

# SYSTEME DE SURVEILLANCE SYNDROMIQUE MULTIVARIE DES MALADIES BOVINES EN SUISSE : SIMULATION D'EPIZOOTIES ET EVALUATION D'ALGORITHMES DE DETECTION\*

Faverjon Céline<sup>1</sup>, Carmo Luís Pedro<sup>1</sup>, Berezowski John<sup>1</sup>



## RESUME

Les systèmes de surveillance syndromique (SSy) multivariés offrent des perspectives intéressantes pour renforcer la détection précoce des épizooties. Cependant, le nombre de ces systèmes opérationnels en santé animal est faible car il reste difficile d'évaluer leurs performances. L'objectif de cette étude était d'évaluer un système de SSy multivarié pour les maladies bovines en Suisse en utilisant une méthode standardisée pour simuler des épizooties multivariées réalistes de différentes maladies et en comparant les performances de détection de deux algorithmes : la carte multivariée de moyenne mobile pondérée exponentiellement (MEWMA), et la carte multivariée de sommes cumulées (MCUSUM). Ces deux algorithmes permettent de combiner quantitativement les informations fournies par plusieurs séries temporelles en un unique indicateur facilement interprétable. Douze séries temporelles de syndromes ont été extraits de deux bases de données nationales suisses pour la période 2013-2017. Des épizooties de quatre maladies bovines (la diarrhée bovine virale (BVD), la rhinotrachéite infectieuse bovine (IBR), la fièvre catarrhale du mouton (FCO) et l'infection à virus Schmallenberg (SV)) ont ensuite été simulées à l'aide d'informations obtenues par consultation d'experts. Les deux algorithmes ont permis de détecter toutes les épizooties simulées environ 4,5 mois après le début de l'épizootie, avec une spécificité de 95 %. Cependant, les résultats variaient selon l'algorithme et la maladie considérée. Nos résultats montrent que les deux cartes de contrôle directionnelles multivariées sont des méthodes prometteuses pour combiner les informations provenant de plusieurs séries temporelles afin de détecter rapidement les changements s'opérant dans une population tout en maintenant le nombre de fausses alarmes positives à un niveau raisonnable.

**Mots-clés :** surveillance syndromique, MEWMA, MCUSUM, séries temporelles, cartes de contrôle multivariées, simulation d'épizooties.

## ABSTRACT

*Multivariate syndromic surveillance systems offer interesting opportunities to strengthen early detection of epidemics. However, the number of operational multivariate SSy systems in animal health is low because it remains difficult to assess the performance of these systems. The objective of this study was to evaluate a multivariate SSy system for bovine diseases in Switzerland using a standardized method to simulate realistic multivariate epidemics of different diseases and comparing the detection performance of two algorithms: the Multivariate Exponentially Weighted Moving Average (MEWMA), and the Multivariate Cumulative Sum (MCUSUM).*

.../..

Reçu le 12 juin 2019, accepté le 25 août 2019

\* Texte de la communication orale présentée au cours de la Journée scientifique AEEMA, 24 mai 2019

<sup>1</sup> Institut de santé publique vétérinaire, Faculté Vétérinaire, Université de Berne, Liebefeld, Suisse

.../..

*These two algorithms make it possible to combine quantitatively the information provided by several time series into a single easily interpretable indicator. Twelve time series of syndromes were extracted from two Swiss national databases for the 2013-2017 period. Then outbreaks of four bovine diseases (i.e., viral bovine diarrhoea (BVD), infectious bovine rhinotracheitis (IBR), bluetongue virus (BTV) and Schmallenberg virus (SV)) were simulated, using information obtained through experts' consultation. Both algorithms detected all simulated epidemics about 4.5 months after the beginning of the epidemic, with a specificity of 95%. However, the results varied depending on the algorithm and the disease considered. Our results show that the two multivariate directional control charts are promising methods for combining information from several time series to quickly detect changes in a population, while maintaining the number of false positive alarms at a reasonable level.*

**Keyword:** *Syndromic surveillance, MEWMA, MCUSUM, Time series, Directional multivariate control charts, Epidemic simulation.*




---

## I - INTRODUCTION

---

La surveillance syndromique (SSy) est fondée sur l'analyse en temps réel, ou quasi réel, des données relatives à la santé qui sont disponibles avant la confirmation en laboratoire. Depuis le début du 21<sup>e</sup> siècle, la SSy a été utilisée pour renforcer la surveillance événementielle des maladies, améliorer les systèmes d'alerte précoce et mieux contrôler les maladies émergentes ou ré-émergentes. Les systèmes de SSy multivariés qui évaluent et combinent simultanément les informations provenant de différentes sources de données peuvent être utiles pour renforcer les systèmes de détection précoce des épizooties de maladies infectieuses. En effet, une maladie cause souvent une grande variété de symptômes et/ou affecte différentes sous-populations. Par exemple, les avortements sont un symptôme qui ne concerne que les femelles en âge de procréer. Une série temporelle (ST) relative à l'évolution temporelle d'un unique syndrome ne peut donc pas saisir tous ces différents types d'informations. Utiliser des systèmes de SSy multivariés offre donc des perspectives intéressantes et a le potentiel d'accroître la sensibilité et la spécificité de détection des épizooties [Dorea et Vial, 2016 ; Hopkins *et al.*, 2017].

Le renforcement de la surveillance pour la détection précoce des maladies animales est une priorité pour la Suisse qui est indemne de nombreuses maladies encore enzootiques dans une grande majorité de pays européens. La Suisse est officiellement indemne de rhinotrachéite infectieuse bovine (RIB) depuis 1990 et a lancé un programme d'éradication de la diarrhée virale bovine (BVD) en 2008, qui a réduit considérablement le nombre de cas dans le pays. Détecter rapidement, et avec certitude, les réémergences de ces maladies est donc d'une importance primordiale pour le pays. Les deux dernières grandes épizooties bovines en Europe, la fièvre catarrhale ovine (FCO) en 2006 et l'infection par le virus Schmallenberg (SV) en 2011, n'ont provoqué qu'un petit nombre de cas en Suisse. Cependant, le risque de nouvelles épizooties de ces maladies en Europe est loin d'être négligeable, comme l'illustre la récente réémergence de la FCO en France [Courtejoie *et al.*, 2018]. Les maladies telles que l'RIB, la BVD, la FCO et la SV produisent généralement des signes cliniques non spécifiques, ce qui peut rendre plus difficile la détection précoce des épizooties avec les systèmes traditionnels de surveillance événementielle [Doherr et Audigé, 2001]. Les SSy multivariées ont le potentiel de renforcer les systèmes de détection précoce de ces maladies bovines ré-émergentes en Suisse.

De nombreuses méthodes temporelles ont été proposées pour la SSy multivariés. Les approches temporelles peuvent être classées en trois grandes catégories [Sonesson et Frisé, 2005 ; Vial *et al.*, 2016] : surveillance parallèle, réduction

dimensionnelle et accumulation de vecteurs. La surveillance parallèle est une méthode qui combine plusieurs algorithmes de détection uni-variés appliqués séparément à chaque ST. Une alerte est déclenchée en fonction du nombre de ST produisant une alerte [Sonesson et Frisé, 2005 ; Vial *et al.*, 2016]. Les méthodes de réduction dimensionnelle résument toutes les informations disponibles en une seule statistique et ce pour chaque période de temps considérée. Une méthode populaire de réduction est la statistique Hotelling  $T^2$ , mais d'autres approches utilisant des tests d'hypothèse bayésiens ont également été proposées. Les approches d'accumulation vectorielle sont des méthodes dans lesquelles l'information de chaque ST est accumulée et transformée en une statistique d'alarme scalaire. Les cartes de contrôle multivariées comme les sommes cumulées multivariées (MCUSUM) et la moyenne mobile pondérée exponentiellement (MEWMA) appartiennent aux méthodes de réduction dimensionnelle. En général les cartes de contrôle multivariées peuvent détecter un changement de la valeur moyenne des STs dans n'importe quelle direction. Dans le cas de la détection précoce des maladies, l'intérêt est souvent de détecter un changement dans une seule direction, habituellement une augmentation du nombre de cas cliniques. Des cartes de contrôle multivariées directionnelles ont ainsi été développées et sont recommandées pour améliorer la performance de détection des algorithmes de détection. Par rapport à Hotelling  $T^2$ , MEWMA et MCUSUM sont plus performants pour détecter de petits changements s'opérant dans les différentes STs du système de surveillance et devraient être plus efficaces pour détecter les épizooties le plus précocement possible, notamment en début d'épizootie ou le nombre de cas est en général faible.

Malgré les nombreux avantages potentiels à mettre en place des systèmes de surveillance multivariés et les nombreuses méthodes déjà disponibles, le nombre de systèmes SSy multivariés opérationnels en médecine vétérinaire est encore très faible [Dorea et Vial, 2016]. L'une des raisons possibles est l'écart toujours présent entre la recherche et la surveillance

[Hopkins *et al.*, 2017], notamment en ce qui concerne les méthodes d'évaluation des performances de détection de ces systèmes de surveillance. En effet, les exemples dans la littérature en santé publique sont rares, et encore plus en santé publique vétérinaire. Lorsqu'ils existent, ces exemples ne portent que sur l'aide à la détection d'une seule maladie et/ou un nombre limité de sources de données réelles [Brouwer *et al.*, 2015 ; Burkom *et al.*, 2011 ; Miekley *et al.*, 2013 ; Siegrist et Pavlin, 2004 ; Vial *et al.*, 2016] ou simulées [Tokars *et al.*, 2009 ; Xing *et al.*, 2011 ; Yahav et Shmueli, 2014). L'utilisation de données réelles soulève la question de savoir comment le système de surveillance fonctionnera lorsqu'une épizootie ayant des caractéristiques différentes apparaîtra dans la population. Lorsque des épizooties multivariées (c.-à-d. des épizooties affectant plusieurs STs) sont simulées, les auteurs comparent souvent la performance des algorithmes dans des conditions normalisées en utilisant, par exemple, un seul pic épizootique apparaissant au même moment et ayant la même forme et la même ampleur dans toutes les STs surveillées. Ces épizooties simulées sont loin d'être représentatives du comportement des épizooties réelles, et les résultats des évaluations ne peuvent donc pas être généralisés à des situations réelles. Certains auteurs ont proposé de simuler des épizooties multivariées plus réalistes en utilisant des élicitations d'experts, des données historiques ou des modèles de compartiment [Colón-González *et al.*, 2018 ; Faverjon *et al.*, 2015 ; Lotze *et al.*, 2007]. Cependant, le nombre d'exemples dans la littérature est encore faible et la simulation d'épizooties plus réalistes dans plusieurs ST à la fois reste un défi.

L'objectif de cette étude était d'évaluer un système de surveillance syndromique multivarié des maladies bovines en Suisse en utilisant des épizooties simulées. Nous présentons les performances de détection de deux cartes de contrôle directionnelles (MEWMA et MCUSUM) appliquées sur données d'épizooties simulées et plus réalistes de quatre maladies d'intérêt pour la Suisse : la BVD, RIB, FCO et SV.

---

## II - MATÉRIEL ET MÉTHODES

---

### 1. SOURCE DE DONNEES ET CREATION DES SERIES TEMPORELLES

Deux bases de données contenant des données du cheptel bovin suisse ont été utilisées : la base de données suisse sur les transports d'animaux (AMD) et une base de données appartenant à l'Association suisse des éleveurs de bovins (ASR).

La base de données AMD a fait l'objet d'études précédentes et a été signalée comme ayant une valeur potentielle pour la surveillance syndromique en raison de sa qualité relativement élevée en termes de représentativité de la population [Struchen *et al.*, 2015]. AMD centralise les données sur la mortalité des bovins. Tous les décès observés dans les élevages et déclarés entre le 1er janvier 2009 et le 31 décembre 2017 ont été extraits de l'AMD. Quatre STs ont été créées en fonction de l'âge des animaux : mortinatalité (c.-à-d. 0-1 jour) (AMD\_stillbirth), jusqu'à six mois (AMD\_mortality\_calves), de 6 mois à deux ans (AMD\_mortality\_young), et plus de deux ans (AMD\_mortality\_adults).

L'un des objectifs de l'Association suisse des éleveurs de bovins (ASR) (<http://asr-ch.ch/en/asr/>) est de coordonner les activités de ses membres. Cette base de données n'a pas encore été étudiée en Suisse, mais des données similaires ont été rapportées comme étant utiles pour SSy dans d'autres pays [Dórea *et al.*, 2014]. Depuis 2013, l'ASR a développé et mis en place une base de données homogénéisée contenant les symptômes observés par les éleveurs et les vétérinaires. Tous les cas sont signalés à l'aide d'un système de codage à quatre niveaux allant du moins spécifique (c.-à-d. organe touché) au plus spécifique (p. ex. agent infectieux isolé). Les données étaient disponibles pour les trois races les plus courantes en Suisse : Braunvieh, Fleckvieh et Holstein, qui représentent la majorité des bovins laitiers suisses. Les données étaient disponibles du 1er janvier 2014 au 31 décembre 2017. Dans la base de données ASR, les veaux sont définis comme des bovins âgés de moins de 6 mois. Tous les autres animaux sont définis comme adultes. Trois ST fondées sur les syndromes les plus fréquents de la base de données ont été créés pour chaque catégorie d'âge : les syndromes gastro-intestinaux (c.-à-d. ASR\_GI\_calves et ASR\_GI\_adults), les syndromes respiratoires (ASR\_RESPI\_calves et ASR\_RESPI\_adults) et les bovins ayant une classification "autres" dans le schéma de classification ASR (c.-à-d., ASR\_OTHER\_calves et ASR\_OTHER\_adults). La catégorie "autres" englobe divers symptômes non spécifiques tels que la fièvre, l'anorexie, un changement de comportement ou une réduction de la

production. Deux autres syndromes, troubles de fertilité (ASR\_FERTILITY) et de locomotion (ASR\_LOCO), ont été créés, mais les données ne permettent pas de distinguer les veaux des adultes.

Au total, les données de six syndromes ont été extraites des deux bases de données et converties en 12 STs hebdomadaires (figure 1).

### 2. SIMULATION DE DONNEES

#### 2.1 SIMULATION DE DONNEES DE REFERENCE EXEMPTES D'ÉPIZOOTIE

Nous avons simulé des séries temporelles sans épizootie pour chaque ST à l'aide de prédictions obtenues à partir du lissage exponentiel généralisé de Holt-Winters (HW) [Chatfield, 1978]. HW est une méthode populaire de modélisation de STs qui n'a besoin que de deux années de données historiques pour être utilisée. Toutes les données disponibles de 2014 à 2015 ont été utilisées pour déterminer les paramètres du modèle en minimisant l'erreur quadratique de prédiction [Kalekar, 2004]. Les données disponibles de 2016 à 2017 ont été utilisées pour valider les modèles et pour l'estimation de leurs performances de prédiction. La valeur moyenne de prédiction obtenue avec le meilleur modèle HW a été utilisée comme moyenne pour une distribution de Poisson. Cette distribution de Poisson a ensuite été échantillonnée de manière aléatoire pour simuler 300 STs pour chacun des 12 syndromes évalués dans cette étude.

#### 2.2 SIMULATION D'ÉPIZOOTIE

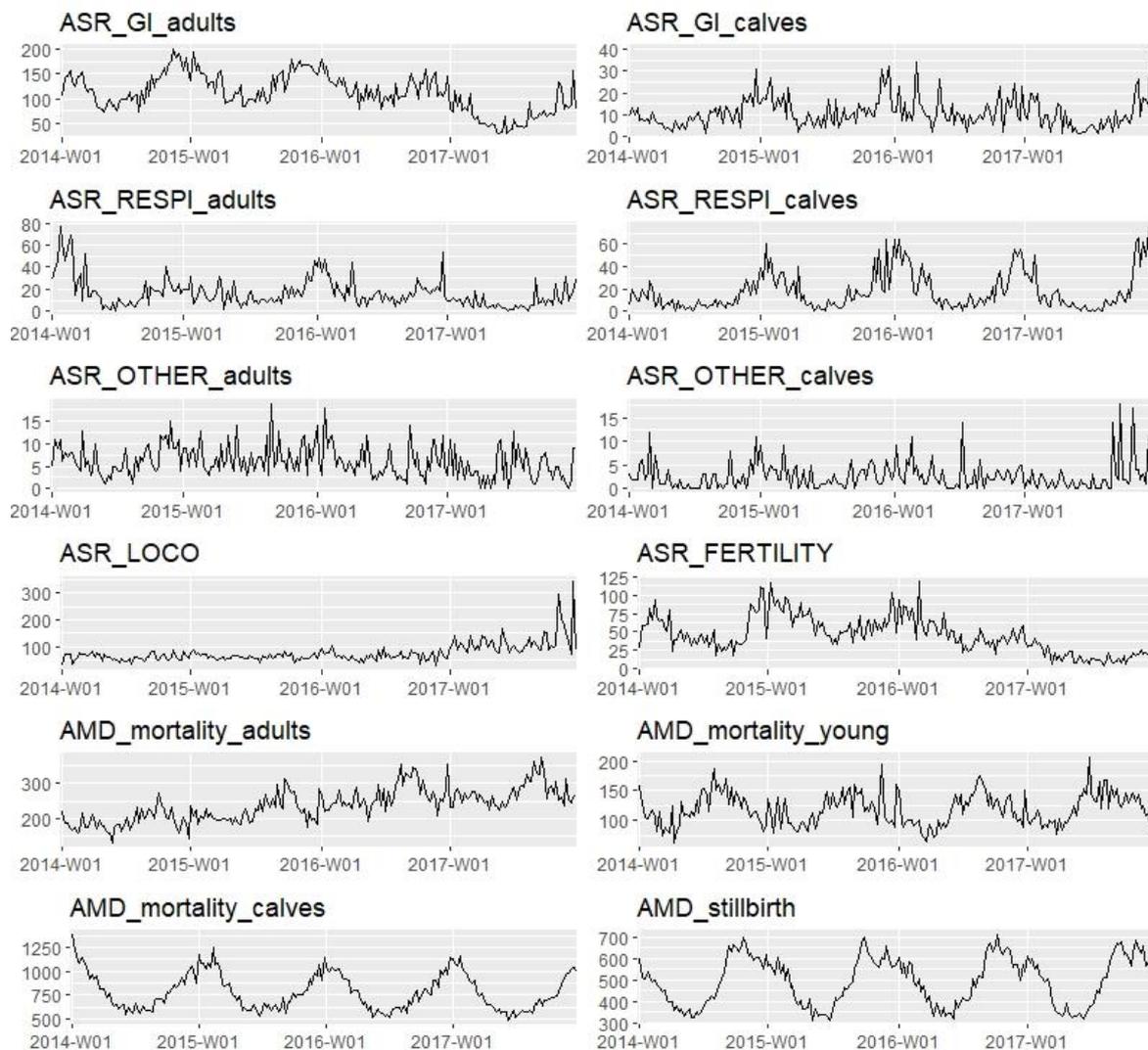
Aucune véritable épizootie de FCO, SV, RIB ou BVD n'était présente dans les données, car il n'y a pas eu d'épizootie de ces maladies en Suisse pendant la période étudiée. Afin d'évaluer les performances de détection des algorithmes, nous avons simulé des épizooties fondées sur une élicitation d'experts pour chacune de ces maladies. L'élicitation d'experts est une méthode acceptée pour estimer des informations lorsque d'autres données ne sont pas disponibles ou qu'il est difficile de les recueillir. Trois experts vétérinaires appartenant à la Clinique des animaux d'élevage de l'Université de Berne et à l'Office fédéral de la sécurité alimentaire et des affaires vétérinaires Suisse (OSAV) ont été interviewés. Les informations obtenues ont ensuite été utilisées pour simuler 300 épizooties pour chacune des quatre maladies incluses dans cette étude et pour chacune des 12 STs.

Le processus de simulation d'épizooties comportait quatre étapes consécutives (voir tableau 1). Pour chaque épizootie simulée, des tirages aléatoires ont

été effectués dans les distributions bêta PERT estimées *via* élicitation d'experts (voir étape 3).

**Figure 1**

**Séries temporelles extraites des deux bases de données nationales suisses AMD et ASR**



Dans certaines des STs, aucun cas épizootique n'a été ajouté parce que les experts consultés ne s'attendaient pas à ce que des cas cliniques apparaissent dans ces STs en cas d'épizootie. De plus, les augmentations du nombre de cas n'ont pas toujours été insérées au même moment dans toutes les STs parce que les experts ont indiqué qu'il pourrait y avoir un retard dans l'apparition de certains syndromes. Par exemple, certains symptômes n'apparaissent pas en tout début d'épizootie.

### 3. ANALYSE

On a supposé que les 12 STs de l'étude étaient conditionnellement indépendantes. Pour utiliser les cartes de contrôle, il est nécessaire d'avoir des STs sans saisonnalité ni tendance. Nous avons donc utilisé une méthode de différenciation d'une semaine pour éliminer les effets temporels présents dans les données brutes : nous avons calculé la différence entre deux observations consécutives. Les résidus ainsi différenciés ont ensuite été sauvegardés en tant que nouvelles STs. La distribution normale multivariée a ensuite été évaluée et confirmée à l'aide du test HenzeZirkler [Henze et Zirkler, 1990].

**Tableau 1**  
**Simulation d'épizooties multivariées**

Étapes	Paramètres	Estimations
<b>Étape 1. Estimer le nombre total de nouvelles infections pour chaque maladie à chaque pas de temps dans le cheptel bovin suisse.</b>	Forme de la courbe épizootique	linéaire pour les quatre maladies
	Durée de l'épizootie (t_max)	t_max= 36 semaines d'après l'élucitation d'experts
	Nombre d'animaux infectés à la fin de la période épizootique	400 animaux nouvellement infectés à t_max
<b>Étape 2. Estimer la proportion de nouvelles infections chez les veaux et les bovins adultes à chaque étape</b>	Nombre de veaux vs bovins adultes dans le cheptel	¾ du cheptel bovin était composé d'animaux adultes et que ¼ était composé de veaux
	Animaux affectés	les quatre maladies infecteraient également les bovins adultes et les veaux.
<b>Étape 3. Estimer la proportion de bovins et de veaux adultes infectés qui devraient présenter des signes cliniques.</b>	Pourcentage d'animaux infectés présentant le syndrome A en début d'épizootie	Distribution bêta PERT estimée via élucitation d'experts
	Pourcentage d'animaux infectés présentant le syndrome A en milieu d'épizootie	
	Pourcentage d'animaux infectés présentant le syndrome A en fin d'épizootie	
	Pourcentage d'animaux infectés présentant le syndrome B en début d'épizootie	
	<i>Etc.</i>	
<b>Étape 4. Animaux avec des signes cliniques reportés dans les séries temporelles</b>	Pourcentage d'animaux infectés présentant des symptômes reportés dans chaque ST	AMD = 100 % ASR = 50 %

Les deux algorithmes ont été mis en œuvre dans R x64 version 3.0.2 [R Development Core Team, 2008]. La distribution normale multivariée a été évaluée à l'aide du package R {MVN} [Korkmaz *et al.*, 2014]. La matrice de covariance et le vecteur moyen ont été estimés à partir des données historiques disponibles de 2013 à 2015, en utilisant la fonction 'mult.chart' du package R {MSQC} [Montgomery, 2009 ; Santos-Fernandez, 2013].

### 3.1 MEWMA DIRECTIONNELLE

MEWMA a été proposée par Lowry [Lowry *et al.*, 1992] et est une carte de contrôle Hotelling  $T^2$  appliquée aux statistiques de l'EWMA au lieu des données originales. La MEWMA est fondée sur les différences cumulatives entre les données observées dans une certaine fenêtre temporelle. Joner *et al.* [2008] ont proposé une version directionnelle de l'algorithme fondée sur l'équation :

$$\text{Si } t > 0 : Z_t = \max \{ 0, \lambda (X_t - \mu) + (1 - \lambda)Z_{t-1} \}$$

$$\text{Si } t = 0 : Z_t = 0$$

où  $\lambda$  est un paramètre de lissage ( $0 \leq \lambda \leq 1$ ) qui détermine le poids relatif des valeurs observées actuelles dans les différentes STs par rapport aux valeurs passées, et  $\mu$  est la moyenne cible du processus. Les valeurs observées dans les différentes STs à l'instant  $t$  sont représentées par le vecteur  $X_t$ .

$Z_t$  est donc le vecteur des moyennes pondérées des observations actuelles standardisées autour de zéro et a une matrice de covariance à l'instant  $t$  connue et égale à  $\sum_{Z_t} = \lambda(1-(1-\lambda)^{2t})/(2-\lambda) * \Sigma$ .  $\Sigma$  est la matrice de covariance de  $X_t$ .

Lorsque  $t \rightarrow \infty$ , alors  $\sum_{Z_\infty}$  est égale à  $\frac{\lambda}{2-\lambda} * \Sigma$ . L'inverse de  $\sum_{Z_\infty}$ ,  $\sum_{Z_\infty}^{-1}$ , correspond à la matrice de corrélation partielle des variables et est utilisée pour le calcul de la statistique du graphique MEWMA :

$$MEW_t = Z_t' \sum_{Z_\infty}^{-1} Z_t$$

Où  $Z_t'$  est le vecteur transposé de  $Z_t$ .

La procédure débute avec  $Z_0 = 0$ . Cinq valeurs de  $\lambda$  ont été évaluées : 0,1, 0,2, 0,3, 0,4 et 0,5.

### 3.2 MCUSUM DIRECTIONNELLE

Plusieurs versions de MCUSUM ont été proposées. Dans cette étude, nous avons choisi la méthode proposée par Crosier [Crosier, 1988] et adaptée à la sensibilité directionnelle par Fricker [Fricker, 2007]. Cette MCUSUM est fondée sur l'équation :

$$\text{Si } C_t > k : S_t = \max\{0, (S_{t-1} + X_t - \mu) * (1 - k/C_t)\}$$

$$\text{Si } C_t \leq k : S_t = 0$$

où  $k$  représente l'amplitude attendue de la distance entre la moyenne cible du processus et la moyenne réelle du processus,  $\mu$  est la moyenne cible du processus, et  $C_t = [(S_{t-1} + X_t - \mu)' \Sigma^{-1} (S_{t-1} + X_t - \mu)]^{1/2}$  avec  $\Sigma^{-1}$  l'inverse de la matrice de covariance du vecteur  $X_t$ .

La procédure commence avec  $S_0 = 0$  et est calculée séquentiellement. Cinq valeurs différentes de  $k$  ont été évaluées : 0,1, 0,2, 0,5, 0,7 et 0,8.

### 4. ÉVALUATION DE LA PERFORMANCE DE DETECTION DES ABERRATIONS

Chaque semaine d'observation a été classée comme « vrai positif » (VP) si une certaine limite supérieure de contrôle (UCL) a été dépassée pendant une semaine faisant partie d'une période épizootique simulée. Une semaine d'épizootie non détectée a été classée dans la catégorie des « faux négatifs » (FN). Chaque semaine appartenant à une période non épizootique a été considérée comme « vraiment négative » (VN) si aucune alerte n'était générée et comme « faux positif » (FP) si une alerte était générée. L'exactitude a été évaluée en utilisant : la sensibilité fondée sur le nombre d'épizooties détectées parmi toutes les épizooties insérées ( $Se_{out}$ ), la sensibilité fondée sur le nombre de semaines pendant lesquelles une alarme a été

déclenchée ( $Se_{wk}$ ), la spécificité ( $Sp$ ), la valeur prédictive positive (PPV) et la valeur prédictive négative (VNP).  $Se_{out}$  a été calculé pour l'ensemble des 300 épizooties simulées pour chacune des quatre maladies considérées dans cette étude. Les paramètres  $Se_{wk}$ ,  $Sp$ , PPV, NPV ont été calculés comme suit :

- $Se_{out} = \text{épizooties détectées} / \text{nombre total d'épizooties introduites}$
- $Se_{wk} = VP / (VP + FN)$
- $Sp = VN / (VN + FP)$
- $PPV = VP / (VP + FP)$
- $NPV = FN / (FN + FP)$

Pour la détection de la maladie, il n'est pas important que toutes les semaines d'une épizootie soient reconnues, mais il est crucial qu'une épizootie soit détectée au moins une fois et qu'elle le soit au début de l'épizootie. Par conséquent,  $Se_{out}$  a été considérée comme plus importante que  $Se_{wk}$ . La rapidité de détection d'une épizootie a été calculée comme étant le décalage (en semaines) entre le début de l'épizootie et la première alarme produite par le système de détection. Les valeurs moyennes (Tmoyen), médiane (Tmed), minimale et maximale (Tmin et Tmax) et l'écart-type (Tsd) du temps de détection ont été calculés en utilisant les 300 STs épizooties simulées de chaque maladie.

Un grand nombre de fausses alarmes positives deviendrait rapidement ingérable dans la pratique de la surveillance. Nous avons donc choisi de fixer les seuils d'alerte de telle sorte que le système de détection produise au maximum 5 % de fausses alarmes positives. Ces seuils d'alerte ont été considérés comme les seuils d'alerte optimaux pour chaque algorithme et ont été notés UCL5 %.

---

## III - RÉSULTATS

---

### 1. ÉLICATION D'EXPERTS ET SIMULATION D'ÉPIZOOTIES MULTIVARIÉES

Les experts ont estimé la proportion d'animaux malades présentant des signes cliniques ou syndromes individuels à différents stades d'une épizootie. Les syndromes respiratoires ont été estimés comme fréquents chez les veaux et les bovins adultes en cas d'épizootie de RIB. L'anorexie, la perte de poids et l'apathie étaient plus fréquentes

chez les animaux infectés par SV. Cependant, pour la plupart des autres signes cliniques, les différences entre les maladies n'étaient pas très importantes. C'est notamment le cas pour la mortalité qui a été estimée comme faible pour les quatre maladies. La diminution des performances de reproduction a été signalée pour les quatre maladies mais a été estimée comme plus importante en cas d'épizootie de BVD.

## 2. PERFORMANCES GLOBALES DE DETECTION ESTIMEE A PARTIR DE DONNEES SIMULEES

Les UCL5 % ont été calculées pour chaque algorithme et chaque ensemble de paramètres testés (c.-à-d.  $k$  pour MCUSUM et  $\lambda$  pour MEWMA) en considérant toutes les épizooties de maladies simulées. À chaque fois que les statistiques de MEWMA ou MCUSUM dépassaient l'UCL5 %, une alarme était produite. Pour un seuil d'alarme défini à l'UCL5 %, les deux algorithmes ont été capables de détecter correctement plus de 97 % des épizooties insérées avec des VPP variant entre 86,9 et 95,0. Se\_wk et VPN étaient globalement faibles et n'ont

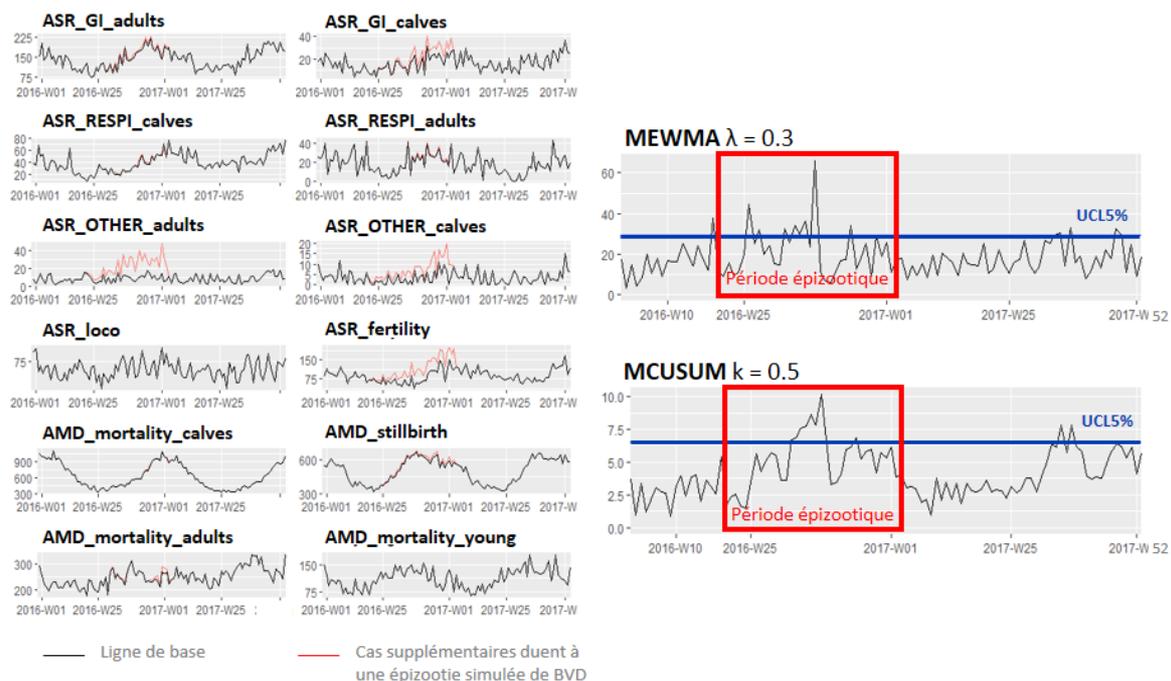
jamais dépassé respectivement 40 % et 62 %. Le temps moyen de détection (Tmean) varie beaucoup selon l'algorithme. La MEWMA a toujours obtenu des temps de détection plus courts que la MCUSUM, indépendamment des paramètres utilisés. Lorsque  $\lambda$  et  $k$  étaient respectivement égaux à 0,3 et 0,5, les temps de détection de MEWMA et MCUSUM étaient les plus courts. Notre objectif étant de favoriser la détection précoce, nous avons fixé les paramètres  $\lambda$  et  $k$  de MEWMA et MCUSUM respectivement à 0,3 et 0,5 pour la suite de l'analyse. Un exemple de résultat obtenu pour une épizootie simulée de BVD est présenté en figure 2.

Figure 2

### Exemple d'évolution des statistiques de MEMWA et MCUSUM pour une des épizooties simulées de BVD.

Les cas supplémentaires simulés et insérés dans les différentes séries temporelles pour cette épizootie simulée de BVD sont présentés à gauche de la figure.

Les statistiques de MEWMA et MCUSUM sont présentées à droite de la figure.



## 3. PERFORMANCES DE DETECTION POUR LES QUATRE MALADIES CONSIDEREES

Les performances de détection ont été calculées séparément pour les quatre maladies pour lesquelles des épizooties ont été simulées. Les résultats obtenus avec ces paramètres au seuil d'alarme UCL5 % précédemment défini sont présentés dans le tableau 2.

Les deux algorithmes ont détecté 100 % des épizooties simulées. Les épizooties de RIB ont été détectées plus tôt en moyenne que les épizooties des

autres maladies. Les épizooties de SV ont été détectées une à trois semaines plus tard que les épizooties de RIB. Les épizooties de FCO et de BVD étaient les plus difficiles à détecter.

L'algorithme MEWMA a toujours surpassé le MCUSUM, en particulier pour la rapidité de détection. Le délai moyen et le délai médian de détection (Tmean et Tmedian, respectivement) ainsi que l'écart-type (Tsd) et le délai maximal de détection (Tmax) étaient plus courts.

Tableau 2

**Performances de détection spécifiques obtenues avec MEWMA et MCUSUM pour les valeurs des paramètres minimisant le temps de détection et pour chaque maladie (c.-à-d. rhinotrachéite infectieuse bovine (RIB), virus de Schmallerberg (SV), virus de la fièvre catarrhale du mouton (FCO) et diarrhée virale bovine (BVD)).**

UCL5 % = limites supérieures de contrôle produisant 5 % de fausses alarmes positives,  
Tmin = temps minimal de détection, Tmax = temps maximal de détection, Tmean = temps moyen de détection,  
Tmedian = temps médian de détection, Tsd = écart type du temps de détection,  
Se\_wk = sensibilité hebdomadaire, Sp = spécificité, VPN = valeur prédictive négative,  
VPP = valeur prédictive positive

Algorithme	UCL5 %	Maladie	Tmin	Tmax	Tmean	Tmedian	Tsd	Se_wk	Sp	VPN	VPP
MCUSUM, K = 0,5	6,5	RIB	1	29	11,4	11	5,5	49,6	96,8	65,1	94,0
		SV	1	24	13,6	15	7,8	44,1	97,0	62,8	93,9
		FCO	1	35	16,3	18	9,3	22,6	96,8	55,7	87,5
		BVD	1	35	16,9	19	9,5	22,1	96,9	54,8	88,0
MEWMA, λ = 0,3	29,5	RIB	1	14	7,2	9	2,8	38,7	96,7	60,5	92,4
		SV	1	22	8,1	7	5,2	35,0	97,0	59,2	92,4
		FCO	1	28	9,0	8	6,1	19,7	96,6	54,7	85,2
		BVD	1	28	10,2	9	6,5	20,0	96,8	54,1	86,5

#### IV - DISCUSSION

À notre connaissance, il s'agit de la première évaluation de cartes de contrôle multivariées directionnelles en santé animale [Dorea et Vial, 2016]. La méthode était facile à mettre en œuvre et les résultats étaient faciles à interpréter parce que toutes les informations contenues dans les différents STs ont été combinées en un seul indicateur. Cependant, le fait de combiner les informations de tous les STs en un seul indicateur constitue également une limite de cette approche. Il n'a pas été possible de déterminer quelle ST a le plus contribué aux alarmes déclenchées. Il ne s'agit pas d'une limitation majeure, mais c'est une préoccupation pratique qui devrait être prise en compte lors de la mise en œuvre de cartes de contrôle multivariées sur le terrain.

Les cartes de contrôle multivariées évaluées dans notre étude ont détecté toutes les épizooties simulées de BVD, RIB, FCO et SV entre 2 et 3,6 mois après le début de l'épizootie et ce avec une spécificité de 95 %. Il est difficile de comparer les performances de détection obtenues avec celle du système de surveillance actuellement mis en place en Suisse car il existe peu de données ou d'informations sur les épizooties de ces maladies dans le pays. À l'exception de Schmallerberg, des systèmes de

surveillance active sont actuellement en place en Suisse. Des programmes visant à garantir l'absence de RIB et de BVD sont en vigueur. Une sélection d'élevages fondée sur les risques est effectuée chaque année pour la RIB. L'échantillonnage est effectué à la fois dans les fermes laitières (par échantillonnage du lait entre janvier et avril) et dans les fermes non laitières (les échantillons de sang sont prélevés entre janvier et mai). En 2008, un programme d'éradication de la BVD a été lancé en Suisse. Actuellement, la procédure d'échantillonnage (y compris la fréquence et le type d'échantillonnage) dépend du type d'exploitation bovine (laitière ou non laitière) et de la présence d'animaux infectés de façon persistante dans l'exploitation au cours des 36 derniers mois. Les experts consultés durant cette étude ont estimé à neuf mois (durée maximale des épizooties simulées dans notre étude) le temps maximal nécessaire, en moyenne, pour identifier une épizootie de RIB, BVD, FCO ou SV en Suisse avec les systèmes actuels de surveillance programmée. Les cartes de contrôle multivariées mises en œuvre dans notre étude présentent donc le potentiel de renforcer la surveillance de la détection précoce en Suisse en se fondant sur cette norme de comparaison. Cependant, notre étude est fondée sur la simulation et il serait

nécessaire d'obtenir plus d'informations avant de pouvoir tirer des conclusions définitives.

Les deux algorithmes ont obtenu la même sensibilité et la même spécificité globale, mais des performances différentes en termes de rapidité de détection. La MEWMA a toujours détecté les épizooties plus tôt que la MCUSUM. Ces résultats sont en contradiction avec ceux de Fricker [Fricker *et al.*, 2008] qui ont rapporté que les deux algorithmes avaient des performances très similaires. Ces différences peuvent s'expliquer par l'approche utilisée dans notre étude pour simuler des épizooties dans plusieurs STs à la fois. Fricker a simulé des épizooties multivariées très simples qui présentaient une augmentation linéaire du nombre de cas et qui avaient été insérées au même moment dans toutes les STs. Dans notre étude, les épizooties simulées étaient plus subtiles et probablement plus semblables aux épizooties réelles. Par exemple, dans certaines de nos STs, aucun cas épizootique n'a été ajouté parce que les experts consultés ne s'attendaient pas à ce que des cas cliniques apparaissent dans ces STs en cas d'épizootie. De plus, les augmentations du nombre de cas n'ont pas toujours été insérées au même moment dans toutes les STs parce que les experts ont indiqué qu'il pourrait y avoir un retard dans l'apparition de certains syndromes. Les cartes de contrôle multivariées sont connues pour être affectées par le problème dit d'inertie [Woodall et Mahmoud, 2005]. Ce problème découle du fait que ces cartes de contrôle accumulent de l'information au fil du temps et ont tendance à détecter les changements qui se produisent dans les données avec un certain retard. Ceci est particulièrement vrai lorsque seuls de petits changements se produisent dans les données. D'après la littérature, les cartes de contrôle multivariées testées dans cette étude seraient moins sévèrement affectées par le problème d'inertie que les autres cartes de contrôle multivariées [Joner *et al.*, 2008]. Nos résultats suggèrent que le MEWMA est plus robuste que le MCUSUM au problème de l'inertie étant donné que le temps nécessaire pour obtenir la première vraie alarme positive avec MCUSUM était presque le double de celui avec MEWMA. Fricker [Fricker *et al.*, 2008] a suggéré que la MEWMA devrait être choisie plutôt que le MCUSUM parce qu'il est plus facile de développer une appréciation intuitive de la valeur de  $\lambda$  que de  $k$ . Nous suggérons que cette approche devrait aussi être choisie parce qu'elle est moins sensible au problème de l'inertie, surtout lorsque des changements subtils surviennent dans les STs.

Des différences dans les performances de détection ont été observées entre les quatre maladies étudiées. Le délai de détection était plus court et la  $Se_{wk}$  était plus élevée pour les deux algorithmes pour les épizooties de RIB et de SV que pour les épizooties de BVD et de FCO. Nous nous attendions à voir des différences, car nos experts s'attendaient à ce que la FCO et la BVD produisent des signes cliniques plus subtils que les autres maladies.

Nous avons présenté dans cette étude une approche standardisée pour simuler des épizooties se manifestant de manière simultanée dans plusieurs STs plus proches de la réalité. Les simulations ont l'avantage de permettre une analyse de sensibilité complète et de tester plusieurs paramètres d'épizootie, ce qui est essentiel pour l'évaluation du système de surveillance. Cependant, l'utilisation d'épizooties simulées soulève toujours des questions quant à la validité et à la fiabilité des résultats. Dans notre étude, nous avons supposé que certains paramètres épizootiques étaient fixes (p. ex. le taux de sous-déclaration, la forme et la durée des épizooties ainsi que le pourcentage de veaux et de bovins adultes dans la population) et que le pourcentage d'animaux infectés présentant certains signes cliniques était estimé selon l'avis d'experts. Il n'a en effet pas été possible de trouver des informations standardisées dans la littérature. Par ailleurs, les informations disponibles sur les signes cliniques étaient rares. Pour ces raisons, nous avons abandonné la recherche documentaire au profit de l'éllicitation d'experts, méthode acceptée pour obtenir des informations lorsque les données ne sont pas disponibles ou sont difficiles à recueillir. Cependant, il est important de souligner que ce processus peut introduire un biais. L'information dont nous avons besoin est très spécialisée. Elle est liée non seulement aux maladies étudiées, mais aussi aux caractéristiques de chaque maladie au sein du cheptel bovin suisse et à la manière dont les données sont collectées en Suisse. Plus important encore, parce que ces paramètres étaient constants, les différences de performances de détection observées entre les maladies sont uniquement dues à des différences en termes de pourcentage d'animaux infectés présentant certains signes cliniques. Nous avons décidé de fixer ces paramètres par souci de simplicité, mais il serait facile de les modifier en utilisant notre approche de simulation d'épizootie. Dans de futures études, il serait intéressant d'étudier comment la modification de ces paramètres affecterait la performance de détection des algorithmes et surtout les différences observées entre les quatre maladies.

## BIBLIOGRAPHIE

- Brouwer H., Stegeman J.A., Straatsma J.W., Hooijer G.A., van Schaik G. - The validity of a monitoring system based on routinely collected dairy cattle health data relative to a standardized herd check. *Prev. Vet. Med.*, 2015, **122**(1-2), 76-82.
- Burkom Howard S., Ramac-Thomas L., Babin S., Holtry R., Mnatsakanyan Z., Yund C. - An integrated approach for fusion of environmental and human health data for disease surveillance. *Stat. Med.*, 2011, **30**(5), 470-479.
- Chatfield C. - The holt-winters forecasting procedure. *Applied Statistics*, 1978, **27**(3), 264-279.
- Colón-González F.J., Lake I.R., Morbey R.A., Elliot A.J., Pebody R., Smith G.E. - A methodological framework for the evaluation of syndromic surveillance systems: A case study of England. *BMC Public Health*, 2018, **18**(1), 1-14.
- Courtejoie N., Sailleau C., Zientara D.V.S., Gourmelen F.B.C., Milard H.A.C., Pagneux S.P.C., Zanella C.V.G. - Circulation of bluetongue virus 8 in French cattle, before and after the re-emergence in 2015, *Transbound. Emerg. Dis.*, 2018, **65**(1), 281-284.
- Crosier R.B. - Multivariate Generalizations of Cumulative Sum Quality-Control Schemes. *Technometrics*, 1988, **30**(3), 291-303.
- Doherr M.G., Audigé L. - Monitoring and surveillance for rare health-related events: a review from the veterinary perspective. *Series B, Biological Sciences*, 2001, **356**(1411), 1097-110.
- Dórea F.C., Lindberg A., McEwen B.J., Revie C.W., Sanchez J. - Syndromic surveillance using laboratory test requests: A practical guide informed by experience with two systems. *Prev. Vet. Med.*, 2014, **116**(3), 313-324.
- Dórea F., Vial F. - Animal health syndromic surveillance : a systematic literature review of the progress in the last 5 years (2011-2016). *Veterinary Medicine: Research and Reports*, 2016, **7**, 157-170.
- Faverjon C., Andersson M.G., Tapprest J., Sala C., Decors A., Triz P., Leblond A. - Evaluation of a multivariate syndromic surveillance system for West Nile virus. *Vector-Borne and Zoonotic Dis.*, 2015, **16**(6), 382-390.
- Fricker R.D., Knitt M.C., Hu C.X. - Comparing directionally sensitive MCUSUM and MEWMA procedures with application to biosurveillance. *Quality Engineering*, 2008, **20**(4), 478-494.
- Fricker R.D. - Directionally Sensitive Multivariate Statistical Process Control Methods. *Advances in Disease Surveillance*, 2007, **3**(1), 1-17.
- Henze N., Zirkler B. - A Class of Invariant Consistent Tests for Multivariate Normality. *Communication in Statistics- Theory and Methods*, 1990, **19**(10), 3595-3617.
- Hopkins R.S., Tong C.C., Burkom H.S., Akkina J.E., Berezowski J., Shigematsu M., Streichert L.C. - A Practitioner-Driven Research Agenda for Syndromic Surveillance. *Public Health Reports*, 2017, **132**, 116S-126S.
- Joner M.D., Woodall W.H., Reynolds M.R., F.R.D. - A One-sided MEWMA Chart for Health Surveillance. *Qual. Reliab. Eng. Int.*, 2008, **24**, 503-518.
- Kalekar P.S. - Time series forecasting using holt-winters exponential smoothing. *Kanwal Rekhi School of Information Technology*, 2004, **4329008**, 1-13.
- Korkmaz S., Goksuluk D., Zararsiz G. - MVN: An R Package for Assessing Multivariate Normality. *The R Journal*, 2014, **6**(2), 151-162.
- Lotze T., Shmueli G., Yahav I. - Simulating Multivariate Syndromic Time Series and Outbreak Signatures. *Working Paper, RHS-06-054, Robert H. Smith School, University of Maryland*, 2007.
- Lowry C.A., Woodall W.H., Champ C.W., Rigdon S.E., Lowry C.A. - A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics*, 1992, **34**(1), 46-53.
- Miekley B., Stamer E., Traulsen I., Krieter J. - Implementation of multivariate cumulative sum control charts in mastitis and lameness monitoring. *J. Dairy Sci.*, 2013, **96**(9), 5723-5733.
- Montgomery D. - Introduction to statistical quality control. *John Wiley et Sons Inc.*, 2009.
- R Development Core Team. - R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2008.
- Santos-Fernandez E. - Multivariate Statistical Quality Control Using R. Ed. Springer Briefs in Statistics. Universidad Central Marta, Villa Clara, Cuba, 2013, 140 pages.
- Siegrist D., Pavilin J.- Bio-ALERT biosurveillance detection algorithm evaluation. *Mmwr*, 2004,

53(Suppl.), 152-158.

Sonesson C., Frisé M. - Multivariate Surveillance.  
In A. B. Lawson et K. Kleinman (Eds.), *Spatial  
and Syndromic Surveillance for Public Health*.  
John Wiley et Sons, Ltd. 2005 (pages 153-166).

Struchen R., Reist M., Zinsstag J., Vial F. -  
Investigating the potential of reported cattle  
mortality data in Switzerland for syndromic  
surveillance. *Prev. Vet. Med.*, 2015, **121**(1), 1-7.

Tokars J.I., Burkom H., Xing J., English R., Bloom  
S., Cox K., Pavlin J.A. -Enhancing time-series  
detection algorithms for automated  
biosurveillance. *Emerg. Inf. Dis.*, 2009, **15**(4),  
533-539.

Vial F., Wei W., Held L. - Methodological

challenges to multivariate syndromic  
surveillance: a case study using Swiss animal  
health data. *BMC Vet. Res.*, 2016, **12**(1), 288.

Woodall W.H., Mahmoud M.A. -The Inertial  
Properties of quality control charts.  
*Technometrics*, 2005, **47**(4), 425-436.

Xing J., Burkom H., Tokars J. - Method selection  
and adaptation for distributed monitoring of  
infectious diseases for syndromic surveillance. *J.  
Biomed. Inf.*, 2011, **44**(6), 1093-1101.

Yahav I., Shmueli G. - Directionally sensitive  
multivariate control charts in practice:  
Application to biosurveillance. *Qual. Reliab.  
Eng. Int.*, 2014, **30**(2), 159-179.



### État des conflits d'intérêt

Aucun conflit d'intérêt n'est à signaler pour cette publication.