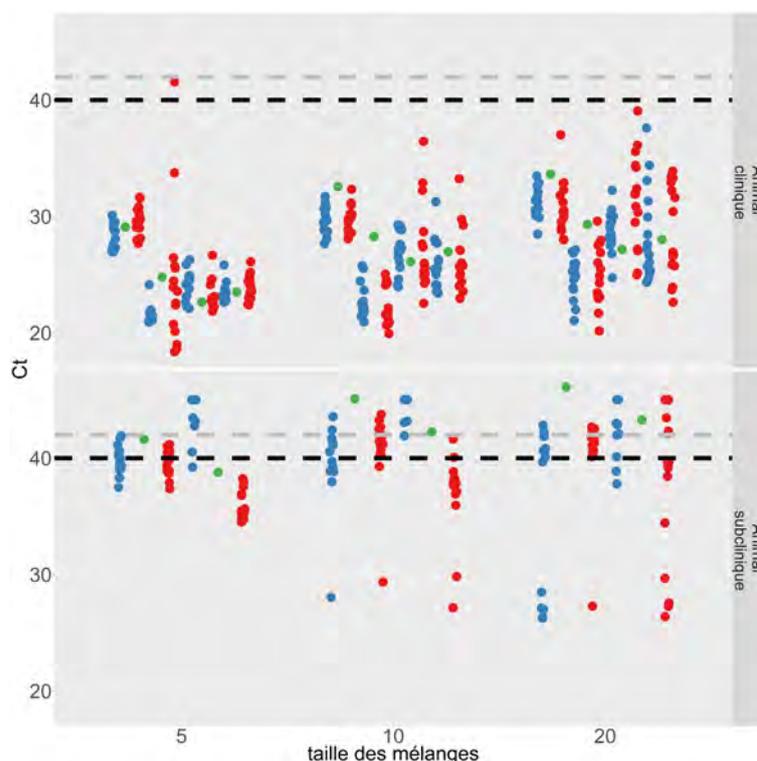


FIGURE 19 – Résultats des analyses de qPCR sur des mélanges de fèces

Chaque mélange analysé est représenté par un point bleu ou rouge.

Pour chaque taille, les mélanges d'une brebis sont groupés autour du Ct théorique représenté par un point vert. Les points bleus sont les mélanges réalisés à partir de 3 g de fèces, les points rouges sont ceux réalisés à partir de 10 g de fèces.

Le panel de haut de la figure regroupe les résultats des quatre animaux forts excréteurs, celui du bas ceux des animaux faibles excréteurs.

Les seuils de positivité de la qPCR sont représentés par deux lignes horizontales en pointillées. En noir, le seuil de Ct=40 recommandé par le fabricant, en gris le seuil de Ct=42 calculé comme le seuil maximum de détection.

2.2 Analyse par sérologie ELISA de mélanges de sérums

Nous avons analysé des mélanges de sérums avec les deux trousse de sérologie ELISA utilisées au chapitre précédent. Pour cela, nous avons catégorisé les sérums en quatre classes en fonction de la valeur S/P obtenue en analyse individuelle, selon les seuils définis dans le tableau 6 :

Une expérience de dilution sériée a été réalisée à partir de 63 sérums ayant des valeurs S/P comprises entre 30 et 220% pour la trousse ID Screen Paratuberculosis Indirect® (ELISA A) et entre 30 et 125% pour la trousse IDEXX paratuberculosis screening® (ELISA B). Cet ensemble de sérums se répartissait comme suit :

- Échantillons négatifs forts (High negative, HN) : 7 pour l'ELISA A, 15 pour l'ELISA B ;

Tableau 6 – Classes de sérum en fonction de la valeur S/P

	Classes des sérums (Valeur S/P)			
	Négatif faible (Low negative, LN)	Négatif fort (High negative, HN)	Positif faible (Low positive, LP)	Positif fort (High positive, HP)
ELISA A	[0-30[[30-60[[60-120[>120
ELISA B	[0-22,5[[22,5-45[[45-90[>90

Définition des seuils S/P pour les deux trousse de sérologie ELISA étudiées. Le seuil inférieur a été défini comme la moitié du seuil correspondant à un résultat douteux tel que recommandé par les fabricants pour l'analyse individuelle. Le seuil supérieur a été défini comme le double du seuil de positivité recommandé par les fabricants pour l'analyse individuelle. ELISA A : ID Screen Paratuberculosis Indirect[®], lot 800, IdVet; ELISA B : IDEXX paratuberculosis screening[®], lot 5074

- Échantillons positifs faibles (Low positive, LP) : 16 pour l'ELISA A, 32 pour l'ELISA B;
- Échantillons positifs forts (High positive, HP) : 40 pour l'ELISA A, 16 pour l'ELISA B.

Nous avons choisi une approche par dilutions sériées dans du sérum négatif faible (Low negative, LN), assurant une très bonne répétabilité des résultats permise par la facilité de manipulation. Nous avons choisi des dilutions sériées de facteur trois permettant d'apporter des résultats proches des taux de dilution opérationnels recherchés (mélange de 5 et 10 sérums). Comme le montre les figures 20 and 21, les valeurs S/P suivent dans l'ensemble une décroissance proche de la linéarité.

En conservant le seuil de positivité appliqué aux échantillons individuels ($S/P > 70\%$ pour l'ELISA A, et $S/P > 55\%$ pour l'ELISA B), les échantillons avec un résultat négatif et une valeur S/P initiale proche du seuil de positivité sont restés négatifs lors des dilutions successives. A l'inverse, aucun des échantillons faiblement positifs n'a été retrouvé positif par les deux trousse à la dilution au 1/9. Pour la dilution au 1/3, six échantillons faiblement positifs sur 16 demeuraient positifs (37,5%) pour la trousse A, contre 15 sur 32 (46,9%) pour la trousse B. Les échantillons avec une forte valeur S/P initiale étaient, à une exception près observée avec la trousse A, toujours positifs à la dilution au tiers. En dilution au 1/9; 40,0% (16/40) de ces échantillons demeurent positifs en utilisant la trousse A contre 68,7% (11/16) pour la trousse B. A la dilution au 1/81, aucun échantillon n'était positif aux seuils définis par les fabricants.

Une forte réduction des valeurs S/P a été observée dès la dilution au 1/3, indiquant

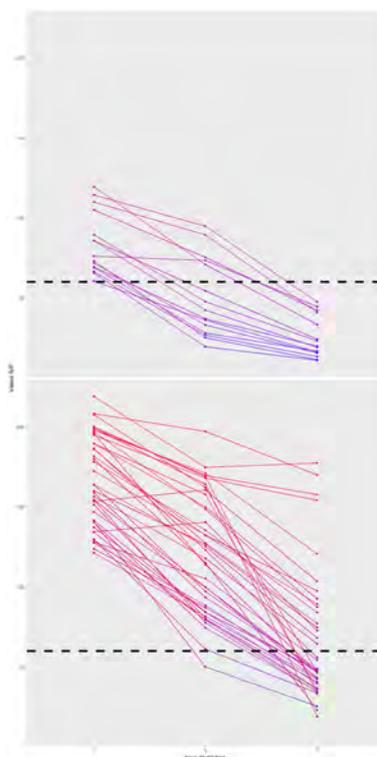


FIGURE 20 – Résultats des dilutions sériées pour le kitID Screen Paratuberculosis Indirect®

Les sérums sont regroupés en fonction de la valeur S/P comme défini au tableau 6
Le seuil de positivité est représenté par une ligne pointillée.

l'obligation de modifier les seuils de décision pour les deux trousse. La trousse A semblait par ailleurs moins adaptée à une utilisation sur des mélanges d'échantillons, puisque les sérums fortement positifs demeuraient positifs uniquement dans 40% des cas à la dilution au 1/9.

La deuxième partie des travaux expérimentaux avait pour objectif de définir de nouveaux seuils de décision pour des mélanges de sérums de taille 5 et 10, et de déterminer les performances diagnostiques relatives associées. Contrairement à l'expérience de dilutions sériées, dans laquelle un seul sérum négatif faible a été utilisé comme diluant, les mélanges étaient ici composés de différents sérums individuels issus du même élevage pour chaque mélange, de valeurs S/P variables mais connues. L'objectif était de déterminer les seuils de décision maximisant le couple $Se_D - Sp_D$ en fonction de la composition du mélange. Contrairement aux analyses de mélanges de fèces, la problématique de la spécificité se pose pour les analyses de mélanges de sérums. En effet, les interactions entre échantillons de sérums au sein d'un même mélange pourraient conduire à une valeur

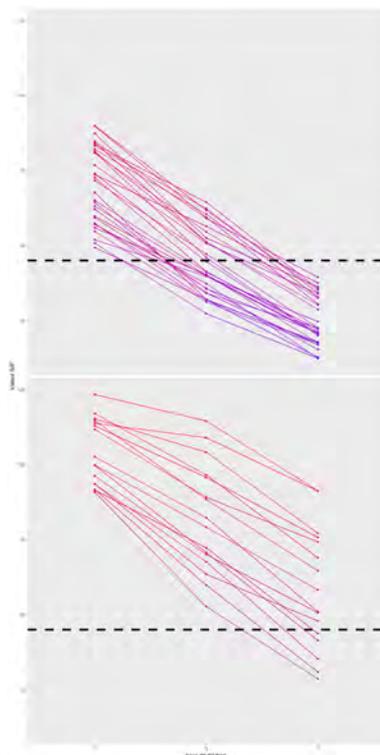


FIGURE 21 – Résultats des analyses en dilutions sériées pour la trousse IDEXX paratuberculosis screening®

Les sérums sont regroupés en fonction de la valeur S/P comme défini au tableau 6
Le seuil de positivité est représenté par une ligne pointillée.

S/P du mélange considérée comme positive si un seuil de décision trop bas était sélectionné. Afin d'estimer la Sp_D relative cette expérience comportait un certain nombre de mélanges composés d'un ou plusieurs échantillons négatifs dont les valeurs S/P étaient relativement élevées (échantillons négatifs forts).

Les résultats sont présentés dans le manuscrit en cours de préparation pour la trousse IDEXX paratuberculosis screening®. Les résultats obtenus avec la trousse ID Screen Paratuberculosis Indirect®, lot 800, IDVET, sont présentés ci-dessous dans le tableau 7 pour les mélanges de taille 5, et le tableau 8 pour les mélanges de taille 10. Pour la trousse IDVET, la Sp_D relative est restée parfaite pour des mélanges d'échantillons LN jusqu'au seuil de 25% quelle que soit la dilution. La présence d'un échantillon HN dans le mélange ne diminuait pas la Sp_D relative au seuil de 25%. En revanche, la présence de deux échantillons HN dans un mélange de 5 échantillons semblait dégrader beaucoup la Sp_D relative du mélange jusqu'au seuil de 30%. Les résultats pour ce type de mélanges sont cependant basés sur des effectifs très faibles ($n=6$) et apparaissent incohérents avec ceux obtenus

avec la trousse IDEXX. Malheureusement ils n'ont pas pu être confirmés ou corrigés, car le nombre et la quantité de sérums disponibles pour faire cette vérification étaient insuffisants. Sans tenir compte des mélanges contenant deux échantillons négatifs forts qui apparaissent étonnants, le seuil de décision de 25% apparaît optimal pour les mélanges de taille 5. Il est en effet associé à une Sp_D relative et à une Se_D relative parfaites pour les mélanges contenant un échantillon HP, ou deux échantillons LP. Pour les mélanges contenant un seul échantillon LP, ce seuil correspond à une Se_D relative de 78,0% (32/41 mélanges testés). Pour les mélanges de taille 10, le seuil de décision de 25% est associé à des performances moins satisfaisantes en termes de Se : seuls 18/43 mélanges contenant un échantillon faible positif (41,9%) et 37/41 mélanges contenant un échantillon positif fort (90,2%) étaient détectés positifs à ce seuil. Le seuil de décision de 20% améliorerait légèrement la Se_D relative, mais est associé à une Sp_D relative dégradée.

Tableau 7 – Résultats des analyses de mélanges de sérums de taille 5 (trousse ID Screen Paratuberculosis Indirect[®], lot 800, IdVet)

Composition du mélange*	# testés	# positif au seuil S/P (%)					
		15%	20%	25%	30%	35%	60%
5 LN	32	2	0	0	0	0	0
1 HN	15	1	0	0	0	0	0
2 HN	6	6	6	5	3	0	0
1 LP	41	40	36	32	30	26	10
2 LP	8	8	8	8	8	8	5
1 HP	38	38	38	38	38	37	30

* LN : $S/P < 30\%$; HN : $30.0\% \leq S/P < 60.0\%$; LP : $60 \leq S/P < 120\%$ and HP : $S/P \geq 120\%$.

Tableau 8 – Résultats des analyses de mélanges de sérums de taille 10 (trousse ID Screen Paratuberculosis Indirect[®], lot 800, IdVet)

Composition du mélange*	# testés	# positif au seuil S/P (%)					
		15%	20%	25%	30%	35%	60%
5 LN	43	1	0	0	0	0	0
1 HN	24	1	1	0	0	0	0
2 HN	6	5	2	0	0	0	0
1 LP	43	27	21	18	13	10	3
2 LP	8	8	8	8	7	4	3
1 HP	41	39	38	37	34	36	25

* LN : S/P < 30%; HN : 30.0% ≤ S/P < 60.0%; LP : 60 ≤ S/P < 120% and HP : S/P ≥ 120%.

3 Conclusions du chapitre

Performances relatives des analyses de mélanges de sérums ou de fèces :

- Pour des mélanges de 20 échantillons, la qPCR sur mélanges de fèces a une Se_D relative quasiment parfaite si les mélanges contiennent des fèces d'ovins forts excréteurs;
- La détection des mélanges contenant des fèces d'ovins faibles excréteurs est plus faible mais reste acceptable pour les mélanges de 5 ou 10 échantillons;
- La sérologie ELISA appliquée à des mélanges de sérums nécessite la redéfinition des seuils de positivité associés à une Se_D relative intéressante, quoique non parfaite pour certaines compositions de mélanges.

Performances diagnostiques des analyses de mélange de sérums ou de fèces à l'échelle des troupeaux :

- Les spécificités diagnostiques individuelles non parfaites des deux trousse ELISA étudiées rendent leur utilisation difficile dans le cadre d'une stratégie de dépistage à grande échelle.
- La qPCR appliquée à des mélanges de fèces semble plus adaptée à une utilisation à grande échelle pour le dépistage de la paratuberculose ovine, malgré la nécessité de réaliser des prélèvements sur un grand nombre d'animaux pour détecter les troupeaux à faible prévalence d'infection et une estimation de la prévalence d'infection qui ne peut pas être extrêmement précise.

Chapitre 5

Maîtrise de la paratuberculose par la vaccination

1 Effets de la vaccination vis-à-vis de la paratuberculose en élevages ovins allaitants : premiers résultats en France

Nous avons réalisé une étude transversale dans des troupeaux ovins de race Causse du Lot où la paratuberculose était enzootique, comprenant neuf troupeaux vaccinés et 10 troupeaux non vaccinés. Cette étude a permis (i) d'évaluer les conséquences de la vaccination sur la réponse sérologique chez les animaux vaccinés (ii) d'évaluer et de comparer les prévalences apparentes d'infection pré et post-vaccination dans les troupeaux vaccinés et (iii) d'évaluer et de comparer les fréquences et les niveaux d'excrétion individuels dans les troupeaux vaccinés et les troupeaux non vaccinés. Les résultats de cette étude ont fait l'objet de plusieurs communications, sous forme d'un poster lors de l'ICP 2016 (annexe III), de présentations orales, dont une lors des Rencontres autour des Recherches sur les Ruminants en 2015, qui a fait l'objet d'un article rédigé présenté à la suite.

La vaccination contre la paratuberculose ovine : développements méthodologiques et évaluation de son efficacité

MATHEVON Y. (1), LAFORT M-P. (2), TASCA C. (1), GALAN E. (3), FALGUIERES R. (3), CORBIERE F. (1,2), FOUCRAS G. (1,2)

(1) INRA-ENVT, UMR 1225 IHAP, 31076 Toulouse, France

(2) ENVT, 23 chemin des Capelles, 31076 Toulouse, France

(3) CAPEL Ovilot, 237 avenue Pierre Semard, 46000 Cahors, France

RESUME

Afin d'évaluer l'efficacité et la pertinence de la vaccination contre la paratuberculose, nous avons réalisé une étude portant sur 19 cheptels ovins dans lesquels l'infection par *Mycobacterium avium* subsp. *paratuberculosis* (*Map*) était démontrée. Parmi ces 19 élevages, les agnelles de renouvellement de 9 d'entre eux ont été vaccinées tous les ans depuis 2009, les 10 autres troupeaux n'ayant pas eu recours à la vaccination. L'excrétion fécale et la réponse sérologique vis-à-vis de *Map* ont été évaluées respectivement par PCR quantitative et ELISA anticorps. Pour cela, des prélèvements de fèces et de sérum ont été réalisés sur les cohortes nées entre 2010 et 2013. Une analyse préliminaire des résultats sérologiques dans 5 élevages vaccinés montre une réponse anticorps très élevée chez la quasi-totalité des animaux, sans décroissance avec l'âge. La sérologie n'apparaît donc pas utilisable pour le suivi du statut infectieux vis-à-vis de *Map* dans les troupeaux vaccinés. Les résultats de qPCR une réduction, mais sans annulation, de la prévalence d'excrétion, suggérant un effet bénéfique de la vaccination sur la contamination de l'environnement. La mise au point d'une méthode pour qualifier le statut des élevages vis-à-vis de la paratuberculose permettra de rationaliser les choix en matière de vaccination.

Ovine Johne's disease vaccination: Methodological developments and efficiency measurement

MATHEVON Y. (1), LAFORT M-P. (2), TASCA C. (1), GALAN E. (3), FALGUIERES R. (3), CORBIERE F. (1,2), FOUCRAS G. (1,2)

(1) INRA-ENVT, UMR 1225 IHAP, 31076 Toulouse, France

SUMMARY

To assess ovine Johne's disease vaccination efficiency and relevance we conducted a study on 19 ovine flocks in which *Map* infection was already demonstrated. Nine of the 19 herds had vaccinated replacement female lambs each year since 2009, the other half did not change their control practices. Fecal shedding and serological response were evaluated respectively by quantitative PCR and ELISA. Faeces and serum samples were collected on cohorts born between 2010 and 2013. A preliminary analysis of ELISA's results in 5 vaccinated herds shows that the immune response was high in almost all vaccinated sheep, without any decrease with ageing, proving that serology is useless to track evolution of *Map* infection in vaccinated herds. qPCR's results suggest a decrease of fecal shedding prevalence, indicating that vaccination may contribute to reduce environmental contamination. The development of a method to qualify the status of herds will streamline vaccination's choices

INTRODUCTION

La paratuberculose est une maladie infectieuse chronique provoquée par *Mycobacterium avium* subsp. *paratuberculosis* (*Map*) qui affecte toutes les espèces de ruminants domestiques. L'absence de traitement, l'impossibilité d'obtenir un diagnostic de certitude et la longue période d'incubation rendent le contrôle de cette maladie difficile. Jusqu'alors les moyens de gestion se limitaient à des mesures de biosécurité (exclusion des animaux suspects ou cliniquement atteints) rarement mises en œuvre en élevage ovin. La vaccination, devenue possible à l'aide d'un vaccin inactivé (SILIRUM® pour les bovins, GUDAIR® pour les ovins et caprins), offre une alternative nouvelle et attractive. Cependant une évaluation de son efficacité en termes bactériologiques n'a pas encore été réalisée dans les élevages ovins français. Dans cette étude, nous avons cherché d'une part à mettre en évidence l'efficacité de la vaccination contre la paratuberculose en cheptels ovins et d'autre part à développer une méthode de suivi de l'infection dans les élevages vaccinés.

1 MATERIEL ET METHODES

1.1 POPULATION ETUDIEE

Les 61 éleveurs adhérents à l'organisation de producteurs Ovilot (Lot, France) constituent la population étudiée. L'échantillon d'étude est représenté par 19 élevages sélectionneurs (en race Causse du Lot) dans lesquels l'infection par *Map* était démontrée (sérologies positives ou résultats d'autopsie évocateurs confirmés par PCR et/ou histologie). La taille des troupeaux constituant l'échantillon d'étude varie de 500 à 1900 brebis adultes (moyenne 893±386). Dans 9 de ces élevages, généralement les plus affectés (mortalité, réformes précoces) une vaccination avec le vaccin Gudair® a été mise en œuvre depuis 2009 (millésime de naissance 2010). Le protocole vaccinal consiste en une injection sous-cutanée d'une dose entre 4 et 6 mois d'âge sur l'ensemble des agnelles de renouvellement chaque année.

1.2 DONNEES COLLECTEES

1.2.1 Sondage sérologique préalable à la vaccination
 Entre 2009 et 2014, un sondage sérologique annuel a été réalisé dans les élevages suivis. Après la mise en place de la vaccination, seuls les élevages non vaccinés étaient concernés. Le sérum d'une cinquantaine d'animaux de 3 ans était prélevé dans chaque élevage. Au total 2791 résultats individuels étaient disponibles dans les 19 élevages suivis permettant une évaluation de la prévalence sérologique observée vis-à-vis de *Map*.

1.2.2 Sondage sérologique et excrétion fécale post-vaccinale

Dans les 9 élevages vaccinés, un objectif de 50 animaux prélevés dans les classes d'âge de 1 à 4 ans a été fixé (n=1589 au total). Dans les 10 élevages non vaccinés, seuls les ovins de 3 ans ont été prélevés, éventuellement complétés par des animaux de 2 ans afin d'atteindre 100 brebis par élevage (n=906 au total). Chaque animal inclus a fait l'objet d'une prise de sang sur tube sec et d'un prélèvement de fèces directement dans le rectum à l'aide de gants à usage unique. Après identification, les prélèvements ont été congelés à -20°C jusqu'à leur analyse.

1.3 ANALYSE BIOLOGIQUES ET TRAITEMENT DES DONNEES

1.3.1 Analyses sérologiques

Les analyses sérologiques de sondage avant la vaccination ont été réalisées par le Laboratoire Départemental d'Analyses Vétérinaires du Lot à l'aide du kit Idexx Paratuberculosis Screening Ab Test®. La limite de positivité est fixée, par le fabricant, à une valeur S/P égale à 55%. Les analyses sérologiques post-vaccinales ont été réalisées à l'UMR 1225 IHAP à l'aide du kit ID Screen Paratuberculosis Indirect® (ID.Vet, Montpellier, France). La limite de positivité est fixée, par le fabricant, à une valeur S/P égale à 70%.

1.3.2 qPCR quantitative sur fèces

L'extraction d'ADN a été réalisée à partir de 10 g de fèces à l'aide du kit d'extraction (Macherey Nagel). La quantification à l'aide d'Adiavet Paratb Realtime® a été réalisée par comparaison à une gamme de 4 points de concentration de 10⁴ à 10¹ équivalents génomes sur chaque plaque préparée à partir d'ADN purifié de *Map*, ainsi qu'un contrôle positif. Le seuil de positivité a été fixé à 40 cycles de PCR, correspondant à la détection d'une copie d'ADN dans la réaction, c'est-à-dire, en moyenne, à une charge d'excrétion de 2.3 bactéries par gramme de fèces.

1.3.3 Analyses statistiques

Afin d'évaluer l'effet de la vaccination sur la persistance de la réponse sérologique, la prévalence observée des ovins séropositifs (rapport S/P >70%) ainsi que la valeur moyenne du rapport S/P ont été calculées et comparées pour chaque cohorte de naissance dans 4 élevages des 9 élevages vaccinés ayant fait l'objet de d'analyses sérologiques.

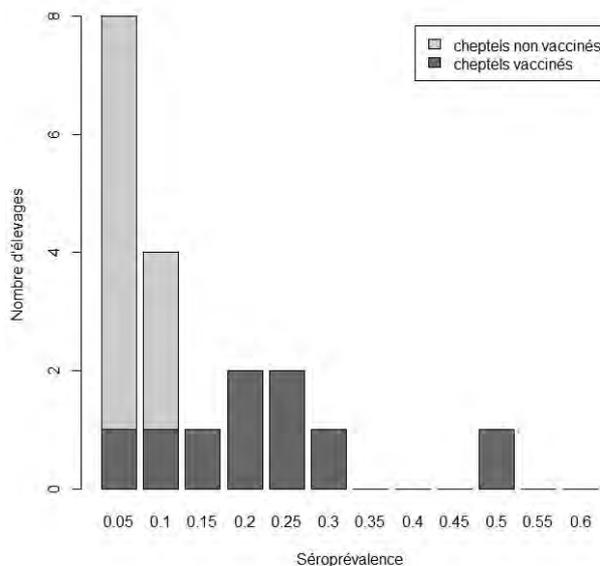
Afin d'évaluer de manière indirecte l'efficacité de la vaccination, nous avons effectué des comparaisons de prévalences sérologiques et des quantités de *Map* excrétées observées, entre les troupeaux vaccinés et non vaccinés avant et après la mise en place de la vaccination. La situation initiale a été établie à partir de la séroprévalence observée calculée avec les données du sondage sérologique. Au vu de la réponse sérologique post-vaccinale (cf infra), nous avons utilisé les résultats d'excrétion pour comparer les deux populations de troupeaux. Les analyses statistiques ont été réalisées avec le logiciel R.

2 RESULTATS

2.1 SITUATION DES CHEPTELS PREALABLEMENT A LA VACCINATION

La séroprévalence observée estimée à partir des données du sondage sérologique variait de 0 à 46% (médiane = 7,8%). La séroprévalence observée dans les troupeaux ayant été ultérieurement vaccinés était significativement plus élevée que celle des cheptels n'ayant pas mis en place la vaccination (test de Mann-Whitney-Wilcoxon p<0,001) (Figure 1).

Figure 1 : Dispersion des séroprévalences observées avant vaccination (résultats des sondages sérologiques)



2.2 SEROLOGIE POST-VACCINALE

Dans un premier temps, les analyses sérologiques ont été réalisées dans 4 des 9 troupeaux vaccinés. La séroprévalence observée chez les ovins vaccinés était très élevée, comprise entre 92% et 100% selon les cohortes de naissance et les élevages. Aucune différence significative du niveau de séroprévalence observée n'a pu être mise en évidence entre élevages ou entre cohortes de naissance. Dans tous les élevages, le rapport S/P des animaux vaccinés était très largement supérieur au seuil de positivité de 70% (moyenne = 158%) et aucune décroissance significative n'a pu être mise en évidence en fonction de l'âge des animaux (Figure 2). Ces résultats suggèrent une persistance forte et durable de la réponse sérologique post-vaccinale. A la vue de ces résultats, il a été décidé de ne pas réaliser (pour des raisons de coût) les analyses sérologiques sur les 5 autres élevages vaccinés.

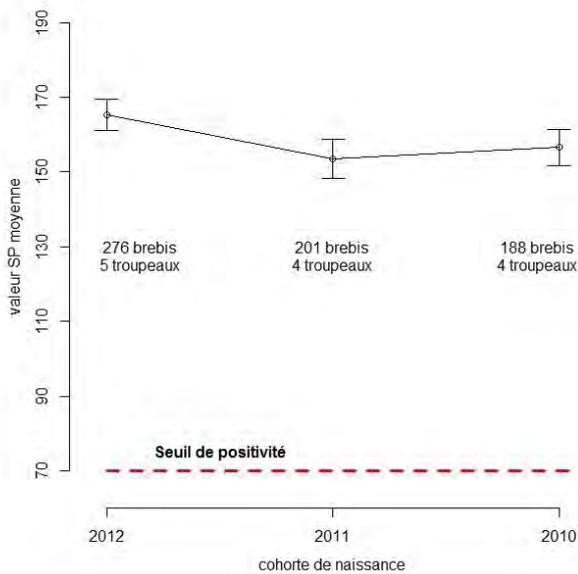
2.3 EXCRETION DE MAP

2.3.1 Troupeaux non vaccinés

La prévalence observée dans la population ne vaccinant pas variait de 0 à 30% d'animaux excréteurs par troupeau (Figure 4). Sur les 75 brebis positives en qPCR dans ces troupeaux, la valeur médiane des charges d'excrétion estimées était de 13.2 bactéries/gramme de fèces (1^{er} quartile : 5.4 ; 3^{ieme} quartile : 60.5).

1. EFFETS DE LA VACCINATION : PREMIERS RÉSULTATS EN FRANCE

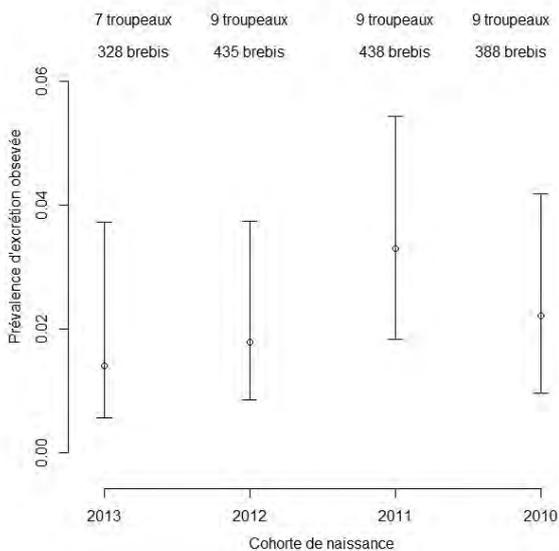
Figure 2 : Valeur S/P par cohorte de naissance dans les troupeaux vaccinés



2.3.2 Troupeaux vaccinés

La prévalence d'excrétion observée dans les troupeaux vaccinés variait de 1 à 7%. Les prévalences d'excrétion observée étaient globalement faibles ($\leq 4\%$). La prévalence d'excrétion n'était par ailleurs pas significativement différente entre les cohortes de naissance (test $\text{Chi}^2=0.39$) (Figure 3). Sur les 35 brebis positives en qPCR dans ces troupeaux, la valeur médiane des charges d'excrétion estimées était de 7.6 bact/gramme de fèces (1^{er} quartile : 4.0 ; 3^{ème} quartile : 31.4).

Figure 3 : Prévalence d'excrétion de *Map* observée par cohorte de naissance dans les troupeaux vaccinés



2.3.3 Comparaison en fonction du statut vaccinal

La prévalence d'excrétion observée n'est pas différente entre troupeaux vaccinés et troupeaux non vaccinés (test de Mann-Whitney-Wilcoxon $p=0,17$). De même le niveau d'excrétion individuel estimé n'est pas différent entre les brebis issues d'élevages vaccinés et ne vaccinant pas (Test de Mann-Whitney Wilcoxon $p=0,15$). La figure 5 montre la dispersion

des animaux positifs en qPCR en fonction de leur niveau d'excrétion estimé. Nous retrouvons un profil similaire dans les troupeaux vaccinés ou ceux ne vaccinant pas avec environ la moitié des brebis faiblement excrétrices (moins de 10 bactéries par gramme de fèces). Les brebis fortes excrétrices (plus de 10000 bactéries par grammes de fèces) représentent 5% des individus excréteurs.

Figure 4 : Prévalence d'excrétion de *Map* observée et intervalle de confiance à 95%

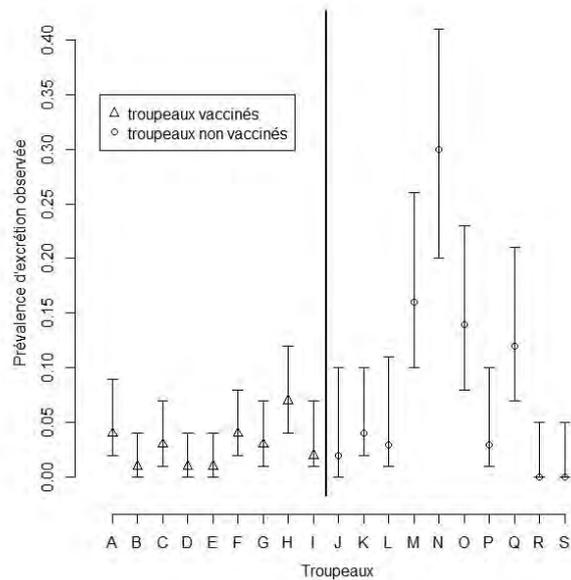
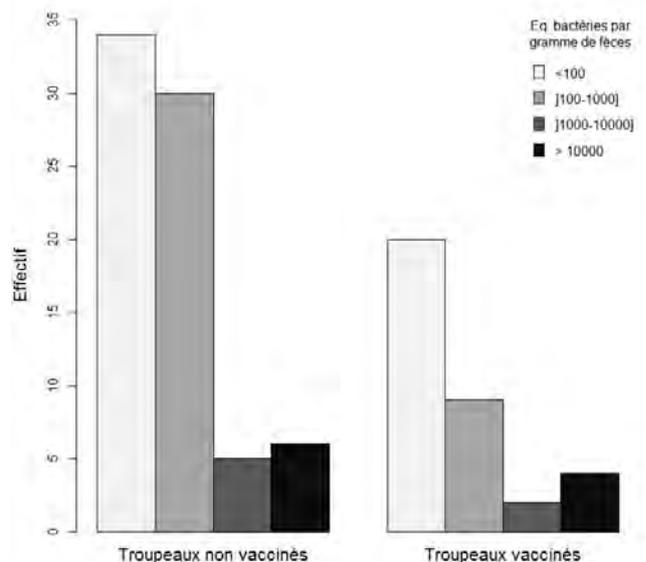


Figure 5 : Dispersion des niveaux d'excrétion en fonction du statut du troupeau



3 DISCUSSION ET CONCLUSION

Le premier objectif de cette étude était d'évaluer la pertinence de l'utilisation de la sérologie pour le suivi des élevages ovins vaccinés contre la paratuberculose. Nos résultats indiquent clairement une persistance à un niveau élevé et durable de la réponse sérologique post-vaccinale, rendant l'utilisation de

l'outil sérologique non pertinente pour le suivi des nouvelles infections dans les troupeaux vaccinés. Ces résultats diffèrent quelque peu des observations d'autres auteurs (Reddacliff et al. 2006; Dhand et al. 2013; Windsor et al. 2014; Bastida et Juste 2011) selon lesquelles la séroprévalence baisse fortement après 2 ans chez les animaux vaccinés. La stabilité de la réponse sérologique et le maintien d'un rapport S/P élevé au cours du temps pourrait s'expliquer par le protocole de vaccination mis en place par Ovilot, où les agnelles ont été vaccinées entre 4 et 6 mois d'âge, soit entre 1 et 3 mois plus tard que pour les animaux inclus dans les études déjà publiées. Il pourrait en résulter une meilleure stimulation du système immunitaire expliquant les différences de niveau et de persistance des réponses anticorps post-vaccinales observées.

Nous avons donc utilisé la qPCR pour répondre au second objectif de notre étude qui était d'évaluer l'efficacité de la vaccination vis-à-vis de l'excrétion fécale de *Map*. Cette méthode reste utilisable en milieu vacciné, rapide d'application et plus sensible que la culture sur fèces congelés (Kawaji et al. 2011); elle fournit en outre des résultats quantitatifs. L'analyse de 4 cohortes de naissance n'a pas montré de classes d'animaux à privilégier pour assurer le suivi des élevages vaccinés. A partir de deux populations d'élevages distinctes, qui présentent une séroprévalence significativement différente avant la mise en place de la vaccination, nos résultats indiquent que la prévalence d'excrétion n'était pas différente entre les élevages non vaccinés (faiblement infectés) et ceux vaccinés depuis 5 ans (ayant initialement des séroprévalences significativement plus élevées). En faisant l'hypothèse qu'une séroprévalence initiale plus élevée dans les troupeaux ayant mise en œuvre la vaccination correspondait à une prévalence ou un niveau d'excrétion aussi plus élevés, et en l'absence d'autres mesures de maîtrise, ceci suggère donc une réduction de la prévalence d'excrétion après la mise en place de la vaccination. Ce schéma d'analyse, imposé par les données disponibles (absence de données d'excrétion avant la mise en place de la vaccination) repose cependant sur l'hypothèse selon laquelle la prévalence d'excrétion et la séroprévalence à l'échelle d'un élevage sont corrélées positivement. Une des prochaines étapes dans notre étude sera d'établir la forme de cette relation. Nos observations suggérant un bénéfice lié à la vaccination sont en accord avec le ressenti des éleveurs qui témoignent d'une forte réduction de la mortalité et des réformes précoces liées à la paratuberculose. Cependant malgré des prévalences faibles, des animaux excréteurs, dont certains à des niveaux élevés, ont été retrouvés dans tous les élevages vaccinés, supportant les résultats d'autres études (Reddacliff et al. 2006; Dhand et al. 2013) qui ont montré une réduction, mais pas d'extinction, de l'excrétion fécale de *Map* à la suite de la vaccination.

Dans ce contexte il apparaît essentiel, pour la maîtrise de la paratuberculose ovine dans les troupeaux infectés, de maintenir une vaccination sur le long terme, même après réduction ou disparition des troubles cliniques.

Le diagnostic de la paratuberculose repose sur des tests ayant une bonne spécificité (comprise entre 95 et 99%) avec une sensibilité médiocre et très variable selon l'espèce et la cible recherché (Nielsen et Toft 2008). La prévalence observée sous-estime la prévalence réelle, empêchant une description précise de la situation des élevages d'Ovilot. Pour les mêmes raisons, il ne nous est pas possible de quantifier directement l'effet de la vaccination. L'évaluation des tests utilisés pour déterminer leurs caractéristiques nous permettra non seulement d'affiner notre estimation de la prévalence mais également d'analyser l'impact économique de la vaccination. Le coût d'une stratégie de gestion étant un frein majeur à sa mise en place, il est nécessaire d'évaluer le bénéfice de la vaccination retiré par les éleveurs, en cherchant également à diminuer les coûts liés au suivi.

Nous tenons à remercier les éleveurs d'Ovilot, les techniciens d'élevage, les vétérinaires et les étudiants vétérinaires ayant contribué à la réalisation de cette étude. Cette étude bénéficie du financement conjoint de l'INRA (Métaprogramme GISA, projet PICSAR) et de la région Midi-Pyrénées (projet PAROVLOT).

- Bastida, Felix, et Ramon A. Juste. 2011.** *Journal of immune based therapies and vaccines* 9 (1): 8.
- Dhand, Navneet K., Wesley O. Johnson, Jeff Eppleston, Richard J. Whittington, et Peter A. Windsor. 2013.** *Preventive Veterinary Medicine* 111 (1-2): 81-91.
- Kawaji, Satoko, Douglas J. Begg, Karren M. Plain, et Richard J. Whittington. 2011.** *Veterinary Microbiology* 148 (1): 35-44.
- Nielsen, S.S., et N. Toft. 2008.** *Veterinary Microbiology* 129 (3-4): 217-35.
- Reddacliff, L, J Eppleston, P Windsor, R Whittington, et S Jones. 2006.** *Veterinary Microbiology* 115 (1-3): 77-90.
- Windsor, Pa, J Eppleston, Nk Dhand, et Rj Whittington. 2014.** *Australian Veterinary Journal* 92 (7): 263-68.

2 Suivi de la prévalence d'infection dans les troupeaux vaccinés par utilisation de la qPCR appliquée à des mélanges de fèces

Les résultats de notre étude transversale dans des élevages vaccinés sont en accord avec ceux publiés dans des études réalisées dans des élevages australiens (Dhand et al., 2013; 2016). Ils décrivent la persistance d'animaux excréteurs, certes en nombre réduit, mais ayant des niveaux d'excrétion fécale semblables à ceux mesurés dans les élevages non vaccinés. D'autre part notre étude démontre clairement l'impossibilité d'utiliser un test sérologique ELISA pour le suivi de la paratuberculose en élevage vacciné, et ce quel que soit l'âge des animaux.

Les résultats présentés au chapitre 4 permettent d'envisager le suivi des élevages vaccinés par qPCR sur mélanges de fèces. En utilisant les mêmes distributions des différents niveaux d'excrétion pour les animaux excréteurs et les mêmes estimateurs de performances diagnostiques des analyses de mélanges, nous avons simulé un suivi annuel basé sur des analyses à l'aide de la qPCR de mélange de fèces pendant cinq ans dans des troupeaux vaccinés. Ces simulations nous ont permis d'évaluer la capacité *cumulative* d'un suivi longitudinal à mettre en évidence des prévalences de plus en plus faibles. La figure 22 donne le nombre de mélanges positifs en qPCR obtenu au bout de trois ans dans le cadre d'un suivi annuel en réalisant 10 mélanges de 5 échantillons pour différents niveaux de prévalence d'infection intra-troupeau. Les résultats indiquent notamment que si aucun mélange positif n'a été détecté au bout de trois ans, nous pouvons affirmer avec une probabilité de 95% que l'élevage a une prévalence d'infection inférieure ou égale à 2,5%. De même si un ou deux mélanges positifs sont détectés durant ces trois années, nous pouvons affirmer avec une probabilité de 95% que l'élevage a une prévalence d'infection inférieure à 10%.

Les résultats au bout de cinq ans de suivi sont plus discriminants entre les élevages dont la prévalence est différente. Nous avons choisi de comparer deux stratégies impliquant le même nombre de mélanges testés chaque année, basées sur 50 animaux (échan-

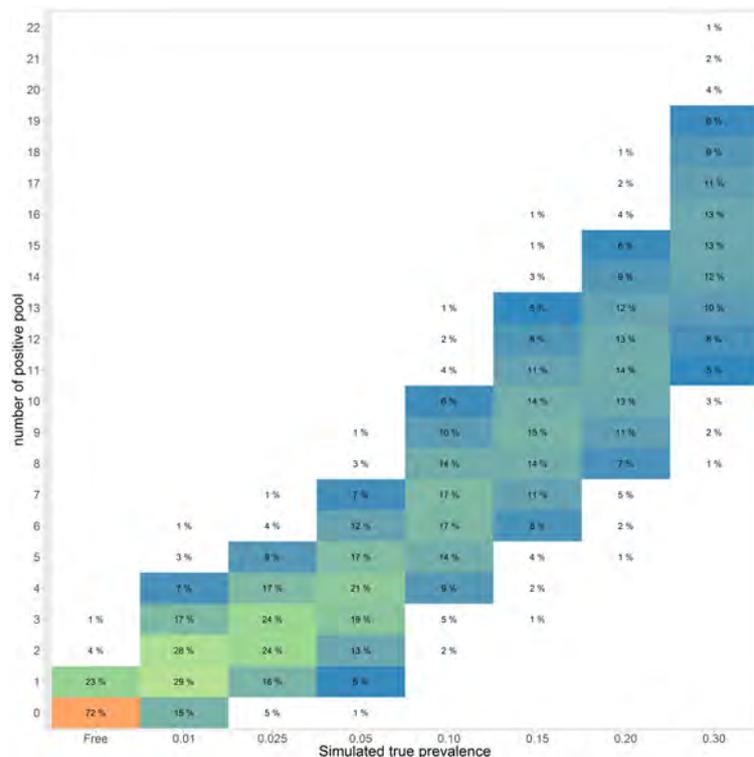


FIGURE 22 – Nombre de mélanges de fèces détectés par qPCR après un suivi annuel d’une durée de trois ans et de 50 animaux (10 mélanges de cinq échantillons)

Le nombre de mélanges positifs détectés est représenté en abscisse pour des élevages indemnes jusqu’à une prévalence d’infection de 30%. Pour une prévalence d’infection donnée, les pourcentages indiqués correspondent à la proportion d’élevages pour lesquels le nombre de mélanges positifs indiqué a été détecté.

tillonnage minimum requis dans le cadre de la prophylaxie obligatoire de la brucellose en France), et sur 100 animaux. Les résultats sont présentés dans la figure 23.

Le suivi par qPCR sur 10 mélanges de 10 animaux permet d’obtenir une meilleure estimation de la prévalence intra-troupeau après cinq ans. Dans le but d’identifier les élevages indemnes, ce suivi permet d’affirmer que si moins de deux mélanges positifs sont détectés, nous pouvons être sûrs à 95% que la prévalence réelle est inférieure à 1%. Dans le cadre d’un suivi annuel de 50 animaux pendant cinq ans nous pouvons être sûrs à 95% que la prévalence réelle est inférieure à 1% uniquement si aucun mélange n’a été détecté au cours des cinq ans. Sachant que les coûts liés au prélèvement sont négligeables par rapport au coût analytique de la qPCR, le suivi annuel de 100 animaux analysés en 10 mélanges de 10 échantillons semble pouvoir fournir des informations plus précises pour le suivi de la paratuberculose en élevage qu’il soit vacciné ou non.

Ces simulations reposent cependant sur plusieurs hypothèses qui sont difficiles à véri-

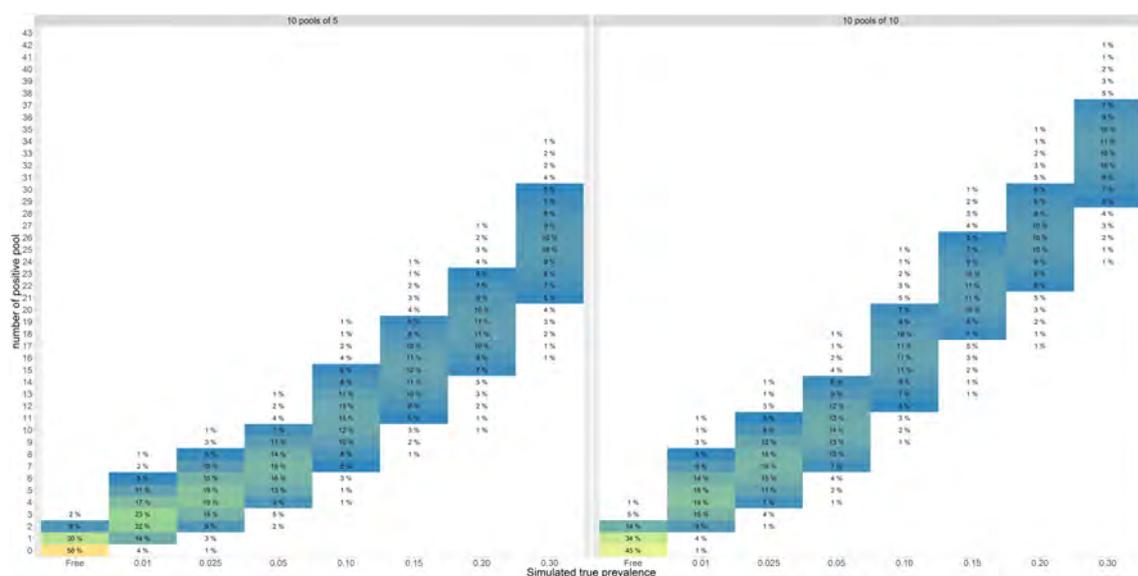


FIGURE 23 – Nombre de mélanges de fèces détectés par qPCR après un suivi annuel de cinq ans sur 50 (10 mélanges de cinq échantillons) ou 100 animaux (10 mélanges de 10 échantillons)

Le nombre de mélanges positifs détectés est représenté en abscisse pour des élevages indemnes jusqu'à une prévalence d'infection de 30%. Pour une prévalence d'infection donnée, les pourcentages indiqués correspondent à la proportion des élevages pour lesquels le nombre de mélanges positifs indiqué a été détecté.

fier. La première est que la prévalence d'infection est constante au cours du temps et de la période de suivi. Ce protocole est donc uniquement applicable après plusieurs années de vaccination. La seconde hypothèse est que la distribution des niveaux d'excrétion parmi les animaux vaccinés excréteurs demeure la même que dans les troupeaux non vaccinés (environ 5% de forts excréteurs). Enfin la dernière hypothèse est que les performances diagnostiques à l'échelle individuelle de la qPCR sont les mêmes pour le dépistage des animaux infectés non vaccinés et infectés vaccinés. Nos observations réalisées dans les élevages ayant eu recours à la vaccination depuis plus de cinq ans sont en faveur de niveaux d'excrétion fécale semblables entre les animaux infectés non vaccinés et ceux qui sont vaccinés et infectés. Cependant s'il s'avérait que la vaccination conduit à une excrétion fécale intermittente, alors la sensibilité de l'analyse sur fèces serait moindre et les résultats des simulations présentés seraient trop optimistes.

3 Conclusions du chapitre

- La réponse anticorps du vaccin GUDAIR[®] est présente chez la quasi-totalité des animaux ayant reçu une injection, rendant donc la sérologie inopérante pour le suivi des troupeaux vaccinés, et ce quelque soit l'âge des animaux après la vaccination.
- Le suivi du statut sanitaire des troupeaux ovins vaccinés peut être réalisé par qPCR sur mélange de fèces.
- La réalisation d'un suivi annuel permet d'affiner l'estimation de la prévalence intra-troupeau résiduelle, rationalisant ainsi le recours à la vaccination.
- La stratégie d'analyse la plus performante, tout étant réalisable sur le terrain, est celle se basant sur l'analyse de 10 mélanges constitués de 10 échantillons de fèces.

Chapitre 6

Discussion générale et perspectives

Nous vivons sur une île placide d'ignorance au milieu des mers noires de l'infini et rien ne nous prédispose à voyager très loin.

H.P. Lovecraft

L'Appel de Cthulhu

Ce travail de thèse est le fruit d'une collaboration entre l'organisation de producteurs CAPEL/Ovilot et l'UMR 1225 IHAP pour l'étude des stratégies de lutte et de maîtrise de la paratuberculose dans les troupeaux allaitants ovins. Centrée autour des thématiques du dépistage de l'infection par *Map* et de la vaccination contre cette maladie, cette collaboration a permis d'étudier les performances diagnostiques des tests dans des conditions de terrain relativement homogènes, les élevages étant choisis parmi les élevages sélectionneurs du schéma pyramidal de production de la race Causse du Lot. Ce plan d'analyse permet d'assurer un transfert des connaissances produites vers les élevages et les organismes qui sont les utilisateurs finaux, la contrepartie étant que les résultats obtenus doivent être extrapolés avec précaution dans des conditions d'élevage qui s'écartent significativement de celles étudiées dans le but d'améliorer leur statut sanitaire.

La maîtrise d'une infection diffusant entre les élevages revient à connaître tout d'abord la prévalence de l'infection et oblige donc à une première étape de diagnostic/dépistage des animaux. Dans le cas de la paratuberculose, le diagnostic de certitude n'existe que pour les animaux malades, en s'appuyant sur la reconnaissance des signes cliniques et

la confirmation par un ou plusieurs examens de laboratoire. Du fait des piètres performances diagnostiques des tests utilisables du vivant de l'animal, un résultat négatif est entaché d'un risque d'erreur élevé sur le statut infectieux réel de l'animal. Les tests diagnostiques utilisés dans ce travail de thèse n'avaient pas été évalués dans l'espèce ovine en France, et afin de d'obtenir des estimateurs précis des sensibilités et spécificités des tests, la première étape a été de produire ces estimateurs. Notre échantillonnage s'appuyait sur des élevages sélectionneurs inscrits à l'organisation de producteurs Ovilot dans une zone géographique de taille réduite et bien délimitée (département du Lot). Les conduites d'élevages étaient sensiblement les mêmes dans ces élevages, à l'exception d'une différence forte, la prévalence de la paratuberculose. Celle-ci était approchée par la prévalence de la séropositivité observée lors du suivi sérologique annuel mis en place précédemment sans que les facteurs de risque associés à cette infection/maladie n'aient été recherchés. Dans tous les élevages visités, les brebis étaient sélectionnées en se basant sur deux critères : l'absence de signes cliniques et la cohorte de naissance à laquelle elles appartenaient, respectant les hypothèses nécessaires pour l'analyse statistique en classes latentes.

1 Estimation des performances diagnostiques à l'échelle individuelle

1.1 Méthodes d'estimation et principaux résultats

Nos estimations sont basées sur l'hypothèse que les performances diagnostiques des tests utilisés sont les mêmes quelle que soit la prévalence de la paratuberculose dans les différents élevages. Cette hypothèse, définie par (Hui and Walter, 1980) est largement utilisée dans le cadre de l'estimation des performances diagnostiques des tests (Kostoulas et al., 2017). Elle peut cependant être remise en cause du fait de la variation possible de la Se_D et de la Sp_D en fonction de la prévalence (Brenner and Gefeller, 1997). Elle pourrait par ailleurs être contournée en modélisant différemment les performances diagnostiques comme la somme d'une moyenne et d'un effet aléatoire attaché à chaque popu-

lation, comme proposé récemment par Bermingham et al. (2015). Néanmoins cette modélisation augmente le nombre de paramètres à estimer et nécessite donc d'augmenter le nombre de degrés de liberté disponibles pour rendre le modèle identifiable, ou bien de fixer un certain nombre de paramètres. Par ailleurs, l'utilisation de ce type d'approche requiert des données peu déséquilibrées, c'est-à-dire la présence de résultats positifs et négatifs en nombres suffisants dans chaque population. Les faibles prévalences apparentes et estimées dans les élevages investigués, avec globalement peu de résultats positifs, limitaient donc l'utilisation de telles approches. Par ailleurs, la vérification par la méthode de ré-échantillonnage Jackknife, utilisée dans l'article 1 présenté au chapitre 3, de l'absence d'influence notable d'une des populations ou de l'un des élevages sur les estimateurs des paramètres soutient l'hypothèse de performances diagnostiques comparables entre les populations où la prévalence est différente.

Les méthodes d'estimation utilisées dans ce travail reposent sur des résultats binaires (positif / négatif), nous obligeant par ailleurs à considérer les résultats douteux en sérologie ELISA, soit comme positifs, soit comme négatifs dans deux analyses séparées, selon les recommandations de Kostoulas et al. (2017). Le fait de remplacer une variable réponse continue (valeur S/P pour la sérologie ELISA, et valeur Ct pour la qPCR sur fèces) par une variable binaire représente une perte d'information. En effet, une valeur Ct faible en qPCR sur fèces correspond très vraisemblablement à un animal réellement infecté et n'est pas sujet à caution. A l'inverse, un Ct très tardif peut être attribué à des erreurs analytiques (défaut d'extraction, anomalie de l'amplification, présence résiduelle d'inhibiteurs...) aussi bien qu'à une très faible excrétion (excréteur passif, interruption temporaire de l'excrétion), rendant l'interprétation du résultat beaucoup plus hasardeuse. Ce même raisonnement peut être appliqué à la sérologie ELISA, les résultats douteux posant souvent plus de difficultés d'interprétation que les résultats avec une très faible ou à l'inverse, une très forte valeur S/P. La prise en compte d'une réponse quantitative, pondérant la confiance apportée aux résultats binaires, pourrait être une amélioration particulièrement utile lors de l'estimation des performances des tests diagnostiques visant des maladies à longue période d'incubation. Cette approche a été appliquée par Nielsen et al. (2007) pour l'analyse des résultats de sérologie ELISA sur le lait de vaches laitières, dans

un modèle de mélange bayésien incluant par ailleurs des variables explicatives (âge, saison). Les résultats obtenus sont en faveur d'une estimation plus précise de la prévalence réelle intra-troupeau. Cette approche n'a cependant pas été utilisée par d'autres auteurs et son intérêt reste à vérifier dans d'autres contextes. Là encore, cette méthode nécessite, pour fournir des intervalles de confiance étroits, un grand nombre d'animaux testés et de nombreux résultats positifs, et ne paraissait pas adaptée aux données recueillies dans les élevages du Lot.

Les approches statistiques bayésiennes et fréquentistes appliquées aux données des élevages infectés du Lot ont fourni des résultats proches en terme de corrélation entre les trois tests évalués. Quelle que soit l'approche utilisée, fréquentiste ou bayésienne, nos résultats montrent clairement que les hypothèses faites sur la structure de dépendance entre les tests influencent fortement les estimateurs obtenus. L'exemple le plus marquant est fourni par les résultats des modèles assumant une indépendance complète entre tous les tests, pour lesquels les estimateurs de performances diagnostiques étaient peu crédibles (sensibilités et spécificités des trousseaux ELISA très supérieures à celles de la qPCR). Comme indiqué précédemment, une mauvaise prise en compte de la structure de dépendance peut ainsi conduire à des estimateurs fortement biaisés (Vacek, 1985 ; Georgiadis et al., 2003 ; Albert and Dodd, 2004 ; Pepe and Janes, 2007). Une étape rigoureuse de sélection des modèles est donc essentielle dans ce type d'études portant sur l'évaluation des performances diagnostiques des tests en l'absence de test de référence parfait basée sur des modèles à classes latentes.

L'approche fréquentiste présentée au chapitre 3 fournit des résultats concordants avec une méthode bayésienne largement utilisée par la communauté scientifique et apporte des améliorations utiles au domaine de l'évaluation des performances diagnostiques des tests biologiques. Contrairement à l'approche bayésienne, la fonction d'estimation que nous proposons évalue en une seule étape les modèles reposant sur différentes structures de corrélation entre tests et les hiérarchise selon le critère de vraisemblance ou de l'AIC, et facilite donc l'étape de sélection du modèle. De plus dans le cas de modèles non identifiables, l'évaluation fréquentiste fournit l'ensemble des solutions permises par les données, les résultats obtenus pouvant alors servir comme distributions *a priori* à des

modèles bayésiens.

D'autre part, nous avons remarqué une différence pour les estimateurs proches des limites de l'espace des possibles. L'approche bayésienne ne semble pas en mesure de produire des estimateurs des paramètres égaux aux bornes de l'espace de probabilité (0 ou 1), contrairement à l'approche fréquentiste. Nous avons observé ce phénomène pour la Sp_D de la qPCR sur fèces, évaluée dans un modèle faisant l'hypothèse d'indépendance conditionnelle de la qPCR vis à vis des deux tests sérologiques à 100% (IC 95% : [99,6 - 100,0]) par l'approche fréquentiste (modèle 4) et à 99,4% (PCI 95% : [98,4 - 99,9]) par l'approche bayésienne (modèle 2). De manière étonnante, l'utilisation comme valeurs de départ de l'algorithme EM des estimateurs bayésiens conduisait systématiquement aux estimateurs fréquentistes, associés à une log-vraisemblance supérieure et donc à un meilleur ajustement aux données. De plus amples travaux, au-delà du cadre de ce travail de thèse, seraient nécessaires pour comprendre ces différences d'estimation entre les approches fréquentiste et bayésienne. Bien que ce modèle n'ait pas été retenu, ni par l'approche fréquentiste (sur la base du critère de vraisemblance), ni par l'approche bayésienne (sur la base du DIC), les différences observées sur les estimateurs de la Sp_D de la qPCR ne sont cependant pas sans conséquences pratiques. En effet, avec une Sp_D estimée à 99,0% par l'approche bayésienne, le risque de résultats faux positifs à l'échelle individuelle n'est pas négligeable, surtout si de nombreux animaux sont testés, alors que ce risque peut être considéré comme très faible voire nul, selon l'approche fréquentiste. Le modèle retenu, sur la base du meilleur ajustement aux données (modèle avec une dépendance conditionnelle entre les tests ELISA sur les paramètres de sensibilité et une dépendance conditionnelle entre les trois tests sur les paramètres de spécificité) fournit en revanche des estimateurs proches : 99,0% (IC 95% : [97,9 - 99,6]) pour l'approche fréquentiste et 99,0% (PCI 95% : [97,9 - 99,7]) pour l'approche bayésienne.

Au final les 2 trousse de sérologie ELISA et la qPCR sur fèces évaluées dans ce travail ont montré des performances diagnostiques médiocres pour l'identification des ovins en phase sub-clinique de l'infection par *Map*. Comme pour toute maladie avec une longue période d'incubation, le diagnostic précoce de ces animaux en phase d'incubation est un atout pour la maîtrise de l'infection en élevage. A l'heure actuelle, il n'existe pas de

test ante-mortem commercialisé, et dont le coût serait abordable, présentant des performances diagnostiques suffisamment grandes pour assurer cette identification.

1.2 Résultats comparés entre trousse ELISA et qPCR sur fèces

Les résultats obtenus pour les tests sérologiques ELISA et la qPCR sur fèces ont révélé des discordances entre les tests, comme nous l'avons présenté dans notre premier article. Seuls 30,5% (32 animaux) des 105 ovins avec une qPCR sur fèces positive avaient aussi une sérologie positive à au moins un des deux tests sérologiques. Les causes possibles de ces discordances qui n'ont pas été abordées dans la partie discussion de cet article, sont cependant intéressantes à évaluer. La première hypothèse permettant d'expliquer ces discordances est que la Se_D est plus faible pour les trousse ELISA comparée à la qPCR. Ces résultats sont en accord avec l'hypothèse d'une apparition tardive de la séroconversion lors de l'infection par *Map*, celle-ci apparaissant généralement après le début de l'excrétion fécale (Gillan et al., 2010 ; Clarke, 1997). Ces discordances pourraient par ailleurs être expliquées par le phénomène d'excrétion passive de *Map* (Sweeney et al., 1992b ; Moloney and Whittington, 2008). Cette possibilité est cohérente chez les animaux faiblement positifs en qPCR ($Ct > 40$) qui représentent tout de même 47% des résultats qPCR positifs de notre étude ($n=51$) car ils évoluent dans un environnement dont on sait qu'il est contaminé.

Le taux de discordance des animaux avec un résultat de qPCR négatif et au moins un résultat de sérologie ELISA positif était plus faible, du fait en partie du plus grand nombre de résultats négatifs (1090 animaux) dans notre population d'étude. Une petite fraction des animaux (6,7%, $n=73$) était cependant positive à au moins un test ELISA alors que le résultat de la qPCR était négatif. Ces résultats pourraient s'expliquer par une excrétion fécale intermittente de bacilles, en particulier chez les animaux en début d'évolution de la maladie, ou à la répartition inégale de *Map* dans les fèces des animaux excréteurs (Christopher-Hennings et al., 2003). Un défaut de Sp_D des tests ELISA ne peut là aussi être exclu.

Concernant plus précisément les résultats de sérologie ELISA, les deux trousse montrent également un taux de discordance élevé : 4,69% sur l'ensemble des animaux, mais 59% si

seuls les échantillons ayant donné au moins un résultat positif sont pris en compte. Le coefficient Kappa de Cohen, permettant de mesurer l'accord entre deux tests a été calculé entre ces deux trousse (coefficient Kappa : -0,23 ; IC 95% : [- 0,39 ; - 0,071]). Selon les règles d'interprétation proposées par Landis and Koch (1977), une valeur négative du coefficient Kappa de Cohen permet de conclure que les deux trousse ELISA sont en désaccord. Des discordances sont aussi rapportées dans l'espèce bovine (S. L. B. McKenna et al., 2005 ; Diéguez et al., 2009). Les différences entre les dispositifs d'étude et les séroprévalences observées rendent cependant difficiles les comparaisons avec nos résultats. Plusieurs facteurs peuvent expliquer les discordances entre tests provenant de fabricants différents, le plus probable étant une différence dans les antigènes utilisés pour sensibiliser les microplaques. Un grand nombre d'antigènes de natures différentes sont utilisés pour la fabrication des tests ELISA pour le diagnostic de la paratuberculose. Malgré l'absence de communication par les laboratoires d'informations sur la nature des antigènes utilisés dans les trousse utilisées ici, il est raisonnable de penser que les antigènes diffèrent entre les trousse.

Par ailleurs, les résultats d'un second lot de la trousse ELISA ID Screen Paratuberculosis Indirect[®] (lot de fabrication n° 800, résultats non montrés) ont été comparés avec le premier lot de cette même trousse (lot de fabrication n° 602). Le coefficient Kappa calculé était meilleur que celui obtenu entre les trousse IDvet et IDEXX (Kappa : 0,38, IC 95% : [0,2 - 0,57]) mais correspond toujours à un accord faible à moyen. Très peu de données sont disponibles sur les discordances existant entre les différents lots de fabrication d'une même trousse, et dans la plupart des études publiées, le lot de fabrication n'est pas rapporté. Dans une communication affichée proposée par le laboratoire IDvet au XIIe Colloque International sur la Paratuberculose à Parme en 2014 , un certain nombre d'éléments descriptifs sont fournis sur ces discordances inter-trousse et inter-lots de fabrication pour une même trousse. Dans cette étude, les auteurs présentent des résultats de sérologie ELISA obtenus à partir de trois trousse de fabricants différents (quatre lots de fabrication pour la trousse A, deux pour la trousse B et un lot pour la trousse C). Ainsi sur 33 bovins issus de deux troupeaux reconnus infectés, 18 seulement (54,5%) obtiennent des résultats concordants et positifs aux sept trousse évaluées. Pour la trousse A, quatre

échantillons sur 33 (12,1%) ont des résultats différents avec au moins un des quatre lots de fabrication. Pour la trousse B, seuls deux échantillons fournissent des résultats discordants entre les deux lots de fabrication. Des discordances peuvent aussi être observées à partir d'échantillons obtenus dans un contexte supposé indemne. Dans cette même étude, 648 bovins ont été testés dans six troupeaux considérés comme indemnes de paratuberculose et 38 animaux (5,9%) fournissent un résultat positif ou douteux avec au moins une des six ou sept trousse testées. Pour la trousse A, 16 échantillons sont discordants avec au moins un des quatre lots de fabrication (2,46%). Pour la trousse B, cinq échantillons sont discordants entre les deux lots de fabrication. Notons par ailleurs que sur 11 échantillons fournissant des résultats concordants positifs avec les deux lots de fabrication de la trousse B, seuls six obtiennent un résultat positif avec au moins un des lots de fabrication des autres trousse testées. Ainsi, même si le pourcentage de discordances entre lots de fabrication d'une même trousse apparaissait moins grand dans cette étude que dans la nôtre, leur fréquence apparaît donc comme non négligeable (entre 2 et 10% selon les trousse) et doit être pris en compte dans l'utilisation/l'interprétation des résultats d'analyse.

Ainsi les trois lots de fabrication de deux trousse différentes testés dans nos travaux ne détectent probablement pas exactement les mêmes anticorps spécifiques de *Map*. Il est possible que certains résultats discordants proviennent d'animaux en début de séroconversion et produisant au moment du test, une faible quantité d'anticorps détectée par une trousse et une plus grande quantité détectée par un autre test. Cette hypothèse est appuyée par le fait que les valeurs S/P des tests discordants négatifs étaient en moyenne significativement plus proches des seuils fixés par les fabricants que celles des tests concordants, indiquant ainsi la détection d'anticorps vis-à-vis de *Map* à un niveau supérieur à celui des tests concordants, mais encore insuffisants pour dépasser le seuil de positivité. Il en est de même pour les valeurs S/P des résultats discordants positifs qui sont elles aussi plus proches du seuil de positivité que celles des tests concordants.

Une autre hypothèse permettant d'expliquer la présence de résultats discordants entre les trousse de sérologie serait cette fois un **défaut de spécificité** des tests, ce que notre étude transversale en élevages infectés (article 1) a clairement montré (Sp_D vis-à-vis de

l'infection des 2 trousses évaluées aux alentours de 95%). Des mycobactéries environnementales sont connues pour positiver les tests ELISA de détection de la paratuberculose (Nielsen et al., 2002). La plupart des trousses intègrent une étape de pré-incubation des échantillons à évaluer avec *Mycobacterium phlei* afin de réduire ces défauts de spécificité. Cependant la nature des antigènes de *M. phlei* utilisés dans les différentes trousses, ainsi que les éventuelles différences entre lots n'étant pas connues, il ne peut être exclu qu'une part des anticorps induits par des mycobactéries environnementales ne sont pas captés au cours de cette étape et provoque des réactions faussement positives aux tests ELISA. Les différences marquées entre les valeurs des résultats discordants dans les élevages investigués (entre 43% et 100% parmi les échantillons positifs à au moins un test dans neuf élevages avec plus de quatre résultats positifs) pourraient être dues à la présence dans certains élevages d'espèces ou de souches de mycobactéries particulières.

Le typage (méthode Miru-VNTR puis identification des souches S par PCR LSPA 20 et 4 (Alexander et al., 2009)) de souches de *Map* a été réalisé à l'UMR INRA Infectiologie et santé Publique (ISP) de Tours à partir de prélèvements d'iléon de deux brebis en phase clinique issues de deux troupeaux différents du Lot (brebis par ailleurs fortement séropositives avec les deux trousses ELISA). Les résultats indiquent qu'il s'agissait dans les deux cas de souches de type S (Frank Biet, communication personnelle). Au-delà du fait que, à notre connaissance, la présence de *Map* de type S chez des ovins en France est démontrée pour la première fois, ces résultats pourraient expliquer les discordances entre trousses ELISA, voire pour partie les faibles sensibilités diagnostiques évaluées. En effet, la composition de l'enveloppe lipidique de *Map*, principal élément immunogène, varie selon les souches (Bannantine et al., 2017), et les souches utilisées pour la conception des trousses ELISA sont des souches de type C. Ainsi, les anticorps ovins, dont la spécificité est dirigée contre des souches S, pourraient avoir des affinités variables vis-à-vis des antigènes présents dans les différentes trousses ELISA et expliquer les discordances observées. De plus amples travaux sont cependant nécessaires pour vérifier cette hypothèse et sont en cours d'élaboration en collaboration avec l'UMR ISP.

Par ailleurs signalons que dans un des troupeaux supposés indemnes de paratuberculose, parmi 169 brebis prélevées, cinq résultats positifs en sérologie ELISA et deux ré-

sultats douteux ont été obtenus avec la trousse ID Screen Paratuberculosis Indirect[®] (lot 602). Sur les 54 échantillons (dont les sept échantillons positifs ou douteux précédemment cités) de ce même troupeau qui ont été testés une seconde fois avec la trousse IDEXX paratuberculosis screening[®], (lot 5074), trois résultats positifs et deux douteux ont été obtenus (dont les deux déjà détectés avec la trousse IDvet). Par ailleurs les 169 brebis échantillonnées ont fait l'objet d'une qPCR sur fèces, et tous les résultats ont été négatifs. Ce troupeau était un élevage de brebis Lacaune viande sélectionneur, installé depuis plus de 15 ans, sans aucune suspicion de paratuberculose clinique et considéré comme fermé car relativement isolé des autres troupeaux ovins. Ces résultats n'ont pas été intégrés dans nos plans de simulation car ils apparaissaient trop éloignés des résultats habituellement obtenus dans des troupeaux indemnes et auraient pénalisé les paramètres de spécificité de la sérologie ELISA dans les troupeaux indemnes (dans ce troupeau : $162/169 = 95,8\%$ pour la trousse ID Screen Paratuberculosis Indirect[®] et $49/54 = 90,7\%$ pour la trousse IDEXX paratuberculosis screening[®]). Ces observations sont cependant intéressantes à prendre en considération. Si cet élevage était réellement indemne de *Map*, elles pourraient notamment témoigner d'un défaut de Sp_D parfois majeur des trusses ELISA, peut-être retrouvé localement dans certains élevages en raison de réactions croisées avec des Mycobactéries environnementales. Une autre hypothèse est que cet élevage a réellement été exposé à *Map*, et que certains animaux ont séroconverti, sans pour autant développer la maladie. Ceci serait en faveur d'une résilience ou résistance forte des brebis de race Lacaune face à la contamination par *Map* (de Silva et al., 2018). Des travaux sont en cours en collaboration avec l'UMT ISP afin de mieux caractériser les échantillons de sérum et de vérifier la spécificité de la réponse anticorps vis-à-vis de *Map*.

2 Estimation des performances relatives à l'échelle des mélanges

Pour des maladies à faible prévalence, il est raisonnable de faire l'hypothèse que les mélanges d'échantillons contiennent peu d'échantillons positifs à l'échelle individuelle. De ce fait, les performances diagnostiques des tests en mélanges d'échantillons ne peuvent

être considérées identiques à celles observées à l'échelle individuelle. Un test considère un échantillon comme positif si la cible recherchée est présente en quantité supérieure à la limite de détection du test, et la réalisation d'un mélange va consister en une dilution de cette cible si un seul échantillon est positif à l'échelle individuelle. D'autre part, la Sp_D peut également être dégradée si plusieurs échantillons à la limite du seuil de détection sont inclus dans le mélange.

2.1 Performances relatives des trousse ELISA appliquées à des mélanges de sérums

Les trousse ELISA ne ciblant pas un antigène particulier mais un mélange d'antigènes, nous avons fait l'hypothèse que les anticorps présents dans le sérum des échantillons peuvent avoir des avidités variables, pouvant modifier les performances diagnostiques des analyses de mélanges. Les analyses en dilutions sériées (figures 20 and 21) ont montré une décroissance de la valeur S/P très variable en fonction des échantillons. Ce phénomène peut être lié à une avidité différente d'un sérum à l'autre. L'avidité caractérise la force de l'interaction entre sérum et antigènes cibles. Elle est égale à la somme des affinités qui sont les forces d'interactions entre épitope et paratope. Les trousse ELISA étudiées dans nos travaux ne ciblent pas un épitope ou antigène particulier mais un mélange d'antigènes fixés au fond des puits des plaques ELISA. Les anticorps vont avoir une affinité propre pour chacun des épitopes de ce mélange contre lesquels ils sont dirigés, créant une variabilité de l'avidité de chaque sérum.

Pour évaluer ces potentielles interactions, une catégorisation plus fine des sérums a été réalisée en se basant sur la valeur S/P individuelle initiale. L'objectif principal était de définir le seuil de positivité des mélanges en fonction de leur composition et de leur taille. Nous n'avons pas opté pour une approche de modélisation (par modèle linéaire généralisé par exemple), car l'obtention d'intervalles de confiance et de prédiction étroits aurait nécessité un plus grand nombre d'échantillons analysés. Nos résultats sont en faveur de l'existence d'interactions entre les sérums positifs. En effet alors que la Se_D relative de la trousse IDEXX paratuberculosis screening[®]) pour la détection d'un seul échan-

tillon faiblement positif était de 62,2% pour les mélanges de cinq échantillons (résultats semblables pour les mélanges de 10 échantillons), elle était de 100% pour les mélanges contenant deux échantillons faiblement positifs. Pour ces mélanges nous avons calculé la sensibilité théorique, obtenue par la relation

$$Se_n = 1 - (1 - Se_1)^n \quad (26)$$

avec Se_n la sensibilité des mélanges contenant n échantillons positifs, et Se_1 la sensibilité des analyses en mélanges contenant un seul échantillon positif. La Se_2 calculé était de 0,8571, bien inférieure à ce que nous avons observé. Le même phénomène semble exister pour la trousse ID Screen Paratuberculosis Indirect[®] avec une Se_D observée de 41,9% pour les mélanges de taille 10 contenant un seul échantillon faiblement positif, contre 100% des mélanges contenant deux échantillons faiblement positifs, alors que la Se_D théorique est de 66,2%. Cependant ces observations sont basées sur un faible nombre de mélanges contenant deux échantillons faiblement positifs ($n=8$) pour cette trousse, et mériteraient d'être consolidées.

Quelle que soit la composition des mélanges, nous avons observé une forte dégradation de la sensibilité des trousse ELISA lorsque le seuil de positivité défini pour les analyses individuelle était conservé. Pour contrebalancer la diminution de Se_D à l'échelle du mélange d'échantillons, plusieurs approches étaient possibles. Nous avons retenu la modification du seuil de positivité comme décrit par Rovira et al. (2008), car cette approche nous permet d'assurer une Sp_D relative parfaite des analyses en mélanges. La Sp_D des tests ELISA à l'échelle individuelle étant médiocre nous ne souhaitons pas dégrader cette performance diagnostique à l'échelle des mélanges. Une autre approche consisterait à diminuer le volume de pré-dilution tel qu'appliqué par Brinkhof et al. (2007) pour la détection des anticorps dirigés contre le virus Maedi-Visna. Dans cette étude, les auteurs ont évalué les performances diagnostiques de plusieurs protocoles de pré-dilution depuis le facteur de dilution recommandé par le fabricant de la trousse (facteur de dilution de 100) jusqu'à l'analyse d'échantillons non dilués. Une pré-dilution à un facteur de 12,5 permet de discriminer correctement les échantillons négatifs de ceux qui sont positifs. Dans notre

cas, les protocoles des deux trousse ELISA recommandent aussi des pré-dilutions d'un facteur 12,5, que nous avons appliquées. La modification de cette étape de pré-dilution aurait par ailleurs nécessité un travail supplémentaire pour calibrer les étapes suivantes, notamment la pré-adsorption par *M. phlei*, ce qui n'était pas l'objectif principal de notre travail. Cette approche mériterait cependant un travail spécifique pour optimiser la détection des sérums positifs de faible valeur S/P, sans dégrader la Sp_D.

2.2 Performances relatives de la qPCR dans des mélanges de fèces

La qPCR utilisée ciblant uniquement le gène *IS900* nous avons considéré, contrairement à la sérologie ELISA, qu'il n'y avait pas d'interactions entre les différents échantillons composant les mélanges. Nous avons catégorisé de manière plus fine les échantillons de fèces en fonction de la valeur de Ct donc de la quantité d'ADN de *Map* présente dans chaque échantillon individuel. Les fèces étant riches en inhibiteurs de réaction d'amplification provenant de la dégradation des aliments (Acharya et al., 2017), le mélange de fèces augmente le risque d'inhibition. Plusieurs auteurs recommandent la dilution des échantillons dans un solvant neutre afin de réduire ce risque, cette approche nous semble difficilement applicable à l'analyse de mélanges car elle augmente artificiellement la dilution déjà existante. Les résultats obtenus dans nos travaux indiquent que la qPCR sur mélange de fèces possède de bonnes performances diagnostiques, en particulier pour les mélanges contenant des fèces à forte concentration en ADN de *Map*. La principale limite de cette approche se situe dans la réalisation pratique des mélanges. Dans un souci de diminuer le coût et le temps nécessaires à la réalisation des mélanges, nous avons choisi de réaliser ces mélanges sous forme solide et manuellement. De plus, pour conserver une bonne représentativité des prélèvements initiaux composant le mélange une masse de 10g de fèces a été utilisée ce qui représente une masse totale de 100 g de fèces à homogénéiser pour un mélange de 10 échantillons. Comparativement aux fèces bovins, la consistance très sèche et ferme des fèces ovins rend leur homogénéisation techniquement plus difficile, ce qui est d'autant plus critique dans le cas de la paratuberculose car les mycobactéries ne sont pas réparties de manière homogène dans les fèces mais sous forme d'amas (Christopher-Hennings et al., 2003). Comme suggéré par

(Whittington et al., 2000a) dans le cas de la culture fécale, une homogénéisation insuffisante peut nuire à de bonnes performances diagnostiques des analyses de mélanges, en particulier pour les fèces issus d'animaux faibles excréteurs. Pour explorer ce problème d'homogénéisation des fèces, une série d'expériences complémentaires a été menée. Des échantillons de fèces provenant de 54 caprins issus d'élevages infectés par la paratuberculose ont été sélectionnés sur la base du résultat individuel de qPCR (Ct variant de 30,6 à 38,0) afin de réaliser des mélanges de taille 5 et 10 contenant 1, 2 ou 4 échantillons positifs. Les mélanges n'ont pas été réalisés en mélangeant *sensu stricto* les fèces, mais en mélangeant les surnageants obtenus après réhydratation selon le protocole de la trousse ADIAVET™ paraTB Real Time, BioX (10 g de fèces dans 70 ml d'eau bi-distillée). L'efficacité du processus d'homogénéisation de fèces a été évaluée en comparant la Se_D obtenue pour des mélanges contenant plusieurs échantillons positifs avec la sensibilité théorique présentée à l'équation (26). Les résultats obtenus à partir des fèces caprins sont présentés dans le tableau 9. Les sensibilités observées pour les mélanges contenant deux ou quatre

Tableau 9 – Résultats des analyses de mélanges de fèces en phase liquide par qPCR

Nombre de positif	Mélanges de 5			Mélanges de 10		
	# pos/# ¹	Se (%)	Se_t ² (%)	# pos/#	Se (%)	Se_t (%)
1	25/36	69,44	-	23/36	63,89	-
2	32/36	88,89	90,66	45/54	83,33	86,96
4	-	-	-	18/18	100,00	98,30

¹ nombre de mélanges positifs/nombre de mélanges testés

² Se_t théorique calculée selon l'équation (26)

échantillons positifs sont très proches des sensibilités théoriques et semblent donc en faveur d'une très bonne homogénéisation des échantillons lorsque les mélanges sont réalisés à partir des surnageants de fèces. Le surcoût engendré par le temps nécessaire à la préparation et les consommables supplémentaires pénalisent cette méthode d'analyse, mais devra être étudié avec les laboratoires susceptibles de réaliser les analyses de mélanges afin de rationaliser son utilisation.

D'autres procédés d'homogénéisation sont possibles tels que ceux recommandés par Eamens et al. (2015) dans un rapport visant à standardiser les méthodes pour le diagnostic de la paratuberculose en Australie et en Nouvelle-Zélande. Les auteurs préconisent

une homogénéisation mécanique des fèces solide en utilisant un mixeur ou l'utilisation d'un broyeur homogénéisateur (Bag Mixer) après addition de sérum physiologique dans un rapport 1 : 1 afin de réhydrater les fèces. Le procédé d'homogénéisation par broyeur-homogénéisateur a été appliqué aux fèces solides (sans ajout de sérum physiologique) dans le cadre de nos expériences, mais n'a pas été retenu par la suite, le procédé se révélant inefficace. Des travaux supplémentaires sont envisagés afin d'évaluer l'intérêt de l'utilisation des broyeurs homogénéisateurs mécaniques avec ajout de liquide.

3 Définition et suivi du statut sanitaire sur la base de résultats de mélanges d'échantillons

L'utilisation des mélanges d'échantillons pour le dépistage de la paratuberculose à l'échelle collective a été étudiée en se basant sur la simulation des résultats des tests diagnostiques à l'échelle individuelle pour différentes prévalences d'animaux infectés et une réalisation aléatoires des mélanges. De manière empirique, nous avons choisi de simuler 1000 élevages pour chaque combinaison de paramètres (prévalence d'infection, taille et nombre de mélanges). Le choix des animaux sélectionnés pour la constitution des mélanges était également répété sur 1000 itérations pour tenir compte de la variabilité liée à l'échantillonnage. Nous avons fait le choix de simuler différentes prévalence d'infection et non les prévalences d'excrétion ou de séropositivité classiquement utilisées pour rendre compte du statut d'un troupeau vis-à-vis de la paratuberculose (Dhand et al., 2010b ; Raizman et al., 2011 ; Tavoranpanich et al., 2008 ; Dhand et al., 2016). La prévalence apparente (P_A) peut se déduire de la prévalence réelle (P_R) en connaissant la Se_D et la $Sp - D$ du test utilisé selon la relation $P_A = (Se_D * P_R) + (1 - Sp_D) * (1 - P_R)$. En utilisant les performances diagnostiques évaluées au chapitre 3, une prévalence réelle de 10% correspond à une prévalence apparente d'animaux positifs en qPCR comprise entre 2,90% et 8,80% (entre 1,45% et 5,40%, pour une prévalence réelle de 5%) et une prévalence apparente d'animaux positif à l'ELISA IDEXX paratuberculosis screening[®] comprise entre 4,60% et 9,70% (entre 4,30% et 8,85% pour une prévalence réelle de 5%). D'après nos résultats de simulation, l'utilisation de la sérologie ELISA pour la détection des élevages infectés à

partir de mélanges de sérums est associée à une Se_T plus élevée que celle obtenue pour la qPCR pour des prévalences d'infection comprises entre 1% et 10%. A l'inverse, pour des prévalences d'infection supérieures à 10%, la qPCR présentait une Se_T supérieure à la sérologie ELISA. La bonne Se_T de la sérologie ELISA est cependant contrebalancée par une Sp_T très dégradée (entre 41% et 62% selon la taille de mélanges et le nombre de mélanges simulés), limitant très fortement son intérêt pratique. Bien que le seuil de positivité de la sérologie ELISA ait été adapté aux analyses de mélanges afin d'assurer une Sp_D relative parfaite, la mauvaise spécificité individuelle de la sérologie ELISA (estimée à 95% dans les troupeaux infectés et à 99% dans les troupeaux indemnes) entraîne un nombre important de résultats individuels faussement positifs, à l'origine à la fois de la bonne sensibilité de détection des troupeaux infectés et de la faible spécificité vis-à-vis des troupeaux indemnes. Dans les élevages effectivement infectés, une proportion non négligeable de mélanges positifs contiennent des échantillons individuels faussement positifs, augmentant artificiellement la Se_T . La figure 24 montre la répartition des mélanges en fonction de leur composition et de la taille des mélanges d'échantillons pour différentes prévalences d'infection intra-élevage, obtenus à partir de nos simulations du chapitre 4. La proportion des mélanges faux positifs est plus faible pour les mélanges de taille 5 que pour les mélanges de taille 10. De plus cette proportion diminue lorsque la prévalence augmente (de 18,37% à 3,20% pour les mélanges de taille 5 et de 31,95% à 0,93% pour les mélanges de taille 10).

De ce fait l'utilisation de la sérologie ELISA appliquée à des mélanges de sérums ne semble pas indiquée dans les plans de dépistage à grande échelle car elle conduirait à considérer à tort comme infectés un trop grand nombre d'élevages indemnes. Par ailleurs, la reprise des mélanges positifs (jusqu'à deux mélanges positifs pour 50 animaux échantillonnés, et trois mélanges positifs pour 100 animaux prélevés pour les élevages indemnes) pour la réalisation d'analyses sérologiques individuelles ne permettrait pas d'améliorer la distinction entre troupeaux indemnes et troupeaux réellement infectés, car les échantillons individuels eux-mêmes sont à l'origine du défaut de Sp_T . En revanche dans un contexte de suspicion de paratuberculose, l'obtention de 4 ou 5 mélanges positifs sur 10 pour 50 animaux prélevés permettrait de confirmer la suspicion à moindre coût, sans

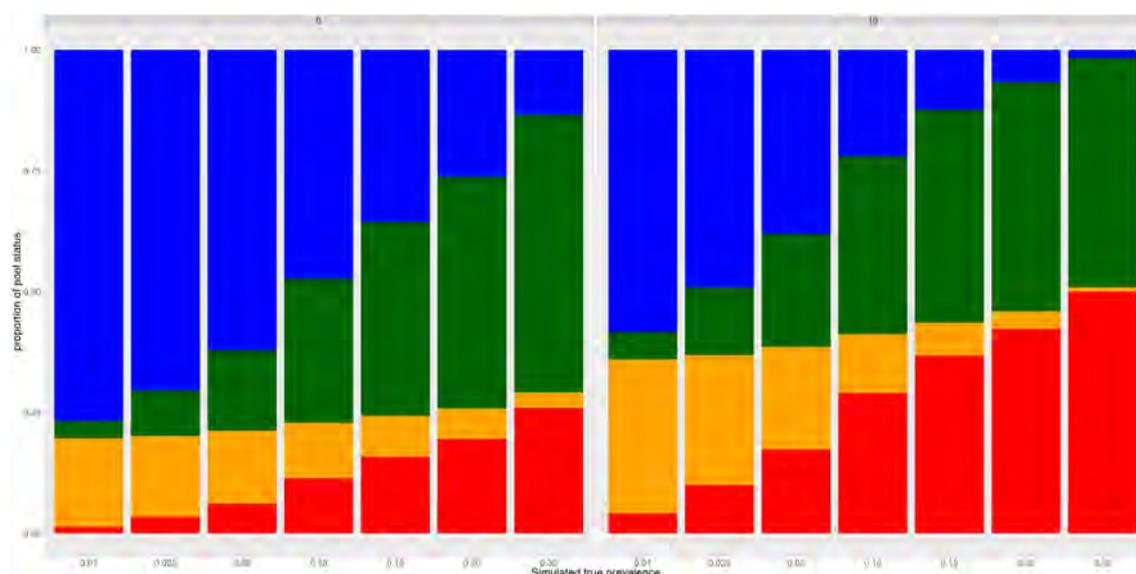


FIGURE 24 – Statut des mélanges de taille cinq et 10 analysés en ELISA en fonction de la prévalence d’infection

Le statut des mélanges est défini en fonction des échantillons individuels. Bleu : vrai négatif : résultat négatif sur un mélange ne contenant que des échantillons d’animaux non infectés ; Orange : faux positif : résultat positif sur un mélange ne contenant que des échantillons d’animaux non infectés ; Rouge : vrai positif : résultat positif sur un mélange contenant au moins un échantillon d’animal infecté ; Vert : faux négatif : résultat négatif sur un mélange contenant au moins un échantillon d’animal infecté. Les analyses ont été réalisées en utilisant la trousse IDEXX paratuberculosis screening®

avoir recours aux analyses individuelles. Cette approche ne permettrait cependant pas une évaluation satisfaisante de la prévalence d’infection intra-troupeau, le même nombre de mélange positifs pouvant être obtenu pour des prévalences d’infection très variées (par exemple trois mélanges positifs peuvent être obtenus dans 19% à 26% des cas pour des prévalences d’infection comprises entre 1% et 30%).

La qPCR sur mélanges de fèces semble plus adaptée pour le dépistage et le suivi de la paratuberculose dans les troupeaux ovins notamment dans l’objectif d’estimer la prévalence d’infection plutôt qu’une indication de présence/absence de la paratuberculose dans l’élevage. Nous avons observé une meilleure discrimination des élevages avec des prévalences différentes en tenant compte du nombre de mélanges positifs. Ainsi, lorsque 10 mélanges de 5 échantillons de fèces sont analysés par qPCR, trois mélanges ou plus peuvent être détectés dans seulement 2% des élevages avec une prévalence d’infection de 1%. Cette probabilité augmente avec la prévalence intra-élevage pour arriver à 96% des cas lorsque la prévalence est de 30%.

4 Application au contexte des troupeaux vaccinés

Le vaccin GUDAIR® a fait l'objet de peu d'études longitudinales à grande échelle et sur une période suffisamment étendue pour pouvoir évaluer son efficacité et ses effets à long terme dans les troupeaux. Les résultats rapportés par (Dhand et al., 2013; 2016) ne sont pas nécessairement transposables aux troupeaux français. Dans les troupeaux australiens, seuls les agneaux (mâles et femelles, et non pas uniquement les agnelles de renouvellement) ont été vaccinés, avant l'âge de six mois (très majoritairement entre un et quatre mois, sans que les distributions des données ne soient communiquées). Dans les 13 élevages inclus dans notre étude, tous les reproducteurs, adultes compris, ont été vaccinés dans les deux premières années suivant la mise en place de la vaccination. Les années suivantes, seules les agnelles de renouvellement ont été vaccinées, à un âge compris entre quatre et six mois, au moment du sevrage et du choix des agnelles qui étaient conservées. Par ailleurs, si la culture appliquée à des mélanges d'échantillons issus de 25 à 50 animaux permet la détection quasi-systématique des animaux forts excréteurs (multibacillaires), elle est associée à une Se_D réduite (de l'ordre de 23% pour les mélanges de taille 20 et de 8% pour des mélanges de taille 40) pour les animaux faibles excréteurs (paucibacillaires) (Whittington et al., 2000a ; Dhand et al., 2010b). S'il s'avérait que la vaccination réduit fortement les niveaux d'excrétion fécale, les approches par culture sur des mélanges de fèces de grande taille, telles qu'utilisées dans l'étude australienne citée plus haut, pourraient avoir une capacité limitée à mettre en évidence une excrétion résiduelle dans les troupeaux vaccinés.

Pour cette raison, nous avons privilégié la détection par qPCR sur fèces individuelles, maximisant ainsi la Se_D et permettant d'obtenir des estimateurs de la prévalence individuelle d'excrétion ainsi que des niveaux d'excrétion des animaux vaccinés. Cependant, la mise en place de la vaccination ayant débuté avant nos travaux, un suivi longitudinal n'a pas pu être réalisé. L'approche que nous avons retenue, basée sur la comparaison de la séroprévalence apparente avant vaccination et celle de la prévalence apparente d'excrétion fécale après quatre à cinq ans nécessite de faire l'hypothèse que ces deux prévalences sont corrélées. La figure 25 montre les niveaux de prévalence d'excrétion et de séroprévalence

dans les élevages non vaccinés de notre étude.

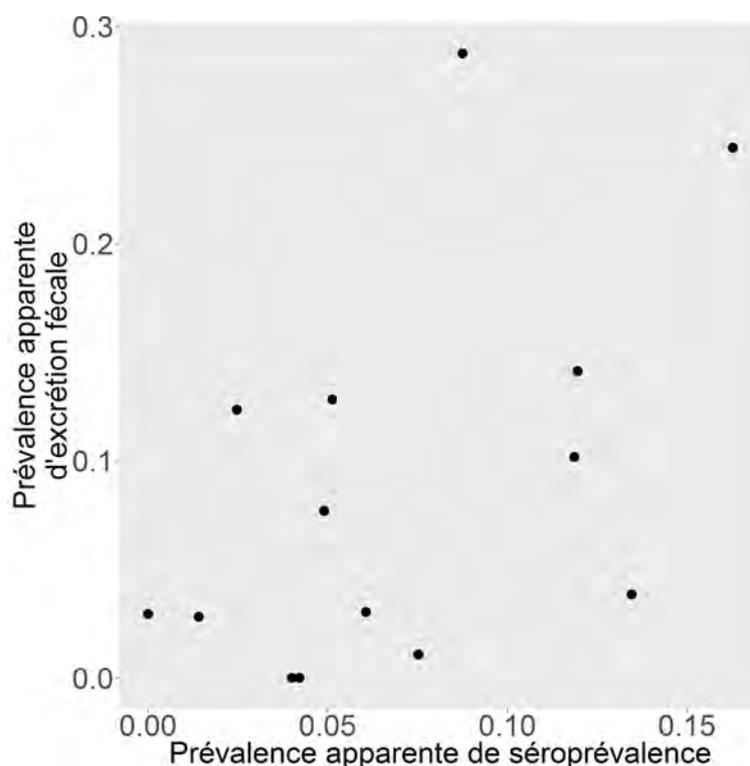


FIGURE 25 – Relation entre séroprévalence apparente et prévalence apparente d'excrétion fécale dans 14 troupeaux ovins non vaccinés du Lot

Chaque point représente un élevage. Les résultats de sérologie ELISA ont été obtenus en utilisant le kit ID Screen Paratuberculosis Indirect[®], lot 602, IdVet. Les résultats de qPCR sur fèces ont été obtenus en utilisant le kit ADIAVET[™]paraTB Real Time, BioX.

Une corrélation positive significative ($p=0,032$) entre ces deux indicateurs existe, mais elle est modérée (coefficient de corrélation de Pearson $r = 0,51$). Cette corrélation imparfaite diminue cependant la puissance statistique de notre analyse. Le schéma d'étude utilisé dans notre étude, comparant les élevages vaccinés à ceux non vaccinés, peut également réduire la puissance statistique. En effet la vaccination a été très majoritairement réalisée dans les élevages où la prévalence de la paratuberculose était la plus élevée, avec un fort impact clinique. À l'inverse 4 des 10 élevages non-vaccinés avaient une prévalence sérologique apparente très faible ($< 5\%$) que ce soit en 2009 ou en 2014 et avec des prévalences apparentes d'excrétion fécale faibles ($< 5\%$) et correspondaient probablement à des élevages faiblement contaminés. De ce fait, la comparaison des deux groupes d'élevages est *a priori* en défaveur des élevages vaccinés, qui souffraient de conditions initiales plus défavorables.

Contrairement à l'étude de Reddacliff et al. (2006), nos résultats indiquent clairement la persistance d'une réponse sérologique forte jusqu'à quatre ans après vaccination chez la quasi totalité des animaux vaccinés. Il n'a pas été possible de vérifier si les quelques animaux séronégatifs, par ailleurs présents dans toutes les cohortes d'âge et négatifs en qPCR sur fèces, avaient effectivement été vaccinés correctement (respect des doses et de la voie d'administration) ou s'il s'agissait d'un échappement à la vaccination (animaux non répondeurs). A l'inverse les résultats obtenus par Reddacliff et al. (2006) montrent une diminution de la séroprévalence deux ans après le vaccin, mais assez variable entre les trois élevages étudiés. Dans cette étude, les agneaux étaient vaccinés plus jeunes (entre un et trois mois) que dans notre étude (entre quatre et six mois). Cette différence dans le protocole vaccinal pourrait pour partie expliquer les résultats contradictoires obtenus. A la naissance le système immunitaire n'est pas mature; Chase et al. (2008) conseillent, chez les bovins, une vaccination tardive pour obtenir une meilleure efficacité. Nous pouvons faire l'hypothèse que le même processus et donc les mêmes recommandations sont valables chez les ovins, une vaccination tardive pourrait être à l'origine d'une meilleure efficacité. Nos résultats ne permettent pas de conclure dans un sens ou un autre, de plus amples travaux expérimentaux seraient nécessaires afin d'explorer cette hypothèse.

Le suivi sanitaire des troupeaux vaccinés par l'utilisation de la sérologie n'apparaît donc pas pertinent dans le contexte des élevages du Lot. Cependant les très fortes séroprévalences observées post-vaccination pourraient faire de la sérologie ELISA un bon indicateur de la vaccination effective des animaux. Par ailleurs dans le cadre du suivi de la paratuberculose dans le contexte de troupeaux vaccinés, l'élaboration d'un test ELISA différenciant les animaux infectés des animaux vaccinés (ELISA DIVA), comme c'est le cas dans le cadre de l'IBR chez les bovins (Muratore et al., 2017), serait une avancée majeure.

Nous avons exclu de nos analyses un élevage dont seulement la moitié des agnelles de renouvellement était vaccinée, l'autre moitié étant constituée d'agnelles non vaccinées achetées. Sur les 171 brebis prélevées dans cet élevage, 93 n'étaient pas vaccinées et 12 étaient positives en qPCR sur fèces (12,90%). Parmi les 78 brebis vaccinées prélevées dans cet élevage, 11 présentaient une qPCR positive (14%). Ces prévalences apparentes d'excrétion fécale sont très supérieures à celles observées dans les autres élevages où la

totalité des agnelles de renouvellement était vaccinée. Ces résultats sont cependant en accord avec ceux de Windsor et al. (2014) qui ont identifié l'introduction d'animaux non vaccinés ou le pâturage en commun avec des troupeaux non vaccinés comme des facteurs de risque d'échec de la vaccination. Ces observations militent pour la mise en place de la vaccination sur la totalité des agnelles de renouvellement.

Une persistance de l'excrétion fécale a été retrouvée dans les élevages vaccinés après quatre ans de vaccination ininterrompue. Les sept animaux positifs en qPCR sur fèces dans quatre des cinq troupeaux étaient tous fortement positifs en sérologie ELISA. Ces résultats ne permettaient pas de savoir s'il s'agissait d'animaux effectivement correctement vaccinés ou non, ou si la réponse sérologique observée correspondait à une réponse à une infection naturelle. Par ailleurs nous avons présenté les résultats de prévalences apparentes d'excrétion fécale et non les prévalences vraies. Sous l'hypothèse que les performances diagnostiques à l'échelle individuelle de la qPCR sont les mêmes pour le dépistage des animaux infectés non vaccinés et infectés vaccinés (de l'ordre de 50%), les prévalences d'infection seraient en réalité le double de celles rapportées, et demeurent donc élevées dans certains élevages.

Les résultats de simulation de suivi annuel basé sur des mélanges de fèces, présentés au chapitre 5 indiquent que la qPCR est un bon outil d'évaluation du statut vis-à-vis de la paratuberculose. Ce test présente des performances diagnostiques suffisantes pour une utilisation en mélanges d'échantillons dans le but de discriminer des élevages indemnes des élevages faiblement infectés sur la base de plusieurs années de suivi. Les résultats obtenus dépendent cependant de plusieurs hypothèses dont la principale est celle de l'homogénéité de la prévalence d'infection au cours des années. Cette hypothèse est sans doute vérifiée dans le cadre d'un maintien de la vaccination contre la paratuberculose mais pourrait s'avérer fautive après l'arrêt de la vaccination. Un nouveau cycle de suivi longitudinal sur à minima trois à cinq ans serait donc à prévoir dans les élevages arrêtant la vaccination, permettant de maintenir une bonne connaissance de leur statut. De même, l'hypothèse d'une Se_D à l'échelle individuelle et des mélanges équivalente entre animaux vaccinés et non vaccinés est importante. Nos observations réalisées dans les élevages ayant eu recours à la vaccination depuis plus de cinq ans sont en faveur de

niveaux d'excrétion fécale semblables entre les animaux infectés non vaccinés et ceux qui sont vaccinés et infectés. Cependant s'il s'avérait que la vaccination conduit à une excrétion fécale intermittente, alors la sensibilité de l'analyse sur fèces serait moindre et les résultats des simulations seraient trop optimistes.

Compte tenu de la persistance d'animaux excréteurs dans les troupeaux vaccinés, il apparaît nécessaire de maintenir une vaccination sur le long terme afin d'éviter la résurgence de la maladie à l'arrêt de la vaccination. Toutefois, nous pouvons nous demander si l'arrêt de la vaccination ne peut pas être réalisé dans les élevages où la prévalence et les niveaux d'excrétion sont faibles. Cet arrêt éventuel de la vaccination ne pourrait être envisagé qu'en maintenant des mesures de surveillance adéquates (évaluation de nouvelles infections naturelles par réalisation de sondage et d'analyse en mélange, en ciblant les générations non vaccinées dans ces troupeaux) et des mesures sanitaires strictes (réforme précoce des animaux excréteurs identifiés après reprise des mélanges positifs en qPCR, contrôle des animaux introduits). Le maintien d'une surveillance épidémiologique dans ces troupeaux, sur la base de la qPCR appliquée à des mélanges de fèces, permettrait par ailleurs une reprise de la vaccination dans les plus brefs délais s'il s'avérait que l'incidence de nouvelles infections naturelles était importante.

Nous avons cependant pu constater que la vaccination ne permet pas d'assainir totalement les élevages dans lesquels elle est utilisée. Cette mesure étant coûteuse, un arrêt de la vaccination peut être envisagé si le ratio coût/bénéfice est considéré négatif. Il n'existe à ce jour aucune connaissance permettant d'évaluer l'évolution de l'infection dans les élevages après l'arrêt de la vaccination dans le contexte des élevages français. La présentation des résultats obtenus dans nos travaux aux éleveurs du Lot lors de différentes réunions de restitution a par ailleurs convaincu les éleveurs de maintenir la vaccination, y compris ceux souhaitant initialement l'arrêter, sans que nous puissions leur indiquer combien d'années seraient encore nécessaires.

Conclusion générale

Les résultats obtenus dans le cadre de ce travail de thèse ont permis de montrer que l'établissement et le suivi du statut sanitaire vis-à-vis de la paratuberculose des troupeaux ovins ovin vaccinés ou non vaccinés est possible par analyses de mélanges d'échantillons de fèces en utilisant la qPCR. Nos résultats nous conduisent par ailleurs à déconseiller l'utilisation de la sérologie ELISA sur des mélanges de sérums dans le cadre d'un diagnostic collectif pour des troupeaux de statut sanitaire *a priori* inconnu, en raison du fort risque de résultat faussement positif, principalement lié à la spécificité insuffisante de ce test à l'échelle individuelle. Malgré des performances diagnostiques imparfaites à l'échelle individuelle ou des mélanges, la réduction des coûts obtenue par l'analyse de mélanges d'échantillons de fèces rend envisageable la mise en place d'un suivi longitudinal des troupeaux ovins souhaitant connaître ou suivre leur statut vis-à-vis de la paratuberculose. Dans le cadre d'un tel plan de suivi, la répétition des mesures permettrait d'augmenter la précision de l'estimation de la prévalence intra-troupeau et d'identifier de manière plus certaine les élevages indemnes ou à faible risque d'infection.

De manière plus générale, la démarche déployée au long de ce travail de thèse peut être appliquée pour le développement de méthodologie diagnostiques à l'échelle collective pour d'autres maladies infectieuses pour lesquelles le statut du troupeau prévaut sur le statut individuel (Græsbøll et al., 2017). Cependant quelles qu'elles soient, les stratégies diagnostiques basées sur l'analyse de mélanges d'échantillons doivent agrégées des connaissances préalables relevant de plusieurs niveaux d'échelle :

- Connaissance ou évaluation des performances diagnostiques individuelles vis-à-vis d'un état cible;
- Connaissance ou évaluation des performances diagnostiques relatives des mélanges,

c'est à dire évaluation de l'effet de la dilution des échantillons et des interactions possibles entre échantillons composant le mélange, sur le niveau de réponse du test appliqué au mélange. Cette étape doit aussi permettre d'évaluer l'impact des niveaux de réponse des échantillons individuels sur la réponse du mélange;

- Connaissance ou évaluation de la distribution des niveaux de réponse individuelle dans la population cible, permettant d'appliquer les résultats de l'étape précédente et de tenir compte des variabilité possibles entre populations ou contextes d'élevage.

Cette démarche permet de tenir compte d'un grand nombre de facteurs de variation et d'incertitude. Certaines étapes peuvent cependant être ignorées, par exemple si les performances diagnostiques à l'échelle individuelle sont considérées comme parfaites, ou si l'effet de dilution des échantillons dans les mélanges est nul, ou encore si la variabilité d'intensité des réponses individuelles est peu marquée. Certaines de ces simplifications conduisent à une expression algébrique des performances des analyses de mélanges et de la prévalence intra-troupeau rendant leur estimation possible par des modèles fréquentistes ou bayésiens (Toribio and Sergeant, 2007 ; Dhand et al., 2010b ; Speybroeck et al., 2012). En dehors de ces conditions, et comme rencontré dans le cadre de notre travail sur la paratuberculose, cette estimation requiert le passage par un phase exploratoire de terrain et l'utilisation de simulations, afin de rendre compte au mieux du degré d'incertitude associé aux estimateurs.

Bibliographie

- Acharya K. R, Dhand N. K, Whittington R. J et Plain K. M. **2017**. PCR Inhibition of a Quantitative PCR for Detection of Mycobacterium avium Subspecies Paratuberculosis DNA in Feces : Diagnostic Implications and Potential Solutions. *Frontiers in Microbiology*, 8. ISSN 1664-302X. doi : 10.3389/fmicb.2017.00115.
- AEEMA. Définition d'une espèce réservoir, 2017. URL : <http://aeema.vet-alfort.fr/index.php/component/glossary/Glossaire-1/R/RESERVOIR-%28d%E2%80%99un-agent-pathog%C3%A8ne-biologique%29-284/>. [Consulté le : 2017-11-07].
- AFSSA. Rapport sur la paratuberculose des ruminants. Technical report, 2009. URL : <https://www.anses.fr/fr/system/files/SANT-Ra-Paratuberculose.pdf>.
- Albert P. S et Dodd L. E. **2004**. A Cautionary Note on the Robustness of Latent Class Models for Estimating Diagnostic Error without a Gold Standard. *Biometrics*, 60(2) :427–435. ISSN 0006-341X.
- Alexander D. C, Turenne C. Y et Behr M. A. **2009**. Insertion and Deletion Events That Define the Pathogen Mycobacterium avium subsp. paratuberculosis. *Journal of Bacteriology*, 191(3) :1018–1025. ISSN 0021-9193, 1098-5530. doi : 10.1128/JB.01340-08.
- Allman E. S, Matias C et Rhodes J. A. **2009**. Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics*, 37(6A) :3099–3132. ISSN 0090-5364. doi : 10.1214/09-AOS689. arXiv : 0809.5032.
- Aly S. S, Mangold B. L, Whitlock R. H, Sweeney R. W, Anderson R. J, Jiang J, Schukken Y. H, Hovingh E, Wolfgang D, Van Kessel J. A. S, Karns J. S, Lombard J. E, Smith J. M et Gardner I. A. **2010**. Correlation between Herrold Egg Yolk Medium Culture and Real-Time Quantitative Polymerase Chain Reaction Results for Mycobacterium Avium Subspecies Paratuberculosis in Pooled Fe-

- cal and Environmental Samples. *Journal of Veterinary Diagnostic Investigation*, 22(5) :677–683. ISSN 1040-6387. doi : 10.1177/104063871002200501.
- Aly S, Anderson R, Whitlock R, Fyock T, McAdams S, Adaska J, Jiang J et Gardner I. **2009**. Reliability of environmental sampling to quantify *Mycobacterium avium* subspecies paratuberculosis on California free-stall dairies. *Journal of Dairy Science*, 92(8) :3634–3642. ISSN 00220302. doi : 10.3168/jds.2008-1680.
- Alzuhherri H. M, Woodall C. J et Clarke C. J. **1996**. Increased intestinal TNF- α , IL-1 β and IL-6 expression in ovine paratuberculosis. *Veterinary Immunology and Immunopathology*, 49(4) : 331–345. ISSN 0165-2427. doi : 10.1016/0165-2427(95)05477-4.
- Anderson J. L, Meece J. K, Koziczowski J. J, Clark D. L, Radcliff R. P, Nolden C. A, Samuel M. D et Ellingson J. L. E. **2007**. *Mycobacterium avium* subsp. paratuberculosis in Scavenging Mammals in Wisconsin. *Journal of Wildlife Diseases*, 43(2) :302–308. ISSN 0090-3558. doi : 10.7589/0090-3558-43.2.302.
- Angelidou E, Kostoulas P et Leontides L. **2016**. Bayesian estimation of sensitivity and specificity of a commercial serum/milk ELISA against the *Mycobacterium avium* subsp. Paratuberculosis (MAP) antibody response for each lactation stage in Greek dairy sheep. *Preventive Veterinary Medicine*, 124 :102–105. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2015.12.011.
- Arango-Sabogal J. C, Côté G, Paré J, Labrecque O, Roy J.-P, Buczinski S, Doré E, Fairbrother J. H, Bissonnette N, Wellemans V et Fecteau G. **2016**. Detection of *Mycobacterium avium* subspecies paratuberculosis in tie-stall dairy herds using a standardized environmental sampling technique and targeted pooled samples. *Canadian Journal of Veterinary Research*, 80(3) :175–182. ISSN 0830-9000.
- Arsenault R. J, Maattanen P, Daigle J, Potter A, Griebel P et Napper S. **2014**. From mouth to macrophage : mechanisms of innate immune subversion by *Mycobacterium avium* subsp. paratuberculosis. *Veterinary Research*, 45(1) :54. ISSN 0928-4249. doi : 10.1186/1297-9716-45-54.
- Atreya R, Bülte M, Gerlach G.-F, Goethe R, Hornef M. W, Köhler H, Meens J, Möbius P, Roeb E et Weiss S. **2014**. Facts, myths and hypotheses on the zoonotic nature of *Mycobacterium avium* subspecies paratuberculosis. *International Journal of Medical Microbiology*, 304(7) :858–867. ISSN 1438-4221. doi : 10.1016/j.ijmm.2014.07.006.

- Ayele W. Y, Bartos M, Svastova P et Pavlik I. **2004**. Distribution of *Mycobacterium avium* subsp. paratuberculosis in organs of naturally infected bull-calves and breeding bulls. *Veterinary Microbiology*, 103(3) :209–217. ISSN 0378-1135. doi : 10.1016/j.vetmic.2004.07.011.
- Bannantine J. P, Wu C.-w, Hsu C, Zhou S, Schwartz D. C, Bayles D. O, Paustian M. L, Alt D. P, Sreevatsan S, Kapur V et others. **2012**. Genome sequencing of ovine isolates of *Mycobacterium avium* subspecies paratuberculosis offers insights into host association. *BMC genomics*, 13(1) :89.
- Bannantine J. P, Etienne G, Laval F, Stabel J. R, Lemassu A, Daffé M, Bayles D. O, Ganneau C, Bonhomme F, Branger M, Cochard T, Bay S et Biet F. **2017**. Cell wall peptidolipids of *Mycobacterium avium* : from genetic prediction to exact structure of a nonribosomal peptide. *Molecular Microbiology*, 105(4) :525–539. ISSN 1365-2958. doi : 10.1111/mmi.13717.
- Bastida F et Juste R. A. **2011**. Paratuberculosis control : a review with a focus on vaccination. *Journal of immune based therapies and vaccines*, 9(1) :8.
- Bauman C. A, Jones-Bitton A, Jansen J, Kelton D et Menzies P. **2016**. Evaluation of fecal culture and fecal RT-PCR to detect *Mycobacterium avium* ssp. paratuberculosis fecal shedding in dairy goats and dairy sheep using latent class Bayesian modeling. *BMC Veterinary Research*, 12(1). ISSN 1746-6148. doi : 10.1186/s12917-016-0814-5.
- Beaunée G, Vergu E et Ezanno P. **2015**. Modelling of paratuberculosis spread between dairy cattle farms at a regional scale. *Veterinary Research*, 46. ISSN 0928-4249. doi : 10.1186/s13567-015-0247-3.
- Begg D. J, de Silva K, Carter N, Plain K. M, Purdie A et Whittington R. J. **2011**. Does a Th1 over Th2 dominancy really exist in the early stages of *Mycobacterium avium* subspecies paratuberculosis infections? *Immunobiology*, 216(7) :840–846. ISSN 0171-2985. doi : 10.1016/j.imbio.2010.12.004.
- Begg D. J, Purdie A. C, de Silva K, Dhand N. K, Plain K. M et Whittington R. J. **2017**. Variation in susceptibility of different breeds of sheep to *Mycobacterium avium* subspecies paratuberculosis following experimental inoculation. *Veterinary Research*, 48 :36. ISSN 1297-9716. doi : 10.1186/s13567-017-0440-7.
- Ben Romdhane R, Beaunée G, Camanes G, Guatteo R, Fourichon C et Ezanno P. **2017**. Which phenotypic traits of resistance should be improved in cattle to control paratuberculosis dynamics

- in a dairy herd : a modelling approach. *Veterinary Research*, 48(1) :62. ISSN 1297-9716. doi : 10.1186/s13567-017-0468-8.
- Benedictus G, Verhoeff J, Schukken Y. H et Hesselink J. W. **2000**. Dutch paratuberculosis programme history, principles and development. *Veterinary microbiology*, 77(3) :399–413.
- Berghaus R. D, Farver T. B, Anderson R. J, Jaravata C. C et Gardner I. A. **2006a**. Environmental Sampling for Detection of *Mycobacterium avium* ssp. *paratuberculosis* on Large California Dairies. *Journal of dairy science*, 89(3) :963–970.
- Berghaus R. D, Farver T. B, Anderson R. J, Adaska J. M et Gardner I. A. **2006b**. Use of age and milk production data to improve the ability of enzyme-linked immunosorbent assay test results to predict *Mycobacterium avium* ssp. *paratuberculosis* fecal culture status. *Journal of veterinary diagnostic investigation*, 18(3) :233–242.
- Bermingham M. L, Handel I. G, Glass E. J, Woolliams J. A, Bronsvort B. M. d. C, McBride S. H, Skuce R. A, Allen A. R, McDowell S. W. J et Bishop S. C. **2015**. Hui and Walter’s latent-class model extended to estimate diagnostic test properties from surveillance data : a latent model for latent data. *Scientific Reports*, 5 :11861. ISSN 2045-2322. doi : 10.1038/srep11861.
- Bhattarai B, Fosgate G. T, Osterstock J. B, Fossler C. P, Park S. C et Roussel A. J. **2013**. Comparison of calf weaning weight and associated economic variables between beef cows with and without serum antibodies against or isolation from feces of *Mycobacterium avium* subsp *paratuberculosis*. *Journal of the American Veterinary Medical Association*, 243(11) :1609–1615.
- Bossuyt P. M, Reitsma J. B, Bruns D. E, Gatsonis C. A, Glasziou P. P, Irwig L. M, Moher D, Rennie D, de Vet H. C. W, Lijmer J. G et Standards for Reporting of Diagnostic Accuracy. **2003**. The STARD statement for reporting studies of diagnostic accuracy : explanation and elaboration. *Annals of Internal Medicine*, 138(1) :W1–12. ISSN 1539-3704.
- Botsaris G, Swift B. M, Slana I, Liapi M, Christodoulou M, Hatzitofi M, Christodoulou V et Rees C. E. **2016**. Detection of viable *Mycobacterium avium* subspecies *paratuberculosis* in powdered infant formula by phage-PCR and confirmed by culture. *International Journal of Food Microbiology*, 216 :91–94. ISSN 01681605. doi : 10.1016/j.ijfoodmicro.2015.09.011.
- Brenner H et Gefeller O. **1997**. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Statistics in Medicine*, 16(9) :981–991. ISSN 0277-6715.

- Brinkhof J. M. A, Houwers D. J et van Maanen C. **2007**. Development of a sample pooling strategy for the serodiagnosis of small ruminant lentiviral infections using the ELITEST-MVV ELISA. *Small Ruminant Research*, 70(2-3) :194-199. ISSN 0921-4488. doi : 10.1016/j.smallrumres.2006.03.003.
- Carta T, Álvarez J, Pérez de la Lastra J. M et Gortázar C. **2013**. Wildlife and paratuberculosis : A review. *Research in Veterinary Science*, 94(2) :191-197. ISSN 0034-5288. doi : 10.1016/j.rvsc.2012.11.002.
- Catton B. A. **2002**. Paucibacillary paratuberculosis in a goat. *The Canadian Veterinary Journal*, 43 (10) :787-788. ISSN 0008-5286.
- Chamberlin W. M et Naser S. A. **2006**. Integrating theories of the etiology of Crohn's disease. On the etiology of Crohn's disease : questioning the hypotheses. *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research*, 12(2) :RA27-33. ISSN 1234-1010.
- Chase C. C. L, Hurley D. J et Reber A. J. **2008**. Neonatal Immune Development in the Calf and Its Impact on Vaccine Response. *Veterinary Clinics of North America : Food Animal Practice*, 24(1) : 87-104. ISSN 0749-0720. doi : 10.1016/j.cvfa.2007.11.001.
- Chiodini R. **1992**. History of Paratuberculosis. *International Association for Paratuberculosis, Madison*, page 658.
- Chiodini R. J et Hermon-Taylor J. **1993**. The thermal resistance of Mycobacterium paratuberculosis in raw milk under conditions simulating pasteurization. *Journal of Veterinary Diagnostic Investigation*, 5(4) :629-631.
- Christopher-Hennings J, Dammen M. A, Weeks S. R, Epperson W. B, Singh S. N, Steinlicht G. L, Fang Y, Skaare J. L, Larsen J. L, Payeur J. B et Nelson E. A. **2003**. Comparison of Two DNA Extractions and Nested PCR, Real-Time PCR, a New Commercial PCR Assay, and Bacterial Culture for Detection of Mycobacterium Avium Subsp. Paratuberculosis in Bovine Feces. *Journal of Veterinary Diagnostic Investigation*, 15(2) :87-93. ISSN 1040-6387. doi : 10.1177/104063870301500201.
- Clarke C. J. **1997**. The pathology and pathogenesis of paratuberculosis in ruminants and other species. *Journal of comparative pathology*, 116(3) :217-261.

- Clarke C. J et Little D. **1996**. The pathology of ovine paratuberculosis : Gross and histological changes in the intestine and other tissues. *Journal of Comparative Pathology*, 114(4) :419–437. ISSN 0021-9975. doi : 10.1016/S0021-9975(96)80017-X.
- Collins D. M, Gabric D. M et de Lisle G. W. **1990**. Identification of two groups of Mycobacterium paratuberculosis strains by restriction endonuclease analysis and DNA hybridization. *Journal of Clinical Microbiology*, 28(7) :1591–1596. ISSN 0095-1137.
- Collins D. M, Gabric D. M et De Lisle G. W. **1989**. Identification of a repetitive DNA sequence specific to Mycobacterium paratuberculosis. *FEMS Microbiology Letters*, 60(2) :175–178. ISSN 1574-6968. doi : 10.1111/j.1574-6968.1989.tb03440.x.
- Collins J et Huynh M. **2014**. Estimation of diagnostic test accuracy without full verification : a review of latent class methods. *Statistics in medicine*, 33(24) :4141–4169. ISSN 0277-6715. doi : 10.1002/sim.6218.
- Collins M. T. **2003**. Update on paratuberculosis : 1. Epidemiology of Johne's disease and the biology of Mycobacterium paratuberculosis. *Irish Veterinary Journal*, 56(11) :565–574.
- Collins M. T et Morgan I. R. **1992**. Simulation model of paratuberculosis control in a dairy herd. *Preventive Veterinary Medicine*, 14(1) :21–32. ISSN 0167-5877. doi : 10.1016/0167-5877(92)90081-P.
- Collins M, Eggleston V et Manning E. **2010**. Successful control of Johne's disease in nine dairy herds : Results of a six-year field trial. *Journal of Dairy Science*, 93(4) :1638–1643. ISSN 00220302. doi : 10.3168/jds.2009-2664.
- Cousins D, Whittington R, Marsh I, Masters A, Evans R et Kluver P. **1999**. Mycobacteria distinct from Mycobacterium avium subsp. paratuberculosis isolated from the faeces of ruminants possess IS 900 -like sequences detectable by IS 900 polymerase chain reaction : implications for diagnosis. *Molecular and Cellular Probes*, 13(6) :431–442. ISSN 0890-8508. doi : 10.1006/mcpr.1999.0275.
- Coussens P. M, Verman N, Coussens M. A, Elftman M. D et McNulty A. M. **2004**. Cytokine Gene Expression in Peripheral Blood Mononuclear Cells and Tissues of Cattle Infected with Mycobacterium avium subsp. paratuberculosis : Evidence for an Inherent Proinflammatory

- Gene Expression Pattern. *Infection and Immunity*, 72(3) :1409–1422. ISSN 0019-9567. doi : 10.1128/IAI.72.3.1409-1422.2004.
- Cowling D. W, Gardner I. A et Johnson W. O. **1999**. Comparison of methods for estimation of individual-level prevalence based on pooled samples. *Preventive veterinary medicine*, 39(3) : 211–225.
- Dalziel T. K. **1989**. Chronic interstitial enteritis. *Diseases of the Colon and Rectum*, 32(12) :1076–1078. ISSN 0012-3706.
- David H. L, Clavel S et Clement F. **1980a**. Adsorption and growth of the bacteriophage D29 in selected mycobacteria. *Annales de l'Institut Pasteur / Virologie*, 131(2) :167–184. ISSN 0769-2617. doi : 10.1016/0769-2617(80)90031-3.
- David H. L, Clavel S, Clement F et Moniz-Pereira J. **1980b**. Effects of antituberculosis and anti-leprosy drugs on mycobacteriophage D29 growth. *Antimicrobial Agents and Chemotherapy*, 18 (2) :357–359. ISSN 0066-4804.
- de Silva K, Begg D et Whittington R. **2011**. The interleukin 10 response in ovine Johne's disease. *Veterinary Immunology and Immunopathology*, 139(1) :10–16. ISSN 0165-2427. doi : 10.1016/j.vetimm.2010.07.022.
- de Silva K, Begg D. J, Plain K. M, Purdie A. C, Kawaji S, Dhand N. K et Whittington R. J. **2013**. Can early host responses to mycobacterial infection predict eventual disease outcomes? *Preventive Veterinary Medicine*, 112(3–4) :203–212. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2013.08.006.
- de Silva K, Plain K, Purdie A, Begg D et Whittington R. **2018**. Defining resilience to mycobacterial disease : Characteristics of survivors of ovine paratuberculosis. *Veterinary Immunology and Immunopathology*, 195 :56–64. ISSN 0165-2427. doi : 10.1016/j.vetimm.2017.11.008.
- Delbrouck Y. Valeur informative d'outils utilisables à grande échelle pour déterminer le statut de troupeaux bovins laitiers vis-à-vis de l'infection à *Mycobacterium avium* spp paratuberculosis. Thèse pour le Doctorat Vétérinaire 2008NANT090V, Faculté de Médecine de Nantes, 2008.
- Dendukuri N et Joseph L. **2001**. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*, 57(1) :158–167.

- Dhand N. K, Johnson W. O et Toribio J.-A. L. M. L. **2010a**. A Bayesian Approach to Estimate OJD Prevalence From Pooled Fecal Samples of Variable Pool Size. *Journal of Agricultural, Biological and Environmental Statistics*, 15(4) :452–473. ISSN 1085-7117, 1537-2693. doi : 10.1007/s13253-010-0032-8.
- Dhand N. K, Sergeant E, Toribio J.-A. L et Whittington R. J. **2010b**. Estimation of sensitivity and flock-sensitivity of pooled faecal culture for Mycobacterium avium subsp. paratuberculosis in sheep. *Preventive Veterinary Medicine*, 95(3-4) :248–257. ISSN 01675877. doi : 10.1016/j.prevetmed.2010.03.013.
- Dhand N. K, Johnson W. O, Eppleston J, Whittington R. J et Windsor P.A. **2013**. Comparison of pre- and post-vaccination ovine Johne's disease prevalence using a Bayesian approach. *Preventive Veterinary Medicine*, 111(1-2) :81–91. ISSN 01675877. doi : 10.1016/j.prevetmed.2013.03.003.
- Dhand N. K, Eppleston J, Whittington R. J et Windsor P.A. **2016**. Changes in prevalence of ovine paratuberculosis following vaccination with Gudair® : Results of a longitudinal study conducted over a decade. *Vaccine*, 34(42) :5107–5113. ISSN 0264410X. doi : 10.1016/j.vaccine.2016.08.064.
- Diéguez F. J, González A. M, Menéndez S, Vilar M. J, Sanjuán M. L, Yús E et Arnaiz I. **2009**. Evaluation of four commercial serum ELISAs for detection of Mycobacterium avium subsp. paratuberculosis infection in dairy cows. *The Veterinary Journal*, 180(2) :231–235. ISSN 1090-0233. doi : 10.1016/j.tvjl.2007.11.004.
- Donat K, Hahn N, Eisenberg T, Schlez K, Köhler H, Wilfried W, Rohde M, Pützschel R, Rösler U, Failing K et Zschöck M. **2015**. Within-herd prevalence thresholds for the detection of Mycobacterium avium subspecies paratuberculosis-positive dairy herds using boot swabs and liquid manure samples. *Epidemiology and Infection*, -1 :1–12. doi : 10.1017/S0950268815000977.
- Dorfman R. **1943**. The Detection of Defective Members of Large Populations. *The Annals of Mathematical Statistics*, 14(4) :436–440. ISSN 0003-4851, 2168-8990. doi : 10.1214/aoms/1177731363.
- Eamens G, Marsh I. M, Plain K. M et Whittington R. Paratuberculosis (Johne's Disease). Technical report, Australian and New Zealand Standard Diagnostic Procedure, 2015.
- Eamens G. J, Whittington R. J, Turner M. J, Austin S. L, Fell S. A et Marsh I. B. **2007**. Evaluation of radiometric faecal culture and direct PCR on pooled faeces for detection of Mycobacterium avium

- subsp. paratuberculosis in cattle. *Veterinary Microbiology*, 125(1-2) :22–35. ISSN 03781135. doi : 10.1016/j.vetmic.2007.04.043.
- Eisenberg S. W. F, Nielen M, Hoeboer J, Rutten V, Heederik D et Koets A. P. **2012**. Environmental contamination with Mycobacterium avium subspecies paratuberculosis within and around a dairy barn under experimental conditions. *Journal of Dairy Science*, 95(11) :6477–6482. ISSN 0022-0302. doi : 10.3168/jds.2012-5548.
- Eisenberg S. W. F, Chuchaisangrat R, Nielen M et Koets A. P. **2013**. Relationship between Presence of Cows with Milk Positive for Mycobacterium avium subsp. paratuberculosis-Specific Antibody by Enzyme-Linked Immunosorbent Assay and Viable M. avium subsp. paratuberculosis in Dust in Cattle Barns. *Applied and Environmental Microbiology*, 79(18) :5458–5464. ISSN 0099-2240. doi : 10.1128/AEM.01187-13.
- Ellingson J. L. E, Bolin C. A et Stabel J. R. **1998**. Identification of a gene unique to Mycobacterium avium subspecies paratuberculosis and application to diagnosis of paratuberculosis. *Molecular and Cellular Probes*, 12(3) :133–142. ISSN 0890-8508. doi : 10.1006/mcpr.1998.0167.
- Ellingson J. L. E, Anderson J. L, Koziczowski J. J, Radcliff R. P, Sloan S. J, Allen S. E et Sullivan N. M. **2005**. Detection of viable Mycobacterium avium subsp. paratuberculosis in retail pasteurized whole milk by two culture methods and PCR. *Journal of Food Protection*, 68(5) :966–972. ISSN 0362-028X.
- Elmore R, Hall P et Neeman A. **2005**. An application of classical invariant theory to identifiability in nonparametric mixtures. *Annales de l'institut Fourier*, 55(1) :1–28. ISSN 0373-0956. doi : 10.5802/aif.2087.
- Englund S, Bölske G et Johansson K.-E. **2002**. An IS900-like sequence found in a Mycobacterium sp. other than Mycobacterium avium subsp. paratuberculosis. *FEMS Microbiology Letters*, 209(2) :267–271. ISSN 1574-6968. doi : 10.1111/j.1574-6968.2002.tb11142.x.
- Eppleston J et Whittington R. **2001**. Isolation of Mycobacterium avium subsp paratuberculosis from the semen of rams with clinical Johne's disease. *Australian Veterinary Journal*, 79(11) : 776–777. ISSN 1751-0813. doi : 10.1111/j.1751-0813.2001.tb10898.x.
- Eppleston J et Windsor P. A. **2007**. Lesions attributed to vaccination of sheep with Gudair for

-
- the control of ovine paratuberculosis : post farm economic impacts at slaughter. *Australian Veterinary Journal*, 85(4) :129–133. ISSN 0005-0423. doi : 10.1111/j.0005-0423.2007.00135.x.
- Eppleston J, Begg D. J, Dhand N. K, Watt B et Whittington R. J. **2014**. Environmental Survival of *Mycobacterium avium* subsp. paratuberculosis in Different Climatic Zones of Eastern Australia. *Applied and Environmental Microbiology*, 80(8) :2337–2342. ISSN 0099-2240. doi : 10.1128/AEM.03630-13.
- Falkinham J. O. **2003**. Mycobacterial Aerosols and Respiratory Disease. *Emerging Infectious Diseases*, 9(7) :763–767. ISSN 1080-6040. doi : 10.3201/eid0907.02-0415.
- Fecteau M.-E, Hovingh E, Whitlock R. H et Sweeney R. W. **2013**. Persistence of *Mycobacterium avium* subsp. paratuberculosis in soil, crops, and ensiled feed following manure spreading on infected dairy farms. *The Canadian Veterinary Journal*, 54(11) :1083–1085. ISSN 0008-5286.
- Fernández M, Benavides J, Sevilla I. A, Fuertes M, Castaño P, Delgado L, Marín J. F. G, Garrido J. M, Ferreras M. C et Pérez V. **2014**. Experimental infection of lambs with C and S-type strains of *Mycobacterium avium* subspecies paratuberculosis : immunological and pathological findings. *Veterinary Research*, 45(1) :5. ISSN 0928-4249. doi : 10.1186/1297-9716-45-5.
- Fernández-Silva J. A, Correa-Valencia N. M et Ramírez N. F. **2014**. Systematic review of the prevalence of paratuberculosis in cattle, sheep, and goats in Latin America and the Caribbean. *Tropical Animal Health and Production*, 46(8) :1321–1340. ISSN 0049-4747, 1573-7438. doi : 10.1007/s11250-014-0656-8.
- Fock-Chow-Tho D, Topp E, Ibeagha-Awemu E et Bissonnette N. **2016**. Comparison of commercial DNA extraction kits and quantitative PCR systems for better sensitivity in detecting the causative agent of paratuberculosis in dairy cow fecal samples. *Journal of Dairy Science*. ISSN 00220302. doi : 10.3168/jds.2016-11384.
- Foddai A et Grant I. **2017**. Sensitive and specific detection of viable *Mycobacterium avium* subsp. paratuberculosis in raw milk by the peptide-mediated magnetic separation-phage assay. *Journal of Applied Microbiology*, 122(5) :1357–1367. ISSN 1365-2672. doi : 10.1111/jam.13425.
- Foley-Thomas E. M, Whipple D. L, Bermudez L. E et Barletta R. G. **1995**. Phage infection, transfection and transformation of *Mycobacterium avium* complex and *Mycobacterium paratuberculosis*. *Microbiology*, 141(5) :1173–1181.
-

- Frederic E. Informative value of environmental and pooled samples to evaluate the within-herd prevalence of mycobacterium avium subspecies paratuberculosis in g..., 2016. URL : <https://colloque.inra.fr/icp2016/layout/set/print/Submission/Informative-value-of-environmental-and-pooled-samples-to-evaluate-the-within-herd-prevalence-of-mycobacterium-avium-subspecies-paratuberculosis-in-g...> (offset)/1. [Consulté le : 2018-01-03].
- Frössling J, Bonnett B, Lindberg A et Björkman C. **2003**. Validation of a *Neospora caninum* iscom ELISA without a gold standard. *Preventive Veterinary Medicine*, 57(3) :141–153. ISSN 0167-5877. doi : 10.1016/S0167-5877(02)00216-7.
- Gardner I. A, Nielsen S. S, Whittington R. J, Collins M. T, Bakker D, Harris B, Sreevatsan S, Lombard J. E, Sweeney R, Smith D. R, Gavalchin J et Eda S. **2011**. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. *Preventive Veterinary Medicine*, 101(1–2) :18–34. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2011.04.002.
- Garrido J. M, Vazquez P, Molina E, Plazaola J. M, Sevilla I. A, Geijo M. V, Alonso-Hearn M et Juste R. A. **2013**. Paratuberculosis Vaccination Causes Only Limited Cross-Reactivity in the Skin Test for Diagnosis of Bovine Tuberculosis. *PLOS ONE*, 8(11) :e80985. ISSN 1932-6203. doi : 10.1371/journal.pone.0080985.
- Georgiadis M. P, Johnson W. O, Gardner I. A et Singh R. **2003**. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 52(1) :63–76.
- Gillan S, O'Brien R, Hughes A. D et Griffin J. E. T. **2010**. Identification of Immune Parameters To Differentiate Disease States among Sheep Infected with *Mycobacterium avium* subsp. paratuberculosis. *Clinical and Vaccine Immunology : CVI*, 17(1) :108–117. ISSN 1556-6811. doi : 10.1128/CVI.00359-09.
- Gonda M, Chang Y, Shook G, Collins M et Kirkpatrick B. **2007**. Effect of *Mycobacterium paratuberculosis* infection on production, reproduction, and health traits in US Holsteins. *Preventive Veterinary Medicine*, 80(2-3) :103–119. ISSN 01675877. doi : 10.1016/j.prevetmed.2007.01.011.
- Goodman L. A. **1974**. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61(2) :215–231.

- Grant I. R, Foddai A. C. G, Tarrant J. C, Kunkel B, Hartmann F. A, McGuirk S, Hansen C, Talaat A. M et Collins M. T. **2017**. Viable Mycobacterium avium ssp. paratuberculosis isolated from calf milk replacer. *Journal of Dairy Science*. ISSN 0022-0302. doi : 10.3168/jds.2017-13154.
- Green E. P, Tizard M. L. V, Moss M. T, Thompson J, Winterbourne D. J, McFadden J. J et Hermon-Taylor J. **1989**. Sequence and characteristics of IS900, an insertion element identified in a human Crohn's disease isolate of Mycobacterium paratuberculosis. *Nucleic Acids Research*, 17 (22) :9063–9073.
- Grewal S. K, Rajeev S, Sreevatsan S et Michel F. C. **2006**. Persistence of Mycobacterium avium subsp. paratuberculosis and Other Zoonotic Pathogens during Simulated Composting, Manure Packing, and Liquid Storage of Dairy Manure. *Applied and Environmental Microbiology*, 72(1) : 565–574. ISSN 0099-2240. doi : 10.1128/AEM.72.1.565-574.2006.
- Græsbøll K, Andresen L. O, Halasa T et Toft N. **2017**. Opportunities and challenges when pooling milk samples using ELISA. *Preventive Veterinary Medicine*, 139 :93–98. ISSN 01675877. doi : 10.1016/j.prevetmed.2016.08.001.
- Guatteo R, Joly A, Beaudeau F, Chognard T et Fourichon C. **2012**. EVALUATION OF COMBINED ENVIRONMENTAL AND POOLED ANIMALS SAMPLING TO DETERMINE DAIRY CATTLE HERD INFECTION STATUS FOR MYCOBACTERIUM AVIUM SUBSPECIES PARATUBERCULOSIS (MAP). In *Proceedings of the 11th International Colloquium on Paratuberculosis, Sydney, Australia, 5-10 February 2012*, pages 349–351. International Association for Paratuberculosis.
- Guescini M, Sisti D, Rocchi M. B, Stocchi L et Stocchi V. **2008**. A new real-time PCR method to overcome significant quantitative inaccuracy due to slight amplification inhibition. *BMC Bioinformatics*, 9(1) :326. ISSN 1471-2105. doi : 10.1186/1471-2105-9-326.
- Hahn N, Failing K, Eisenberg T, Schlez K, Zschöck P.-M, Donat K, Einax E et Köhler H. **2017**. Evaluation of different diagnostic methods for the detection of Mycobacterium avium subsp. paratuberculosis in boot swabs and liquid manure samples. *BMC Veterinary Research*, 13. ISSN 1746-6148. doi : 10.1186/s12917-017-1173-6.
- Heid C. A, Stevens J, Livak K. J et Williams P. M. **1996**. Real time quantitative PCR. *Genome Research*, 6(10) :986–994. ISSN 1088-9051.

- Herthnek D, Nielsen S. S, Lindberg A et Bölske G. **2008**. A robust method for bacterial lysis and DNA purification to be used with real-time PCR for detection of *Mycobacterium avium* subsp. paratuberculosis in milk. *Journal of Microbiological Methods*, 75(2) :335–340. ISSN 0167-7012. doi : 10.1016/j.mimet.2008.07.009.
- Hickey S, Morris C, Dobbie J et Lake D. **2003**. Heritability of Johne's disease and survival data from Romney and Merino sheep. In *Proceedings of the New Zealand Society of Animal Production*, volume 63, pages 179–182, Queenstown. New Zealand Society of Animal Production.
- Hilbink F et West D. **1990**. The antibody response of sheep to vaccination against Johne's disease. *New Zealand Veterinary Journal*, 38(4) :168–169. ISSN 0048-0169, 1176-0710. doi : 10.1080/00480169.1990.35646.
- Hines M. E, Stabel J. R, Sweeney R. W, Griffin F, Talaat A. M, Bakker D, Benedictus G, Davis W. C, de Lisle G. W, Gardner I. A, Juste R. A, Kapur V, Koets A, McNair J, Pruitt G et Whitlock R. H. **2007**. Experimental challenge models for Johne's disease : A review and proposed international guidelines. *Veterinary Microbiology*, 122(3) :197–222. ISSN 0378-1135. doi : 10.1016/j.vetmic.2007.03.009.
- Hoogendam K, Richardson E et Mee J. **2009**. Paratuberculosis sero-status and milk production, SCC and calving interval in Irish dairy herds. *Irish Veterinary Journal*, 62(Suppl 4) :265–271. ISSN 0368-0762. doi : 10.1186/2046-0481-62-4-265.
- Hostetter J, Kagan R et Steadham E. **2005**. Opsonization Effects on *Mycobacterium avium* subsp. paratuberculosis-Macrophage Interactions. *Clinical and Diagnostic Laboratory Immunology*, 12(6) :793–796. ISSN 1071-412X. doi : 10.1128/CDLI.12.6.793-796.2005.
- Hui S. L et Walter S. D. **1980**. Estimating the Error Rates of Diagnostic Tests. *Biometrics*, 36(1) :167. ISSN 0006341X. doi : 10.2307/2530508.
- Johnson Y. J, Kaneene J. B, Gardiner J. C, Lloyd J. W, Sprecher D. J et Coe P. H. **2001**. The Effect of Subclinical *Mycobacterium paratuberculosis* Infection on Milk Production in Michigan Dairy Cows. *Journal of Dairy Science*, 84(10) :2188–2194. ISSN 0022-0302. doi : 10.3168/jds.S0022-0302(01)74665-6.

- Jones G, Johnson W. O, Hanson T. E et Christensen R. **2010**. Identifiability of Models for Multiple Diagnostic Testing in the Absence of a Gold Standard. *Biometrics*, 66(3) :855–863. ISSN 0006341X. doi : 10.1111/j.1541-0420.2009.01330.x.
- Jones W. D et David H. L. **1971**. Inhibition by Rifampin of Mycobacteriophage D29 Replication in its Drug-Resistant Host, Mycobacterium smegmatis ATCC 607. *American Review of Respiratory Disease*, 103(5) :618–624. ISSN 0003-0805. doi : 10.1164/arrd.1971.103.5.618.
- Judge J, Davidson R. S, Marion G, White P. C et Hutchings M. R. **2007**. Persistence of Mycobacterium avium subspecies paratuberculosis in rabbits : the interplay between horizontal and vertical transmission. *Journal of Applied Ecology*, 44(2) :302–311. ISSN 1365-2664. doi : 10.1111/j.1365-2664.2007.01282.x.
- Jungersen G, Mikkelsen H et Grell S. N. **2012**. Use of the johnin PPD interferon-gamma assay in control of bovine paratuberculosis. *Veterinary Immunology and Immunopathology*, 148(1) : 48–54. ISSN 0165-2427. doi : 10.1016/j.vetimm.2011.05.010.
- Kaevska M, Lvoncik S, Slana I, Kulich P et Kralik P. **2014**. Microscopy, Culture, and Quantitative Real-Time PCR Examination Confirm Internalization of Mycobacteria in Plants. *Applied and Environmental Microbiology*, 80(13) :3888–3894. ISSN 0099-2240. doi : 10.1128/AEM.00496-14.
- Kalis C, Collins M, Barkema H et Hesselink J. **2004**. Certification of herds as free of Mycobacterium paratuberculosis infection : actual pooled faecal results versus certification model predictions. *Preventive Veterinary Medicine*, 65(3-4) :189–204. ISSN 01675877. doi : 10.1016/j.prevetmed.2004.07.005.
- Kawaji S, Taylor D. L, Mori Y et Whittington R. J. **2007**. Detection of Mycobacterium avium subsp. paratuberculosis in ovine faeces by direct quantitative PCR has similar or greater sensitivity compared to radiometric culture. *Veterinary Microbiology*, 125(1-2) :36–48. ISSN 03781135. doi : 10.1016/j.vetmic.2007.05.002.
- Kawaji S, Begg D. J, Plain K. M et Whittington R. J. **2011**. A longitudinal study to evaluate the diagnostic potential of a direct faecal quantitative PCR test for Johne's disease in sheep. *Veterinary Microbiology*, 148(1) :35–44. ISSN 03781135. doi : 10.1016/j.vetmic.2010.07.022.
- Khare S, Adams L. G, Osterstock J, Roussel A et David L. **2008**. Effects of Shipping and Storage

- Conditions of Fecal Samples on Viability of Mycobacterium paratuberculosis. *Journal of Clinical Microbiology*, 46(4) :1561–1562. ISSN 0095-1137. doi : 10.1128/JCM.02348-07.
- Khare S, Drake K. L, Lawhon S. D, Nunes J. E. S, Figueiredo J. F, Rossetti C. A, Gull T, Everts R. E, Lewin H. A et Adams L. G. **2016**. Systems Analysis of Early Host Gene Expression Provides Clues for Transient Mycobacterium avium ssp avium vs. Persistent Mycobacterium avium ssp paratuberculosis Intestinal Infections. *PLOS ONE*, 11(9) :e0161946. ISSN 1932-6203. doi : 10.1371/journal.pone.0161946.
- Kirkpatrick B. W, Shi X, Shook G. E et Collins M. T. **2011**. Whole-Genome association analysis of susceptibility to paratuberculosis in Holstein cattle. *Animal Genetics*, 42(2) :149–160. ISSN 1365-2052. doi : 10.1111/j.1365-2052.2010.02097.x.
- Kirkpatrick B. W et Shook G. E. **2011**. Genetic Susceptibility to Paratuberculosis. *Veterinary Clinics of North America : Food Animal Practice*, 27(3) :559–571. ISSN 07490720. doi : 10.1016/j.cvfa.2011.07.003.
- Kiser J. N, Neupane M, White S. N et Neibergs H. L. **2017**. Identification of genes associated with susceptibility to *Mycobacterium avium* ssp. *paratuberculosis* (*Map*) tissue infection in Holstein cattle using gene set enrichment analysis–SNP. *Mammalian Genome*, pages 1–11. ISSN 0938-8990, 1432-1777. doi : 10.1007/s00335-017-9725-4.
- Klanicova B, Slana I, Roubal P, Pavlik I et Kralik P. **2012**. Mycobacterium avium subsp. paratuberculosis survival during fermentation of soured milk products detected by culture and quantitative real time PCR methods. *International Journal of Food Microbiology*, 157(2) :150–155. ISSN 0168-1605. doi : 10.1016/j.ijfoodmicro.2012.04.021.
- Koets A. P, Eda S et Sreevatsan S. **2015**. The within host dynamics of Mycobacterium avium ssp. paratuberculosis infection in cattle : where time and place matter. *Veterinary Research*, 46(1). ISSN 1297-9716. doi : 10.1186/s13567-015-0185-0.
- Kostoulas P, Leontides L, Billinis C, Amiridis G et Florou M. **2006a**. The association of sub-clinical paratuberculosis with the fertility of Greek dairy ewes and goats varies with parity. *Preventive Veterinary Medicine*, 74(2-3) :226–238. ISSN 01675877. doi : 10.1016/j.prevetmed.2005.12.001.

- Kostoulas P, Leontides L, Billinis C et Florou M. **2006b**. Application of a semi-dependent latent model in the Bayesian estimation of the sensitivity and specificity of two faecal culture methods for diagnosis of paratuberculosis in sub-clinically infected Greek dairy sheep and goats. *Preventive Veterinary Medicine*, 76(1-2) :121–134. ISSN 01675877. doi : 10.1016/j.prevetmed.2006.04.008.
- Kostoulas P, Leontides L, Enøe C, Billinis C, Florou M et Sofia M. **2006c**. Bayesian estimation of sensitivity and specificity of serum ELISA and faecal culture for diagnosis of paratuberculosis in Greek dairy sheep and goats. *Preventive Veterinary Medicine*, 76(1-2) :56–73. ISSN 01675877. doi : 10.1016/j.prevetmed.2006.04.006.
- Kostoulas P, Nielsen S. S, Branscum A. J, Johnson W. O, Dendukuri N, Dhand N. K, Toft N et Gardner I. A. **2017**. STARD-BLCM : Standards for the Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models. *Preventive Veterinary Medicine*, 138 :37–47. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2017.01.006.
- Kralik P, Babak V et Dziedzinska R. **2014**. Repeated cycles of chemical and physical disinfection and their influence on Mycobacterium avium subsp. paratuberculosis viability measured by propidium monoazide F57 quantitative real time PCR. *The Veterinary Journal*, 201(3) :359–364. ISSN 1090-0233. doi : 10.1016/j.tvjl.2014.05.032.
- Krishnan M. Y, Manning E. J. B et Collins M. T. **2009**. Comparison of three methods for susceptibility testing of Mycobacterium avium subsp. paratuberculosis to 11 antimicrobial drugs. *Journal of Antimicrobial Chemotherapy*, 64(2) :310–316. ISSN 0305-7453. doi : 10.1093/jac/dkp184.
- Kruze J, Monti G, Schulze F, Mella A et Leiva S. **2013**. Herd-level prevalence of Map infection in dairy herds of southern Chile determined by culture of environmental fecal samples and bulk-tank milk qPCR. *Preventive Veterinary Medicine*, 111(3) :319–324. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2013.05.011.
- Kruze J, Salgado M, Paredes E, Mella A et Collins M. T. **2006**. Goat Paratuberculosis in Chile : First Isolation and Confirmation of Mycobacterium Avium Subspecies Paratuberculosis Infection in a Dairy Goat. *Journal of Veterinary Diagnostic Investigation*, 18(5) :476–479. ISSN 1040-6387. doi : 10.1177/104063870601800510.
- Kudahl A et Nielsen S. **2009**. Effect of paratuberculosis on slaughter weight and slaughter value

- of dairy cows. *Journal of Dairy Science*, 92(9) :4340–4346. ISSN 00220302. doi : 10.3168/jds.2009-2039.
- Kuenstner J. T, Naser S, Chamberlin W, Borody T, Graham D. Y, McNees A, Hermon-Taylor J, Hermon-Taylor A, Dow C. T, Thayer W, Biesecker J, Collins M. T, Sechi L. A, Singh S. V, Zhang P, Shafran I, Weg S, Telega G, Rothstein R, Oken H, Schimpff S, Bach H, Bull T, Grant I, Ellingson J, Dahmen H, Lipton J, Gupta S, Chaubey K, Singh M, Agarwal P, Kumar A, Misri J, Sohal J, Dhama K, Hemati Z, Davis W, Hier M, Aitken J, Pierce E, Parrish N, Goldberg N, Kali M, Bendre S, Agrawal G, Baldassano R, Linn P, Sweeney R. W, Fecteau M, Hofstaedter C, Potula R, Timofeeva O, Geier S, John K, Zayanni N, Malaty H. M, Kahlenborn C, Kravitz A, Bulfon A, Daskalopoulos G, Mitchell H, Neilan B, Timms V, Cossu D, Mameli G, Angermeier P, Jelic T, Goethe R, Juste R. A et Kuenstner L. **2017**. The Consensus from the Mycobacterium avium ssp. paratuberculosis (MAP) Conference 2017. *Frontiers in Public Health*, 5. ISSN 2296-2565. doi : 10.3389/fpubh.2017.00208.
- Lambeth C, Reddacliff L, Windsor P, Abbott K, McGREGOR H et Whittington R. **2004**. Intrauterine and transmammary transmission of Mycobacterium avium subsp paratuberculosis in sheep. *Australian Veterinary Journal*, 82(8) :504–508. ISSN 1751-0813. doi : 10.1111/j.1751-0813.2004.tb11171.x.
- Landis J. R et Koch G. G. **1977**. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1) :159. ISSN 0006341X. doi : 10.2307/2529310.
- Lee H, Stabel J. R et Kehrli M. E. **2001**. Cytokine gene expression in ileal tissues of cattle infected with Mycobacterium paratuberculosis. *Veterinary Immunology and Immunopathology*, 82(1) : 73–85. ISSN 0165-2427. doi : 10.1016/S0165-2427(01)00340-3.
- Leeflang M. M. G, Bossuyt P. M. M et Irwig L. **2009**. Diagnostic test accuracy may vary with prevalence : implications for evidence-based diagnosis. *Journal of Clinical Epidemiology*, 62(1) :5–12. ISSN 0895-4356. doi : 10.1016/j.jclinepi.2008.04.007.
- Leeflang M. M, Rutjes A. W, Reitsma J. B, Hooft L et Bossuyt P. M. **2013**. Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ : Canadian Medical Association Journal*, 185 (11) :E537–E544. ISSN 0820-3946. doi : 10.1503/cmaj.121286.
- Lefrançois L. H. **2012**. *Etude des adhésines HBHA et LBP impliquées dans l'interaction de Mycobac-*

terium avium ssp. paratuberculosis avec les cellules épithéliales intestinales, cibles privilégiées de la bactérie in vivo. PhD thesis. URL : <http://www.theses.fr/17861002X>.

Li L, Bannantine J. P, Zhang Q, Amonsin A, May B. J, Alt D, Banerji N, Kanjilal S et Kapur V. **2005**. The complete genome sequence of Mycobacterium avium subspecies paratuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, 102(35) :12344–12349. ISSN 0027-8424. doi : 10.1073/pnas.0505662102.

Lijmer J. G, Mol B. W, Heisterkamp S, Bossel G. J, Prins M. H, Meulen J. H. P. v. d et Bossuyt P. M. M. **1999**. Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests. *JAMA*, 282(11) : 1061–1066. ISSN 0098-7484. doi : 10.1001/jama.282.11.1061.

Litvak E, Tu X. M et Pagano M. **1994**. Screening for the Presence of a Disease by Pooling Sera Samples. *Journal of the American Statistical Association*, 89(426) :424–434. ISSN 0162-1459. doi : 10.2307/2290842.

Lombard J. E, Wagner B. A, Smith R. L, McCluskey B. J, Harris B. N, Payeur J. B, Garry F. B et Salman M. D. **2006**. Evaluation of Environmental Sampling and Culture to Determine Mycobacterium avium subspecies paratuberculosis Distribution and Herd Infection Status on US Dairy Operations. *Journal of Dairy Science*, 89(11) :4163–4171. ISSN 0022-0302. doi : 10.3168/jds.S0022-0302(06)72461-4.

Lombard J. E, Garry F. B, McCluskey B. J et Wagner B. A. **2005**. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. *Journal of the American Veterinary Medical Association*, 227(12) :1975–1981. ISSN 0003-1488. doi : 10.2460/javma.2005.227.1975.

Luu-The V, Paquet N, Calvo E et Cumps J. **2005**. Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction. *BioTechniques*, 38 :287–93. doi : 10.2144/05382RR05.

Maio E, Carta T, Balseiro A, Sevilla I. A, Romano A, Ortiz J. A, Vieira-Pinto M, Garrido J. M, de la Lastra J. M. P et Gortázar C. **2011**. Paratuberculosis in European wild rabbits from the Iberian Peninsula. *Research in Veterinary Science*, 91(2) :212–218. ISSN 0034-5288. doi : 10.1016/j.rvsc.2010.12.014.

- Manning E. J et Collins M. T. **2001**. *Mycobacterium avium* subsp. *paratuberculosis* : pathogen, pathogenesis and diagnosis. *Revue Scientifique Et Technique (International Office of Epizootics)*, 20(1) :133–150. ISSN 0253-1933.
- Marcé C, Ezanno P, Seegers H, Pfeiffer D. U et Fourichon C. **2011**. Within-herd contact structure and transmission of *Mycobacterium avium* subspecies *paratuberculosis* in a persistently infected dairy cattle herd. *Preventive Veterinary Medicine*, 100(2) :116–125. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2011.02.004.
- Mathevon Y, Foucras G, Falguières R et Corbiere F. **2017**. Estimation of the sensitivity and specificity of two serum ELISAs and one fecal qPCR for diagnosis of *paratuberculosis* in sub-clinically infected young-adult French sheep using latent class Bayesian modeling. *BMC Veterinary Research*, 13 :230. ISSN 1746-6148. doi : 10.1186/s12917-017-1145-x.
- Möbius P, Hotzel H, Raßbach A et Köhler H. **2008**. Comparison of 13 single-round and nested PCR assays targeting IS900, ISMav2, f57 and locus 255 for detection of *Mycobacterium avium* subsp. *paratuberculosis*. *Veterinary Microbiology*, 126(4) :324–333. ISSN 0378-1135. doi : 10.1016/j.vetmic.2007.07.016.
- McGregor H, Dhand N. K, Dhungyel O. P et Whittington R. J. **2012**. Transmission of *Mycobacterium avium* subsp. *paratuberculosis* : Dose–response and age-based susceptibility in a sheep model. *Preventive Veterinary Medicine*, 107(1-2) :76–84. ISSN 01675877. doi : 10.1016/j.prevetmed.2012.05.014.
- McGregor H, Abbott K. A et Whittington R. J. **2015**. Effects of *Mycobacterium avium* subsp. *paratuberculosis* infection on serum biochemistry, body weight and wool growth in Merino sheep : A longitudinal study. *Small Ruminant Research*, 125 :146–153. ISSN 0921-4488. doi : 10.1016/j.smallrumres.2015.02.004.
- Mercier P, Baudry C, Beaudeau F, Seegers H et Malher X. **2010**. Estimated prevalence of *Mycobacterium avium* subspecies *paratuberculosis* infection in herds of dairy goats in France. *Veterinary Record*, 167(11) :412–415. ISSN 0042-4900, 2042-7670. doi : 10.1136/vr.c4454.
- Merwe R. G. v. d, Helden P. D. v, M. Warren R, L. Sampson S et Pittius N. C. G. v. **2014**. Phage-based detection of bacterial pathogens. *Analyst*, 139(11) :2617–2626. doi : 10.1039/C4AN00208C.

- Minozzi G, Williams J. L, Stella A, Strozzi F, Luini M, Settles M. L, Taylor J. F, Whitlock R. H, Zanella R et Neibergs H. L. **2012**. Meta-Analysis of Two Genome-Wide Association Studies of Bovine Paratuberculosis. *PLOS ONE*, 7(3) :e32578. ISSN 1932-6203. doi : 10.1371/journal.pone.0032578.
- Mita A, Mori Y, Nakagawa T, Tasaki T, Utiyama K et Mori H. **2015**. Comparison of fecal pooling methods and DNA extraction kits for the detection of Mycobacterium avium subspecies paratuberculosis. *MicrobiologyOpen*, 5(1) :134–142. ISSN 2045-8827. doi : 10.1002/mbo3.318.
- Mitchell R. M, Schukken Y, Koets A, Weber M, Bakker D, Stabel J, Whitlock R. H et Louzoun Y. **2015**. Differences in intermittent and continuous fecal shedding patterns between natural and experimental Mycobacterium avium subspecies paratuberculosis infections in cattle. *Veterinary Research*, 46(1). ISSN 1297-9716. doi : 10.1186/s13567-015-0188-x.
- Münster P, Völkel I, Wemheuer W, Schwarz D, Döring S et Czerny C.-P. **2013**. A Longitudinal Study to Characterize the Distribution Patterns of Mycobacterium avium ssp. paratuberculosis in Semen, Blood and Faeces of a Naturally Infected Bull by IS 900 Semi-Nested and Quantitative Real-Time PCR. *Transboundary and Emerging Diseases*, 60(2) :175–187. ISSN 1865-1682. doi : 10.1111/j.1865-1682.2012.01336.x.
- Moloney B et Whittington R. **2008**. Cross species transmission of ovine Johne's disease from sheep to cattle : an estimate of prevalence in exposed susceptible cattle. *Australian Veterinary Journal*, 86(4) :117–123. ISSN 1751-0813. doi : 10.1111/j.1751-0813.2008.00272.x.
- Momotani E, Whipple D. L, Thiermann A. B et Cheville N. F. **1988**. Role of M Cells and Macrophages in the Entrance of Mycobacterium paratuberculosis into Domes of Ileal Peyer's Patches in Calves. *Veterinary Pathology*, 25(2) :131–137. ISSN 0300-9858. doi : 10.1177/030098588802500205.
- Moravkova M, Babak V, Kralova A, Pavlik I et Slana I. **2012**. Culture- and Quantitative IS900 Real-Time PCR-Based Analysis of the Persistence of Mycobacterium avium subsp. paratuberculosis in a Controlled Dairy Cow Farm Environment. *Applied and Environmental Microbiology*, 78(18) :6608–6614. ISSN 0099-2240. doi : 10.1128/AEM.01264-12.
- Mortier R. A, Barkema H. W, Bystrom J. M, Illanes O, Orsel K, Wolf R, Atkins G et De Buck J. **2013**. Evaluation of age-dependent susceptibility in calves infected with two doses of Mycobacterium

- avium subspecies paratuberculosis using pathology and tissue culture. *Veterinary Research*, 44 (1) :94. ISSN 0928-4249. doi : 10.1186/1297-9716-44-94.
- Mortier R. A, Barkema H. W, Orsel K, Wolf R et De Buck J. **2014**. Shedding patterns of dairy calves experimentally infected with *Mycobacterium avium* subspecies paratuberculosis. *Veterinary Research*, 45(1) :71. ISSN 0928-4249. doi : 10.1186/s13567-014-0071-1.
- Muñoz M, Garcia Marin J, Garcia-Pariente C, Reyes L, Verna A, Moreno O, Fuertes M, Doce J, Puentes E, Garrido J. M et Pérez V. **2005**. IAP : 2005 : Efficacy of a killed vaccine (SILIRUM®) in calves challenged with Map. Copenhagen, Denmark. URL : http://www.paratuberculosis.info/proc8/abst3a_p67.htm.
- Muratore E, Bertolotti L, Nogarol C, Caruso C, Lucchese L, Iotti B, Ariello D, Moresco A, Masoero L, Nardelli S et Rosati S. **2017**. Surveillance of Infectious Bovine Rhinotracheitis in marker-vaccinated dairy herds : Application of a recombinant gE ELISA on bulk milk samples. *Veterinary Immunology and Immunopathology*, 185 :1–6. ISSN 0165-2427. doi : 10.1016/j.vetimm.2017.01.003.
- Nebbia P, Robino P, Zoppi S et Meneghi D. D. **2006**. Detection and excretion pattern of *Mycobacterium avium* subspecies paratuberculosis in milk of asymptomatic sheep and goats by Nested-PCR. *Small Ruminant Research*, 66(1) :116–120. ISSN 0921-4488. doi : 10.1016/j.smallrumres.2005.07.049.
- Nielsen S. S, Thamsborg S. M, Houe H et Bitsch V. **2000**. Bulk-tank milk ELISA antibodies for estimating the prevalence of paratuberculosis in Danish dairy herds. *Preventive Veterinary Medicine*, 44(1) :1–7. ISSN 0167-5877. doi : 10.1016/S0167-5877(00)00098-2.
- Nielsen S. S, Bjerre H et Toft N. **2008**. Colostrum and Milk as Risk Factors for Infection with *Mycobacterium avium* subspecies paratuberculosis in Dairy Cattle. *Journal of Dairy Science*, 91 (12) :4610–4615. ISSN 0022-0302. doi : 10.3168/jds.2008-1272.
- Nielsen S. S, Grønbaek C, Agger J. F et Houe H. **2002**. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. *Preventive Veterinary Medicine*, 53(3) :191–204. ISSN 0167-5877. doi : 10.1016/S0167-5877(01)00280-X.
- Nielsen S. S et Toft N. **2009**. A review of prevalences of paratuberculosis in farmed animals in

-
- Europe. *Preventive Veterinary Medicine*, 88(1) :1–14. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2008.07.003.
- Nielsen S et Toft N. **2008**. Ante mortem diagnosis of paratuberculosis : A review of accuracies of ELISA, interferon- γ assay and faecal culture techniques. *Veterinary Microbiology*, 129(3-4) : 217–235. ISSN 03781135. doi : 10.1016/j.vetmic.2007.12.011.
- Nielsen S et Toft N. **2011**. Effect of management practices on paratuberculosis prevalence in Danish dairy herds. *Journal of Dairy Science*, 94(4) :1849–1857. ISSN 00220302. doi : 10.3168/jds.2010-3817.
- Nielsen S, Toft N, Jørgensen E et Bibby B. **2007**. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. *Preventive Veterinary Medicine*, 81(4) :290–305. ISSN 01675877. doi : 10.1016/j.prevetmed.2007.05.014.
- Norton S, Heuer C et Jackson R. **2009**. A questionnaire-based cross-sectional study of clinical Johne's disease on dairy farms in New Zealand. *New Zealand Veterinary Journal*, 57(1) :34–43. ISSN 0048-0169, 1176-0710. doi : 10.1080/00480169.2009.36866.
- O'Brien L. M, McAloon C. G, Stewart L. D, Strain S. a. J et Grant I. R. Diagnostic potential of the peptide-mediated magnetic separation (PMS)-phage assay and PMS-culture to detect *Mycobacterium avium* subsp. paratuberculosis in bovine milk samples. *Transboundary and Emerging Diseases*, pages n/a–n/a. ISSN 1865-1682. doi : 10.1111/tbed.12794.
- OIE. Système Mondial d'Information Zoosanitaire, 2017. URL : http://www.oie.int/wahis_2/public/wahid.php/Diseaseinformation/Diseasedistributionmap?disease_type_hidden=&disease_id_hidden=&selected_disease_name_hidden=&disease_type=0&disease_id_terrestrial=24&species_t=2&disease_id_aquatic=-999&species_a=0&sta_method=semesterly&selected_start_year=2016&selected_report_period=1&selected_start_month=2&date_submit=0K. [Consulté le : 2017-11-06].
- Okuni J. B. **2013**. Occurrence of Paratuberculosis in African Countries : a Review. *Journal of Veterinary Advances*, 3(1) :1–8. ISSN 2251-7685.
- Olsen J. E, Jørgensen J. B et Nansen P. **1985**. On the reduction of *Mycobacterium paratuberculosis* in bovine slurry subjected to batch mesophilic or thermophilic anaerobic digestion. *Agricultural Wastes*, 13(4) :273–280. ISSN 0141-4607. doi : 10.1016/0141-4607(85)90052-6.
-

- OMS. Lignes directrices de l'OMS pour l'utilisation chez les animaux de rentes destinés à l'alimentation humaine des antimicrobiens importants pour la médecine humaine, 2017. URL : <http://apps.who.int/iris/bitstream/10665/259244/1/WHO-NMH-FOS-FZD-17.4-fre.pdf?ua=1>.
- Parrish N, Vadlamudi A et Goldberg N. **2017**. Anaerobic adaptation of *Mycobacterium avium* subspecies paratuberculosis in vitro : similarities to *M. tuberculosis* and differential susceptibility to antibiotics. *Gut Pathogens*, 9. ISSN 1757-4749. doi : 10.1186/s13099-017-0183-z.
- Pavlík I, Bejčková L, Pavlas M, Rozsypalová Z et Kosková S. **1995**. Characterization by restriction endonuclease analysis and DNA hybridization using IS900 of bovine, ovine, caprine and human dependent strains of *Mycobacterium paratuberculosis* isolated in various localities. *Veterinary Microbiology*, 45(4) :311–318. ISSN 0378-1135. doi : 10.1016/0378-1135(94)00130-O.
- Pepe M. S et Janes H. **2007**. Insights into latent class analysis of diagnostic test performance. *Biostatistics*, 8(2) :474–484. ISSN 1465-4644, 1468-4357. doi : 10.1093/biostatistics/kxl038.
- Pesqueira M. N, Yus E, Factor C, Mato I, Sanjuán M. L, Eiras C, Arnaiz I et Diéguez F. J. **2017**. Short communication : Correlation between within-herd antibody-prevalence and bulk tank milk antibody levels to *Mycobacterium avium* ssp. paratuberculosis using 2 commercial immunoassays. *Journal of Dairy Science*, 100(9) :7544–7548. ISSN 0022-0302. doi : 10.3168/jds.2017-12706.
- Petit H. **2006**. Résultats d'une enquête GDS sur la paratuberculose. *Le point vétérinaire*, (263).
- Pickup R. W, Rhodes G, Arnott S, Sidi-Boumedine K, Bull T. J, Weightman A, Hurley M et Hermon-Taylor J. **2005**. *Mycobacterium avium* subsp. paratuberculosis in the Catchment Area and Water of the River Taff in South Wales, United Kingdom, and Its Potential Relationship to Clustering of Crohn's Disease Cases in the City of Cardiff. *Applied and Environmental Microbiology*, 71(4) : 2130–2139. ISSN 0099-2240. doi : 10.1128/AEM.71.4.2130-2139.2005.
- Pickup R. W, Rhodes G, Bull T. J, Arnott S, Sidi-Boumedine K, Hurley M et Hermon-Taylor J. **2006**. *Mycobacterium avium* subsp. paratuberculosis in Lake Catchments, in River Water Abstracted for Domestic Use, and in Effluent from Domestic Sewage Treatment Works : Diverse Opportunities for Environmental Cycling and Human Exposure. *Applied and Environmental Microbiology*, 72(6) :4067–4077. ISSN 0099-2240, 1098-5336. doi : 10.1128/AEM.02490-05.

- Pillai S. R et Jayarao B. M. **2002**. Application of IS900 PCR for Detection of Mycobacterium avium subsp. paratuberculosis Directly from Raw Milk. *Journal of Dairy Science*, 85(5) :1052–1057. ISSN 0022-0302. doi : 10.3168/jds.S0022-0302(02)74165-9.
- Pillars R. B, Grooms D. L et Kaneene J. B. **2009**. Longitudinal study of the distribution of Mycobacterium avium subsp. paratuberculosis in the environment of dairy herds in the Michigan Johne's disease control demonstration herd project. *The Canadian Veterinary Journal*, 50(10) : 1039–1046. ISSN 0008-5286.
- Portaels F et Pattyn S. R. **1982**. Growth of mycobacteria in relation to the pH of the medium. *Annales De Microbiologie*, 133(2) :213–221. ISSN 0300-5410.
- Poupart P, Coene M, Heuverswyn H. V et Cocito C. **1993**. Preparation of a specific RNA probe for detection of Mycobacterium paratuberculosis and diagnosis of Johne's disease. *Journal of Clinical Microbiology*, 31(6) :1601–1605. ISSN 0095-1137, 1098-660X.
- Pérez V, García Marín J. F et Badiola J. J. **1996**. Description and classification of different types of lesion associated with natural paratuberculosis infection in sheep. *Journal of Comparative Pathology*, 114(2) :107–122. ISSN 0021-9975.
- Purdie A. C, Plain K. M, Begg D. J, de Silva K et Whittington R. J. **2011**. Candidate gene and genome-wide association studies of Mycobacterium avium subsp. paratuberculosis infection in cattle and sheep : A review. *Comparative Immunology, Microbiology and Infectious Diseases*, 34(3) : 197–208. ISSN 0147-9571. doi : 10.1016/j.cimid.2010.12.003.
- Qu Y, Tan M et Kutner M. H. **1996**. Random effects models in latent class analysis for evaluating accuracy of diagnostics tests. *Biometrics*, 52(3) :797–810.
- Raizman E. A, Wells S. J, Muñoz-Zanzi C. A et Tavoranpanich S. **2011**. Estimated within-herd prevalence (WHP) of Mycobacterium avium subsp. paratuberculosis in a sample of Minnesota dairy herds using bacterial culture of pooled fecal samples. *Canadian Journal of Veterinary Research*, 75(2) :112.
- Reddacliff L, Epplenston J, Windsor P, Whittington R et Jones S. **2006**. Efficacy of a killed vaccine for the control of paratuberculosis in Australian sheep flocks. *Veterinary Microbiology*, 115(1-3) : 77–90. ISSN 03781135. doi : 10.1016/j.vetmic.2005.12.021.

- Ritz C et Spiess A.-N. **2008**. qpcR : an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics (Oxford, England)*, 24(13) :1549–1551. ISSN 1367-4811. doi : 10.1093/bioinformatics/btn227.
- Robbe-Austerman S. **2007**. *Evaluation of cell mediated immune diagnostic tests to detect Mycobacterium avium subspecies paratuberculosis*. PhD thesis, Iowa State University.
- Robbe-Austerman S, Gardner I. A, Thomsen B. V, Morrill D. G, Martin B. M, Palmer M. V, Thoen C. O et Ewing C. **2006**. Sensitivity and specificity of the agar-gel-immunodiffusion test, ELISA and the skin test for detection of paratuberculosis in United States Midwest sheep populations. *Veterinary Research*, 37(4) :553–564. ISSN 0928-4249, 1297-9716. doi : 10.1051/vetres:2006018.
- Rossi G, Nigro G, Tattoli I, Vincenzetti S, Mariani P, Magi G. E, Renzoni G, Taccini E et Bernardini M. L. **2009**. Adhesion molecules and cytokine profile in ileal tissue of sheep infected with Mycobacterium avium subsp. paratuberculosis. *Microbes and Infection*, 11(6) :698–706. ISSN 1286-4579. doi : 10.1016/j.micinf.2009.04.006.
- Rothel J, Jones S, Corner L, Cox J et Wood P. **1990**. A sandwich enzyme immunoassay for bovine interferon- γ and its use for the detection of tuberculosis in cattle. *Australian Veterinary Journal*, 67(4) :134–137. ISSN 1751-0813. doi : 10.1111/j.1751-0813.1990.tb07730.x.
- Rovira A, Cano J. P et Muñoz-Zanzi C. **2008**. Feasibility of pooled-sample testing for the detection of porcine reproductive and respiratory syndrome virus antibodies on serum samples by ELISA. *Veterinary Microbiology*, 130(1) :60–68. ISSN 0378-1135. doi : 10.1016/j.vetmic.2007.12.016.
- Rowe M et Grant I. **2006**. Mycobacterium avium ssp. paratuberculosis and its potential survival tactics. *Letters in Applied Microbiology*, 42(4) :305–311. ISSN 1472-765X. doi : 10.1111/j.1472-765X.2006.01873.x.
- Rutjes A. W. S, Reitsma J. B, Coomarasamy A, Khan K. S et Bossuyt P. M. M. **2007**. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technology Assessment (Winchester, England)*, 11(50) :iii, ix–51. ISSN 1366-5278.
- S. L. B. McKenna, D. C. Sockett, G. P. Keefe, J. McClure, J. A. VanLeeuwen et H. W. Barkema. **2005**. Comparison of Two Enzyme-Linked Immunosorbent Assays for Diagnosis of Mycobacterium Avium Subsp. Paratuberculosis. *Journal of Veterinary Diagnostic Investigation*, 17(5) :463–466. ISSN 1040-6387. doi : 10.1177/104063870501700510.

Saah A. J et Hoover D. R. **1997**. "Sensitivity" and "specificity" reconsidered : the meaning of these terms in analytical and diagnostic settings. *Annals of Internal Medicine*, 126(1) :91–94. ISSN 0003-4819.

Saiki R. K, Scharf S, Faloona F, Mullis K. B, Horn G. T, Erlich H. A et Arnheim N. **1985**. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230(4732) :1350–1354. ISSN 0036-8075, 1095-9203. doi : 10.1126/science.2999980.

Schwarz D. G. G, Lima M. C, Barros M, Valente F. L, Scatamburlo T. M, Rosado N, Oliveira C. T. S. A. M, Oliveira L. L et Moreira M. A. S. **2017**. Short communication : Passive shedding of *Mycobacterium avium* ssp. *paratuberculosis* in commercial dairy goats in Brazil. *Journal of Dairy Science*, 100(10) :8426–8429. ISSN 0022-0302. doi : 10.3168/jds.2017-12918.

Sergeant E. S. G, Whittington R. J et More S. J. **2002**. Sensitivity and specificity of pooled faecal culture and serology as flock-screening tests for detection of ovine paratuberculosis in Australia. *Preventive Veterinary Medicine*, 52(3) :199–211. ISSN 0167-5877. doi : 10.1016/S0167-5877(01)00261-6.

Sergeant E et Baldock F. **2002**. The estimated prevalence of Johne's disease infected sheep flocks in Australia. *Australian Veterinary Journal*, 80(12) :762–768. ISSN 1751-0813. doi : 10.1111/j.1751-0813.2002.tb11348.x.

Serrano M, Elguezabal N, Sevilla I. A, Geijo M. V, Molina E, Arrazuria R, Urkitza A, Jones G. J, Vordermeier M, Garrido J. M et Juste R. A. **2017**. Tuberculosis Detection in Paratuberculosis Vaccinated Calves : New Alternatives against Interference. *PLOS ONE*, 12(1) :e0169735. ISSN 1932-6203. doi : 10.1371/journal.pone.0169735.

Shaughnessy L. J, Smith L. A, Evans J, Anderson D, Caldow G, Marion G, Low J. C et Hutchings M. R. **2013**. High prevalence of paratuberculosis in rabbits is associated with difficulties in controlling the disease in cattle. *The Veterinary Journal*, 198(1) :267–270. ISSN 1090-0233. doi : 10.1016/j.tvjl.2013.08.030.

Slana I, Pribylova R, Kralova A et Pavlik I. **2011**. Persistence of *Mycobacterium avium* subsp. *paratuberculosis* at a Farm-Scale Biogas Plant Supplied with Manure from Paratuberculosis-Affected

- Dairy Cattle. *Applied and Environmental Microbiology*, 77(9) :3115–3119. ISSN 0099-2240. doi : 10.1128/AEM.02407-10.
- Smeed J. A, Watkins C. A, Rhind S. M et Hopkins J. **2007**. Differential cytokine gene expression profiles in the three pathological forms of sheep paratuberculosis. *BMC Veterinary Research*, 3 : 18. ISSN 1746-6148. doi : 10.1186/1746-6148-3-18.
- Speybroeck N, Williams C. J, Lafia K. B, Devleeschauwer B et Berkvens D. **2012**. Estimating the prevalence of infections in vector populations using pools of samples. *Medical and Veterinary Entomology*, 26(4) :361–371. ISSN 1365-2915. doi : 10.1111/j.1365-2915.2012.01015.x.
- Stabel J. R. **2000**. Transitions in immune responses to *Mycobacterium paratuberculosis*. *Veterinary Microbiology*, 77(3) :465–473. ISSN 0378-1135. doi : 10.1016/S0378-1135(00)00331-X.
- Stabel J. R. **2006**. Host responses to *Mycobacterium avium* subsp. *paratuberculosis* : a complex arsenal. *Animal Health Research Reviews*, 7(1-2) :61–70. ISSN 1475-2654, 1466-2523. doi : 10.1017/S1466252307001168.
- Stabel J. R et Robbe-Austerman S. **2011**. Early Immune Markers Associated with *Mycobacterium avium* subsp. *paratuberculosis* Infection in a Neonatal Calf Model. *Clinical and Vaccine Immunology : CVI*, 18(3) :393–405. ISSN 1556-6811. doi : 10.1128/CVI.00359-10.
- Stevenson K. **2015**. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* and the influence of strain type on infection and pathogenesis : a review. *Veterinary Research*, 46(1). ISSN 0928-4249. doi : 10.1186/s13567-015-0203-2.
- Sting R, Hrubenja M, Mandl J, Seemann G, Salditt A et Waibel S. **2014**. Detection of *Mycobacterium avium* subsp. *paratuberculosis* in faeces using different procedures of pre-treatment for real-time PCR in comparison to culture. *The Veterinary Journal*, 199(1) :138–142. ISSN 10900233. doi : 10.1016/j.tvjl.2013.08.033.
- Strommenger B, Stevenson K et Gerlach G.-F. **2001**. Isolation and diagnostic potential of ISMav2, a novel insertion sequence-like element from *Mycobacterium avium* subspecies *paratuberculosis*. *FEMS Microbiology Letters*, 196(1) :31–37. ISSN 1574-6968. doi : 10.1111/j.1574-6968.2001.tb10536.x.
- Sung N et Collins M. T. **1998**. Thermal Tolerance of *Mycobacterium paratuberculosis*. *Applied and Environmental Microbiology*, 64(3) :999–1005. ISSN 0099-2240.

-
- Sweeney R. W, Whitlock R. H, Hamir A. N, Rosenberger A. E et Herr S. A. **1992a**. Isolation of *Mycobacterium paratuberculosis* after oral inoculation in uninfected cattle. *American Journal of Veterinary Research*, 53(8) :1312–1314. ISSN 0002-9645.
- Sweeney R. W, Whitlock R. H et Rosenberger A. E. **1992b**. *Mycobacterium paratuberculosis* isolated from fetuses of infected cows not manifesting signs of the disease. *American Journal of Veterinary Research*, 53(4) :477–480. ISSN 0002-9645.
- Sweeney R. W, Uzonna J, Whitlock R. H, Habecker P. L, Chilton P et Scott P. **2006**. Tissue predilection sites and effect of dose on *Mycobacterium avium* subs. *paratuberculosis* organism recovery in a short-term bovine experimental oral infection model. *Research in Veterinary Science*, 80(3) : 253–259. ISSN 0034-5288. doi : 10.1016/j.rvsc.2005.07.007.
- Swift B. M. C, Denton E. J, Mahendran S. A, Huxley J. N et Rees C. E. D. **2013**. Development of a rapid phage-based method for the detection of viable *Mycobacterium avium* subsp. *paratuberculosis* in blood within 48 h. *Journal of Microbiological Methods*, 94(3) :175–179. ISSN 0167-7012. doi : 10.1016/j.mimet.2013.06.015.
- Swift B. M. C, Huxley J. N, Plain K. M, Begg D. J, de Silva K, Purdie A. C, Whittington R. J et Rees C. E. D. **2016**. Evaluation of the limitations and methods to improve rapid phage-based detection of viable *Mycobacterium avium* subsp. *paratuberculosis* in the blood of experimentally infected cattle. *BMC Veterinary Research*, 12. ISSN 1746-6148. doi : 10.1186/s12917-016-0728-2.
- Tavorpanich S, Muñoz-Zanzi C. A, Wells S. J, Raizman E. A, Carpenter T. E, Johnson W. O et Gardner I. A. **2008**. Simulation model for evaluation of testing strategies for detection of *paratuberculosis* in Midwestern US dairy herds. *Preventive Veterinary Medicine*, 83(1) :65–82. ISSN 01675877. doi : 10.1016/j.prevetmed.2007.06.010.
- Taylor A. W. **1953**. The experimental infection of cattle with varieties of *Mycobacterium Johnei* isolated from sheep. *Journal of Comparative Pathology and Therapeutics*, 63(Supplement C) : 368–373. ISSN 0368-1742. doi : 10.1016/S0368-1742(53)80038-X.
- Team R. D. C. R : A Language and Environment for Statistical Computing, 2017. URL : <https://www.r-project.org/>. [Consulté le : 2018-02-01].
- Timms V. J, Mitchell H. M et Neilan B. A. **2015**. Optimisation of DNA extraction and validation of
-

- PCR assays to detect *Mycobacterium avium* subsp. *paratuberculosis*. *Journal of Microbiological Methods*, 112 :99–103. ISSN 0167-7012. doi : 10.1016/j.mimet.2015.03.016.
- Toma B, Dufour B, Bénét J, Sanaa M, Shaw A et Moutou F. *Epidémiologie appliquée à la lutte collective contre les maladies animales transmissibles majeures*. AEEMA, 3ème édition, 2010. ISBN 978-9044-481-7.
- Toribio J.-A et Sergeant E. **2007**. A comparison of methods to estimate the prevalence of ovine Johne's infection from pooled faecal samples. *Australian Veterinary Journal*, 85(8) :317–324. ISSN 0005-0423, 1751-0813. doi : 10.1111/j.1751-0813.2007.00188.x.
- Tse C et Capeau J. **2003**. Quantification des acides nucléiques par PCR quantitative en temps réel. *Annales de Biologie Clinique*, 61(3) :279–293. ISSN 0003-3898.
- Tu X. M, Litvak E et Pagano M. **1995**. On the Informativeness and Accuracy of Pooled Testing in Estimating Prevalence of a Rare Disease : Application to HIV Screening. *Biometrika*, 82(2) : 287–297. ISSN 0006-3444. doi : 10.2307/2337408.
- Turenne C. Y, Collins D. M, Alexander D. C et Behr M. A. **2008**. *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* subsp. *avium* Are Independently Evolved Pathogenic Clones of a Much Broader Group of *M. avium* Organisms. *Journal of Bacteriology*, 190(7) :2479–2487. ISSN 0021-9193. doi : 10.1128/JB.01691-07.
- Twort F. W et Ingram G. L. Y. **1912**. A Method for Isolating and Cultivating the *Mycobacterium enteritidis chronicae pseudotuberculosis bovis*, Johne, and some Experiments on the Preparation of a Diagnostic Vaccine for Pseudo-tuberculous Enteritis of Bovines. *Proceedings of the Royal Society of London B : Biological Sciences*, 84(575) :517–542. ISSN 0962-8452, 1471-2954. doi : 10.1098/rspb.1912.0011.
- Vacek P. M. **1985**. The Effect of Conditional Dependence on the Evaluation of Diagnostic Tests. *Biometrics*, 41(4) :959–968. ISSN 0006-341X. doi : 10.2307/2530967.
- van Schaik G, Haro F, Mella A et Kruze J. **2007**. Bayesian analysis to validate a commercial ELISA to detect paratuberculosis in dairy herds of southern Chile. *Preventive Veterinary Medicine*, 79 (1) :59–69. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2006.11.005.

- VBJDCP. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. Technical Report APHIS 91-45-016, United States Department of Agriculture, 2010. URL : https://www.aphis.usda.gov/animal_health/animal_diseases/johnes/downloads/johnes-ups.pdf.
- Verdugo C, Jones G, Johnson W, Wilson P, Stringer L et Heuer C. **2014a**. Estimation of flock/herd-level true *Mycobacterium avium* subspecies paratuberculosis prevalence on sheep, beef cattle and deer farms in New Zealand using a novel Bayesian model. *Preventive Veterinary Medicine*, 117(3-4) :447-455. ISSN 01675877. doi : 10.1016/j.prevetmed.2014.10.004.
- Verdugo C, Pleydell E, Price-Carter M, Prattley D, Collins D, de Lisle G, Vogue H, Wilson P et Heuer C. **2014b**. Molecular epidemiology of *Mycobacterium avium* subsp. paratuberculosis isolated from sheep, cattle and deer on New Zealand pastoral farms. *Preventive Veterinary Medicine*, 117(3) :436-446. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2014.09.009.
- Verin R, Perroni M, Rossi G, De Grossi L, Botta R, De Sanctis B, Rocca S, Cubeddu T, Crosby-Durrani H et Taccini E. **2016**. Paratuberculosis in sheep : Histochemical, immunohistochemical and in situ hybridization evidence of in utero and milk transmission. *Research in Veterinary Science*, 106(Supplement C) :173-179. ISSN 0034-5288. doi : 10.1016/j.rvsc.2016.04.006.
- Verna A. E, Garcia-Pariente C, Muñoz M, Moreno O, García-Marin J. F, Romano M. I, Paolicchi F et Pérez V. **2007**. Variation in the Immuno-pathological Responses of Lambs after Experimental Infection with Different Strains of *Mycobacterium avium* subsp. paratuberculosis. *Zoonoses and Public Health*, 54(6-7) :243-252. ISSN 1863-2378. doi : 10.1111/j.1863-2378.2007.01058.x.
- Waddell L, Rajić A, Stärk K et McEwen S. A. **2016**. *Mycobacterium avium* ssp. paratuberculosis detection in animals, food, water and other sources or vehicles of human exposure : A scoping review of the existing evidence. *Preventive Veterinary Medicine*, 132(Supplement C) :32-48. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2016.08.003.
- Weber M. F, Verhoeff J, van Schaik G et van Maanen C. **2009**. Evaluation of Ziehl-Neelsen stained faecal smear and ELISA as tools for surveillance of clinical paratuberculosis in cattle in the Netherlands. *Preventive Veterinary Medicine*, 92(3) :256-266. ISSN 0167-5877. doi : 10.1016/j.prevetmed.2009.08.017.

- Whittington R. J, Marsh I, McAllister S, Turner M. J, Marshall D. J et Fraser C. A. **1999**. Evaluation of Modified BACTEC 12b Radiometric Medium and Solid Media for Culture of Mycobacterium avium subsp. paratuberculosis from Sheep. *Journal of Clinical Microbiology*, 37(4) :1077–1083. ISSN 0095-1137.
- Whittington R. J, Fell S, Walker D, McAllister S, Marsh I, Sergeant E, Taragel C. A, Marshall D. J et Links I. J. **2000a**. Use of Pooled Fecal Culture for Sensitive and Economic Detection of Mycobacterium avium subsp. paratuberculosis Infection in Flocks of Sheep. *Journal of clinical microbiology*, 38(7) :2550–2556.
- Whittington R. J, Marsh I. B et Whitlock R. H. **2001**. Typing of IS 1311 polymorphisms confirms that bison (*Bison bison*) with paratuberculosis in Montana are infected with a strain of Mycobacterium avium subsp. paratuberculosis distinct from that occurring in cattle and other domesticated livestock. *Molecular and Cellular Probes*, 15(3) :139–145. ISSN 0890-8508. doi : 10.1006/mcpr.2001.0346.
- Whittington R. J, Begg D. J, de Silva K, Purdie A. C, Dhand N. K et Plain K. M. **2017**. Case definition terminology for paratuberculosis (Johne's disease). *BMC Veterinary Research*, 13. ISSN 1746-6148. doi : 10.1186/s12917-017-1254-6.
- Whittington R. J et Windsor P. A. **2009**. In utero infection of cattle with Mycobacterium avium subsp. paratuberculosis : A critical review and meta-analysis. *The Veterinary Journal*, 179(1) : 60–69. ISSN 10900233. doi : 10.1016/j.tvjl.2007.08.023.
- Whittington R. J, Marshall D. J, Nicholls P. J, Marsh I. B et Reddacliff L. A. **2004**. Survival and Dormancy of Mycobacterium avium subsp. paratuberculosis in the Environment. *Applied and Environmental Microbiology*, 70(5) :2989–3004. ISSN 0099-2240. doi : 10.1128/AEM.70.5.2989-3004.2004.
- Whittington R. J, Marsh I. B et Reddacliff L. A. **2005**. Survival of Mycobacterium avium subsp. paratuberculosis in Dam Water and Sediment. *Applied and Environmental Microbiology*, 71 (9) :5304–5308. ISSN 0099-2240, 1098-5336. doi : 10.1128/AEM.71.9.5304-5308.2005.
- Whittington R. J, Whittington A.-M, Waldron A, Begg D. J, de Silva K, Purdie A. C et Plain K. M. **2013**. Development and Validation of a Liquid Medium (M7h9c) for Routine Culture of Myco-

bacterium avium subsp. paratuberculosis To Replace Modified Bactec 12b Medium. *Journal of Clinical Microbiology*, 51(12) :3993–4000. ISSN 0095-1137. doi : 10.1128/JCM.01373-13.

Whittington R, Reddacliff L, Marsh I, McALLISTER S et Saunders V. **2000b**. Temporal patterns and quantification of excretion of Mycobacterium avium subsp paratuberculosis in sheep with Johne's disease. *Australian Veterinary Journal*, 78(1) :34–37. ISSN 1751-0813. doi : 10.1111/j.1751-0813.2000.tb10355.x.

Wilson D. J, Rood K, Biswas P et Byrem T. M. **2010**. Herd-level prevalence of Johne's disease in Utah and adjacent areas of the intermountain west as detected by a bulk-tank milk surveillance project. *Journal of Dairy Science*, 93(12) :5792–5797. ISSN 0022-0302. doi : 10.3168/jds.2010-3481.

Windsor P, Eppleston J, Dhand N et Whittington R. **2014**. Effectiveness of Gudair™ vaccine for the control of ovine Johne's disease in flocks vaccinating for at least 5 years. *Australian Veterinary Journal*, 92(7) :263–268. ISSN 00050423. doi : 10.1111/avj.12194.

Wolf R, Barkema H. W, De Buck J et Orsel K. **2015**. Sampling location, herd size, and season influence Mycobacterium avium ssp. paratuberculosis environmental culture results. *Journal of Dairy Science*, 98(1) :275–287. ISSN 0022-0302. doi : 10.3168/jds.2014-8676.

Zanetti S, Molicotti P, Cannas S, Ortu S, Ahmed N et Sechi L. A. **2006**. "In vitro" activities of anti-mycobacterial agents against Mycobacterium avium subsp. paratuberculosis linked to Crohn's Disease and Paratuberculosis. *Annals of Clinical Microbiology and Antimicrobials*, 5 :27. ISSN 1476-0711. doi : 10.1186/1476-0711-5-27.

Zare Y, Shook G. E, Collins M. T et Kirkpatrick B. W. **2014**. Genome-Wide Association Analysis and Genomic Prediction of Mycobacterium avium Subspecies paratuberculosis Infection in US Jersey Cattle. *PLoS ONE*, 9(2). ISSN 1932-6203. doi : 10.1371/journal.pone.0088380.

Annexes

I Liste des communications

Communications écrites publiées dans des revues scientifiques à comité de lecture

- MATHEVON Y, FOUCRAS G, FALGUIERES R et CORBIERE F. Estimation of the sensitivity and specificity of two serum ELISAs and one fecal qPCR for diagnosis of paratuberculosis in sub-clinically infected young-adult French sheep using latent class Bayesian modeling. *BMC Veterinary Research*, 13 :230, August 2017. ISSN 1746-6148. doi : 10.1186/s12917-017-1145-x. URL <https://doi.org/10.1186/s12917-017-1145-x>

Manuscrits en préparation pour des revues scientifiques à comité de lecture

- MATHEVON Y, FOUCRAS G, CORBIERE F. Flock sensitivity and specificity of pooled fecal qPCR and pooled serum ELISA for screening ovine paratuberculosis in France (manuscrit en préparation, soumission prévue à la revue *Plos One*)
- MATHEVON Y, CONCORDET D, CORBIERE F. Freak : a frequentist approach for diagnosis accuracy studies (manuscrit en préparation, soumission prévue à la revue *Preventive Veterinary Medicine*)

Communications orales en congrès scientifiques ou techniques

- MATHEVON Y., LAFORT, M.-P., TASCA, C., GALAN, E., FOUCRAS, G., CORBIERE, F. (2015). La vaccination contre la paratuberculose ovine : développements méthodologiques et éva-

luation de son efficacité, Rencontres autour des recherches sur les ruminants, 2-03/12/2015, Paris

- MATHEVON Y, CORBIERE F, LAFORT M-P, TASCA C, GALAN E, FALGUIERES R, FOUCRAS G. Vaccination contre la paratuberculose ovine : évaluation de son intérêt en élevage ovin allaitant. Journées Nationales des Groupements Techniques Vétérinaires, Nantes, 18 mai 2016
- MATHEVON Y, LAFORT M-P, TASCA C, FALGUIERES R, CORBIERE F, FOUCRAS G. Vaccination against Paratuberculosis in commercial French sheep flocks : is it useful? 9th International Sheep Veterinary Congress, Harrogate, 26 Mai 2017
- CORBIERE F, MATHEVON Y, FOUCRAS G. Paratuberculose ovine et caprine : intérêts de la vaccination et stratégies de suivi en élevages vaccinés. Journées Nationales des Groupements Techniques Vétérinaires, Nantes, mai 2018 (accepté)

Communications affichées en congrès scientifiques

- MATHEVON Y. (2016). Detection and quantification of *Mycobacterium paratuberculosis* by qPCR in ovine faeces : methodology and improvement of fluorescence curves interpretation, International colloquium on paratuberculosis, 20-24/06/2016, Nantes
- MATHEVON Y, LAFORT, M.-P., TASCA, C., GALAN, E., FOUCRAS, G., CORBIERE, F. (2016). Vaccination against Paratuberculosis in sheep : methodological developments and evaluation of efficacy, International colloquium on paratuberculosis, 20-24/06/2016, Nantes

Communications écrites dans la presse technique vétérinaire

- FOUCRAS G, MATHEVON Y, CORBIERE F. Maitrise de la paratuberculose chez les ruminants : que peut-on attendre de la vaccination? Bulletin des GTV, 2016, N°84, 39-46

Autres communications orales

- MATHEVON Y, LAFORT M-P, TASCA C, GALAN E, FALGUIERES R, CORBIERE F, FOUCRAS G. La vaccination contre la paratuberculose ovine : développements méthodologiques et évaluation de son efficacité. Comité de programmation et d'orientation de l'UMT Santé des Troupeaux de Petits Ruminants, Toulouse 28 janvier 2016

- CORBIERE F, MATHEVON Y, HUGNET J, LAFORT M-P, TASCA C, GALAN E, FALGUIERES R, FOUCRAS G. Concordance entre trousses sérologiques dans le diagnostic de la paratuberculose ovine : éléments de description sur les données du Lot. Comité de programmation et d'orientation de l'UMT Santé des Troupeaux de Petits Ruminants, Toulouse 28 janvier 2016
- MATHEVON Y, CORBIERE F, FOUCRAS G. Analyses de mélanges de sérums et de fèces pour le diagnostic de la paratuberculose à l'échelle des troupeaux : premiers résultats chez les ovins. Comité de programmation et d'orientation de l'UMT Santé des Troupeaux de Petits Ruminants, Toulouse 23 novembre 2017
- CORBIERE F, MATHEVON Y, FOUCRAS G. Analyses de mélanges de fèces pour le diagnostic de la paratuberculose à l'échelle des troupeaux : premiers résultats chez les ovins et caprins. Journée UMT SPR et OMACAP, Paris, 27 novembre 2017
- FOUCRAS G, MATHEVON Y, CORBIERE F. Paratuberculose et vaccination. Journée technique sur la paratuberculose caprine, Bourges, 01 octobre 2015
- FOUCRAS G, MATHEVON Y, CORBIERE F. Vaccination contre la paratuberculose : Mécanismes, conditions de mise en œuvre et conséquences. GDS France – Commission ovine et caprine – 15 novembre 2016
- MATHEVON Y, LAFORT M-P, TASCA C, GALAN E, FALGUIERES R, CORBIERE F, FOUCRAS G. La vaccination contre la paratuberculose ovine : développements méthodologiques et évaluation de son efficacité. Journées d'animation scientifique du département Santé Animale de l'INRA, Poitiers, 10 mai 2016
- CORBIERE F, MATHEVON Y, FOUCRAS G. Paratuberculose et vaccination. Journée technique sur la paratuberculose caprine et ovine, Montauban – 18 octobre 2016
- CORBIERE F, MATHEVON Y, FOUCRAS G. Analyses de mélanges de sérums et de fèces pour le diagnostic de la paratuberculose à l'échelle des troupeaux : premiers résultats chez les ovins. Journée technique sur la paratuberculose caprine et ovine, Montauban – 18 octobre 2016

Réunions du comité de pilotage du programme PICSAR

- CORBIERE F, MATHEVON Y, FOUCRAS G. Programme PICSAR : état d'avancement des travaux chez les ovins. Comité de pilotage du métaprogramme PICSAR, Paris, 28 octobre 2015

- CORBIERE F, MATHEVON Y, FOUCRAS G. Paratuberculose ovine : concordances entre tests individuels et premiers résultats des analyses de mélanges. Comité de pilotage du métaprogramme PICSAR, Paris, 25 mai 2016
- MATHEVON Y, FOUCRAS G, CORBIERE F. Paratuberculose ovine : estimation des performances diagnostiques des tests à l'échelle individuelle et apports des analyses de mélanges. Comité de pilotage du métaprogramme PICSAR, Paris, 26 janvier 2017.

Présentations aux éleveurs

- Programmes Paratuberculose PICSAR / PAROVLOT : vaccination contre la paratuberculose ovine : premiers résultats dans le Lot. Commission Sanitaire d'Ovilot, Gramat, 26 mars 2015
- Programmes Paratuberculose PICSAR / PAROVLOT : résultats dans le lot, Assemblée générale d'Ovilot, Fontanes du Causse, 20 avril 2016

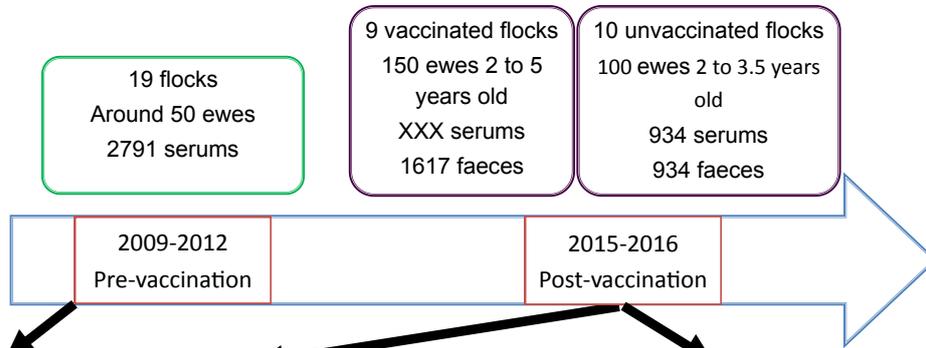
II High throughput qPCR analysis : Application to the detection of Mycobacterium Paratuberculosis in sheep

Vaccination against Paratuberculosis in sheep: methodological developments and evaluation of efficacy

MATHEVON Y.¹, LAFORT M-P.², TASCA C.¹, GALAN E.³, FALGUIERES R.³, CORBIERE F.^{1,2}, FOUCRAS G.^{1,2}

(1) INRA-ENVT, UMR 1225 IHAP, 31076 Toulouse, France
 (2) ENVT, 23 chemin des Capelles, 31076 Toulouse, France
 (3) CAPEL Ovilot, 237 avenue Pierre Semard, 46000 Cahors, France

Since 2009, vaccination against paratuberculosis is available in France for sheep. We evaluate the GUDAIR® vaccination in 19 meat sheep holdings, 9 of them vaccinate replacement ewes later than the manufacturer recommendation, at 6 months of age.



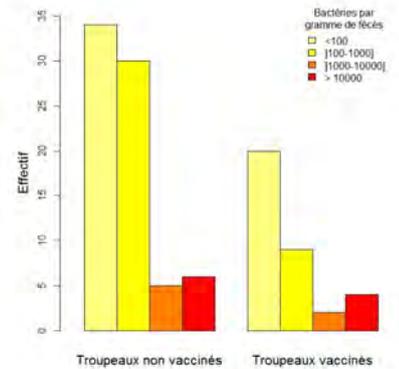
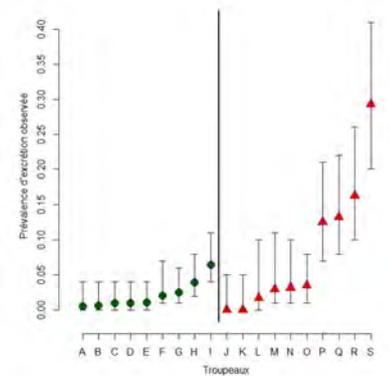
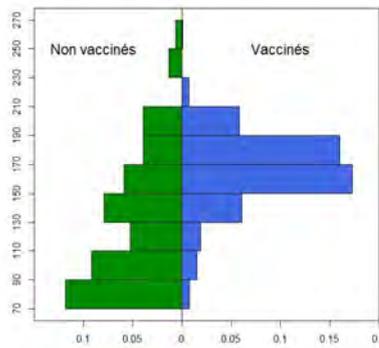
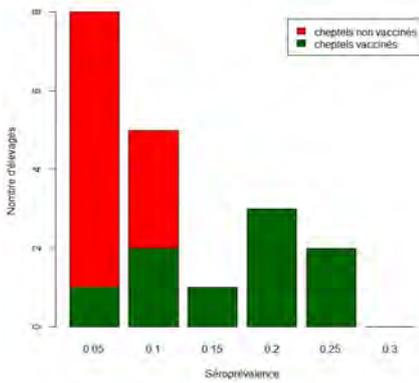
Material and methods

- Idexx Paratuberculosis Screening Ab Test®

Serum : ID Screen Paratuberculosis Indirect, ID.Vet,

Faeces : Adiafilter concentration, Adiavet qPCR (Biomérieux)

age (year)	#	% pos	mean SP*
2	276	0.97	165.26
3	201	0.95	153.48
4	188	0.96	156.61



The apparent **seroprevalence** towards paratuberculosis was significantly **higher** in the nine flocks that implemented vaccination programs than in the ten flocks that did not (median= 12.2% vs 2.9%, Wilcoxon Signed rank p=0.01).

Vaccination the **seroconversion** persists even 4 years after vaccination. **S/P ratios** in vaccinated ewes and naturally infected ewes in non-vaccinated flocks were similar

The **shedding prevalence** in the two populations aren't statistically different. Results of qPCR in vaccinated flocks showed a **low faecal shedding from 0% to 8%** compared to non-vaccinated flocks (0%-30%). **Shedding levels** were similar

The gudair vaccination at 6 month of age; compared to an earlier vaccination; provide a long lasting seroconversion. Due to this strong seroconversion, serology is inadequate in vaccinated area. In highly infected flocks, a 5 year vaccination reduce shedding prevalence to a low and homogenous level but isn't enough to free flocks from paratuberculosis

- Late vaccination protocol reduce shedding
 - 5 years of vaccination isn't enough to break infection cycle



III Vaccination against Paratuberculosis in sheep : methodological developments and evaluation of efficacy

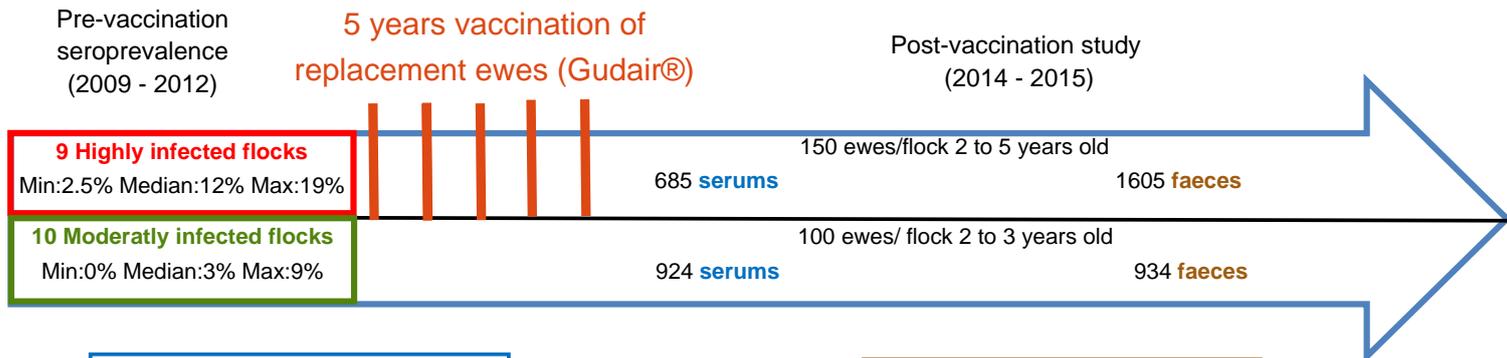
Vaccination against Paratuberculosis in sheep: methodological developments and evaluation of efficacy

MATHEVON Y.¹, LAFORT M-P.², TASCA C.¹, GALAN E.³, FALGUIERES R.³, CORBIERE F.^{1,2}, FOUCRAS G.^{1,2}

Corresponding author: y.mathevon@envt.fr

(1) INRA-ENVT, UMR 1225 IHAP, 31076 Toulouse, France
 (2) ENVt, 23 chemin des Capelles, 31076 Toulouse, France
 (3) CAPEL Ovilot, 237 avenue Pierre Semard, 46000 Cahors, France

Since 2009, vaccination against paratuberculosis is available in France for sheep. We evaluate the GUDAIR® vaccination effects on individual serological response and faecal shedding of Map in 19 commercial meat sheep holdings. Nine highly infected flocks vaccinated replacement ewes between 4 and 6 months old, each year. After a 5 years vaccination program these flocks were compared to 10 other moderately infected ones left without any active control measures during the same period.

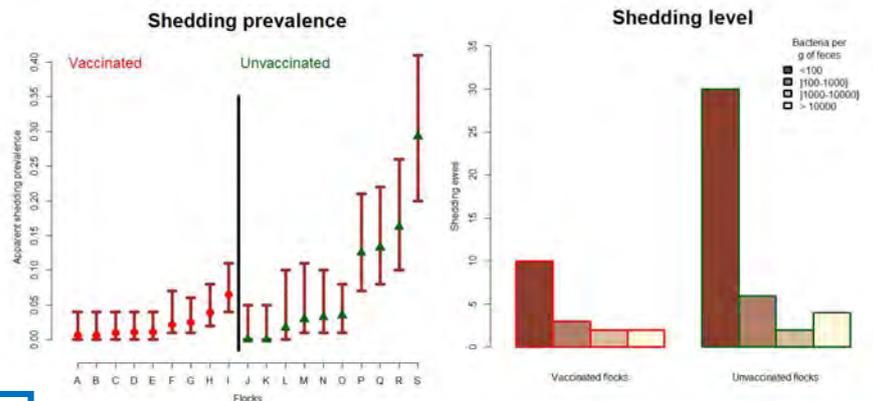


Individual serum
 ID Screen Paratuberculosis Indirect, ID.Vet®
 overnight incubation protocol

Individual faeces
 Adiafilter concentration, Adiavet qPCR®
 (Biomerieux)

	Age (year)	#	% pos	mean S/P*
Vaccinated	2	276	0.97	165.26
	3	201	0.95	153.48
	4	188	0.96	156.61
Unvaccinated	2	881	0.07	135.16
	3	198	0.06	124.59

* mean S/P ratios calculated on positive ewes (>70%)



- Vaccination induces a long lasting **seroconversion** (at least 4years) in almost all ewes.
- **S/P ratios** don't allow to discriminate between vaccinated ewes and naturally infected ewes in non-vaccinated flocks.
- Serology is inappropriate to follow up new infections in vaccinated flocks

- The apparent **shedding prevalence** was consistently low in vaccinated flocks (0 to 8%) with no difference between birth cohorts, while highly variable in non-vaccinated flocks (0%-30%).
- Individual **shedding levels** were similar between flocks with few high shedders (>10⁴ eq bacteria/gramm).

- **Even long lasting vaccination programs aren't successful in stopping the infection cycle of Map.**
- **However vaccination of 4 to 6 months old replacement ewes seems efficient to reduce faecal shedding of Map at flock level.**



La paratuberculose est une maladie enzootique contagieuse des ruminants due à *Mycobacterium avium* ssp. *paratuberculosis* (*Map*). La longue période d'incubation et les faibles performances diagnostiques des tests limitent l'efficacité des plans de maîtrise. Les grands effectifs des troupeaux ovins limitent l'approche par dépistage individuel exhaustif, et les plans de maîtrise s'orientent vers la vaccination, dont les effets n'ont pas été évalués dans les troupeaux français.

À partir d'échantillons de sérum et de fèces de près de 4000 brebis issues de 30 élevages du Lot, les performances diagnostiques de deux trousseaux sérologiques ELISA et d'une qPCR sur fèces ont été estimées par des modèles à classes latentes bayésiens et fréquentistes. Nos résultats confirment la faible sensibilité et le défaut de spécificité des sérologies ELISA pour la détection des ovins infectés; la qPCR présentant de meilleures performances diagnostiques. Par ailleurs nous avons évalué les performances diagnostiques relatives des ELISA et de la qPCR appliquées à des mélanges d'échantillons. Dans les deux cas les animaux forts répondeurs étaient détectés de façon systématique, les animaux faiblement positifs étant détectés de façon moins constante. Sur la base de simulations, nous avons évalué les performances des stratégies de dépistage et de suivi basées sur les analyses de mélanges d'échantillons à l'échelle des troupeaux. Alors que la sérologie ELISA est apparue insuffisamment spécifique, l'analyse de mélange de fèces par qPCR semble être une approche prometteuse, permettant de détecter des faibles prévalences d'infection. Enfin les travaux menés dans les troupeaux vaccinés ont précisé dans quelles mesures leur situation épidémiologique pouvait être approchée par l'emploi d'analyses en mélanges.

Mots clés : Paratuberculose, Ovin, Dépistage, ELISA, qPCR, Analyses de mélange

Paratuberculosis is a contagious enzootic disease in ruminants caused by *Mycobacterium avium* ssp. *paratuberculosis* (*Map*). The long incubation period and the low informative values of ante-mortem diagnostic tests limit the effectiveness of control plans. In large sheep flocks, exhaustive individual testing is unfeasible and control plans mainly focus on vaccination, the effects of which have not yet been evaluated in French flocks.

Using blood and feces samples from nearly 4000 ewes in 30 sheep flocks from the French department of Lot, the diagnostic performances of two serum ELISA and one fecal qPCR kits were estimated using bayesian and frequentist latent class modeling. Our results confirm the low sensitivity and non-perfect specificity of serum ELISA for the detection of subclinically infected animals, while the diagnostic performances of fecal qPCR were better. We also evaluated the relative diagnostic performances of pooled-sample analysis for both tests. Highly qPCR/ELISA positive samples were invariably detected while low positive ones were associated with lower detection rates.

The flock-level epidemiological performances of screening strategies based on pooled-sample analysis were evaluated by simulation studies. Pooled serum ELISA appeared lowly specific. Conversely, pooled fecal qPCR appeared promising, allowing the detection of low infection prevalence. Finally, the work carried out in the vaccinated flocks made it possible to better know their epidemiological status and to specify to what extent it could be approached using pooled-sample analysis.

Keywords : Paratuberculosis, Ovine, Screening, ELISA, qPCR, Pooled sample analysis