



HAL
open science

Evaluation des systèmes d'intelligence épidémiologique appliqués à la détection précoce des maladies infectieuses au niveau mondial.

Philippe Barboza

► **To cite this version:**

Philippe Barboza. Evaluation des systèmes d'intelligence épidémiologique appliqués à la détection précoce des maladies infectieuses au niveau mondial.. Santé publique et épidémiologie. Université Pierre et Marie Curie - Paris VI, 2014. Français. NNT : 2014PA066529 . tel-01133801

HAL Id: tel-01133801

<https://theses.hal.science/tel-01133801v1>

Submitted on 20 Mar 2015

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.

THESE DE DOCTORAT
DE L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité **Epidémiologie**

ECOLE DOCTORALE PIERRE LOUIS DE SANTE PUBLIQUE A PARIS : EPIDEMIOLOGIE ET SCIENCES DE
L'INFORMATION BIOMEDICALE

Présentée par

M. Philippe BARBOZA

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

**Evaluation des systèmes d'intelligence épidémiologique appliqués à la
détection précoce des maladies infectieuses au niveau mondial.**

Soutenue le 16 décembre 2014

devant le jury composé de :

M. le Pr Pascal ASTAGNEAU

Directeur de thèse

M. le Pr Loïc JOSSERAN

Rapporteur

M. Le Pr Jean-Baptiste MEYNARD

Rapporteur

M. Le Pr Gilles BRÜCKER

Examineur

M. le Pr Pierre-Yves BOËLLE

Examineur

M. le Pr Philippe GUERIN

Examineur

A mes Parents

A ma sœur et mon frère

Remerciements

A mon Directeur de thèse, le Pr Pascal Astagneau, pour m'avoir conduit à travers cet exercice long et périlleux.

Aux membres de mon jury, qui ont accepté de juger mon travail.

A Laetitia, Guy, Thierry et Marc qui ont intrigué pour le faire sauter le pas, tout vient à point à qui sait attendre.

A Laetitia pour son aide précieuse durant tout le processus et sa relecture du manuscrit.

A tous les autres qui m'ont soutenu et encouragé, la liste est trop longue mais ils se reconnaîtront

Table des matières

RESUME DE LA THESE	1
ENGLISH SUMMARY	3
ARTICLES	5
Article - 1	6
Article - 2	15
ACRONYMES	24
I. INTRODUCTION	25
I.1 Contexte	25
I.2 Limite de la surveillance épidémiologique conventionnelle	26
I.3 Sources alternatives d'information	26
I.4 Intelligence Epidémiologique	28
I.4.1 Terminologie.....	28
I.4.2 Principes.....	28
I.4.3 Concepts et définitions	28
I.4.4 Processus de l'intelligence épidémiologique.....	30
I.4.5 Sources d'information de la surveillance électronique	32
I.4.6 Les étapes du processus de l'intelligence épidémiologique	34
I.5 Hypothèse.....	36
I.6 Objectifs	37
I.6.1 Objectif Principal	37
I.6.2 Objectifs spécifiques.....	37
II. MATERIEL ET METHODE	38
II.1 Organisation.....	38
II.2 Fonctionnement des systèmes de biosurveillance (Article A).....	38
II.2.1 Définition	38
II.2.2 Principes	38
II.2.3 Le recueil des informations	39
II.2.4 Traitement des informations	39
II.2.5 Ontologie	39
II.2.6 Extraction des évènements	40
II.2.7 Le Dé-doublonnage	40
II.2.8 Analyse.....	40
II.2.9 Dissémination de l'information.....	40
II.3 Choix des systèmes de biosurveillance	41
II.3.1 Le Projet EAR.....	42
II.3.2 Caractéristiques des systèmes de biosurveillance.....	42
II.4 Les indicateurs	43
II.4.1 Revue de la Littérature	43

II.4.2	Indicateurs de la surveillance des maladies transmissibles	44
II.4.3	Indicateurs de résultats	44
II.5	Stratégies d'intelligence épidémiologique.....	47
II.5.1	Diversité des stratégies	47
II.5.2	Impact des stratégies d'intelligence épidémiologique	47
II.6	Evaluation systèmes de biosurveillance (article 1).....	48
II.7	facteurs détection évènements infectieux (article 2).....	48
II.8	réactivité détection et sensibilité systèmes (non-publiées).....	49
III.	RESULTATS.....	50
III.1	Stratégies d'intelligence épidémiologique et impacts	50
III.1.1	Diversité des stratégies	50
III.1.2	Impact de la stratégies et des objectifs.....	53
III.2	Evaluation des systèmes de biosurveillance (article 1).....	59
III.3	Facteurs de détection des évènements infectieux (Article 2).....	60
III.4	réactivité détection et sensibilité systèmes (non-publiées).....	60
IV.	DISCUSSION.....	64
IV.1	Résultats.....	64
IV.1.1	Synthèse des résultats	64
IV.1.2	Sensibilités.....	65
IV.1.3	Valeurs prédictives positives.....	66
IV.1.4	Facteurs influençant la détection :	66
IV.1.5	Réactivité de détection	68
IV.1.6	Complémentarité des systèmes	69
IV.1.7	Rôle de l'intelligence épidémiologique	69
IV.1.8	Délais survenue de l'évènement et de la détection.....	69
IV.2	Avantages de la méthode : une démarche progressive.....	71
IV.3	Application pratique: Le projet EAR.....	71
IV.4	Limites de l'étude	72
IV.4.1	Gold-Standard	72
IV.4.2	Risques sanitaires et champ géographique.....	74
IV.5	Limites de l'utilisation des sources informelles	74
IV.5.1	Réseau sociaux et surveillance participative	74
IV.5.2	Accès à l'information brute.....	75
V.	CONCLUSION ET PERSPECTIVES.....	77
ANNEXE 1 – ARTICLE A.....	80	
ANNEXE 2 – ARTICLE B.....	88	
ANNEXE 3 - PRINCIPAUX SYSTEMES DE BIOSURVEILLANCE	98	
ANNEXE 4 - LA VEILLE INTERNATIONALE A L'INVS.....	106	
BIBLIOGRAPHIE.....	110	

Tables des illustrations

Figures

Figure 1 – Concepts surveillance des indicateurs, surveillance des évènements et intelligence épidémiologique (adapté de l’OMS ²¹)	30
Figure 2 – Etapes du processus l’intelligence épidémiologique	31
Figure 3 – Volume mensuel d’information échangé sur Internet 1990-2010*	35
Figure 4 – Evènements rapportés par l’OMS en 2012, Disease Outbreak News	56
Figure 5 – Evènements rapportés par EpiSouth dans l’e-WEB en 2012	56
Figure 6 – Evènements rapportés par l’ECDC dans le CDTR en 2012.....	57
Figure 7 – Evènements rapportés par l’InVS dans le BHI en 2012.....	57
Figure 8 – Délais entre la détection d’un signal par un système et sa détection par la Veille Internationale de l’InVS	62
Figure 9 – Répartition des utilisateurs d’Internet rapporté à 100 habitants, 2012 (données brutes CIA)	76
Figure 10 – La liberté de la presse dans le monde, source Reporters sans Frontières ⁹¹	76
Figure 11 – Portail d’Argus	98
Figure 12 – Exemple de rapport publié par Argus	98
Figure 13 – Portail BioCaster	99
Figure 14 – Exemple de rapport publié BioCaster.....	99
Figure 15 – Portail Gphin	100
Figure 16 – Exemple de rapport publié par Gphin	100
Figure 17 – Portail HealthMap	101
Figure 18 – Exemple de rapport HealthMap.....	101
Figure 19 – Portail MedISys.....	102
Figure 20 – MedISys, sélection par pathologie.....	102
Figure 21 – Portail ProMED-mail	103
Figure 22 – Exemple de rapport ProMED	103
Figure 23 – PULS : exemple de traitement des informations de MedISys par PULS	104
Figure 24 – Interactions simplifiées entre systèmes de biosurveillance (Source EAR) ⁵¹	105
Figure 25 – Zones prioritaires définies pour la Veille Internationale de l’InVS.....	109

Tableaux

Tableau 1 – Exemple de réseaux de surveillance et d’alerte potentiellement considérés comme des sources officielles dans le cadre de l’intelligence épidémiologique	33
Tableau 2 – Exemples de sources informelles.....	34
Tableau 3 – Systèmes de biosurveillance en 2010	41
Tableau 4 - Participants au projet EAR.....	42
Tableau 5 - Principales caractéristiques des systèmes de biosurveillance inclus dans l’étude, bilan réalisé en 2010	43
Tableau 6 - Principaux indicateurs d’évaluation des systèmes de surveillance des maladies infectieuses et leur application à l’évaluation de l’intelligence épidémiologie et des systèmes de biosurveillance.	46
Tableau 7 – Définitions des indicateurs et des mesures de réactivité	49
Tableau 8 - Domaines couverts en routine par les unités d’intelligence épidémiologique des institutions participant au projet EAR en 2010 (n=10)	51
Tableau 9 – Utilisation des systèmes de biosurveillance par les utilisateurs, Projet EAR 2010 (n=10).....	51
Tableau 10 – Processus d’intelligence épidémiologique, institutions participants au projet EAR en 2010 (n =10)	52
Tableau 11 – Caractéristiques des activités d’intelligence épidémiologique OMS, ECDC, EpiSouth et InVS.	53
Tableau 12 – Caractéristiques des supports de rétro-information élaborés dans le cadre de l’intelligence épidémiologique par l’OMS, l’ECDC, EpiSouth et l’InVS en 2012.	54
Tableau 13 – Pathologies rapportées dans les supports de rétro information l’intelligence épidémiologique dans par l’OMS, l’ECDC, EpiSouth et l’InVS en 2012.	55
Tableau 14 – Sensibilité des systèmes de biosurveillance vis-à-vis de l’équipe de la Veille internationale de l’InVS (n =73).....	61
Tableau 15 – Délais entre la détection d’un signal par un système et sa détection par la Vielle Internationale de l’InVS, sa publication et sa date de survenue (n = 73 évènements).	61
Tableau 16 – Français inscrits au registre mondial au 31/12/2011 (source : MAE)	106
Tableau 17 – Immigrés par pays de naissance, recensement 2009 (source : adapté de INSEE).....	106
Tableau 18 – Evolution simplifiée de l’utilisation des systèmes experts dans le cadre de la veille internationale effectuée par le département international de l’InVS.....	107
Tableau 19 – Critères de sélection des signaux dans le cadre de la Veille internationale mise en œuvre par le Département International de l’InVS.....	108
Tableau 20 – Exemples d’utilisation des critères	108

RESUME DE LA THESE

En dépit des avancées technologiques et scientifiques, le risque posé par les maladies émergentes demeure majeur. Bien qu'efficaces, les systèmes d'alerte précoce traditionnels, basés sur la détection et la notification par des professionnels ou des structures de santé, ont montré leurs limites. Les données produites par ces systèmes, répondent à des objectifs, des procédures et des contraintes spécifiques et ne sont pas toujours transmises dans un délai compatible avec la mise en œuvre rapide des mesures de contrôle adéquates.

A partir des années 1990, le développement exponentiel d'Internet a permis l'accès à des sources alternatives d'information qui ne sont pas soumis à ces contraintes : les sources informelles. L'intelligence épidémiologique et la surveillance des événements ont été mises en œuvre par les institutions de santé publique pour permettre l'exploitation de cette nouvelle source de données. En parallèle, des outils appelés systèmes de biosurveillance, ont été développés pour permettre le traitement de ce volume considérable d'information.

Si, la surveillance épidémiologique traditionnelle est largement décrite dans la littérature scientifique, l'intelligence épidémiologique et la surveillance des événements sont des concepts relativement nouveaux pour lesquels il n'existe pas de définition standardisée ni de procédures d'évaluation. L'évaluation combinée des performances des principaux systèmes de biosurveillance comparées à un même standard n'avait jamais été réalisée.

Notre objectif était donc d'évaluer, les performances des systèmes de biosurveillance utilisés pour la détection précoce des risques sanitaires dans le cadre de l'intelligence épidémiologique, et ce du point de vue des utilisateurs.

Tout d'abord, nous avons pu définir une méthodologie adaptée à l'évaluation des systèmes de biosurveillance. Afin d'identifier les indicateurs les plus pertinents du point de vue des utilisateurs, nous avons réalisé un état des lieux des stratégies d'intelligence épidémiologique utilisées par les principaux acteurs internationaux. Nous avons ainsi documenté la diversité des objectifs, des stratégies et des moyens dédiés à d'intelligence épidémiologique, ainsi que leur influence sur le nombre et la nature des événements détectés et rapportés par ces institutions.

Nous avons documenté, du point de vue des utilisateurs, les performances de sept des principaux systèmes de biosurveillance. Tout d'abord une évaluation qualitative des performances de ces sept systèmes de biosurveillance a été menée auprès des institutions participant au projet Early Alerting and Reporting (EAR). Si, la réactivité et la sensibilité ont été notées avec des scores élevés, des faiblesses en termes de représentativité, d'exhaustivité et de flexibilité ont été rapportées.

Ces résultats ont été corroborés par l'analyse quantitative effectuée sur des signaux potentiellement liés à des épisodes d'influenza A/H5N1. Le taux de détection des systèmes individuels variait de 31% à 38%, alors que les valeurs prédictives positives variaient de 3% à 24%. La sensibilité des systèmes variait de 38% à 72% mais elle augmentait à 93% lorsque le système virtuel combiné (intégrant le meilleur résultat des sept systèmes) était considéré, illustrant clairement les complémentarités entre les systèmes. Le délai moyen entre la détection des épisodes A/H5N1 par les systèmes et leur notification officielle par l'OMS et l'OIE était de 10,2 jours (IC 95%: 6,7 - 13,8) confirmant la valeur ajoutée de l'utilisation des sources informelles pour la détection précoce des alertes sanitaires et les performances des systèmes de biosurveillance dans ce domaine.

Une seconde étude a permis d'explorer les facteurs influençant la détection. Si aucune différence significative n'a été observée en matière de détection brute, la sensibilité intrinsèque des systèmes variait de 43 % à 71 % ($p = 0,001$). De plus, après ajustement pour les autres variables, la sensibilité intrinsèque était significativement associée : aux systèmes, au type de modération, à la langue, à la région d'occurrence et au type de maladie infectieuse. Si tous les systèmes ont montré des performances de détection similaire après ajustement sur les autres variables, les systèmes intégrant une modération humaine ont démontré une sensibilité intrinsèque 53% supérieure à celle des systèmes automatisés ($p < 0,001$). Cette étude a confirmé que les capacités de détection des maladies infectieuses des systèmes sont synergiques.

En dépit des limites, nos travaux ont permis de documenter la complémentarité des systèmes de biosurveillance et d'identifier la meilleure stratégie pour permettre une optimisation de leur utilisation : l'intégration de ces systèmes dans un outil unique. Au-delà de la démarche académique, nos travaux ont déjà eu des implications pratiques, puisqu'un tel outil est en cours de développement.

Au-delà de l'utilisation des outils, c'est bien le concept même de l'intelligence épidémiologique qui a pu être documenté par nos travaux. Ils ont permis de démontrer l'intérêt de l'utilisation des sources informelles pour la détection précoce des événements sanitaires d'un point de vue pratique en prenant en compte le point de vue des utilisateurs (les institutions de santé publique). Nos travaux ouvrent de nouveaux champs d'investigations dont les applications pourraient être importantes pour les utilisateurs comme pour les systèmes.

Les maladies infectieuses, les questions environnementales, et les menaces bioterroristes potentielles continueront de poser des menaces considérables pour la sécurité sanitaire mondiale. Dans ce contexte, les systèmes de détection et d'alerte précoce sont essentiels. Pour les renforcer, l'expertise des institutions de santé publique et des systèmes de biosurveillance devrait être mise en commun, en optimisant les ressources disponibles afin de renforcer la sécurité sanitaire mondiale.

Mots-clés : Intelligence épidémiologique; surveillance des événements; évaluation; détection des risques; détection des menaces; alerte précoce; détection précoce.

ENGLISH SUMMARY

Despite technological and scientific advances, the risk posed by emerging diseases remains major. Although effective, traditional early warning systems based on the detection and reporting by professionals or health facilities have shown their limits. The data produced by these systems comply with specific objectives, procedures and constraints and are not always transmitted in a time frame compatible with the rapid implementation of appropriate control measures.

From the 1990s, the exponential growth of the Internet has enabled access to alternative sources of information that are not affected by these constraints: the informal sources. The epidemic intelligence and the event-based surveillance have been developed by public health institutions in order to utilise this new data source. In parallel tools i.e., the biosurveillance systems, have been developed to enable the processing of this large volume of information.

While traditional epidemiological surveillance is widely described in the scientific literature, epidemic intelligence and event-based surveillance are relatively new concepts for which there is neither standardised definition nor evaluation procedures. The combined evaluation of performance of the main biosurveillance systems compared to the same standard had never been performed.

Our objective was to evaluate the performance of biosurveillance systems used for the early detection of health risks in the context of epidemic intelligence, and from the point of view of users.

First, we were able to define a suitable methodology for the evaluation of the biosurveillance systems. To identify the most relevant indicators from the point of view of users, we conducted an inventory of epidemic intelligence strategies used by the main international actors. We documented the diversity of objectives, strategies and resources dedicated to epidemiological intelligence and their influence on the number and the nature of events detected and reported by these institutions.

Our work allowed documenting the performance of seven key biosurveillance systems. First, a qualitative assessment of the performance of these seven biosurveillance systems was conducted among institutions participating to the Early Alerting and Reporting project (EAR). If the timeliness and sensitivity were rated highly, weaknesses in terms of representativeness, completeness, and flexibility were reported.

These results were corroborated by a quantitative analysis performed on signals potentially related to episodes of influenza A/H5N1. The detection rate of the individual systems ranged from 31% to 38%, while the positive predictive values ranged from 3% to 24%. System sensitivity ranged from 38% to 72% but increased to 93% when the combined virtual system (integrating the best result of the seven systems) was considered, clearly illustrating the systems' complementarity. The average time between the detection of A/H5N1 episodes by systems and their official notification by WHO or OIE was 10.2 days (95% CI: 6.7 to 13.8), confirming the added value of informal sources for the early detection of health threats and the performance of the biosurveillance systems in this domain.

A second study explored the factors influencing detection. If no significant difference was observed in terms of crude detection, the intrinsic sensitivity of the systems varied significantly from 43% to 71% ($p = 0.001$). After adjustment for other variables, the intrinsic sensitivity was significantly associated with: the individual systems, the type of moderation, the language, the region of occurrence and the type of infectious disease. If, after adjustment for other variables, all systems showed similar detection performance, systems incorporating human moderation demonstrated a 53% higher ($p < 0.001$) intrinsic sensitivity than automated systems. This study shows that overall, the systems' ability to detect infectious disease are synergistic.

Despite the limitations, our work allowed documenting the complementary of these biosurveillance systems and identifying the best strategy to enable optimisation of their use: the integration of these systems into a single tool. Beyond the academic approach, our work already has practical implications as such a tool is currently under development.

Beyond the simple use of tools, it is the very concept of epidemiological intelligence that was documented by our work. It highlighted the pertinence of using informal sources for the early detection of health events from a practical point of view taking into account the views of users (i.e. public health institutions). Our work opens new fields of investigation whose applications could be important for both users and systems.

Infectious diseases, environmental issues, and potential bioterrorism threats will continue to pose significant risks to global health security. In this context, detection systems and early warning are determinant. To reinforce them, the expertise of public health institutions and biosurveillance systems should be pooled while optimising available resources in order to strengthen the global health security.

Key words : epidemic intelligence; event-based surveillance; evaluation; risk detection; early warning; early detection

ARTICLES

Articles entrant dans le cadre de la thèse

Article 1 - Barboza P, Vaillant L, Mawudeku A, Nelson NP, Hartley DM, Madoff LC, Linge JP, Collier N, Brownstein JS, Yangarber R, Astagneau P; *Evaluation of Epidemic Intelligence Systems Integrated in the Early Alerting and Reporting Project for the Detection of A/H5N1 Influenza Events*. PLoS ONE. 2013; 8(3): e57252. doi:10.1371/journal.pone.0057252

Article 2 - Barboza P, Vaillant L, Le Strat Y, Hartley DM, Nelson NP, Mawudeku A, Madoff LC, Linge JP, Collier N, Brownstein JS, Astagneau P. *Factors Influencing Performance of Internet-Based Biosurveillance Systems Used in Epidemic Intelligence for Early Detection of Infectious Diseases Outbreaks*. PLoS ONE 2014 ; 9(3): e90536. doi:10.1371/journal.pone.0090536

Articles en relation avec la thèse (en annexe)

Article A - Hartley DM, Nelson NP, Arthur RR, Barboza P, Collier N, Lightfoot N, Linge JP, van der Goot E, Mawudeku A, Madoff LC, Vaillant L, Walters R, Yangarber R, Mantero J, Corley CD, Brownstein JS. *An overview of Internet biosurveillance*. Clin Microbiol Infect. 2013 Nov;19(11):1006-13. doi: 10.1111/1469-0691.12273.

Article B - Vaillant L, Collier N, Nelson NP, Hartley DM, Doherty B, Linge JP, Schnitzler J, Madoff LC, Yangarber R, Mawudeku A, Brownstein JS, Barboza P. *A review of the evaluation processes and performance measures for epidemic intelligence systems mining of public health information from the Web*. En cours de soumission

Autre publication en relation avec la thèse

Organisation Mondiale de la Santé (OMS). *Early detection, assessment and response to acute public health events: Implementation of Early Warning and Response with a focus on Event-Based surveillance* (2014). WHO/HSE/GCR/LYO/2014.4

ARTICLE - 1

OPEN ACCESS Freely available online

PLOS ONE

Evaluation of Epidemic Intelligence Systems Integrated in the Early Alerting and Reporting Project for the Detection of A/H5N1 Influenza Events

Philippe Barboza^{1*}, Laetitia Vaillant¹, Abba Mawudeku², Noele P. Nelson³, David M. Hartley⁴, Lawrence C. Madoff⁵, Jens P. Linge⁶, Nigel Collier⁷, John S. Brownstein⁸, Roman Yangarber⁹, Pascal Astagneau¹⁰, on behalf of the Early Alerting, Reporting Project of the Global Health Security Initiative[†]

1 International Department, French Institute for Public Health Surveillance (Institut de Veille Sanitaire), Saint Maurice, France, **2** Situational Awareness Section, Public Health Agency of Canada, Ottawa, Canada, **3** Department of Pediatrics, Georgetown University Medical Center, Washington, D.C., United States of America, **4** Department of Microbiology and Immunology, Georgetown University Medical Center, Washington, D.C., United States of America, **5** ProMED-mail, International Society for Infectious Diseases, Boston, Massachusetts, United States of America, **6** Joint Research Centre of the European Commission, Ispra, Italy, **7** National Institute of Informatics, Tokyo, Japan, **8** Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **9** Department of Computer Science, University of Helsinki, Helsinki, Finland, **10** Department of Public Health, Pierre et Marie Curie University School of Medicine, Paris, France

Abstract

The objective of Web-based expert epidemic intelligence systems is to detect health threats. The Global Health Security Initiative (GHSI) Early Alerting and Reporting (EAR) project was launched to assess the feasibility and opportunity for pooling epidemic intelligence data from seven expert systems. EAR participants completed a qualitative survey to document epidemic intelligence strategies and to assess perceptions regarding the systems performance. Timeliness and sensitivity were rated highly illustrating the value of the systems for epidemic intelligence. Weaknesses identified included representativeness, completeness and flexibility. These findings were corroborated by the quantitative analysis performed on signals potentially related to influenza A/H5N1 events occurring in March 2010. For the six systems for which this information was available, the detection rate ranged from 31% to 38%, and increased to 72% when considering the virtual combined system. The effective positive predictive values ranged from 3% to 24% and F1-scores ranged from 6% to 27%. System sensitivity ranged from 38% to 72%. An average difference of 23% was observed between the sensitivities calculated for human cases and epizootics, underlining the difficulties in developing an efficient algorithm for a single pathology. However, the sensitivity increased to 93% when the virtual combined system was considered, clearly illustrating complementarities between individual systems. The average delay between the detection of A/H5N1 events by the systems and their official reporting by WHO or OIE was 10.2 days (95% CI: 6.7–13.8). This work illustrates the diversity in implemented epidemic intelligence activities, differences in system's designs, and the potential added values and opportunities for synergy between systems, between users and between systems and users.

Citation: Barboza P, Vaillant L, Mawudeku A, Nelson NP, Hartley DM, et al. (2013) Evaluation of Epidemic Intelligence Systems Integrated in the Early Alerting and Reporting Project for the Detection of A/H5N1 Influenza Events. PLoS ONE 8(3): e57252. doi:10.1371/journal.pone.0057252

Editor: Hiroshi Nishiura, University of Hong Kong, Hong Kong

Received: July 27, 2012; **Accepted:** January 18, 2013; **Published:** March 5, 2013

Copyright: © 2013 Barboza et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: p.barboza@invs.sante.fr

† Membership of the Early Alerting and Reporting Project of the Global Health Security Initiative is provided in the Acknowledgments.

Introduction

Epidemic intelligence provides a new approach to address the challenges of disease globalization [1–3]. It provides an approach that is complementary to countries' national surveillance strategies. Moreover epidemic intelligence was included by the World Health Organization (WHO) in the health threat detection mechanisms integrated into the International Health Regulations [4,5]. While epidemiological indicator-based surveillance relies on regular reporting of a number of well-defined indicators provided mainly by health care facilities, epidemic intelligence focuses on event detection, prior to official health care reporting, laboratory confirmation and eventual official notification. Epidemic intelligence consists of the *ad hoc* detection and interpretation of

unstructured information available in the Internet. This information is very diverse in nature and is generated by multiple types of sources, both official and informal. The information may include unverified rumors from the media or more reliable information from official sources or traditional epidemiological surveillance systems. These raw signals usually contain very little information (e.g. medical or scientific) on which analysis can be performed, and they are often embedded in noise. Epidemic intelligence is a complex, time and resource-intensive process that includes a formalized protocol for event selection, verification of the genuineness of reported events, searches of complementary reliable information, analysis and communication.

Epidemic intelligence is still a relatively new discipline that emerged in the 1990s triggered by the development of the Internet. ProMED-mail [6] was the first Internet-based reporting system to use both formal and informal sources. It was followed by several expert systems developed to detect relevant information from the Internet [7–10]. In parallel, national and international institutions have developed epidemic intelligence capacities to fulfill their own needs [8,11–14]. A number of studies [15–18] have been carried out to assess expert systems' abilities to detect and correctly classify health threats using informal open sources or to present innovative functionalities. These papers rarely address users' viewpoints (i.e., the detection of relevant information by public health institutions). A thorough evaluation of epidemic intelligence information faces major challenges, including the lack of an adequate gold standard and standardized indicators and, but also the type of information collected, which is often not designed for health surveillance purposes.

The development of expert systems and epidemic intelligence took place independently, resulting in both the development of specific expertise among expert systems and institutions and, varying degrees of duplication. This paper aims to present a methodology and results that can be utilized to assess the complementarity of expert systems' capability and epidemic intelligence frameworks.

Methods

The EAR Project

The Global Health Security Initiative (GHSI) is an informal, international partnership among like-minded countries aiming to strengthen global health preparedness and response to chemical, biological, radio-nuclear (CBRN) terrorism and pandemic influenza threats. GHSI was launched in November 2001 by Canada, the European Union (EU), France, Germany, Italy, Japan, Mexico, the United Kingdom and the United States. The WHO serves as an expert advisor to the GHSI [19]. In 2009, an international project called Early Alerting and Reporting (EAR) was established, bringing together end-users (i.e., public health institutions in charge of epidemic intelligence), systems providers, and stakeholders (see Tables 1 and 2). Its objective for 2009–2010 was to assess the feasibility of developing a single web-based platform that would enable partners to access health threats identified from open source web-based public health intelligence systems, as well as to combine risk assessment processes.

Evaluation

The study included a qualitative (questionnaire-based) and a quantitative assessment. The qualitative assessment's goal was to

provide information essential for determining the best strategy for the quantitative part of the study.

Qualitative analysis

A questionnaire was constructed to assess both the type of epidemic intelligence performed by participating public health institutions and their perception of the seven integrated expert systems. The questionnaire, sent to ten EAR points of contact, was self-administrated during the first quarter of 2010. In order to measure the perceived performances of each system that they at least occasionally utilize, users were asked to rate each system through a simple choice (Yes/No). The following pre-defined characteristics were measured: representativeness (of information e.g., geographic coverage, type of diseases, etc.), completeness (or "exhaustivity" of information collected for the detected events), timeliness (of reporting), sensitivity (of the system), usefulness (of provided information), simplicity of use and flexibility of the systems (adaptation to users needs). Results were represented in two-dimensional spider-charts of the proportion of users that selected the corresponding attribute to describe the system (denominator being the number of systems' users). Participants were also requested to describe epidemic intelligence activities implemented in their institution in terms of priority domains of interest (e.g., specific epidemic diseases, plant diseases, CBRN agents, and natural disasters), the selection criteria, and frequency of these activities.

Gold standard

A review was performed to identify a global health threat covered by all participating expert systems, which was frequently occurring and for which a "gold standard" exists. Highly pathogenic avian influenza A/H5N1 (A/H5N1) was selected as the most appropriate. Human and veterinary cases are subjected to mandatory notification through WHO and the World Organisation for Animal Health (OIE). These two international organizations constitute the most reliable and recognized source of information regarding A/H5N1 biologically confirmed human cases and epizootics. Human cases and epizootics (both single cases of animal disease and larger outbreaks) that occurred in March 2010 (date of first symptoms, date of the start of the outbreak) or reported by WHO/OIE in March 2010 were considered as the gold standard.

Quantitative analysis: Database and indicator analyses

Raw data. Despite intrinsic differences, expert systems operate in similar ways. They search the Internet to detect information potentially relevant for epidemic intelligence purposes.

Table 1. Early Alerting and Reporting (EAR), participating systems.

	System name	System owner/developer	Country	Moderation type	n users 2010*	references
Expert systems	Argus	Georgetown University	USA	Human moderation	5	[29–31]
	BioCaster	National Institute of Informatics	Japan	Fully automated	4	[8,16,32]
	GPHIN	Public Health Agency of Canada	Canada	Human moderation	6	[1,9,33]
	HealthMap	Harvard University	USA	Partially moderated	5	[7,34,35]
	MedISys	Joint Research Centre	EU	Fully automated	5	[10,36,37]
	ProMED-mail	International Society of Infectious Diseases	USA	Human moderation	9	[6,17,23]
	Puls	University of Helsinki	Finland	Fully automated	4	[18,38,39]

doi:10.1371/journal.pone.0057252.t001

Table 2. Early Alerting and Reporting (EAR) public health institutions and stakeholders.

	Institution name	Country
Public Health Institutions	Centers for Disease Control and Prevention (CDC)	United States (USA)
	European Centre for Disease Prevention and Control (ECDC)	European Union (EU)
	Health Protection Agency (HPA)	United Kingdom
	Institut de Veille Sanitaire (InVS)	France
	Istituto Superiore di Sanità (ISS)	Italy
	National Institute of Infectious Diseases (NIID)	Japan
	Public Health Agency of Canada (PHAC)	Canada
	Robert Koch Institute (RKI)	Germany
Stakeholders	Ministries of Health	Canada
		France
		Germany
		Italy
		Japan
		Mexico
		United Kingdom
		United States
		Directorate General for Health and Consumers of the European Commission (DG-SANCO)
		European Food Safety Authority (EFSA)
	World Health Organization (WHO) as observer	

doi:10.1371/journal.pone.0057252.t002

es, i.e., raw signals in unprocessed news articles, messages in forums, official press releases, extracts from public official websites, etc. These signals are then stored on dedicated web-based platforms (specific to each system) accessible to end-users for their assessment and verification.

Databases. Two different databases were constituted: The first one (“prospective”) aimed at assessing the event detection process under close to real life conditions (i.e., detecting pertinent signals potentially relevant for the study among a large volume of raw signals). The second (“retrospective”) database aimed at assessing systems’ theoretical performances.

“Prospective” database: Detection rate, Effective Positive Predictive Value (EPPV), F1-score. All raw signals detected between March 1 and March 31, 2010 that potentially referred to an A/H5N1 event (human cases and epizootics) were considered for the analysis. Raw signals were automatically collected (through prospective specifically designed queries), provided directly by the systems (i.e. datasets), or collected manually (through retrospective *ad hoc* queries). From these data, each signal captured by the systems was reviewed and classified as: detected or inadequately detected. A detected (DET) event was defined as the first report mentioning a human case or an epizootic detected by a system in March 2010 and before the reporting of this event by WHO/OIE on their respective websites. An inadequately detected (XDET) was defined as a signal initially tagged A/H5N1, but after verification was found to be not related to the occurrence of confirmed A/H5N1 cases, or an event previously detected by the same system (i.e. duplicate), or an A/H5N1 case report detected by a system after or on the same day as the reporting of this event by WHO/OIE. A not-detected (NDET) event was defined as an event reported by WHO/OIE but not detected by the system in March 2010. True negative events could not be considered because it is not possible to determine the total number of reports issued on the Internet nor those events

discovered but not published by systems. The detection rate (DR) was defined as the ability of a system to detect confirmed A/H5N1 cases before their reporting by WHO/OIE ($DR = DET / (DET + NDET)$). Effective Predictive Positive Value (EPPV) was defined as the probability for the system to timely detect confirmed A/H5N1 cases among all reports ($EPPV = DET / (DET + XDET)$). The F1-score is the harmonic mean of DR and EPPV, weighted equally [16] $F1 = 2 * (EPPV * DR) / (EPPV + DR)$.

“Retrospective” database: Sensitivity and timeliness. For each event included in the gold standard (i.e., reported by WHO or OIE), a specific manual retrospective search was performed on all systems to identify the first report related to this event. No restriction was set on the time period in order to capture both early and late event detection. A true positive (TP) event was defined as the first report mentioning a human case or an epizootic detected by a system before the reporting of the event by WHO/OIE. A false negative (FN) event was defined as an event not detected by the system. Sensitivity ($Se = TP / (TP + FN)$) was defined as the retrospective ability of a system to detect an event included in the gold standard. Timeliness was defined as the delay between official reporting and the detection by a system (date of report on WHO/OIE websites minus date of first detection by the system, in days). Common variables were used for the analyses: mean, median, rates. Box plot graphs were made to display timeliness, statistical measures and the ANOVA test was used to compare mean values. All statistics were computed using Stata 11.0 for Windows.

Type of events. When used with the terms DR, EPPV, timeliness and Se, “overall” refer to animal and human cases.

Virtual combined system. In order to assess the complementarity and added value of combining the systems’ information, a virtual system named “combined system” was constructed by pooling signals detected by all systems.

Results

Qualitative analysis

Ten users from seven countries and EU public health institutions (Centers for Disease Control and Prevention (CDC), European Commission (EC), European Centre for Disease Prevention and Control (ECDC), Health Protection Agency (HPA), Institut de Veille Sanitaire (InVS), Istituto Superiore di Sanità (ISS), National Institute of Infectious Diseases (NIID), Public Health Agency of Canada (PHAC) and Robert Koch Institute (RKI)) participated in the survey. Respondents were either the head of a unit or an epidemiologist in charge of epidemic intelligence related activities within their institution.

Epidemic intelligence focus. According to participants, epidemic intelligence processes varied widely. However, infectious disease was the main focus for most of the experts involved in this survey. All users systematically considered epidemic-prone diseases in general, though, for some institutions the focus was set on specific diseases. Only three countries monitored systematically generic zoonoses in their routine activities. All events involving potential bio-terrorism pathogens were systematically monitored by three countries, while no country included systematically radiological/nuclear and chemical threats in their routine activities. Although, CBRN threats are of interest to all countries, the bio-terrorism aspect (i.e., intentional release) was not considered as pertinent from the detection perspective. Plant diseases were included in the threat detection criteria of one institution. Natural disasters were monitored according to specific criteria (e.g., geographical, type and size of disaster).

Variations in system usage. Not all users had access or used routinely all of the expert systems included in the study. Of the seven systems four are freely accessible (BioCaster, Health-Map, MediSys and ProMED-mail) and three have restricted access (Argus, GPHIN and Puls). Nine of the ten respondents utilized regularly at least one of the seven included systems; the remaining respondent used other expert systems not included in this survey. Users routinely accessed from four to seven different systems and their utilization varied greatly. ProMED-mail was used routinely by all respondents while utilization of the other six systems ranged from 60% to 80%. When routinely used, Argus, GPHIN and ProMED-mail were accessed on a daily basis. GPHIN and ProMED-mail were predominantly used for early prospective alert detection (60%), while others were used mostly as a complementary source of information (e.g. to further document already detected events). Finally, 60% of users also utilized other epidemic intelligence systems that were not integrated into the survey, e.g., RSOE-EDIS (Radio Distress-Signaling and Infocommunications, Emergency and Disaster Information Service) or EpiSPIDER [20].

Systems' users perception. The perceptions of users regarding the system attributes (completeness, flexibility, representativeness, sensitivity, simplicity, timeliness and usefulness) are represented Figure 1. Timeliness scores ranged from 33% to 100% and usefulness scores ranged from 40% to 100%. Simplicity was the highest scored attribute with scores ranging from 60% to 100%. Sensitivity ranged from 0% (i.e., no user qualified the system as sensitive) to 80%. The spider charts also highlighted the global weaknesses perceived by users with lower scores for three attributes: flexibility (17% to 60%), representativeness (25 to 50%) and completeness (0 to 40%). Individual spider graphs tend to have relatively similar surfaces, except for the less utilized systems (<5 users).

Quantitative analysis: A/H5N1 data

Detection rate, positive predictive value & F1-score (Table 3). A total of 1,154 signals potentially relating to A/H5N1 events were collected. For the same period, 29 A/H5N1 events were reported by WHO (14 events) or OIE (15 events) and were included in the gold standard. In regard to large differences in terms of intrinsic systems' characteristics, interface designs, database storage or extraction capacities it has not been possible to implement a homogenous data collection procedure across the seven systems. For one system (GPHIN), the system design did not allow the extraction or collection of data in a format compatible with this analysis and as such 366 signals were excluded from the analysis. As of July 30, 2010, six datasets were collected from the six other systems for a total of 788 signals. Three of these datasets were collected prospectively and three were collected retrospectively.

For the six systems, the overall detection rate (DR) ranged from 31% to 38%, from 29% to 57% for human cases and from 20% to 40% for epizootics. Differences in DR were observed between human cases and epizootic events (the largest being 57% for human cases versus 20% for epizootics). For the combined system (pooled from six systems), the DR increased to 72% overall, to 93% for human cases and to 53% for epizootics. Overall EPPV ranged from 3% to 24% and the F1-score ranged from 6% to 27%. The overall EPPV and F1-scores of the combined system were 3% and 5%, respectively.

Sensitivity and Timeliness. Two events (7%) were not detected by the systems before official notification, 6 (21%) events were detected by only one system and only 2 (7%) were detected by the seven systems (Table 4). Sensitivity ranged from 38% to 72% for overall A/H5N1 events, from 29% to 79% for human cases and from 33% to 67% for epizootics. For five systems the sensitivities were higher for human cases than for epizootics. When considering the virtual combined system (seven systems) overall sensitivity increased to 93%, 100% for human cases and 87% for epizootics (Table 5). Timeliness for human cases detected by the systems varied from 1.9 days (confidence interval 95%: -0.4; 4.1) to 6.1 days (3.1; 9.1) before the reporting by WHO. For epizootics the mean timeliness varied from 2.9 days (-3.9; 9.7) to 12.7 days (3.4; 22.0) before OIE reporting. Overall timeliness ranged from 2.2 days (0.5; 3.8) to 7.8 days (4.0; 11.5) before WHO/OIE reporting. Differences observed among systems were not significant (F-statistic calculated for analysis of variance (ANOVA) = 0.553). For the combined system (pooled from seven systems), events were detected on average 10.2 days (6.7; 13.8) before their reporting by WHO/OIE, while timeliness for human cases was 6.9 days (4.2; 9.5) and 13.5 days (7.1; 19.9) for epizootics (Figure 2).

Discussion

The results highlight how combining the expertise of multiple epidemic intelligence systems could substantially increase sensitivity and timeliness. When the seven systems were pooled, the sensitivity of the combined system increased to 93% and events were detected on average 10.2 days earlier (-1; 44 days), a period of time that would indeed be crucial for implementation of control measures in the case of a potential threat. The results also point out the many challenges faced by the system, including the lack of specificity of raw information, and the advances that need to be achieved in this domain.

The qualities most frequently quoted by users in the qualitative analysis were simplicity, usefulness and timeliness while flexibility, representativeness and completeness received much lower scores. The findings were corroborated by practices as all users routinely

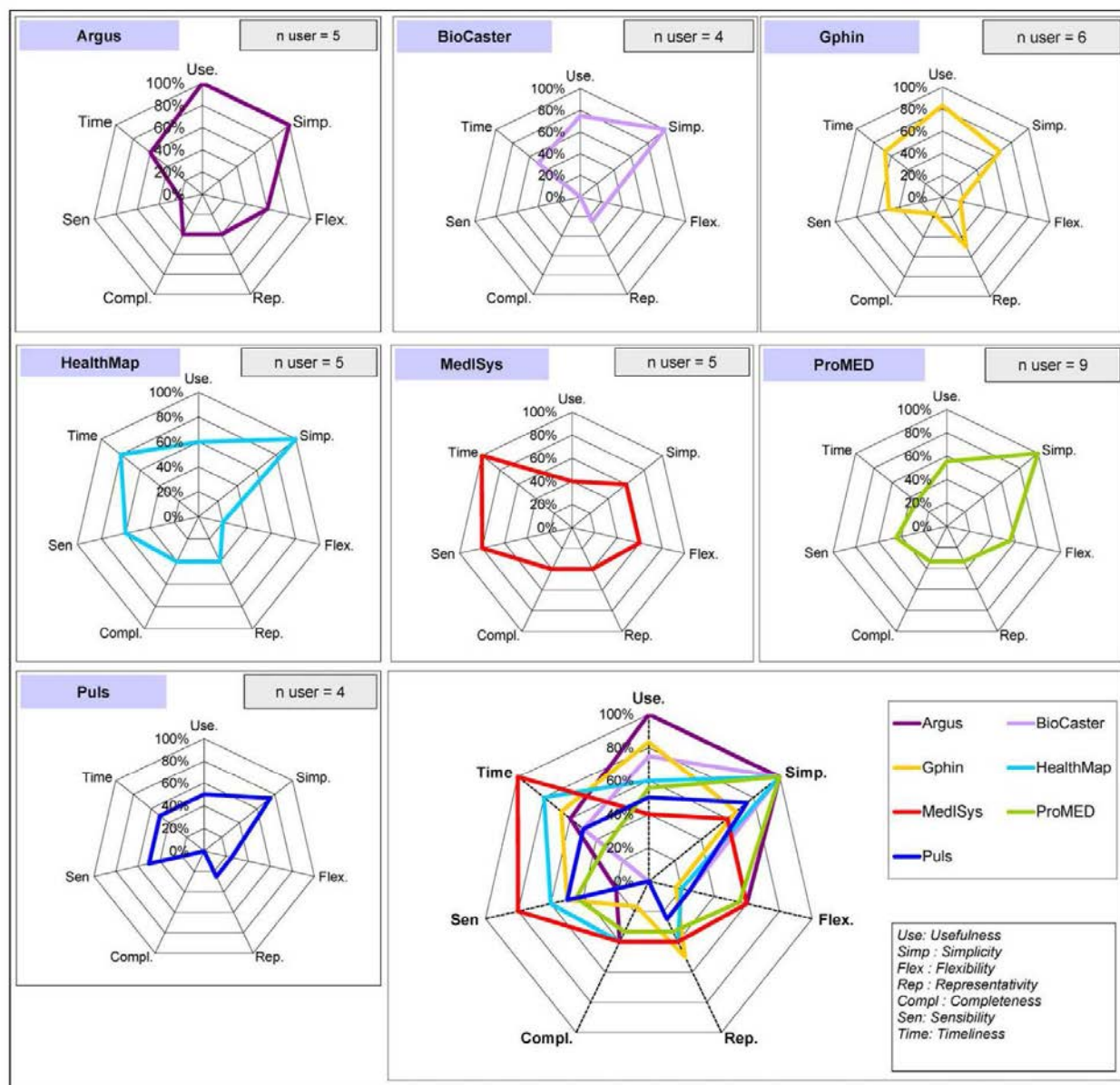


Figure 1. Users' perception regarding systems performances.
doi:10.1371/journal.pone.0057252.g001

utilized four to seven different systems. Not all interviewees were familiar with all systems and their routine utilization varied. Although the global approach was similar, each institution has set objectives and procedures that best suit their specific needs. These differences in both system design and user practice may have influenced the perceptions. Intensively-used systems were more susceptible to being assessed on users' experience and according to their ability to meet an institution's goal (as opposed to their intrinsic performance), while more theoretical opinions might have been applied for seldom used systems. The number of interviewees may appear as a limitation. However, the number of institutions performing structured epidemic intelligence was very limited and the people interviewed were key experts in their domain. It is therefore unlikely that their views substantially differed from those

of the team and the institution they represented. We believe that it is unlikely that adding a few additional people or institutions would have resulted in significantly different results.

It must be stressed that system's designs and functionalities are very different [21]. The quantitative analysis was not designed to compare systems' performances and therefore comparison would be misleading. Rather, the objective was to detect and document diversity and potential complementarities from the end-user perspective. Interpretations of these results should therefore avoid pairwise system comparisons.

The goal of computing the detection rate was to assess the capability of a system and the aptitude for a user to detect relevant information from systems in a situation resembling real life conditions. DR provides an estimate of the events adequately

Table 3. Detection rate, positive predictive value and F1 score for A/H5N1 human cases and epizootic detected by systems from 1st to 31st March 2010.

Systems		Argus	BioCaster	HealthMap	MediSys	ProMED	Puls	Combined system (a)
Collection process		Auto	Auto	Prov	Prov	Auto	Prov	-
n signals		103	95	126	347	37	80	788
A/H5N1 human cases (H)	Detected	5	8	6	5	4	5	13
	Not detected	9	6	8	9	10	9	1
	Inadequately detected (b)	14	20	45	52	14	34	179
	Detection rate	36%	57%	43%	36%	29%	36%	93%
	EPPV	26%	29%	12%	9%	22%	13%	7%
	F1 score	30%	38%	18%	14%	25%	19%	13%
A/H5N1 epizootics (V)	Detected	4	3	5	6	5	6	8
	Not detected	11	12	10	9	10	9	7
	Inadequately detected(d)	66	25	39	227	8	19	384
	Detection rate	27%	20%	33%	40%	33%	40%	53%
	EPPV	6%	11%	11%	3%	38%	24%	2%
	F1 score	9%	14%	17%	5%	36%	30%	4%
Overall A/H5N1 cases (H+V)	Detected	9	11	11	11	9	11	21
	Not detected	20	18	18	18	20	18	8
	Inadequately detected (e)	94	84	115	336	28	69	767
	Detection rate	31%	38%	38%	38%	31%	38%	72%
	EPPV	9%	12%	9%	3%	24%	14%	3%
	F1 score	14%	18%	14%	6%	27%	20%	5%

Auto: Automatically emailed; Prov: Provided by system.

(a) Virtual combined system pooling the 6 systems i.e. event detected by any of the system was considered as detected by the combined system, (d) differs from (b) + (c) because it includes events that could not be categorized in human cases or epizootics.

doi:10.1371/journal.pone.0057252.t003

tagged A/H5N1 during this period and detected as such by the analyst. However, in order to have the same denominator (also used for the EPPV estimation), events that could have been detected before March 2010 were not included. The calculated DRs are likely to be underestimated and hence should not be regarded as a proxy for sensitivity. The overall DRs were very similar (from 31% to 38%). These low scores could be attributed to the non-inclusion of signals detected before March 1, 2010, but also to the difficulty for an end-user to prospectively detect relevant information in a large volume of noise.

The low measured EPPV and F1-scores illustrate the varying ability of systems to adequately detect, efficiently sort-out, and make accessible only the pieces of information relevant for epidemic intelligence purposes while reducing the background noise. The F1-score [16], by weighing them equally, can provide good balance between EPPV and DR. System developers can increase the F1 score by improving signal detection (e.g., expanding geographical coverage, languages, sources, etc.) and/or by reducing background noise (e.g., algorithms for de-duplication). In this study, the F1 score was strongly impacted by the high numbers of XDET and the EPPV, which can lower the sensitivity performance (DR). In a period of one month and considering only one clearly identified topic, A/H5N1, 1,154 documents were detected by the seven systems (on average 37 per day), hence providing an indication of the volume of information to be reviewed routinely when extended to an all hazard approach

(i.e., covering all potential health threats). The EPPV of the virtual combined system was very low, however a genuine operational combined system would include functionalities (e.g., de-duplication) that would substantially reduce the redundant information, hence increasing performance.

No single system was able to detect all events included in the gold standard before their public reporting by WHO or OIE. Sensitivity varied from 38 to 72%. An average difference of 23% was observed between the sensitivities calculated for human cases and epizootics (Table 4) but no explanation was found for such a large difference within and across systems. These findings, however, underline the difference in conceptual design and the associated performance, but also the difficulties met in developing an efficient algorithm covering the different facets of a single disease.

No significant difference (ANOVA = 0.553) was observed between system timeliness. The difference in the number of detected events could have contributed to the observed variation. Systems operated in different time zones and normalizing time proved difficult (because time of posting was not retrievable for all systems). Although for the systems for which information was available, no difference was observed, an effect of the time of posting could not be formally ruled out. Nevertheless, our findings are consistent with other studies: HealthMap detected events around 12 days before WHO publication and ProMED-mail between 2 days and 2 weeks earlier than OIE when events were

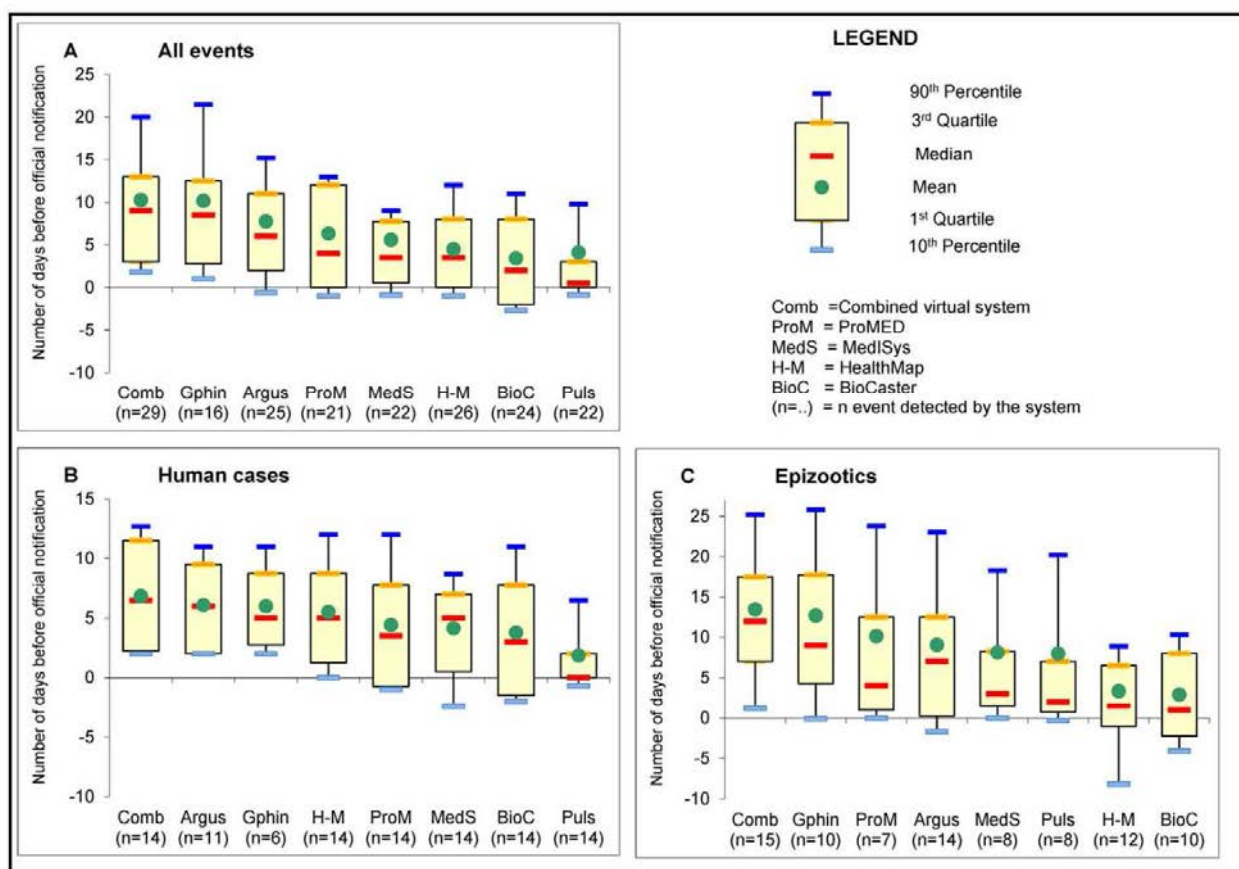


Figure 2. Timeliness of the systems for A/H5N1 cases (total, human, epizootic) reported in March 2010.
 doi:10.1371/journal.pone.0057252.g002

detected by both sources [22], Argus detected confirmed cases of pandemic (A/H1N1) from 1 to 16 days ahead of WHO for 42 countries [40]. No timeliness differences were found between HealthMap, BioCaster and EpiSPIDER [20].

A number of limitations have been identified in this study. The first one concerns the gold standard. The choice of A/H5N1 events was suggested by its public health significance and the existence of an easy to access gold standard. Across affected countries, access to health care, laboratory facilities, surveillance systems, national protocols for biological confirmation (for both human and animal diseases), control strategies vary greatly and not all events will have samples submitted for biological confirmation. Finally, reporting by both the WHO and OIE is subjected to an official notification by a national authority, a process that can take time and that is not always performed. The limits of using WHO and OIE as a gold standard have already

been pointed out by previous studies [22–25], though very few surveys proposed alternatives [26,27]. It is likely that only a portion of genuine A/H5N1 occurrences was effectively reported to WHO or OIE but the magnitude of this bias cannot be estimated. Measured values (DR, EPPV and F1-score) could have been underestimated. Nevertheless, reports classified as XDET were often duplicates (redundant information) or misclassified reports (not related to A/H5N1 cases) as opposed to non-verified events. The impact on the EPPV and F1-score is likely to have been limited while the effect of this potential bias might be more important for DR and Se. It cannot be assumed that the weight of the bias was evenly distributed and that the performances of individual systems were likely to have been affected in different ways.

Despite the heterogeneity of designs [21] the same methodology had to be used for all systems. This uniform approach allowed for

Table 4. Number of gold standard events detected by the systems.

	Not detected	Detected by						
		1 system	2 systems	3 systems	4 systems	5 systems	6 systems	7 systems
n	2	6	3	1	2	8	5	2
%	7%	21%	10%	3%	7%	28%	17%	7%

doi:10.1371/journal.pone.0057252.t004

Table 5. Sensitivity of the systems for A/H5N1 cases (overall, human, epizootic) notified by WHO and OIE in March 2010.

A/H5N1 information (raw signals)		Argus	BioCaster	GPHIN	HealthMap	MedISys	ProMED	Puls	Combined System (a)
A/H5N1 human cases	TP	11	9	4	11	10	9	5	14
	FN	3	5	10	3	4	5	9	0
	Se	79%	64%	29%	79%	71%	64%	36%	100%
A/H5N1 epizootics	TP	10	6	8	7	6	5	6	13
	FN	5	9	7	8	9	10	9	2
	Se	67%	40%	53%	47%	40%	33%	40%	87%
Overall A/H5N1 events	TP	21	15	12	18	16	14	11	27
	FN	8	14	17	11	13	15	18	2
	Se	72%	52%	41%	62%	55%	48%	38%	93%

(a) Virtual combined system pooling the 7 systems i.e. event detected by any of the system was considered as detected by the combined system.

TP = True positive; FN = False Negative, Se = Sensibility.

doi:10.1371/journal.pone.0057252.t005

the provision of a global overview, but did not reflect adequately the large variability of systems' functionalities and genuine performances. The systems are in constant evolution (internal methodology, algorithms, etc.), but for a short study period such changes are likely to be minor and not impact the results. This study was implemented in the scope of EAR and the results had to be delivered within a fixed time frame. The assessment was thus intentionally performed over a short period and was focused on only one topic (A/H5N1) in order to keep the number of signals relatively small. This evaluation could not integrate all potentially important elements, such as languages, geographical distribution, type of sources, interconnections among systems, and others. Excluding such parameters may limit the results generalizability, but despite these limitations and potential biases, the results provided a global perspective and a characterization of the complexity of epidemic intelligence under "real life" conditions. In the scope of EAR, the study results helped to inform future research strategies, i.e., identifying each system's strengths and defining mechanisms that will allow more efficient synergies and cross-fertilization of knowledge and information as opposed to attempting to strengthen "the best of the systems" or to create a "new system".

Conclusions

This study emphasized the added value, and synergistic qualities: between systems, among users and between systems and users. The complexity and the diversity of the epidemic intelligence approaches and the vast expertise developed by the systems are much broader than what could be described in this article [9]. Despite the systems' success, both systems and institutions face major challenges [28] such as the rapidly escalating volume of Internet information, the changing type of communication and information dissemination (i.e., social net-

works and brief, instantaneous communications) and the management of large volumes of data. Levels of duplicative information and noise are very high and international collaboration is still limited. No super-system exists to pool expert systems' expertise and more initiatives must be developed in this direction. More research needs to be carried out, including longer study periods, different types of health events and more robust gold standards. Also additional users and other systems' perspectives should be considered. Overall, this relatively easy to implement study constitutes a first step that will hopefully pave the way for continued exploration in this challenging, but essential component of the global and nations' health security processes and initiatives.

Acknowledgments

Our thanks to the entire EAR working group for their participation and fruitful discussion raised by this manuscript. We gratefully acknowledge GHSI which supports the project and allow countries and institutions to provide human and financial supports for this project.

***EAR project core members:** Ray R. Arthur (CDC), Philippe Barboza (InVS), Mike Barker (HPA), John Brownstein (Children's Hospital Boston), Catherine Chow (CDC), Nigel Collier (National Institute of Informatics), Ricardo Cortes (Ministry of Health of Mexico), Silvia Declich (ISS), Maria Grazia Dente (ISS), Pamela S. Diaz (CDC), Brian Doherty (JRC), Christian Herzog (RKI), Mike Hiley (HPA), Nigel Lightfoot (HPA), Jens Linge (JRC), Lawrence C. Madoff (ProMED-mail), Jas Mantero (ECDC), Abba Mawudeku (PHAC), Noele P. Nelson (Georgetown University), Flavia Riccardo (ISS), Agnes Rortais (EFSA), Mika Shigematsu (NIID), Johannes Schnitzler (WHO), Germain Thinus (DG-SANCO), Laetitia Vaillant (InVS), Roman Yangarber (University of Helsinki).

Author Contributions

Conceived and designed the experiments: PB LV AM PA. Performed the experiments: PB LV. Analyzed the data: PB LV PA. Wrote the paper: PB LV AM NPN DMH LCM JPL NC JSB RY PA.

References

- Chan EH, Brewer TF, Madoff LC, Pollack MP, Sonricker AL, et al. (2010) Global capacity for emerging infectious disease detection. *Proc Natl Acad Sci U S A* 107: 21701–21706.
- Heymann DL, Rodier GR (2001) Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases. *Lancet Infect Dis* 1: 345–353.
- Morse SS (2007) Global infectious disease surveillance and health intelligence. *Health Aff (Millwood)* 26: 1069–1077.
- Formenty P, Roth C, Gonzalez-Martin F, Grein T, Ryan M, et al. (2006) [Emergent pathogens, international surveillance and international health regulations (2005)]. *Med Mal Infect* 36: 9–15.
- World Health Organization (2008) *International Health Regulation (2005)* – 2nd ed. Geneva: WHO Press. 82 p.
- Madoff LC (2004) ProMED-mail: an early warning system for emerging diseases. *Clin Infect Dis* 39: 227–232.
- Brownstein JS, Freifeld CC (2007) HealthMap: the development of automated real-time internet surveillance for epidemic intelligence. *Euro Surveill* 12:

- E071129. Available: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=3322>. Accessed 2012 Jun 12.
8. Collier N, Doan S, Kawazoe A, Goodwin RM, Conway M, et al. (2008) BioCaster: detecting public health rumors with a Web-based text mining system. *Bioinformatics* 24: 2940–2941.
 9. Keller M, Blech M, Tolentino H, Freifeld CC, Mandl KD, et al. (2009) Use of unstructured event-based reports for global infectious disease surveillance. *Emerg Infect Dis* 15: 689–695.
 10. Linge JP, Steinberger R, Weber TP, Yangarber R, van der Gott E, et al. (2009) Internet surveillance systems for early alerting of health threats. *Euro Surveill* 14. Available: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19162>. Accessed 2012 Jun 12.
 11. Bohigas PA, Santos-O'Connor F, Coulombier D (2009) Epidemic intelligence and travel-related diseases: ECDC experience and further developments. *Clin Microbiol Infect* 15: 734–739.
 12. Dente MG, Fabiani M, Gnesotto R, Putoto G, Montagna C, et al. (2009) EpiSouth: a network for communicable disease control in the Mediterranean region and the Balkans. *Euro Surveill* 14. Available: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19113>. Accessed 2012 Jun 12.
 13. Kaiser R, Coulombier D (2006) Different approaches to gathering epidemic intelligence in Europe. *Euro Surveill* 11: E060427. Available: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=2948>. Accessed 2012 Jun 12.
 14. Rotureau B, Barboza P, Tarantola A, Paquet C (2007) International epidemic intelligence at the Institut de Veille Sanitaire, France. *Emerg Infect Dis* 13: 1590–1592.
 15. Chanlekha H, Collier N (2010) A methodology to enhance spatial understanding of disease outbreak events reported in news articles. *Int J Med Inform* 79: 284–296.
 16. Collier N (2010) What's unusual in online disease outbreak news? *Journal of Biomedical Semantics* 1.
 17. Madoff LC, Woodall JP (2005) The internet and the global monitoring of emerging diseases: lessons from the first 10 years of ProMED-mail. *Arch Med Res* 36: 724–730.
 18. Von Etter P, Huttunen S, Vihavainen A, Vuorinen M, Yangarber R (2010) Assessment of Utility in Web Mining for the Domain of Public Health. 29–37.
 19. Global Health Security Initiative (GHSI) Available: <http://www.ghsi.ca/english/index.asp>. Accessed 2012 Jun 12.
 20. Lyon A, Num M, Grosse G, Burgman M (2012) Comparison of Web-Based Biosecurity Intelligence Systems: BioCaster, EpiSPIDER and HealthMap. *Transbound Emerg Dis* 59: 223–232. Doi: 10.1111/j.1865-1682.2011.01258.x.
 21. Hartley DM, Nelson NP, Walters R, Arthur RR, Yangarber R, et al. (2010) The Landscape of International Event-based Biosurveillance. *Emerging Health Threats Journal* 3. Doi: 10.3134/ehj.10.003.
 22. Brownstein JS, Freifeld CC, Reis BY, Mandl K (2007) HealthMap: Internet-based emerging infectious disease intelligence. In: *Global Infectious Disease Surveillance and Detection: Assessing the Challenges – Finding Solutions, Workshop Summary*. Washington: The National Academies Press. 122–136.
 23. Cowen P, Garland T, Hugh-Jones ME, Shimshony A, Handysides S, et al. (2006) Evaluation of ProMED-mail as an electronic early warning system for emerging animal diseases: 1996 to 2004. *J Am Vet Med Assoc* 229: 1090–1099.
 24. Woodall JP (2001) Global surveillance of emerging diseases: the ProMED-mail perspective. *Cad Saude Publica* 17 Suppl: 147–154.
 25. Zhang Z, Chen D, Chen Y, Liu W, Wang L, et al. (2010) Spatio-temporal data comparisons for global highly pathogenic avian influenza (HPAI) H5N1 outbreaks. *PLoS One* 5: e15314. Doi: 10.1371/journal.pone.0015314.
 26. Conway M, Kawazoe A, Chanlekha H, Collier N (2010) Developing a disease outbreak event corpus. *J Med Internet Res* 12: e43. Doi: 10.2196/jmir.1323.
 27. Zeldenrust ME, Rahamat-Langendoen JC, Postma MJ, van Vliet JA (2008) The value of ProMED-mail for the Early Warning Committee in the Netherlands: more specific approach recommended. *Euro Surveill* 13. Available: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=8033>. Accessed 2012 Jun 12.
 28. Wilson K, Brownstein JS (2009) Early detection of disease outbreaks using the Internet. *CMAJ* 180: 829–831.
 29. Nelson NP, Brownstein JS, Hartley DM (2010) Event-based biosurveillance of respiratory disease in Mexico, 2007–2009: connection to the 2009 influenza A(H1N1) pandemic? *Euro Surveill* 15. Available: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19626>. Accessed 2012 Jun 12.
 30. Thomas CS, Nelson NP, Jahn GC, Niu T, Hartley DM (2011) Use of media and public-domain Internet sources for detection and assessment of plant health threats. *Emerging Health Threats Journal* 4.
 31. Torii M, Yin L, Nguyen T, Mazumdar CT, Liu H, et al. (2011) An exploratory study of a text classification framework for Internet-based surveillance of emerging epidemics. *Int J Med Inform* 80: 56–66.
 32. Collier N (2011) Towards cross-lingual alerting for bursty epidemic events. *Journal of Biomedical Semantics* 2 (Suppl 5).
 33. Mawudeku A, Blech M (2008) Global Public Health Intelligence Network (GPHIN).
 34. Brownstein JS, Freifeld CC, Reis BY, Mandl KD (2008) Surveillance Sans Frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS Med* 5: e151.
 35. Freifeld CC, Mandl KD, Reis BY, Brownstein JS (2008) HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports. *J Am Med Inform Assoc* 15: 150–157.
 36. Rortais A, Belyaeva J, Gemo M, van der GE, Linge JP (2012) MedSys: An early-warning system for the detection of (re-)emerging food and feed-borne hazards. *Food Research International* 43: 1553–1556.
 37. Steinberger R, Fuat F, van der Gott E, Best C, von Etter P, et al. (2008) Text Mining from the Web for Medical Intelligence. In: Fogelman-Soulie F, Perrotta D, Piskorski J, Steinberger R, editors. *Mining Massive Data Sets for Security*. Amsterdam: IOS Press. 295–310.
 38. Grishman R, Huttunen S, Yangarber R (2002) Information extraction for enhanced access to disease outbreak reports. *J Biomed Inform* 35: 236–246.
 39. Yangarber R, Best C, von Etter P, Fuat F, Horby D, et al. (2006) Combining Information about Epidemic Threats from Multiple Sources. 295–310.
 40. Nelson NP, Yang L, Reilly AR, Hardin JE, Hartley DM (2012) Event-based internet biosurveillance: relation to epidemiological observation. *Emerg Themes Epidemiol* 9(1): 4. doi: 10.1186/1742-7622-9-4. Available: <http://www.etc-online.com/content/9/1/4>. Accessed 2012 Nov 1.

ARTICLE - 2

OPEN ACCESS Freely available online

PLOS ONE

Factors Influencing Performance of Internet-Based Biosurveillance Systems Used in Epidemic Intelligence for Early Detection of Infectious Diseases Outbreaks

Philippe Barboza^{1*}, Laetitia Vaillant¹, Yann Le Strat², David M. Hartley^{3,4}, Noele P. Nelson⁵, Abba Mawudeku⁶, Lawrence C. Madoff⁷, Jens P. Linge⁸, Nigel Collier^{9,10}, John S. Brownstein¹¹, Pascal Astagneau¹²

1 International Department, French Institute for Public Health Surveillance (Institut de Veille Sanitaire), Saint Maurice, France, **2** Infectious Department, French Institute for Public Health Surveillance (Institut de Veille Sanitaire), Saint Maurice, France, **3** Department of Microbiology and Immunology, Georgetown University Medical Center, Washington, D.C., United States of America, **4** Imaging Science and Information Systems Center, Georgetown University School of Medicine, Washington, D.C., United States of America, **5** Department of Pediatrics, Georgetown University Medical Center, Washington, D.C., United States of America, **6** Centre for Emergency Preparedness and Response, Public Health Agency of Canada, Ottawa, Canada, **7** ProMED-mail, International Society for Infectious Diseases, Boston, Massachusetts, United States of America, **8** Joint Research Centre of the European Commission, Ispra, Italy, **9** National Institute of Informatics, Tokyo, Japan, **10** The European Bioinformatics Institute, Cambridge, United Kingdom, **11** Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **12** École des Hautes Études en Santé Publique (EHESP), University school of public Health, PRES Sorbonne Cité, Paris, France

Abstract

Background: Internet-based biosurveillance systems have been developed to detect health threats using information available on the Internet, but system performance has not been assessed relative to end-user needs and perspectives.

Method and Findings: Infectious disease events from the French Institute for Public Health Surveillance (InVS) weekly international epidemiological bulletin published in 2010 were used to construct the gold-standard official dataset. Data from six biosurveillance systems were used to detect raw signals (infectious disease events from informal Internet sources): Argus, BioCaster, GPHIN, HealthMap, MedISys and ProMED-mail. Crude detection rates (C-DR), crude sensitivity rates (C-Se) and intrinsic sensitivity rates (I-Se) were calculated from multivariable regressions to evaluate the systems' performance (events detected compared to the gold-standard) 472 raw signals (Internet disease reports) related to the 86 events included in the gold-standard data set were retrieved from the six systems. 84 events were detected before their publication in the gold-standard. The type of sources utilised by the systems varied significantly ($p < 0.0001$). I-Se varied significantly from 43% to 71% ($p = 0.0001$) whereas other indicators were similar (C-DR: $p = 0.20$; C-Se, $p = 0.13$). I-Se was significantly associated with individual systems, types of system, languages, regions of occurrence, and types of infectious disease. Conversely, no statistical difference of C-DR was observed after adjustment for other variables.

Conclusion: Although differences could result from a biosurveillance system's conceptual design, findings suggest that the combined expertise amongst systems enhances early detection performance for detection of infectious diseases. While all systems showed similar early detection performance, systems including human moderation were found to have a 53% higher I-Se ($p = 0.0001$) after adjustment for other variables. Overall, the use of moderation, sources, languages, regions of occurrence, and types of cases were found to influence system performance.

Citation: Barboza P, Vaillant L, Le Strat Y, Hartley DM, Nelson NP, et al. (2014) Factors Influencing Performance of Internet-Based Biosurveillance Systems Used in Epidemic Intelligence for Early Detection of Infectious Diseases Outbreaks. PLoS ONE 9(3): e90536. doi:10.1371/journal.pone.0090536

Editor: Vishnu Chaturvedi, California Department of Public Health, United States of America

Received: September 28, 2013; **Accepted:** February 1, 2014; **Published:** March 5, 2014

Copyright: © 2014 Barboza et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: barbozap@who.int

Introduction

Emerging and re-emerging infectious diseases continue to pose major threats to global health security [1,2]. The Internet provides information that can be used to detect health threats early [3]. Epidemic intelligence (EI) relies mainly on *event-based* biosurveillance, i.e. the *ad hoc* detection and interpretation of unstructured information originating from multiple and not predefined sources on the Internet. Sources are varied but typically include the electronic news media and official governmental and

non-governmental organisations [4]. Internet-based biosurveillance systems have been developed to monitor this large volume of information [5]. Despite substantial inherent differences, these systems all scan the Internet to detect reported related to infectious disease that could represent potential health threats, and filter unstructured information through complex algorithms. Select relevant information is stored on dedicated web-based platforms and disseminated. Information collected is then further filtered, verified, and analysed by end-users (i.e. national or international institution and stakeholders involved in EI management).

Table 1. Biosurveillance systems included in the study.

System name	System owner/developer	Country	Starting date	Type of Moderation	n languages	references
Argus	Georgetown University Medical Center	USA	2004	Human moderation	50	[22]
BioCaster	National Institute of Informatics	Japan	2006	Fully automated	7	[23]
GPHIN	Public Health Agency of Canada	Canada	1997	Human moderation	9	[24]
HealthMap	Harvard University	USA	2006	Automated*	7	[25]
MediSys	Joint Research Centre	EU	2004	Fully automated	60	[26]
ProMED	International Society of Infectious Diseases	USA	1994	Human moderation	7	[27]

*Partially moderated.

doi:10.1371/journal.pone.0090536.t001

One of the limitations of event-based biosurveillance is the difficulty of applying traditional epidemiological parameters (e.g. sensitivity, specificity, positive predictive value, etc.), due to the lack of accessible data on verified outbreaks (i.e. a gold standard). The use of Internet-based biosurveillance systems is still maturing and its assessment is on-going [6,7]. Most of the available scientific literature focuses either on the assessment of biosurveillance system performance regarding detection and adequate classification of health-related information using informal open sources, or on the presentation of innovative functionalities. An important topic yet to be elucidated in the literature concerns the performance of biosurveillance systems relative to end-user needs and expectations.

This study aims at providing a quantitative evaluation of multiple biosurveillance systems' performance compared to a gold-standard.

Methods

Epidemic intelligence in France

France is a medium sized country made up of metropolitan France and eleven overseas territories scattered over Africa,

America, and Oceania. French surveillance has been focused traditionally on the detection of unusual health events occurring in the national territory [8,9]. In 2002 the International Department of the French institute for public health surveillance (InVS) developed EI to detect internationally emerging health threats that could affect the French population living in France and abroad [10]. The process was formalized into five steps: detection of informal disease reports (e.g. using biosurveillance systems), selection of disease events (through a set of defined criteria), validation of the event (through a network of contacts, available official information, etc.), analysis, and communication [11].

Events targeted to the InVS public health network are integrated into the weekly international epidemiological bulletin (BHI) available on the InVS website every Wednesday [12]. Only verified events are reported in the BHI, and events are usually reported only once. Updates can occasionally be integrated but only if major epidemiological changes occur.

Table 2. Definition of indicators and rates.

Indicators	Abrev.	Definition
Crude Detected Event	CDET	First signal relating to a health event included in the gold-standard
Intrinsically Detected Event	IDET	First signal detected primarily by the system (excluding signal originating from another system included in the study)
Not-Detected Event	NDET ₁	Event not detected by the system
	NDET ₂	Event not detected by the system or not primarily detected by the system
Crude True Positive event	CTP	First signal related to an event included in the gold-standard and detected by a system before the reporting of the event in the BHI
Intrinsic True Positive event	ITP	First signal related to an event included in the gold-standard, detected by a system before its reporting in the BHI and primarily detected by the system
False Negative	FN ₁	Event not detected by the system before its reporting in the BHI
	FN ₂	Event not detected by the system before its reporting in the BHI or not detected primarily by the system
Rates	Abrev.	Definition
Crude Detection Rate	C-DR	Ability of a system to detect confirmed infectious disease outbreaks. $C-DR = (C-DET)/(C-DET+NDET_1)$
Intrinsic Detection Rate	I-DR	Ability of a system to detect confirmed infectious disease outbreaks independently from other systems. $I-DR = (IDET+NDET_2)$
Crude Sensitivity Rate	C-Se	Ability of a system to detect confirmed infectious disease outbreaks prior to the publication in the BHI. $C-Se = (CTP)/(CTP+FN_1)$.
Intrinsic Sensitivity Rate	I-Se	Ability of a system to detect confirmed infectious disease outbreaks independently from other systems and before their publication in the BHI. $I-Se = (ITP)/(ITP+FN_2)$.

doi:10.1371/journal.pone.0090536.t002

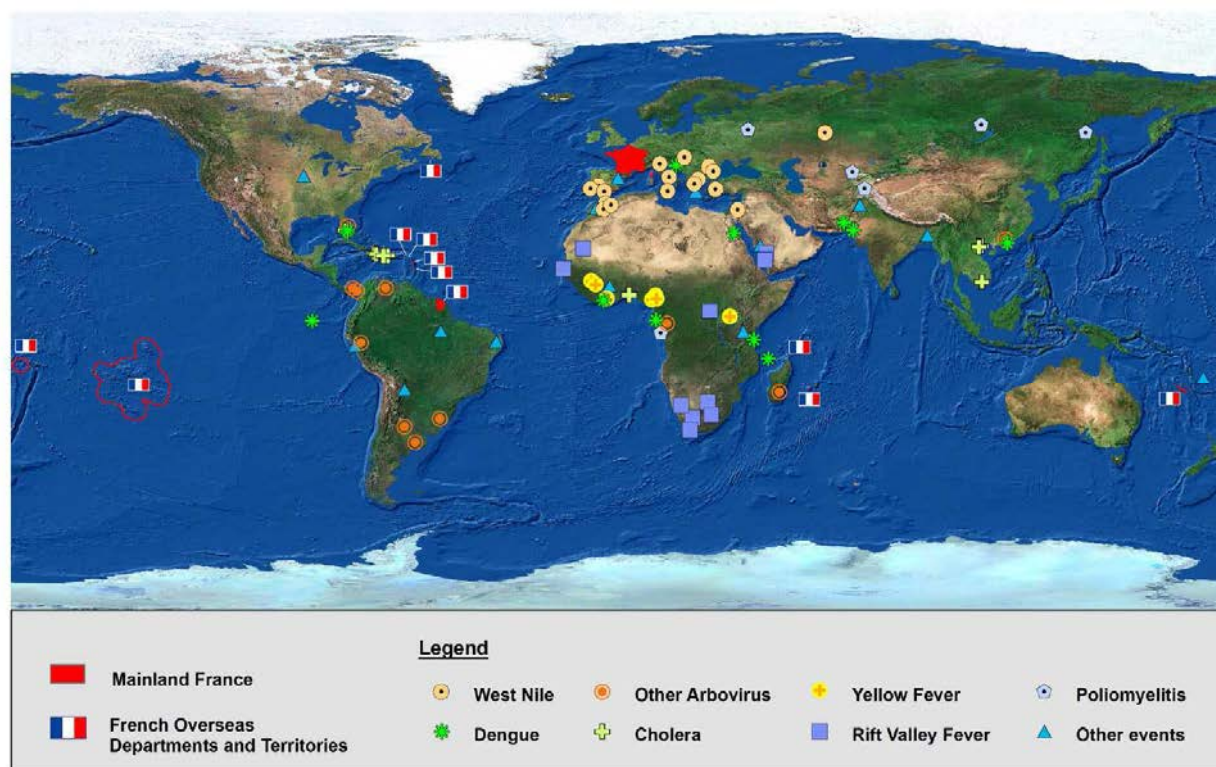


Figure 1. Type and geographical distribution of events published in the BHI in 2010.
doi:10.1371/journal.pone.0090536.g001

Definitions

Events. Events were defined as a verified infectious disease occurrence resulting from the EI process (i.e. including verification and analysis). Events were defined by the disease, the type of cases (human or animal), the country, the province or state, and the month of occurrence regardless of the number of cases concerned. All events are considered to be independent.

Signals. Signals were defined as unverified raw infectious disease information (in relation to an event included in the gold-standard data set) collected from biosurveillance systems (informal sources). Biosurveillance systems are therefore a source of signal, but are not the source of the events included in the BHI (i.e. gold-standard data set).

Biosurveillance systems are used only for the detection of signals. Once selected signals are fully processed (i.e. systematically verified, analysed, characterised, etc.), signals can be classified as an “event” or “discarded”. If biosurveillance systems are the main sources of signals, in the EI process they are not the “source of the event”.

Gold-Standard. Ideally, infectious disease signals identified by Internet biosurveillance methods should be compared to official event reports (gold-standard) during the same time period. The EI performed at InVS, for which necessary information (rational, selection and validation processes, etc.) was readily accessible, was chosen as the best source of gold-standard data for this study.

Infectious disease events (human cases and zoonosis epizootics) reported in the BHI in 2010 were included in the gold-standard data set. A/H5N1 influenza has been considered a health threat for a long time though it is subject to substantial under-reporting [13]. Hence, A/H5N1 cases were excluded from the study.

Non-infectious disease events, not systematically monitored by all systems, were also excluded from our study.

Biosurveillance Systems. Six biosurveillance systems contributing to the Early Alerting and Reporting (EAR) project launched under the Global Health Security Initiative (GHSI) [14] were used to detect raw signals (informal reports of infectious disease): Argus, BioCaster, GPHIN, HealthMap, MedISys and ProMED (Table 1). Of those, three are fully moderated (i.e. include human analysts in the selection, sorting and/or translation processes) while the remaining three systems are based mainly on automated processes for detection, sorting, and translation (Table 1).

Data. Signals relating to events included in the gold-standard data set were retrospectively searched on all six biosurveillance systems through *ad hoc* queries using keywords or a series of keywords. Searches were performed by two InVS epidemiologists (i.e. independent from the six biosurveillance systems). Discordant pairs were reviewed and the most relevant signal was kept in the final database.

Rates. The crude detection rate (C-DR) was defined as the ability of a system to detect an infectious disease event, the intrinsic detection rate (I-DR) was defined as the ability of a system to detect outbreaks independently from other systems. The crude sensitivity rate (C-Se) was defined as the ability of a system to detect outbreaks prior to their publication in the BHI. The intrinsic sensitivity rate (I-Se) was defined as the ability of a system to detect outbreaks independently from other systems and before their publication in the BHI (see table 2 for details).

Associated factors. Bivariable and multivariable modified Poisson regressions (well suited to produce rate ratios) were used to

Table 3. Nature and geographic distribution of events reported in InVS weekly international bulletin (BHI), 2010.

Diseases	America		Sub-Saharan Africa		Asia		Europe		North-Africa Middle-East		Total	
	n	%	n	%	n	%	n	%	n	%	n	%
West Nile	-	-	-	-	-	-	13	(76%)	5	(50%)	18	(21%)
Dengue	4	(17%)	4	(18%)	2	(14%)	1	(6%)	1	(10%)	12	(14%)
Cholera	5	(22%)	1	(5%)	4	(29%)	-	-	-	-	10	(12%)
Rift Valley Fever	-	-	6	(27%)	-	-	-	-	2	(20%)	8	(9%)
Yellow fever	-	-	6	(27%)	-	-	-	-	-	-	6	(7%)
Poliomyelitis	-	-	1	(5%)	4	(29%)	1	(6%)	-	-	6	(7%)
Chikungunya	-	-	2	(9%)	1	(7%)	-	-	-	-	3	(3%)
Plague	2	(9%)	1	(5%)	-	-	-	-	-	-	3	(3%)
Malaria	-	-	-	-	-	-	2	(12%)	1	(10%)	3	(3%)
Saint Louis enc.	2	(9%)	-	-	-	-	-	-	-	-	2	(2%)
Mayaro	2	(9%)	-	-	-	-	-	-	-	-	2	(2%)
Measles	2	(9%)	-	-	-	-	-	-	-	-	2	(2%)
Venezuelan Eq. Enc.	2	(9%)	-	-	-	-	-	-	-	-	2	(2%)
Eastern Eq. Enc.	1	(4%)	-	-	-	-	-	-	-	-	1	(1%)
Oropuche	1	(4%)	-	-	-	-	-	-	-	-	1	(1%)
Crimean-Congo HF	-	-	-	-	1	(7%)	-	-	-	-	1	(1%)
Nipah	-	-	-	-	1	(7%)	-	-	-	-	1	(1%)
Alkhurma	-	-	-	-	-	-	-	-	1	(10%)	1	(1%)
Influenza	1	(4%)	-	-	-	-	-	-	-	-	1	(1%)
Typhoid	-	-	-	-	1	(7%)	-	-	-	-	1	(1%)
Diphtheria	1	(4%)	-	-	-	-	-	-	-	-	1	(1%)
Anthrax	-	-	1	(5%)	-	-	-	-	-	-	1	(1%)
Total	23	(100%)	22	(100%)	14	(100%)	17	(100%)	10	(100%)	86	(100%)

Eq. Enc. = Equine encephalitis HF = Haemorrhagic fever.
doi:10.1371/journal.pone.0090536.t003

assess the associated factors [15]. Rate Ratios (RR), 95% confidence intervals (CI) and p values were computed to assess the strength of these associations. The same variables were included in the different regression models. Potential interactions were tested for their significance at the 005 level. All analyses were performed using Stata 12.1 (StataCorp LP, USA).

Results

Gold-Standard

In 2010, 132 events were reported in the BHI. 46 (35%) were excluded (26 global overviews or long-lasting events, 12 follow-ups of previously reported events and 8 non-infectious health occurrences). The 86 events included in the gold-standard data set occurred in 46 countries; 23 (27%) in the Americas, 22 (25%) in Africa, 17 (20%) in Europe, 14 (16%) in Asia and 10 (12%) in Near-East and North-Africa (Figure 1). Twenty-two disease events were identified. West Nile virus (WNV) infection and dengue infections represented 21% (n = 18) and 14% (n = 12) of the events, respectively. Important variations were observed according to the region of occurrence, e.g. WNV infection represented 76% of events reported in Europe, 50% in the North-Africa and Near-East, but none of the events reported in America or in Asia (Table 3).

Signals Detected

A total of 472 raw signals relating to events included in the gold-standard data set were retrieved from the six biosurveillance systems. The language of the initial source was English for 53% (n = 249), Spanish for 15% (n = 72), French for 11% (n = 53) and other languages for 20% (n = 94). The remaining 4 events were detected through several sources in different languages and information was insufficient to rank them. Language sources varied according to the systems (p = 0.063) and to the region of occurrence of the event (p < 0.001). All 86 events included in the gold-standard data set were detected by at least one of the systems and 57 (66%) were detected by all six systems. When early detection was considered, 84 (98%) events were detected before their publication in the BHI, 30 (35%) were detected by the six systems, 26 (30%) by five systems, 10 (12%) by four systems, 7 (8%) by three systems, 6 (7%) by two systems, 5 (6%) by a single system and 2 events (2%) were not detected prior to their publication in the BHI. According to the system the proportion of media sources utilised by the systems ranged from 44% to 73%, official sources from 6% to 32%. Raw signals originating from another system included in the study ranged from 1% to 49% (p < 0.001) (Table 4).

The systems' C-DR ranged from 83% to 95% (p = 0.20), I-DR ranged from 47% to 92% (p < 0.001), C-Se ranged from 71% to 85% (p = 0.13), and I-Se ranged from 43% to 71% (p = 0.001) (Table 5).

Table 4. Distribution Variables.

Types source	Argus		BioCaster		GPHIN		HealthMap		MedISys		ProMED		Total		p value
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	
Media [§]	53	67%	35	44%	66	93%	40	49%	42	52%	58	73%	294	62%	
Systems *	1	1%	39	49%	1	1%	32	39%	33	41%	1	1%	107	23%	<0.001
Official [#]	25	32%	5	6%	4	6%	10	12%	6	7%	21	26%	71	15%	
Total	79	100%	79	100%	71	100%	82	100%	81	100%	80	100%	472	100%	
Languages	Argus		BioCaster		GPHIN		HealthMap		MedISys		ProMED		Total		p value
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	
English	37	49%	46	49%	33	46%	39	48%	51	63%	43	54%	249	53%	
Spanish	7	9%	11	9%	12	17%	16	20%	8	10%	18	23%	72	15%	
French	9	12%	6	12%	11	15%	7	9%	8	10%	12	15%	53	11%	
Russian	7	9%	4	9%	4	6%	8	10%	4	5%	3	4%	30	6%	0.063
Portuguese	1	1%	4	1%	5	7%	7	9%	3	4%	2	3%	22	5%	
Other	14	19%	8	19%	6	8%	5	6%	7	9%	2	3%	42	9%	
Total	75	100%	79	100%	71	100%	82	100%	81	100%	80	100%	468	100%	
Regions of occurrence	Argus		BioCaster		GPHIN		HealthMap		MedISys		ProMED		Total		p value
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	
Sub-Sah. Afr [£]	22	28%	21	27%	20	28%	21	26%	22	27%	22	28%	128	27%	
America	16	20%	21	27%	19	27%	22	27%	22	27%	22	28%	122	26%	
Europe	17	22%	15	19%	11	15%	17	21%	15	19%	16	20%	91	19%	1.00
Asia	14	18%	14	18%	14	20%	14	17%	14	17%	14	18%	84	18%	
N. Afr.- M.East [‡]	10	13%	8	10%	7	10%	8	10%	8	10%	6	8%	47	10%	
Total	79	100%	79	100%	71	100%	82	100%	81	100%	80	100%	472	100%	
Types of disease	Argus		BioCaster		GPHIN		HealthMap		MedISys		ProMED		Total		p value
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	
Encephalitis	23	29%	19	24%	16	23%	22	27%	19	23%	18	23%	117	25%	
Dengue-like	16	20%	21	27%	15	21%	20	24%	21	26%	20	25%	113	24%	
VHF [¥]	7	9%	7	9%	7	10%	7	9%	7	9%	7	9%	42	9%	1.00
Diarrhoeal	10	13%	10	13%	11	15%	10	12%	10	12%	11	14%	62	13%	
Other	23	29%	22	28%	22	31%	23	28%	24	30%	24	30%	138	29%	
Total	79	100%	79	100%	71	100%	82	100%	81	100%	80	100%	472	100%	
Types of case	Argus		BioCaster		GPHIN		HealthMap		MedISys		ProMED		Total		p value
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	
Human	64	81%	67	85%	61	86%	68	83%	70	86%	68	85%	398	84%	
Animal	15	19%	12	15%	10	14%	14	17%	11	14%	12	15%	74	16%	0.94
Total	79	100%	79	100%	71	100%	82	100%	81	100%	80	100%	472	100%	
Regions of occurrence	English		Spanish		French		Russian		Portuguese		Other		Total		p value
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	
Sub-Sah. Afr [£]	75	30%	1	1%	44	83%	3	10%	1	5%	3	7%	127	27%	
America	51	20%	46	64%	2	4%	0	0%	20	91%	1	2%	120	26%	
Europe	52	21%	21	29%	1	2%	9	30%	1	5%	7	17%	91	19%	<0.001
Asia	45	18%	4	6%	1	2%	15	50%	0	0%	18	43%	83	18%	
N. Afr.- N.East [‡]	26	10%	0	0%	5	9%	3	10%	0	0%	13	31%	47	10%	
Total	249	100%	72	100%	53	100%	30	100%	22	100%	42	100%	468	100%	

[§]Media = Press, news aggregators and blogs.

* Systems = another biosurveillance system included in the study.

[#]Official = official sources and expert contributions.

Table 4. Cont.

[‡]Sub-Sah. Afr = Sub-Saharan Africa.

[†]N. Afr.- N.East = North Africa and Near-East.

^{*}VHF = Viral Haemorrhagic Fever.

doi:10.1371/journal.pone.0090536.t004

Factors associated with Crude or Intrinsic detection

In the bivariable regression analysis individual systems, moderation of system, languages, regions of occurrence, types of disease, and types of case were significantly associated (Table 6). No significant differences were observed across and within systems for C-DR; I-DR and I-Se varied significantly across and within systems; one system (HealthMap) showed a significantly higher C-Se than moderated systems, but across systems the difference was not significant ($p = 0.13$) (Table 5).

From the multivariable Poisson regression models, no statistical difference in C-Se was observed after adjustment for individual systems, moderation, languages, types of diseases, regions of occurrence, and type of cases. Conversely, systems, moderation, languages, regions of occurrence, and types of disease were all significantly associated with I-Se (Table 7). Potential interactions were tested, and none were found significant at the 0.05 level.

Discussion

The systems' characteristics (type of moderation, sources accessed, diseases, languages, and regions covered) were found to significantly influence disease detection performance. This

highlights the differences in conceptual design used to develop the biosurveillance systems, and the importance of taking advantage of synergies through combining systems' data for infectious diseases detection.

C-DR was used to evaluate the global detection rate independently from the EI quality and type. With C-DR values ranging from 83% to 95%, all systems were found to have a similar ability to detect infectious events, findings consistent with other studies [23-34]. However, C-DR is a very crude indicator that does not take into consideration the main EI objective of early detection.

C-Se provides a better estimation of the systems' ability to detect infectious diseases outbreaks early in a given framework defined by the chosen gold-standard. Although biosurveillance systems were originally designed to detect relevant information through informal sources, they now include a noteworthy proportion of early released official information. C-Se of early published official information was 27% lower than C-Se of media sources ($p = 0.001$) underlining the usefulness of media sources in the detection of communicable diseases outbreaks. Despite their different conceptual designs and notably the type of sources used, all systems demonstrated remarkably similar early detection capacities as C-

Table 5. Detection, sensitivity rates and rate ratios from bivariable Poisson regressions

		Argus	BioCaster	GPHIN	HealthMap	MedISys	ProMED	p value
Crude Detection	CDET	79	79	71	82	81	80	0.20
	C-DR	92%	92%	83%	95%	94%	93%	
	RR	-	1.00	0.90	1.04	1.03	1.01	
	CI	-	0.91–1.09	0.80–1.00	0.95–1.12	0.94–1.11	0.92–1.10	
	p value	-	1.00	0.07	0.35	0.55	0.77	
Intrinsic Detection	IDET	78	40	70	50	48	79	<0.001
	IDR	91%	47%	81%	58%	56%	92%	
	RR	1.11	0.57	-	0.71	0.69	1.13	
	CI	0.99–1.26	0.45–0.73	-	0.58–0.88	0.55–0.85	1.00–1.27	
	p value	0.08	<0.001	-	0.001	0.001	0.05	
Crude Sensitivity	CTP	61	63	62	73	67	62	0.13
	C-Se	71%	73%	72%	85%	78%	72%	
	RR	0.84	0.86	0.85	-	0.92	0.85	
	CI	0.71–0.98	0.73–1.00	0.72–0.99	-	0.79–1.05	0.72–0.99	
	p value	0.03	0.06	0.04	-	0.24	0.04	
Intrinsic Sensitivity	ITP	60	37	62	43	44	61	0.001
	I-Se	70%	43%	72%	50%	51%	71%	
	RR	0.98	0.61	1.02	0.70	0.72	-	
	CI	0.81–1.19	0.46–0.80	0.84–1.23	0.55–0.91	0.56–0.92	-	
	p value	0.87	<0.001	0.87	0.006	0.01	-	

RR= Rate Ratio; CI= 95% confidence interval.

None of the system can be considered as a reference and similar results were found using different combinations. For the table, moderated systems were alternatively chosen (alphabetic order) as reference and HealthMap was selected as the reference for CTP to improve readability.

doi:10.1371/journal.pone.0090536.t005

Table 6. Factors associated with crude or intrinsic detection from bivariable Poisson regressions.

		Crude detection			Intrinsic detection				
		RR	95% CI	p value	p value	RR	95% CI	p value	p value
Systems	ProMED	-	-	-	-	-	-	-	-
	Argus	0.98	0.81–1.19	0.87	-	0.98	0.81–1.19	0.87	-
	BioCaster	1.02	0.85–1.22	0.86	-	0.61	0.46–0.80	0.00	-
	Gphin	1.00	0.83–1.20	1.00	0.131	1.02	0.84–1.23	0.87	<0.001
	HealthMap	1.18	1.00–1.38	0.04	-	0.70	0.55–0.91	0.01	-
	MedSys	1.08	0.91–1.29	0.38	-	0.72	0.56–0.92	0.01	-
Moderation of systems	Automated	-	-	-	-	-	-	-	-
	Moderated	0.91	0.82–1.01	0.07	-	1.48	1.27–1.71	0.0001	-
Types of diseases	Encephalitis	-	-	-	-	-	-	-	-
	Dengue-like	1.15	0.97–1.36	0.12	-	1.16	0.91–1.48	0.22	-
	VHF [‡]	1.45	1.23–1.71	<0.001	<0.001	1.69	1.33–2.15	<0.001	<0.001
	Diarrhoeal	1.39	1.18–1.63	<0.001	-	1.73	1.40–2.15	<0.001	-
	Others	1.30	1.12–1.52	0.001	-	1.33	1.07–1.66	0.01	-
Regions of occurrence	America	-	-	-	-	-	-	-	-
	Sub-Sah. Afr [£]	1.16	1.02–1.33	0.03	-	1.33	1.09–1.63	0.01	-
	Europe	0.90	0.75–1.08	0.24	<0.001	1.02	0.80–1.31	0.85	<0.001
	Asia	1.27	1.12–1.45	<0.001	-	1.57	1.29–1.91	<0.0001	-
	N. Afr.-N.East [‡]	0.94	0.76–1.16	0.55	-	0.92	0.67–1.26	0.61	-
Types of case	Human	-	-	-	-	-	-	-	-
	Animal	0.72	0.59–0.86	<0.001	-	0.77	0.61–0.97	0.03	-
Types of source	Media [§]	-	-	-	-	-	-	-	-
	Systems [*]	0.86	0.77–0.97	0.01	0.001	-	-	-	-
	Official [#]	0.79	0.67–0.92	0.004	-	-	-	-	-

RR = Rate Ratio; 95% CI = 95% confidence interval.

[§]Media = Press, news aggregators and blogs.

^{*}Systems = another biosurveillance system included in the study.

[#]Official = official and expert contributions.

[£]Sub-Sah. Afr = Sub-Saharan Africa.

[‡]N. Afr.-N.East = North Africa and Near-East.

[‡]VHF = Viral Haemorrhagic Fever.

doi:10.1371/journal.pone.0090536.t006

Se remained comparable even after adjusting for other variables ($p = 0.13$).

A high level of cross-feeding (i.e., a system using another fellow system as a source) was documented, further emphasising the synergistic qualities of the systems. For automated systems, 39% of early detected signals were collected from another system versus 1% for moderated systems ($p < 0.001$). I-Se provides an appropriate proxy to assess a system's detection rate. Individually (data not shown) and collectively (Table 6) moderated systems were found to have a 53% higher I-Se ($p < 0.001$) than automated systems after adjustment for the other variables included in the model. This increased I-Se of moderated systems can be attributed directly to their common characteristic: the human moderation. The difficulties met by the systems in developing an efficient algorithm covering the different facets of a single disease have been demonstrated by a previous study [7]. Hence, our study illustrates the significant added value resulting from the input of human analysts and their ability to balance the limits inherent to a fully automated detection. Yet, I-Se should be considered as a lower limit, because for each system only the first detected signal was considered. Therefore, it cannot be excluded that some signals primarily collected via a fellow system may have been later

detected through another source, but still earlier than the gold-standard. All systems integrate a de-duplication module, aiming at reducing the volume of redundant information; de-duplication performance varies according to systems. A consistent collection of the second detected signals was not possible across all systems, and the weight of this potential bias cannot be estimated.

At the time of the study, none of the systems were able to detect early all events included in the gold-standard data set, substantiating the necessity for end-users to use several systems in parallel. The purpose of cross-feeding is to increase sensitivity by utilising all available pieces of information. However, such cross-feeding matters when several systems are used in parallel. In a previous study, some authors documented that major EI stakeholders routinely accessed four to seven different systems for event detection [7]. In this context, cross-feeding generates a substantial level of duplication for end-users. In the current study, 43% of signals detected by automated systems would have already been seen on a fellow system. This stresses the importance of developing a common tool that would combine system outputs and specificities while reducing duplication.

Overall 97% of detected signals were published in seven languages (English, Spanish, French, Russian, Portuguese, Arabic

Table 7. Factors associated with crude or intrinsic detection from multivariable Poisson regression models.

Variable	Crude detection			Intrinsic Detection		
	RR	95% CI	p value	RR	95% CI	p value
Types of System						
Automated	-	-	-	-	-	-
Moderated	0.96	0.88–1.04	0.27	1.53	1.34–1.75	<0.001
Languages of detection						
English	-	-	-	-	-	-
Spanish	1.10	0.96–1.27	0.18	1.22	0.99–1.51	0.06
French	0.99	0.86–1.14	0.93	1.06	0.85–1.33	0.58
Other	1.08	0.98–1.20	0.13	1.21	1.01–1.44	0.04
Regions of occurrence						
America	-	-	-	-	-	-
Sub-Sah. Afr. [‡]	1.10	0.95–1.28	0.20	1.25	0.98–1.60	0.07
Europe	0.95	0.79–1.13	0.54	1.18	0.91–1.52	0.22
Asia	1.09	0.97–1.24	0.15	1.30	1.07–1.59	0.01
N. Afr.-N.East [†]	1.16	0.98–1.37	0.08	1.15	0.86–1.53	0.35
Types of disease						
Encephalitis	-	-	-	-	-	-
Dengue-like	0.95	0.80–1.12	0.53	1.11	0.84–1.47	0.46
VHF [‡]	1.06	0.86–1.31	0.60	1.39	0.99–1.94	0.06
Diarrheal	1.09	0.91–1.30	0.34	1.55	1.16–2.06	0.003
Other	1.01	0.85–1.21	0.90	1.14	0.87–1.49	0.33
Types of cases						
Human	-	-	-	-	-	-
Animal	0.84	0.68–1.02	0.08	1.08	0.84–1.37	0.55

RR = Rate Ratio; 95% CI = 95% confidence interval.

[‡]Sub-Sah. Afr = Sub-Saharan Africa.

[†]N. Afr.-N.East = North Africa and Near-East.

[‡]VHF = Viral Haemorrhagic Fever.

doi:10.1371/journal.pone.0090536.t007

and Chinese), findings consistent with another study [16]. These languages were the first integrated by the systems and as such the linguistic methodology (i.e. ontology) might be better developed for those languages than for languages incorporated more recently. Signals in all languages incorporated in the systems were systematically considered. Yet, it cannot be formally excluded that it might have been easier for analysts to detect information published in English, French, Spanish, or Portuguese than in information published in languages requiring systematic translation (e.g., Arabic, Chinese) or rarely used. Despite this potential bias, the results underline the importance of multiple languages tools.

The lack of recognised and consistently available reference sources across diseases and regions represents a major challenge to the evaluation of EI and biosurveillance systems [17,18]. The choice of the gold-standard for this study (BHI) might have impacted the results. The disease and the location of occurrence are among the selection criteria used for EI at InVS. Events occurring in an area close to a French territory were more likely to be retained. Arboviruses represent a risk of exportation especially in overseas territories where competent vectors are present [19], as illustrated in 2006 by the outbreak of chikungunya virus, which affected over one third of the population of both Reunion and Mayotte islands [20]. For France, the circulation of major arthropod-borne infections in previously non-endemic areas is perceived as a threat to the blood supply, and therefore is a topic of high interest. This may partially explain both the high proportion (64%) of vector-borne diseases among reported events and the specific attention placed on those events. For instance, 2010 was a year marked by unprecedented WNV circulation in the Mediterranean area [21] and all 18 WNV infection events

reported in 2010 occurred in Mediterranean countries not previously considered as endemic. Hence, other EI stakeholders, with different objectives or disease distribution, would certainly select different events resulting in another gold-standard data set. The performance of the EI process might also affect the results. Unfortunately, in the absence of a recognised international gold-standard, the sensitivity of the chosen gold-standard could not be assessed. This assessment should be carried out, but was beyond the current studies objectives. Although, these results are not fully representative, France's large geographic distribution (spread over four continents) and systematic and stable approach to EI suggest that the results from this study provide a larger overview that can be transposed to other contexts. Others studies using other gold-standards should be implemented to better assess and ensure generalizability of the results.

The number of events included in the gold-standard was too limited to allow stratified analysis and may have undermined potential associations. For example, it is likely that some systems might have developed specific competences in specific languages, for certain geographic regions, or specific diseases, but this aspect could not be explored since biosurveillance systems are in constant evolution, and as such, the short time frame was chosen to limit intra-system variability. An extended study time period to generate more disease events could be considered for a future study.

The retrospective search for information might have influenced the results. In a previous study, authors found that prospective detection rates were 17% lower than the retrospective sensitivity rate [7]. However; this bias applies equally to all systems and does not affect the overall findings. Similarly, in this study no significant difference ($p=0.53$) was found between the two InVS epidemiologists suggesting that the user-bias was limited.

In conclusion, infectious diseases, environmental issues, and potential bioterrorist threats will continue to pose major risks for global health security and epidemic intelligence is now an essential component of early warning systems. Overall, the systems' disease detection capabilities are complementary (synergistic) with demonstrated timeliness and sensitivity [7]. The output from these systems and also the expertise of the public health institutions responsible for EI should be pooled for optimal early detection. Internet biosurveillance systems have evolved substantially; sufficient data is now available to implement robust validation studies using epidemiological approaches against an official comparison data set. Moreover, larger scope studies should be implemented that would prospectively involve major stakeholders,

increase the number of epidemiologists involved, and enable implementation of innovative strategies to pool the expertise developed by the different systems. Especially, a more robust composite gold-standard that pools information and expertise from national and international institutions in charge of EI activities should be developed.

Author Contributions

Conceived and designed the experiments: PB LV PA. Performed the experiments: PB LV. Analyzed the data: PB LV YL. Wrote the paper: PB LV YL DH NN AM LM JL NC JB PA.

References

- Heymann DL, Rodier GR (2001) Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases. *Lancet Infect Dis.* 1(5):345–353.
- Formenty P, Roth C, Gonzalez-Martin F, Grein T, Ryan M, et al. (2006) Emergent pathogens, international surveillance and international health regulations (2005). *Med Mal Infect* 36: 9–15.
- Morse SS (2012) Public health surveillance and infectious disease detection. *Biosecur Bioterror.* 10(1):6–16.
- Paquet C, Coulombier D, Kaiser R, Ciotti M (2012) Epidemic intelligence a new framework for strengthening disease surveillance in Europe. *Euro Surveill.* 12: 212–214.
- Hartley DM, Nelson NP, Walters R, Arthur RR, Yangarber R, et al. (2010) The Landscape of International Event-based Biosurveillance. *Emerg Health Threats J.* 3:e3. doi: 10.3134/ehj.10.003.
- Hartley DM, Nelson NP, Arthur RR, Barboza P, Collier N, et al. (2013) An overview of Internet biosurveillance. *Clin Microbiol Infect* 19(11): 1006–1013. doi: 10.1111/1469-0691.1227.
- Barboza P, Vaillant L, Mawudeku A, Nelson NP, Hartley DM, et al. (2013) Evaluation of Epidemic Intelligence Systems Integrated in the Early Alerting and Reporting Project for the Detection of A/H5N1 Influenza Events. *PLoS One* 8: 1–9. doi:10.1371/journal.pone.0057252.
- Bitar D, Tarantola A, Capek I, Barboza P, Che D (2009) Risk of importation of tropical diseases in metropolitan France: detection, alert, response. *Bull Acad Natl Med.* 193(3):1847–59; discussion 1859–60.
- Josseran L, Foullet A, Caillère N, Brun-Ney D, Illeff D, et al. (2010) Assessment of a syndromic surveillance system based on morbidity data: results from the aerte/venue-santaire-internationale Accessed 13/02/2013.
- Institut de Veille Sanitaire website. Bulletin hebdomadaire international, <http://www.invs.sante.fr/fr/Publications-et-outils/Bulletin-hebdomadaire-international/Tous-les-numeros/>; Accessed 12/02/2013
- Zhang Z, Chen D, Chen Y, Liu W, Wang L, et al. (2010) Spatio-temporal data comparisons for global highly pathogenic avian influenza (HPAI) H5N1 outbreaks. *PLoS One* 5: e15314. Doi: 10.1371/journal.pone.0015314.
- Global Health Security Initiative (GHSI) website. <http://www.ghsi.ca/english/index.asp>; accessed 15/02/2013.
- Zou G (2004) A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *American Journal of Epidemiology.* 159(7), 702–706. doi:10.1093/aje/kwh090.
- Lyon A, Nunn M, Grossel G, Burgman M (2012) Comparison of Web-Based Biosecurity Intelligence Systems: BioCaster, EpiSPIDER and HealthMap. *Transbound Emerg Dis.* 59(3):223–32.
- Cowen P, Garland T, Hugh-Jones ME, Shimshony A, Handysides S, et al. (2006) Evaluation of ProMED-mail as an electronic early warning system for emerging animal diseases: 1996 to 2004. *J Am Vet Med Assoc* 229: 1090–1099.
- Woodall JP (2001) Global surveillance of emerging diseases: the ProMED-mail perspective. *Cad Saude Publica.* 17 Suppl: 147–154.
- Barboza P, Quatresous I (2007) Globalization, emergence and importation of infectious diseases. *Rev Prat.* 30:57(8):867–3.
- Ledrans M, Quatresous I, Renault P, Pierre V (2007) Outbreak of chikungunya in the French Territories, 2006: lessons learned. *Euro Surveill.* 12(9):E070906.3.
- EpiSouth Website (2010) West Nile virus circulation in the EpiSouth countries and neighbouring areas (Hungary and Russia), Update 6th of October 2010. <http://www.episouth.org/download.php?&id=12182>; accessed 18/02/2013;
- Nelson NP, Brownstein JS, Hartley DM (2010) Event-based biosurveillance of respiratory disease in Mexico, 2007–2009: connection to the 2009 influenza A(H1N1) pandemic? *Euro Surveill.* 15(30):2007–9.
- Collier N, Doan S, Kawazoe A, Goodwin RM, Conway M, et al. (2008) BioCaster: detecting public health rumors with a Web-based text mining system. *Bioinformatics* 24: 2940–2941. doi:10.1093/bioinformatics/btn1129.
- Linge JP, Steinberger R, Weber TP, Yangarber R, van der Gott E, et al. (2009) Internet surveillance systems for early alerting of health threats. *Euro Surveill* 14. Available: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19162>.
- Madoff LC (2004) ProMED-mail: an early warning system for emerging diseases. *Clin Infect Dis.* 39: 227–232.

ACRONYMES

ASPC	Agence de la santé publique du Canada
BHI	Bulletin Hebdomadaire International
CDC	Center for Disease prevention and Control
CDTR	Communicable Disease Threats Reports
CE	Commission Européenne
DI	Département International (de l'InVS)
DON	Disease Outbreak News
EAR	Early Alerting and Reporting
EEE	Espace Economique Européen
ECDC	Centre européen de prévention et contrôle des maladies
eWEB	EpiSouth Weekly Epidemiological Bulletin
GHSI	Global Health Security Initiative
HPA	Health Protection Agency
IE	Intelligence Epidémiologique
InVS	Institut de Veille Sanitaire
ISS	Istituto Superiore di Sanità
FAO	Organisation des Nations Unies pour l'alimentation et l'agriculture
MAE	Ministère des affaires étrangères
NIID	Institut national des maladies infectieuses
NRBC	Nucléaire, Radiologique, Biologique et Chimique
OIE	Organisation mondiale de la santé Animale
OMS	Organisation Mondiale de la Santé
ONG	Organisations Non Gouvernementales
RKI	Robert Koch Institut
RSI	Règlement Sanitaire International
SBE	Surveillance des Evénements
SBI	Surveillance des indicateurs
SRAS	Syndrome Respiratoire aigu sévère
UE	Union Européenne
VI	Veille Internationale
VPP	Valeur Prédicative Positive

I. INTRODUCTION

I.1 CONTEXTE

La mondialisation, en particulier l'accroissement du volume des échanges de personnes, de marchandises et les impacts majeurs de l'homme sur son écosystème ont considérablement modifié la dynamique de la transmission des maladies infectieuses^{1,2} et de l'exposition à des substances toxiques. La stratégie séculaire visant à limiter la propagation internationale des infections par des mesures de quarantaine a depuis longtemps cessé de fonctionner. Disposer d'un système performant de surveillance de santé publique, n'est plus suffisant pour assurer la sécurité sanitaire de la population. Pour prévenir la propagation internationale des maladies, les États et les organismes supranationaux doivent être en mesure de détecter le plus précocement possible les risques sanitaires internationaux susceptibles de les affecter.^{3,4}

La globalisation du risque épidémique n'est pas un phénomène nouveau, les migrations ont influencé l'histoire de l'humanité depuis ses origines.^{5,6} La colonisation par des Européens des Amériques et îles du Pacifique Sud à partir du XVIe siècle illustre l'impact dévastateur de l'importation de nouvelles maladies (coqueluche, grippe, rougeole, variole). Les populations autochtones non-immunes ont été décimées et certains auteurs estiment qu'en quelques décennies près de 90% de la population d'origine aurait été anéantie.^{7,8,9} Depuis le XIXe siècle, les avancées technologiques modifient la nature du risque épidémique. Les progrès dans le domaine du transport sont à l'origine de la première pandémie de choléra en 1817, une maladie précédemment endémique que dans le sous-continent indien. Depuis sept pandémies ont été recensées¹⁰. C'est à l'issue des épidémies de choléra qui ont touché l'Europe entre 1830 et 1850, que la nécessité d'une approche internationale du contrôle des maladies infectieuses a été formalisée. La première conférence internationale de la santé qui s'est tenue à Paris en 1851 a ouvert la voie au processus qui aboutira à l'adoption en 1951 du premier Règlement Sanitaire International (RSI) sous l'égide de l'Organisation Mondiale de la Santé (OMS).

En dépit des avancées scientifiques, le risque posé par les maladies émergentes demeure majeur. La première pandémie du XXIe siècle, le SRAS a illustré à la fois la nature et la dimension possible de ces nouveaux risques. De novembre 2002 à juillet 2003, 8 096 cas dont 774 décès ont été notifiés à l'OMS par 29 pays ou territoires¹¹. A l'échelle de la planète, ce bilan peut paraître limité, mais avec une létalité estimée de 15%, cette pandémie, si elle n'avait pas été contrôlée, aurait eu un impact sérieux et durable à l'échelle mondiale. De novembre 2002 à février 2003, l'épidémie est restée circonscrite à Canton et à cinq municipalités de la province chinoise du Guangdong, mais il a suffi d'un seul cas pour que l'épidémie se répande dans le monde. Le 21 février 2003, un médecin chinois qui avait soigné des malades atteints de SRAS dans la province du Guangdong arrive à l'hôtel Métropole de Hong Kong. Son séjour à l'hôtel a duré moins de 48 heures durant lesquelles il a contaminé 14 clients et visiteurs de l'hôtel qui seront à leur tour à l'origine d'épidémies à Hanoï, Toronto et Singapour, d'au moins quatre des foyers hospitaliers de Hong Kong et de cas isolés à Vancouver, en Irlande et aux États-Unis¹². L'épidémie a généré d'importantes perturbations sociales et des pertes économiques considérables. Elle a mis en évidence trois grandes leçons :

- Une surveillance et des capacités d'intervention inadéquates dans un seul pays mettent en danger la sécurité sanitaire et sociale dans le monde entier.¹³
- Le contrôle d'une pandémie est possible. Ce succès a été possible grâce à une coopération internationale exemplaire coordonnée par l'OMS.
- Internet peut jouer un rôle déterminant dans la détection des risques sanitaires. En l'absence de communication fiable de la part des autorités sanitaires concernées, c'est l'utilisation de sources informelles d'information qui ont permis la détection puis le contrôle de l'épidémie de SRAS.

I.2 LIMITE DE LA SURVEILLANCE EPIDEMIOLOGIQUE CONVENTIONNELLE

Classiquement, les systèmes d'alerte précoce reposaient sur les données produites par les systèmes traditionnels de surveillance épidémiologique. Les données produites par ces systèmes, basés sur la détection et la notification par des professionnels ou des structures de santé, sont souvent fiables, mais répondent à des objectifs, des procédures et des contraintes spécifiques. Ces informations sont rarement transmises dans un délai compatible avec la mise en œuvre rapide de mesures de contrôle adéquates, les processus d'analyse, de vérification sont souvent longs et l'obtention d'une autorisation formelle par les autorités compétentes est souvent nécessaire avant que ces résultats ne soient communiqués. Il est également encore trop fréquent que cette communication officielle soit retardée ou empêchée, en particulier en raison de préoccupations concernant d'éventuelles conséquences économiques (par exemple, impact sur le tourisme, mesures d'embargo, etc.) ou de la crainte d'une possible stigmatisation. Enfin, dans de nombreuses régions du globe, souvent plus susceptibles d'être affectées par des phénomènes émergents, l'accès aux structures de soins et plus encore aux capacités diagnostiques est limité.

Des systèmes d'alerte précoce performants existent au niveau national, régional ou international, et leur rôle déterminant dans la détection de risques sanitaires ne doit pas être déprécié. Toutefois, pour être pleinement efficace à l'échelle mondiale, tous les États devraient disposer de systèmes de surveillance permettant une détection précoce performante de tous les risques potentiels sur l'ensemble de leur territoire.

I.3 SOURCES ALTERNATIVES D'INFORMATION

A partir des années 1990, le développement exponentiel d'Internet a permis l'accès à des sources alternatives d'information qui ne sont pas soumises à ces contraintes : les sources informelles c'est-à-dire des sources non non-officielles et/ou ne provenant pas de sources médicales autorisées.

Depuis le début des années 1990, le besoin de disposer à l'échelle internationale, d'outils de détection et de notification des crises sanitaires internationales tenant compte les défis d'un monde de plus en plus interconnecté et interdépendant a émergé. En 1995, l'OMS a initié le processus de révision du RSI. Le RSI (2005) adopté par 193 pays lors de l'Assemblée mondiale de la santé de mai 2005 est entré en vigueur le 15 juin 2007. Il définit les pratiques et les développements de la surveillance des maladies tant aux niveaux nationaux qu'internationaux. Le précédent RSI (1969), basé sur la notification obligatoire de maladies incluses dans une liste préétablie, a montré ses limites. Une telle liste deviendrait très rapidement obsolète et nécessiterait des mises à jour fréquentes (un mécanisme peu compatible avec un système de surveillance et d'alerte réactif et performant). De plus, La survenue de cas d'une maladie même si elle présente un fort potentiel épidémique, ne constitue pas systématiquement une menace de portée internationale.

C'est sur la base de ces constats¹⁴, que le RSI (2005) a été élaboré par l'OMS. Les États signataires s'engagent à notifier toute crise sanitaire de portée internationale définie comme un évènement de santé publique inhabituel constituant un risque de santé publique pour d'autres États du fait de la propagation internationale de la maladie et/ou pouvant éventuellement exiger une riposte internationale coordonnée. Les États sont également tenus de répondre aux demandes de vérification de l'OMS y compris lorsque l'information émane d'une source non officielle, y compris des médias et de la presse.

Si le RSI est un outil essentiel de la sécurité sanitaire mondiale, sa mise en œuvre est encore très récente et les impacts devront être mesurés à long terme. A l'heure actuelle, le RSI ne constitue pas encore l'outil d'alerte précoce de signalement des alertes sanitaires de portées internationale. Des efforts notables ont été réalisés depuis la mise en place du RSI, mais restent néanmoins, insuffisant. En 2012, parmi les 150 États Membres (sur les 194 qui composent l'OMS) qui ont fourni les données de leur autoévaluation, le score moyen mondial concernant les capacités essentielles n'était que de 81% pour la surveillance, 75% pour les laboratoires et de 50% pour la détection des événements chimiques¹⁵. Entre 2001 et avril 2012, parmi les 574 événements suivis par le bureau Europe de l'OMS, le signalement initial provenait pour 39% d'une source presse, pour 28% d'une source officielle, pour 13% de l'OMS ou d'une autre agence des Nations Unies et pour 20% d'une autre source. Depuis la mise en place du RSI et surtout depuis la pandémie de grippale A/H1N1, l'OMS observe une augmentation du nombre de notifications spontanées par les États Membres¹⁶. Ce nombre d'événements rapportés originellement par les points focaux RSI reste faible ; ils ne représentaient que 36% des 340 événements suivis par le bureau Europe de l'OMS en 2006-2012.

Le RSI est basé sur la confidentialité des notifications et l'OMS n'est donc en mesure de communiquer des informations que si un événement répond aux critères définis pour une urgence sanitaire de portée internationale ou si l'État approuve formellement la communication. En pratique très peu d'informations sont partagées et quand elles le sont, c'est souvent de manière très tardive.

Dans une étude portant sur 266 épidémies, Mondor et al.¹⁷ a montré que le délai médian entre le début des événements et la communication publique initiale était inférieur de 10 jours pour les sources non-gouvernementales : 23 jours (IC 95% ; 20-32) contre 33 jours (30-45) pour les sources gouvernementales. Récemment, les études menées suite à la pandémie grippale due au virus A/H1N1¹⁸ ont également permis de montrer la pertinence de l'utilisation des sources informelles dans la détection et le suivi de la pandémie mais également leur utilité pour décrire les principales caractéristiques et les facteurs de risque des 574 premiers décès rapportés dans 28 pays¹⁹.

L'utilisation de l'ensemble des sources d'information disponible au travers d'Internet demeure donc essentielle pour la détection précoce des épidémies et des risques sanitaires internationaux. En revanche, le processus, les dispositifs et les outils mis en place pour collecter ces informations n'ont été que très peu documentés et encore moins évalués. Ce domaine de l'intelligence épidémiologique est encore très récent et les domaines de recherche très vastes²⁰.

I.4 INTELLIGENCE EPIDEMIOLOGIQUE

I.4.1 TERMINOLOGIE

L'intelligence épidémiologique et la surveillance des événements sont des concepts relativement nouveaux. Suivant les utilisateurs (et les institutions), les mêmes termes peuvent avoir des conceptions différentes et un même concept peut avoir des dénominations différentes, par exemple :

- En anglais, les termes « *epidemic intelligence* » et « *medical intelligence* » sont synonymes mais le terme « *medical intelligence* » est plus souvent utilisé dans le contexte militaire.
- En fonction du contexte, « *Epidemic intelligence* », peut référer à la détection des épidémies (maladies transmissibles) ou à l'ensemble des problèmes de santé publique.

Jusqu'à une date très récente, il n'existait pas de document de référence proposant l'ensemble des définitions, la présentation des concepts et leur organisation. En avril 2014, l'OMS a publié un guide intitulé « *Early detection, assessment and response to acute public health events: Implementation of Early Warning and Response with a focus on Event-Based Surveillance; Interim Version.* »²¹ dont l'objectif est d'une part, de fournir des définitions pour les différents concepts et d'autre part, de proposer un modèle organisationnel pour la mise en œuvre des activités de détection précoce et de réponse aux crises sanitaires. Ce document devrait permettre une harmonisation de l'utilisation des différents termes. Ce guide devrait être traduit en français, mais en attendant, il existe encore moins de références concernant l'utilisation de ces termes en français. La terminologie utilisée dans les rares documents en français est dérivée de concepts anglo-saxons. Ce sont les définitions proposées par l'OMS qui ont été utilisées dans cette section. Le choix des termes français utilisés dans ce document a été fait en prenant en considération deux exigences i) rester le plus proche possible de la définition utilisée dans le guide de l'OMS (et donc de sa probable traduction) ii) tenir compte de la terminologie déjà utilisée en français. Ce sont donc les termes l'intelligence épidémiologique, la surveillance des indicateurs et de la surveillance des événements qui ont été retenus.

I.4.2 PRINCIPES

Pour assurer l'efficacité du recueil des données et de leurs analyses, les systèmes d'alerte précoce doivent pouvoir se baser sur un processus formalisé : l'intelligence épidémiologique et sur deux canaux d'information : la surveillance des indicateurs (surveillance de routine) et de la surveillance des événements.

I.4.3 CONCEPTS ET DEFINITIONS

I.4.3.1 Surveillance des indicateurs

La surveillance basée sur les indicateurs (SBI), « *indicator-based* » en anglais, ou plus simplement « la surveillance des indicateurs » a été définie par l'OMS comme : « *le recueil systématique (régulier), le suivi, l'analyse et l'interprétation de données structurées - d'indicateurs- produits par un nombre bien identifié de sources formelles constituées principalement de structures sanitaires* »²¹.

Les données utilisées dans le cadre de la surveillance des indicateurs proviennent principalement des structures sanitaires (structures de soins, professionnels de santé, laboratoires, etc.) mais peuvent également être non-médicales par exemple dans le cas de données entomologiques, météorologiques, etc. La surveillance des indicateurs regroupe tous les mécanismes de recueil systématique de données, qu'il s'agisse de notifications obligatoires de cas, de surveillance syndromique, sentinelle, de données de laboratoire, de données des registres (cancers, maladies rares), de mortalité, etc. Le recueil des données s'effectue à partir de définitions de cas précises et l'analyse repose sur le suivi des tendances ou le franchissement de seuils préétablis.

La surveillance des indicateurs répond à deux objectifs :

- L'alerte précoce pour les pathologies sévères ou à potentiel épidémique.
- Le suivi à plus long terme (morbidity, suivi des programmes, planification) pour des pathologies infectieuses, les maladies non-transmissibles et/ou chroniques et les expositions à des toxiques.

1.4.3.2 Surveillance des Événements

La surveillance basée sur les événements (SBE), « event-based surveillance » en anglais, ou plus simplement la « surveillance des événements » est définie comme : « *le recueil organisé, le suivi, l'évaluation et l'interprétation d'information ad hoc, principalement non structurées, et concernant des événements ou des risques, qui peuvent représenter un menaç aiguë pour la santé humaine* ». ²¹

Les informations recueillies sont de nature très diverse et proviennent de sources multiples officielles ou non-officielles et souvent non préétablies. Il peut s'agir de rumeurs provenant des médias, d'informations *ad hoc* provenant de réseaux informels. Le processus de recueil des informations est principalement actif et réalisé dans un cadre formalisé. ²¹

En recueillant des informations avant la survenue des premiers cas humains ou avant la détection, la confirmation ou la notification d'un événement, la surveillance des événements permet de contourner les contraintes de temps liées à la surveillance des indicateurs et ainsi d'augmenter la sensibilité des mécanismes de détection précoce et donc la mise en œuvre plus rapide des mesures de contrôle.

De nombreuses sources d'information sont utilisées dans le cadre de la surveillance des événements, elles peuvent être schématiquement classées en deux grandes catégories :

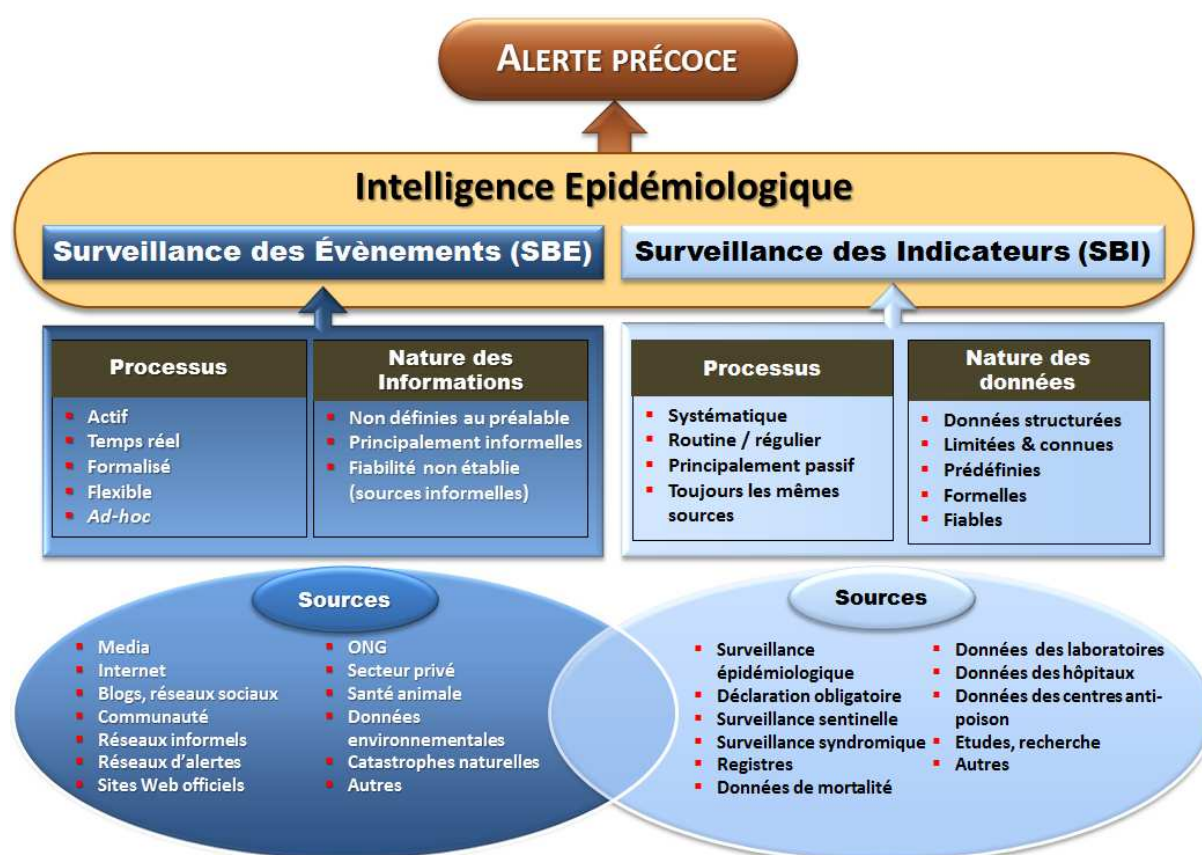
- Les données provenant de la « surveillance électronique » basée sur les données disponibles sur Internet. Il s'agit des informations provenant des médias, des blogs, des réseaux sociaux, mais également de données officielles disponibles sur les sites web institutionnels nationaux et internationaux. C'est dans le traitement de ce type de donnée qu'interviennent les systèmes de biosurveillance.
- Les sources provenant directement de la communauté. Il peut s'agir d'information provenant de la surveillance communautaire, d'information *ad hoc* recueillie en dehors de tout cadre formalisé, d'informations provenant par exemple des Organisation Non-Gouvernementales (ONG), de la société civile (enseignants, professionnels de santé notamment du secteur privé, etc.), de la médecine alternative, du monde associatif et religieux ou encore de l'analyse des données de certaines « hotlines ».

1.4.3.3 Intelligence Epidémiologique

L'OMS définit l'intelligence épidémiologique (IE) comme « *le recueil systématique, l'analyse et la communication d'informations destinées à la détection, la vérification, l'évaluation et l'investigation d'événements et de risques sanitaires avec un objectif d'alerte précoce* ». ²¹ L'intelligence épidémiologique n'a pas pour vocation le suivi à long terme de tendances ou de la morbidité. Elle intègre les informations provenant de la surveillance des événements et des données de la surveillance des indicateurs (par exemple des données issues d'autres systèmes d'alerte précoce, des systèmes de surveillance).

La surveillance des indicateurs et la surveillance des événements sont des sources d'information complémentaires qui contribuent toutes deux à la fonction d'alerte précoce déterminante pour la mise en œuvre d'une réponse rapide et proportionnée. Toutes deux présentent des caractéristiques intrinsèques en matière de type de données / informations recueillies et de processus. Le processus de la surveillance des indicateurs est défini par les professionnels de santé publique pour leur propre usage programmatique spécifique, les données et les indicateurs sont élaborés en conséquence et sont recueillis et transmis en routine. En revanche, dans le cadre de la surveillance des événements, le contenu et le format non-structuré des informations recueillies n'ont pas été initialement conçus pour leur utilisation dans le cadre de la détection des événements sanitaires, l'information est dite « non-structurée » et son recueil est principalement actif (Figure 1).

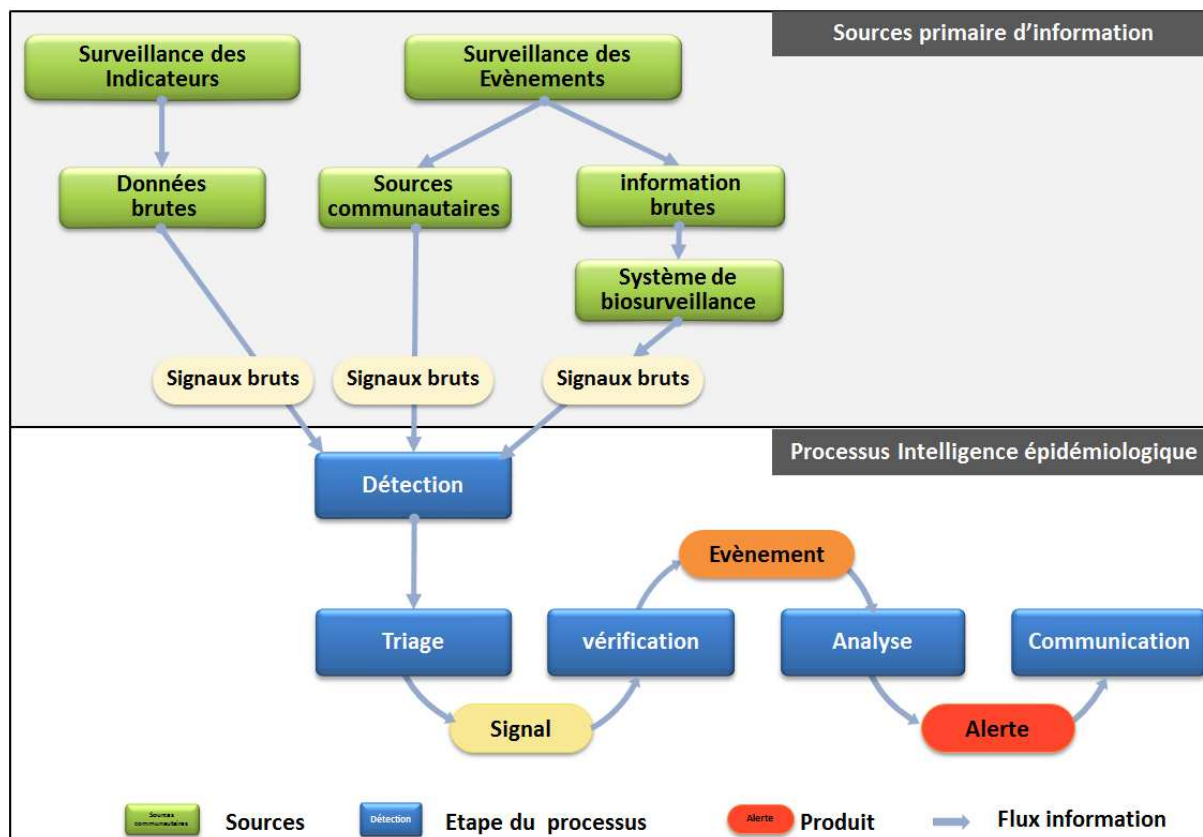
Figure 1 – Concepts surveillance des indicateurs, surveillance des événements et intelligence épidémiologique (adapté de l'OMS ²¹)



I.4.4 PROCESSUS DE L'INTELLIGENCE EPIDEMIOLOGIQUE

L'intelligence épidémiologique implique la mise en œuvre d'un processus formalisé de traitement des données qui comprend 5 étapes : la détection des données/informations brutes, le tri des signaux bruts au travers de critères spécifiques, la vérification des signaux, l'analyse de risque et la communication. Ces étapes sont schématisées dans la Figure 2 et détaillées dans les chapitres suivants.

Figure 2 – Etapes du processus l'intelligence épidémiologique



Les produits du processus sont définis comme suit :

- **Information brute** : toute information disponible sur Internet ou dans les réseaux et avant tout traitement par les systèmes de biosurveillance et/ou les instituts de santé publique.
- **Donnée brute** : toute donnée recueillie par les systèmes conventionnels de surveillance (ISB) avant leur traitement dans le cadre de la détection précoce des évènements sanitaires graves.
- **Signal** : information/donnée traitée - triée et identifiée - comme contenant des informations en relation avec un éventuel évènement ou un risque sanitaire.
- **Signal brut** : toute information fournie par un système de biosurveillance comme pouvant contenir des informations en relation avec un évènement ou un risque sanitaire potentiel et avant son traitement dans le cadre de l'intelligence épidémiologique.
- **Evènement** : tout épisode ayant ou pouvant avoir des répercussions sur la santé humaine : cas sporadique, épidémies, cas animaux de zoonoses, détection de nouveaux pathogènes ou de nouvelles caractéristiques chez des pathogènes connus, intoxication, pollution, exposition à des produits contaminés (quelle que soit la nature de la contamination), catastrophe naturelle (inondation, tremblement de terre, etc.), acte de bioterrorisme, etc. Une fois sélectionné (c'est-à-dire qu'il contient des informations pertinentes correspondant aux critères de sélection définis par l'institution en charge de l'intelligence épidémiologique) puis vérifié (c'est-à-dire que la véracité de la survenue et des éléments connexes a été établie auprès de sources formelles), un signal devient un évènement.
- **Alerte** : tout évènement nécessitant la mise en œuvre, par les autorités sanitaires, de mesures de contrôle ou de mitigation (quelle qu'en soit la nature).

I.4.5 SOURCES D'INFORMATION DE LA SURVEILLANCE ELECTRONIQUE

I.4.5.1 Définitions

- **Source informelle** : source n'émanant pas d'un organisme public ou d'une source autorisée ou formelle. Les sources informelles incluent la presse et les autres médias (radio, télévision, etc.), les blogs, twitter[®], les réseaux sociaux (type Facebook[®]).
- **Source officielle** : source émanant d'une institution gouvernementale nationale ou internationale. Il s'agit principalement des États et des organismes qui en dépendent (ministères de la santé, de l'agriculture, des affaires étrangères, instituts nationaux de santé publique, laboratoires, etc.) et des organisations internationales et supranationales telles que l'OMS pour la santé humaine, l'Organisation mondiale santé animale (OIE) et l'Organisation des Nations unies pour l'alimentation et l'agriculture (FAO) pour les pathologies vétérinaires et les zoonoses, d'autres agences des Nations Unies, d'institutions supranationales comme le Centre européen pour la prévention et le contrôle des maladies (ECDC), de réseaux régionaux (comme le réseau EpiSouth²² pour la Méditerranée), de réseaux de surveillances spécifiques à des maladies (voir Tableau 1).
- **Source autorisée** : ce sont les sources en contact direct avec l'évènement, mais qui ne dépendent pas d'un organisme gouvernemental. On retrouve dans cette catégorie les Organisations Non Gouvernementales (ONG) (ex : Médecins sans Frontières, Action contre la Faim, Save the Children, les Croix Rouges nationales et internationale, etc.), les sources hospitalières et médicales, les laboratoires, les coopérants, etc.
- **Source formelle** : les sources formelles correspondent aux sources officielles plus les sources autorisées.

I.4.5.2 Sources officielles d'information

Les sources officielles d'information qui permettent une détection précoce de risques sanitaires sont relativement peu nombreuses²³. Ces sources fournissent des informations généralement fiables. Issues des systèmes de surveillance, elles contiennent généralement tout ou partie des éléments indispensables à l'analyse du risque (description temps, lieu et personne, données biologiques, etc.). Ces sources peuvent être publiques (disponibles sur les sites Internet des institutions) ou confidentielles notamment quand elles sont partagées dans des réseaux spécialisés.

Les informations issues des sources officielles présentent néanmoins certaines limites. Les informations ne sont pas toujours fiables (exemple SRAS¹²) notamment en début d'épidémie et ou pour des raisons politiques. Les processus sont souvent longs et particulièrement concernant le partage des informations à l'échelle mondiale et leur accès est souvent difficile. Il s'agit souvent de « littérature grise » non publiée, d'information publiée uniquement dans les langues officielles des différents pays ou d'information communiquée confidentiellement.

Tableau 1 – Exemple de réseaux de surveillance et d’alerte potentiellement considérés comme des sources officielles dans le cadre de l’intelligence épidémiologique

Type	Institution	Nom	Acronyme	Couverture	Activité principale	
Mondial	Organisation Mondiale de la Santé (OMS)	Règlement Sanitaire International	RSI	Mondiale	Alerte Précoce	
		Réseau mondial d’alerte et d’action en cas d’épidémie	GOARN www.who.int/csr/outbreaknetwork/fr/	Mondiale	Alerte Précoce	
		Réseau des autorités de sécurité sanitaire des aliments	INFOSAN www.who.int/foodsafety/fs_management/infosan/en/	Mondiale	Alerte Précoce	
	Organisation mondiale santé animale (OIE)	Base de données mondiale d’informations sanitaires	WAHID www.oie.int/fr/	Mondiale	Alerte	
Réseaux régionaux	OMS- EMRO	Réseau d’alerte de surveillance et de détection d’épidémies	RASDON http://gis.emro.who.int/RASDON/Main/Forms/MainPage.aspx	Méditerranée orientale	Surveillance	
	OMS- EURO	Système centralisé d’information des maladies infectieuses	CISID http://data.euro.who.int/cisid/?TabID=307355	Europe	Surveillance	
	ECDC	Système d’alerte précoce européen	EWRS https://ewrs.ecdc.europa.eu/	UE	Alerte Précoce	
		Alerte rapide denrées alimentaires et aliments animaux	RASFF www.ec.europa.eu/food/food/rapidalert/index_en.htm	UE	Alerte Précoce	
	EpiSouth	Système d’alerte précoce méditerranéen	EpiSouth www.episouthnetwork.org/	Méditerranée	Alerte Précoce	
	EpiNorth	Maladies transmissibles en Europe du Nord	EpiWatch http://www.epinorth.org	Europe du Nord	Alerte Précoce	
	Commission Océan Indien	Réseau surveillance des épidémies et gestion des alertes	SEGA www.reseusega-coi.org/	Océan indien	Surveillance	
	Communauté du Pacifique	Réseau océanien de surveillance de la santé publique	ROSSP www.spc.int/phs/ROSSP/index.htm	Pacifique	Alerte Précoce	
	Caribbean Community	Caribbean Public Health Agency	CARPHA http://carpha.org/articles	Caraiïbe	Alerte Précoce	
	ASEAN+3	Centre d’information maladies infectieuses émergentes	ICEID www.aseanplus3-eid.info/	ASEAN+3	Surveillance	
	Réseaux spécifiques pathologies	ECDC/CE	Surveillance de la grippe	EISN www.ecdc.europa.eu/en/activities/surveillance/EISN/Pages/index.aspx	UE	Surveillance
			Maladies virales importées	ENIVD www.enivd.org/	UE	Surveillance
Infection nosocomiale			HAI-Net www.ecdc.europa.eu/en/activities/surveillance/hai/pages/default.aspx	UE	Surveillance	
Soutenus par l’OMS		Dengue	DENGUE-NET www.who.int/denguenet	Mondiale	Surveillance	
		Grippe	FLU-NET www.who.int/flu-net	Mondiale	Surveillance	
		Rage	RABNET www.who.int/rabnet	Mondiale	Surveillance	

1.4.5.3 Sources informelles

Ce sont essentiellement les médias et les réseaux d’information et d’échange accessibles par Internet. Ces informations sont très nombreuses, en principe d’un accès facile, mais présentent deux limites majeures : elles sont souvent peu fiables et leur volume est considérable. Elles constituent néanmoins la source principale des signaux bruts. On peut distinguer plusieurs grandes catégories :

- Les médias qu’il s’agisse de la presse (presse écrite, Internet, radio ou télévision), ou des agences de presse (Agence France-Presse, Reuter, etc.).
- Les sites spécialisés dédiés à des pathologies spécifiques ou des publics particuliers (voyageurs...).
- Les forums qui collectent les informations de leurs membres et des informations publiques
- Les blogs qui reprennent le plus souvent les informations disponibles sur d’autres supports.

Des exemples de ces sites sont donnés dans le Tableau 2. A cette liste, il convient d’ajouter les nouveaux supports de communication tels que les réseaux sociaux (par exemple Facebook® ou Twitter®) dont l’utilisation est encore très expérimentale.

Tableau 2 – Exemples de sources informelles

	Nom	Domaine principal	Pays/ couverture	Type information	lien
Presse traditionnelle	The Jakarta Post	National	Indonésie	Presse	http://www.thejakartapost.com/
	Than Nien News	International	Vietnam	Presse	http://www.thanhniennews.com/
	South China Morning Post	International	Hong Kong	Presse	http://www.scmp.com/
	Le Soleil	National	Sénégal	Presse	http://www.lesoleil.sn/
	O Globo	National	Brésil	Presse	http://oglobo.globo.com/
	Le Monde	National	France	Presse	http://www.lemonde.fr/
Presse Internet	Google News	International		Presse	http://news.google.fr/
	Yahoo News	International		Presse	http://fr.news.yahoo.com/
	News Now	International		Presse	http://www.newsnow.co.uk
Agence de presse	Agence France Presse	International	International	Presse	http://www.afp.com/
	Reuters	International	International	Presse	http://fr.reuters.com/
	Fars New Agency	International	Iran	Organe officiel	http://english.farsnews.com
	Granma	National	Cuba	Organe officiel	http://www.granma.cubaweb.cu/
	Xinhua	International	Chine	Presse + officiel	http://www.xinhuanet.com/english/
	IRIN	Humanitaire	International	Presse + officiel	http://www.irinnews.org/fr
Relief Web	Humanitaire	International	Presse + officiel	http://reliefweb.int/	
	Nom	Domaine	Type information	lien	
Sites spécialisés	CIDRAP	Epidémies	Officiel, presse, scientifique	http://www.cidrap.umn.edu/	
	Avian Flu Diary	Grippe aviaire	Presse+ officiel	http://afludiarary.blogspot.fr/	
	Pandemic Information News	Mal. émergentes	Presse+ blog	http://pandemicinformationnews.blogspot.fr/	
	Novel Coronavirus	Coronavirus	Presse+ officiel	http://www.ncov.co/	
Forum	Flutrackers	Mal. émergentes	Contributions, presse & officiel	http://www.flutrackers.com	
	Forum Influenza H5N1	Grippe	Contributions	http://www.influenzah5n1.fr/index.php	
Blog	Nigeria Health Watch	Généraliste	Blog	http://www.nigeriahealthwatch.com	
	Recombinomics	Grippe	Blog	http://www.recombinomics.com/	
	H5N1	Mal. émergentes	Presse+ blog	http://crofsblogs.typepad.com/h5n1	

I.4.6 LES ETAPES DU PROCESSUS DE L'INTELLIGENCE EPIDEMIOLOGIQUE

I.4.6.1 Détection des signaux bruts

Les mécanismes de détection des événements inhabituels sont tributaires de la nature de la source de l'information. Ils diffèrent selon que l'acquisition des données provienne de la surveillance des indicateurs ou de la surveillance des événements. Pour la surveillance des indicateurs, cette détection repose sur les mécanismes de la surveillance traditionnelle c.-à-d. sur l'élaboration du type et des modalités de recueil des données destinées à d'alerte précoce (format et périodicité du recueil, établissement de seuil, etc.) puis leur recueil et leur analyse. Pour la surveillance des événements, le processus est plus complexe puisque les données proviennent de sources officielles et informelles. Les modes de détections sont donc dépendants du type de source. Il peut s'agir de :

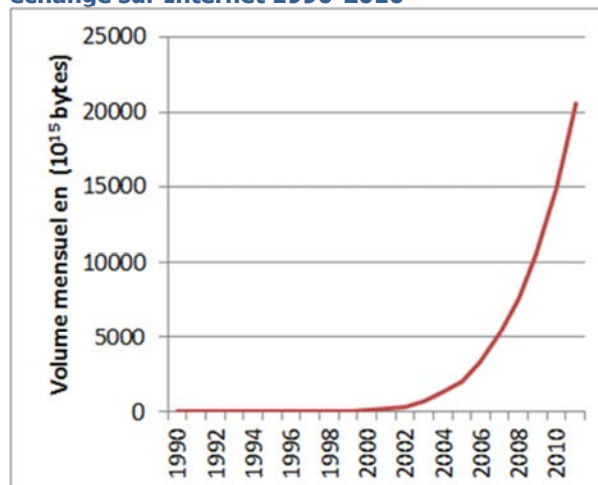
- **Signalement institutionnel** par une autre institution d'un phénomène inhabituel qu'il s'agisse de communication directe dans le cadre de collaborations multilatérales ou de la consultation systématique de sites dédiés à l'alerte précoce tels que le RSI ou l'EWRS pour l'Europe.
- **Recherche active des informations officielles** comme la consultation systématique de sites Internet (ministères de la santé, instituts nationaux de santé publique, agences internationales) mais également la revue de la littérature scientifique.
- **Signalement direct** tel que le signalement « spontané » (hors cadre défini par la SBI) d'événements inhabituels ; par exemple de l'augmentation de la fréquence d'une souche bactérienne spécifique, l'apparition de signes cliniques/sévérité inhabituels, etc.
- **La recherche active d'informations disponibles sur Internet.** C'est principalement dans ce cadre qu'interviennent les systèmes de biosurveillance.

Les mécanismes de recueil sont multiples, il peut s'agir de communications directes (appels directs et spécifiques ou au travers de l'analyse de données des « hotlines » spécialisées), de la consultation systématique (et manuelle) de sites Internet, de la souscription à des listes de diffusion ou de l'utilisation de dispositifs comme les fils RSS (« Really Simple Syndication ») dont le contenu est produit automatiquement en fonction des mises à jour d'un site Web.

1.4.6.2 Détection des signaux bruts à partir d'Internet

Depuis les années 1990, l'utilisation d'Internet et le volume d'information qui y circule augmentent exponentiellement (Figure 3). Les moteurs de recherche classiques tels que Google®, Yahoo® ou Bing® fonctionnant à partir de mots-clés choisis par l'utilisateur sont simplement inadaptés à la recherche d'information pouvant être utilisée dans la détection de risques sanitaires. Par exemple, en avril 2013, la recherche sur Google®, des termes [Bird + flu] donnait potentiellement accès à 607 millions de documents auxquels il convient d'ajouter les 10 millions correspondant à [avian + influenza], les 11 millions pour [禽 + 流 + 感], les 2,5 millions pour [grippe + aviaire].

Figure 3 – Volume mensuel d'information échangé sur Internet 1990-2010*



* En pétaoctet (1 pétaoctet=10¹⁵ octets). Source : adapté de Wikipédia http://en.wikipedia.org/wiki/Internet_traffic (accès le 08/04/2013)

1.4.6.3 Tri des signaux

Le tri des signaux consiste à sélectionner les informations potentiellement pertinentes pour la détection précoce des événements sanitaires nécessitant une réponse. Cette étape est cruciale pour éviter de saturer le processus d'intelligence épidémiologique. Tous les signaux bruts ne correspondent pas à un événement qui pourrait nécessiter une réponse. Il peut s'agir d'information concernant une maladie bénigne, d'un artefact de surveillance mais également de doublons et de fausses rumeurs. Le tri a donc pour objectif de limiter le nombre de signaux qui devront être vérifiés et/ou investigués. Ce triage sera influencé par l'institution qui l'effectue. Les domaines d'application et le volume de signaux bruts sont extrêmement importants, les institutions doivent donc adapter le champ d'application de l'intelligence épidémiologique à leur mandat et à leurs besoins. De fait, la stratégie adoptée par l'OMS dont la mission est de renforcer la sécurité sanitaire mondiale sera différente de l'approche choisie par une institution dont est mandat est national ou d'une organisation comme l'OIE dont l'objectif est la santé animale.

Parmi ces facteurs influençant la stratégie, on trouve le mandat géographique de l'institution (mondial (OMS, OIE), régional (ECDC) ou national (InVS)), le mandat thématique (tous les risques sanitaires (OMS, InVS), les maladies infectieuses uniquement (ECDC)) et d'autres facteurs comme la taille du pays, sa localisation et son histoire. Par exemple, les besoins d'un grand pays avec de longues frontières terrestres sont différents de ceux d'un pays insulaire ou d'un petit pays enclavé au sein dans une entité supranationale comme l'UE. Selon leur organisation structurelle (gouvernement fédéral versus centralisé) et la performance de leur système de surveillance, les pays peuvent avoir besoin de consacrer une part plus importante de leurs capacités d'intelligence épidémiologique à la détection d'événements de santé survenant sur leur territoire.

Un exemple de formalisation des critères de sélection, élaboré par le Département International (DI) de l'Institut de Veille Sanitaire (InVS) est présenté en Annexe 4 ; p 106.

1.4.6.4 Vérification

Cette étape, la plus compliquée du processus, est essentielle. En raison de la sensibilité désirée, les signaux collectés, quelle que soit leur origine, ne peuvent être considérés comme fiables et leur vérification est indispensable. En particulier, les signaux provenant de sources informelles sont susceptibles de contenir des informations erronées, des fausses rumeurs ou des canulars. La vérification consiste à confirmer la réalité (c.-à-d. l'authenticité et la conformité) du signal et de ses caractéristiques. Cette phase consiste à vérifier et à compléter les informations disponibles auprès de sources fiables telles que les instituts de santé publique, les ministères (santé, agriculture, environnement, ...), les laboratoires des pays concernés, l'OMS, les réseaux régionaux, les ONG, les ambassades, la source à l'origine de l'information, les autorités locales, etc. Un travail en réseau est souvent déterminant pour assurer cet échange d'informations, l'analyse et la confirmation des événements. Les signaux bruts ne contiennent en général qu'une partie de l'information nécessaire. La phase de vérification représente également une opportunité de recueillir des informations complémentaires nécessaires pour la vérification, mais également l'analyse du risque.

1.4.6.5 Analyse du risque

L'analyse doit permettre de caractériser l'évènement et d'estimer l'importance du risque qu'il représente en matière de santé publique. Elle repose sur :

- La description de l'évènement en termes de temps, lieu, personne, degrés de connaissance du phénomène, dynamique lorsqu'il s'agit d'un type d'évènement connu (incidence attendue, saisonnalité, distribution géographique connue, etc.)
- La disponibilité d'éventuelles mesures de contrôle.

Elle constitue l'une des principales valeurs ajoutées de l'intelligence épidémiologique. Le processus d'analyse de risque ne diffère pas fondamentalement de ce qui est réalisé classiquement dans le cadre de la surveillance conventionnelle. En revanche, il couvre à la fois des événements infectieux et non-infectieux et les situations où des cas sont déjà survenus et celles pour lesquelles ils ne sont pas encore survenus (introduction d'un vecteur, épidémie dans des zones proches et ou présentant des risques d'importation de cas, produits contaminés, exposition à des agents toxiques, ...).

1.4.6.6 Communication

En fonction de leur nature, les alertes détectées par l'intelligence épidémiologique doivent être communiquées au travers de supports adaptés aux différents acteurs et autorités concernés : ministère de la santé et autres ministères (notamment en cas d'évènement d'origine non-infectieuse), autorités sanitaires internationales (ex : OMS, ECDC), réseau national de santé publique, laboratoires, structures de soins et le cas échéant le grand public.

I.5 HYPOTHESE

Pour couvrir l'ensemble des risques sanitaires, notamment ceux non-couverts par les systèmes conventionnels de surveillance, de plus en plus de pays développent une stratégie dite de veille sanitaire qui intègre la surveillance des indicateurs, la surveillance non-spécifique et l'intelligence épidémiologique. Si, l'intelligence épidémiologique intègre des signaux provenant des systèmes de surveillance conventionnels, elle repose particulièrement sur l'analyse des signaux produits par la surveillance des événements générés par l'utilisation des données informelles non-structurées. La phase de détections des signaux pertinents correspondants est principalement assurée par des systèmes de biosurveillance dont le développement est relativement récent. La majorité de ces systèmes ont été développés sans associer des utilisateurs pourtant essentiels : les institutions de santé publique impliquées dans la détection et la réponse aux crises sanitaires.

L'intelligence épidémiologique et la surveillance des événements sont des concepts relativement nouveaux pour lesquels il n'existe pas encore de définition standardisée reconnue ni de procédures d'évaluation. L'évaluation combinée des performances des principaux systèmes de biosurveillance comparées à un même standard n'a jamais été réalisée et/ou publiée. Ce type d'évaluation est confronté à des obstacles importants : l'absence d'indicateurs d'évaluation spécifiques, le volume et la variabilité de l'information traitée, la complexité du processus et enfin l'absence d'un gold-standard internationalement reconnu et accessible. L'emploi des systèmes de biosurveillance est fondé sur le recueil et le traitement d'un volume très important d'informations provenant de sources informelles dont la fiabilité est très souvent questionnable et dont l'objectif primaire n'est ni la surveillance épidémiologique ni la détection des risques sanitaires.

L'intérêt de l'intelligence épidémiologique pour la détection des crises sanitaires mondiales est reconnu depuis plus d'une décennie. Cette perception repose, et sur des données empiriques et la pratique quotidienne des utilisateurs. En particulier, la surveillance des événements et notamment la surveillance d'Internet et des médias électroniques n'est basée que sur des données indirectes concernant la performance théorique des systèmes de biosurveillance.

L'hypothèse générale de recherche peut être formulée dans ces termes : l'utilisation des systèmes de biosurveillance, dans les conditions du terrain, permet une détection plus précoce des événements sanitaires que les données de la surveillance traditionnelle. Les questions spécifiques qu'elle soulève sont :

- Les informations recueillies sont-elles pertinentes ?
- Les différents systèmes utilisés sont-ils complémentaires ou redondants ?
- Quels sont les facteurs qui peuvent influencer sur la détection des événements ?

I.6 OBJECTIFS

I.6.1 OBJECTIF PRINCIPAL

L'objectif principal de cette étude était d'évaluer, du point de vue des utilisateurs, les performances des systèmes de biosurveillance utilisés dans le cadre de l'intelligence épidémiologique pour la détection précoce des risques sanitaires au niveau mondial.

I.6.2 OBJECTIFS SPECIFIQUES

Les objectifs spécifiques étaient de :

- Réaliser un état des lieux des performances des principaux systèmes de biosurveillance existants du point de vue des utilisateurs
- Déterminer la meilleure stratégie en termes de performance et de valeur ajoutée pour la détection précoce des risques sanitaires.
- Identifier les systèmes les plus performants en prenant en compte les besoins et spécificités des utilisateurs.

II. MATERIEL ET METHODE

II.1 ORGANISATION

La méthode et la présentation des résultats s'articulent autour de quatre phases principales qui ont donné lieu à des articles scientifiques.

- **Première phase** : Définition des modalités de fonctionnement des systèmes de biosurveillance (article A) et sélection des systèmes de biosurveillance à intégrer à l'analyse.
- **Seconde phase** : Identification des indicateurs appropriés pour l'évaluation des systèmes de biosurveillance à partir de la revue des indicateurs utilisés dans le cadre de l'évaluation des systèmes de surveillance des indicateurs (ISB), et d'une revue systématique des méthodes et des indicateurs déjà utilisés dans l'évaluation des systèmes de biosurveillance (article B).
- **Troisième phase** : Estimation de la variabilité des stratégies d'intelligence épidémiologique mise en œuvre par les institutions de santé publique et estimation de l'impact des objectifs et la stratégie sur le type d'évènements détectés et rapportés, le type de support et la fréquence de publication.
- **Quatrième phase** : Réalisation d'un état des lieux des principales performances des systèmes de biosurveillance au travers d'une enquête qualitative menée auprès des principaux acteurs du domaine de l'intelligence épidémiologique et d'une étude quantitative afin : i) d'estimer la sensibilité des différents systèmes, la réactivité et la valeur prédictive positive des signaux détectés ; ii) d'identifier les éventuelles complémentarités ou les redondances entre ces systèmes ; iii) d'identifier une stratégie permettant une optimisation de leur utilisation en prenant en compte les besoins et spécificités des utilisateurs (Article 1).
- **Cinquième phase** : Basée sur les résultats de l'étude précédente, exploration des facteurs influençant les performances en termes de détection des évènements sanitaires (Article 2).

II.2 FONCTIONNEMENT DES SYSTEMES DE BIOSURVEILLANCE (ARTICLE A)

II.2.1 DEFINITION

Les systèmes de biosurveillance ont été définis comme : des systèmes qui utilisent principalement les données non-structurées provenant de diverses sources Internet et dont objectif est la détection précoce des risques sanitaires (qu'il s'agisse de maladies infectieuses humaines, animales ou végétales, de risques chimiques, radiologiques, nucléaires ou de risques naturels). Ils doivent également couvrir différents risques sanitaires (au moins plusieurs pathologies), réaliser un traitement des données collectées et disposer d'une application Internet pour la dissémination des résultats.

II.2.2 PRINCIPES

Les systèmes de biosurveillance ont été conçus et développés de manière indépendante. Ils diffèrent du point de vue de leurs fonctionnalités : nombre de langues (7 à 50 langues), domaines couverts (les maladies infectieuses, risques nucléaires, radiologiques, biologiques et chimiques (NRBC), maladies animales et végétales, catastrophes naturelles, ...), nombre de signaux bruts, type d'accès (restreint ou public), et de leur couverture géographique. Certains systèmes sont entièrement automatisés, d'autres bénéficient d'une modération humaine. Néanmoins, ils présentent tous des caractéristiques similaires : ils collectent automatiquement, à partir d'Internet, en utilisant des mots-clefs ou des séquences de mots-clefs, des informations pouvant être en relation avec une alerte sanitaire. Ces

informations sont traitées et triées grâce à des algorithmes complexes et avec l'apport d'une analyse humaine pour les systèmes modérés. Les signaux ainsi détectés sont restitués (24H/24 – 7j/7) aux utilisateurs au travers d'une plateforme accessible depuis Internet.

II.2.3 LE RECUEIL DES INFORMATIONS

Les informations brutes sont collectées à partir d'Internet à travers deux principales modalités : les agrégateurs de média et des applications spécifiques à chaque système. Ces derniers peuvent utiliser différentes techniques de datamining comme l'indexation « web-crawling »^a conçu pour collecter les informations afin de permettre à un moteur de recherche de les indexer ou le « web-scraping »^b une technique d'extraction du contenu des sites web dont le but de le transformer pour permettre son utilisation dans un autre contexte. Les systèmes revisitent les sites sélectionnés à intervalles réguliers.

II.2.4 TRAITEMENT DES INFORMATIONS

Les informations collectées sont converties dans un format unique afin de permettre les traiter. Le traitement des informations varie en fonction des systèmes et de leurs objectifs. Néanmoins, les catégories suivantes représentent des étapes importantes dans le traitement des données de biosurveillance: traduction, classement par pertinence, le regroupement, extraction et déduplication. Pour les systèmes qui intègrent un module de traduction, es stratégies de traitement linguistique sont multiples : traiter les langues prises en charge séparément, traduire chaque langue source vers une langue cible commune ou traiter les informations avant ou après traduction. Le classement par pertinence consiste à évaluer la pertinence du signal brut en fonction des objectifs. Il s'agit d'un processus complexe dont la première phase consiste à élaborer un révérenciel qui permettra d'évaluer les algorithmes. Elle est suivie de la mise en œuvre de techniques d'intelligence artificielle telles que « l'apprentissage automatique supervisé »^c, les « machines à vecteurs de support »^d, des classifications naïves bayésiennes ou des recherches booléennes incluant des opérateurs logiques tels que « ET » et « OU ».

II.2.5 ONTOLOGIE

En informatique et en sciences de l'information, l'ontologie est l'ensemble structuré des termes et concepts représentant le sens d'un champ d'information. Plus simplement, on peut aussi dire que « l'ontologie est aux données ce que la grammaire est au langage »^e.

La biosurveillance nécessite une connaissance conceptuelle des maladies, des agents pathogènes, des symptômes et de la géographie. Des ressources ontologiques ont été développées ou réutilisées pour la santé publique et plus spécifiquement par des systèmes de biosurveillance. Ces ontologies fournissent les connaissances nécessaires aux systèmes de biosurveillance pour le traitement des termes figurant dans les signaux bruts fournis aux utilisateurs par les systèmes. Par exemple, la mention des mots *Yersinia pestis* peut signifier que la maladie considérée est la peste. Cependant,

^a Un robot d'indexation (« web crawler ») est un logiciel qui explore automatiquement le Web. Il est généralement conçu pour collecter les ressources (pages Web, images, vidéos, documents Word, PDF ou Postscript, etc.), afin de permettre à un moteur de recherche de les indexer. Source Wikipédia http://fr.wikipedia.org/wiki/Robot_d%27indexation accédée 10/04/2013

^b Le Web scraping (parfois appelé harvesting) est une technique d'extraction du contenu de sites Web, via un script ou un programme, dans le but de le transformer pour permettre son utilisation dans un autre contexte, par exemple le référencement. Source définition Wikipédia accédée 10/04/2013 http://fr.wikipedia.org/wiki/Web_scraping.

^c L'apprentissage automatique (« machine learning »), un des champs d'étude de l'intelligence artificielle, est la discipline scientifique concernée par le développement, l'analyse et l'implémentation de méthodes automatisables qui permettent à une machine (au sens large) d'évoluer grâce à un processus d'apprentissage, et ainsi de remplir des tâches qu'il est difficile ou impossible de remplir par des moyens algorithmiques plus classiques. Source définition Wikipédia accédée le 10/04/2013 http://fr.wikipedia.org/wiki/Apprentissage_automatique

^d Les machines à vecteurs de support (SVM) sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination et de régression. Source Wikipédia accédée le 10/04/2013 http://fr.wikipedia.org/wiki/Machine_%C3%A0_vecteurs_de_support

^e En informatique et en science de l'information l'ontologie, est l'ensemble structuré des termes et concepts représentant le sens d'un champ d'informations, que ce soit par les métadonnées d'un espace de noms, ou les éléments d'un domaine de connaissances. L'ontologie constitue en soi un modèle de données représentatif d'un ensemble de concepts dans un domaine, ainsi que des relations entre ces concepts (...). Source définition Wikipédia accédée le 10/04/2013 [http://fr.wikipedia.org/wiki/Ontologie_\(informatique\)](http://fr.wikipedia.org/wiki/Ontologie_(informatique))

toutes les ambiguïtés ne peuvent être résolues aussi facilement. L'un des problèmes les plus pratiques est homonymie toponyme (c.-à-d. noms de lieux). Par exemple, le nom «Cambridge» correspond au moins à une trentaine de localités en Australie, Jamaïque, Nouvelle-Zélande et dans plusieurs états ou provinces du Canada, du Royaume-Uni et des États-Unis.

II.2.6 EXTRACTION DES EVENEMENTS

Une fois l'ensemble des sujets d'intérêt potentiel identifié, les éléments spécifiques sont extraits des données au travers d'algorithmes de reconnaissance de mots clés, de combinaisons booléennes ou de recherches de proximité. Il peut s'agir des informations telles que le nom de la maladie, des espèces concernées, du nombre de cas et de décès, ainsi que de la localisation de l'épidémie. Cette extraction des événements est une procédure très complexe. Si, comme mentionné précédemment, la mention des mots *Yersinia pestis* indique qu'il est très certainement question de la peste, elle n'indique pas s'il s'agit d'un événement nouveau ou d'un article sur la peste noire au Moyen-Âge.

II.2.7 LE DE-DOUBLONNAGE

Le dé-doublonnage (déduplication) consiste à identifier et à supprimer les documents identiques ou dont le contenu est identique. Le processus est basé sur des techniques linguistiques de regroupement (« clustering »). Cette étape est une des plus complexes puisqu'il s'agit de supprimer les informations redondantes sans pour autant supprimer celles qui présentent des variations mineures d'un point de vue informatique mais qui peuvent être essentielles d'un point de vue sanitaire comme l'augmentation du nombre de cas ou le changement de l'état d'un patient.

II.2.8 ANALYSE

Un système de biosurveillance produit un ensemble de documents concernant des événements potentiellement pertinents pour les utilisateurs finaux. Toutefois, seul un sous-ensemble de ceux-ci peut être très utile pour un utilisateur particulier. Compte tenu du conflit entre le volume de données à analyser et la capacité limitée d'un individu d'examiner rapidement de grandes quantités d'informations, des modules complémentaires ont été développés pour augmenter la rapidité de traitement. Le défi consiste à modéliser ce qui est déjà connu (ce qui est normal ou attendu), et à déterminer si l'évènement en cours nécessite un traitement plus urgent. L'analyse de séries chronologiques et d'autres techniques de traitement du signal permettent l'analyse temporelle des indicateurs d'épidémie. La méthode de détection des anomalies tente d'expliquer les caractéristiques de l'évènement dans son contexte afin de déterminer un niveau de signification.

II.2.9 DISSEMINATION DE L'INFORMATION

Les systèmes de biosurveillance utilisent différentes modalités de diffusion de l'information en fonction des besoins, des ressources, de la nature des informations et des utilisateurs. La plupart des systèmes utilisent une combinaison de solutions. Les méthodes actives (transmission systématique à l'utilisateur sans intervention de celui-ci) utilisent les différents supports de communication électronique (email, SMS, twitter). Les modalités « passives » permettent aux utilisateurs d'extraire des informations conformément à leurs besoins. Elles peuvent également permettre d'effectuer des recherches spécifiques sur les sites Internet dédiés des systèmes de biosurveillance, de paramétrer les catégories de maladies ou les zones géographiques. Suivant les systèmes, l'accès est libre (public) ou restreint à des groupes ou des utilisateurs sélectionnés (ex : les institutions de santé publique).

II.3 CHOIX DES SYSTEMES DE BIOSURVEILLANCE

Si, de nombreux sites Internet contiennent des informations sanitaires^{24,25} peu correspondent aux critères définis (§ II.2.1, p 38) : permettre le traitement de données non-structurées pour la détection précoce de différents risques sanitaires dans une application web. En particulier, les réseaux permettant le partage d'informations sanitaires structurées comme les réseaux spécialisés (ex : GeoSentinel <http://www.istm.org/geosentinel/main.html>) ou les sites institutionnels dédiés à l'alerte précoce (ex : Health Emergency & Diseases Information System (<http://hedis.jrc.ec.europa.eu/>) de l'Union Européenne (UE) ne répondent pas à ces critères. Dix systèmes présentant les caractéristiques d'un système de biosurveillance ont été identifiés en 2009-2010 (Tableau 3).

Tableau 3 – Systèmes de biosurveillance en 2010

Système	Développeur	Pays	Accès	Cible principale	Cible	Références
Argus	Université de Georgetown	USA	Restreint	Infectieux	Agences gouvernementales US	[26, 27, 28]
BioCaster	Institut National d'Informatique	Japon	Public	Infectieux	Tout public	[29, 30, 31]
GPHIN	Agence de santé publique du Canada	Canada	Restreint	Tous risques	Instituts de santé publique	[32, 33, 34]
HealthMap	Université d'Harvard	USA	Public	Infectieux	Tout public	[35, 36, 37]
MedISys	Joint Research Centre	UE	Public & Restreint	Tous risques	tout public & Instituts santé pub.	[38, 39, 40]
ProMED-mail	International society of infectious diseases	USA	Public	Infectieux	Tout public	[41, 42, 43]
Puls	Université d'Helsinki	Finlande	Restreint	Infectieux	Tout public	[44, 45,46,]
RSOE-EDIS	RSOE*	Hongrie	Public	Catastrophe naturelle Risques industriels	Tout public	https://hisz.rsoe.hu
EpiSPIDER	EpiSPIDER	USA	Public	Infectieux	Tout public	[47, 48]
VERATEC	Veratect Corporation	USA	Payant	Infectieux	Secteur privé	http://www.veratect.com/

* RSOE = national association of radio distress-signaling & info communications, EDIS=emergency and disaster information service

Leur développement résulte d'initiatives indépendantes, leurs modalités d'accès sont très différentes et leur pérennité n'est pas garantie. Ainsi, fin 2010, deux systèmes avaient cessé de fonctionner par manque de clients (Veratect) ou d'être maintenu (EpiSPIDER). Si pour la moitié des systèmes, l'accès à la plateforme Internet où sont regroupées les informations brutes est librement accessible, l'accès à leur méthodologie et aux informations-systèmes sont restreints. De plus, les modalités de recherches rétrospectives sont souvent limitées et les données brutes ne sont pas toujours exportables dans un format compatible avec les analyses.

Afin de déterminer les systèmes qui entreraient dans l'étude, la première étape a été d'identifier un cadre institutionnel permettant un accès à l'ensemble des données durant la période d'étude. C'est le projet Early Alerting and Reporting (EAR) qui a fourni ce cadre opérationnel. Par ailleurs, afin d'assurer une comparabilité des résultats, les systèmes pour lesquels la détection des événements sanitaires d'origine infectieuse n'était pas un objectif prioritaire ont été exclus.

II.3.1 LE PROJET EAR

La « Global health security initiative » (GHSI)⁴⁹, ou initiative mondiale pour la sécurité sanitaire a été fondée en novembre 2001 à la suite des attentats du 11 septembre par l'Allemagne, le Canada, la France, les États-Unis, l'Italie, le Japon, le Mexique, le Royaume-Uni et l'Union européenne, l'OMS est associée en tant qu'expert. La GHSI est un partenariat international informel visant à renforcer la préparation et la réponse aux menaces sanitaires liées à l'utilisation intentionnelle d'agents NRBC et aux risques liés à la grippe pandémique.⁵⁰ C'est dans ce cadre, qu'en 2009, face au manque de collaboration internationale, le projet EAR a été lancé afin d'évaluer la faisabilité de développer un outil commun spécifiquement dédié à l'intelligence épidémiologique. EAR est une initiative qui rassemble, les utilisateurs (à savoir des instituts de santé publique en charge de l'intelligence épidémiologique), des institutions et centres de recherches à l'origine du développement des systèmes de biosurveillance et les ministères de la santé des pays concernés.

L'objectif de la première phase de EAR (2009-2010) était d'évaluer la faisabilité et la pertinence du développement d'une plateforme web unique qui, d'une part, permettrait d'optimiser la détection d'éventuelles menaces sanitaires à partir des informations produites par les systèmes de biosurveillance, et d'autre part, de mutualiser le processus d'évaluation des risques entre les différents pays participants. Le projet était constitué de 7 groupes de travail, dont un portant sur l'évaluation des systèmes de biosurveillance a été confié à l'InVS (Tableau 4). EAR est autofinancé par les pays et les institutions participantes.

Tableau 4 - Participants au projet EAR

Systèmes de biosurveillance et centres de recherche		
Nom	Développeur	Pays
ARGUS	Université de Georgetown	États-Unis
BioCaster	National Institute of Informatics	Japon
GPIN	Agence de santé publique du Canada	Canada
HealthMap	Children's Hospital Boston	États-Unis
MedISys	Joint Research Centre	Union Européenne (UE)
ProMED-mail	International Society for Infectious Diseases.	Italie
PULS	Université d'Helsinki	Finlande
Instituts de Santé publique		
Acronyme	Nom	Pays
RKI	Institut Robert Koch	Allemagne
ASPC	Agence de la santé publique du Canada	Canada
CDC	Center for Disease prevention and Control	USA
InVS	Institut de Veille Sanitaire	France
HPA	Health Protection Agency	Royaume Uni
ISS	Istituto Superiore di Sanità	Italie
NIID	Institut national des maladies infectieuses	Japon
ECDC	Centre européen de prévention et de contrôle des maladies	UE
Décideurs		
Ministères de la Santé Allemagne, Canada, États-Unis, France, Italie, Japon, Mexique et Royaume-Uni ; Direction générale de la Santé et des Consommateurs de l'Union Européenne (DG SANCO) ; Organisation mondiale de la Santé (Observateur)		

II.3.2 CARACTERISTIQUES DES SYSTEMES DE BIOSURVEILLANCE

Cette section présente les systèmes de biosurveillance qui ont été utilisés dans les deux études (cf. résultats). Les principales caractéristiques sont présentées dans le Tableau 3. Les systèmes sont présentés individuellement en Annexe 3.

Tableau 5 - Principales caractéristiques des systèmes de biosurveillance inclus dans l'étude, bilan réalisé en 2010⁵¹

Systèmes	ARGUS	BioCaster	GPHIN	HealthMap	MedISys	ProMED	PULS
Caractéristiques générales							
Mode de gestion des données (Automatique ou Modéré*)	M	A	M	A	A	M	A
n moyen de documents par jour	150	90	2000	360	>1000	7	150
n langues sources	40	13	9	7	50	5	2
n langues disponibles sur l'interface	1	13	9	7	50	5	2
Langues disponibles (par traduction humaine ou interface modifiable #)	An	Non	An, Ar, C-T C-S, Es, Fr, Po, Pe, Ru,	Non	Non	An, Es, Fr, Po, Ru,	An, Fn
Type de risque (Humain, Végétal, Animal, Environnemental)	H.V.A.E	H.V.A.E	H.V.A.E	H.V.A.E	H.V.A.E	H.V.A	H.V.A
Thématique (NRBC, Syndromique)	B	B.S	NRBC.S	B.S	NRBC. S	BC.S	B
Couverture géographique	Monde	Monde	Monde	Monde	Monde	Monde	Monde
Processus							
Flux d'acquisition des données (Automatique / Manuel £)	A/M	A	A/M	A/M	A	M	A
Méthodologie de sélection : Taxonomie, Ontologie, Manuel, Autre	T,O	A	T,O,A	T,O,A	O,A	M	MedISys
Déduplication (Automatique / Manuel)	M	A	A et M	A	A	M	A
Evaluation de risque	Oui	Non	Oui	Oui	Non	Oui	Non
Validation du contenu des informations	Oui	Non	Oui	Oui	Non	Oui	Oui
Analyse humaine	Oui	Non	Oui	parfois	Non	Oui	Non
Méthodologie de recherche: mots clés (1), texte libre (2), numéro de référence (3), autres (4)	1.2	1.2	1. 2. 3. 4	1.2.3	1.2.3	1.2.3	1.2.3
Archivage	Oui	Non	Oui	Oui	Oui	Oui	Oui
Communication de l'information utilisateurs (hors accès portail Internet)	Email	Email, flux RSS, twitter	Email,	Email, flux RSS, twitter	Email, flux RSS, sms	Email, flux RSS, twitter	Flux RSS
Mode opératoire	Oui	Oui	Oui	Oui	Non	Oui	Non
Fonctions supplémentaires							
Intervention humaine	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Interface personnalisable	Oui	Non	Oui	Oui	Oui	Non	Oui
Carte interactive	Oui	Oui	Non	Oui	Oui	Oui	Oui
Données du trafic aérien	Oui	Non	Non	Non	Non	Non	Non
Rapports spéciaux	Non	Non	Oui	Oui	Oui	Oui	Non
Graphiques	Non	Oui	Non	Non	Oui	Non	Oui
Alerte email personnalisable	Oui	Oui	Oui	Oui	Oui	Non	Non
Analyse statistique	Non	Oui	Non	Oui	Oui	Non	Oui

* Modération = intervention humaine (analystes) dans les processus

An= Anglais ; AR= Arabe ; C-T=Chinois traditionnel, C-S=Chinois simplifié ; Es=Espagnol, Fr=Français ; Po=Portugais ; Pe=Persan ; Ru=Russe.

£ Sélection manuelle effectuée par des modérateurs (pas de processus automatisé)

II.4 LES INDICATEURS

II.4.1 REVUE DE LA LITTÉRATURE

Une revue de la littérature a été effectuée afin d'identifier les indicateurs déjà utilisés dans la littérature pour l'évaluation des outils d'intelligence épidémiologique utilisant Internet comme source d'information. Les documents se rapportant à l'intelligence épidémiologique et utilisant des méthodes statistiques standards et des attributs qualitatifs ont été systématiquement recherchés et les résultats des évaluations ont été résumés. Les études ont été recueillies à partir de la littérature scientifique et d'autres sources (rapports techniques, des résumés, des actes de conférence, etc.).

Un total de 31 documents scientifiques relatifs au processus d'évaluation des outils d'intelligence épidémiologique a été identifié : 10 enquêtes portant sur la perspective des utilisateurs (évaluation extrinsèque), 17 enquêtes de performance des systèmes (évaluation intrinsèque) et 4 études combinant ces deux aspects de l'évaluation. Les performances intrinsèques (classification de texte et la capacité de géocodage) étaient mesurées avec des paramètres connus: la précision (« precision », la sensibilité (« recall rate » ou « sensitivity ») ou, exactitude « accuracy », le F-score, et la spécificité. Les évaluations extrinsèques étaient principalement axées sur la mesure de la réactivité (« timeliness »), la sensibilité et l'exactitude (« accuracy »).

En conclusions, les deux modes d'évaluation - extrinsèques et intrinsèques - sont essentiels pour le développement de la technologie et pour maintenir en adéquation les besoins des utilisateurs et les informations produites par les systèmes. Néanmoins, d'importants défis subsistent en termes d'évaluation d'autant que ces systèmes doivent s'adapter à l'évolution d'Internet. Les fournisseurs de systèmes ont effectué des évaluations régulières afin d'évaluer la performance de leurs systèmes, ils ont également exprimé la nécessité de l'implication des utilisateurs dans les processus d'évaluation. Plusieurs axes doivent être développés pour améliorer les processus d'évaluation : l'élaboration d'un gold-standard, l'utilisation de mesures standardisées, l'utilisation combinée des évaluations intrinsèques et extrinsèques, en utilisant une approche d'évaluation multi-systèmes, et en insistant sur l'utilité pour la santé publique (cf. Annexe 2 – Article B, p 88).

II.4.2 INDICATEURS DE LA SURVEILLANCE DES MALADIES TRANSMISSIBLES

Dans son guide sur l'évaluation et le monitoring des systèmes de surveillance des maladies transmissibles et de réponse⁵², l'OMS classe les indicateurs utilisés pour l'évaluation des systèmes de surveillance en cinq catégories :

- **Les indicateurs d'intrant** (« inputs ») sont utilisés pour évaluer les ressources nécessaires pour mettre en œuvre le système (ressources humaines, financières, matérielles, logistiques, etc.).
- **Les indicateurs de processus** (« process ») sont utilisés pour estimer et suivre la mise en œuvre des activités essentielles à la surveillance (formation, supervision, procédures, outils, etc.).
- **Les indicateurs de produit** (output) sont les mesures des résultats immédiats des activités (rapports issus des données de surveillance, la rétro-information aux fournisseurs de données, proportion de personnel de santé formé, proportion des visites de supervision prévues et réalisées, etc.).
- **Les indicateurs de résultat** (« outcome ») sont des mesures de la qualité du système de surveillance. Ils peuvent inclure des indicateurs permettant d'évaluer l'utilité du système, l'exhaustivité de l'information, l'utilisation des données de surveillance dans le cadre des politiques de santé et la gestion des programmes ou la pertinence de la réponse aux épidémies.
- **Les indicateurs d'impact** (« impact ») mesurent la manière dont les objectifs globaux de la surveillance sont atteints (évolution de la morbidité et/ou de la mortalité de maladies à potentiel épidémique, changements dans les comportements liés à la santé de la population cible, etc.).

Les indicateurs d'intrant et de processus sont surtout utilisés lors de la mise en œuvre des systèmes de surveillance, alors que l'importance des indicateurs de résultat et d'impact augmente dès lors que ces systèmes se stabilisent.

II.4.3 INDICATEURS DE RESULTATS

Dans le cadre de l'évaluation des performances des systèmes de biosurveillance, les indicateurs de résultats sont les plus pertinents. Une partie de ces indicateurs est adaptable pour l'évaluation globale d'un dispositif d'intelligence épidémiologique et pour l'évaluation des systèmes de biosurveillance. En revanche, très peu d'études ont été menées pour adapter et mesurer l'ensemble de ces indicateurs et leur utilisation dans ce cadre n'est pas formellement définie.

En raison de leur potentiel épidémique et de la disponibilité en mesures de contrôle, les systèmes d'alerte précoce consacrent une large place aux maladies transmissibles. Bien que les domaines d'application de l'intelligence épidémiologique et des systèmes de biosurveillance ne soient pas limités à ce champ, c'est l'évaluation des systèmes de surveillance des maladies infectieuses qui présente le plus d'analogie. D'autre part, si certains systèmes de biosurveillance intègrent également d'autres types de risques sanitaires (chimiques, radionucléaires), les maladies transmissibles sont un objectif prioritaire et constituent un dénominateur commun entre tous les systèmes intégrés dans notre étude.

Les indicateurs utilisés par l'OMS et le CDC pour l'évaluation des systèmes de surveillance des maladies infectieuses ont été revus. Parmi, tous ces indicateurs, les plus pertinents du point de vue de l'utilisateur ont été retenus, il s'agit de la représentativité, la complétude, la réactivité, la sensibilité, l'utilité, la simplicité, la flexibilité et la valeur prédictive positive (VPP). Ces indicateurs ont ensuite été adaptés au contexte de l'évaluation des systèmes de biosurveillance (Tableau 6).

On notera que la spécificité n'a pas été retenue comme un indicateur potentiel de l'évaluation des systèmes de biosurveillance. En effet, pour pouvoir estimer la spécificité d'un système de biosurveillance, il faudrait pouvoir estimer le nombre de vrais négatifs. C'est-à-dire le volume des informations disponibles sur Internet et traitées par les systèmes qui ne correspondaient pas à un potentiel « signal brut ». En pratique, il est impossible de déterminer le nombre de vrais négatifs d'une part en raison du volume considérable d'information et d'autre part car ces informations ne sont ni stockées par les systèmes ni considérées par les utilisateurs.

Tableau 6 - Principaux indicateurs d'évaluation des systèmes de surveillance des maladies infectieuses et leur application à l'évaluation de l'intelligence épidémiologie et des systèmes de biosurveillance.

	Définition des indicateurs utilisés pour l'évaluation des systèmes de surveillance (des maladies infectieuses) par l'OMS et le CDC *		Application à l'évaluation des systèmes de biosurveillance
Représentativité	OMS †	Capacité du système à décrire avec précision la survenue d'un événement sanitaire en termes de temps, lieu et personne.	La représentativité peut être définie en termes de i) couverture géographique (mondiale ou par région) ; ii) Sujets: maladies infectieuses, zoonoses, risques chimiques, etc.
	CDC ‡	Lorsque la détection des cas dans une population est incomplète, la représentativité reflète si un système décrit avec précision la répartition des cas en termes de temps, lieu et personne.	
Complétude	OMS †	La complétude concerne l'intégralité de la surveillance et peut avoir des dimensions variables telles que: complétude des sites de surveillance, des rapports de surveillance, de la déclaration obligatoire des cas, des données de surveillance...	Complétude peut être définie comme la fidélité de l'information produite par les systèmes de biosurveillance : i) les informations essentielles contenues dans la source d'origine sont conservées par les systèmes ii) des données concernant un même événement sont correctement collectées et groupées
	CDC ‡	La complétude est liée à la qualité des données elle est définie comme La fréquence des données « inconnues »	
Réactivité	OMS †	La réactivité reflète la vitesse entre les étapes d'un système de surveillance de santé publique. La mesure la plus importante vise à déterminer si les données sont soumises à temps pour initier les enquêtes et mettre en œuvre des mesures de contrôle.	La réactivité peut être définie comme l'intervalle de temps entre la détection d'un événement par un système de biosurveillance et la notification de ce même événement aux autorités sanitaires en charge de la mise en œuvre des mesures de contrôle
	CDC ‡	La réactivité de la détection des épidémies est mesurée par la période de temps entre l'exposition à l'agent pathogène et l'initiation d'une intervention de santé publique.	
Sensibilité	OMS †	La sensibilité correspond à la proportion de cas réels dans une population qui sont détectés et notifiés par le système. Elle se décline en : i) sensibilité de la définition de cas ; ii) sensibilité de la détection des événements pour les interventions de santé publique ; iii) sensibilité du système de notification	La sensibilité peut être définie comme la proportion d'événements sanitaires (confirmés) détectée par un système de biosurveillance
	CDC ‡	La sensibilité est le pourcentage d'épidémie se produisant dans une juridiction et détecté par le système de surveillance.	
VPP *	OMS †	La proportion de personnes que le système classe comme ayant une pathologie qui présente effectivement cette pathologie	La VPP peut être définie la probabilité qu'un signal détecté par un système corresponde à un événement réel (proportion d'événements parmi les signaux).
	CDC ‡	La probabilité qu'un signal détecté corresponde à une épidémie	
Utilité	OMS †	Capacité du système de surveillance à atteindre l'objectif (s) pour lequel il a été conçu.	L'utilité peut être définie comme i) La capacité du système de biosurveillance à collecter des données appropriées aux besoins de l'utilisateur: pertinence et/ou exactitude de l'information ii) L'intérêt des outils qui sont disponibles au sein des systèmes.
	CDC ‡	L'utilité est la capacité de contribuer à la détection d'événements qui pourrait affecter la population. Un système de surveillance peut également être utile s'il permet de déterminer qu'un événement que l'on croyait sans importance est réellement important	
Simplicité	OMS †	Simplicité se réfère à la structure du système de surveillance et sa facilité de mise en œuvre, la simplicité de la circulation de l'information de son point de production à l'utilisateur final doit être envisagée	La simplicité peut être définie comme la structure et la facilité d'utilisation du système de biosurveillance (utilisation quotidienne, l'utilisation des fonctions spécifiques, la présence module d'aide de l'utilisateur).
	CDC ‡	La structure et la facilité d'utilisation du système reflètent la simplicité du système de surveillance	
Flexibilité	OMS †	La flexibilité désigne la capacité du système à s'adapter à l'évolution des besoins tels que la suppression ou l'inclusion d'autres maladies, la modification de la fréquence des rapports, des besoins, etc.	La flexibilité peut être définie comme i) la mesure des capacités des systèmes de biosurveillance à détecter de nouveaux risques sanitaires; ii) la capacité à s'adapter aux besoins des utilisateurs (personnalisation).
	CDC ‡	La flexibilité d'un système de surveillance se réfère à la capacité du système à changer à mesure que les besoins changent.	

* VPP : Valeur Prédictive Positive

† **Adapté de :** Organisation Mondiale de la santé, Communicable disease surveillance and response systems; Guide to monitoring and evaluating. WHO/CDS/EPR/LYO (2006)⁵² http://whqlibdoc.who.int/hq/2006/WHO_CDS_EPR_LYO_2006_2_eng.pdf

‡ **Adapté de:** Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V; CDC Working Group. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. MMWR Recomm Rep. 2004 May 7; 53(RR-5):1-11.⁵³ <http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5305a1.htm>

II.5 STRATEGIES D'INTELLIGENCE EPIDEMIOLOGIQUE

II.5.1 DIVERSITE DES STRATEGIES

Afin de documenter la diversité des stratégies d'intelligence épidémiologique, une enquête a été menée auprès de neuf institutions (CDC, CE, ECDC, HPA, InVS, ISS, NIID, ASPC et RKI) pour collecter des informations détaillées sur le type d'activité mis en place. Les principales rubriques portaient sur : les domaines d'intérêt, les moyens dédiés à l'intelligence épidémiologique, les modalités de sélection des signaux bruts, les procédures de validation, le type de production, et les actions générées par la détection d'une alerte. Les 10 utilisateurs-répondants étaient soit les responsables de l'unité d'intelligence épidémiologique soit des épidémiologistes en charge des activités d'intelligence épidémiologique au sein de leur institution.

II.5.2 IMPACT DES STRATEGIES D'INTELLIGENCE EPIDEMIOLOGIQUE

Aux vues de la diversité des stratégies mise en place par les institutions pour répondre à leurs objectifs spécifiques, il était nécessaire de déterminer dans quelle mesure les objectifs et la stratégie affectent le type d'évènements détectés et rapportés, le type de support et la fréquence de publication.

L'évaluation et la comparaison de différents dispositifs d'intelligence épidémiologique sont des domaines de recherche à part entière qui dépassent largement le cadre de cette thèse. Néanmoins, afin d'estimer l'impact, les productions résultant des activités d'intelligence épidémiologique menées en 2012 par quatre institutions internationale (OMS), régionales (ECDC et EpiSouth*) et nationale (InVS) ont été comparées. Pour estimer dans quelle mesure les objectifs et la stratégie de ces institutions affectent le type d'évènements détectés et rapportés, le type de support et la fréquence de publication, les données concernant leurs productions ont été comparées. Pour permettre une comparaison des résultats, seuls les supports publics de communication (accessible en ligne) et dédiés à l'intelligence épidémiologique ont été pris en considération. Les évènements rapportés dans ces différents supports ont été classés en fonction de leur lieu de survenue (pays et/ou province), de la pathologie et du type d'information : nouvel évènement (nouvel évènement ou nouvel élément épidémiologique pour un évènement déjà communiqué), bilans systématiques (mis à jour systématiquement qu'il y ait ou non de nouveaux éléments épidémiologiques) par pathologie ou type d'évènement (ex : grands rassemblements).

* **Note sur EpiSouth.** EpiSouth était un réseau cofinancé par la Commission Européenne et le ministère italien de la santé. Son objectif était de consolider la sécurité sanitaire autour du bassin méditerranéen par la création d'une plateforme de travail dans le domaine de l'épidémiologie afin de renforcer la surveillance des maladies transmissibles, la communication et la formation. 27 pays méditerranéens (Europe du Sud, Balkans, Maghreb et Moyen-Orient) y participaient. Un programme de travail, intitulé « intelligence épidémiologique et alertes transfrontalières », avait été confié à l'InVS et était piloté par le département international (DI)⁵⁴. Il avait pour objectif i) la détection et la caractérisation des alertes internationales pouvant avoir un impact de santé publique pour les pays EpiSouth ii) l'échange d'alertes nationales à potentiel transfrontalier. L'intelligence épidémiologique EpiSouth répondait à la même méthodologie que la veille internationale réalisée par le DI de l'InVS. Cependant, les critères de sélection des évènements étaient différents et les activités étaient réalisées par du personnel dédié au projet EpiSouth. Les deux équipes VI France et EpiSouth travaillaient en étroite collaboration et synergie et sous une même coordination. Les informations de l'intelligence épidémiologique étaient diffusées (jusqu'à Avril 2013) en anglais sur un bulletin hebdomadaire d'alerte (eWEB: EpiSouth Weekly Epidemiological Bulletin) disponible sur les sites EpiSouth.^{55,56} Faute de financements adéquats EpiSouth a cessé d'être opérationnel en janvier 2014.

II.6 EVALUATION SYSTEMES DE BIOSURVEILLANCE (ARTICLE 1)

L'objectif de cette étude, composée d'une partie qualitative et d'une partie quantitative, était d'évaluer les principaux indicateurs de performance des systèmes de biosurveillance et d'estimer la complémentarité des systèmes de biosurveillance.

Etude qualitative. Un questionnaire a été élaboré pour mesurer la perception des utilisateurs, les points focaux du projet EAR, vis-à-vis des performances des sept systèmes de biosurveillance intégrés dans le projet. Sept indicateurs préalablement définis ont été retenus : la complétude, la flexibilité, la réactivité, la représentativité, la sensibilité, la simplicité et l'utilité.

Etude quantitative:

Gold-standard :

Une revue a été effectuée pour identifier une menace sanitaire mondiale, survenant fréquemment, couverte par tous les systèmes de biosurveillance et pour laquelle un gold standard existe. C'est l'influenza aviaire A/H5N1 qui a été sélectionné comme le plus approprié. Les cas humains et vétérinaires sont soumis à la notification obligatoire à l'OMS ou à l'OIE). Ces deux organisations internationales constituent la source d'informations la plus fiables concernant la survenue de cas humains et d'épizooties A/H5N1 biologiquement confirmés. Les cas humains et les épizooties (cas isolés ou épidémies plus importantes) survenus en mars 2010 (date des premiers symptômes, date du début de l'épizootie) ou rapporté par l'OMS / OIE en Mars 2010 ont constitué le gold-standard.

Analyse quantitative et indicateurs

Deux bases de données ont été créées: la première "prospective" afin d'évaluer le processus de détection précoce dans une situation proche des conditions réelles d'utilisation (i.e., détection de signaux potentiellement pertinents dans un large volume d'information brute). La seconde base de données "rétrospective" avait pour objectif l'évaluation des performances théoriques des systèmes. La base de données "prospective" a été utilisée pour calculer les taux de détection et les valeurs prédictives positives (VPP) et la base de donnée "rétrospective" pour calculer la sensibilité et la réactivité.

II.7 FACTEURS DETECTION EVENEMENTS INFECTIEUX (ARTICLE 2)

Cette étude vise à documenter les facteurs influençant les performances des systèmes de biosurveillance en matière de détection des événements sanitaires d'origine infectieuse.

Gold-standard : les événements sanitaires détectés et vérifiés par l'équipe de la veille internationale de l'InVS sont publiés chaque mercredi dans le Bulletin Hebdomadaire International (BHI), accessible en ligne sur le site de l'InVS⁵⁷. Les événements ne sont en principe rapportés qu'une seule fois, des mises à jour sont néanmoins intégrées en cas de changements épidémiologiques majeurs. Tous les événements d'origine infectieuse (sauf les cas humains et aviaires de grippe A/H5N1) publiés dans le BHI en 2010 ont été inclus dans le gold-standard.

Analyse quantitative et indicateurs : Six systèmes de biosurveillance ont été utilisés pour détecter les signaux bruts (Argus, BioCaster, GPHIN, HealthMap, MedISys et ProMED). Les signaux en relation avec les événements inclus dans le gold-standard ont été recherchés rétrospectivement sur les six systèmes au moyen de requêtes *ad hoc* utilisant des mots-clés ou des séries de mots-clés. Les recherches ont été effectuées par deux épidémiologistes indépendants des six systèmes. Les paires discordantes ont été analysées et la plus pertinente a été conservée dans la base de données finale. Les taux de détection-brute (C-DR), taux de sensibilité brute (C-SE) et le taux de sensibilité intrinsèque (I-Se) ont été utilisés comme indicateurs pour évaluer les performances des systèmes. Les régressions de Poisson modifiées (bi et multivarié) adaptées aux événements rares et aux données binaires ont été utilisées pour évaluer les facteurs influençant la détection. Les taux de ratios (RR), les

intervalles de confiance à 95% (IC) et les valeurs de p ont été utilisés pour quantifier la force de ces associations. Les mêmes variables ont été incluses dans les différents modèles de régression. Les interactions potentielles ont été testées pour leur importance à 0,05.

II.8 REACTIVITE DETECTION ET SENSIBILITE SYSTEMES (NON-PUBLIEES)

Dans le cadre de la même étude (Article - 2), des données complémentaires ont été recueillies pour permettre de mesurer leurs réactivités et leurs sensibilités en termes de détection précoce et vis-à-vis des activités d'intelligence épidémiologique.

Pour chacun des 86 évènements inclus dans le gold-standard (BHI), les signaux bruts ont été recherchés rétrospectivement sur six systèmes de biosurveillance et indépendamment par deux épidémiologistes. Les paires discordantes ont été analysées et la plus pertinente a été conservée dans la base de données finale. Pour les 472 signaux bruts détectés les dates de détection, de parution dans le BHI et de survenue de l'évènement ont été recueillies. A partir de ces dates, trois mesures de réactivité ont été calculées pour chaque système et pour un système virtuel combiné (qui résulterait de l'intégration des signaux détectés par tous les systèmes). Les trois réactivités ont été calculées pour mesurer : le délai entre la publication dans le BHI et la détection par le système, le délai entre la détection de l'évènement par l'équipe de la VI de l'InVS et la détection par le système et le délai entre la survenue de l'évènement et la détection par le système. Enfin, deux sensibilités ont été calculées : la sensibilité de détection vis-à-vis de l'équipe de la VI (Se-VI) et la sensibilité de détection précoce (Se-DP).

Tableau 7 – Définitions des indicateurs et des mesures de réactivité

Indicateurs	Définitions
Date de détection-système	Date de détection par un système du premier signal correspondant à un évènement inclus dans le gold-standard
Date parution-BHI	Date de parution (mise en ligne) du BHI rapportant l'évènement en question
Date VI-InVS	Date à laquelle l'équipe VI a détecté le signal pour la 1 ^{ère} fois dans les conditions de normale de routine (en recherche prospective)
Date survenue	Date de survenue déterminée en fonction des données disponibles pour chaque évènement : date de détection du premier cas, date de début des signes, date d'hospitalisation, date de diagnostic, date de décès ou date de notification
Réactivités	Définition
Détection-InVS* – Détection-système	Le délai entre la publication dans le BHI et la détection par le système (date BHI moins date détection-système, en jour)
Publication- BHI – Détection-système	Le délai entre la détection de l'évènement par l'équipe VI et la détection par le système (date VI-InVS moins date détection-système, en jour).
Survenue-évènement – Détection-système #	Le délai entre la survenue de l'évènement et la détection par le système (date de survenue moins date détection-système, en jour).
Sensibilités	Définition
Sensibilité VI (Se-VI)	Capacité d'un système à détecter un signal avant ou le même jour que la détection de l'évènement par l'équipe de la VI.
Sensibilité DP (Se-DP)	Capacité d'un système à détecter un signal au moins un jour avant la détection de l'évènement par l'équipe de la VI

* Parmi les 73 évènements (85%) pour lesquels la date de détection de l'évènement était disponible

La date de survenue de l'évènement en fonction des données disponibles (dates de détection du premier cas, de début des signes, d'hospitalisation, de diagnostic, de décès, ou de notification)

III. RESULTATS

III.1 STRATEGIES D'INTELLIGENCE EPIDEMIOLOGIQUE ET IMPACTS

III.1.1 DIVERSITE DES STRATEGIES

Les données recueillies permettent de décrire la diversité des modalités de l'intelligence épidémiologique mise en œuvre par les institutions participant au projet EAR. Ces différences, concernent à la fois le type d'évènements recherchés par les utilisateurs (Tableau 8), le processus de sélection et de validation des signaux bruts, les modes de communication (Tableau 9) et le mode d'utilisation des systèmes de biosurveillance (Tableau 10). Les résultats illustrent l'importance et l'hétérogénéité des critères de sélection. Les domaines couverts sont extrêmement vastes : pathologies humaines, animales mais également végétales (pour une institution). Si les risques infectieux sont prédominants, les risques chimiques radionucléaires et environnementaux sont couverts par la plupart des institutions participantes. En revanche, l'approche NRCB (bioterrorisme), n'était pas privilégiée par les participants. Les critères de sélection sont extrêmement variables, et n'étaient formalisés que pour 6 des 10 répondants.

Certains critères de gravité (comme la sévérité des signes cliniques, la létalité, le mode de transmission) étaient quasiment systématiquement considérés. L'utilisation d'autres critères était très hétérogène. Par exemple, la moitié des utilisateurs incluaient dans leurs critères les produits toxiques, les risques environnementaux, les pays de forte immigration ou émigration, 70% incluaient la grippe pandémique et plus de 80 % les risques globaux (SRAS, grippe aviaire) et les pathologies à risque d'importation. L'alerte (des autorités sanitaires compétentes) est l'objectif prioritaire de toutes les institutions participantes et la moitié des répondants étaient impliqués directement ou indirectement dans la mise en œuvre des mesures de contrôle (Tableau 8)

Globalement, les utilisateurs utilisaient entre 4 et 7 systèmes en routine. ProMED était le système le plus utilisé (9/10 répondants), Argus et PULS les moins utilisés (5/10), cette faible utilisation peut être expliquée par l'accès très restreint d'Argus et le caractère encore expérimental de PULS. Les modalités d'utilisation des systèmes étaient très variables. Les e-mails générés automatiquement par les systèmes et les recherches manuelles définies par les utilisateurs étaient les plus utilisées (Tableau 9).

Tableau 8 - Domaines couverts en routine par les unités d'intelligence épidémiologique des institutions participant au projet EAR en 2010 (n=10)

Domaines d'intérêt	Sous-domaines	Tous	Événement retenu / évalué			
			La plupart	Suivant critères	Rarement	Aucun
Événements biologiques	Toute pathologie épidémique	20%	0%	40%	10%	30%
	Pathologies spécifiques	30%	23%	20%	10%	7%
	Pathologies humaines	20%	33%	20%	0%	17%
	Zoonoses	10%	3%	30%	10%	17%
	Pathologies végétales	0%	0%	10%	10%	50%
	Agents NRBC	50%	0%	30%	10%	0%
Événements chimiques	Tout évènement chimique	10%	0%	30%	20%	40%
	NRBC uniquement	20%	0%	0%	10%	50%
	Événements spécifiques*	0%	3%	0%	10%	67%
Événements radiologiques / nucléaires	Tout évènement	10%	0%	30%	20%	30%
	NRBC uniquement	20%	0%	10%	0%	60%
	Événements spécifiques*	0%	0%	10%	10%	60%
Désastres naturels	Inondations	10%	0%	30%	30%	30%
	Tremblements de terre	10%	0%	30%	20%	40%
	Activité volcanique	10%	0%	30%	20%	40%
	Autres désastres naturels	0%	0%	40%	30%	30%

* Événements spécifiques retenus par des d'autres critères non-spécifiquement NRBC (ex : géographique,...)

Tableau 9 – Utilisation des systèmes de biosurveillance par les utilisateurs, Projet EAR 2010 (n=10)

	Argus	BioCaster	GPHIN	HealthMap	MedISys	ProMED	PULS	Autres
Utilisateurs	5/10	6/10	6/10	7/10	8/10	9/10	5/10	6/10
<i>Dont usage quotidien</i>	5/5	4/6	6/6	3/7	4/8	9/9	2/5	3/6
Type d'utilisation (10 répondants)								
Alerte	4	2	6	3	4	6	2	5
Rapports	2	1	2	3	1	3	1	3
Surveillance	3	4	4	3	5	6	1	4
Sujets suivis spécifiquement (10 répondants)								
Maladies transmissibles	5	5	6	6	7	9	2	3
Chimique	1	0	4	0	3	3	0	0
Radio- Nucléaire	1	0	4	0	3	2	0	1
Type d'alerte-système utilisé (10 répondants)								
e-mail automatique	5	0	4	0	1	6	0	2
Flux RSS	0	1	1	0	3	2	1	2
Recherches manuelles	2	3	4	4	5	5	2	3
Alerte e-mail personnalisée	3	0	2	0	2	1	0	1
Aucune	5	5	3	4	2	2	5	1

Tableau 10 – Processus d'intelligence épidémiologique, institutions participants au projet EAR en 2010 (n =10)

Moyens dédiés	Equipe spécifique dédiée à l'intelligence épidémiologique (IE)		50%
	Nombre de personnes dédiées spécifiquement à l'IE (moyenne/équipe)		2,7
Sélection des signaux bruts	Critères spécifiques de sélection des signaux		60%
	Critères géographiques et critères de risque	Zone géographique limitrophe (proche) du pays	60%
		Pays de forte immigration ou émigration	50%
		Zones touristiques	40%
		Grands rassemblements	60%
		Pas de zone spécifique	20%
		Autres	60%
		Menaces globales (SRAS, grippe aviaire)	90%
	Nature de l'évènement	Risques d'importation	80%
		Risques NRBC spécifiques	90%
		Grippe pandémique	70%
		Toxiques	50%
		Pollution	20%
		Menaces environnementale sur la santé	50%
Produits (médicaments, nourriture, effets secondaires)		60%	
Mode de transmission		100%	
Sévérité		90%	
Létalité		80%	
Critères de gravité	Morbidité	80%	
	Emergence	90%	
	Changement épidémiologique	90%	
	Distribution géographique	100%	
	Procédures Standardisées	70%	
Modalité de validation	Type	Validation en groupe	80%
		Validation individuelle	20%
	Moyens	Réseau	80%
		Personne dédiée	70%
	Nombre de personnes impliquées dans la validation (moyenne/ équipe)		4,5
	Possibilité de partage du réseau		50%
	Base de référence pour la vérification		30%
Evaluation du risque NRBC		50%	
Rapports	Diffusion des rapports	Accès limité	90%
		Accès libre	40%
	Type de rapport	<i>Ad hoc</i>	70%
		Bulletin hebdomadaire	30%
Action	Bulletin journalier		30%
	Rapport		70%
	Alerte		100%
	Déploiement d'équipe sur le terrain		50%
	Coordination de la réponse		50%
Echange international		60%	

III.1.2 IMPACT DE LA STRATEGIES ET DES OBJECTIFS

III.1.2.1 Caractéristique activités d'intelligence épidémiologique

Les principales caractéristiques des activités d'intelligence épidémiologique d'intelligence menées en 2012 par les quatre institutions internationale (OMS), régionales (ECDC et EpiSouth) et nationale (InVS) sont résumées dans le Tableau 11.

Tableau 11 – Caractéristiques des activités d'intelligence épidémiologique OMS, ECDC, EpiSouth et InVS.

Institution	OMS	ECDC	EpiSouth*	InVS
Objectif	Global	Régional	Régional	National
Mandat	Mondial RSI (2005)	UE	Réseau (volontaire)	loi santé publique (2004)
Références	[58, 59]	[60, 61, 62]	[22, 63]	[64, 65, 66]
Couverture principale	Monde	28 pays UE	27 pays Méditerranée	France
Type de risque	Tous risques	Maladies transmissibles	Tous risques	Tous risques
Type de publication	<i>ad hoc</i>	Hebdo et <i>ad hoc</i>	Hebdo et <i>ad hoc</i>	Hebdo et <i>ad hoc</i>
Type de communication	Site web	Bulletin Site web	Bulletin Site web	Bulletin Site web
Type d'information publiée	Uniquement si accord pays	Presse et information validée	Information validée	Information validée
Stratégie communication	Nouveaux événements	Evènements + mises à jour+ suivis hebdo systématiques	Nouveaux événements	Nouveaux événements

III.1.2.2 Types d'évènements rapportés

L'OMS ne rapporte que des informations officielles concernant de nouvelles crises sanitaires, ou de nouveaux éléments les concernant. Les informations sont publiées au fur et à mesure de leur disponibilité (pas de périodicité définie)⁶⁷ mais seulement avec l'accord des pays concernés. En raison de cette contrainte, seule une petite partie des évènements ayant fait l'objet d'une évaluation de risque sont communiqués publiquement. De nombreuses autres données sanitaires, y compris concernant des épidémies, font l'objet d'une communication par l'OMS. Elles sont publiées sous d'autres supports dédiés à la surveillance épidémiologique (et pas spécifiquement à l'alerte précoce). On peut citer le Relevé Epidémiologique Hebdomadaire (REB-WER) (www.who.int/wer/fr/) ou le site dédié à la surveillance des épidémies de méningite dans 19 pays d'Afrique subsaharienne⁶⁸ qui réalise un bulletin hebdomadaire de retro-information durant toute la période épidémique. Ces supports de communication n'ont pas été intégrés dans l'étude.

Les modalités de communication des résultats de l'intelligence épidémiologique mise en œuvre par EpiSouth et l'InVS sont très proches : elles combinent un bulletin hebdomadaire ne contenant que des nouveaux évènements (ou les changements épidémiologiques majeurs) et des notes d'information complémentaires pour le suivi des crises complexes (pas de périodicité définie). Seuls les évènements vérifiés sont intégrés dans ces supports. Pour l'InVS, les évènements rapportés sont systématiquement vérifiés, leur inclusion dans le bulletin ne fait jamais l'objet d'une demande d'autorisation préalable auprès des autorités des pays concernés. Pour EpiSouth, tous les évènements sont systématiquement vérifiés, mais les informations concernant l'un des 27 pays méditerranéens participants au réseau font l'objet d'une demande d'autorisation auprès des pays participants.

Le bulletin de l'ECDC⁶⁹, est le support le plus complexe. Il intègre à la fois des éléments d'intelligence épidémiologique et de surveillance épidémiologique. D'un point de vue intelligence épidémiologique, c'est le seul qui intègre à la fois des événements vérifiés et des signaux bruts non-vérifiés (médias). C'est également le seul qui comporte des bilans systématiques pour certaines crises sanitaires (qu'il y ait eu ou non des éléments nouveaux au cours de la semaine), un bilan systématique pour certaines pathologies (dengue, chikungunya, rougeole, rubéole, polio, grippe aviaire A/H5N1, grippe saisonnière) et des bilans systématiques pour certains grands rassemblements (ex : les Jeux Olympiques Londres 2012).

Sur l'année 2012, le nombre d'évènement définis par un même lieu et une même pathologie, rapportés par l'OMS, EpiSouth et l'InVS sont comparables : en moyenne 128 par an (Tableau 12). En revanche, le nombre d'évènements rapportés dans le bulletin de l'ECDC était 5 fois plus élevé (651), et ce chiffre n'inclut pas les évènements rapportés dans les suivis hebdomadaires: dengue, chikungunya, grippe saisonnière, rubéole et ni ceux inclus dans les sections spécialement consacrées aux Jeux Olympiques de Londres ni à la coupe d'Europe de football 2012. De plus, parmi les évènements rapportés plus d'un tiers (229) portaient sur des mises à jour d'évènements précédemment rapportés (sans notion de changement épidémiologique majeur).

Tableau 12 – Caractéristiques des supports de rétro-information élaborés dans le cadre de l'intelligence épidémiologique par l'OMS, l'ECDC, EpiSouth et l'InVS en 2012.

Institution	OMS	ECDC	EpiSouth	InVS
Bulletin/site	DON ^(a) 67	CDRT ^(b) (c) (d) 69	eWEB ^(e) 55	BHI ^(f) 70
Périodicité	<i>ad hoc</i> (jour)	hebdomadaire	hebdomadaire	hebdomadaire
n documents postés	85	46	52	52
n épisodes rapportés	106	651	151	156
n mises-à-jour §	-	229	-	-
n épisodes retenus	101	406	138	147

^(a) Disease Outbreak News; OMS site Alerte et action au niveau mondial (GAR);

^(b) CDTR =Communicable disease threat report;

^(c) Pour le CDTR, données disponibles que du 6 février au 31 décembre 2012 ;

^(d) Les suivis hebdomadaires: dengue, chikungunya, grippe saisonnière, rubéole et N° spéciaux J.O. Londres et coupe Europe football (2012) ne sont pas inclus;

^(e) eWEB= EpiSouth weekly epidemiological bulletin;

^(f) BHI = Bulletin Hebdomadaire International.

La distribution géographique des évènements rapportés en 2012 est également très différente. La majorité des évènements rapportés par l'OMS sont concentrés en Afrique subsaharienne et ceux rapportés en Asie du Sud-Est ne concernent que des cas humains de grippe A/H5N1 (Figure 4). Les évènements rapportés par EpiSouth sont fortement concentrés dans le bassin méditerranéen et les zones limitrophes et ceux rapportés en Asie ne concernent que des cas humains et les épizooties de grippe A/H5N1 (Figure 5). Les évènements rapportés par l'ECDC sont fortement concentrés en Europe et notamment les pays de l'UE. Les évènements rapportés hors Europe concernent principalement la poliomyélite (Nigéria, Pakistan, Afghanistan), le West Nile (Amérique du Nord) et la grippe aviaire (Asie du Sud-Est) (Figure 6). Enfin, Les évènements rapportés dans le BHI par l'InVS sont distribués plus globalement, et principalement sur les zones proches de la métropole et des collectivités françaises d'outre-mer d'Amérique, de l'Océan Indien et du Pacifique (Figure 7).

Tableau 13 – Pathologies rapportées dans les supports de rétro information l'intelligence épidémiologique dans par l'OMS, l'ECDC, EpiSouth et l'InVS en 2012.

Catégorie	Pathologie	OMS DON ^(a)		ECDC CDRT ^{(b)(c)(d)}		EpiSouth eWEB ^(e)		InVS BHI ^(f)	
		n	%	n	%	n	%	n	%
Influenza	A(H3N2)v	n		11		1		5	
	A/H5N1 (humain)	n	33	23		32		35	
	A/H5N1 (animal)	n				14		13	
	Grippe saisonnière	n		1					
	Total	n	33 (33%)	35 (9%)		47 (34%)		53 (36%)	
Transmission vectorielle	Chikungunya					3		7	
	Dengue	1		10		10		22	
	Paludisme			12		8		11	
	Autres arboviroses					3		3	
	West Nile			135		43		8	
	Total	n	1 (1%)	157 (39%)		67 (49%)		51 (35%)	
Fièvres Hémorragiques Virales	Ebola		21			1		4	
	Lassa		1					1	
	Marburg		10		3	2		2	
	Fièvre Vallée du Rift		1			1		2	
	Fièvre Jaune		7					2	
	CCHF			1		2		1	
	Junin							1	
	Total	n	40 (40%)	4 (1%)		6 (4%)		13 (9%)	
Pathologies épidémiques et à prévention vaccinale	Choléra		6		3	1		5	
	Méningites à méningo.		5			4		9	
	Diphthérie							1	
	Rougeole				47	1			
	Polio				84				
Total	n	11 (11%)	134 (33%)		6 (4%)		15 (10%)		
Autres	Anthrax				12				
	MERS-CoV		10		9	5		7	
	Hantavirus		2		2			2	
	HFMD (EV-71)		4		1			1	
	Légionellose				7			1	
	Gastroentérites				6				
	Monkey Pox							1	
	Salmonelloses				20				
	Cryptosporidiose				5				
	Autres		0		13		7		3
	Total	n	16 (16%)	75 (19%)		12 (9%)		15 (10%)	
Total		101 (100%)	405 (100%)		138 (100%)		147 (100%)		

^(a) Disease Outbreak News; OMS site Alerte et action au niveau mondial (GAR);

^(b) CDTR =Communicable disease threat report;

^(c) Pour le CDTR, données disponibles que du 6 février au 31 décembre 2012 ;

^(d) Les suivis hebdomadaires: dengue, chikungunya, grippe saisonnière, rubéole et N° spéciaux J.O. Londres et coupe Europe football (2012) ne sont pas inclus;

^(e) eWEB= EpiSouth weekly epidemiological bulletin;

^(f) BHI = Bulletin Hebdomadaire International.

Figure 4 – Evènements rapportés par l’OMS en 2012, Disease Outbreak News



Figure 5 – Evènements rapportés par EpiSouth dans l’e-WEB en 2012

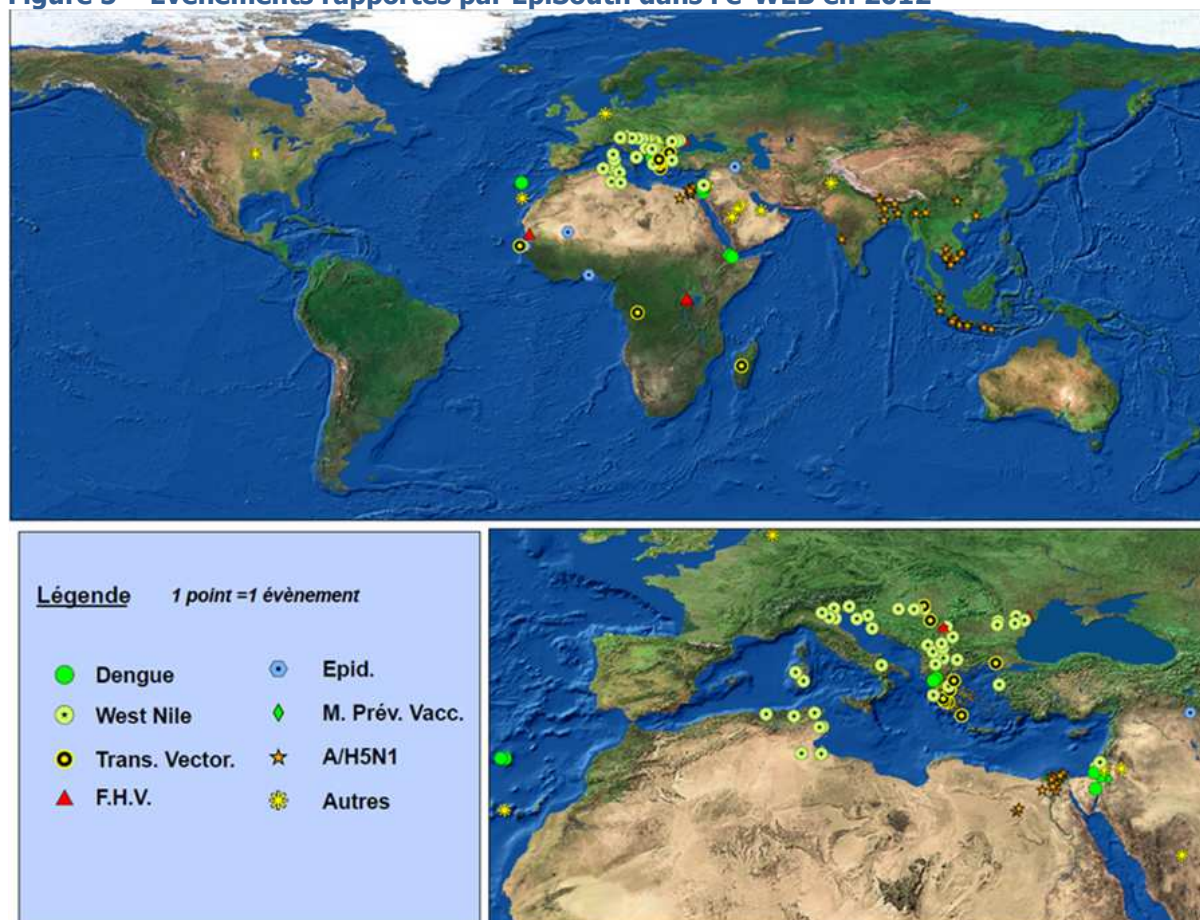


Figure 6 – Evénements rapportés par l'ECDC dans le CDTR en 2012

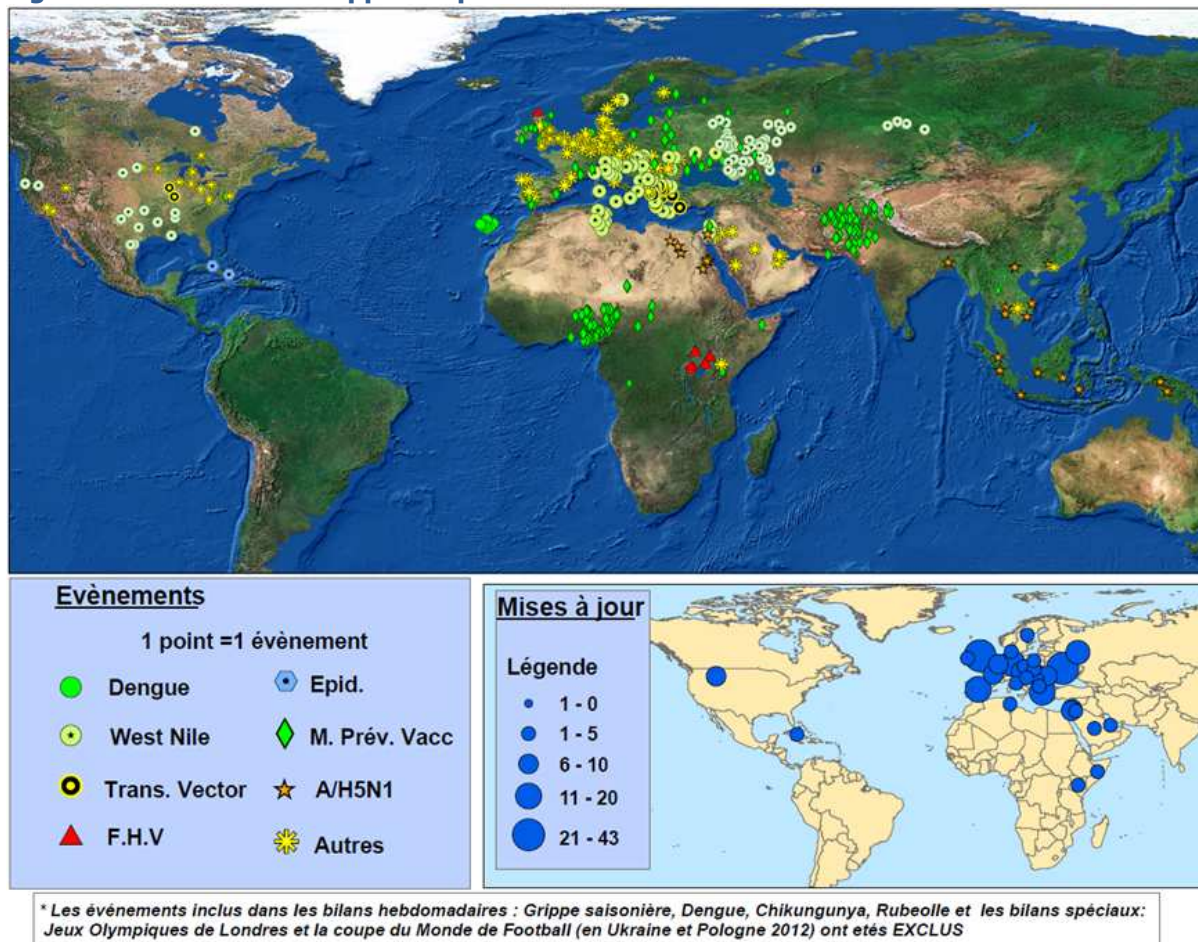
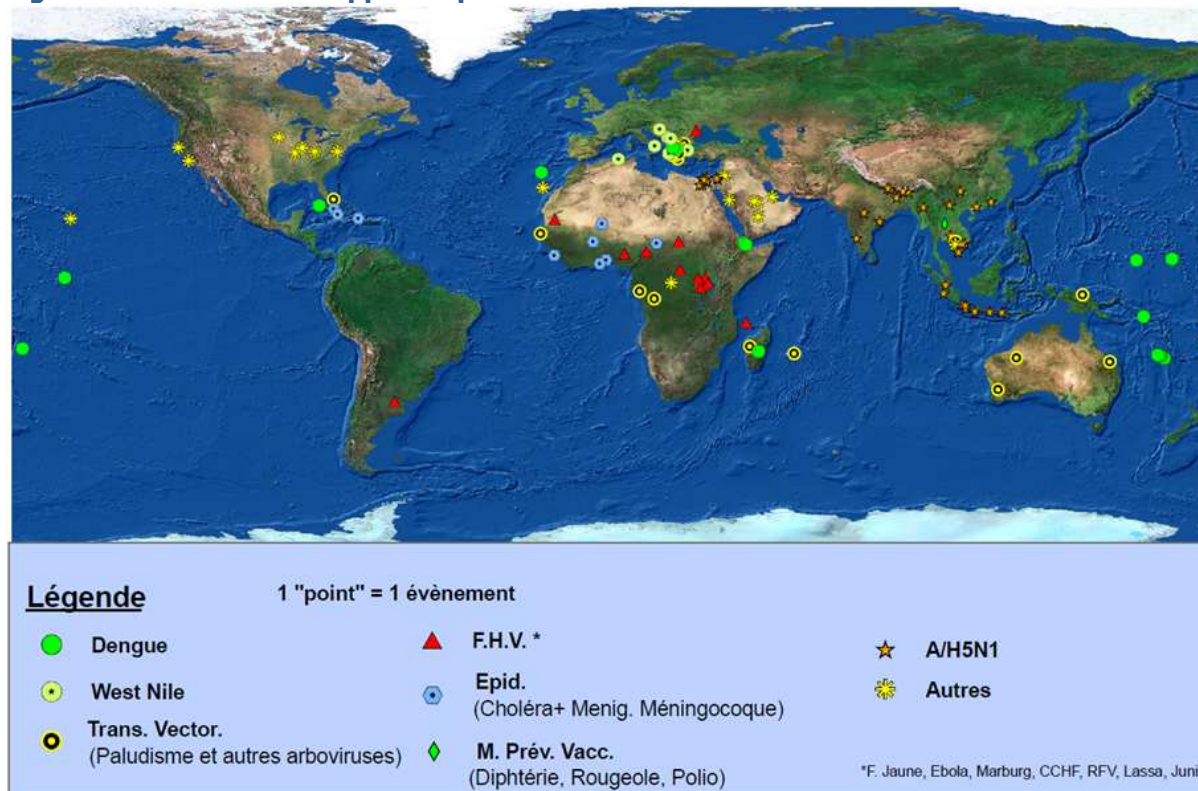


Figure 7 – Evénements rapportés par l'InVS dans le BHI en 2012



Ces résultats illustrent, l'importance de la stratégie d'intelligence épidémiologique sur la détection des événements sanitaires. Le type de pathologie et la répartition géographique des événements rapportés illustrent les priorités des institutions.

Les pathologies à transmission vectorielle apparaissent particulièrement importantes pour l'ECDC, EpiSouth et l'InVS⁷¹, alors qu'elles sont très peu représentées dans les productions de l'OMS. Cette différence s'explique par le mandat de ces institutions. Les arboviroses sont endémo-épidémiques dans de larges parties du monde. En dehors des zones non-endémiques, l'impact potentiel des événements survenus durant la période était relativement limité. Ce type d'événement a donc été peu rapporté par l'OMS. En revanche, les arboviroses sont peu présentes dans le bassin méditerranéen. Les impacts potentiels ou les mesures en matière de diagnostic, d'exclusion du don du sang ou de lutte anti-vectorielle sont beaucoup plus importants en Europe et dans le bassin méditerranéen d'où l'intérêt de l'ECDC, EpiSouth et l'InVS pour ces pathologies. Les plupart des épisodes concernant des arboviroses rapportées par ces institutions sont liés à : la détection d'événements sanitaires nécessitant une réponse liée à l'émergence de nouvelle souche virale (émergence d'une souche virale West Nile appartenant au Lignée 2 dans le bassin méditerranéen⁷²), de la présence de vecteurs compétents ou de l'introduction récente d'un virus (ex : dengue aux Açores) ou aux risques particuliers dans les territoires d'outre-mer (ex : dengue dans le Pacifique). En parallèle, la proportion des FHV et des nouveaux pathogènes comme le MERS-CoV (Middle East respiratory syndrome coronavirus) est beaucoup plus importante parmi les épisodes rapportés par l'OMS reflétant l'approche globale de cette institution. Enfin, de par son mandat portant sur le contrôle des maladies infectieuses dans l'UE⁶⁰, l'ECDC porte une attention particulière à certaines pathologies infectieuses qui ne rentrent pas dans les critères des autres institutions comme les salmonelloses ou la rubéole.

Cette étude présente une limite majeure, elle n'est basée que sur les informations publiques et sous-estime certainement le nombre d'événements détectés mais non publiés en particulier dans le cas de l'OMS qui demeure tributaire d'une autorisation du pays concerné pour rendre public les événements. Ainsi, en 2012 parmi les 292 crises sanitaires¹⁵ investiguées par l'OMS, seules une vingtaine ont fait l'objet d'une communication sur le site Disease Outbreak News. Ce biais est certainement beaucoup moins important pour l'ECDC (données non disponibles) et très faible pour l'InVS et EpiSouth. Ces différences importantes limitent considérablement les possibilités de comparaison entre institutions.

Cette étude illustre la diversité des stratégies d'intelligence épidémiologique et leur impact sur la détection et la communication des événements sanitaires pouvant représenter un risque pour la santé publique. Elle souligne également, la difficulté d'identifier une source de données qui pourrait servir de « gold-standard » à l'évaluation des systèmes de biosurveillance. Enfin, elle réaffirme l'importance d'effectuer cette évaluation des systèmes de biosurveillance dans un référentiel intégrant les objectifs régissant la construction du gold-standard.

III.2 EVALUATION DES SYSTEMES DE BIOSURVEILLANCE (ARTICLE 1)

Les sources informelles accessibles par Internet ne sont pas conçues à des fins d'intelligence épidémiologique. L'intelligence épidémiologique et notamment l'indispensable vérification de l'authenticité des événements détectés sont des processus très lourds, et qui requièrent beaucoup de temps et de ressources. L'évaluation de la pertinence des signaux détectés par les systèmes de biosurveillance est essentielle pour pouvoir optimiser les activités. Comme indiqué précédemment, cette évaluation de l'intelligence épidémiologique et des systèmes de biosurveillance fait face à des défis majeurs en particulier l'absence de gold-standard et d'indicateurs standardisés.

Le projet EAR a fourni le cadre formel souhaité en permettant la mobilisation et la participation des différents partenaires. Entreprendre d'emblée une évaluation exhaustive aurait nécessité des moyens non-disponibles dans le cadre du projet EAR. L'étude a été construite autour de deux composantes : une analyse qualitative et une évaluation quantitative. L'objectif n'était pas de réaliser une évaluation exhaustive des systèmes de biosurveillance et des dispositifs d'intelligence épidémiologique, mais plutôt de fournir des éléments rationnels pour évaluer la faisabilité d'une telle étude et de documenter les éventuelles complémentarités ou la redondance des systèmes. Cet article vise à présenter la méthodologie et les résultats utilisés pour évaluer la complémentarité des systèmes de biosurveillance et à fournir une base pour les recherches ultérieures qui doivent être menées pour évaluer et améliorer la structure des dispositifs d'intelligence épidémiologique.

Les participants à l'étude qualitative ont noté la réactivité (« timeliness ») et la sensibilité avec des scores élevés, illustrant la valeur des systèmes dans l'intelligence épidémiologique. Les faiblesses identifiées portaient sur le manque de représentativité, d'exhaustivité et de flexibilité.

Ces résultats ont été corroborés par l'analyse quantitative effectuée sur des signaux potentiellement liés à des épisodes de grippe A/H5N1 survenus en mars 2010. Pour les six systèmes pour lesquels cette information est disponible, le taux de détection variait de 31% à 38%, et augmentait à 72% lorsque l'on considérait le système virtuel combiné (intégrant les meilleures performances des systèmes individuels). Les valeurs prédictives positives variaient de 3% à 24% et les F1-scores de 6% à 27%. La sensibilité des systèmes variait de 38% à 72%. Une différence moyenne de 23% a été observée entre les sensibilités calculées pour les cas humains et les épizooties, soulignant les difficultés à développer un algorithme efficace pour une seule pathologie. Cependant, l'augmentation de la sensibilité système virtuel combiné à 93%, illustre clairement les complémentarités entre les systèmes. Le délai moyen entre la détection des épisodes A/H5N1 par les systèmes et leur notification officielle par l'OMS et l'OIE était de 10,2 jours (IC 95%: 6,7 - 13,8) confirmant la valeur ajoutée de l'utilisation des sources informelles pour la détection précoce des alertes sanitaires et les performances des systèmes de biosurveillance dans ce domaine.

Cette étude a souligné la valeur ajoutée, et les synergies entre les systèmes, entre utilisateurs et entre systèmes et utilisateurs. La complexité et la diversité des approches de l'intelligence épidémiologique et l'expertise développée par les systèmes sont beaucoup plus larges que ce qui pourrait être décrit dans cet article. En dépit des résultats obtenus, les systèmes et les institutions font face à des défis majeurs, tels que l'augmentation exponentielle du volume d'information postée sur Internet, le changement de type de communication et de diffusion de l'information (réseaux sociaux et communications instantanées) et la gestion de grands volumes de données. Les niveaux d'information redondante et le bruit de fond sont très élevés et la collaboration internationale est encore limitée. Aucun « supra-système » n'existe pour permettre la mise en commun de l'expertise des systèmes de biosurveillance et des initiatives doivent être développées dans ce sens. D'autres recherches doivent être réalisées, sur des périodes d'étude plus longues, avec différents types d'événements et un gold-standard plus robuste. De même, des utilisateurs supplémentaires et l'expérience d'autres systèmes devraient être pris en considération. Cette étude constitue une première étape qui, ouvre la voie à l'exploration de cette composante difficile, mais essentielle pour la sécurité sanitaire.

III.3 FACTEURS DE DETECTION DES EVENEMENTS INFECTIEUX (ARTICLE 2)

Cette étude vise à documenter les facteurs influençant les performances des systèmes de biosurveillance en matière de détection des événements sanitaires d'origine infectieuse.

472 signaux bruts liés aux 86 événements inclus dans le gold-standard ont été extraits des six systèmes. 84 événements ont été détectés avant leur publication dans le gold-standard. En fonction des systèmes, la I-Se variait de significativement passant de 43 % à 71 % ($p = 0.001$), tandis que les autres indicateurs étaient similaires (C-DR : $p = 0,20$; C- Se, $p = 0,13$). La I-Se était significativement associée : aux systèmes individuels, au type de système, à la langue, à la région d'occurrence et au type de maladie infectieuse après ajustement pour les autres variables. En revanche, aucune différence statistique de la C-DR n'a été observée après ajustement pour les autres variables.

Bien que les différences puissent résulter de la conception des systèmes de biosurveillance, les résultats suggèrent que l'expertise combinée entre systèmes améliore les performances de détection précoce des maladies infectieuses. Tous les systèmes ont montré des performances de détection précoce similaires. En revanche, après ajustement pour les autres variables, les systèmes intégrant une modération humaine ont démontré une I-SE 53% supérieure ($RR=1.53 [1.34-1.75]$; $p < 0.001$) à celle des systèmes automatisés.

L'intelligence épidémiologique est désormais une composante essentielle des systèmes d'alerte précoce. Dans l'ensemble, les capacités de détection des maladies infectieuses des systèmes sont complémentaires avec des performances en termes de réactivité et de sensibilité démontrées. Les produits de ces systèmes, ainsi que l'expertise des systèmes de biosurveillance et des institutions de santé publique responsables de l'intelligence épidémiologique devraient être mis en commun pour assurer la détection précoce optimale des menaces sanitaires. Les systèmes de biosurveillance ont considérablement évolué et suffisamment de données sont maintenant disponibles pour leur évaluation. D'autres études associant prospectivement les principales parties prenantes doivent être réalisées. Enfin, des stratégies innovantes doivent être mises en œuvre pour permettre la mise en commun de l'expertise développée par les différents systèmes et les institutions de santé publique, de même que la construction d'un gold-standard robuste.

III.4 REACTIVITE DETECTION ET SENSIBILITE SYSTEMES (NON-PUBLIEES)

Dans le cadre de la même étude (Article - 2), des données complémentaires ont été recueillies pour permettre de mesurer leurs réactivités et leurs sensibilités en termes de détection précoce et vis-à-vis des activités d'intelligence épidémiologique.

Pour les 73 événements pour lesquels l'information était disponible, 177 signaux bruts (40,4%) n'ont pas été détectés ou ont été détectés **après** la détection de l'événement par l'équipe de la VI, 111 (25,3%) l'ont été le même jour et 150 (34,3%) signaux ont été détectés **avant** leur détection par l'équipe VI.

En fonction des systèmes, la Se-VI variait de 48% à 79% et elle était de 100% pour le système combiné virtuel. La Se-DP (au moins 1 jour avant la VI) variait de 24 à 44% en fonction des systèmes et elle était de 86% pour le système combiné virtuel (Tableau 14).

Tableau 14 – Sensibilité des systèmes de biosurveillance vis-à-vis de l'équipe de la Veille internationale de l'InVS (n = 73)

Signaux	Argus	BioCaster	GPHIN	HealthMap	MedISys	ProMED	Combiné
Non-détectés/ détectés après	31	38	15	24	35	34	0
Détectés le même jour	18	17	27	17	18	14	10
Détectés avant la VI	24	18	31	32	20	25	63
Total	73	73	73	73	73	73	73
Sensibilité VI (Se-VI)	58%	48%	79%	67%	52%	53%	100%
Sensibilité DT (Se-DP)	33%	25%	42%	44%	27%	34%	86%

La réactivité de la détection des événements par les systèmes vis-à-vis de leur publication dans le BHI variait en moyenne de 8,4 jours (IC 95% : 4,8; 12,0) à 10,8 jours (7,1; 14,6). Pour le système combiné les événements auraient été détectés en moyenne 17,3 jours (12,3; 22,2) avant leur publication dans le BHI (Tableau 15).

Quand la date de détection par l'équipe de la VI était prise en considération, le délai moyen de détection variait de -3,7 jours (-8,7 ; +1,3) à -0,9 (-5,6 ; +3,8) pour les systèmes individuels (les systèmes détectaient en moyenne les événements entre 0,9 et 3,7 jours après la VI). Ce délai était de +5,5 jours (+1,7 ; +9,3) pour le système combiné (Tableau 15 et Figure 8). Enfin, quand la date de survenue est prise en considération, le délai moyen de détection variait de -23,5 jours (-32,4; -14,6) à -20,9 jours (-29,0; -12,8) pour les systèmes individuels et il était de -14,8 jours (-20,4; -9,2) pour le système combiné.

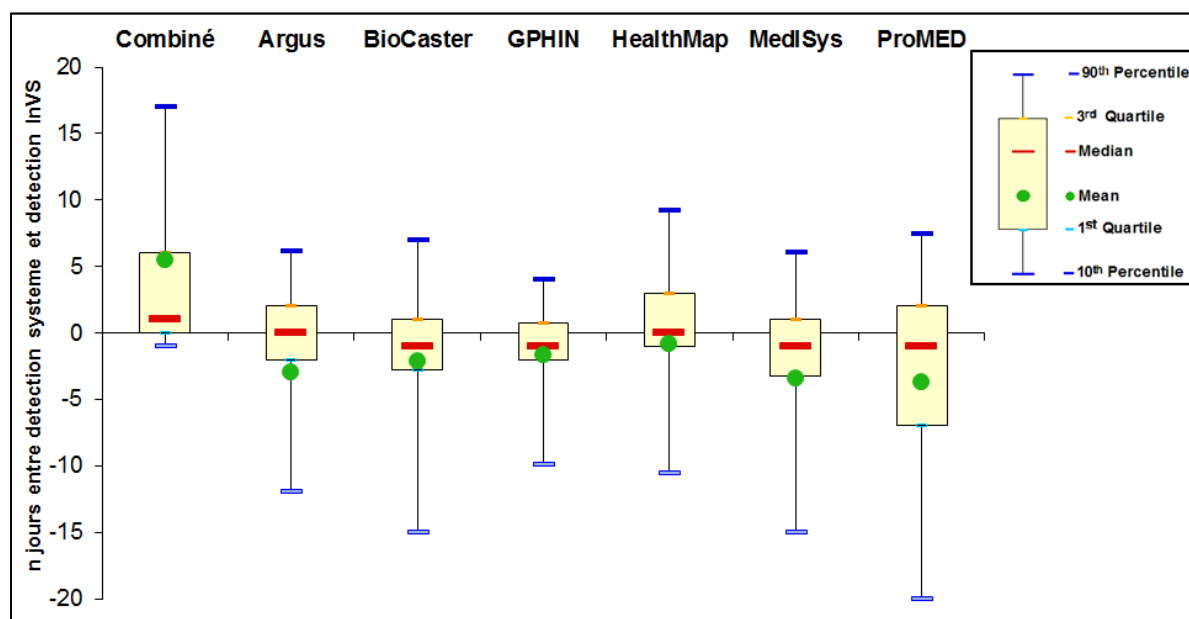
Tableau 15 – Délais entre la détection d'un signal par un système et sa détection par la Veille Internationale de l'InVS, sa publication et sa date de survenue (n = 73 événements).

Réactivité	Argus	BioCaster	GPHIN	HealthMap	MedISys	ProMED	Combiné	
Détection-InVS* – Détection-système	Moyenne	-2,9	-2,2	-1,7	-0,9	-3,5	-3,7	5,5
	IC 95%	±4,4	±5,3	±2,6	±4,7	±4,7	±5,0	±3,8
	<i>n événements</i>	(70)	(66)	(58)	(69)	(68)	(67)	(71)
Publication- BHI – Détection-système	Moyenne	9,1	10,0	8,9	10,8	8,7	8,4	17,3
	IC 95%	±5,2	±4,5	±4,1	±3,8	±3,7	±3,6	±4,9
	<i>n événements</i>	(82)	(79)	(79)	(71)	(81)	(80)	(84)
Survenue-événement – Détection-système#	Moyenne	-21,8	-20,9	-23,5	-22,3	-22,3	-22,4	-14,8
	IC 95%	±7,4	±8,1	±8,9	±8,9	±8,3	±8,4	±5,6
	<i>n événements</i>	(61)	(66)	(49)	(59)	(64)	(63)	(84)

* Parmi les 73 événements (85%) pour lesquels la date de détection de l'événement était disponible

La date de survenue de l'événement en fonction des données disponibles (dates de détection du premier cas, de début des signes, d'hospitalisation, de diagnostic, de décès, ou de notification)

Figure 8 – Délais entre la détection d'un signal par un système et sa détection par la Veille Internationale de l'InVS



Ces résultats apportent des éléments complémentaires qui illustrent les capacités de détection précoce des systèmes de biosurveillance, et les gains importants en matière de rapidité de détection et de sensibilité qui résulteraient de la mise en commun de ces systèmes. Ils confirment la nécessité pour une équipe en charge de l'intelligence épidémiologique d'utiliser en parallèle plusieurs de ces systèmes. Ils permettent également de fournir des premières indications sur deux aspects non encore explorés : le délai écoulé entre la survenue d'un événement et sa détection et la capacité de détection d'une équipe en charge de l'intelligence épidémiologique.

Dans cette étude, les 1^{ers} signaux ont, en moyenne, été détectés entre 8,4 et 10,8 jours avant leur parution dans le BHI pour les systèmes individuels et 17,3 jours pour le système combiné. Si, ces résultats illustrent la capacité de détection précoce, ils surestiment les performances des systèmes vis-à-vis du gold-standard. En effet, le BHI, est une parution hebdomadaire, et la date de parution ne reflète donc pas fidèlement la date de détection de l'évènement par l'équipe de VI puisqu'un évènement publié dans un BHI a pu être détecté au cours de n'importe lequel des 7 jours précédents sa publication. L'utilisation d'un bulletin hebdomadaire constitue donc une limite importante pour estimer la réactivité de détection. Cependant, les bulletins hebdomadaires représentent d'un mode de communication fréquemment utilisé par les institutions en charge de l'intelligence épidémiologique, et souvent les seules sources de données publiques comme dans le cas de l'ECDC ou d'EpiSouth (Tableau 12). Ce type de données ne peut donc pas être exclu a priori des études futures, mais les comparaisons entre les études doivent tenir compte de ce biais. Dans notre première étude (Article 2), la complétude globale des systèmes individuels variait de 2,2 (0,5; 3,8) à 7,8 jours (4,0; 11,5) et 10,2 jours (6,7; 13,8) pour le système combiné. En tenant compte du biais relatif à des données issues de bulletins hebdomadaires, les résultats sont comparables.

Les évènements pour lesquels les informations étaient disponibles, ont été détectés en moyenne 22 jours après la survenue de l'évènement avec un maximum de 179 jours. Ce délai n'était plus que de 15 jours pour le système virtuel. Nos résultats sont consistants avec ceux de l'étude réalisée par Chan et al³³ sur 281 épidémies vérifiées par l'OMS et survenues entre 1996 et 2009. Cette étude avait montré que le temps médian entre la survenue d'une épidémie et sa communication publique avait été réduit de 29,5 jours (IC 95% : 13,5; 59,0) en 1996 à 13,5 jours (3,5; 44,5) en 2009. Des variations avaient également été observées en fonction des régions géographiques : de 30 (24 ; 41) jours dans la Région OMS Afrique à 4 (3 ; 7) jours dans la Région du Pacifique Occidental.

Ces résultats illustrent les progrès qui peuvent encore être accomplis en matière de détection précoce par les systèmes eux-mêmes et au travers de leur association. Ces données doivent cependant être interprétées avec précaution. Les informations utilisées pour estimer la date de survenue sont très hétérogènes (date de détection du premier cas, date de début des signes, date d'hospitalisation, date de diagnostic, date de décès, date de notification). Il s'agit de données brutes disponibles au moment de la détection, on ne peut donc exclure que des investigations poussées auraient pu permettre d'identifier des cas plus anciens. Enfin, ces données sont tributaires de l'accessibilité aux données y compris informelles.

Enfin, quand la date de détection par l'équipe de la VI était prise en considération, le délai moyen de détection des 438 signaux par les systèmes individuels était de -2,5 jours pour une médiane de -1,0 jour (Figure 8), soit 2,5 jours plus tard que leur détection par l'équipe de la VI. Ces résultats doivent bien sûr être mis en relation avec la sensibilité (Se_{VI}) des systèmes individuels : en moyenne 60% des signaux ont été détectés avant ou le même jour que par l'équipe VI. En revanche, l'analyse illustre clairement le gain de performance qui résulterait de l'utilisation d'un système combiné puisque si un tel système avait existé les 73 signaux (pour lesquels les informations étaient disponibles) auraient pu être détectés en moyenne 5 jours (+1,7 ; +9,3) plus tôt.

Les informations concernant la date de détection d'un événement étaient collectées en routine par l'équipe de la VI avec un objectif de documentation des activités mais pas dans le cadre d'une utilisation de ces données pour une étude scientifique. On ne peut pas exclure que certains signaux aient, en fait, pu être détectés plus tôt mais qu'ils aient fait l'objet d'une première évaluation avant d'être effectivement considérés comme un événement potentiellement important (et enregistré comme tel dans la base de données de routine). Les dates de détection par l'équipe VI n'étaient pas disponibles pour 15% (13/86) des événements.

Dans l'article 2, nous avons montré que la sensibilité intrinsèque (I_{Se}) des systèmes modérés (Argus, GPHIN et ProMED) ajustée pour les autres facteurs était 1,53 (1,34 ; 1,75 ; $p < 0.001$) fois supérieure à celle des systèmes automatisés. Ces différences de performance contribuent aux différences observées en termes de Se_{VI} et Se_{DP} et notamment aux performances élevées observées pour GPHIN (dont la Se_{VI} était de 21% à 31% supérieure à celle des autres systèmes) ou pour HealthMap et GPHIN dont les Se_{DP} étaient supérieures d'au moins 7 à 17%. Néanmoins, les six systèmes ne sont pas tous utilisés en routine de manière équivalente et les modalités de leur utilisation auraient pu influencer les résultats. On ne peut notamment pas exclure que pour les systèmes les plus utilisés en routine (Argus, GPHIN, HealthMap et ProMED) les signaux soient plus facilement et précocement détectés par l'équipe VI du fait de leur maîtrise des systèmes plutôt que des performances intrinsèques de l'outil. Les informations concernant l'origine de la détection d'un événement n'étant pas disponibles, l'ampleur de ce biais potentiel n'a pas pu être estimée.

En dépit de ces limites, ces résultats illustrent la complexité du processus de détection par l'équipe chargée de l'intelligence épidémiologique, de leur dépendance vis-à-vis des systèmes et de la nécessité d'utiliser plusieurs systèmes en dépit de l'importante duplication que cela génère. Ils soulignent la nécessité de développer des solutions techniques permettant d'intégrer les systèmes individuels dans un système unique et confortent la faible valeur ajoutée des signaux bruts. Si d'autres études, notamment prospectives, doivent être menées pour confirmer ces résultats, ils constituent une première base pour l'évaluation des performances d'un dispositif d'intelligence épidémiologique et pour l'estimation du délai de détection-survenue des événements sanitaires.

IV. DISCUSSION

IV.1 RESULTATS

IV.1.1 SYNTHÈSE DES RESULTATS

L'objectif principal de nos travaux était d'évaluer, du point de vue des utilisateurs, les performances des systèmes de biosurveillance utilisés pour la détection précoce des risques sanitaires dans le cadre de l'intelligence épidémiologique.

Nos travaux ont démontré les performances des systèmes individuels en matière de détection précoce des événements sanitaires infectieux. Ils illustrent notamment la valeur ajoutée spécifique de chaque système, tous (sauf un) ayant été en mesure de détecter précocement des événements non-détectés par les autres systèmes, la plus grande sensibilité intrinsèque des systèmes modérés et la variabilité du type de sources d'information utilisés.

Au-delà de la description des performances individuelles, nos résultats démontrent la complémentarité de ces systèmes, tant du point de vue de la sensibilité que de la réactivité. L'augmentation de la sensibilité rétrospective de la détection des épisodes A/H5N1 par le système virtuel combiné à 93% (contre en moyenne 38% à 72% pour les systèmes individuels) et la détection des événements par le système virtuel combiné 17,3 jours avant leur publication dans le BHI (contre en moyenne 8,4 à 10,8 jours pour les systèmes individuels) illustrent clairement les gains potentiels qui résulteraient de l'intégration de ces systèmes individuels dans un « supra-système ».

En mesurant la réactivité dans des conditions proches de la routine, nos résultats illustrent l'importance que peut jouer l'intelligence épidémiologique dans le cadre du renforcement de la sécurité sanitaire et de la mise en œuvre des mesures de contrôle des épidémies. En permettant la détection d'événements d'origine infectieuse d'une dizaine de jours à plusieurs semaines avant leur notification officielle, l'intelligence épidémiologique permet d'alerter les autorités sanitaires et ainsi d'anticiper la mise en œuvre d'éventuelles mesures de contrôle et de mitigation.

Nos résultats ont également permis de documenter les limites de ces outils et en particulier la faible VPP des signaux bruts détectés (3% à 24% pour les épisodes A/H5N1), la variabilité des capacités de détection pour une même pathologie, mais également l'influence significative jouée par le type de pathologie, la langue et la région de survenue sur les capacités de détection des événements infectieux.

Ils ont permis de documenter la grande diversité des stratégies d'intelligence épidémiologique mises en œuvre par les institutions de santé publique pour répondre à leurs besoins spécifiques et l'impact de ces stratégies sur la nature, l'origine géographique et le nombre des événements rapportés.

Bien que ces attributs n'aient pas été quantitativement mesurés au cours de notre étude, la documentation de la perception des utilisateurs vis-à-vis des systèmes de biosurveillance, nous a permis d'illustrer d'une part la bonne acceptabilité (les utilisateurs utilisaient entre 4 et 7 systèmes en routine et le plus souvent quotidiennement) mais également leur manque de flexibilité (exprimé par les utilisateurs). De ces constats, il résulte que la flexibilité des systèmes, c'est-à-dire leur adaptabilité aux besoins de différents utilisateurs, devrait être considérée comme un axe de recherche essentiel.

IV.1.2 SENSIBILITES

Parmi les 31 documents scientifiques en relation avec l'évaluation des systèmes de biosurveillance, identifiés par la revue de la bibliographie, seuls cinq présentaient des résultats de sensibilité brute (sensibilité et ou F score) et tous correspondaient à des études rétrospectives. Or, nos résultats ont démontré que si la sensibilité brute est un bon indicateur pour mesurer les performances théoriques d'un système, elle est inadaptée pour l'estimation des performances dans des conditions réelles d'utilisation et encore moins pour comparer les performances des systèmes.

Pour la détection des épisodes A/H5N1, le taux de sensibilité brute (C-Se) des différents systèmes variait de 38 à 72%. De cet unique constat, on pourrait à tort, conclure que certains systèmes sont plus performants que d'autres. La réalité est bien plus complexe.

Tout d'abord, la C-Se constitue une mesure uniquement théorique, visant à déterminer rétrospectivement si un système avait ou non détecté un signal brut. Elle présente des limites importantes. Tout d'abord, elle sous-estime un facteur essentiel : le facteur humain. Nos résultats ont mis en évidence des différences importantes entre la C-Se et le taux de détection (DR) qui lui prend en considération la capacité d'un utilisateur à détecter effectivement un signal brut dans des conditions proches de la routine. Bien que ces différences observées soient au moins en partie directement attribuables à la difficulté rencontrée par l'utilisateur à détecter prospectivement des informations pertinentes dans un volume important de bruit de fond, les taux de détections ne peuvent être directement comparés aux sensibilités brutes. En effet, afin d'obtenir le même dénominateur, les événements survenus avant mars 2010, n'ont pas été inclus dans les calculs. Les taux de détection sont donc probablement sous-estimés. Néanmoins, ce biais potentiel a affecté tous les systèmes de la même manière. Ce taux de détection des systèmes variait entre 31% à 38%, on peut donc en conclure que, dans les conditions de routine, les performances globales des systèmes sont équivalentes.

La C-Se, présente une autre limite, elle ne tient pas compte de l'importance du syntrophisme⁶. Pour 39 % des signaux détectés précocement par les systèmes automatisés, la source était l'un des autres systèmes inclus dans l'étude. Ce résultat démontre pour la première fois l'importance de ce phénomène d'informations croisées. Il illustre également la vulnérabilité des systèmes intégrant une forte proportion de syntrophie. Dans un contexte stable, où la pérennité des différents systèmes serait assurée, l'impact global du syntrophisme serait limité, ce n'est malheureusement pas le cas. Depuis la mise en œuvre de notre étude, trois des systèmes ont soit cessé d'être utilisable soit d'être maintenu. Depuis 2012, les activités d'Argus (financées par le gouvernement Américain) ont été réorientées vers d'autres types d'activités de surveillance et ne couvrent plus la détection des risques sanitaires ; BioCaster et PULS, développés dans le cadre de recherches universitaires ont été contraints de suspendre leur maintenance et/ou leur développement. En d'autres termes, seuls quatre des sept systèmes demeurent opérationnels en 2014.

La mesure de la sensibilité intrinsèque (I-Se) représente un meilleur indicateur de performance. En se concentrant sur les performances propres des systèmes, elle permet de mesurer l'impact de la syntrophie. Nous avons pu démontrer que si les autres indicateurs étaient similaires (C-DR: $p = 0,20$; C-Se, $p = 0,13$), la I-Se variait significativement de 43 % à 71 % ($p = 0,001$), mettant ainsi en évidence les performances intrinsèques des systèmes et en particulier l'importance du facteur humain dans le processus de détection. En effet, nos résultats ont montré, pour la première fois que la I-Se des systèmes modérés (qui incluent une intervention humaine) était 53% ($p < 0,001$) plus élevée que celle des systèmes automatisés après ajustement pour les autres variables.

⁶ Le terme syntrophie est adapté de sa définition en biologie où il décrit le mécanisme qui explique la croissance dans des conditions apparemment défavorables de bactéries exigeantes autour de colonies d'une autre espèce. Ce phénomène s'explique par le fait que la bactérie exigeante tire parti des facteurs métaboliques qu'elle ne peut synthétiser elle-même et qui ont été libérés par l'autre bactérie. <http://dictionnaire.sensagent.com/syntrophie/fr-fr/>

Les systèmes modérés sont beaucoup plus chers à développer et à maintenir que les systèmes automatisés en raison des coûts liés aux ressources humaines (en termes de nombre, mais également de la diversité de l'expertise linguistique) qui y sont associés. En démontrant que l'analyse humaine permettait d'augmenter significativement les performances, nos résultats apportent des éléments objectifs pour permettre aux systèmes modérés de plaider pour la sécurisation des ressources nécessaires à la maintenance et au développement de leur système. Les implications potentielles vont bien au-delà de la survie des systèmes individuels puisqu'en raison de la syntrophie des systèmes automatisés, l'arrêt d'un ou de plusieurs systèmes modérés entraînerait mécaniquement une baisse des performances des systèmes automatisés.

IV.1.3 VALEURS PREDICTIVES POSITIVES

En s'attachant à mesurer la valeur ajoutée des systèmes du point de vue des utilisateurs, notre étude illustre les contraintes posées par le volume d'information, la faible VPP des informations brutes détectées et l'importance de la duplication.

L'évaluation quantitative de la détection des épisodes d'influenza aviaire A/H5N1 a mis en évidence que sur une période d'un mois, 1 154 documents, en relation avec un épisode A/H5N1 avaient été détectés par les systèmes, soit en moyenne 37 documents par jour. D'autre part, elle illustre la faible VPP globale des signaux détectés (3% to 24% en fonction des systèmes). L'importance du syntrophisme illustre l'importance de la duplication entre les systèmes puisque au moins 39 % des signaux détectés précocement par les systèmes automatisés constituaient un doublon (signal déjà détecté par un autre système). Et ce chiffre sous-estime largement l'importance de la duplication puisqu'il ne concerne que les signaux correspondant à un vrai évènement. Il ne prend donc pas en compte la duplication du volume total d'information collecté par tous les systèmes ni le nombre de rapports similaires détectés par chaque système pour un même évènement.

La proportion d'évènements détectés par les systèmes avant leur détection par l'équipe de la VI fournit une autre indication du volume de bruit de fond. Le délai moyen de détection des 438 signaux par les systèmes individuels était de -2,5 jours (médiane : -1,0 jour), soit en moyenne 2,5 jours plus tard que leur détection par l'équipe de la VI. Seul 34% (150) des évènements ont été détectés par les systèmes avant leur détection par l'équipe de la VI. En pratique, les deux tiers des signaux recueillis n'étaient donc plus pertinents puisqu'ils correspondaient à des évènements déjà détectés.

Ces chiffres permettent de mieux appréhender le volume considérable d'information à traiter quotidiennement pour l'ensemble des risques sanitaires potentiels. En raison des contraintes importantes (absence de base de données adéquate, volume d'information, disponibilité en ressources humaines, durée de la période d'étude, etc.), la mesure de la VPP n'a pu être effectuée que sur l'une des deux études, sur une période courte et en ne prenant en compte qu'une seule pathologie. D'autres études devront être menées pour affiner la mesure la VPP de l'ensemble des signaux bruts.

IV.1.4 FACTEURS INFLUENÇANT LA DETECTION :

Nos travaux ont permis d'établir que les performances de détection précoce étaient significativement influencées par le type de pathologie, la langue et le lieu de survenue des épisodes, mais également qu'elles variaient pour différentes formes d'une même pathologie. Ces résultats ouvrent de nouveaux champs d'investigations dont les applications pourraient être importantes pour les utilisateurs comme pour les systèmes. Pour les utilisateurs, le fait que certains systèmes soient plus performants pour la détection de certaines pathologies ou pour certaines régions pourrait permettre d'optimiser leur utilisation. Pour les systèmes, le fait qu'ils soient plus performants pour certaines langues, zones géographiques ou pathologies pourrait leur permettre d'identifier les mécanismes de sélection (par exemple algorithmes) les plus performants et de prioriser les axes de recherche opérationnelle. Dans le cadre d'une coopération formalisée entre les systèmes, les implications potentielles seraient plus

importantes. En particulier, l'intégration des produits des différents systèmes dans un outil unique pourrait tirer parti de leur diversité et ainsi améliorer la sensibilité tout en augmentant la spécificité.

En revanche, nos résultats ne permettent pas d'aller plus loin dans la description de l'impact des facteurs associés à la performance. L'échantillon d'événements inclus dans le gold-standard de notre seconde étude était malheureusement insuffisant pour permettre des analyses stratifiées notamment en raison de l'existence de potentiels facteurs de confusion. Par exemple, la langue source et la zone géographique pourraient constituer un éventuel facteur de confusion (on peut imaginer que le français soit plus susceptible d'être la langue source d'événements survenant en Afrique francophone et au Maghreb, l'arabe en Afrique du Nord et au Moyen-Orient, l'espagnol en Amérique latine, le russe en Asie Centrale, etc.). Le type de pathologie et la zone de survenue pourraient également être liés. On ne peut formellement exclure, des différences dans les performances de détection d'une même pathologie (ex : la dengue ou le paludisme) en fonction de son lieu de survenue : dans une zone de forte endémie ou dans une zone exempte. A ce stade, il n'est pas possible de déterminer dans quelle direction ce biais pourrait infléchir les résultats.

Néanmoins, en raison de la force des associations et de l'absence d'interaction significative dans les modèles uni et multivarié, nos résultats ne sont probablement pas affectés. Par ailleurs, ils sont étayés par l'existence d'éléments liés à la conception des systèmes.

IV.1.4.1 La langue

L'anglais, le français et l'espagnol, sont des langues communes à tous les systèmes alors que le persan, le vietnamien ou le roumain ne sont pris en charge au maximum que par deux systèmes. La diversité des langues traitées par les systèmes, mais également l'ancienneté de leur traitement, sont certainement en mesure d'influencer les performances de traitement des informations dans ces idiomes. Les objectifs que se sont fixés les systèmes permettent également d'expliquer, au moins en partie, l'influence de la langue sur les performances de détection. Par exemple, BioCaster portait une attention particulière sur les pays de l'Asie du Sud, et traitait des articles en 13 langues dont le chinois, le japonais, le thaï, le coréen et le vietnamien. Nos résultats semblent indiquer de meilleures performances pour BioCaster en chinois que dans des langues comme le portugais, le russe ou même l'anglais. Ces différences ne sont cependant pas significatives, probablement en raison du faible nombre d'événements survenus en Asie inclus dans le BHI. D'autres études devraient donc être menées pour étudier ces potentialités.

D'autres facteurs doivent également être pris en considération et notamment la nature même de la structure linguistique, prenons un exemple. Le chinois est une langue monosyllabique à tons dans laquelle i) une syllabe (un son) est égale à un mot et ii) pour une même syllabe il existe quatre mots caractérisés (à l'oral) par leur accent tonique. En chinois, le concept de nom propre n'existe pas et cette particularité rend difficile l'identification automatisée des lieux, des individus et des nouveaux mots. En l'absence de système orthographique, la création d'un idéogramme pour un nouveau mot ou un nom propre résulte de l'association du mot, d'un concept et d'un idéogramme existant. Par exemple, le mot « France ». En raison du caractère monosyllabique un « son » le plus proche possible de la consonance « France » a été choisi, le son « fa ». Parmi les idéogrammes correspondant au son « fa » c'est l'idéogramme 法 (fǎ) correspondant au mot « loi » qui a été choisi. L'idéogramme 国 (guó) signifie « pays ». Le mot « France » s'écrit donc 法国 (fǎ guó). On comprendra que suivant le contexte de la phrase et de leur positionnement dans la phrase, ces deux idéogrammes peuvent signifier soit « la France » soit « le pays de la loi » (où les mots pays et loi sont génériques). Si, les systèmes sont capables d'identifier des mots aussi fréquents, c'est loin d'être le cas pour les noms de provinces, de localités ou les noms propres. Ainsi, le nom de l'ancien président américain Bush se transforme fréquemment en « buisson » la traduction littérale de son nom. Le système se complique encore quand on prend en considération qu'il existe deux modes d'écriture ; le chinois simplifié (utilisé

en Chine continentale et à Singapour) et le chinois traditionnel (utilisé à Hong Kong, Macao, Taïwan et dans de nombreuses communautés expatriées). Ainsi en idéogramme simplifié le mot « guó » (pays) s'écrit 国 et 國 en idéogramme traditionnel. Là encore, on mesure bien l'importance de l'analyse humaine en termes d'expertise linguistique, dans les processus de sélection et de traduction et la complexité d'inclure de nouvelles langues.

IV.1.4.2 Type de pathologie

Nos travaux ont permis d'établir l'impact significatif du type de pathologie sur les performances de détection mais également que ces performances variaient pour une même pathologie. En l'état actuel des connaissances, c'est l'élément le plus complexe à interpréter. Notre première étude portant sur la détection des épisodes d'influenza aviaire A/H5N1 a mis en évidence les difficultés à développer un algorithme efficace pour une seule pathologie. Une différence moyenne de 23% a été mesurée entre les sensibilités calculées pour les cas humains et les épizooties, et d'autre part des différences très importantes entre les systèmes ont été observées (certains détectant mieux les cas humains d'autres mieux les épizooties). Dans la seconde étude (excluant les épisodes d'influenza aviaire A/H5N1), les détections brutes et intrinsèques étaient significativement associées au type de cas (humain/animal) (C-DR ; $p > 0,001$ et I-DR ; $p = 0,03$) dans le modèle uni-varié mais pas dans le modèle multivarié (C-DR ; $p = 0,08$ et I-DR ; $p = 0,55$). Néanmoins, le fait que les résultats ne soient pas statistiquement significatifs dans le modèle multivarié n'exclut pas forcément une éventuelle association entre le type de cas et la détection. D'une part, dans notre gold-standard seul 17% ($n=15$) des événements concernaient des épizooties et d'autre part, elles concernaient quatre pathologies dont les manifestations cliniques chez l'animal sont différentes. Le type de cas (humain/animal) n'est donc qu'un des éléments à prendre en considération, la fréquence de la pathologie, sa sévérité, la spécificité de certains signes cliniques, la rapidité de propagation, le mode de transmission, etc. sont autant d'éléments qui sont potentiellement susceptibles d'influencer les capacités de détection.

IV.1.5 REACTIVITE DE DETECTION

Nos résultats démontrent que l'utilisation des sources informelles dans des conditions proches de la routine, permet une détection beaucoup plus précoce des événements sanitaires que l'utilisation de sources officielles et publiques, et ce, dans des délais qui permettent d'anticiper significativement la mise en œuvre de mesures de mitigation ou de contrôle. Dans le cas de la grippe aviaire A/H5N1, ce délai variait entre 2.2 jours (0,5; 3,8) et 7.8 jours (4,0; 11,5) en fonction des systèmes. Dans le cas des risques sanitaires correspondant aux critères de l'InVS, les premiers signaux bruts ont, en moyenne, été détectés entre 8,4 et 10,8 jours pour les systèmes individuels.

La relation entre la détection précoce et la mise en œuvre plus précoce de mesures de contrôle ne faisait pas l'objet de nos travaux et l'impact en matière de santé publique n'a donc pas pu être objectivé. Cependant, l'épidémie de SRAS illustre particulièrement bien cet impact potentiel. La détection de cette épidémie a reposé exclusivement sur les activités d'intelligence épidémiologique (son existence n'ayant été reconnue par les autorités chinoises que plusieurs semaines après sa détection). En dehors de la Chine continentale, seuls quatre pays et territoires : le Canada (Toronto), Hong Kong, Singapour, Taïwan et le Vietnam (Hanoi) ont été affectés par une épidémie soutenue. Or dans toutes ces zones (sauf Taïwan), l'importation des premiers cas est survenue avant la détection de l'épidémie alors que dans les 23 autres pays et territoires, l'importation de cas est principalement survenue après la détection de l'émergence de ce nouveau virus¹². C'est donc bien la détection de cette émergence et la mise en œuvre de mesures de contrôle qui en a résulté qui ont permis de prévenir l'établissement de cycles autochtones de transmission et au final le contrôle de la pandémie.

IV.1.6 COMPLEMENTARITE DES SYSTEMES

La complémentarité de ces systèmes a clairement pu être mesurée grâce à la construction d'un système virtuel dont les performances (théoriques) mettent en évidence les gains considérables en termes de sensibilité qui résulterait de leur intégration dans un « supra-système ». Un tel système aurait en effet permis d'augmenter le DR des épisodes A/H5N1 à 72%, la sensibilité brute à 93%. De même pour le système virtuel combiné, la réactivité de détection des épisodes A/H5N1 était de 10,2 jours (6,7 ; 13,8) avant leur notification par l'OMS/OIE (contre entre 2,2 et 7,8 jours pour les systèmes individuels). Lorsque les signaux bruts correspondant aux critères établis pour la VI étaient considérés, le délai moyen entre la détection des événements par le système virtuel combiné et leur parution dans le BHI était de 17,3 jours (contre entre 8,4 et 10,8 jours pour les systèmes individuels). Si un tel système avait existé, les 73 signaux (pour lesquels les informations étaient disponibles) aurait pu être détecté en moyenne 5 jours (+1,7; +9,3) plus tôt par l'équipe de la VI. Par ailleurs, nos résultats ont également démontré que tous les systèmes (sauf un, dans une étude) ont été capables de détecter des signaux qu'aucun autre système n'avait identifiés précocement.

IV.1.7 ROLE DE L'INTELLIGENCE EPIDEMIOLOGIQUE

Tous les événements inclus dans ces études n'ont pas été détectés par les systèmes de biosurveillance ; dans la première étude, 2 événements (7%) n'ont pas été détectés avant la notification officielle par l'OMS et/ou l'OIE et dans la seconde 2 événements (4%) n'ont pas été détectés avant leur publication dans le BHI. Ces chiffres peuvent paraître faibles, mais ils sous-estiment très certainement le rôle de la détection par d'autres sources que les systèmes de biosurveillance. En effet, la détection d'un événement au travers d'une source informelle (quelle qu'en soit la source) génère un processus de vérification auprès de sources officielles. D'une part, la détection d'un événement par une institution notamment internationale induit une accélération du processus de communication officielle et d'autre part, on ne peut pas savoir si ces mêmes événements auraient bien été détectés par les systèmes conventionnels de surveillance (SBI) des pays concernés. En l'absence d'activité d'intelligence épidémiologique, il est donc probable que la notification officielle d'une partie de ces événements ait été bien plus tardive. Cette hypothèse est confortée l'étude de Chan et al³² qui avait démontré que le délai médian entre la détection d'une épidémie et sa communication officielle avait réduit depuis la mise en œuvre du RSI (qui entre autres, officialise l'utilisation par l'OMS de l'intelligence épidémiologique). Le délai médian entre la détection d'une épidémie et sa communication officielle était passé de 33 jours (29; 40) avant la mise en œuvre effective du RSI en 2007 à 23 jours (17; 43) entre 2007 et 2009.

Au-delà de la pertinence de l'utilisation des systèmes de biosurveillance pour la détection précoce des événements sanitaires, nos travaux illustrent bien l'importance d'un dispositif intégral d'intelligence épidémiologique ne reposant pas exclusivement sur l'utilisation des systèmes de biosurveillance et en particulier de disposer d'un réseau.

IV.1.8 DELAIS SURVENUE DE L'EVENEMENT ET DE LA DETECTION

Les événements pour lesquels les informations étaient disponibles, ont été détectés en moyenne 22 jours après le début supposé de leur survenue avec un maximum de 179 jours et ce délai n'était plus que de 15 jours pour le système virtuel. Ces valeurs sous-estiment certainement le délai entre le début de la survenue de l'évènement et sa détection.

Tout d'abord, les informations utilisées pour estimer la date de survenue étaient très hétérogènes et certaines (comme les dates de diagnostic, de décès ou de notification) ne reflètent que partiellement la date de survenue du premier cas puisqu'elles incluent intrinsèquement un délai. Idéalement, il serait souhaitable de pouvoir disposer de données standardisées pour tous les cas. Même dans le cadre de la surveillance conventionnelle (SBI), il est souvent difficile d'obtenir l'ensemble des données surtout en début d'épidémie. Quand ces données sont disponibles, elles le sont généralement au niveau des institutions nationales et sont rarement accessibles publiquement.

Les données publiques présentent autre limite majeure, il s'agit de données brutes disponibles au moment de la détection de l'évènement. Ce sont donc les dates concernant les premiers cas détectés qui sont communiquées et pas forcément celles des vrais cas index. La pandémie d'influenza A/H1N1 fournit un bon exemple de l'importance de ce biais :

- Les 15 et 17 avril 2009, le CDC confirmait les deux 1^{ers} cas humains d'infection par une nouvelle souche d'influenza A/H1N1 d'origine porcine chez deux enfants californiens sans lien épidémiologique et n'ayant pas voyagé. Le diagnostic est réalisé de manière fortuite lors d'un exercice de préparation à la pandémie grippale⁷³.
- Les 2 enfants ont présenté les premiers signes cliniques respectivement les 29 et 30 mars.
- Les 23 et 24 avril, le CDC confirmait 6 nouveaux cas d'infection par la nouvelle souche A/H1N1.
- Cet isolement chez l'homme d'une nouvelle souche d'influenza porcine a été détecté par la VI de l'InVS le 23 avril. Cet évènement a fait l'objet d'une communication aux autorités sanitaires le même jour et d'une communication publique le 25 avril⁷⁴.
- Le 24 avril l'OMS annonçait 7 cas confirmés aux États-Unis et 18 au Mexique⁷⁵ ;
- Le 25 avril 2009, pour la 1^{ère} fois depuis la mise en place du RSI (et de son histoire) l'OMS déclarait une Urgence de Santé Publique de Portée Internationale⁷⁶.

Si on définit comme évènement « l'isolement d'une nouvelle souche d'influenza porcine chez l'homme » le délai entre les 1^{ères} informations disponibles (seconde vague) et le début de l'évènement était inférieur à 1 jour. Par contre, ce délai était de 8 jours si la date de la confirmation biologique du 1^{er} des deux cas est prise en considération, et de 39 jours pour la date de survenue des 1^{ers} signes chez le premier cas.

Si on définit comme évènement « la survenue des 1^{er} cas de A(H1N1) porcine chez l'homme », ce délai est d'au moins 67 jours. En effet, les investigations menées au Mexique indiquent que les premiers cas suspects seraient survenus au Mexique (La Gloria, Veracruz) vers la mi-février, hypothèse vérifiée par la première estimation phylogénétique qui date un premier ancêtre de la souche A(H1N1) d'origine porcine vers le 12/01/2009 (extrêmes : novembre 2008 - mars 2009)⁷⁷.

Cet exemple illustre la complexité d'évaluer avec précision le délai entre la survenue d'un évènement et sa détection. En pratique, il est illusoire d'espérer obtenir le même niveau d'information pour tous les évènements sanitaires aigus survenant dans le monde. Si nos résultats sous-estiment très certainement le délai de détection, ils fournissent une première base. D'autres études devaient être menées sur de plus larges échantillons et en standardisant les informations collectées pour mesurer ce délai et fournir une base solide pour estimer dans le temps les progrès accomplis.

IV.2 AVANTAGES DE LA METHODE : UNE DEMARCHE PROGRESSIVE

La mise en œuvre de l'évaluation des capacités de détection était confrontée à plusieurs obstacles, en particulier, l'absence d'un gold-standard reconnu et accessible, l'absence de méthodologie adaptée à ce nouveau domaine et l'inadaptation des indicateurs classiquement utilisés dans le cadre de l'évaluation de la surveillance des indicateurs. Par ailleurs, bien qu'en augmentation, le nombre d'acteurs est encore restreint, l'expertise dans ce domaine et les ressources notamment humaines demeurent limitées. Très peu d'études, toutes uniquement rétrospectives, ont été menées. Elles ne répondent pas aux besoins des utilisateurs en particulier, car elles n'ont pas considéré les éléments essentiels tels que le volume d'information, la redondance et la VPP des signaux détectés.

Notre objectif n'était pas d'évaluer les performances théoriques des systèmes mais bien d'intégrer le point de vue des utilisateurs dans la conception et l'analyse. Au vu des contraintes précédemment mentionnées, il était essentiel de privilégier une approche progressive dont les principales étapes ont consisté à définir les concepts, les besoins des utilisateurs et les indicateurs devant être pris en considération pour l'évaluation des systèmes de biosurveillance. L'étape suivante (Article 1) a consisté à estimer les perceptions des principaux utilisateurs vis-à-vis des principaux systèmes de biosurveillance, à mesurer les principales caractéristiques des systèmes en matière de détection précoce (sensibilité, PPV et réactivité), à déterminer les éventuels potentiels de synergie entre les systèmes et enfin à identifier la meilleure approche pour optimiser leur utilisation. C'est sur cette base que la seconde étude (Article 2) a pu être construite afin d'investiguer les facteurs influençant la détection des événements sanitaires par les systèmes de biosurveillance. Forts des enseignements de la 1^{ère} étude, cette étude devait : inclure des événements sanitaires différents en termes de type de pathologie et de distribution géographique de survenue et utiliser un gold-standard qui ne serait pas influencé par la notification officielle des événements par le pays où ils surviennent. A notre connaissance, c'est la première étude multi-systèmes visant à mesurer les sensibilités de six systèmes vis-à-vis d'un même standard multi-pathologies n'intégrant pas uniquement des données publiques.

IV.3 APPLICATION PRATIQUE: LE PROJET EAR

Au-delà de la démarche académique nos résultats ont déjà eu des implications pratiques dans le cadre du projet EAR (Cf. II.3.1 p 42). L'un des objectifs initiaux de EAR était d'évaluer la pertinence et la faisabilité du développement d'une plateforme web unique qui permettrait d'optimiser la détection de risques sanitaires à partir des informations produites par les systèmes de biosurveillance. En 2009, quand EAR a été initié⁷⁸, la coopération dans le domaine de l'intelligence épidémiologique était très limitée qu'il s'agisse de la collaboration i) entre institutions de santé publique, ii) entre les systèmes de biosurveillance ou ii) entre institutions de santé publique et systèmes de biosurveillance. EAR a donc été la première initiative concrète et formalisée réunissant ces partenaires.

La plupart des systèmes de biosurveillance ayant été initialement développés sans associer les institutions de santé publique, EAR a permis de combler le fossé en réunissant utilisateurs et fournisseurs de systèmes de biosurveillance. Nos travaux ont notamment contribué à documenter la diversité des objectifs et des stratégies utilisées par les institutions de santé publique, de caractériser et de documenter leurs contraintes lors de l'utilisation des systèmes de biosurveillance fournissant ainsi aux systèmes des indications précises sur les améliorations à considérer.

Mais surtout, nos résultats ont permis d'orienter la stratégie de développement du projet⁷⁹ en répondant à une question essentielle : fallait-il améliorer le meilleur des systèmes ou intégrer les différents systèmes dans un outil unique? En démontrant les qualités intrinsèques des systèmes, leur complémentarité, et surtout en établissant les gains significatifs en termes de réactivité et de sensibilité qui résulteraient de la synergie des systèmes, nos résultats ont orienté la stratégie vers le développement d'une plateforme incorporant ces différents systèmes.

C'est sur cette base que lors de la seconde phase (2011-2012) un prototype intégrant les produits issus des différents systèmes de biosurveillance a été construit. Cette plate-forme commune, limitée d'abord à un nombre restreint d'agents biologiques, a permis d'explorer les mécanismes d'intégration des informations fournies par les systèmes participants tout en assurant de partager de l'évaluation de risque.⁸⁰ Le nombre d'agents biologiques a ensuite été progressivement augmenté. Une liste pilote d'agents chimiques a également été incluse afin de tester leur détection⁸¹.

En 2013, l'efficacité de la mutualisation de la détection de l'évaluation des risques d'origine biologique a été reconnue par les membres de la GHSI ; la pertinence de développer de maintenir et d'opérer une telle plateforme a été entérinée⁸². L'objectif de EAR est désormais de développer une plateforme opérationnelle permettant la détection de tous les types de risques (y compris les nouveaux agents)⁸³ avec un objectif de pérennisation à long terme. Les développements de la plateforme EAR se poursuivent en 2014.

IV.4 LIMITES DE L'ÉTUDE

IV.4.1 GOLD-STANDARD

IV.4.1.1 Données publiques Institutions Internationales

Dans le cadre de l'évaluation des systèmes de biosurveillance, le gold-standard utilisé devrait idéalement être représentatif de l'ensemble des menaces sanitaires survenues sur une période donnée, il ne devrait pas être tributaire de la volonté des pays de partager des informations et il devrait être disponible, sinon en continu, au moins dans un délai de temps raisonnable après la survenue des événements. Actuellement un tel gold-standard n'existe pas.

Si des données concernant les crises sanitaires existent, même en limitant le champ aux maladies infectieuses, la première difficulté réside dans la définition même d'une menace sanitaire. Nous avons vu que les objectifs de l'intelligence épidémiologique étaient très variables en fonction des institutions et que de nombreux facteurs sont à prendre en compte : les modalités techniques (nombre de systèmes, modalité d'utilisation, ressources, etc.), les critères de sélection des signaux bruts (pathologies infectieuses uniquement ou tous risques sanitaires), le type de pathologie, le lieu de survenue de la crise sanitaire ou les caractéristiques épidémiologiques du pays concerné.

La stratégie retenue par chaque institution influence fortement la nature des événements détectés et rapportés (cf. II.5.2, p 47). En 2012, sur son site dédié, l'OMS a publié des informations concernant 101 événements sanitaires dont seuls 1% concernaient un pays européen. En comparaison, sur une période comparable (février - décembre 2012), l'ECDC a communiqué sur plus de 400 événements sanitaires dont 50% concernaient un ou plusieurs pays européens. De même si les FHV représentaient 40% des événements rapportés par l'OMS et les maladies à prévention vaccinale 11%, ces dernières représentaient respectivement 1% et 33%% des événements rapportés par l'ECDC.

Les limites ne sont pas uniquement liées à la stratégie, mais également à l'accessibilité de ces données. Théoriquement, une institution comme l'OMS devrait être en mesure de fournir une base publique de référence, malheureusement en pratique cela n'est pas le cas.

En 2010-11, 122 événements survenus dans la région Europe de l'OMS ont été investigués par l'OMS, dont 73 (60%) dans des pays UE/EEE¹⁶. Sur cette même période seuls 8 ont fait l'objet d'une publication sur le site de référence de l'OMS (DON). Les raisons de ces différences sont multiples, d'une part tous les événements investigués ne nécessitaient probablement pas une communication internationale, mais surtout l'OMS applique des conditions de confidentialités extrêmement strictes. L'OMS ne publie des informations qu'avec l'accord du pays concerné, et c'est

certainement ce point qui explique en grande partie le manque de représentativité des données publiques de l'OMS. Le même phénomène s'applique aux autres organisations internationales comme l'OIE la FAO et dans une moindre mesure l'ECDC. Par ailleurs, ces données demeurent tributaires des caractéristiques sanitaires (accès aux soins, aux capacités diagnostiques, performance du système de surveillance, etc.) très variables suivant les pays affectés et qui impactent leur capacité à détecter, investiguer et confirmer les événements.

Ces données publiques ne doivent pas pour autant être systématiquement exclues. Dans ce contexte d'absence de « gold-stand reconnu », elles présentent certains avantages : elles sont facilement accessibles et elles constituent une source d'information officielle dûment vérifiées et fiables. En dépit de leurs limites et dans certaines conditions, elles peuvent donc être utilisées. En particulier, les données concernant certaines crises sanitaires majeures (grippe aviaire A/H5N1, pandémie grippale,...) peuvent constituer une option pour des études ponctuelles et ciblées. Durant la phase aiguë de ces crises, l'impact global et la pression internationale contribuent certainement à réduire la sous-notification. D'autre part, l'importance du volume d'information sanitaire produite par de nombreux pays contribue à réduire l'impact de la non-communication par certains pays. En revanche, ces données ne permettent d'évaluer la détection que d'une seule pathologie. Or, nous avons démontré que pour une même pathologie (la grippe aviaire A/H5N1) la sensibilité variait de manière importante entre les systèmes. En dehors de certaines crises sanitaires majeures, les seules données publiques des organisations internationales sont donc peu utilisables pour construire un gold-standard.

43,84,85,86

IV.4.1.2 Données Institutions Nationales

L'alternative utilisée pour contourner cette difficulté a consisté à utiliser comme gold-standard les données disponibles auprès d'une institution nationale. La même stratégie avait été utilisée par *Zeldenrust et al.*⁸⁷, pour estimer la valeur ajoutée de ProMED-mail dans le cadre du système d'alerte précoce aux Pays-Bas. Les données disponibles au niveau d'un institut de santé publique présentent plusieurs avantages : elles ne sont pas tributaires de la volonté des pays de partager les informations, les critères de sélection des événements et la méthodologie utilisée sont disponibles (si l'institution contribue à l'étude). En revanche, les données issues d'une seule institution nationale, sont tributaires des objectifs nationaux et des performances intrinsèques de l'équipe en charge de l'intelligence épidémiologique.

A ce titre, la comparaison des résultats obtenus en comparant les événements publiés dans le BHI dont la cible est le réseau français de santé publique et ceux publiés dans l'eWEB au bénéfice des 27 pays du réseau EpiSouth sont très illustratifs. On rappellera, que si ces deux activités VI-France et intelligence-épidémiologique EpiSouth, étaient réalisées par du personnel dédié, elles étaient mises en œuvre dans un même département par deux équipes travaillant en étroite coordination (pas de duplication). La méthodologie utilisée est donc exactement la même à l'exception d'un aspect important les critères de sélection des événements. Par exemple si une épidémie de dengue dans un des États insulaires du Pacifique présente un intérêt pour la France en raison du risque de diffusion aux trois territoires français du Pacifique, cet événement ne constitue pas une menace pour les pays EpiSouth. De même, une épidémie de choléra au Kurdistan Irakien représente risque sanitaire direct pour les pays frontaliers EpiSouth (Jordanie et Turquie) mais pas pour la France.

Les différences observées en matière de type d'évènements rapportés dans le BHI et l'eWEB sont donc uniquement attribuables aux critères de sélection des évènements. S'il y a peu de différence en termes de type de pathologie, les différences géographiques sont importantes : 73% des évènements (hors A/H5N1) rapportés dans l'eWEB concernent un des 27 pays EpiSouth contre 23% pour le BHI, 32% des évènements rapportés dans le BHI sont survenus en Amérique ou dans le Pacifique contre seulement 2% de ceux rapportés dans l'eWEB. Il faut noter que ces résultats sous-estiment largement l'impact de critères de sélection, la France faisant partie du réseau EpiSouth, plus de 20% des évènements étaient pertinents et pour la France et pour EpiSouth.

Dans notre étude, seules les données de l'InVS ont été utilisées pour construire le gold-standard, ces résultats ne peuvent donc pas être considérés comme extrapolables à toutes les situations. Un gold-standard élaboré à partir des données d'autres institutions aurait certainement été différent (en termes de type d'évènements, de leur distribution géographique, etc.). Comme nous avons pu démontrer, la nature des pathologies et leur lieu de survenue influencent significativement les performances de détection des systèmes de biosurveillance, les mesures de sensibilité auraient probablement été différentes. Si ce biais limite la reproductibilité des résultats avec un autre gold-standard et l'extrapolabilité des résultats, il n'en diminue pas pour autant leur pertinence. Les résultats reflètent simplement les performances des systèmes dans un cadre donné et surtout ils fournissent les seules indications actuellement disponibles permettant de comparer simultanément les performances de six systèmes utilisant un même gold-standard multi-pathologie.

IV.4.2 RISQUES SANITAIRES ET CHAMP GEOGRAPHIQUE

Notre étude ne couvre qu'une partie du domaine potentiellement couvert par l'intelligence épidémiologique et la surveillance des évènements. En raison de la nature des activités menées par les institutions participantes et d'autre part des activités de VI du DI de l'InVS, seul l'aspect risque sanitaire international a pu être évalué. Néanmoins, le champ d'application n'est pas limité à ce type de risque. Plusieurs institutions, comme l'ISS en Italie ou l'ECDC, appliquent l'intelligence épidémiologique à la détection des risques sanitaires sur leur territoire national (ou à l'échelle européenne pour l'ECDC).

Par ailleurs, sur les 10 systèmes identifiés en 2010, 7 ont pour objectif prioritaire la détection des menaces sanitaires d'origine infectieuse (cf. Tableau 3, p 41). Dans notre étude, le plus petit dénominateur commun a été retenu afin d'assurer une comparabilité des résultats, les systèmes pour lesquels la détection des évènements d'origine infectieuse n'était pas un objectif prioritaire ont donc été exclus. L'évaluation de la détection des risques non-infectieux, notamment chimiques et environnementaux qui font partie intégrante de l'intelligence épidémiologique, reste donc à réaliser.

IV.5 LIMITES DE L'UTILISATION DES SOURCES INFORMELLES

IV.5.1 RESEAU SOCIAUX ET SURVEILLANCE PARTICIPATIVE

Les informations issues des réseaux sociaux (notamment Twitter[®] et Facebook[®]) n'ont pas pu être prises en compte dans nos travaux. L'utilisation de ce nouveau mode de communication est encore récente mais suit une courbe exponentielle. D'après Wikipédia, le nombre de « tweets » est passé de 400,000 « tweets » par trimestre en 2007, à 65 millions de « tweets » par jour en mars 2010, et à 400 millions en septembre 2013 (source Wikipédia http://en.wikipedia.org/wiki/Twitter#cite_note-34). Actuellement, aucun système de biosurveillance n'est en mesure de proposer un traitement même rudimentaire des données issues des réseaux sociaux. Des projets visant à explorer le traitement de ce nouveau type de données existent, mais ils sont encore limités au domaine de la recherche.

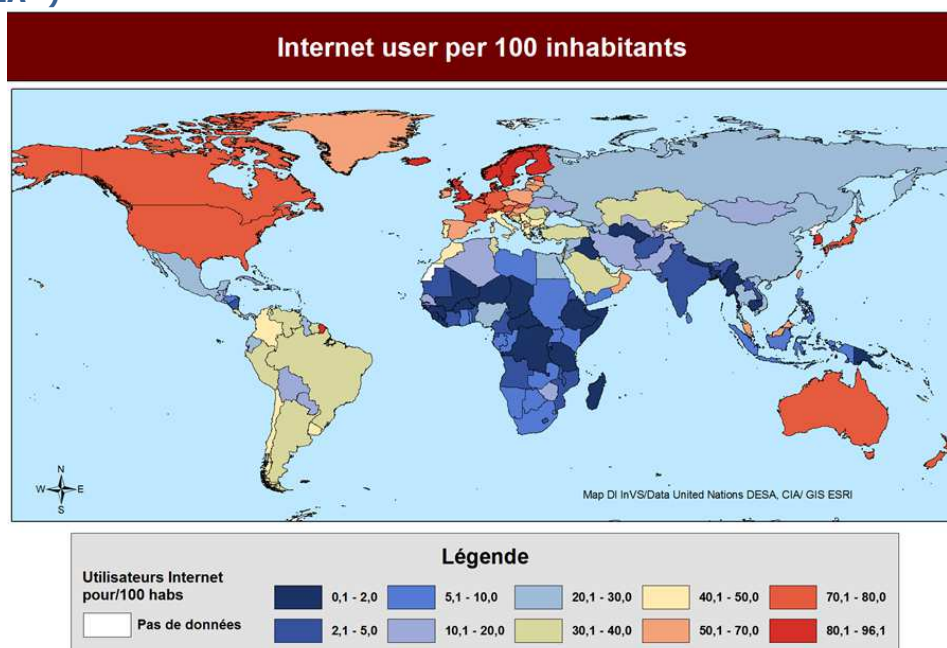
En revanche, des initiatives dites de surveillance participative ont déjà été initiées notamment dans le cadre de la surveillance de la grippe saisonnière. Ces projets visent à recueillir directement auprès de la population (recueil volontaire et anonyme) des données épidémiologiques sur la grippe. Ces données recueillies, via un compte Internet personnalisé et sécurisé sont destinées à compléter les données de la surveillance traditionnelle provenant des systèmes de soins (qui ne tiennent pas compte des malades qui ne consultent pas). Ces projets de recherche tels que « Gripnet » (<https://www.grippenet.fr/>) mis en place par l'Inserm et l'Université Pierre et Marie Curie, en coordination avec l'InVS, « Influenzanet » (<https://www.influenzanet.eu/>) à l'échelle Européenne, ou « Flu Near You » (<https://flunearyou.org/>) développé par HealthMap pour les États-Unis s'ajoutent à d'autres initiatives comme « Google Flu Trends » (<http://www.google.org/flutrends/>) qui utilise les fortes corrélations entre les recherches effectuées sur le moteur de recherche Google® et le pourcentage de patients consultant un généraliste pour un syndrome grippal.⁸⁸

Si, ces projets semblent très prometteurs notamment pour compléter la SBI pour des pathologies bien connues et fréquentes, l'utilisation des réseaux sociaux et de la surveillance participative dans la détection précoce d'autres événements sanitaires reste encore à démontrer. La nature même de ces nouveaux modes de communication les rend actuellement inutilisables. Le traitement de ce type d'information présente encore plus de contraintes que celle des médias traditionnels. Par exemple, les messages Twitter sont limités à 140 caractères, ce qui laisse très peu de place pour l'analyse du contenu, Facebook contient de nombreuses données personnelles ce qui soulève la question de la confidentialité. Le langage souvent familier est également générateur d'un volume important de faux signaux (par exemple dans de nombreux langages le mot peste a un sens très différent suivant le contexte médical ou familial). Il est également vraisemblable que pour ce type de communication destinée à un « réseau personnel », l'idiome utilisé soit plus fréquemment la langue ou le dialecte maternel que cela n'est le cas actuellement sur Internet. Ces nouveaux modes de communications sont néanmoins une réalité et pourraient potentiellement fournir des informations utiles dans le cadre de la détection d'événements sanitaires. Dans le futur, la pertinence de leur utilisation devra être considérée et évaluée dans le cadre d'études spécifiques.

IV.5.2 ACCES A L'INFORMATION BRUTE

L'intelligence épidémiologique en général et l'utilisation des systèmes de biosurveillance en particulier sont tributaires d'Internet. Or, à l'échelle planétaire, l'accessibilité à Internet est très inégalement répartie. Dans de larges zones du globe, en particulier en Afrique, en Asie du Sud et en Amérique Centrale, l'accès à Internet est très restreint (Figure 9). Ces zones de faible couverture Internet se superposent souvent à des zones rurales et/ou isolées où l'accès aux soins de santé primaire et aux capacités diagnostiques sont les plus faibles et donc où la surveillance épidémiologique traditionnelle est la moins à même de détecter précocement événements sanitaires graves. Or, c'est précisément dans ces régions que l'intérêt de l'intelligence épidémiologique est le plus grand. La plupart des régions mal couvertes par Internet sont situées dans la zone intertropicale où l'écologie et les interactions entre l'homme, les environnements encore « vierges » sont les plus favorables à l'émergence de risques sanitaires majeurs (ex : virus Ebola en Afrique équatoriale, nouveau virus grippal en Asie du Sud, etc.). C'est également dans ces zones que le nombre de conflits armés est le plus important. La déstructuration des structures de soins, le déplacement massif de population et la détérioration des conditions de vie qu'ils engendrent sont autant de facteurs favorisant l'apparition d'épidémies. En raison du nombre peu élevé d'utilisateurs, la presse et les autres supports accessibles par Internet (ex : les blogs) sont très peu développés voire inexistantes. Quand elles sont disponibles, les informations ne couvrent souvent qu'une partie du territoire national (la capitale et les grandes zones urbanisées).

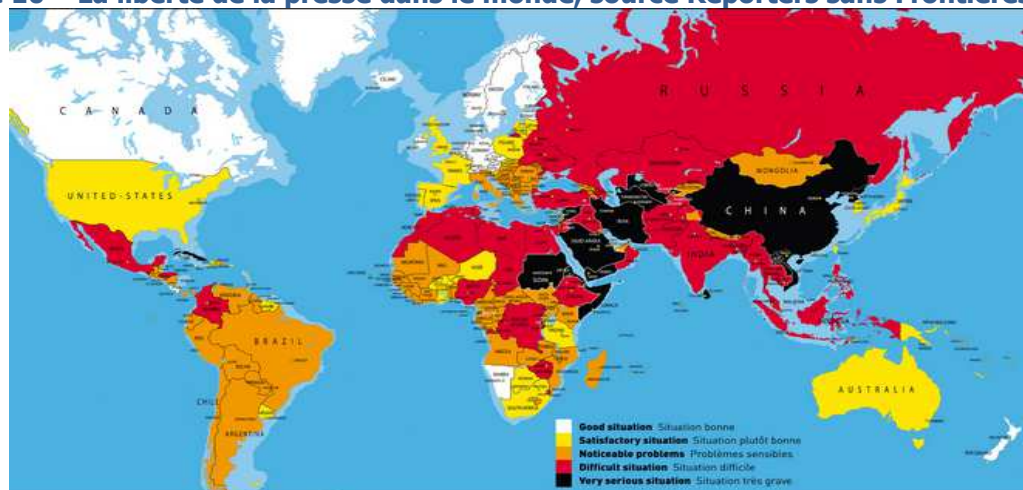
Figure 9 – Répartition des utilisateurs d'Internet rapporté à 100 habitants, 2012 (données brutes CIA⁸⁹)



La seconde limite est constituée par la liberté de communication et notamment la liberté de la presse. Dans de trop nombreux pays (Figure 10), la presse est contrôlée par les gouvernements et les données sanitaires n'échappent pas à la censure.

Prenons un exemple : dans l'étude portant sur les facteurs influençant la performance des systèmes de biosurveillance (Article - 2) un évènement a été détecté seulement 169 jours après son début présumé. Il s'agissait d'une épizootie de Fièvre de la Vallée du Rift en Arabie Saoudite qui n'a été communiquée que le jour de la détection d'un cas humain. La sous-détection ou l'absence de communication concernant les épizooties de Fièvre de la Vallée du Rift n'est pas un phénomène nouveau⁹⁰. Dans ce pays classé 163^e sur 179 dans le classement mondial de la liberté de la presse 2013 établi par Reporters Sans Frontières⁹¹, on ne peut exclure que des informations concernant l'épizootie aient été disponibles bien plus tôt mais qu'elles n'aient pas été rendues publiques. Enfin certains pays cumulent une absence de liberté de la presse et un accès très limité à Internet et constituent des zones particulièrement préoccupantes. Dans ces zones, les sources primaires d'information sont souvent réduites à quelques médias internationaux dont l'accès au terrain est limité ce qui constitue une limite importante pour la détection des éventuelles menaces sanitaires.

Figure 10 – La liberté de la presse dans le monde, source Reporters sans Frontières⁹¹



V. CONCLUSION ET PERSPECTIVES

Nos résultats ont permis d'établir les fondations de l'évaluation des dispositifs d'intelligence épidémiologique et des performances des systèmes de biosurveillance. Ils ne constituent cependant qu'une première étape. D'autres études devront être menées pour étudier en profondeur les différentes facettes de ce domaine complexe.

gold-standard. Les futures études devront s'attacher à construire un gold-standard multi-pathologie intégrant les données de différentes institutions nationales et internationales. Les obstacles qui devront être surmontés sont nombreux : accessibilité aux informations concernant les évènements, mode de sélection, compatibilités des données, confidentialité, transparence. Etant donné la multiplicité des besoins, et les moyens limités des utilisateurs, une approche internationale permettant de formaliser la collaboration entre les institutions internationales (comme l'OMS, l'OIE, ou l'ECDC) et des institutions nationales représente la meilleure stratégie afin d'obtenir un gold-standard représentatif tant du point de vue géographique que du type d'évènement sanitaire.

Volume d'information. Si nos résultats illustrent les performances honorables des systèmes en matière de sensibilité, ils illustrent aussi les contraintes posées par le volume d'information et la faible VPP des informations brutes détectées. Or, pour être effectivement applicable, l'amélioration de la sensibilité des systèmes devrait impérativement s'accompagner d'une amélioration importante de la VPP des signaux bruts détectés. La perspective du développement d'un système combiné intégrant les principaux systèmes de biosurveillance existants représente une réelle solution d'avenir. Si nous avons pu démontrer qu'un tel système permettrait des gains importants en termes de sensibilité et de réactivité, le gain spécificité ne pourra être mesuré que lorsqu'un tel système sera opérationnel.

Etudes prospectives. Nos résultats mettent en évidence l'importance d'envisager la mise en œuvre d'études prospectives en dépit des difficultés qu'elles présentent et notamment des ressources nécessaires. En effet, seules des études prospectives permettront de mesurer avec précision les performances des systèmes dans des conditions réelles d'utilisation.

Approche « tous risques » sanitaires. Les évènements notamment chimiques et radiologiques et nucléaires représentent également des risques majeurs pour la santé publique et la performance de leur détection n'a pas pu être évaluée dans nos travaux. Là encore, une approche multicentrique et internationale devra être utilisée et de nouveaux acteurs, par exemple l'Agence Internationale de l'Energie Atomique (AIEA), les centres antipoison, les centres d'excellence, ou les instituts chargés de la pharmacovigilance devront y être associés afin de disposer des données de référence nécessaires.

Type de source. Dans les zones où différents types de sources informelles (notamment médias) sont disponibles à l'échelle internationale, nationale, régionale et locale, les sources régionales et locales semblent plus réactives. Cette perception ne repose cependant que sur la pratique et n'a pas pu être mesurée, une étude approfondie devrait être envisagée afin de mesurer la réactivité des différents types d'information.

Langue de publication. Nos travaux ont démontré que la langue de publication était l'un des facteurs qui influençait les performances de détection. Malheureusement, cet aspect n'a pas pu être étudié en détail et seules les grandes langues internationales ont pu être prises en considération. La prise en considération des nouvelles langues par les systèmes est un processus complexe et coûteux. Avant de la recommander, l'estimation de la disponibilité des sources potentielles et de leur utilisation dans le domaine de l'intelligence épidémiologique devrait être considérée comme un prérequis.

Accès à l'information brute. La radio constitue une source alternative d'information qui n'est actuellement pas exploitée. Dans de nombreuses zones reculées, notamment en Afrique, la radio constitue toujours un support de communication essentiel y compris au niveau local. Les techniques de reconnaissance de la parole ont beaucoup évolué et leur utilisation devrait donc être considérée comme un axe de développement prioritaire par les systèmes de biosurveillance. Elles constituent à court terme la seule alternative réaliste pour palier la pénétration largement insuffisante d'Internet dans de larges parties du monde. Ces technologies pourraient également permettre l'intégration de langues régionales. Certaines langues (comme le swahili en Afrique de l'Est, le malais/indonésien ou le bengali en Asie) dépassent très largement les frontières ou sont la langue maternelle de millions de locuteurs dans des zones particulièrement vulnérables.

Evaluation des dispositifs d'intelligence épidémiologique. L'évaluation des performances des dispositifs d'intelligence épidémiologique mis en œuvre par les institutions de santé publique devrait être réalisée. Elle permettrait d'identifier les forces, les faiblesses et d'optimiser l'utilisation des ressources. Plus encore que pour l'évaluation des systèmes de biosurveillance, l'évaluation de l'intelligence épidémiologique est confrontée à l'absence de gold-standard. Les informations disponibles actuellement permettent tout au plus d'identifier rétrospectivement des manquements importants (par exemple la non-détection d'un événement grave). Elles ne permettent pas d'envisager la mise en œuvre unilatérale d'une évaluation formelle. Une telle évaluation ne pourrait être envisagée que lorsque plus de données seront disponibles dans le cadre d'une collaboration internationale.

Evaluation globale et mesure d'impact. L'évaluation de l'ensemble du processus d'intelligence épidémiologique devrait également être envisagée afin d'évaluer : les ressources nécessaires pour mettre en œuvre le système (indicateurs d'intrant), la mise en œuvre (indicateurs de processus) et de ses résultats (indicateurs de produit), les performances du dispositif d'intelligence épidémiologique (indicateurs de résultat) et enfin de l'impact sur la santé publique (indicateurs d'impact). La mesure de l'impact de l'intelligence épidémiologique sur la santé publique devrait être mesurée au niveau de chaque institution/pays. Les délais entre d'une part, la détection et la notification aux autorités sanitaires concernées et d'autre part les délais entre la notification et la mise en œuvre des mesures de contrôle ou de mitigation sont des indicateurs d'impact relativement simples à mesurer. Ces évaluations dépassent largement le cadre des seules équipes en charge de l'intelligence épidémiologique et doivent intégrer l'ensemble des acteurs impliqués dans la détection et la réponse aux menaces sanitaires au niveau national.

Collaboration Internationale. Un partage d'information est régi par des législations internationales (RSI pour la l'OMS ou la directive 1082/2013⁹² pour l'échange d'informations au sein de l'UE), mais nous avons vu qu'il était encore très insuffisant. Des initiatives existent, par exemple le projet EAR, la collaboration EpiSouth-InVS (II.5.2), la collaboration tripartite OMS-OIE-FAO (projet GLEW). D'autres collaborations plus informelles peuvent exister notamment entre les instituts nationaux, mais ces dernières, souvent basées sur la confiance mutuelle sont trop personnes-dépendantes pour être considérée comme pérennes. Toutes ces initiatives sont louables, mais demeurent largement insuffisantes pour d'optimiser l'utilisation des ressources et réduire l'inutile duplication des activités.

Lors de l'émergence d'un nouveau virus ou d'une crise sanitaire majeure, et alors que les systèmes conventionnels de surveillance (SBI) ne sont pas encore opérationnels, l'intelligence épidémiologique joue un rôle déterminant pour recueillir des informations cruciales à l'adaptation des mesures de contrôle et de mitigation comme la diffusion géographique, les caractéristiques épidémiologiques, l'identification de groupes à risque, etc. Ces informations sont essentielles pour prioriser le diagnostic des patients les plus susceptibles d'être infectés, élaborer les mesures de contrôle adaptées (isolement des contacts, restriction des voyages, etc.). A l'heure d'Internet, ce type d'activité devrait pouvoir être mutualisé.

Aujourd'hui, du fait de la rapidité de la circulation de l'information et des pressions des autorités politiques, les institutions chargées de l'intelligence épidémiologique doivent être mesure de répondre « quasiment » en temps réel aux multiples sollicitations. Elles ne peuvent donc plus rester dépendantes des seules informations et des analyses fournies par une autre institution, fusse l'OMS. Faute de mécanisme permettant une mutualisation sécurisée des contributions dans ce domaine, les institutions demeureront contraintes de continuer à dupliquer inutilement des activités chronophages. Il est notamment absurde que toutes les institutions comptent « les mêmes cas d'influenza survenant dans le monde ». Cette duplication est sans nul doute, moins efficace et plus coûteuse que la mutualisation des moyens. Malgré, les leçons de la pandémie A/H1N1, on peut déplorer le manque d'initiatives qui permettraient une mutualisation des moyens et des informations ainsi que l'optimisation de l'utilisation des ressources, ce qui au final serait le meilleur garant d'un renforcement pérenne de la sécurité sanitaire mondiale.

Publications scientifiques et partage d'information. Le nombre d'articles publiés dans des revues scientifiques fait partie des indicateurs de performance utilisés notamment pour déterminer les ressources allouées à de nombreuses structures (laboratoires, hôpitaux universitaires, centres de recherche, etc.). La publication d'articles est donc essentielle pour ces structures. On peut néanmoins déplorer que le partage de l'information avec la communauté scientifique soit conditionné à la parution d'un article scientifique. Si des procédures accélérées existent au niveau de certains journaux pour permettre une publication plus rapide, elles ne permettent pas une mise à disposition des informations en temps réel. La publication de résultats importants, notamment pour la mise en œuvre de mesures de contrôle, est souvent retardée par ces impératifs de publication. Des mécanismes devraient pouvoir être mis en place pour permettre d'une part un partage sécurisé de l'information vers les autorités sanitaires sans pour autant obérer une éventuelle publication scientifique ultérieure.

En conclusion, au-delà de la simple utilisation des outils, c'est bien le concept même de l'intelligence épidémiologique qui a pu être documenté par nos travaux. En anticipant substantiellement la détection des événements sanitaires graves, l'intelligence épidémiologique complète efficacement la surveillance traditionnelle et contribue effectivement à la consolidation des dispositifs de surveillance nationaux et internationaux et ainsi concourt au renforcement de la sécurité sanitaire mondiale. Aucune institution ne dispose à elle seule des moyens nécessaires pour couvrir l'ensemble des menaces sanitaires potentielles. La mise en œuvre de stratégies de coopération internationale devrait être considérée comme une priorité afin de bénéficier au mieux de ces synergies tout en assurant une meilleure utilisation des ressources disponibles. Nos résultats suggèrent des applications concrètes. Les maladies infectieuses, les questions environnementales, et les menaces bioterroristes potentielles continueront de poser des menaces importantes pour la sécurité sanitaire nationale et mondiale. L'intelligence épidémiologique est désormais une composante essentielle des systèmes d'alerte précoce. Pour répondre à ces enjeux, la poursuite de la recherche dans ce domaine est nécessaire.

ANNEXE 1 – ARTICLE A

REVIEW

10.1111/1469-0691.12273

An overview of Internet biosurveillance

D. M. Hartley^{1,2,*}, N. P. Nelson^{3,*}, R. R. Arthur⁴, P. Barboza⁵, N. Collier^{6,7}, N. Lightfoot⁸, J. P. Linge⁹, E. van der Goot⁹, A. Mawudeku¹⁰, L. C. Madoff¹¹, L. Vaillant⁵, R. Walters¹², R. Yangarber¹³, J. Mantero¹⁴, C. D. Corley¹⁵ and J. S. Brownstein¹⁶

1) *Imaging Science and Information Systems Center, Georgetown University School of Medicine*, 2) *Department of Microbiology and Immunology, Georgetown University Medical Center*, 3) *Department of Pediatrics, Georgetown University Medical Center, Washington, DC*, 4) *Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA*, 5) *International Department, French Institute for Public Health Surveillance (InVS), Saint Maurice, France*, 6) *The National Institute of Informatics, Tokyo, Japan*, 7) *The European Bioinformatics Institute, Hinxton, Cambridge, UK*, 8) *Connecting Organizations for Regional Disease Surveillance (CORDS), Lyon, France*, 9) *Joint Research Centre (JRC) of the European Commission, Ispra, Italy*, 10) *Public Health Agency of Canada (PHAC), Ottawa, ON, Canada*, 11) *University of Massachusetts Medical School, Worcester, MA*, 12) *Pacific Northwest National Laboratory, Richland, WA, USA*, 13) *Department of Computer Science, University of Helsinki, Helsinki, Finland*, 14) *Surveillance and Response Support Unit, European Centre for Disease Prevention and Control, Stockholm, Sweden*, 15) *Pacific Northwest National Laboratory, Richland, WA* and 16) *Harvard-MIT Division of Health Sciences and Technology, Children's Hospital Boston, Harvard Medical School, Boston, MA, USA*

Abstract

Internet biosurveillance utilizes unstructured data from diverse web-based sources to provide early warning and situational awareness of public health threats. The scope of source coverage ranges from local media in the vernacular to international media in widely read languages. Internet biosurveillance is a timely modality that is available to government and public health officials, healthcare workers, and the public and private sector, serving as a real-time complementary approach to traditional indicator-based public health disease surveillance methods. Internet biosurveillance also supports the broader activity of epidemic intelligence. This overview covers the current state of the field of Internet biosurveillance, and provides a perspective on the future of the field.

Keywords: Digital disease detection, digital epidemiology, electronic surveillance, epidemic intelligence, event-based surveillance, Internet-based surveillance, participatory epidemiology, web-based text mining

Clin Microbiol Infect

Corresponding author: D. M. Hartley, Imaging Science and Information Systems Center, Georgetown University Medical Center, 2115 Wisconsin Avenue NW, Suite 603, Washington, DC 20057, USA

E-mail: hartley@isis.georgetown.edu

*These authors contributed equally to this study.

Introduction

Internet biosurveillance, or digital disease detection [1], utilizes unstructured data from diverse web-based sources to provide early warning and situational awareness of human, animal and plant infectious diseases, as well as chemical, radiological and nuclear threats [2]. The discipline emerged in the mid-1990s, relying primarily on text media for its information, and has evolved into a globally recognized field [3,4]. With the increasing volume of information and new media types available via the Internet, the field has grown to include social

media, participatory sources, and non-text-based sources. The scope of source coverage ranges from local media in the vernacular to international media in widely read languages. Online official reporting sources are typically used to supplement and verify such informal Internet sources.

Internet biosurveillance is a timely modality that is available to government and public health officials, healthcare workers, and the public and private sector, serving as a real-time complementary approach to traditional indicator-based public health disease surveillance methods [5,6]. Internet biosurveillance also supports the broader activity of epidemic

intelligence (EI). This review covers the current state of the field, and provides a perspective on its future.

Methods

This is not a 'systematic review'; rather, this article outlines a general process of Internet biosurveillance according to established best practices, and discusses common technologies employed in extant systems. Each step of the process is collectively described, drawing upon personal experiences of system builders and practitioners, as well as published studies. The authors contributing to this article are either affiliated with Internet biosurveillance systems, are end-users of Internet biosurveillance systems, and/or have published recently in the field. Authors from the following active Internet biosurveillance systems are represented: BioCaster [7], the Global Public Health Intelligence Network (GPHIN) [8], HealthMap [9], the Medical Information System (MedISys) (Steinberger *et al.*, IDRC, 2008, Short and Extended Abstracts, pp. 612–614, <http://publications.jrc.ec.europa.eu/repository/handle/111111111/13078> (accessed 9 February 2013)), the Program for Monitoring Emerging Diseases (ProMED-mail) [10], and the Pattern Understanding and Learning System (PULS) [11].

Results

The process of Internet biosurveillance varies, but, in general, includes: (i) the collection and storage of data from the Internet; (ii) processing those data to produce information; (iii) assembling that information into analyses; and (iv) dissemination of analyses to end-users (Fig. 1). Each part of the process can entail many technical steps, which are described below. Information vetting can occur through fully automated, human-moderated or partially moderated approaches throughout the process. Multilingual data are managed via human linguists, machine translation, and natural language-processing technology.

Collection and storage

Data sources. Internet biosurveillance systems rely on data from a variety of sources. Publicly available, informal sources include text-based news sites (e.g. *New York Times* and *Thanh Nien News*) and social media sources (e.g. Twitter [12], Facebook, and blogs); more recently, sources that utilize public input (e.g. FluTrackers, Flu Near You, and crowd-sourcing platforms [13]) have gained popularity and credibility. Information from these sources is often available in real time as an event is developing. This information is validated

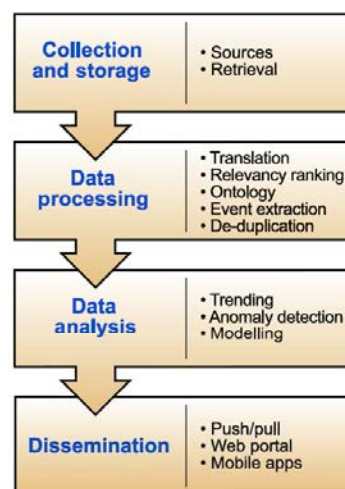


FIG. 1. The general process of Internet-based biosurveillance. Human input from information technology, public health and other experts can occur at any step.

and supplemented by official, publically available information sources (e.g. public health agencies, ministries of health, the WHO, the World Organization for Animal Health, and the Food and Agriculture Organization). Systems also may utilize sources with paid content (e.g. newswires and news aggregators). Audio and video sources provide non-text-based information. Sources range widely in geographical coverage, from local to international, and cover all languages with publicly available media.

Data retrieval. Data are retrieved from the Internet via two predominant modalities: media aggregators and system-specific web monitoring. As an example of the latter, Internet biosurveillance systems monitor the web by scraping (that is, specific web pages are accessed and stored) or crawling (that is, in addition to storing one specific web page, links on that page and links of links are accessed and stored).

Systems re-visit a list of predefined sites at regular intervals (typically, once to several times each day) in order to process data in a timely manner for early alerting. For paid or access-limited content, items might be accessed via a secure connection. News items from online news sites and social media are converted to a common format after retrieval, to enable searching and content mining. Public health agencies and ministries of health often provide their own feeds with official information. Feeds from aggregator news sites (e.g. Google and Yahoo) can be used to provide additional coverage. Content is extracted from the HTML code, with proper removal of advertisements and any other irrelevant text.

Social media data stem mostly from Twitter [14] and Facebook, which can be retrieved via their application programming interface. Access may be limited to a certain volume, and is subject to change according to the provider's Terms of Service. As some social media users are unaware that they publish their opinions worldwide, privacy issues arise under some jurisdictions, even with the publicly available data. Participatory data can be included via dedicated apps (e.g. iPhone and Android) or websites where users can leave comments (e.g. <http://www.flutrackers.com/>; <http://www.healthmap.org/outbreaksnearme/>) [15,16].

Data processing

Once data are retrieved from the Internet, they must be processed to make them amenable for analysis. We emphasize that, because different types of users have different needs, there is no single, overarching goal for the data-processing step. Nevertheless, the following categories represent important steps in biosurveillance data processing: translation, relevancy ranking, ontology, event extraction, and de-duplication.

Translation. Although Arabic, Chinese, English, French, Spanish and Portuguese dominate the world's online news media, news of an outbreak event can appear in any language, and is often reported first in a local language. Systems have choices to make regarding the approach to translation. For example, they can build customized pipelines for a few languages, or they can translate each source language into a common target language. The decision is influenced by factors such as the availability of resources in each language, the time available to maintain each resource, and the translation quality required. For example, BioCaster employs full text translation first and uses only English language selection algorithms, whereas MedSys and HealthMap are language-specific in terms of the keywords employed to search Internet data. GPHIN employs both language-specific keywords and algorithms to extract relevant data from the Internet and news aggregator databases [17], whereas PULS employs language-specific linguistic analysis and ontologies and inference rules to extract relevant data.

Relevancy ranking. The next stage in processing is to assess the relevancy of the report according to some measure of the user's interest. Defining the user's interest as a set of guidelines, a decision tree or as a collection of examples is a crucial stage in system building, and provides a reference standard against which to evaluate various algorithms. Once this has been done, various approaches can be implemented, including supervised classifiers such as Naïve Bayes or Support Vector Machines with learn-to-rank, and Boolean keyword

searches, which include logical operators such as AND and OR [18]. These techniques are language-specific, but it is also possible to deploy automated methods that are language-independent, such as clustering followed by automated labelling.

Ontology. Ontologies have proven useful in many domains (e.g. the life sciences) for structuring relationships between concepts. Biosurveillance requires a conceptual knowledge of diseases, microorganisms, signs and symptoms, and geography. A number of ontological resources have been developed or re-used for public health, although these are not generally as well known as those in experimental biology or clinical fields, such as the Unified Medical Language System. Among those developed specifically for public health are GIDEON (commercial, openly available), BioCaster (open source), and GPHIN (non-commercial, limited access). Such ontologies provide knowledge needed by Internet biosurveillance systems to make intelligent judgements about the terms appearing in news reports. For example, a mention of *Yersinia pestis* may imply that the disease under consideration is bubonic plague. However, not all ambiguities can be resolved with the static knowledge contained in an ontology. One of the most practical problems is toponym disambiguation (i.e. place names). For example, a mention of a disease outbreak in 'Cambridge' might resolve to any of several places worldwide, including the UK or the USA.

Event extraction. Once a set of topics of potential interest has been identified, specific biological events are extracted from the data. This can be accomplished in different ways. As one example, simple keyword recognition algorithms are often used to categorize incoming news items. In this approach, an article is categorized according to predefined keywords (see example in Table 1). Boolean combinations (e.g. AND, OR, NOT) and proximity searches (i.e. search for articles where two or more separately matching term occurrences are within

TABLE 1. Examples of multilingual keywords used for identification of dengue fever in MedSys

Keywords
dengue
Denguefieber
дэнге́ра.ик%+деиге
登革熱
دنگی
δάγκειος+πυρετός
Ιός+δάγγειου
knokkelkoorts
febre+hemorragica
fiebre+hemorragica
hemorrhagic+fever
...

a specified word or character distance) can then be applied [19].

More detailed aspects of an outbreak can be extracted by event meta-data extraction, in which the aspects of interest are known and defined *a priori*. Examples of commonly detected aspects include the name of the disease, the species affected, the date of the outbreak, the numbers of cases and deaths, and the location of the outbreak. Event meta-data extraction uses the extensively researched technology known as information extraction, which is the basis of PULS and BioCaster [11]. Less common aspects include distal indicators of political and social response, such as ward closures or the deployment of international organizations to the affected region. Often, the techniques used are linguistic patterns developed with specific rule systems, but supervised, semi-supervised and unsupervised machine-learning approaches have also been evaluated [20].

De-duplication. Effective de-duplication is essential for events with wide coverage, so that nearly identical stories appearing in many sources do not overwhelm the user. De-duplication may involve the detection of reports that are identical in content, which are handled in practice with clustering techniques as outlined above. Reports may also be identical in the aspects of the outbreak that they report. De-duplicating these reports in practice is challenging, and can require deeper-meaning analysis. Nevertheless, there are often subtle but important aspects of an event that may not be easily captured, such as the revision of victim numbers, the change in a patient's condition, or a comparison between a novel and a known agent. De-duplication should ideally be sensitive to these grey areas, and pass forward such articles for human analysis.

Data analysis

At this stage of the process, a biosurveillance system will have produced a structured collection of events that are potentially relevant to end-users. However, only a subset of these may be highly useful, given a particular user's interests. For example, a case of seasonal influenza in a celebrity, although widely reported, may be less relevant than a few reports of a cluster of novel influenza among farmers. Given the conflict between the volume of data to be analysed and the limited ability of humans to review large amounts of information quickly, it is often desirable to process the articles through an automated trend and anomaly detection capability in order to increase throughput and timeliness. The objective is to infer which events are more urgent or unusual in a timely manner, so that the user can investigate further and potentially initiate risk analysis. The challenge is to model what is already known (i.e. what is normal or expected), and to decide whether the

current event is significantly at variance as early as possible. We focus on two complementary classes of approach in this section: trend analysis and anomaly detection.

Trend analysis. The temporal nature of Internet biosurveillance data produces longitudinal patterns and trends. Precursors and indicators of outbreaks can be tracked over time to show the precedence of an event before symptoms or the populace pass thresholds for warning. Timelines can also be used to track classifiers, keywords, locations, or terms, and indicate temporal traces of events for significance against predefined baselines. Visualizing topical trends and shifts over time based on such lexicons can facilitate the detection of unexpected disease events. Standard time-series algorithms and other signal-processing techniques are often used to model these temporal trends [21–23].

Anomaly detection. Anomaly detection attempts to put the features of the event into context in order to determine some level of significance. Context is usually considered to be spatial and/or temporal or a mixture of the two, and can be based on simple event counts of a particular disease type or on multiple features of the event. However, in situations where terminology begins to specialize or diverge (e.g. 'mad cow' to 'bovine spongiform encephalopathy', or 'swine flu' to 'H1N1'), the anomaly detection can be attenuated.

Dissemination

Achieving the ultimate public health goals of biosurveillance systems—to facilitate early outbreak detection, thereby allowing timely interventions, limiting the severity and extent of spread—depends on the clear and rapid distribution of information. Internet-based biosurveillance systems use different means of disseminating information, depending on user needs and resources and the nature of the information.

Most systems use a combination of actively 'pushing' material to users and allowing users to 'pull' material when desired. ProMED-mail, one of the earliest Internet-based biosurveillance systems, uses mailing lists (e-mail) and listserv software, where users can subscribe to specific resources (e.g. animal or plant diseases). GPHIN uses a pushing function to send alerts about events that have been identified as significant to subscribers. Some services (e.g. HealthMap) allow users to specify parameters for pushed information, such as specific diseases, categories of disease, and geographical locations. SMS text messages, mobile telephone networks and social networks (e.g. Twitter) actively send information to anyone subscribing to a feed.

In addition, most Internet biosurveillance systems have a dedicated website where users may query and filter material on demand. Although they are passive, websites allow users to

obtain specific information when it is needed, and they usually provide the capacity to search for specific data (e.g. specific disease categories, locations, or time periods). Geographical mapping, which is automatically generated and displayed by several current systems, allows users to visualize clustering of events over time and space. More recently, smartphone apps have been developed that allow a combination of active and passive dissemination of information (and also allow users to report data back to the system).

With the rationale that it is not always possible to predict who will need a specific piece of information, many systems make their data available freely to anyone. Other systems make their information available to selected groups or individuals. Selectivity of dissemination may be based on the need to restrict access to confidential information (e.g. the Epi-X system of the US CDC, which is available only to vetted public health officials), or a paid subscription model may be used in order to recoup the costs of creating and maintaining the system.

Illustration of Internet biosurveillance: Madeira Island dengue fever outbreak, October 2012

To illustrate how an event is detected and observed to evolve through the lens of an Internet biosurveillance system, consider the October 2012 dengue fever outbreak in the Autonomous Region (island) of Madeira, a Portuguese territory located approximately 1000 km from the mainland [24]. It was the first dengue outbreak in Europe since 1928. With the keyword-based approach outlined in Table 1, MedISys [25] identified several Portuguese media articles on 5 September 2012, reporting that 'the mosquito *Aedes aegypti* struck again in force on Madeira' and 'left pharmacies without repellents and ointments' (peak A in Fig. 2) [3,26].

The data showed a sudden increase in dengue fever reporting in the Portuguese press, and MedISys issued an alert on Wednesday 3 October 2012 (peak B in Fig. 2). In more than 40 news articles, two confirmed and 22 suspected cases of dengue were reported. The story was run in newspapers in other European Union (EU) countries (Spain, Finland, etc.) on 4 October (peak C). On 5 October, 34 cases were reported as confirmed. The story was reported in the French and Belgian press on 10 October and in the UK press on 12 October, following a Reuters news wire story. An update from the Portuguese health authorities (Direcção-Geral da Saúde) was broadly discussed in the news on 8 November (peak D), and 517 confirmed cases were mentioned. The publication of the European Centre for Disease Prevention and Control (ECDC) Rapid Risk Assessment (RRA) update on 20 November met wide coverage, with over 80 articles being published within and outside the EU on 21 November (peak E).

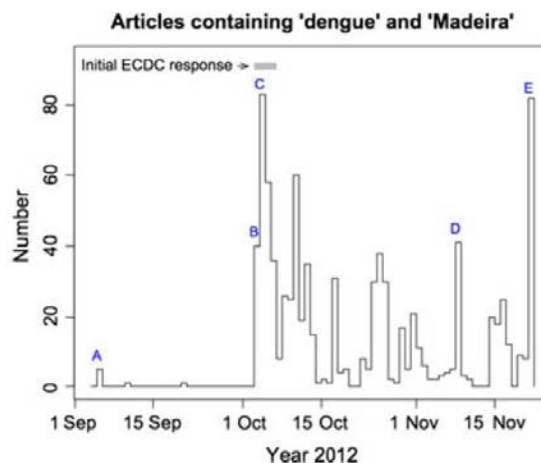


FIG. 2. Media reports on dengue fever on Madeira (number of articles per day, from 5 September to 21 November, 2012). The grey bar denotes the initial European Centre for Disease Prevention and Control (ECDC) response to the first alert, issued on 3 October 2012 (described in the text).

Internet biosurveillance played an important role in triggering an early public health response to this event (the grey bar in Fig. 2). On 3 October, the ECDC noticed a MedISys automated alert, and immediately began the process of verification by contacting the national health authorities of Portugal and gathering additional information from external experts in order to finalize an RRA for the EU population. Following this action, on 4 October, preliminary information about the outbreak was confidentially shared by the Portuguese health authorities with the EU/European Economic Area member states through the Early Warning and Reporting System (EWRS). The EWRS is the EU official communication restricted web-platform, and enables national authorities to exchange information on confirmed communicable disease events of potential international concern [27].

Early in the outbreak (near peak C in Fig. 2) on 6 October, the first ECDC RRA was internally finalized, and it was shared a few days later (10 October) with the EU/European Economic Area national health authorities through the EWRS. On 11 October, as agreed with the Portuguese authorities, the ECDC RRA was also made available online for the general public on the ECDC website [28]. In this outbreak, Internet biosurveillance played an important role in making international public health agencies aware of a potential outbreak earlier than would have been the case otherwise. This resulted in an early warning about the risk of infection in travellers returning from Madeira, where tourism is an important part of the economy. It also highlighted the risk of importation of

dengue virus to continental Europe via air and sea cargo at the onset of the outbreak [29].

Discussion

Outbreak data for human, animal and plant disease, available through informal media channels via the Internet, have been demonstrated to provide detection of anomalous disease events prior to official reporting [30–32]. In general, Internet media have the advantage of being timely, comprehensive, and available in any language from local and international sources. Such information can help to focus traditional surveillance efforts, and provides key data that can be used for a range of important public health purposes [33]. The value and pertinence of Internet biosurveillance have been demonstrated [34–36], and the approach has been integrated into the revised International Health Regulations [37]. Internet biosurveillance therefore contributes to early warning and situational awareness, and aims to trigger public health responses to mitigate outbreaks of infectious disease.

Biosurveillance as an input to EI

Internet biosurveillance has influenced the way in which EI is gathered. To meet its objective of early warning, EI typically combines one or more Internet biosurveillance systems that are complementary to one another, in order to gain a broad view of topics and regions of interest. EI is widely used by national and trans-national public health organizations (e.g. the US CDC, the ECDC, the Public Health Agency of Canada, the French Institute for Public Health Surveillance (InVS), and the WHO) to strengthen their early detection functions [38–40]. The scope of EI and its final objective are broad, and vary according to the mandate and objectives of the implementing institution. For example, EI can be adapted to specific goals, including the early detection of public health emergencies, of specific infectious diseases only [1], and of public health events during mass gatherings [41]. Nevertheless, core functions and EI can be defined as the process of early detection, collection, verification, analysis and organization of information in relation to public health events [42,43]. EI processes integrate both formal and informal sources of information (e.g. Internet biosurveillance and traditional public health surveillance).

From the end-user perspective, the first EI step is the detection of pertinent raw signals. Official sources of health information (e.g. ministries of health, and surveillance networks) are typically easily identified, and their content is meant to support public health analysis. However, access to these may be difficult and constrained (for example, the information may be available only in the national language, and access to the

information may be restricted), and their frequency of publication may not be appropriate for early disease detection. Therefore, informal sources (e.g. Internet media, discussion forums, and social networks) often represent the main source of signals. To collect and process large volumes of such material requires the use of Internet biosurveillance systems.

From the many raw signals observed from Internet biosurveillance systems, EI teams select information according to selection criteria defined by their public health institution. Following this, signals are verified; it is this verification phase that discriminates biosurveillance from EI. Verification consists of confirming and supplementing available information from additional and reliable sources, which are mainly networks of public health experts such as public health institutes, international institutions such as the WHO, World Organization for Animal Health, and ECDC, regional networks such as EpiSouth, laboratories, and non-governmental organizations.

Once verified, events are analysed to assess potential public health significance and potential national and/or international implications. Each is considered within its context and in the light of available scientific knowledge regarding spread, severity, and the efficacy of appropriate control measures [44]. Finally, following this analysis, the detected health threats are communicated to alert health authorities and to inform the public health community.

Needs for future research

Above, we have described the current state of the field of Internet biosurveillance, from data collection to data utilization for EI. Internet technology has significantly advanced the disease surveillance landscape; however, gaps in biosurveillance processes exist, and many challenges lie ahead in the field; some of those are described below.

Real-time signal detection. Sifting through the vast array of multimedia information on the Internet in real time is challenging. The noise of non-specific reports and misinformation complicates signal detection. Moreover, identifying anomalous activity without an established multi-year baseline of reporting for a given disease in a particular region is an obstacle. Anomaly detection is a capability in some biosurveillance systems at present, but there is a need for more robust anomaly detection approaches, including better entity extraction, visual analytical modalities, clustering methods, etc. [45]. Moreover, more work is needed on capturing and analysing the data from multilingual sources through linguistic algorithms or automated translation.

Data analysis. Internet biosurveillance data typically cannot be analysed with traditional epidemiological approaches, owing to

a lack of timely data verification and validation. For example, recognizing false-positive and false-negative events is problematic, owing to the lack of official comparison data or delays in diagnostic testing [33]. Frequencies of reports or events are often used for anomaly detection. However, identifying a common denominator (e.g. reports, events, articles, and sources) for analysis, and assigning a weight to sources based on accuracy, scope, and publication frequency, are not well established.

Collaboration, networking, and participatory epidemiology. Public self-reporting of events is increasingly recognized as benefiting disease detection. Extracting the data from participatory platforms (e.g. FluNearYou, Twitter, and Facebook) and utilizing it for early detection and surveillance is a critical area of current focus. For example, DIZIE, a project developed at the National Institute of Informatics in Tokyo, Japan, is used to visualize the extent to which Twitter data can detect/track infectious disease outbreaks [46]. More work is needed in this area, as health information sharing on social networking platforms has become prolific [47]. Users and public health experts can utilize this data in real time to track and assess disease situations [48].

Platforms with user-customizable features based on their specific needs and interests may make participatory modalities more attractive to a wider range of users. Also, more interactive functions for users (e.g. scoring option and comment field), may facilitate user interactions and information dissemination. An example of sharing and networking is the fully functional system for early alerting and reporting of potential chemical, biological, radiological, and nuclear events that has been developed by the Global Health Security Action Group through an extensive collaboration between the Joint research Centre of the European Commission and a team of risk assessment specialists from the G7+ Mexico countries [49].

Transparency Declaration

The authors declare that they have no conflicts of interest.

References

- Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. *N Engl J Med* 2009; 360: 2153–2155, 2157.
- Walters RA, Harlan PA, Nelson NP, Hartley DM. Data sources for biosurveillance. In: Voeller JG, ed. *Wiley handbook of science and technology for homeland security*, vol. 4. Hoboken: Wiley, 2010; 2431–2447.
- Hartley DM, Nelson N, Walters R et al. The landscape of international event-based biosurveillance. *Emerg Health Threats J* 2010; 3: e3. doi 10.3134/ehthj.10.003.
- Chunara R, Freifeld CC, Brownstein JS. New technologies for reporting real-time emergent infections. *Parasitology* 2012; 5: 1–9.
- European Centre for Prevention and Disease Control. *Framework for a strategy for infectious disease surveillance in Europe*, 2006. Available from: http://www.ecdc.europa.eu/en/activities/surveillance/documents/0806_framework_surveillance_strategy_in_europe.pdf (last accessed 17 January 2013).
- The White House. *Homeland security presidential directive 21 (HSPD-21)*. Public Health and Medical Preparedness. October 18, 2007. Available from: <http://www.fas.org/irp/offdocs/nspd/hspd-21.htm> (last accessed 8 February 2013).
- Collier N, Doan S, Kawazoe A et al. BioCaster: detecting public health rumors with a Web-based text mining system. *Bioinformatics* 2008; 24: 2940–2941.
- Mykhalovskiy E, Weir L. The Global Public Health Intelligence Network and early warning outbreak detection: a Canadian contribution to global public health. *Can J Public Health* 2006; 97: 42–44.
- Freifeld CC, Mandl KD, Reis BY, Brownstein JS. HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports. *J Am Med Inform Assoc* 2008; 15: 150–157.
- Cowen P, Garland T, Hugh-Jones ME et al. Evaluation of ProMED-mail as an electronic early warning system for emerging animal diseases: 1996 to 2004. *J Am Vet Med Assoc* 2006; 229: 1090–1099.
- Yangarber R, Jokipii L, Rauramo A, Huttunen S. Extracting information about outbreaks of infectious epidemics. In: Proceedings of the Human Language Technology Conference/Conference on Empirical Methods in Natural Language Processing: HLT/EMNLP-2005. Vancouver, Canada, 2005; pp. 22–23. Available at: <http://aclweb.org/anthology/H/H05/H05-2012.pdf> [accessed 19/06/2013].
- St Louis C, Zorlu G. Can Twitter predict disease outbreaks? *BMJ* 2012; 344: e2353.
- Morse SS. Public health surveillance and infectious disease detection. *Biosecur Bioterror* 2012; 10: 6–16.
- Collier N, Son NT, Nguyen NM. OMG U got flu? Analysis of shared health messages for bio-surveillance. *J Biomed Semantics* 2011; 2(suppl 5): S9. doi: 10.1186/2041-1480-2-S5-S9.
- Salathé M, Bengtsson L, Bodnar TJ et al. Digital epidemiology. *PLoS Comput Biol* 2012; 8: e1002616.
- Freifeld CC, Chunara R, Mekar SR et al. Participatory epidemiology: use of mobile phones for community-based health reporting. *PLoS Med* 2010; 7: e1000376.
- Mawudeku A, Lemay R, Werker D, Andraghetti R, St John R. The Global Public Health Intelligence Network. In: M'ikanatha N, Lynfield R, Van Beneden CA, de Valk H, eds. *Infectious disease surveillance*. Oxford: Blackwell Publishing, 2008; 304–317.
- Torii M, Yin L, Nguyena T et al. An exploratory study of a text classification framework for Internet-based surveillance of emerging epidemics. *Int J Med Inform* 2011; 80: 56–66.
- Mantero J, Belyaeva J. How to maximise event-based surveillance web systems: the example of ECDC/JRC collaboration to improve the performance of MedSys. JRC European Commission Publication Repository, 2011. Available from: <http://publications.jrc.ec.europa.eu/repository/bitstream/11111111/16206/1/lb-na-24763-en-c.pdf> (last accessed 27 March 2013).
- Keller M, Freifeld CC, Brownstein JS. Automated vocabulary discovery for geo-parsing online epidemic intelligence. *BMC Bioinformatics* 2009; 10: 385. doi:10.1186/1471-2105-10-385.
- Box GEP, Jenkins GM. *Time series analysis, forecasting and control*. San Francisco, CA: Holden-Day, 1970.

22. Wiener N. *Extrapolation, interpolation, and smoothing of stationary time series*. Cambridge, MA: MIT Press, 1964.
23. Gilbert P. *Dynamic system estimation (time series package)*. Available from: <http://cran.r-project.org/web/packages/dse/index.html> (last accessed 27 March 2013).
24. Sousa CA, Clairouin M, Seixas G *et al*. Ongoing outbreak of dengue type I in the Autonomous Region of Madeira, Portugal: preliminary report. *Euro Surveill* 2012; 17: pii=20333. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20333>.
25. Jens L, Belyaeva J, Steinberger R *et al*. MedISys: Medical Information System. In: Asimakopoulou E, Bessis N, eds. *Advanced ICTs for disaster management and threat detection: collaborative and distributed frameworks*. Hershey, PA: IGI Global, 2010; 131–142.
26. Jens L, Steinberger R, Weber T *et al*. Internet surveillance systems for early alerting of health threats. *Euro Surveill* 2009; 14: 1–2.
27. Guglielmetti P, Coulombier D, Thinus G, Van Loock F, Schreck S. The Early Warning and Response System for communicable diseases in the EU: an overview from 1999 to 2005. *Euro Surveill* 2006; 11: pii=666. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=666> (last accessed 27 March 2013).
28. ECDC Rapid Risk Assessment. Autochthonous dengue cases in Madeira, Portugal—10 October 2012. Available from: <http://ecdc.europa.eu/en/publications/Publications/Dengue-Madeira-Portugal-risk-assessment.pdf> (last accessed 15 March 2013).
29. ECDC. Mission Report 'Dengue outbreak in Madeira, Portugal: October–November 2012'. Available from: <http://ecdc.europa.eu/en/publications/Publications/dengue-outbreak-madeira-mission-report-nov-2012.pdf> (last accessed 6 May 2013).
30. Mondor L, Brownstein JS, Chan E *et al*. Timeliness of nongovernmental versus governmental global outbreak communications. *Emerg Infect Dis* 2012; 18: 1184–1187.
31. Chan EH, Brewer TF, Madoff LC *et al*. Global capacity for emerging infectious disease detection. *Proc Natl Acad Sci USA* 2010; 107: 21701–21706.
32. Woodall J. Official versus unofficial outbreak reporting through the Internet. *Int J Med Inform* 1997; 47: 31–34.
33. Nelson NP. Advantages and challenges of using Internet media for disease detection and tracking. *Phytopathology* 2012; 102: 161–162.
34. Heymann DL, Rodier GR. Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases. *Lancet Infect Dis* 2001; 1: 345–353.
35. Wilson K, Brownstein JS. Early detection of disease outbreaks using the Internet. *CMAJ* 2009; 180: 829–831.
36. Hoen AG, Keller M, Verma AD, Buckeridge DL, Brownstein JS. Electronic event-based surveillance for monitoring dengue, Latin America. *Emerg Infect Dis* 2012; 18: 1147–1150.
37. World Health Assembly. Revision of the International Health Regulations. World Health Assembly Resolution 58.3. 23 May 2005. Available from: http://who.int/csr/ihr/IHRWHA58_3-en.pdf (last accessed 27 March 2013).
38. Kaiser R, Coulombier D, Baldari M, Morgan D, Paquet C. What is epidemic intelligence, and how is it being improved in Europe? *Euro Surveill* 2006; 11: E060202.4. Available from: <http://www.eurosurveillance.org/ew/2006/060202.asp#4> (last accessed 27 March 2013).
39. Blench M. Global Public Health Intelligence Network (GPHIN). In: *Proceedings of the Eighth Conference of the Association for Machine Translation in the Americas*. Waikiki, HI: Elsevier, 2008; 1–5.
40. ECDC Epidemic Intelligence Group. *ECDC Epidemic Intelligence e-tutorial, a tool to learn more about the detection and assessment of public health threats*. 2011. Available from: http://external.ecdc.europa.eu/El_Tutorial/course.htm (last accessed 27 March 2013).
41. Khan K, Freifeld CC, Wang J *et al*. Preparing for infectious disease threats at mass gatherings: the case of the Vancouver 2010 Olympic Winter Games. *CMAJ* 2010; 182: 579–583.
42. Paquet C, Coulombier D, Kaiser R, Ciotti M. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. *Euro Surveill* 2006; 11: 212–214. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=665> (last accessed 27 March 2013).
43. Rotureau B, Barboza P, Tarantola A, Paquet C. International epidemic intelligence at the Institut de Veille Sanitaire, France. *Emerg Infect Dis* 2007; 13: 1590–1592.
44. World Health Organization. *Rapid risk assessment of acute public health events*. Available from: http://whqlibdoc.who.int/hq/2012/WHO_HSE_GAR_ARO_2012.1_eng.pdf (last accessed 27 March 2013).
45. Nelson NP, Brownstein JS, Hartley DM. Event-based biosurveillance of respiratory disease in Mexico, 2007–2009: connection to the 2009 influenza A(H1N1) pandemic? *Euro Surveill* 2010; 15: pii=19626. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19626> (last accessed 27 March 2013).
46. Collier N, Doan S. Syndromic classification of Twitter messages. In: Akan O, Bellavista P, Cao J, Dressler F, Ferrari D, Gerla M, Kobayashi H, Palazzo S, Sahni S, Shen X, Stan M, Xiaohua J, Zomaya A, Coulson G, eds. *Lecture Notes of the Institute for Computer Science, vol. 91, Social Informatics and Telecommunications Engineering*. Berlin: Springer, 2012; 186–195.
47. Fox S. *The Social Life of Health Information, 2011*. Report of the Pew Research Center, Washington, DC, 2011. Available from: <http://pewinternet.org/Reports/2011/Social-Life-of-Health-Info.aspx> (last accessed 13 March 2013).
48. Anonymous. *Bulletin: London 2012*. HPA, ECDC, WHO, 2012. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317135289768 (last accessed 13 March 2013).
49. Barboza P, Vaillant L, Mawudeku A *et al*. Evaluation of epidemic intelligence systems integrated in the Early Alerting and Reporting project for the detection of A/H5N1 influenza events. *PLoS ONE* 2013; 8: e57252.

ANNEXE 2 – ARTICLE B

A review of the evaluation processes and performance measures for epidemic intelligence systems mining of public health information from the Web

Soumis à : International Journal of Medical Informatics

Vaillant L.¹, Collier N.², Nelson N.P.³, Hartley D.M.⁴, Doherty B.⁵, Linge J.P.⁵, Schnitzler J.⁶, Madoff L.C.⁷, Yangarber R.⁸, Mawudeku A.⁹, Brownstein J.S.¹⁰, Barboza P.¹

On behalf of the Early Alerting and Reporting Project of the Global Health Security Initiative*

Affiliations:

¹ International Department, French Institute for Public Health Surveillance (InVS), Saint Maurice, France

² Department of information science, National Institute of Informatics, Tokyo, Japan

³ Department of Pediatrics, Georgetown University Medical Center, Washington, DC 20007, USA

⁴ Department of Microbiology and Immunology, Georgetown University Medical Center, Washington DC, 20007

⁵ Joint Research Centre (JRC) of the European Commission, Ispra, Italy

⁶ World Health Organization (WHO), Alert and Response Operations, Geneva, Switzerland

⁷ ProMED-mail, International Society of Infectious Diseases, USA

⁸ Department of Computer Science, University of Helsinki, Finland

⁹ Situational Awareness Section, Public Health Agency of Canada (PHAC), Ottawa, Canada

¹⁰ Children's Hospital Boston, Harvard Medical School, Harvard-MIT Division of Health Sciences and Technology

Abstract

Background: Information exchange providing timely and pertinent health related information has grown dramatically as the Internet has advanced. To optimise the use of this new vast source of information, tools mining public health information have been developed to assist epidemic intelligence.

Objective: This study aimed to identify and review literature relating to evaluations applied to epidemic intelligence tools using Internet information. Epidemic intelligence measures using standard statistical methods and qualitative attributes have been exhaustively identified and results of evaluation have been summarised.

Methods: Evidence was collected from peer-reviewed literature and other sources (e.g., technical reports, abstracts, conference proceedings) meeting specific criteria: event-based system, epidemic intelligence, assessment process.

Results: A total of 31 scientific documents relating to the evaluation process of epidemic intelligence tools was identified: 10 surveys dedicated to the users' perspective (extrinsic evaluation), 17 surveys related to systems' performance (intrinsic evaluation) and 4 studies combined both evaluations. Intrinsic performances (text classification and geo-encoding capability) were measured with well-known parameters: precision, recall (or sensitivity), accuracy, F-score, specificity. Extrinsic evaluations were mainly focused on timeliness, sensitivity, and accuracy.

Conclusions: Both extrinsic and intrinsic evaluations are essential for the development of technology and maintaining fitness between end users requirements and systems' output. Although Internet-based systems are becoming critical tools for epidemic intelligence, important challenges still remain in terms of evaluation, which needs to keep up with Internet advances. Therefore, system providers conducted regular evaluations to provide measures reflecting the performance of their systems; they also expressed the need of end user involvement for a complete evaluation. Several challenges remain in regard to this review to enhance evaluation processes: establishment of gold standard, utilisation of standardised metrics, combining of intrinsic and extrinsic evaluations, utilising a system-mixed evaluation approach, and emphasising the utility for public health.

Keywords: epidemic intelligence, evaluation methodology, data mining, Internet

Introduction

The Internet has had a substantial impact effect on early detection of global public health events in recent years by accelerating information dissemination. Frequently events are posted on the Internet before official publications are published [1]. Public health institutions and organizations hence monitor news média and specialist sites on the Internet to detect outbreaks and public health events [2-4]. Nonetheless, even if countries and international organizations tend to publish more outbreaks and surveillance reports on the Internet, there are still gaps and delays due to many reasons among which political issues or bureaucratic clearance. The revised International Health Regulations (IHR) (2005) aim to improve the capacity to detect and respond to public health events of concern in a timely manner and support the possibility to take into account reports from sources other than notifications or consultations by state parities. Information retrieved from the World Wide Web has become an important source for the detection of public health events [5]. Epidemic intelligence (EI) [3-4] can be defined as the operation of early detection, collection, verification, and organization of often-informal information (e.g. news media) in relation with health events, and provides a new approach allowing national and supranational institutions to have a more complete surveillance approach. While the Internet facilitates accessibility to information, the large amount of false information or rumours, which can be detrimental to the detection of relevant health related information, remains a challenge.

To meet these challenges, medical information mining systems have been developed to address the need of retrieval and classification of large volumes of information from the Internet. The Global Public Health Intelligence Network (GPHIN) [6] was developed in 1996 by the Public Health Agency of Canada and since its inception has been one of the precursor and most robust systems used by public health experts. While GPHIN is a human moderated system, fully automated tools have been successively developed [7-9]. The tools are designed to cope with the growing amount of information and diversity of sources available on the Internet including news média, social networks, blogs and others. Internet-based systems are continuously evolving; their performances are improving and complement each other.

The basis for automated systems [10] is the monitoring and automatic classification of a large number of news

articles in real-time. In practice, many articles can mention diseases or public health issues without being related specifically to acute public health risks (e.g. general information on health topics). Therefore, classifier requirements should integrate the capability to distinguish general diseases or health oriented news from infectious disease outbreak reports and ongoing health events and allow to deal with redundancy and a high volume. Systems are also expected to be as sensitive and specific in a timely manner.

While traditional surveillance systems are commonly assessed using defined indicators and standardised processes [11;12], an assessment process for methodologies utilised by event-based systems used for EI (i.e. early detection of outbreaks) has not yet been well developed. A key question is whether extrinsic evaluation or intrinsic evaluation or both is the appropriate evaluation approach. To fulfil the need to quantitatively characterize EI systems, we set out to investigate what has been done to date using key EI measures that can be assessed using standard statistical methods.

Methods

Epidemic intelligence systems

EI systems have already been presented in the literature [10; 13] and their processes have been well described. Some systems include human intervention (Argus, GPHIN, ProMED-mail) [14-16], while others are fully automated such as BioCaster Global Health Monitor [17] and MedISys (Medical Information System) [10], or partially automated as HealthMap [18]. Puls (Pattern-based Understanding and Learning System) [9] is also automated but differs from other by its ability to cluster and analyse the information. Other Internet event-based systems like EpiSPIDER (Semantic Processing and Integration of Distributed Electronic Resources for Epidemics [and disasters]) [19] or the Hungarian National Association of RSOE-EDIS (Radio Distress-Signalling and Info communications, Emergency and Disaster Information Service, available from <http://hisz.rsoe.hu/alertmap/index2.php#>) also monitor and document public health events (including environmental events which may cause a disaster or emergency). They differ in terms of the language and range of information sources, their fields of focus (e.g. human, animal, plant, environmental, infectious disease, etc.), their geographical coverage and their additional functionalities (e.g. mapping, statistics, alerting mechanisms).

Table I: Epidemic intelligence systems in the literature.

Systems name	System owner / developer	Country	Moderation Type	Date of creation
Argus [16]	Georgetown University	USA	Human moderated	2004
BioCaster [17]	National Institute of Informatics	Japan	Fully automated	2006
GPHIN [15]	Public Health Agency of Canada	Canada	Human moderated	1997
HealthMap [7]	Harvard University	USA	Partially moderated	2006
MedISys [12]	Joint Research Centre	EU	Fully automated	2004
ProMED-mail [14]	International Society of Infectious Diseases	USA	Human moderated	1994
Puls [9]	University of Helsinki	Finland	Fully automated	2007

Data Sources

The objective was to review research conducted on evaluation of EI systems using open source and social média. A systematic review of the literature was conducted utilizing the bibliographic database PubMed enriched by Google bibliographic search and grey literature (Figure 1). Keywords and MeSH terms used to search PubMed were divided into two categories:

(1) Free text searches for systems mining information from the Internet for early detection of public health events: Argus, BioCaster, EpiSPIDER, GPHIN, HealthMap, MedISys, ProMED-mail, and Puls.

(2) MeSH (Medical Subject Headings) Terms and key words: Evaluation Methodology (“evaluation studies” as topic [MeSH Terms]), epidemic intelligence (“epidemics”[MeSH Terms] AND (“intelligence”[MeSH Terms]), systems (“event-based bio/system”, “Internet-based system”), Internet (“Internet/instrumentation”[Mesh] OR “Internet/utilization”[Mesh]).

Unfortunately, terms such as epidemic intelligence have not yet been coded in these databases. The search covered peer-reviewed articles, conference proceedings and reports published until September 2011.

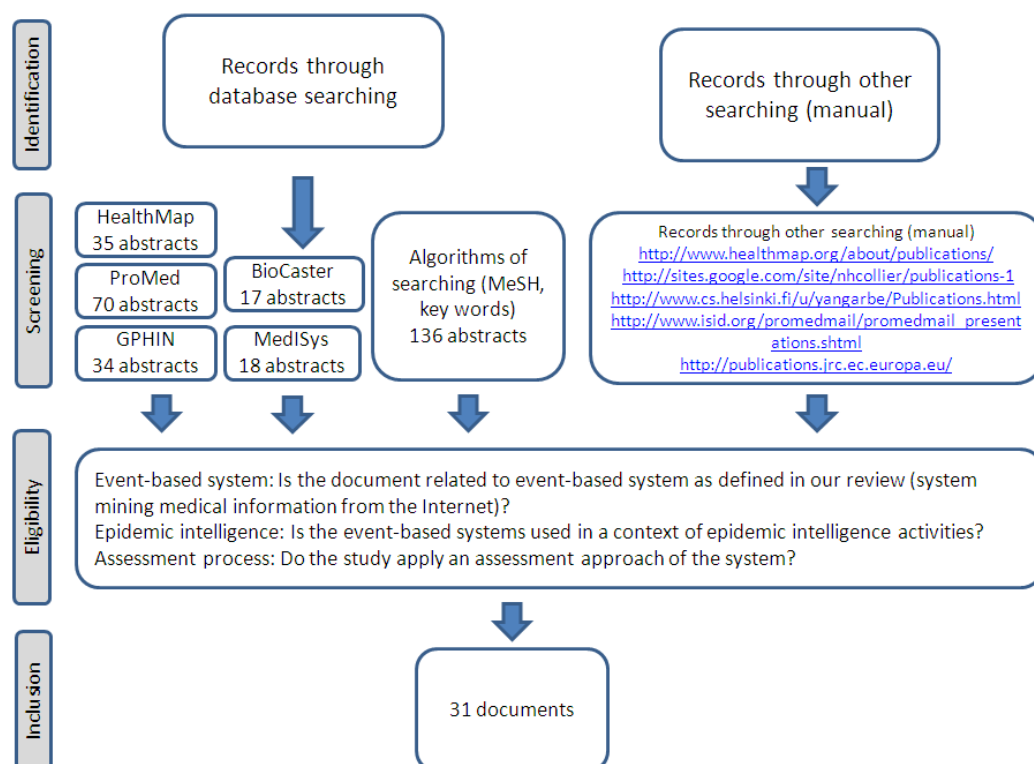
Data Extraction and Synthesis

Using the search terms described above, a total of 275 items was retrieved (Figure 1). Publications that related to survey data, physician records, medical records and search engines were excluded. Thirty-one publications (including conference proceedings and posters) focusing on evaluations conducted on systems mining outbreak and public health information data were reviewed and classified according to their perspective of evaluation (intrinsic/system or extrinsic/user).

Analysis process

All publications focusing on mining Internet information for EI and public health event early detection purposes have been included. Descriptive analysis was conducted throughout 2 perspectives [20]: primarily focused on surveys related to the evaluation from a systems’ perspective (intrinsic evaluation) and secondarily focused on the evaluation from the end-users perspective (extrinsic evaluation). These perspectives are frequently used for evaluation and represent an easy and well-understanding way of classification.

Figure 1: Flow diagram of the study selection process.



Results

General results

Among the 31 selected documents, ten surveys were dedicated to the users’ perspective (extrinsic evaluation), 17 surveys to systems’ performance (intrinsic evaluation) and 4 studies combined both evaluations (See Table II). Most of the evaluation studies analysed one system.

Systems

Text classification.

To classify texts in the context of unstructured data for early epidemic alerting, systems need to capture semantic characteristics of the text to perform the most accurate classification and limit false positive results. Precise definitions have been developed in studies

dedicated to evaluation of text classifiers [38]. Different text classifiers were classically used and evaluated, including two machine learning algorithms that are often used by surveillance systems [37]: Naive Bayes (NB) and Support Vector Machines (SVM).

Methods (e.g. combinations of algorithms and specific natural language processing (NLP) techniques) have been implemented to improve systems performance and several evaluations were performed using well-estimated parameters described in previous studies: precision, recall (or sensitivity), accuracy, F-score, specificity (See Table III). Several machine learning algorithms have been commonly assessed and compared using this methodology. Annotation scheme and corpora that can be used for the evaluation of outbreak or health event extraction algorithms have been developed [43].

Researchers working on the BioCaster project developed evaluations based on the F-score, accuracy and precision values to improve the performance of the system. NB and SVM methods have been successively evaluated and improved by adding specific algorithms and features [31;34; 37-38]. The accuracy measures varied from 85% to 95% (See table IV). Zhang and Liu [35] explored the performance of classifiers for sentences containing disease outbreak information in ProMED-mail. NB and SVM classifiers (using extraction of sentences, word n-grams and other features) were evaluated and were found to be competitive on ProMED-mail sentences. The performance of both types of classifiers was improved by combining features: the F-score was improved from 62% to 76% and the highest scores were related to SVM with n-grams > 1.

Scientists from the University of Helsinki [44] developed a new approach for the evaluation of 3 different types of text classifiers (NB, SVM and SVM with the Radial Basis Function (RBF) kernel) based on the relevance prediction determined by users and developers (non-users) of Puls. They concluded that SVM performance (79%) was higher than the NB classifier (76%) at the event level, with specific lexical combination.

In 2007, Freifeld et al. [36] analysed HealthMap classification accuracy, which was 84% for overall system (655 on 778 reports over a one-month period). As their information came mainly from ProMED-mail and Google News, they also determined the specific accuracy for both sources: performance on ProMED-mail alerts was substantially better than on Google News reports (91% vs. 81%).

In the framework of the Argus project applied to disease surveillance, Torii et al. [45] assessed machine learning classifiers trained on relevant and randomly sampled unlabelled articles to study the use of classifiers in a real-time, human based surveillance system. Daily averages of areas under Receiver Operating Characteristic (ROC) curves analysed over the 15-day evaluation period were

84% and 87% respectively for the NB and SVM classifier. The study concluded that the performance observed for the two algorithms, NB and SVM, was competitive, and together yield a robust classifier that can be applied to a range of relevant data.

Geo-encoding capability.

Geocoding is one aspect of natural language processing and text classification used for the recognition of geographical information. It can also be applied to additional applications, such as mapping events or showing the evolution of geographical coverage over time. Same parametric indicators (Accuracy, Precision, F-score) have been used to evaluate different levels of geographical information and their interactions.

In 2004, researchers from JRC (Joint Research Centre), working on the EMM (European Média Monitor) technology (basis of MedISys) [20] aimed to assess the systems' ability to recognize places and countries using expert judgement. Geographical coding correctness was measured at an average of 98% precision and 88% recall. Nevertheless, the results of the same evaluation using the same tools can be lower when using a test with more ambiguous geographical names (based on newspaper stories out of the EMM): 77% for both precision and recall [33].

BioCaster [39; 40] also focused some of their evaluations on geographical and temporal classification capabilities and developed a more accurate classification with a high level of granularity (86% for precision and 85% for recall). Spatial attribute recognition performance is very valuable for systems and users since event mapping is often an important feature of event-based systems. BioCaster's mapping capability was evaluated in 2008 [8] and the percentage of correct pairing of localisation had a precision of 93%. Other studies [40] have analysed textual spatial information from different perspectives that will not be covered in this report. For example, Chaudet et al. [32] proposed a formal representation, which combined spatial and temporal attributes, evaluating ProMED-mail's data.

Other topics.

Finally, another aspect explored within the framework of intrinsic evaluation aimed to examine the capability of anomaly detection in Internet information. Collier et al. [41] evaluated and compared 5 different aberration detection algorithms (EARS C2, C3, W2, F-statistic and EWMA) toward implementing a threshold based on the input from the systems for outbreak detection. Results emphasized that the W2 method had the best balance of F-score, timeliness and low number of false alarms. To improve the MedISys system, studies [30] have been performed with a "human filter" to increase sensitivity, specificity, and PPV and decrease the false positive rate. This evaluation was based on users' analysis to determine the best model of performance.

Table II: Manuscripts and reports related to the evaluation process of EI tools.

1 st author Pub.	Year of	Systems evaluated	Reference (GS)	Evaluation perspective	Indicators
Heymann	2001 [2]	GPHIN	WHO	U	Utility / Usefulness
Woodall	2001 [21]	ProMED-mail	WHO	U	Timeliness, Accuracy, Comprehensiveness, Friendliness
Cowen	2006 [22]	ProMED-mail	OIE	U	Timeliness, Accuracy (correction rate)
Brownstein	2007 [18]	HealthMap	WHO	U	Timeliness, Sensitivity
Zeldenrust	2008 [23]	ProMED-mail	NEWC	U	Utility / Usefulness
Keller	2009 [19]	GPHIN, EpiSPIDER, HealthMap	NA	U	Completeness, Performance
Rortais	2010 [24]	MedISys, ProMED-mail	RASFF	U	Efficiency (time delays)
Mondor	2011 [25]	HealthMap	WHO	U	Timeliness
Lyon	2011 [26]	BioCaster, EpiSPIDER, HealthMap	NA	U	Timeliness information characterisation (overlap, sources, language, geographical coverage)
Thomas	2011 [27]	Argus	official reports & peer-reviewed articles	U	Accuracy
Grishman	2002 [28]	Proteus-BIO	Web search engine	S, U	Accuracy (Recall + Precision), Effectiveness (relevance)
Steinberger	2008 [9]	Puls, MedISys	NA	S, U	Accuracy, Precision
Chan	2010 [29]	GPHIN, ProMED-mail, HealthMap	WHO	S, U	Timeliness
Mantero	2011 [30]	MedISys	NA	S, U	Sensitivity, specificity, PPV, False Positive rate
Pouliquen	2004 [20]	EMM	NA	S	Precision, Recall
Kawazoe	2006 [31]	BioCaster	NA	S	F-score
Chaudet	2006 [32]	ProMED-mail	ProMED-mail	S	Accuracy, Concordance
Pouliquen	2006 [33]	EMM	NA	S	Precision, Recall
Doan	2007 [34]	BioCaster	NA	S	Precision, Recall, Accuracy, and F-score
Zhang	2007 [35]	ProMED-mail	NA	S	F-score
Doan	2008 [8]	BioCaster	NA	S	Precision
Freifeld	2008 [36]	HealthMap, ProMED-mail	NA	S	Accuracy
Conway	2009 [37]	BioCaster	NA	S	Accuracy, Precision, Recall, Specificity, F-score
Doan	2009 [38]	BioCaster	NA	S	Precision, Recall, Accuracy, and F-score
Chanlekha	2010 [39]	BioCaster	WHO	S	Precision, Recall and F-score
Chanlekha	2010 [40]	GPHIN, Argus, HealthMap, BioCaster	NA	S	Reliability (geo) = mapping capability
Collier	2011 [41]	BioCaster	ProMED-mail	S	Sensitivity, specificity, PPV, NPV, F1-score
Collier	2010 [42]	BioCaster	ProMED-mail	S	Sensitivity, F1-score, timeliness
Conway	2010 [43]	BioCaster	ProMED-mail	S	Accuracy
Von Etter	2010 [44]	Puls	NA	S	Precision, Recall, Accuracy, and F-score
Torii	2011 [45]	Argus	NA	S	Accuracy, sensitivity, relevancy

GS: Gold standard, FP: False Positive, NEWC: Netherlands' Early Warning Committee, NPV: Negative Predictive Value, OIE: World Organisation of Animal Health, PPV: Predictive Positive Value, RASFF: Rapid Alert System for Food and Feed, S: Systems, U: Users, WHO: World Health Organisation

Table III: Performance measures used for text classifier.

Precision OR PPV (Predictive Positive value)	$P = a / (a+b)$
Recall (or Sensitivity)	$Rc = a / (a+c)$
Specificity	$Sp = b / (b+d)$
NPV (Negative Predictive Value)	$NPV = c / (c+d)$
Accuracy	$Ac = (a+d)/(a+b+c+d)$
F-score	$F = (2.P.Rc) / (P + Rc)$

a: True positive (event correctly classified as type X), b: False positive event incorrectly classified as type X), c: False negative (event incorrectly classified as not of type X) d: True negative (event correctly classified as not of type X)

Table IV: Main measures of metrics applied to evaluation of EI systems

Date, author	System	Measures	Description	Ref.
2011, Torii et al.	Argus	AUC = 87%	SVM classifier	[45]
2006, Kawazoe et al.	BioCaster	F-score = 80%	SVM classifier	[31]
2007, Doan et al.	BioCaster	Accuracy = 84%, Precision = 75%	NB classifier	[34]
2008, Doan et al.	BioCaster	Precision = 93%	mapping capability	[8]
2009, Doan et al.	BioCaster	Accuracy = 91%	SVM classifier	[38]
2009, Conway et al.	BioCaster	Accuracy = 95%, F-score = 93%, Recall = 97%, Precision = 89%	NB classifier	[37]
2010, Chanlekha et al.	BioCaster	Precision = 86%, Recall = 85%	geographical and temporal classifier	[39]
2010, Collier et al.	BioCaster	Timeliness = 3-6 days	detection before ProMED-mail	[41]
2004, Pouliquen et al.	EMM	Precision = 98%, Recall = 88%	Geographical coding correctness	[20]
2010, Chan et al.	GPHIN	Timeliness = 23 days	Detection before WHO notification	[29]
2007, Brownstein et al.	HealthMap	Timeliness = 12 days	Detection before WHO notification	[18]
2007, Brownstein et al.	HealthMap	Sensitivity = 58%		[18]
2010, Chan et al.	HealthMap	Timeliness = 32 days	Detection before WHO notification	[29]
2001, Woodall et al.	ProMED	Accuracy = 2.6%	Error rate of reporting	[21]
2006, Cowen et al.	ProMED	Accuracy = 1 to 2%	Error rate of reporting	[22]
2006, Cowen et al.	ProMED	2 days and 2 weeks	Detection before OIE notification	[22]
2006, Cowen et al.	ProMED	Sensitivity = 47%		[22]
2007, Zhang and Liu	ProMED	F-score = 76%	SVM classifier	[35]
2010, Chan et al.	ProMED	Timeliness = 48 days	Detection before WHO notification	[29]
2005, Grisham et al.	Proteus BIO	Precision = 79%, Recall = 41%	classification accuracy	[28]
2010, Von Etter et al.	Puls	Accuracy = 79%	SVM classifier	[44]
2007, Freifeld et al.	HealthMap	Accuracy = 84%	Classifier	[36]

Users

Accuracy.

Pouliquen et al. [20] mentioned the advantage of involving a willing user group. This approach has been used in an evaluation study [9] involving MedISys and Puls to measure accuracy. Accuracy is an indicator often used for intrinsic evaluation as well as extrinsic evaluation. While it is well defined for intrinsic evaluation (see systems), it is more variable for extrinsic evaluation and dependant on the system assessed. Accuracy of ProMED-mail according to Woodall et al. [21] and Cowen et al. [22] corresponds to the error rate of reporting provided by media and contributors, excluding official reports that could have been denied later, and were respectively measured with 2.6% (September 1999 – March 2000) and from 1 to 2% for animal diseases (1996-2004). In 2002, Grisham et al. [28] worked on a system mining medical and epidemic information, Proteus BIO. They used ProMED-mail and WHO as data and evaluated the classification accuracy (Precision: 79%, Recall: 41%) and the search effectiveness (compared to Web search engine). They also used an innovative approach involving users, who were asked to identify documents relevant to 2 different outbreaks to evaluate the effectiveness of the system. Thus, they showed that the extraction database was a valuable tool for users, because it is more relevant than other standard Web search approaches (like Google): using the database (based on 2 test subjects: dengue and Rift Valley fever), the median number of relevant articles was 18.5 (range: 10 to 29); using standard Web search, the median was 12.5 (range: 4 to 34). Thomas

et al. [27] proposed an evaluation approach using a case study of a plant pest and pathogen epidemic in China and Vietnam from February 2006 to December 2007. Events detected by the Argus system during this study period were compared to Official government or international (NGO) reports and peer-reviewed scientific journal articles to verify information accuracy.

Timeliness.

Timeliness and sensitivity are critical attributes of systems focused on early detection. WHO (4 surveys) and OIE (1 survey) events were compared to ProMED-mail and HealthMap. Brownstein et al. [18] found that the HealthMap system successfully detected events with Google News approximately 12 days before WHO publication with a moderate sensitivity (58%). Cowen et al. [22] also conducted an evaluation of ProMED-mail compared to OIE. A total of 35 events (47%) were detected by ProMED-mail between 2 days and 2 weeks earlier than OIE. In 2001, Woodall [21] showed the benefit in terms of timeliness of using unofficial Internet sources to disseminate information rather than wait for official posts. In another report [46], the delay observed between ProMed mail and reports by WHO in 1999 has been analysed and ranged from 2 days to 8 weeks according to the outbreak. Part of the difference was explained by diverse reporting delay at that time. Chan et al. [29] analysed the WHO public record of Disease Outbreak News reports from 1996 to 2009 to characterize spatial-temporal trends (GPHIN, HealthMap and ProMED-mail) in the timeliness of outbreak discovery and public communication about the

outbreak relative to the estimated outbreak start date. Cox proportional hazards regression analyses showed that the timeliness of public communication improved by 6.2% per year. Stratified by WHO region, the Eastern Mediterranean (Hazard Ratio (HR) = 1.201 per year) and Western Pacific regions (HR = 1.119 per year) showed improvement in public communication. These findings provided a quantitative historical assessment of timeliness in infectious disease detection and public reporting of outbreaks.

Mondor et al. [25] analysed the timeliness of HealthMap reporting from 1996-2009 by governmental and non-governmental sources, geographical region and the time of year. No statistically significant differences (RR=0.93) were found in HealthMap' reporting timeliness. Collier et al. [41] calculated timeliness for monolingual news (information coming from a monolingual newswire like English newswire) and multilingual news compared to the ProMED-mail silver standard. Silver standard is a proxy of gold standard providing the best estimation of official data based on publically available information focused on specific subject matter but lacking of exhaustivity (impossible to verify). The mean number of days that BioCaster's alerts were disseminated before ProMED-mail reports varied from 3 days to 6 days.

A recent study [26] compared the timeliness of reports for the period 2–30 August 2010 from BioCaster, EpiSPIDER and HealthMap between each other but found no significant difference after taking into account the different time zones. EFSA initiated the development and assessment of automated monitoring tools based on MedISys to collect food- and feedborne hazards. One of the evaluations [24] was based on a retrospective case-study approach, which consisted in comparing time delays for notifications selected from the RASFF reporting system (Rapid Alert System for Food and Feed), from January 2007 to March 2009, with MedISys and ProMED-mail. The results of this study showed that MedISys performed well as an early-warning system by detecting hazards earlier (i.e. from a few days or a few weeks to a few months earlier than RASFF).

Sensitivity, utility.

The utility of the EI systems can be defined by their contribution to the early detection of events of public health significance [26]. An assessment of utility can go beyond detection and analyse the impact or value added by its application. Hence evaluation should begin with a review of the objectives of the system and should consider the priorities. Actually, utility of event-based tools is often approached by timeliness (paragraph above) and sensitivity defined by the number of events detected before official notification. A GPHIN study [15] compared GPHIN reporting to WHO official reports and showed that from July 1998 - August 2001, 56% of events verified by WHO were initially picked up by GPHIN. In a HealthMap [18] study in comparison to WHO, they defined sensitivity as the proportion of WHO alerts detected by news média and calculated a value close to GPHIN's: 58% of the WHO alerts were reported in the news (WHO, 2007). In a different perspective, Zeldenrust and al. [23] assessed events identified as possible threat for Netherlands in

NEWC (May 2006 - June 2007) and originating from ProMED-mail. They found that five (20%) of the events were considered as possible threats for the Netherlands.

Other indicators.

Comprehensiveness and friendliness [21] are qualitative characteristics that illustrate the perception of users. These are important criteria to measure how helpful and efficient a system is for end users. Pouliquen et al. [20] evocated the value and challenges of evaluating the positive impact of maps and animation on the reactivity of experts. Keller et al. [19] proposed to evaluate several different EI tools (GPHIN, HealthMap and EpiSPIDER project) and described the processes of data acquisition and data dissemination. They presented clearly to the public health community the processes involved in these systems, by listing and comparing the characteristics of each systems. Moreover they highlight the systems similarities and complementarities.

Most studies presented in this review are focused on the intrinsic evaluation of a unique system. Public health experts' needs are broader and encompass public health utility such as early warning, preparedness or response. Even if some studies tend to assess systems' utility, there are still gaps for end-users who are not familiar with all these tools.

Lack of a "gold standard"

A gold standard would be an official (e.g. WHO, FAO, OIE, Ministries of Health) and complete and real time list of all public health events of interest. Such a data set could be used as reference for indicators such as sensitivity and timeliness. These data are also indispensable for measuring and evaluating quantitative EI indicators. For instance, sensitivity is a well-known indicator used for traditional epidemiological surveillance. Evaluation of EI is more complicated compared to the evaluation of traditional indicator based surveillance systems:

- First, EI is relatively new in the public health domain and there is no established standard evaluation method.
- Second, there is no gold standard available allowing for reliable benchmarking.
- Third, EI involves detecting public health events using structured and more often unstructured data which is more difficult to analyse in a standardized way.

Fourth, EI often tries to identify the unexpected and unknown events, which definitions depend on the mandate of the organization (e.g. national / international). Evaluation of utility in terms of Public Health requires also standards specific to public health concerns and depends on institution and organisation mandates.

Outbreaks reports are increasingly posted on the Internet by multiple official sources, hence providing some information that can be used for the evaluation: World Health Organization (reports, Weekly Epidemiological Record [47]), OIE (Disease Information weekly, periodic disease Alert Message reports) [48], CDC (Morbidity and Mortality Weekly Report) [49], HPA (Health Protection Report) [50], ECDC (Communicable Disease Threat Report) [51], InVS (Weekly epidemiological bulletin) [52], EpiSouth (eWeb,

[53]). However, none of these sources include all of the elements necessary to construct a fully exhaustive “gold standard” required to assess global epidemic activities that need to be detected by the systems.

Discussion

Official sources are mainly used as gold standard and WHO data were used several times in this regard [2; 18; 21; 29; 39]. However, ProMED-mail has also been used as a “silver standard” benchmark [32;41-43], which according to Collier has several advantages including ProMED-mails’ availability, high level of global coverage on most global public health concerns, broad community, and timeliness of reporting. Moreover, exhaustive “gold standard” will not be available in the near future, although more official and timely information is nowadays available. Hence, we could discuss a “silver standard” definition including EI information.

Intrinsic evaluation versus extrinsic evaluation

Several intrinsic evaluations have been conducted to improve the performance of the EI systems in terms of detection and classification of information. A correct classification enhances traceability and early detection of the information. Studies on text classifiers and improvement of the NLP capability to clearly identify the pathogen, filter irrelevant reports and duplicates and to enhance the spatial resolution of location have been performed. The characterisation of the information is continuously being improved by enhancing work on the NLP capability but also by introducing other functionalities that must be assessed (usefulness of statistics and other metrics to define the detection threshold. Numerous algorithms were processed and evaluated by standard and robust metrics known and adapted to the NLP field, though some difficulty persists to classify ambiguous biomedical names by reporting F-scores [31]. Another main difficulty is the integration of different languages while performing the evaluation [20]. For instance, it is well known that Japanese is a complex language to integrate, due to an intermingling of multiple alphabets. Therefore, these evaluations provide valuable insights for system developers; but the best approach to evaluate the relevancy of information or the public health utility remains unclear for public health experts and decision makers.

An extrinsic evaluation was performed in the Netherlands whereby the utility of subject matter expert reports from ProMED-mail (informal data) for early warning was evaluated by end users of the Netherlands’ Early Warning Committee [23] through comparison with public health data. The study was specific to the Netherlands and a broader assessment of the utility of integrating end user objectives in system evaluation is needed. Using unstructured data as sources of surveillance information is not efficient where a robust epidemiological surveillance system is already implemented. Even if extrinsic evaluation helps to inform public health experts of the value of EI systems, they will need to use it on a daily basis to realize the early warning potential. Both extrinsic and intrinsic evaluations have limitations. Gaps in geography and

languages covered by systems require assessment, better understanding and likely system modifications. To fill these gaps and to improve, systems may introduce new sources such as blogs, discussion groups and aggregators. The usefulness of these new sources remains poorly understood and requires further evaluation. Lack of a gold standard of information also inhibits the ability of systems and users to conduct robust assessments.

In summary, both extrinsic and intrinsic evaluations are essential for the development of technology and maintaining fitness between end users requirement and systems’ output. These evaluations complement each other and lead to an efficient improvement of the systems in a public health utility perspective. Developers enhance engine performances of systems and users provide an iterative process of feedback to the developers to adapt systems to the public health needs and improve global usefulness. The combination of an extrinsic evaluation and an intrinsic evaluation focused on technical aspects of the systems is the ideal approach for epidemic intelligence system evaluation because it has the best chance of capturing all of the benefits and weaknesses of the different systems whereby user feedback informs system development. Therefore such evaluation could be useful for enhancing systems and to help public health experts in choosing the most appropriate system for a specific problem.

Conclusion

This study aims to provide an overview of systems, and evaluation tools and methods associated, used to identify public health risks from informal sources. It does not provide a critical review of the individual studies nor the evaluation of the public health impact of the study results. A significant delay is usually apparent when the detection dates of an outbreak (human or animal) by EI-systems are compared to ProMED-mail, OIE or WHO’s publication (i.e. official reports), confirming the utility of using unofficial sources for earlier detection. More generally, information from official institutions is often provided after the information is available through informal sources [25]. Nevertheless this does not allow to draw any conclusions on the impact of public health action as public health measures are often initiated before information is officially communicated. While their performance in terms of timeliness and sensitivity has been demonstrated, their utilization by public health experts and their effect on public health measures remain under-documented and require careful evaluation. Some analyses have shown that the detection time was dependent on geographical coverage but also that it has improved between 1996 and 2009 [29]. Intrinsic systems’ performances (based on text classifiers, natural language processing and algorithms) are measured with defined indicators (accuracy, precision, recall, F-score). These parameters allow researchers to improve disease detection by performing correct text categorisation of information retrieved from the Internet. Even if the same indicators were used across surveys, the results would not be comparable due to different data structure and

methodologies.

The studies reviewed have yielded informative methodologies and concepts but they remain heterogeneous and difficult to compare. The research and investment made on media sources and the Internet to improve early detection of public health events highlights an important shift in disease surveillance, evolving far beyond traditional public health systems. Internet-based systems are more widely known and are now essential tools. Nevertheless, they ultimately must remain adapted in a context of wide Internet development and, important challenges still remain for their evaluation. To address these challenges, system providers often conduct regular evaluations to provide measures reflecting the performances of their system; they also expressed the need of end-users involvement for a

complete evaluation.

On-going evaluation of systems remains essential to improve systems performances and to support public health expertise and decision-making. Each stakeholder level requires specific, representative and clear indicators, to fulfil their goals, i.e. the early health threat detection in order to anticipate and respond to public health events in a timely manner. Naturally, synergy between end-users and systems' providers must be emphasised at several stages of the surveillance process, including fruitful interaction to improve individual systems, to maintain a common understanding, to optimise the utilisation. Collaboration between and among end-users and systems' providers is key to the success and optimal development of an effective global EI system.

Authors' contribution

LV drafted this review manuscript and the other authors completed literature and critically revised the manuscript.

Acknowledgements

We gratefully acknowledge the entire EAR working group for their contribution and assistance for this manuscript. We sincerely thank GHSI which supports the project and allows countries and institutions to provide human and financial supports for this project.

***EAR project core members:** Ray R. Arthur (CDC, US), Philippe Barboza (InVS, France), Mike Barker (HPA, UK), John Brownstein (Children's Hospital Boston, US), Catherine Chow (CDC, US), Nigel Collier (NII, Japan), Ricardo Cortes (Ministry of Health of Mexico), Silvia Declich (ISS, Italy), Maria Grazia Dente (ISS, Italy), Pamela S. Diaz (CDC, US), Brian Doherty (JRC, EC), Christian Herzog (RKI, Germany), Mike Hiley (HPA, UK), Nigel Lightfoot (HPA, UK), Jens Linge (JRC, EC), Lawrence C. Madoff (ProMED-mail, US), Jas Mantero (ECDC, EC), Abla Mawudeku (PHAC, Canada), Noele P. Nelson (Georgetown University, US), Flavia Riccardo (ISS, Italy), Agnes Rortais (EFSA, EC), Mika Shigematsu (NIID, Japan), Johannes Schnitzler (WHO), Germain Thinus (DG-SANCO, EC), Laetitia Vaillant (InVS, France), Roman Yangarber (University of Helsinki, Finland).

Statement on conflicts of interest

No competing interest has to be declared.

Reference

- [1] Wilson K, Brownstein JS. *Early detection of disease outbreaks using the Internet*. CMAJ. 2009;180(8):829-31.
- [2] Heymann DL, Rodier GR. *Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases*. Lancet Infect Dis. 2001;1(5):345-53.
- [3] Paquet C, Coulombier D, Kaiser R, Ciotti M. *Epidemic intelligence: a new framework for strengthening disease surveillance in Europe*. Euro Surveill. 2006;11(12):212-4.
- [4] Rotureau B, Barboza P, Tarantola A, Paquet C. *International epidemic intelligence at the Institut de Veille Sanitaire, France*. Emerg Infect Dis. 2007;13(10):1590-2.
- [5] Formenty P, Roth C, Gonzalez-Martin F, Grein T, et al. *Emergent pathogens, international surveillance and international health regulations (2005)*. Med Mal Infect. 2006;36(1):9-15.
- [6] Mykhalovskiy E, Weir L. *The Global Public Health Intelligence Network and early warning outbreak detection: a Canadian contribution to global public health*. Can J Public Health. 2006;97(1):42-4.
- [7] Brownstein JS, Freifeld CC. *HealthMap: the development of automated real-time Internet surveillance for epidemic intelligence*. Euro Surveill. 2007;12(11):E071129.
- [8] Doan S, Hung-Ngo Q, Kawazoe A, Collier N. *Global Health Monitor - A Web-based System for Detecting and Mapping Infectious Diseases*. Proceedings of the 3rd international joint conference on natural language processing (IICNLP08); 2008 Jan 7-12; Hyderabad, India.
- [9] Steinberger R, Fuat F, van der GE, Best C, von Etter P, et al. *Text Mining from the Web for Medical Intelligence*. In: Fogelman-Soulié F, Perrota D, Piskorski J, Steinberger R, editors. Mining Massive Data Sets for Security. Amsterdam: IOS Press; 2008. p. 295-310.
- [10] Linge JP, Steinberger R, Weber TP, Yangarber R, Van der Goot E, et al. *Internet surveillance systems for early alerting of health threats*. Euro Surveill. 2009;14(13).
- [11] Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V. *Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group*. MMWR Recomm Rep 2004;53(RR-5):1-11.
- [12] World Health Organization. *Communicable disease surveillance and response systems. Guide to monitoring and evaluating*. WHO/CDS/EPR/LYO/2006.2. 2006. Geneva, WHO Press.
- [13] Hartley DM, Nelson NP, Walters R, Arthur RR, Yangarber R, Madoff LC, et al. *Landscape of international event-based biosurveillance*. Emerg. Health Threats J. 2010 ;3:e3.
- [14] Madoff LC. *ProMED-mail: an early warning system for emerging diseases*. Clin Infect Dis. 2004;39(2):227-32.
- [15] Mawudeku A, Blench M. *Global Public Health Intelligence Network (GPHIN)*. Proceedings of the eighth Conference of the Association for Machine Translation in the Americas ; 2008 Oct 21-25; Hawai, USA. Elsevier:2008.
- [16] Nelson NP, Brownstein JS, Hartley DM. *Event-based biosurveillance of respiratory disease in Mexico, 2007-2009: connection to the 2009 influenza A(H1N1) pandemic?* Euro Surveill. 2010;15(30).
- [17] Collier N, Doan S, Kawazoe A, Goodwin RM, Conway M, et al. *BioCaster: detecting public health rumors with a Web-based text mining system*. Bioinformatics. 2008;24(24):2940-1.
- [18] Brownstein JS, Freifeld CC, Reis BY, Mandl K. *HealthMap: Internet-based emerging infectious disease intelligence. Global Infectious Disease Surveillance and Detection: Assessing the Challenges—Finding Solutions*. Workshop Summary. Washington: The National Academies Press; 2007. p. 122-36.
- [19] Keller M, Blench M, Tolentino H, Freifeld CC, Mandl KD, et al. *Use of unstructured event-based reports for global infectious disease surveillance*. Emerg Infect Dis 2009;15(5):689-95.

- [20] Pouliquen B, Steinberger R, Ignat C, De Groeve T. *Geographical Information Recognition and Visualisation in Texts Written in Various Languages*. Proceedings of the 19th Annual ACM Symposium on Applied Computing; 2004 March 14; Nicosia, Cyprus.
- [21] Woodall JP. *Global surveillance of emerging diseases: the ProMED-mail perspective*. Cad Saude Publica. 2001;17 Suppl:147-54.
- [22] Cowen P, Garland T, Hugh-Jones ME, Shimshony A, Handsides et al. *Evaluation of ProMED-mail as an electronic early warning system for emerging animal diseases: 1996 to 2004*. J Am Vet Med Assoc. 2006;229(7):1090-9.
- [23] Zeldenrust ME, Rahamat-Langendoen JC, Postma MJ, van Vliet JA. *The value of ProMED-mail for the Early Warning Committee in the Netherlands: more specific approach recommended*. Euro Surveill 2008;13(6).
- [24] Rortais A, Belyaeva J, Gemo M, van der GE, Linge JP. *MedISys: An early-warning system for the detection of (re-)emerging food and feed-borne hazards*. Food Research International. 2012; 43:1553-1556.
- [25] Mondor L, Brownstein JS, Chan EH, Sonricker AL, Madoff LC. *Does source matter? Comparing the timeliness of outbreak reports from governmental and nongovernmental sources*. Proceedings of the International Meeting on Emerging Diseases and Surveillance conference; 2011 Feb. 4; Vienna, Austria.
- [26] Lyon A, Nunn M, Gossel G, Burgman M. *Comparison of Web-Based Biosecurity Intelligence Systems: BioCaster, EpiSPIDER and HealthMap*. Transbound Emerg Dis. 2012;59(3):223-32.
- [27] Thomas CS, Nelson NP, Jahn GC, Niu T, Hartley DM. *Use of media and public-domain Internet sources for detection and assessment of plant health threats*. Emerging Health Threats Journal 4(7157). 2011.[PMID: 3168368] [PubMed - in process]
- [28] Grishman R, Huttunen S, Yangarber R. Information extraction for enhanced access to disease outbreak reports. J Biomed Inform. 2002;35(4):236-46.
- [29] Chan EH, Brewer TF, Madoff LC, Pollack MP, Sonricker AL, et al. *Global capacity for emerging infectious disease detection*. Proc Natl Acad Sci USA. 2010;107(50):21701-6.
- [30] Mantero J, Belyaeva E. *How to maximise event-based surveillance web-systems: the example of ECDC/JRC collaboration to improve the performance of MedISys*. Publications Office of the European Union; 2011. Report No: JRC63805.
- [31] Kawazoe A, Jin L, Shigematsu M, Bekki D, Barrero R, et al. *The development of a schema for the annotation of terms in the BioCaster disease detection/tracking system*. Second International Workshop on Formal Biomedical Knowledge Representation (KR-MED 2006); 2006 Nov 8; Baltimore, USA.
- [32] Chaudet H. *Extending the event calculus for tracking epidemic spread*. Artif Intell Med. 2006;38(2):137-56.
- [33] Pouliquen B, Kimler M, Steinberger R, Ignat C, Oellinger T, et al. *Geocoding Multilingual Texts: Recognition, Disambiguation and Visualisation*. Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC-2006) conference; 2006 May 26-27; Genoa, Italy. US: Cornell University Library; 2006.
- [34] Doan S, Kawazoe A, Collier N. *The Role of Roles in Classifying Annotated Biomedical Text*. Proceedings of the BioNLP 2007 conference; 2007 Apr 7; Prague, Czech Republic. Association for Computational Linguistics; 2007.
- [35] Zhang Y, Liu B. *Semantic Text Classification of Emergent Disease Reports*. Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD); 2007 Sept 21-27; Warsaw, Poland. USA: Springer; 2007.
- [36] Freifeld CC, Mandl KD, Reis BY, Brownstein JS. *HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports*. J Am Med Inform Assoc. 2008;15(2):150-7.
- [37] Conway M, Doan S, Kawazoe A, Collier N. *Classifying disease outbreak reports using n-grams and semantic features*. Int J Med Inform. 2009;78(12):e47-e58.
- [38] Doan S, Kawazoe A, Conway M, Collier N. *Towards role-based filtering of disease outbreak reports*. J Biomed Inform. 2009;42(5):773-80.
- [39] Chanlekha H, Collier N. *A methodology to enhance spatial understanding of disease outbreak events reported in news articles*. Int J Med Inform. 2010;79(4):284-96.
- [40] Chanlekha H, Kawazoe A, Collier N. *A framework for enhancing spatial and temporal granularity in report-based health surveillance systems*. BMC Med Inform Decis Mak. 2010 Jan 12;10:1.
- [41] Collier N. *What's unusual in online disease outbreak news?* Journal of Biomedical Semantics. 2010 March 31;1(2).
- [42] Collier N. *Towards cross-lingual alerting for bursty epidemic events*. Journal of Biomedical Semantics. 2011 Oct 6; 2 (Suppl 5):S10.
- [43] Conway M, Kawazoe A, Chanlekha H, Collier N. *Developing a disease outbreak event corpus*. J Med Internet Res. 2010;12(3):e43.
- [44] Von Etter P, Huttunen S, Vihavainen A, Vuorinen M, Yangarber R. *Assessment of Utility in Web Mining for the Domain of Public Health*. Proceedings of the NAACL-HLT 2010 Second Louhi Workshop on Text and Data Mining of Health Documents; 2010 Jun 5-6; Los Angeles, USA.
- [45] Torii M, Yin L, Nguyen T, Mazumdar CT, Liu H, et al. *An exploratory study of a text classification framework for Internet-based surveillance of emerging epidemics*. Int J Med Inform. 2011;80(1):56-66.
- [46] Woodall J. *Official versus unofficial outbreak reporting through the Internet*. Int J Med Inform. 1997;47(1-2):31-4.
- [47] World Health Organization. *Weekly Epidemiological Record*. Geneva, Switzerland. 2011, [cited 2012 Jun 01]. Available from: <http://www.who.int/wer/en/>
- [48] World Organisation for Animal Health (OIE). *The World Animal Health Information System (WAHIS)*. Paris, France: 2011 [cited 2012 Jun 01]; Available from: <http://www.oie.int/en/animal-health-in-the-world/the-world-animal-health-information-system/the-oie-data-system/>
- [49] Centers for Disease Control and Prevention (USA). *Morbidity and Mortality Weekly Report*. Atlanta: 2011 [cited 2011 Dec 21]. Available from: <http://www.cdc.gov/mmwr/>
- [50] Health Protection Agency, United Kingdom. *Health Protection Report*: 2011 [cited 2012 Jun 01]. Available from: <http://www.hpa.org.uk>
- [51] European Centre for Disease Prevention and Control (ECDC). *Communicable Disease Threat Report*. Stockholm: 2011 [cited 2012 Jun 01]. Available from: <http://ecdc.europa.eu/en/activities/epidemicintelligence/Pages/EpidemicIntelligenceThreatCommunicationCDTR.aspx>
- [52] Institut de Veille Sanitaire (InVS). *Bulletin Hebdomadaire International*. France 2011 [cited 2012 Jun 01]. Available from: <http://www.invs.sante.fr/Publications-et-outils/Bulletin-hebdomadaire-international>
- [53] Dente MG, Fabiani M, Gnesotto R, Putoto G, Montagna C, et al. *EpiSouth: a network for communicable disease control in the Mediterranean region and the Balkans*. Euro Surveill. 2009;14(5).
- [54] Vaillant L, Nys J, Barboza P. *Epidemic Intelligence: Assessing event-based tools and users' perception in the GHSAG community*. *International Meeting on Emerging Diseases and Surveillance*; 2011 Feb 4-6. Vienna, Austria.

ANNEXE 3 - PRINCIPAUX SYSTEMES DE BIOSURVEILLANCE

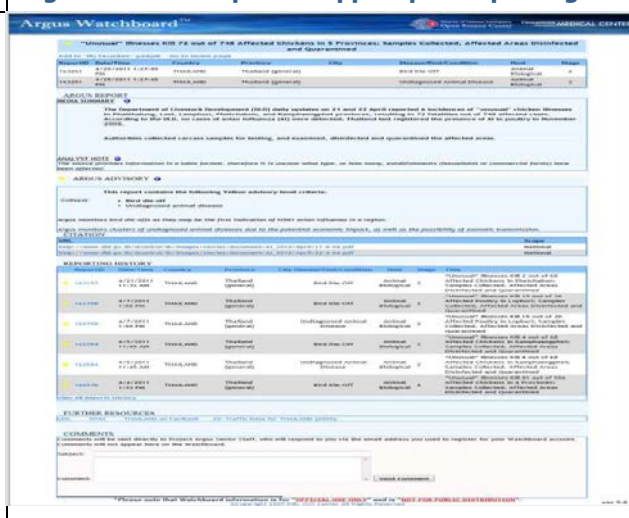
ARGUS

Argus, est un système développé par le Centre médical de l'Université de Georgetown (États-Unis) et financé par le gouvernement des États-Unis. Il était initialement conçu pour détecter et de suivre l'évolution des événements biologiques (santé humaine, animale et végétale) survenant dans le monde entier (sauf les États-Unis) et. Depuis le 1^{er} février 2012, les activités ont été réorientées vers la surveillance des événements sociaux et politiques et ne couvrent plus la détection des risques sanitaires). Argus collectait à l'aide d'un processus automatisé, des informations dans leur langue originale de publication, qu'il s'agisse de médias, les blogs ou de sources officielles. Les événements étaient classés en fonction d'un ensemble de concepts et d'une taxonomie spécifique. Les éléments de taxonomie définissaient des indicateurs directs (concernant la maladie) et six catégories d'indicateurs indirects : reconnaissance officielle de l'évènement, mise en œuvre de mesure de contrôle, impact sur les infrastructures, perception locale de la menace, impact sur les structures de soins et stocks de médicaments et impact économique lié à l'évènement. L'équipe d'analystes d'Argus, était composée d'environ 40 spécialistes multilingues qui vérifiaient quotidiennement plusieurs milliers de sources Internet dans une quarantaine de langues différentes. A partir d'une base de données dynamique mise à jour six fois par jour, les signaux pertinents étaient sélectionnés par des recherches booléennes et un outil bayésien. Des rapports d'évènements basés sur les sources médias, étaient régulièrement mises à jour. Un score était attribué à chaque évènement en fonction son évolution, de la dynamique connue ou précédemment décrite pour le même type d'évènement. Ces rapports étaient postés sur le portail Internet d'Argus à l'intention des utilisateurs.⁹³

Figure 11 – Portail d'Argus



Figure 12 – Exemple de rapport publié par Argus



BioCaster

BioCaster^{29,30,31} (<http://born.nii.ac.jp/index.php>) a été créé en 2006, à l'institut national d'informatique du Japon (National Institute of Informatics (NII)), en collaboration avec 5 autres instituts de 3 pays (dont le National Institute of Infectious Diseases du Japon). BioCaster mettait à disposition un portail destiné aux chercheurs, cliniciens et personnels de santé. Ce système, complètement automatisé, utilisait des flux RSS sans intervention humaine. Sa cible principale était les maladies transmissibles et BioCaster portait une attention particulière sur les pays de l'Asie du Sud, et traite des articles dans 13 langues dont le chinois, le japonais, le thaï, le coréen et le vietnamien. Les principales sources étaient Google News et l'European Média Monitor (EMM). BioCaster était le seul des 7 systèmes dont la stratégie linguistique consistait à d'abord traduire intégralement le texte original en anglais puis à appliquer les algorithmes de sélection aux textes traduits. Un module d'extraction de texte (datamining) basé sur un processus d'apprentissage automatique à partir d'une collection multilingue d'articles annotés par les analystes a été développé. BioCaster a cessé d'être opéré en 2013.

Figure 13 – Portail BioCaster

Figure 14 – Exemple de rapport publié BioCaster

Gphin

Le Réseau mondial d'information en santé publique plus connu sous son nom anglais Global Public Health Intelligence Network (Gphin)^{33,94,95,96} a été créé en 1997 par le centre de préparation et réponse aux urgences de l'Agence de santé publique du Canada en collaboration avec l'OMS. GPHIN est un système sécurisé, modéré par une équipe d'analystes multilingues et son accès est restreint aux organismes publics. Il collecte analyse et trie, en temps réel (24 h/7j), des signaux provenant principalement d'agrégateurs de médias et en particulier de Factiva (<http://www.dowjones.com/factiva/>). Ces sources sont complétées par des recherches *ad hoc* effectuées manuellement par les analystes. 9 langues (anglais, arabe chinois simplifié, chinois traditionnel, espagnol, français, persan, portugais et russe) sont prises en charge. Gphin est le seul système qui permet une traduction multilingue de toutes les langues prises en charge. La qualité de la traduction automatique est revue par des analystes et éventuellement améliorée manuellement. Tous les risques sanitaires sont couverts : infectieux, NRBC, environnementaux et produits de santé (médicaments, vaccins). La sélection automatique des articles est basée sur l'utilisation de mots clés (ontologie) et de leurs synonymes y compris en langage familier. Le traitement de l'information est effectué systématiquement dans la langue d'origine de publication. L'information est d'abord filtrée en fonction de sa pertinence par un processus automatisé complété par une analyse humaine. Les informations sont classées et stockées dans une base de données accessible aux utilisateurs. Les informations répondant aux critères d'urgence définis dans l'algorithme (correspondant potentiellement à des événements majeurs de santé publique) sont transmises par email automatiquement par le système ou manuellement par les analystes.

Figure 15 – Portail Gphin

Articles 1 to 20	Page 1 / 17				
2011-04-25 19:27 GMT	Killing deer to control ticks is ridiculous	The Patriot Ledger	EN	26	REAP
2011-04-25 19:27 GMT	Cervical cancer vaccines, tests donated to Rwanda	Randev Health E-Line	EN	64	HMAECCOP
2011-04-25 19:27 GMT	Ministry backs call to defer nuclear plants	Bangkok Post	EN	40	MSE
2011-04-25 19:27 GMT	Worker testifies about key pipe's break before deadly blaze at toxic '07 ground zero	Associated Press Newswire	EN	52	MSECP
2011-04-25 19:27 GMT	Issue	IT&T-TASS World Service	EN	80	HMAECCOP
2011-04-25 19:27 GMT	Karachi CM observes fast to enforce ban on endosulfan	Business Line (The Hindu)	EN	38	HMAECCOP
2011-04-25 19:27 GMT	Sandoz-Aventis gets expanded FDA approval for meningitis vaccine use in young infants	Associated Press Newswire	EN	37	HMAECCOP
2011-04-25 19:27 GMT	St. Augustine Gold and Copper Limited Announces Developments at the King-King	Marketwire	EN	35	HMAECCOP
2011-04-25 19:27 GMT	UN envoy says \$5 billion malaria fight has saved several thousand lives in recent years	Associated Press Newswire	EN	32	HMAECCOP
2011-04-25 19:27 GMT	Colombia launches major emergency ops after floods	Agence France Presse	EN	30	HMAECCOP
2011-04-25 19:27 GMT	UAE BANS LIVE BIRDS IMPORTS FROM S. AFRICA	States News Service	EN	34	HMAECP
2011-04-25 19:27 GMT	BRBP: 22,400 gallons of untreated sewage spills into Bradley Creek	Honoring Star (NCT)	EN	50	SEC
2011-04-25 19:27 GMT	Flash flood warning issued for Illinois river basin and 13 Oklahoma counties	The Daily Oklahoman (NCT)	EN	64	SCP
2011-04-25 19:27 GMT	Vegetables poisoned by chemicals	The Financial Express (Bangladesh)	EN	58	NDCOP
2011-04-25 19:27 GMT	US FDA sees safety issue with Merck hepatitis drug	Randev Health E-Line	EN	34	HMAECCOP
2011-04-25 19:27 GMT	TB study points to drugs to halt lung destruction	Randev Health E-Line	EN	26	HMAECCOP
2011-04-25 19:27 GMT	INDONESIAN QUAKE	Comicon Post	EN	50	E
2011-04-25 19:27 GMT	Official: Iran Has Uncovered 2nd Cyber Attack -AP	Div Jones News Service	EN	48	NDCP
2011-04-25 19:27 GMT	South Carolina emergency management officials say rural dam could fail, flooding some areas	Associated Press Newswire	EN	50	MSECP
2011-04-25 19:27 GMT	SHALLOW QUAKE	The Press (Christchurch)	EN	50	E

Figure 16 – Exemple de rapport de rapport publié par Gphin

The screenshot shows a report page with a world map at the top. Below the map, there is a section titled 'Article' and a large block of text. The text discusses the distribution of the last known species of a certain virus, mentioning the WHO and the impact of the virus on the population. It also mentions the need for a vaccine and the role of the WHO in coordinating the response. The text is in French and appears to be a translation of an article from Factiva.

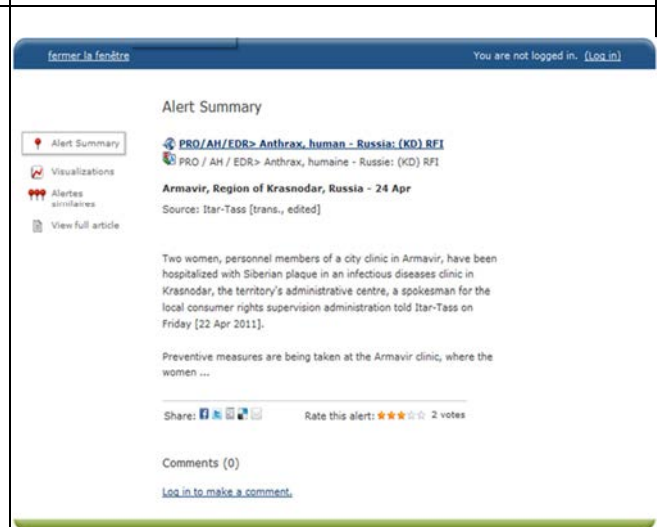
HealthMap

HealthMap^{35, 36, 37}, <http://healthmap.org/fr/> est un système en accès libre développé depuis 2006, par une équipe de chercheurs, d'épidémiologistes et de développeurs de logiciels de l'Hôpital pour enfants de Boston (États Unis). La cible principale d'HealthMap est les maladies infectieuses (humaine et zoonoses). Ses principales sources (environ 20 000) sont constituées de : ProMED-mail, d'agrégateurs de média en ligne en particulier Google News, des groupes de discussion, de rapports officiels et publics (notamment ceux provenant de l'OMS, OIE ou la FAO). HealthMap permet également des contributions directes des utilisateurs. Le processus de détection et de traitement est automatisé (24h/7j). Sept langues sont prises en charge (anglais, arabe, chinois, espagnol, français, russe, et portugais) et un module de traduction automatique est intégré. En 2010, HealthMap générait environ 300 alertes par jour couvrant plus de 140 agents pathogènes dans 174 pays. Les informations filtrées et triées mises à disposition du public sur une carte mondiale basée sur la technologie Google.

Figure 17 – Portail HealthMap



Figure 18 – Exemple de rapport HealthMap



MedISys

Le Medical Information System MedISys (MedISys)⁴⁰ <http://medusa.jrc.it/medisys/> est une initiative de la Direction générale de la santé et des consommateurs de la commission européenne (DG SANCO), en collaboration avec Joint Research Centre (JRC) en Italie. Les champs couverts par MedISys portent sur les événements de type NRBC. Ce système entièrement automatisé. En 2012 environ 20 000 articles provenant ~10 000 flux RSS et pages HTML de 3 750 sites génériques et 20 sites commerciaux d'agrégation de nouvelles en 60 langues étaient traités quotidiennement⁹⁷. Les articles sont classés automatiquement dans des catégories prédéfinies sur la base 350 mots-clés mis à jour régulièrement. L'information est disponible dans le format initial (identique à la source). MedISys n'intègre pas de module de traduction automatique, mais d'un lien vers Google traduction. Les informations collectées sont analysées en fonction de la zone géographique et par type de pathologie (analyse de tendance) pour détecter une augmentation du volume d'articles dans l'une des catégories de pathologie et pour un pays spécifique. Il intègre des fonctionnalités permettant aux utilisateurs de développer des filtres adaptés à leurs besoins par exemple de sélectionner les rapports en fonction de la pathologie, de la zone géographique ou de la langue. Le niveau de sensibilité des alertes envoyées automatiquement par email peut être adapté.

Figure 19 – Portail MedISys

The screenshot shows the MedISys home page with a sidebar on the left containing navigation menus for 'Home', 'Alert Statistics', 'Diseases', 'Symptoms', and 'External Links'. The main content area features 'Rubriques les plus actives' with sections for Tuberculosis, Malaria, and Meningitis. A 'Recent disease incidents provided by the University of Helsinki' table is visible, listing diseases like Hepatitis, Typhoid, and Cholera with their locations and case counts.

Disease	Time	Location	Cases
Hepatitis	2011-11	Austria	George Michael
Hepatitis	---	Spain	third British passenger
Hepatitis	---	Spain	more than 20 people
Typhoid	2012-01-16-2012-01-22	Dominika	279 typical cases
Cholera	2011-12	UK	various

Figure 20 – MedISys, sélection par pathologie

The screenshot shows the MedISys interface with a filter applied to 'Grippe aviaire'. The main content area displays 'Dernières nouvelles - Grippe aviaire' with several news items, including 'Avian flu outbreak leads to ban on poultry imports' and '02/02/2012: Low pathogenic avian influenza (poultry), Sri Lanka'. A 'Show World Map' section at the bottom right shows a map of Asia with a bar chart indicating the number of articles published in various regions.

ProMED-mail

Fondé en 1994, ProMED-mail (Program for Monitoring Emerging Diseases) ^{41, 98}, est le premier système basé sur Internet à avoir utilisé des sources officielles et officieuses pour la détection de d'évènements sanitaires. Depuis 1999 – il est opéré en tant que programme de l'International Society for Infectious Diseases (ISID) <http://www.promedmail.org/>. Ces champs de détection couvrent les maladies infectieuses, les expositions aiguës à des substances toxiques, mais également les maladies animales et les maladies affectant des plantes cultivées pour l'alimentation humaine ou animale. Les sources d'information comprennent les médias, les rapports officiels, des résumés en ligne, des informations fournies par les correspondants locaux et notamment les abonnés. Une des originalités de ProMED-mail est justement d'inciter les 40 000 abonnés présents dans 185 pays à partager les informations dont ils disposent. Une équipe d'une cinquantaine de spécialistes de différents domaines (maladies infectieuses, santé humaine, animale et végétale) trient et examinent les rapports avant leur diffusion aux abonnés (par email et mise en ligne sur le site ProMED-mail). ProMED-mail est décliné en six programmes régionaux multilingues : Amérique Latine (espagnol et portugais), Afrique de l'Ouest (français), Afrique de l'Est (anglais), pays de l'ex-Union Soviétique (Russe) et Pays du bassin du Mékong (anglais).

Figure 21 – Portail ProMED-mail



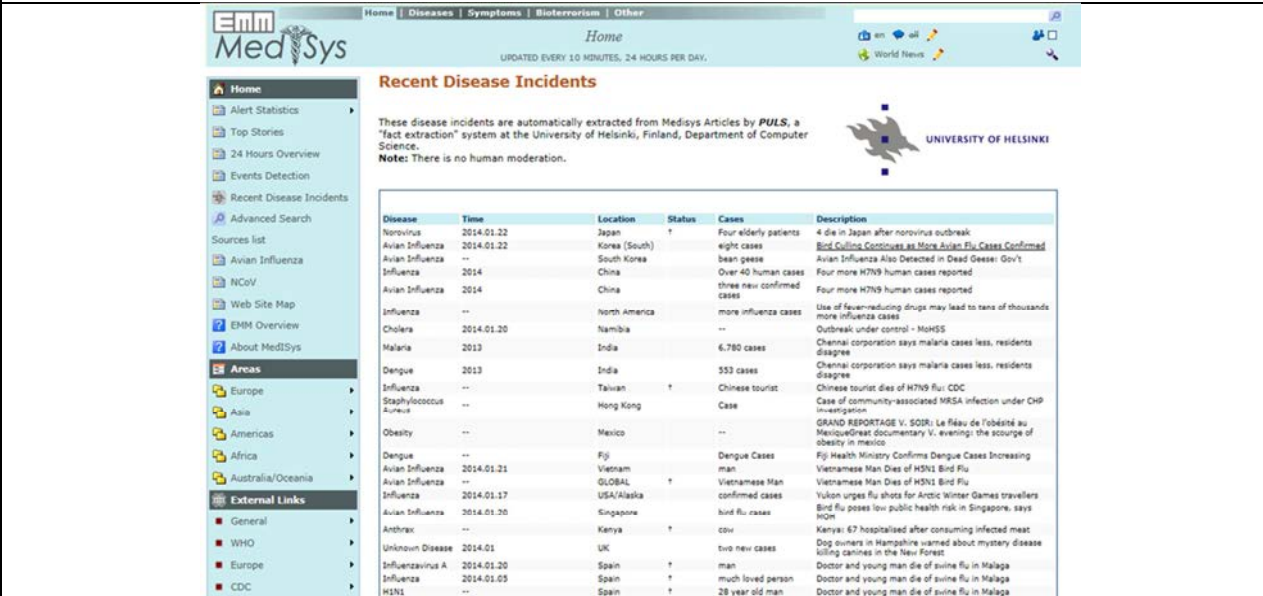
Figure 22 – Exemple de rapport ProMED



PULS

Le pattern-based Understanding and Learning System (PULS) (<http://sysdb.cs.helsinki.fi/puls/jrc/table/all>)^{44,45,46} est un projet développé par l'Université d'Helsinki en collaboration avec l'ECDC et le ministère finlandais de la santé. Il s'agit d'un système dont l'objectif primaire n'est pas la détection de signaux bruts, mais leur traitement linguistique basé sur la technique du traitement automatique des langues (NLP natural language processing) utilisé pour la recherche, l'extraction, l'agrégation et la visualisation des informations. PULS est complètement automatisé et utilise principalement ProMed et MedISys comme sources d'information. En particulier, PULS capture et analyse les documents identifiés par MedISys. Il les sélectionne, les tri et les regroupe par évènement. Une analyse linguistique automatique est effectuée et les informations importantes contenues dans le texte sont extraites automatiquement (par exemple nombre de cas localisation géographique pathologie, etc.). Les résultats de traitement est ensuite réinjecté dans MedISys. En 2010, Les données extraites par PULS représentaient environ 300 000 articles. Les informations sont présentées sous un format de bases de données sur Google Earth.

Figure 23 – PULS : exemple de traitement des informations de MedISys par PULS



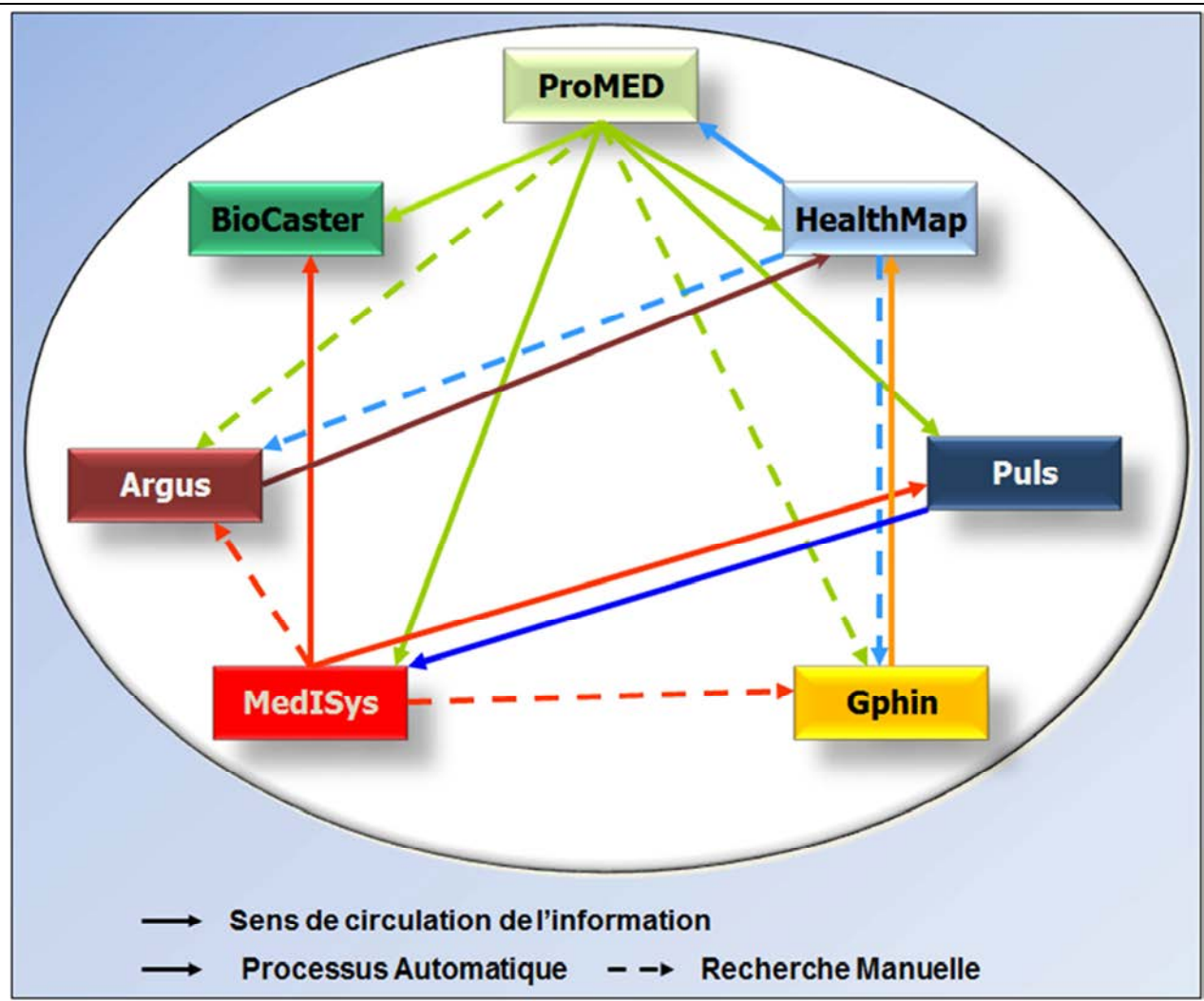
The screenshot shows the MedISys website interface. The main content area is titled "Recent Disease Incidents" and contains a table of disease incidents. The table has columns for Disease, Time, Location, Status, Cases, and Description. Below the table, there is a note stating that these incidents are automatically extracted from MedISys articles by PULS, a "fact extraction" system at the University of Helsinki, Finland, Department of Computer Science. A note also states that there is no human moderation.

Disease	Time	Location	Status	Cases	Description
Norovirus	2014.01.22	Japan	†	Four elderly patients	4 die in Japan after norovirus outbreak
Avian Influenza	2014.01.22	Korea (South)		eight cases	Bird Culling Continues as More Avian Flu Cases Confirmed
Avian Influenza	--	South Korea		bean geese	Avian Influenza Also Detected in Dead Geese: Gov't
Influenza	2014	China		Over 40 human cases	Four more H7N9 human cases reported
Avian Influenza	2014	China		three new confirmed cases	Four more H7N9 human cases reported
Influenza	--	North America		more influenza cases	Use of fever-reducing drugs may lead to tens of thousands more influenza cases
Cholera	2014.01.20	Namibia		--	Outbreak under control - MoHSS
Malaria	2013	India		6,780 cases	Chennai corporation says malaria cases less, residents disagree
Dengue	2013	India		553 cases	Chennai corporation says malaria cases less, residents disagree
Influenza	--	Taiwan	†	Chinese tourist	Chinese tourist dies of H7N9 flu: CDC
Staphylococcus aureus	--	Hong Kong		Case	Case of community-associated MRSA infection under CHP investigation
Obesity	--	Mexico		--	GRAND REPORTAGE V. SQR: Le fléau de l'obésité au Mexique/Great documentary V. evening: the scourge of obesity in Mexico
Dengue	--	Fiji		Dengue Cases	Fiji Health Ministry Confirms Dengue Cases Increasing
Avian Influenza	2014.01.21	Vietnam		man	Vietnamese Man Dies of H5N1 Bird Flu
Avian Influenza	--	GLOBAL	†	Vietnamese Man	Vietnamese Man Dies of H5N1 Bird Flu
Influenza	2014.01.17	USA/Alaska		confirmed cases	Yukon urges flu shots for Arctic Winter Games travellers
Avian Influenza	2014.01.09	Singapore		bird flu cases	Bird flu poses low public health risk in Singapore, says MoH
Anthrax	--	Kenya	†	cow	Kenya: 67 hospitalised after consuming infected meat
Unknown Disease	2014.01	UK		two new cases	Dog owners in Hampshire warned about mystery disease killing canines in the New Forest
Influenza A	2014.01.20	Spain	†	man	Doctor and young man die of swine flu in Malaga
Influenza	2014.01.03	Spain	†	much loved person	Doctor and young man die of swine flu in Malaga
H1N1	--	Spain	†	28 year old man	Doctor and young man die of swine flu in Malaga
Polio	2013	Pakistan		83 new polio cases	Deadly Pakistani airstrikes target militants believed responsible for recent attacks

Interactions entre systèmes

Différents circuits d'information coexistent. Certains systèmes partagent et/ou échangent leurs informations (ProMed-HealthMap, HealthMap-ARGUS, PULS-MedISys, PULS-ProMED). Certains systèmes sont quasiment indépendants c.-à-d. qu'ils ne dépendent pas d'autres systèmes de biosurveillance. C'est le cas d'Argus, Gphin et ProMED. Ils peuvent occasionnellement intégrer des signaux provenant d'autres systèmes, mais dans ce cas la détection résulte de recherches manuelles effectuées par les analystes. Les quatre autres systèmes intègrent comme sources de signaux bruts d'autres systèmes de biosurveillance. Le processus de détection est alors entièrement automatisé. ProMED, en raison de son ancienneté et de son accès libre est particulièrement utilisé. Les interactions peuvent être complexes comme dans le cas de l'interaction PULS-MedISys, les principales sont résumées dans la Figure 24

Figure 24 – Interactions simplifiées entre systèmes de biosurveillance (Source EAR) ⁵¹



ANNEXE 4 - LA VEILLE INTERNATIONALE A L'INVS

A. CONTEXTE FRANÇAIS

- **Complexité du territoire national** : avec 11 départements et collectivité d'outre-mer répartis sur 4 continents (Afrique, Amérique, Europe, Océanie), la France est le seul pays de l'Union Européenne à avoir des frontières terrestres avec des pays situés hors de l'espace euro-méditerranéen (Brésil et Surinam).
- **La démocratisation du tourisme** est une source d'exposition des citoyens aux nouveaux risques sanitaires et elle augmente le risque d'importer des maladies absentes sur le territoire national
- **Flux migratoires**

- **Emigration.** Au 31 déc. 2011, près d'1,6 millions de personnes étaient inscrits au registre mondial des Français établis hors de France⁹⁹. Plus de la moitié (51%) résidait en Europe, 19% en Amérique, 14% en Afrique, 9% au Proche et Moyen-Orient et 7% en Asie-Océanie. (Tableau 16)

Tableau 16 – Français inscrits au registre mondial au 31/12/2011 (source : MAE¹⁰⁰)

zones	n	% total	% hors UE
U.E.	619 712	39%	-
Afrique du Nord	94 789	6%	10%
Afrique Francophone	116 862	7%	12%
Afrique non-francophone	18 621	1%	2%
Proche et Moyen-Orient	137 679	9%	14%
Asie et Océanie	116 729	7%	12%
Europe non-UE (dont CEI)	190 780	12%	20%
Amérique Centrale et du Sud	97 117	6%	10%
Amérique du Nord	202 014	13%	21%
Total	1 594 303	100%	
Total hors UE	2 447 862		100%

- **Immigration.** D'après l'INSEE, en 2009, 5,4 millions d'immigrés soit 8,4% de la population ont été recensés en France métropolitaine, dont 3,8 millions de personnes de nationalité étrangère (5,9% de la population). Parmi ces dernières, 36% avaient la nationalité de l'un des pays de l'UE, 28% de l'un des 3 pays francophones du Maghreb, 11% d'un pays d'Afrique francophone, 7% d'un pays du Proche et Moyen-Orient.

- De par son **histoire**, la France a conservé des contacts privilégiés avec certaines parties du monde. L'origine des flux migratoires et des populations étrangères résidant en France sont très différentes de celle des autres grands pays européens. En 2009, les ressortissants des ex-colonies françaises représentaient 60% des immigrés en métropole (Tableau 17)

Tableau 17 – Immigrés par pays de naissance, recensement 2009 (source : adapté de INSEE¹⁰¹)

zones	n	% Total	% hors UE
UE	1 323 279	35%	-
Europe hors UE (dont CEI)	183 532	5%	7%
Maghreb	1 052 223	28%	43%
Afrique francophone	422 236	11%	17%
Afrique non-francophone	45 202	1%	2%
Proche et Moyen-Orient	266 584	7%	11%
Asie	245 699	7%	10%
Amériques	227 052	6%	9%
Océanie	5 335	0,1%	0,2%
Total	3 771 141	100%	
Total hors UE	2 447 862		100%

B. CREATION DE LA VEILLE INTERNATIONALE (VI)

En 2002, l'analyse de la situation a mis en évidence que les systèmes d'alerte précoce internationaux existant ne couvraient que partiellement les aux besoins spécifiques de la France. La VI a donc été conçue par le Département International de l'InVS avec pour objectif de détecter toutes les menaces survenant à l'étranger et pouvant affecter la population vivant sur le territoire national ou la population française vivant à l'étranger. Une méthodologie stricte a été élaborée pour obtenir les informations les plus fiables possibles. Elle consiste à détecter des signaux qui pourraient révéler une menace sanitaire, à les trier, les vérifier et les valider puis analyser et interpréter les données disponibles et communiquer quand nécessaire.

C. DETECTION DES SIGNAUX SYSTEMES DE BIOSURVEILLANCE

La VI a été développée par à une période où très peu de systèmes de biosurveillance existaient et où le nombre d'institutions nationales pratiquant ce type d'activité était extrêmement limité. En 2002,

date du début des activités, seul système disponible était ProMED. En dépit de ses qualités indéniables ProMED est de tous les systèmes présentés ici le moins structuré d'un point de vue technologique. Sa valeur ajoutée est ailleurs. Elle réside dans son important réseau d'abonnés (qui constitue une part essentielle des sources d'information) et dans son équipe d'experts. Dès 2003, GPHIN est venu compléter le dispositif cette fois-ci avec une vraie technologie permettant d'envisager des recherches systématisées sur Internet. Jusqu'en 2009, GPHIN et ProMED étaient les seuls systèmes utilisés systématiquement en routine. Si d'autres systèmes étaient disponibles sur cette période (notamment MedISys), ils n'étaient pas utilisés en routine en raison de leurs faibles performances (en particulier la très faible spécificité des signaux bruts détectés) ou de leur caractère encore trop expérimental.

En 2007, une analyse des événements sanitaires détectés et vérifiés par l'équipe de VI du DI de l'InVS a été réalisée.⁶⁴ Elle portait sur 200 événements publiés dans le bulletin hebdomadaire international (BHI) et survenus entre le 17 mai et le 27 décembre 2007. Parmi les 176 événements pour lesquels les données étaient disponibles, la première source d'information était GPHIN avec 36% (63/176) des événements détectés, suivi par l'OMS 29% (51/176) et de ProMED 17% (30/176). Pour 60% (105/176) des événements détectés le premier signal provenant d'une source informelle et pour 40% (71/176) d'une source formelle. Parmi les 176 événements 20% (35/176) avaient été détectés par des recherches non-spécifiques réalisées de manière *ad hoc* sur Internet (hors systèmes de biosurveillance). Cette part importante d'événements non-détectés par les systèmes illustrant à la fois les limites des systèmes existants en 2007, les besoins de développement dans ce domaine et l'impératif d'utiliser différents systèmes.

Depuis 2007 d'autres systèmes sont venus enrichir le paysage de l'intelligence épidémiologique. Ils ont ainsi progressivement été inclus dans les activités VI de routine en fonction de leur développement et ou de leur accessibilité (ex : début du développement de HealthMap en 2006 ; Argus n'a été accessible à l'InVS que dans le cadre du projet EAR en 2009). Les inclusions ont été progressives, une évolution simplifiée de l'utilisation des systèmes dans le cadre de la VI sur la période 2011-2012 est présenté dans le Tableau 18.

En l'absence d'évaluations adéquates publiées dans la littérature scientifique, l'utilisation des différents systèmes était basée sur la perception des spécificités intrinsèques des différents systèmes par exemple : leur couverture géographique, les langues prises en charge, les domaines couverts, les processus de sélection des signaux bruts (automatisés ou modéré), etc. Cette perception des spécificités des systèmes était basée exclusivement sur l'expérience pratique des épidémiologistes qui les utilisaient en routine. Cette connaissance, bâtie sur une base principalement empirique, est très personne-dépendante. Les systèmes étant en perpétuelle évolution, il est très difficile avec une équipe réduite de concilier d'une part des activités de routine et de maintenir un niveau de connaissance approfondie des systèmes et de leurs fonctionnalités.

Tableau 18 – Evolution simplifiée de l'utilisation des systèmes experts dans le cadre de la veille internationale effectuée par le département international de l'InVS

Type d'utilisation	2010	2011	2012
Systématique et quotidienne	ARGUS	ARGUS	GPHIN
	GPHIN	GPHIN	ProMED-mail
	ProMED-mail	ProMED-mail HealthMap	HealthMap RSOE-EDIS
En complément	BioCaster	BioCaster	BioCaster
	HealthMap	MedISys	MedISys
	MedISys	RSOE-EDIS	
	RSOE-EDIS		

Nb les systèmes sont présentés par ordre alphabétique.

D. CRITERE DE SELECTION DES SIGNAUX

Afin de remplir la mission de la VI des critères ont été définis afin de sélectionner les événements devant être analysés et suivis dans le cadre de la VI à INVS. Ces critères permettent de répondre à deux questions :

- La menace est-elle sérieuse ?
- La menace sanitaire peut-elle toucher le territoire national ou la population française à l'étranger?

Trois types de critères sont utilisés de façon conjointe : géographiques et populationnels, liés à la gravité de la menace et enfin aux caractéristiques de l'agent en cause (Tableau 19 & Tableau 20).

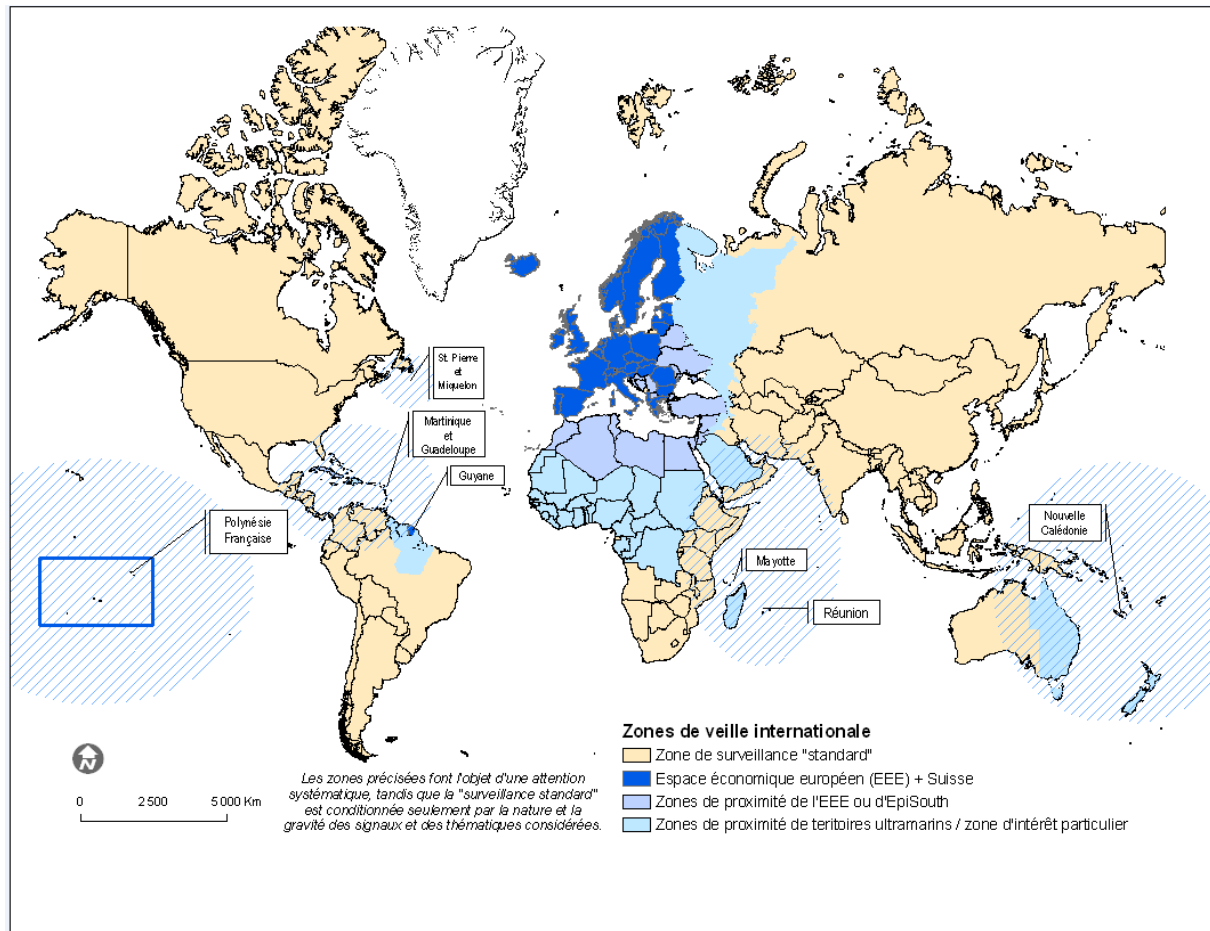
Tableau 19 – Critères de sélection des signaux dans le cadre de la Veille internationale mise en œuvre par le Département International de l'InVS

Géographique / populationnel	Sévérité / dynamique	Agent
<ul style="list-style-type: none"> Risque d'atteinte du territoire national (outremer inclus) Risque d'importation de cas sur le territoire national (outremer inclus) Atteinte d'une région proche de la France (outremer inclus) Atteinte d'un des principaux pays d'origine des migrants en France Atteinte d'un pays à forte communauté française expatriée Atteinte d'une destination des touristes et voyageurs français Grands rassemblements, pèlerinages.... 	<ul style="list-style-type: none"> Morbidité, Incidence, Létalité, Nombre de décès, Mortalité Gravité, fréquence des hospitalisations Séquelles Temps (durée, dynamique de l'épidémie) Localisation : rural/urbain; isolement de la zone touchée; densité de la population ;... Groupes à risque Population particulière (personnel soignant, hôpital...) 	<ul style="list-style-type: none"> Virulence Pathogénicité Transmissibilité Mode de transmission Capacité de diffusion Existence de moyens de prévention et de contrôle ; facilité et disponibilité de leurs mises en œuvre Changements épidémiologiques et biologiques (résistances, distribution géographique,...) Degré de connaissance sur l'agent
<ul style="list-style-type: none"> Toute émergence d'un nouvel agent est considérée comme potentiellement dangereuse 		
<ul style="list-style-type: none"> Evènements médiatiques qui requièrent d'être remis dans leur contexte épidémiologique 		

Tableau 20 – Exemples d'utilisation des critères

Nature du Signal	Critères considérés dans le tri	Résultat tri
Nombreux cas humains de rage dans une province du centre de la Chine	<ul style="list-style-type: none"> ↗ Létalité très élevée ↘ Pas de transmission interhumaine; pathologie endémique; hors zone prioritaire 	Non retenu
Quelque cas humains de rage à Bali (en 2008)	<ul style="list-style-type: none"> ↗ Rage non documentée à Bali auparavant; zone fortement touristique ; risque non-négligeable même pour les touristes (nombreux chiens errants) 	Retenu
2 cas autochtones de paludisme aux Bahamas	<ul style="list-style-type: none"> ↗ Paludisme autochtone absent des Caraïbes; modification recommandations voyageurs; zone touristique; espace régional DOM; existence de mesure de prophylactiques 	Retenu
1000 cas de méningite à méningocoque A au Mali entre janvier et juin	<ul style="list-style-type: none"> ↗ Zone prioritaire (forte communauté française et pays d'immigration); zone d'endémicité connue et bien signalée ↘ Chiffres conformes aux nombres attendus en regard des données historiques ; mesures prophylactiques déjà recommandées 	Non retenu
150 cas de choléra dans la capitale d'un pays Afrique francophone	<ul style="list-style-type: none"> ↘ Zone d'endémie connue ↗ Risque de propagation (zone urbaine); forte communauté française et pays d'immigration; risque de cas importés; Fréquence liaisons aériennes 	Retenu
Cas suspects de chikungunya dans un pays de l'espace régional métropole, DOM ou COM	<ul style="list-style-type: none"> ↘ Transmission vectorielle; très faible létalité ↗ Pathologie absente de l'espace régional; Risque de propagation important; Risque de cas importés; Mesure de contrôle à mettre en place 	Retenu
Epidémie de fièvre de la vallée du Rift liée à des inondations en Afrique de l'Est	<ul style="list-style-type: none"> ↘ Phénomène connu, pas de transmission interhumaine ni vectorielle documentée ↗ Espace régional DOM COM; Risque de diffusion régional (via le commerce de petit bétail) 	Retenu
Epidémie de gastro-entérite à <i>norovirus</i> dans un pays frontalier	<ul style="list-style-type: none"> ↘ Endémique en France; Pas de caractère nouveau rapporté; Très faible létalité; Symptômes modérés; peu d'hospitalisation. 	Non retenu
Lait de soja à très forte teneur en iode produit au Japon	<ul style="list-style-type: none"> ↗ Produit susceptible d'être disponible dans certains espaces régionaux DOM COM ; malades signalés chez des personnes ayant consommé ce produit ; mesures de contrôle simples (retrait des lots contaminés) 	Retenu

Figure 25 – Zones prioritaires définies pour la Veille Internationale de l'InVS



BIBLIOGRAPHIE

- ¹ Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. *Nature*. 2004;430(6996):242–9.
- ² Feldmann H, Czub M, Jones S, Dick D, Garbutt M, Grolla A, et al. Emerging and re-emerging infectious diseases. *Med Microbiol Immunol*. 2002; 191(2):63-74.
- ³ Heymann DL, Rodier GR. Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases. *Lancet Infect Dis*. 2001; 1(5):345-353.
- ⁴ Formenty P, Roth C, Gonzalez-Martin F, Grein T, Ryan M, Drury P, et al. Emergent pathogens, international surveillance and international health regulations (2005). *Med Mal Infect*. 2006; 36:9-15.
- ⁵ O'Fallon BD, Fehren-Schmitz L. Native Americans experienced a strong population bottleneck coincident with European contact. *Proc Natl Acad Sci*. 2011;108(51): 20444-8.
- ⁶ Morens DM; Folkers KF; Fauci AS. Emerging infections: a perpetual challenge. *Lancet Infect Dis*. 2008; 8(11), 710–719
- ⁷ Métais P. Démographie des Néo-Calédoniens. *Journal de la Société des océanistes*. 1953, Tome 9,. pp. 99-128. http://www.persee.fr/web/revues/home/prescript/article/jso_0300-953x_1953_num_9_9_1773 Consulté le 17 août 2012
- ⁸ Rallu JL. Situation démographique de la Polynésie française. *Population*, 35e année, n°2, 1980 pp. 385-415. http://www.persee.fr/web/revues/home/prescript/article/pop_0032-4663_1980_num_35_2_18311 Consulté le 17 août 2012
- ⁹ Rumyantsev NS. Dating of first emergence of human epidemics. *J. Med. Med. Sci*. 2012; 3(6): 423-433,
- ¹⁰ Morris JG Jr. Cholera - modern pandemic disease of ancient lineage. *Emerg Infect Dis*. 2011 Nov;17(11):2099-104.
- ¹¹ Organisation Mondiale de la Santé. Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003 http://www.who.int/csr/sars/country/table2004_04_21/en/index.html. Consulté le 18 mai 2010
- ¹² Barboza P, Syndrome respiratoire aigu sévère (SRAS), *Flammarion 2004*; Traité de santé publique (Bourdillon F, Brückner G, Tabuteau D. Chapitre 42; (309-319).
- ¹³ Heymann DL. SARS and emerging infectious diseases: a challenge to place global solidarity above national sovereignty. *Ann Acad Med Singapore*. 2006; 35(5):350-3.
- ¹⁴ Grein TW, Kamara KB, Rodier G, Plant AJ, Bovier P, Ryan MJ, Ohyama T, Heymann DL. Rumors of disease in the global village: outbreak verification. *Emerg Infect Dis*. 2000; 6(2):97-102.
- ¹⁵ Chungong S. Global implementation of Surveillance Core Capacity under the IHR. EBS Surveillance Meeting, March 2013, Lyon, France.
- ¹⁶ Khalakdina, A. Event Management within the IHR Framework in the WHO European Region, EpiSouth Workshop, 4-7 Juin 2012, Istanbul, Turquie
- ¹⁷ Mondor, L., Brownstein, J. S., Chan, E., Madoff, L. C., Pollack, M. P., Buckeridge, D. L., & Brewer, T. F. Timeliness of Nongovernmental versus Governmental Global Outbreak Communications. *Emerg Infect Dis*. 2012; 18(7), 1184–1187.
- ¹⁸ Brownstein JS, Freifeld CC, Chan EH, Keller M, Sonricker AL, et al. Information technology and global surveillance of cases of 2009 H1N1 influenza. *N Engl J Med*. 2010; 362: 1731–1735.
- ¹⁹ Vaillant L, La Ruche G, Tarantola A, Barboza P. Epidemiology of fatal cases associated with pandemic H1N1 influenza 2009. *Euro Surveill*. 2009; 449: 1–6.
- ²⁰ Kman NE, Bachmann DJ (2012) Biosurveillance: a review and update. *Adv Prev Med*. 2012: 301408. Available: <http://www.hindawi.com/journals/apm/2012/301408/>. Consulté 5 May 2014.
- ²¹ Organisation Mondiale de la Santé (OMS). Early detection, assessment and response to acute public health events: Implementation of Early Warning and Response with a focus on Event-Based surveillance. (2014). WHO/HSE/GCR/LYO/2014.4 http://www.who.int/ihr/publications/WHO_HSE_GCR_LYO_2014.4/en/
- ²² Dente MG, Fabiani M, Gnesotto R, Putoto G, Montagna C, Pando CM De. EpiSouth: A network for communicable disease control in the mediterranean region and the balkans. *Euro Surveill*. 2009; 3–6.
- ²³ Ferguson NE, Steele L, Crawford CY, Huebner NL, Fonseca JC, et al. Bioterrorism web site resources for infectious disease clinicians and epidemiologists. *Clin Infect Dis*. 2003; 36: 1458–1473.

- ²⁴ Brownstein JS, Freifeld CC, Madoff LC. Digital Disease Detection—Harnessing the Web for Public Health Surveillance. *N. Engl. J. Med.* 2009;360(21):2153–7.
- ²⁵ Hartley DM, Nelson NP, Walters R, Arthur RR, Yangarber R, Madoff LC, et al. Landscape of international event-based biosurveillance. *Emerg. Health Threats J.* 2010 ;3:e3.
- ²⁶ Nelson NP, Yang L, Reilly AR, Hardin JE, Hartley DM. Event-based Internet biosurveillance: relation to epidemiological observation. *Emerg Themes Epidemiol.* 2012; 9: 4.
- ²⁷ Thomas CS, Nelson NP, Jahn GC, Niu T, Hartley DM. Use of media and public-domain Internet sources for detection and assessment of plant health threats. *Emerg Health Threats J.* 2011; 4: 1–15.
- ²⁸ Torii M, Yin L, Nguyen T, Mazumdar CT, Liu H, et al. An exploratory study of a text classification framework for Internet-based surveillance of emerging epidemics. *Int J Med Inform.* 2011; 80: 56–66
- ²⁹ Collier N, Doan S, Kawazoe A, Goodwin RM, Conway M, et al. BioCaster: detecting public health rumors with a Web-based text mining system. *Bioinformatics.* 2008; 24: 2940–2941.
- ³⁰ Collier N. What’s unusual in online disease outbreak news? *J Biomed Semantics.* 2010; 1: 2.
- ³¹ Collier N. Uncovering text mining: A survey of current work on web-based epidemic intelligence. *Glob Public Health.* 2012; 7: 731–749.
- ³² Chan EH, Brewer TF, Madoff LC, Pollack MP, Sonricker AL, et al. Global capacity for emerging infectious disease detection. *Proc Natl Acad Sci U S A.* 2010; 107: 21701–21706.
- ³³ Keller M, Blench M, Tolentino H, Freifeld CC, Mandl KD, et al. Use of unstructured event-based reports for global infectious disease surveillance. *Emerg Infect Dis.* 2009; 15: 689–695.
- ³⁴ Mawudeku A, Blench M. Global Public Health Intelligence Network (GPHIN). Eighth Conference of the Association for Machine Translation in the Americas, Hawai. 2008; p. 7–11. <http://www.mt-archive.info/AMTA-2008-TOC.htm>
- ³⁵ Brownstein JS, Freifeld CC. HealthMap: the development of automated real-time Internet surveillance for epidemic intelligence. *Euro Surveill.* 2007; 12(48):13–6.
- ³⁶ Brownstein JS, Freifeld CC, Reis BY, Mandl KD. Surveillance Sans Frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS Med.* 2008; 5(7):e151.
- ³⁷ Freifeld CC, Mandl KD, Reis BY, Brownstein JS. HealthMap: Global infectious disease monitoring through automated classification and visualization of internet media reports. *JAMIA.* 2006; 15(2):150–7.
- ³⁸ Linge JP, Steinberger R, Weber TP, Yangarber R, van der Gott E, et al. Internet surveillance systems for early alerting of health threats. *Euro Surveill.* 2009; 14.
- ³⁹ Rortais A, Belyaeva J, Gemo M, van der GE, Linge JP. MedISys: An early-warning system for the detection of (re-) emerging food and feed-borne hazards. *Food Research International.* 2012; 43: 1553–1556.
- ⁴⁰ Steinberger R, Fuart F, van der GE, Best C, von Etter P, Yangarber R. Text Mining from the Web for Medical Intelligence. In: Fogelman-Soulié F, Perrota D, Piskorski J, Steinberger R, editors. Mining Massive Data Sets for Security. 10.3233/978-2-58603-898-4-295 ed. IOS Press, Amsterdam, The Netherlands; 2008. p. 295-310
- ⁴¹ Madoff LC. ProMED-mail: an early warning system for emerging diseases. *Clin Infect Dis.* 2004; 39: 227–232.
- ⁴² Madoff LC, Woodall JP. The Internet and the global monitoring of emerging diseases: lessons from the first 10 years of ProMED-mail. *Arch Med Res.* 2005; 36: 724–730.
- ⁴³ Cowen P, Garland T, Hugh-Jones ME, Shimshony A, Handysides S, et al. Evaluation of ProMED-mail as an electronic early warning system for emerging animal diseases: 1996 to 2004. *J Am Vet Med Assoc.* 2006; 229: 1090–1099.
- ⁴⁴ Von Etter P, Huttunen S, Vihavainen A, Vuorinen M, Yangarber R. Assessment of Utility in Web Mining for the Domain of Public Health. *Proceedings of LOUHI-2010.* 2010;29–37.
- ⁴⁵ Grishman R, Huttunen S, Yangarber R Information extraction for enhanced access to disease outbreak reports. *J Biomed Inform.* 2002; 35: 236–246.
- ⁴⁶ Yangarber R, Best C, von Etter P, Fuart F, Horby D, et al. (2006) Combining Information about Epidemic Threats from Multiple Sources. Mining Massive Data Sets for Security, *NATO Advanced Study Institute.* 2007; 295–310.
- ⁴⁷ Tolentino H, Kamadjeu R, Fontelo P, Ms FL, Matters M, Pollack M, et al. Scanning the Emerging Infectious Diseases Horizon - Visualizing ProMED Emails Using EpiSPIDER. *Adv. Dis. Surveill.* 2007;2(169).
- ⁴⁸ Lyon a, Nunn M, Gossel G, Burgman M. Comparison of web-based biosecurity intelligence systems: BioCaster, EpiSPIDER and HealthMap. *Transbound. Emerg. Dis.* 2012 ;59(3):223–32.
- ⁴⁹ The Global Health Security Initiative (GHSI) Web site <http://www.ghsi.ca/english/index.asp>

- ⁵⁰ The Global Health Security Initiative (GHSI). 10 years of collaborative action <http://www.ghsi.ca/Documents/AnniversaryPublicationFinal.pdf>
- ⁵¹ Vaillant L, Barboza P : WP2: Evaluation of the current source systems EAR-GHSAG project (rapport non-publié)
- ⁵² Organisation Mondiale de la santé. Communicable disease surveillance and response systems; Guide to monitoring and evaluating. *WHO/CDS/EPR/LYO. 2006; Consulté 10/12/2011* http://whqlibdoc.who.int/hq/2006/WHO_CDS_EPR_LYO_2006_2_eng.pdf
- ⁵³ Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V; CDC Working Group. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. *MMWR Recomm Rep.* 2004 May 7; 53(RR-5):1-11.
- ⁵⁴ Fatima Aït-Belghiti, Nathalie El Omeiri and Philippe Barboza; On behalf of the WP6 Steering Team. WP 6 Epidemic Intelligence & Cross-Border in the Mediterranean Countries and Balkans. EpiSouth. consulté le 03/05/2013. http://www.episouth.org/outputs/wp6/3_EpiSouth%20Strategic%20document%20on%20Cross-Border%20Ep%20Int%20Rev%20luglio%202010.pdf
- ⁵⁵ EpiSouth, EpiSouth Weekly Epi Bulletin (e-WEB). The electronic Weekly Epidemiological Bulletin <http://www.episouthnetwork.org/content/episouth-weekly-epi-bulletin-e-web> consulté le 03/05/2013
- ⁵⁶ EpiSouth, EpiSouth Weekly Epi Bulletin (e-WEB). The electronic Weekly Epidemiological Bulletin <http://www.episouth.org/cgi-bin/searchbull?TEMP= 2>. Consulté le 15/02/2013
- ⁵⁷ Institut de Veille Sanitaire. Bulletin hebdomadaire international, <http://www.invs.sante.fr/fr/Publications-et-outils/Bulletin-hebdomadaire-international/Tous-les-numeros/> ; Consulté le 03/05/2013
- ⁵⁸ Epidemic alert and verification: summary report for 2005. *Wkly Epidemiol Rec* 2006;81(38):357-62. <http://www.who.int/wer/2006/wer8138.pdf>
- ⁵⁹ Organisation Mondiale de la Santé ; Epidemic intelligence - systematic event detection. WHO 2011 <http://www.who.int/csr/alertresponse/epidemicintelligence/en/index.html>. Consulté 11 /11/2011
- ⁶⁰ Commission Européenne. Règlement (CE) no 851/2004 du parlement européen et du conseil du 21 avril 2004 instituant un Centre européen de prévention et de contrôle des maladies ; <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:142:0001:0011:FR:PDF>
- ⁶¹ European Centre for Disease Prevention and Control; Role and mandate of ECDC in detecting emerging health threats ECDC. 2005; http://www.ecdc.europa.eu/en/activities/epidemicintelligence/Pages/EpidemicIntelligence_ThreatDetection.aspx
- ⁶² Coulombier D, Ciotti M, Freitas G, Frota A, Varela C, et al. Strengthening Europe's epidemic intelligence capacity: the first collaboration between a European Union Member State and the European Centre for Disease Prevention and Control. *Eurosurveillance* (2008) 13: 20–21.
- ⁶³ Aït-Belghiti F, El Omeiri N and Philippe Barboza, On behalf of the WP6 Steering Team; Strategic document; Epidemic Intelligence & Cross-Border in the Mediterranean Countries and Balkans. EpiSouth 2010; http://www.episouth.org/outputs/wp6/3_EpiSouth%20Strategic%20document%20on%20Cross-Border%20Ep%20Int%20Rev%20luglio%202010.pdf. consulté le 10/02/2011
- ⁶⁴ Rotureau B, Barboza P, Tarantola A, Paquet C. International epidemic intelligence at the Institut de Veille Sanitaire, France. *Emerg Infect Dis.* 2007 Oct 13(10):1590-2.
- ⁶⁵ LOI n° 2004-806 du 9 août 2004 relative à la politique de santé publique (1) Titre III : modernisation du système de veille, d'alerte et de gestion des situations d'urgence sanitaire ; Chapitre Ier : Veille et alerte ; Article 15 *Journal officiel de la République Française. JORF* n°185 du 11 août 2004
- ⁶⁶ Institut de Veille Sanitaire. Veille sanitaire internationale. <http://www.invs.sante.fr/Dossiers-thematiques/Veille-et-alerte/Veille-sanitaire-internationale> consulté le 15/02/2013.
- ⁶⁷ Organisation mondiale de la santé, Global Alert and Respons, Disease Outbreak News (DON) <http://www.who.int/csr/don/en/>. Consulté le 03/05/2013
- ⁶⁸ Organisation Mondiale de la santé, Bulletin Hebdomadaire de retro-information sur la méningite cérébrospinale <http://www.who.int/csr/disease/meningococcal/epidemiological/en/index.html#> . Consulté le 12/05/2013
- ⁶⁹ European Centre for Disease Prevention and Control, The ECDC communicable disease threat report (CDTR), http://ecdc.europa.eu/en/activities/epidemicintelligence/Pages/EpidemicIntelligence_ThreatCommunicationCDTR.aspx . Consulté le 03/05/2013
- ⁷⁰ Institut de Veille Sanitaire, Bulletin hebdomadaire international (BHI), <http://www.invs.sante.fr/Publications-et-outils/Bulletin-hebdomadaire-international>. Consulté le 03/05/2013

- ⁷¹ Barboza P, Quatresous I. Globalization, emergence and importation of infectious diseases. *Rev Prat.* 2007; 30;57(8):867-3.
- ⁷² EpiSouth Website. West Nile virus circulation in the EpiSouth countries and neighbouring areas; Seasons 2010, 2011 and 2012, update 20th June 2013. Consulté 30/06/2013
http://www.episouthnetwork.org/sites/default/files/outputs/note_west_nile_episouth_2010_2011_2012_june2013.pdf
- ⁷³ Tarantola A; Ait Belghiti F; Barboza P; Baudon C; Cohuet S; Degail MA; Dejour Salamanca D; Gauthier V; Gueguen J; La Ruche G; Rachas A; Vaillant L; Gastellu Etchegorry M. Grippe A(H1N1)v : naissance d'une pandémie; Bilan mondial au 5 juin 2009. *BEHWeb [Internet]*. 2009; 1:1–5. Consulté le 18/08/2013
http://www.invs.sante.fr/beh/actualite_epidemiologique/a-9.html
- ⁷⁴ Département International InVS, Cas humains de grippe porcine A(H1N1), USA, 25 avril 2009
http://www.invs.sante.fr/international/notes/cas_humains_grippe_porcine_250409.pdf. Consulté le 03/05/2013
- ⁷⁵ Organisation Mondiale de la Santé. Influenza-like illness in the United States and Mexico,
http://www.who.int/csr/don/2009_04_24/en/index.html. Consulté le 18/08/2013.
- ⁷⁶ Organisation Mondiale de la Santé, Swine influenza - Statement by WHO Director-General, Dr Margaret Chan. 2009. http://www.who.int/csr/don/2009_04_25/en/index.html#. Consulté le 03/05/2013
- ⁷⁷ Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Kerkhove D Van, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science.* 2009 ;19 ;324(5934) :1557–61.
- ⁷⁸ Tenth Ministerial Meeting Global Health Security Initiative (GHSI) ; Ministerial Statements, London, United Kingdom - 4 December 2009. <http://www.ghsi.ca/english/statementlondon2009.asp>
- ⁷⁹ Vaillant L, Barboza P, Arthur RR. Epidemic Intelligence: Assessing event-based tools and users' perception in the GHSAG community. Présentation au congrès IMED, Vienne, Autriche. 2011.
http://www.isid.org/events/archives/IMED2011/Downloads/IMED2011_Presentations/IMED2011_Vaillant.pdf
- ⁸⁰ Twelfth Ministerial Meeting Global Health Security initiative (GHSI) ; Ministerial Statements. Paris, France – 9 December, 2011. <http://www.ghsi.ca/english/statementparis2011.asp>
- ⁸¹ Eleventh Ministerial Meeting of The Global Health Security Initiative (GHSI); Ministerial Statements. Mexico City, Mexico - 2-3 December 2010. <http://www.ghsi.ca/english/statementmexicocity2010.asp>
- ⁸² Thirteenth Ministerial Meeting Global Health Security Initiative (GHSI); Ministerial Statements. Berlin, Germany – 7 December, 2012. <http://www.ghsi.ca/english/statementgermany2012.asp>
- ⁸³ Fourteenth Ministerial Meeting Global Health Security Initiative (GHSI) ; Ministerial Statements Rome, Italy – 13 December, 2013. <http://www.ghsi.ca/english/statementitaly2013.asp>
- ⁸⁴ Brownstein JS, Freifeld CC, Reis BY, Mandl K. HealthMap: Internet-based emerging infectious disease intelligence. *Global Infectious Disease Surveillance and Detection: Assessing the Challenges—Finding Solutions, Workshop Summary.* Washington: *The National Academies Press*, 2007. p. 122-36.
- ⁸⁵ Woodall JP. Global surveillance of emerging diseases: the ProMED-mail perspective. *Cad Saude Publica* 2001;17 Suppl:147-54.
- ⁸⁶ Zhang Z, Chen D, Chen Y, Liu W, Wang L, et al. (2010) Spatio-temporal data comparisons for global highly pathogenic avian influenza (HPAI) H5N1 outbreaks. *PLoS One* ; 5: e15314..
- ⁸⁷ Zeldenrust ME, Postma MJ, Vliet JA Van. The value of ProMED-mail for the Early Warning Committee in the Netherlands: more specific approach recommended. *Eurosurveillance* ;2007;13(6):1–4. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=8033>
- ⁸⁸ Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature.* 2009 Feb 19 [cited 2013 Dec 11];457(7232):1012–4.
<http://www.netmarketshare.com/report.aspx?qprid=88&sample=51&qpmr=1000&qpct=2#>
- ⁹⁰ MMWR. Outbreak of Rift Valley fever--Saudi Arabia, August–October, 2000. *Morb. Mortal. Wkly. Rep.* 2000 Oct 13;49(40):905–8.
- ⁹¹ Reporters Sans Frontières. Classement mondial 2013 Liberté de la Presse, consulté le 10/10/2013
<http://fr.rsf.org/press-freedom-index-2013,1054.html>
- ⁹² Decision no 1082/2013/eu of the European Parliament and of the Council of 22 October 2013 on serious cross-border threats to health and repealing Decision No 2119/98/EC. Consulté le 24/02/2014
http://ec.europa.eu/health/preparedness_response/docs/decision_serious_crossborder_threats_22102013_en.pdf
- ⁹³ Nelson NP, Brownstein JS, Hartley DM. Event-based biosurveillance of respiratory disease in Mexico, 2007–2009: connection to the 2009 influenza A(H1N1) pandemic?. *Euro Surveill.* 2010;15(30):pii=19626
<http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19626>

- ⁹⁴ Mawudeku A, Lemay R, Werker D, Andraghetti R, St. John R. The Global Public Health Intelligence Network. In: M'ikanatha NM, Lynfield R, Van Beneden CA, de Valk H, editors. *Infectious disease surveillance*, 1st ed. Lynn (MA): *Blackwell Publishing*; 2007
- ⁹⁵ Mykhalovskiy E, Weir L. The Global Public Health Intelligence Network and early warning outbreak detection: a Canadian contribution to global public health. *Can J Public Health* ;2006;97(1):42-4.
- ⁹⁶ Mawudeku A, Blench M. Global Public Health Intelligence Network (GPHIN). In: Proceedings of the 7th Conference of the Association for Machine Translation in the Americas 2006. Consulté le 24/03/2009 <http://www.mt-archive.info/MTS-2005-Mawudeku.pdf>
- ⁹⁷ <http://emm.newsbrief.eu/overview.html#medisys> consulté le 17 avril 2012
- ⁹⁸ Woodall J. Official versus unofficial outbreak reporting through the Internet. *Int J Med Inform.* 1997;47(1-2):31-4.
- ⁹⁹ Ministère des affaires étrangères http://www.diplomatie.gouv.fr/fr/les-francais-etranger_1296/les-francais-etranger_5294/les-francais-etablis-hors-france_4182/population-francaise-inscrite-au-31-decembre-2010_89269.html consulté le 29/04/2013
- ¹⁰⁰ Ministère des affaires étrangères Français inscrits au registre mondial au 31/12/2011, http://www.diplomatie.gouv.fr/fr/IMG/pdf/Inscrits_zone_pays_2011_cle896b45-1.pdf; consulté le 29/04/2013
- ¹⁰¹ Institut National de la Statistique et des Etudes Economiques (INSEE). Répartition des étrangers par pays de naissance <http://www.insee.fr/fr/ffc/figure/etrangersnat.xls> consulté le 29/04/2013