UTILISATION DES MÉTHODES DE CAPTURE-RECAPTURE POUR L'ÉVALUATION DES SYSTÈMES DE SURVEILLANCE EN SANTÉ ANIMALE : INTÉRÊTS ET LIMITES*

Timothée Vergne[1,2], Vladimir Grosbois[1], Benoît Durand[2], François Roger[1] et **Barbara Dufour[3]**

RÉSUMÉ

Bien qu'exploitées très largement dans le domaine de la santé publique, les méthodes de capture-recapture (CR) ne sont utilisées que depuis peu de temps pour évaluer les systèmes de surveillance des maladies animales. En nous fondant sur la littérature existante, nous proposons une revue des méthodes disponibles, ainsi qu'une discussion de leurs avantages et de leurs limites dans ce contexte. Ces méthodes semblent relativement faciles à conduire dans la plupart des situations et permettent à faible coût d'estimer l'importance réelle d'une maladie sur un territoire quand celle-ci est surveillée de manière imparfaite. Elles permettent aussi d'étudier les facteurs pouvant agir sur les processus de détection et de déclaration. Il semble cependant que les pratiques de surveillance et de contrôle des maladies animales limitent les applications à l'échelle de l'animal infecté, et nécessitent d'élargir l'unité épidémiologique à une échelle supérieure (troupeau, commune, *etc.*). Cet élargissement introduit de nouvelles contraintes (notamment l'hétérogénéité d'abondance) qu'il est nécessaire de prendre en compte pour ne pas biaiser les estimations finales.

Mots-clés : capture-recapture, surveillance, épidémiologie, santé animale, sensibilité.

SUMMARY

Although capture-recapture techniques have long been exploited in public health investigations, their use in surveillance of animal infectious diseases is only recent. Based on a review of the literature, we present various available, techniques and discuss their respective advantages and limitations in investigations on animal health. These techniques are relatively easy to implement and, with little additional input, provide an evaluation of the real significance of a disease in a given territory where proper surveillance processes are lacking. In addition, they can be of help in evaluating factors that affect detection and reporting processes. However, current procedures in animal disease surveillance and control tend to restrict the application of such techniques to individual animals. The evaluation of, surveillance systems should be extended to larger epidemiological units (holdings or communities for example). Such extensions create new constraints (such as higher heterogeneity) that have to be taken into account in order to obtain unbiased estimates.

Keywords: Capture-recapture, Surveillance, Epidemiology, Animal health, Sensitivity.

---

* Texte de la communication orale présentée au cours des Journées scientifiques AEEMA, 1er juin 2012

[1] CIRAD, UR22-AGIRs, TA C22/E, Campus international de Baillarguet, 34398 Montpellier cedex 5, France
[2] Anses, Laboratoire de Santé animale, 23 avenue du Général de Gaulle, 94706 Maisons-Alfort, France
[3] ENVA Maisons-Alfort, Unité EPIMAI, 23 avenue du Général de Gaulle, 94706 Maisons-Alfort, France

I - INTRODUCTION

A ce jour, les méthodes de capture-recapture sont largement utilisées en santé publique pour corriger les prévalences apparentes issues de systèmes de surveillance imparfaits. Les premières applications ont été initiées par les travaux de Wittes *et al.* [Wittes, 1974 ; Wittes et Sidel, 1968], puis ces méthodes sont devenues très populaires dès la parution des trois articles de référence sur l'utilisation de ces méthodes en santé publique [Hook et Regal, 1995 ; IWGDMF, 1995a ; b].

En surveillance des maladies animales, les méthodes de capture-recapture sont encore peu développées [Drewe *et al.*, 2012]. L'idée d'utiliser les approches de capture-recapture semble remonter à 1999 lorsque Cameron essaya d'estimer le nombre de foyers de fièvre aphteuse dans trois provinces du nord de la Thaïlande [Cameron, 1999]. Ce n'est qu'en 2005 qu'est apparue la première publication dans la revue *Preventive Veterinary Medicine* [Del Rio Vilas *et al.*, 2005]. Il s'agissait d'estimer au Royaume-Uni le nombre réel d'élevages ovins infectés par l'agent de la tremblante classique pour en déduire la sensibilité du système de surveillance à l'échelle des élevages, en utilisant une approche multiliste à trois listes. Del Rio Vilas *et al.* [2008] ont ensuite proposé une alternative à cette approche multiliste qui se révélait peu adaptée, en appliquant une approche uniliste au même contexte. L'exemple de la tremblante ovine au Royaume-Uni a par la suite été largement utilisé pour illustrer divers développements méthodologiques [Böhning et Del Rio Vilas, 2008 ; Böhning et Del Rio Vilas, 2009 ; Böhning *et al.*, 2011 ; Navaratna *et al.*, 2008]. Suivant l'approche uniliste, Vergne *et al.* [2012a] ont utilisé les données de la surveillance de la tremblante classique en France pour proposer une modélisation paramétrique des données de surveillance dans l'objectif d'estimer le nombre d'élevages ovins français infectés par l'agent de la tremblante classique. En Tanzanie, une approche multiliste à deux listes a été utilisée pour estimer l'incidence de la theileriose et sa mortalité associée [Kivaria et Noordhuizen, 2009]. Enfin, Vergne *et al.* [2012b] ont utilisé une approche multiliste à deux listes pour corriger l'estimation du nombre de villages infectés par la fièvre aphteuse issue de la surveillance événementielle au Cambodge. Les études de capture-recapture appliquées en santé animale et publiées jusqu'en juin 2012 sont présentées dans le tableau 1.

L'objectif de cet article est, après avoir présenté les différentes méthodes de capture-recapture existantes potentiellement utilisables pour évaluer les systèmes de surveillance des maladies animales, de discuter les intérêts et les limites de ces méthodes dans le contexte de la santé animale.

Tableau 1

Etudes de capture-recapture appliquées aux maladies infectieuses animales publiées au 1 juin 2012

Référence	Maladie	Pays	Méthode
Del Rio Vilas *et al.* [2005]	Tremblante classique	Royaume-Uni	CR 3 sources
Del Rio Vilas et Böhning [2008]	Tremblante classique	Royaume-Uni	CR uniliste
Böhning et Del Rio Vilas [2008]	Tremblante classique	Royaume-Uni	CR uniliste
Navaratna *et al.* [2008]	Tremblante classique	Royaume-Uni	CR uniliste
Böhning et Del Rio Vilas [2009]	Tremblante classique	Royaume-Uni	CR uniliste
Kivaria et Noordhuizen [2010]	Theileriose	Tanzanie	CR 2 sources
Böhning *et al.* [2011]	Tremblante classique	Royaume-Uni	CR uniliste
Vergne *et al.* [2012a]	Tremblante classique	France	CR uniliste
Vergne *et al.* [2012b]	Fièvre aphteuse	Cambodge	CR 2 sources

II - LES MÉTHODES DE CAPTURE-RECAPTURE

En épidémiologie, les méthodes de capture-recapture sont fondées sur les détections multiples des unités épidémiologiques infectées. Dans cette partie, les différentes définitions de l'unité épidémiologique que l'on peut rencontrer dans les applications de capture-recapture en surveillance des maladies animales sont présentées, puis les principes des approches de capture-recapture multilistes et unilistes sont décrits.

1. L'UNITÉ ÉPIDÉMIOLOGIQUE

L'unité épidémiologique est définie comme l'unité dont on souhaite estimer l'abondance.

L'unité épidémiologique peut être l'unité surveillée, comme dans la plupart des applications en santé publique. En épidémiologie vétérinaire, l'unité surveillée est souvent l'animal, comme c'est le cas pour la tremblante classique du mouton en France [Morignat *et al.*, 2006]. Cependant, dans certains contextes de surveillance de maladies très contagieuses (telles que la fièvre aphteuse) ou de maladies affectant des espèces de faible valeur économique individuelle (telles que les volailles), l'unité surveillée est souvent définie à l'échelle de l'élevage (voire du village dans les pays du Sud).

L'unité épidémiologique peut aussi être définie à l'échelle d'un regroupement d'unités surveillées. L'unité épidémiologique peut alors être détectée plusieurs fois si plusieurs unités surveillées la composant sont détectées. C'est la démarche considérée dans chacune des applications à la tremblante classique où l'unité épidémiologique était représentée par l'élevage infecté alors que l'unité surveillée était l'animal infecté [Del Rio Vilas et Böhning, 2008 ; Del Rio Vilas *et al.*, 2005 ; Vergne *et al.*, 2012a]. Ce type d'unité épidémiologique est, dans la suite de cet article, appelé « unité épidémiologique élargie ».

2. LES APPROCHES MULTILISTES

Les approches multilistes sont utilisables lorsque les unités épidémiologiques infectées (les cas) peuvent être détectées par plusieurs protocoles de surveillance distincts (par exemple un protocole de surveillance événementielle et un protocole de surveillance active). Il est admis qu'un certain nombre d'unités épidémiologiques infectées ne sont détectées par aucun des protocoles. Ces approches sont qualifiées de *multilistes* car chaque protocole de détection génère une liste de cas. Les détections multiples des unités sont issues du croisement de ces listes de cas. Les données de détections croisées issues de *k* protocoles peuvent être consignées dans un tableau de contingence à 2^k dimensions dont la cellule associée aux cas non détectés par l'ensemble des protocoles est structurellement vide. La fréquence (non observée) associée à cette cellule est l'objet de l'inférence. Le tableau 2 présente un tel tableau pour trois protocoles de détection.

Tableau 2

Tableau de contingence représentant le croisement de trois protocoles de détection

		Protocole 1			
		Détecté		Non détecté	
		Protocole 3		**Protocole 3**	
		Détecté	Non détecté	Détecté	Non détecté
Protocole 2	Détecté	*a*	*d*	*c*	*g*
	Non détecté	*b*	*e*	*f*	-

En présence de seulement deux protocoles de détection, une estimation du nombre de cas non détectés peut être obtenue simplement en se servant de l'estimateur de Chapman [Chapman, 1951]. Un certain nombre d'hypothèses doivent cependant être vérifiées [Hook et Regal, 1995]. Les deux hypothèses les plus restrictives en épidémiologie concernent l'indépendance directe et l'indépendance indirecte entre les protocoles. L'indépendance directe suppose que le fait d'être détecté par un protocole ne modifie pas la probabilité d'être détecté par l'autre protocole. L'indépendance indirecte suppose que si les deux protocoles détectent les cas de manière hétérogène (certains cas sont plus facilement détectés que d'autres), les facteurs sous-tendant les hétérogénéités dans chacun des protocoles ne sont pas corrélés. Une violation de ces hypothèses engendre un biais dans l'estimation qu'il n'est pas possible de corriger [Hook et Regal, 1995]. En surveillance des maladies animales, cette méthode à deux sources a été appliquée à la theileriose en Tanzanie à l'échelle de l'élevage [Kivaria et Noordhuizen, 2009], et à la fièvre aphteuse au Cambodge à l'échelle des villages [Vergne *et al.*, 2012b].

En présence de trois protocoles ou plus ($k \geq 3$), les détections croisées sont généralement modélisées par les modèles log-linéaires introduits par Fienberg [1972]. L'avantage de ces applications à trois listes ou plus est que les dépendances directe et indirecte entre les protocoles de détection peuvent être prises en compte dès lors que les combinaisons d'interactions entre les sources sont testées dans le modèle et que l'interaction d'ordre k est supposée absente [Hook et Regal 1995]. En surveillance des maladies animales, cette approche à trois sources n'a été appliquée qu'au contexte de la surveillance de la tremblante classique des ovins en Grande-Bretagne à l'échelle des élevages [Del Rio Vilas *et al.*, 2005].

Il peut être parfois intéressant de démontrer que certaines sous-populations infectées souffrent de problèmes de détection plus importants que d'autres sous-populations. Stratifier la population générale et réaliser une analyse dans chaque strate peut permettre d'atteindre cet objectif. Cependant, une telle approche nécessite une importante quantité de données pour maintenir une puissance suffisante. Introduire les covariables pertinentes dans les modèles log-linéaires peut être une alternative intéressante [Tilling et Sterne, 1999].

3. LES APPROCHES UNILISTES

En présence d'un seul protocole de détection permettant des détections multiples des unités épidémiologiques, les approches unilistes peuvent être utilisées. L'information modélisée n'est plus les détections croisées entre protocoles, mais le nombre de détections de chaque unité par le protocole (par exemple le nombre d'animaux infectés détectés par élevage infecté). Il est admis qu'un certain nombre d'unités épidémiologiques infectées sont détectées zéro fois. En présence de telles données, deux grandes approches peuvent être utilisées pour estimer ce nombre d'unités épidémiologiques infectées ayant été détectées zéro fois : l'approche tronquée en zéro et l'approche enflée en zéro.

Dans l'approche tronquée en zéro, la population cible est composée uniquement des unités épidémiologiques infectées (comme dans l'approche multiliste). Les données tronquées en zéro donnent accès à la fréquence f_i des unités épidémiologiques infectées qui ont été détectées i fois, pour $i>0$. La donnée manquante est la fréquence f_0 (le nombre d'unités infectées qui ont été détectées zéro fois). L'objectif des approches unilistes tronquées en zéro est précisément d'estimer la fréquence de ces zéros, pour pouvoir estimer le nombre total d'unités infectées. Un exemple de données tronquées en zéro dans le contexte de la surveillance des maladies animales est celui s'intéressant au nombre de brebis infectées par l'agent de la tremblante qui ont été détectées dans les élevages infectés [Del Rio Vilas et Böhning, 2008]. Comme le nombre d'élevages infectés où aucune brebis infectée n'a été détectée est inconnu, la fréquence associée au comptage 0 n'est pas accessible. Une estimation de ce nombre d'unités infectées où aucune détection n'a eu lieu peut être faite soit en utilisant des estimateurs non paramétriques [Böhning *et al.*, 2011 ; Del Rio Vilas et Böhning, 2008], soit en utilisant des modèles paramétriques tels que les modèles de Poisson tronqués en zéro ou binomial négatif tronqués en zéro [Vergne *et al.*, 2012a].

Dans l'approche enflée en zéro, la population cible est cette fois l'ensemble de la population (composée des unités épidémiologiques infectées et des unités épidémiologiques non infectées). Le nombre d'unités qui ont été détectées zéro fois est donc connu, mais est composé d'unités non infectées et d'unités infectées et non détectées (figure 1). L'objet de l'analyse est donc d'estimer la proportion d'unités infectées parmi les unités où aucune

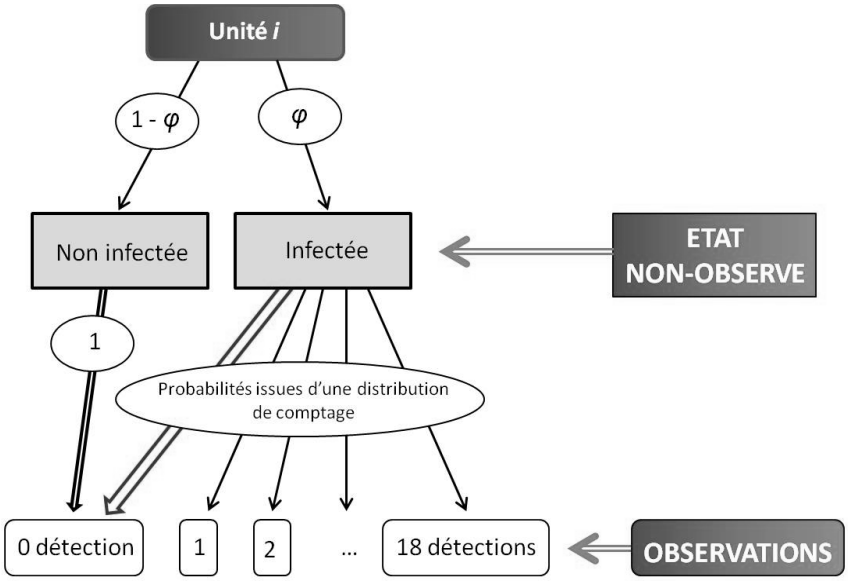
détection n'a eu lieu. La probabilité d'être infecté et non détecté (faux négatif) peut aussi être intéressante, et est accessible par ce genre de modèle. Les modèles enflés en zéro sont en cours d'adaptation pour évaluer les systèmes de surveillance [Bronner *et al.*, 2012].

La principale hypothèse de ces approches est que chaque unité épidémiologique infectée a la même probabilité d'être détectée et que cette probabilité est constante au cours du temps. Si cette hypothèse n'est pas vérifiée, ces modèles peuvent être généralisés pour prendre en compte des variables expliquant ces hétérogénéités entre les unités épidémiologiques [Cameron et Trivedi, 1998].

Figure 1

Représentation schématique d'un processus générant des données de comptage enflées en zéro en épidémiologie.

L'état infecté/non infecté de l'unité épidémiologique est un état latent (non observé) déterminé par la probabilité φ ; le nombre de détections d'une unité infectée est déterminé par un processus stochastique de comptage (distribution de Poisson ou binomiale négative) ; les cercles sur les flèches indiquent les probabilités associées à l'état indiqué.



III - INTÉRÊTS DES MÉTHODES DE CAPTURE-RECAPTURE EN SURVEILLANCE DES MALADIES ANIMALES

Les méthodes de capture-recapture sont intéressantes pour plusieurs raisons : elles peuvent permettre de corriger les prévalences apparentes issues de systèmes de surveillance imparfaits, elles permettent de décrire les facteurs influençant le processus de détection, et il s'avère qu'elles sont d'une utilisation suffisamment flexible pour s'adapter à la plupart des données générées par les systèmes de surveillance des maladies animales.

1. ESTIMATION DE LA SENSIBILITÉ DE LA SURVEILLANCE

1.1. ESTIMATION DE LA SENSIBILITE DE LA SURVEILLANCE A L'ECHELLE DE L'UNITE SURVEILLEE

Dans certaines situations, les analyses de capture-recapture peuvent permettre d'estimer le nombre réel d'unités surveillées infectées, et par conséquent le taux de sous-détection directement à l'échelle de l'unité surveillée.

L'interprétation d'une telle analyse en termes d'évaluation de la surveillance en est donc facilitée.

C'est la situation rencontrée dans la majorité des applications en santé publique, ainsi que dans l'application à la fièvre aphteuse au Cambodge [Vergne *et al.*, 2012b]. Cette application a permis d'évaluer le système de surveillance de la fièvre aphteuse à l'échelle de l'unité surveillée, le village : le système de surveillance de la fièvre aphteuse n'a vraisemblablement déclaré qu'un village infecté sur 20.

1.2. ESTIMATION DE LA SENSIBILITE DE LA SURVEILLANCE A L'ECHELLE DE L'UNITE EPIDEMIOLOGIQUE ELARGIE

Une analyse de capture-recapture réalisée à l'échelle d'une unité épidémiologique élargie permet d'estimer le nombre total d'unités épidémiologiques élargies infectées et ainsi la sensibilité de la surveillance à cette échelle. En revanche, elle ne permet pas d'estimer la sensibilité de la surveillance à l'échelle de l'unité surveillée.

D'un côté, c'est précisément pour cela que de telles analyses peuvent être intéressantes : alors que l'objectif principal du système de surveillance de la tremblante classique en France est d'estimer la prévalence à l'échelle individuelle, les méthodes de capture-recapture peuvent apporter une information complémentaire sur le nombre d'*élevages* infectés et en déduire la sensibilité de la surveillance à l'échelle des élevages [Vergne *et al.*, 2012a].

D'un autre côté, l'interprétation épidémiologique d'une analyse réalisée à l'échelle d'une unité épidémiologique élargie peut parfois devenir délicate. Prenons l'exemple de l'application à la surveillance de la tremblante classique au Royaume-Uni [Del Rio Vilas et Böhning, 2008]. L'analyse (réalisée à l'échelle des élevages alors que l'unité surveillée est l'animal) révèle que pendant la période considérée, la sensibilité de la surveillance était proche de 30% à l'échelle des élevages, c'est-à-dire que la surveillance a détecté *au moins* un animal infecté dans 30% des élevages dans lesquels *au moins un* animal infecté était présent. Il n'est en aucun cas possible de transposer cette sensibilité à l'échelle des animaux. Il est même probable que la sensibilité à l'échelle des individus est beaucoup moins bonne que la sensibilité à l'échelle des élevages car tous les animaux infectés d'un élevage détecté n'ont sans doute pas été détectés.

Considérer une unité épidémiologique élargie dans une application de capture-recapture permet donc les détections multiples, mais il est important de noter que cela se fait potentiellement au détriment d'une interprétation épidémiologique pertinente. En effet, suite à une application à cette échelle, l'estimation ne concerne que le nombre d'unités épidémiologiques infectées. Il est donc important que l'unité épidémiologique soit définie de façon judicieuse pour permettre une représentation épidémiologique appropriée de la population malade, et pour que l'estimation soit utile en termes de mesures de contrôle et de surveillance. Ainsi, plus l'unité épidémiologique est élargie (notamment à des échelles administratives telles que les districts ou les provinces), plus l'interprétation biologique est délicate. Pour des maladies très contagieuses telles que l'IAHP H5N1 ou la fièvre aphteuse, une unité épidémiologique définie à l'échelle d'une unité administrative (par exemple la commune ou le sous-district) serait tout à fait recevable alors que pour des maladies très peu diffusibles telles que la tremblante classique, une telle définition n'aurait pas de sens épidémiologique en termes de surveillance et de contrôle.

2. COMPRÉHENSION DES FACTEURS INFLUENÇANT LA DÉTECTION

L'autre intérêt majeur des analyses de capture-recapture (multilistes ou uniliste) est qu'elles peuvent permettre de mettre en évidence l'influence de certains facteurs sur le processus de détection des cas d'une maladie par le système de surveillance (les unités infectées présentant le facteur *k* sont plus détectables que celles ne le présentant pas). Ainsi, elles peuvent devenir un outil d'aide à la décision pour choisir les zones géographiques (ou les strates de la population) où il serait nécessaire de concentrer les efforts de surveillance pour améliorer la détectabilité. La mise en évidence de ces facteurs peut être faite soit en stratifiant l'analyse en fonction de ces facteurs [Böhning *et al.*, 2011 ; Gill *et al.*, 2003] soit en introduisant ces facteurs dans une régression [Tilling et Sterne, 1999 ; Van der Heijden *et al.*, 2003].

Au départ, la prise en compte de ces facteurs dans les analyses n'était motivée que par la volonté d'obtenir une estimation moins biaisée de la taille totale de la population infectée. Mais l'identification de ces facteurs peut

devenir l'objectif principal pour mettre en évidence des unités à risque d'infection et de non-détection, et proposer des recommandations de surveillance.

A ce jour, très peu d'études utilisent ces approches dans ce but (même dans le domaine de la santé publique), bien que les méthodes de capture-recapture semblent particulièrement adaptées. Les approches unilistes enflées en zéro, permettant de modéliser non seulement la présence de l'événement de santé mais aussi son processus de détection, nous apparaissent comme un outil de choix pour mener à bien cet objectif, comme suggéré par Bronner *et al.* [2012].

3. ADAPTABILITÉ DES MÉTHODES AUX DONNÉES GÉNÉRÉES

Quel que soit le type de système de surveillance envisagé, il s'avère donc qu'une évaluation de sa sensibilité peut systématiquement être réalisée par l'une ou l'autre des méthodes de capture-recapture disponibles, la plupart du temps sans nécessiter de coûts supplémentaires liés à la création d'une source de détection additionnelle.

Ainsi, en présence de données provenant d'un système de surveillance complexe (faisant intervenir plusieurs protocoles de surveillance), les approches multilistes peuvent être classiquement utilisées [Del Rio Vilas *et al.*, 2005], mais peuvent aussi être remplacées par les approches unilistes [Del Rio Vilas et Böhning., 2008 ; van Hest *et al.*, 2008 ; Vergne *et al.*, 2012a]. Le choix d'une approche plutôt qu'une autre dépendra des contraintes liées à la collecte des données qu'il est nécessaire de prendre en compte dans l'analyse (dépendance entre sources, hétérogénéité de la détection).

En présence de données provenant d'un système de surveillance simple (ne faisant intervenir qu'un seul protocole de surveillance), les approches unilistes peuvent être directement employées, pour peu que les détections multiples des unités épidémiologiques pertinentes soient possibles [Bronner *et al.*, 2012].

Si le système de surveillance ne permet pas la détection multiple des unités surveillées, il est possible d'élargir l'unité épidémiologique et d'étudier la sensibilité de la surveillance à cette échelle (si elle a un sens épidémiologique). Sinon, si un financement est disponible, il est toujours possible de créer spécifiquement un protocole rétrospectif de détection des cas, et d'utiliser une approche multilistes à deux listes pour estimer la sensibilité du système de surveillance existant à l'échelle des unités surveillées [Vergne *et al.*, 2012b]

Cependant, la précision de l'estimation de la sensibilité dépend beaucoup de la qualité et de la quantité des données issues de la surveillance : plus la qualité et la quantité des données est importante, plus il est facile de prendre en compte des facteurs supposés influencer la surveillance pour affiner les estimations ou mettre en évidence le rôle de ces facteurs. Paradoxalement, il ressort donc qu'en présence de peu de données de détection (soit parce que la maladie est présente à un très bas niveau, soit parce que la surveillance est très déficiente), les analyses de capture-recapture sont d'un intérêt limité [Vergne *et al.*, 2012a].

IV - LIMITES DES MÉTHODES DE CAPTURE-RECAPTURE EN SURVEILLANCE DES MALADIES ANIMALES

Bien que les méthodes de capture-recapture soient parmi les seuls outils formalisés permettant l'évaluation quantitative des systèmes de surveillance, un certain nombre de caractéristiques propres à la surveillance des maladies animales en restreignent l'utilisation ou détériorent l'exactitude de l'inférence.

1. RARETÉ DES DÉTECTIONS MULTIPLES DE L'UNITÉ SURVEILLÉE

En surveillance des maladies animales, les pratiques de détection et de contrôle des maladies limitent les possibilités de détections multiples des unités surveillées, restreignant l'applicabilité des méthodes de capture-recapture. Deux principales raisons peuvent être invoquées : la brièveté de l'infection et les pratiques de dépistage post-mortem.

1.1. BRIEVETE DE L'INFECTION

Dans le domaine de la santé publique, les applications des méthodes de capture-recapture aux maladies infectieuses sont beaucoup moins courantes que celles concernant des affections chroniques telles que le diabète, le syndrome de Down ou les cancers [van Hest, 2007]. Ceci est dû au fait que les individus atteints d'affections chroniques ou permanentes peuvent être détectés plus longtemps et donc possiblement plus de fois que les individus atteints de maladies infectieuses qui elles, peuvent être plus facilement traitées et donc disparaitre. Satisfaisant de manière plus sûre les hypothèses d'utilisation, il est donc plus aisé d'appliquer des méthodes de capture-recapture aux individus atteints de conditions chroniques ou permanentes que de maladies infectieuses aiguës [van Hest, 2007].

En surveillance des maladies infectieuses *animales*, cet argument est d'autant plus valable que les méthodes utilisées pour faire disparaitre l'infection peuvent être beaucoup plus radicales qu'en santé publique. En effet, bien que le traitement médical soit aussi pratiqué, l'abattage des animaux infectés est très souvent utilisé dans le but d'éviter la contagion à d'autres animaux.

Ainsi, dans le cadre du contrôle de certaines maladies infectieuses très contagieuses, l'abattage des animaux infectés est souvent la règle (notamment dans les pays développés). Dans ces contextes-là, lorsqu'un animal infecté est détecté, il est abattu. Par conséquent, l'animal infecté ne peut alors plus être détecté une nouvelle fois par le système de surveillance. Ces pratiques sont courantes en Europe pour contrôler des maladies infectieuses contagieuses telles que la tuberculose [Cousins, 2001] ou la peste porcine classique [Elbers *et al.*, 1999].

1.2. DEPISTAGE POST-MORTEM

Lorsque l'animal représente l'unité surveillée, une des approches pour détecter l'agent pathogène d'intérêt est de prélever l'animal dans des situations rendant l'échantillonnage pratique. Ces situations peuvent être rencontrées à l'abattoir, où les animaux de rente destinés à la consommation sont systématiquement conduits, ou à l'équarrissage, où les animaux morts en élevage sont regroupés.

Le problème est que l'animal est prélevé une fois mort. Donc, quand il est détecté (il existe un certain délai entre le prélèvement et le résultat du test), l'animal ne fait déjà plus partie de la population surveillée, empêchant donc les possibilités de détection multiple du même animal.

Ces pratiques de surveillance sont effectivement très utiles pour détecter la présence de maladies à l'échelle individuelle, et sont donc assez souvent mises en œuvre dans les systèmes de surveillance. A titre d'exemple, elles sont rencontrées dans le cadre de la surveillance de l'influenza porcin pandémique H1N1 (2009) au Vietnam [Trevennec *et al.*, 2012], ou dans le cadre de la surveillance de la tremblante classique ovine en France [Morignat *et al.*, 2006].

1.3. CONSEQUENCES SUR L'APPLICABILITE DES METHODES DE CAPTURE-RECAPTURE

La détection multiple des unités surveillées étant un phénomène relativement rare en surveillance des maladies infectieuses animales, l'applicabilité des méthodes de capture-recapture utilisant les données de surveillance à cette échelle est par conséquent très limitée.

Pour contourner ce problème, une première solution est d'utiliser une enquête ponctuelle extérieure au système de surveillance dans l'objectif de corriger la prévalence apparente issue du système de surveillance existant [Cameron, 1999 ; Vergne *et al.*, 2012b]. Le problème d'une telle approche est le coût inhérent à la mise en place d'une telle étude, et notamment le coût des analyses de laboratoire qu'il est indispensable de réaliser pour limiter au maximum la présence de faux positifs dans cette source de détection extérieure.

L'autre solution pour contourner le problème dû à l'absence de détections multiples des unités surveillées, est d'élargir l'unité épidémiologique à une échelle comprenant plusieurs unités surveillées sous réserve que l'unité épidémiologique ainsi définie conserve une signification épidémiologique en termes de surveillance ou de contrôle (voir *supra*) [Del Rio Vilas et Böhning, 2008 ; Del Rio Vilas *et al.*, 2005 ; Vergne *et al.*, 2012a].

2. HÉTÉROGÉNÉITE D'ABONDANCE DE LA MALADIE

Définir l'unité épidémiologique à une échelle supérieure à l'unité surveillée est une

approche possible pour permettre les détections multiples. Cependant, cette démarche introduit un biais spécifique dans les applications de capture-recapture en surveillance des maladies infectieuses animales : le biais provenant de l'hétérogénéité des unités épidémiologiques liée à l'abondance de l'infection, que l'on appellera par la suite « l'hétérogénéité d'abondance » et qui conduit à une hétérogénéité de la probabilité de détecter chaque unité épidémiologique.

Ceci signifie que dans les unités épidémiologiques où beaucoup d'unités surveillées sont infectées (c'est-à-dire où l'abondance de l'infection est grande), la probabilité de détecter au moins une unité surveillée infectée est plus grande que dans les unités épidémiologiques où peu d'unités surveillées sont infectées. Si elle n'est pas prise en compte, cette hétérogénéité conduit à une estimation biaisée du nombre total de cas que ce soit en utilisant les approches multilistes [Hook et Regal, 1995] ou les approches unilistes [Royle et Nichols, 2003].

A moins de supposer que le nombre d'unités surveillées infectées est le même dans toutes les unités épidémiologiques, il est important d'essayer de tenir compte de cette hétérogénéité. Le problème est qu'elle est difficile à apprécier quantitativement, car elle n'est pas observable.

La manière la plus simple de prendre en compte cette hétérogénéité d'abondance est de la modéliser indirectement en introduisant de manière traditionnelle (stratification ou régression) des covariables influant sur l'abondance de l'infection dans les unités épidémiologiques (et donc indirectement sur la probabilité de détection). Ces covariables peuvent par exemple être le nombre d'unités surveillées dans chaque unité épidémiologique ou les facteurs de risque associés à la maladie d'intérêt [Bronner *et al.*, 2012]. Cette solution est valable pour les approches multilistes ou les approches unilistes.

Une autre possibilité est de modéliser explicitement l'hétérogénéité d'abondance par une distribution paramétrique discrète à inclure dans un modèle hiérarchique à état latent [Royle et Nichols, 2003 ; Royle *et al.*, 2005]. Cette approche, qui n'est évidemment valable que pour les approches unilistes (tronquées ou enflées en zéro), nécessite encore des développements méthodologiques car elle produit souvent des estimations instables.

A défaut de pouvoir modéliser explicitement l'hétérogénéité d'abondance dans les approches unilistes, il est simplement possible d'ajuster le modèle en fonction de cette hétérogénéité, en utilisant des distributions de comptage flexibles telles que la binomiale négative. C'est ce qui a été fait dans Vergne *et al.* [2012a] où le peu de données disponibles empêchait une modélisation complexe.

3. POLICE SANITAIRE ET DÉPENDANCE NÉGATIVE DES DÉTECTIONS

Au-delà d'empêcher la détection multiple d'une unité surveillée infectée (*cf. supra*), les mesures de police sanitaire limitent aussi les possibilités de détections multiples des unités épidémiologiques élargies. En effet, l'abattage total d'un troupeau consiste à éliminer l'ensemble des animaux du troupeau (ou l'ensemble des animaux à risque du troupeau) dès lors qu'au moins un animal infecté par l'agent pathogène d'intérêt a été détecté dans ce troupeau. Une fois que les mesures de police sanitaire ont été mises en place, l'unité épidémiologique d'intérêt (le troupeau infecté) n'est plus présente, empêchant les futures détections.

Les détections multiples s'en trouvent le plus souvent limitées mais pas impossibles. En effet, les animaux étant quelquefois testés par lots (à l'abattoir ou à l'équarrissage), plusieurs animaux infectés peuvent être détectés en même temps. De plus, le délai entre la suspicion de l'infection chez cet animal (ou ce groupe d'animaux suspectés simultanément) et la confirmation de son (leur) infection n'étant pas nul, d'autres animaux infectés du même troupeau peuvent quand même être aussi détectés avant la mise en place des mesures de police sanitaire. Il est donc important de remarquer que les détections multiples d'un élevage infecté restent possibles [Vergne *et al.*, 2012a].

Cependant, le délai entre la suspicion et la mise en place des mesures de police sanitaire étant court, la probabilité qu'un animal infecté soit détecté pendant ce délai est sans doute beaucoup plus faible que la probabilité qu'un animal infecté soit détecté en l'absence de mesures de police sanitaire (comme pour le premier animal détecté). Il en résulte une dépendance négative des détections croisées (dans les approches multilistes) ou des détections successives (dans les approches unilistes) conduisant nécessairement à une surestimation du nombre total d'unités épidémiologiques infectées [Chao, 2001].

A notre connaissance, les outils méthodologiques permettant de surmonter cette limite ne sont pas encore disponibles. C'est la raison pour laquelle, les applications de capture-recapture à des maladies dont le contrôle passe par des mesures de police sanitaire doivent être discutées avec attention.

V - CONCLUSION

Si les méthodes de capture-recapture (notamment les approches multilistes) ont déjà fait leurs preuves dans le domaine de la santé publique en tant qu'outil d'évaluation de la surveillance, leur développement en surveillance des maladies animales n'en est encore qu'à ses débuts.

Les méthodes de capture-recapture permettent à un moindre coût (la plupart du temps) d'estimer effectivement l'importance de la sous-détection d'une maladie animale sur un territoire, voire d'étudier les facteurs impliqués dans le processus de détection. Leur grande force est que la diversité des approches disponibles permet de s'adapter à la diversité des systèmes de surveillance existants en proposant pour la plupart des situations une méthode adaptée.

Dans des contextes aux ressources limitées où la surveillance des maladies infectieuses contagieuses est souvent d'une faible efficacité, il est possible et souhaitable d'envisager ponctuellement la création d'un protocole de détection additionnel pour corriger la prévalence apparente de la maladie grâce à une approche capture-recapture à deux sources. Malgré le coût inhérent à cette méthode, elle apparait d'une bonne pertinence pour les prises de décision sanitaire car elle permet d'évaluer la situation épidémiologique « réelle » des maladies contagieuses d'intérêt mondial en corrigeant les prévalences apparentes issues de systèmes de surveillance imparfaits.

Les réflexions présentées dans cet article n'ont porté que sur l'utilisation des méthodes de capture-recapture en tant qu'outil d'évaluation des systèmes de surveillance des maladies animales : ces méthodes ont jusqu'à présent été utilisées sur des données préexistantes de surveillance, dans l'objectif d'estimer la fraction non détectée. Il serait intéressant de commencer à réfléchir si, à l'avenir, ces méthodes ne pourraient pas être considérées comme des outils de surveillance à part entière, directement intégrées dans les systèmes de surveillance. Ce travail n'est donc qu'une introduction au domaine riche et prometteur qu'est l'utilisation des méthodes de capture-recapture en surveillance des maladies infectieuses animales et ouvre potentiellement la voie à d'importants développements méthodologiques en modélisation et en statistique.

BIBLIOGRAPHIE

Böhning D., Del Rio Vilas V. - Estimating the hidden number of scrapie affected holdings in Great Britain using a simple, truncated count model allowing for heterogeneity. *J. Agr. Biol. Envir. Stat.*, 2008, **13**, 1-22.

Böhning D., Del Rio Vilas V.J. - On the question of proportionality of the count of observed scrapie cases and the size of holding. *BMC Vet. Res.*, 2009, **5**, 17.

Böhning D., Kunhert R., Del Rio Vilas V.J. - Capture-recapture estimation by means of empirical Bayesian smoothing with an application to the geographical distribution of hidden scrapie in Great Britain. *Applied statistics*, 2011, **60**, 723-741.

Bronner A., Gay E., Vergne T., Hendrikx P., Calavas D. - Evaluation du dispositif de surveillance évènementielle de la brucellose bovine à partir des méthodes de capture-recapture unilistes. *Epidémiologie et santé animale*, 2012, **61**, 79-94.

Cameron A. - Survey Toolbox for livestock diseases: a practical manual and software package for active surveillance in developing countries. *http://aciar.gov.au/publication/MN054*, 1999.

Cameron A., Trivedi P. - Regression analysis of count data. *Economtric Society Monographs*, 1998, N°30. Cambridge University Press.

Chao A. - An overview of closed capture-recapture models. *J. Agr. Biol. Envir. Stat.*, 2001, **6**, 158-175.

Chapman C. - Some properties of the hypergeometric distribution with applications to zoological censuses. *U California Public Stat.*, 1951, **1**, 131 - 160.

Cousins D. - *Mycobacterium bovis* infection and control in domestic livestock. *Rev. Sci. Tech.*, 2001, **20**, 71-85.

Del Rio Vilas V.J., Böhning D. - Application of one-list capture-recapture models to scrapie surveillance data in Great Britain. *Prev. Vet. Med.*, 2008, **85**, 253-266.

Del Rio Vilas V.J., Sayers R., Sivam K., Pfeiffer D., Guitian J., Wilesmith J.W. - A case study of capture-recapture methodology using scrapie surveillance data in Great Britain. *Prev. Vet. Med.*, 2005, **67**, 303-317.

Drewe J.A., Hoinville L.J., Cook A.J., Floyd T., Stark K.D. - Evaluation of animal and public health surveillance systems: a systematic review. *Epidemiol. Infect.*, 2012, **140**, 575-590.

Elbers A.R., Stageman A., Moser H., Ekker H., Smak J., Pluimers F. - The classical swine fever epidemics 1997-1998 in the Netherlands: descriptive epidemiology. *Prev. Vet. Med.*, 1999, **42**, 157-184.

Fienberg S. - The multiple-recapture census for closed populations and incomplete 2^k contingency tables. *Biometrika*, 1972, **59**, 591-603.

Gill G.V., Ismail A.A., Beeching N.J., Macfarlane S.B., Bellis M.A. - Hidden diabetes in the UK: use of capture-recapture methods to estimate total prevalence of diabetes mellitus in an urban population. *J. R. Soc. Med.*, 2003, **96**, 328-332.

Hook E.B., Regal R.R. - Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol. Rev.*, 1995, **17**, 243-264.

IWGDMF. - Capture-recapture and multiple-record systems estimation I: History and theoretical development. International Working Group for Disease Monitoring and Forecasting. *Am. J. Epidemiol.*, 1995a, **142**, 1047-1058.

IWGDMF. - Capture-recapture and multiple-record systems estimation II: Applications in human diseases. International Working Group for Disease Monitoring and Forecasting. *Am. J. Epidemiol.*, 1995b, **142**, 1059-1068.

Kivaria F.M., Noordhuizen J.P. - Capture-recapture analysis of East Coast fever in smallholder dairy herds in the Dar es Salaam region of Tanzania. *Vet. J.*, 2009, **184**, 187-193.

Morignat E., Cazeau G., Biacabe A.G., Vinard J.L., Bencsik A., Madec J.Y., Ducrot C., Baron T., Calavas D. - Estimates of the prevalence of transmissible spongiform encephalopathies in sheep and goats in France in 2002. *Vet. Rec.*, 2006, **158**, 683-687.

Navaratna W.C., Del Rio Vilas V.J., Bohning D. - Extending Zelterman's approach for robust estimation of population size to zero-truncated clustered Data. *Biom. J.*, 2008, **50**, 584-596.

Royle J.A., Nichols J.D. - Estimating abundance from repeated presence-absence data data or point counts. *Ecology*, 2003, **84**, 777-790.

Royle J.A., Nichols J.D., Kéry M. - Modelling occurrence and abundance of species when detection is imperfect. *Oikes*, 2005, **110**, 353-359.

Tilling K., Sterne J.A. - Capture-recapture models including covariate effects. *Am. J. Epidemiol.*, 1999, **149**, 392-400.

Trevennec K., Leger L., Lyazrhi F., Baudon E., Cheung C.Y., Roger F., Sriyal Peiris J.M., Garcia J.M. - Transmission of pandemic influenza H1N1 (2009) in Vietnamese swine in 2009-2010. *Influenza Other Respi Viruses*, 2012, DOI: 10.1111/j.1750-2659.2011.00324.x.

Van der Heijden P.G., Bustami R., Cruyff M.J., Engbersen G., Houwelingen H. - Point and interval estimation of the population size using the truncated Poisson regression model. *Statistical Modelling*, 2003, **3**, 305-322.

van Hest N.A., Grant A.D., Smit F., Story A., Richardus J.H. - Estimating infectious diseases incidence: validity of capture-recapture analysis and truncated models for

incomplete count data. *Epidemiol. Infect.*, 2008, **136**, 14-22.

van Hest R. - Capture-recapture methods in surveillance of tuberculosis and other infectious diseases (thesis). *Rotterdam: Erasmus University*, 2007.

Vergne T., Calavas D., Cazeau G., Durand B., Dufour B., Grosbois V. - A Bayesian zero-truncated approach for analysing capture-recapture count data from classical scrapie surveillance in France. *Prev. Vet. Med.*, 2012a, **105**, 127-135.

Vergne T., Grosbois V., Durand B., Goutard F., Bellet C., Holl D., Roger F., Dufour B. - A capture-recapture analysis in a challenging environment: Assessing the epidemiological situation of foot-and-mouth disease in Cambodia. *Prev. Vet. Med.*, 2012b, **105**, 235-243.

Wittes J. - Applications of a multinomial capture-recapture model to epidemiological data. *J. Am. Stat.*, 1974, **69**, 93-107.

Wittes J., Sidel V.W. - A generalization of the simple capture-recapture model with applications to epidemiological research. *J. Chronic Dis.*, 1968, **21**, 287-301.