

Le défi : la détection des maladies et le déluge de « mégadonnées »

Traditionnellement, le signalement et le suivi rapides des maladies infectieuses émergentes étaient mis en cause par des questions relatives au manque de données et à l'accès aux données, ainsi que par le comportement des gouvernements dans le cadre de la déclaration de flambées épidémiques de maladies. Par exemple, en 1967–68, lors d'une importante flambée épidémique de poliovirus sauvage, l'Organisation mondiale de la santé (OMS) n'avait pas été en mesure d'obtenir des données épidémiologiques et de surveillance exactes de la part du gouvernement de la part du gouvernement de la Guinée. Malgré des rapports faisant état de plus de 200 cas de polio paralysante survenus dans le pays, les efforts déployés par l'OMS pour obtenir des informations

supplémentaires de la part des autorités sanitaires guinéennes au cours de l'épidémie se sont révélés infructueux en raison des retards bureaucratiques et des préoccupations en matière de souveraineté nationale. Il a finalement fallu sept mois à une équipe de consultants de l'OMS pour obtenir l'autorisation d'entrer en Guinée afin d'évaluer l'urgence.

Près de quatre décennies

après l'épidémie de polio survenue en Guinée, l'essor de l'ère des mégadonnées et les diffusions de sources de données numériques hétérogènes ont contribué à transformer la pratique et la portée de la surveillance de la santé publique à l'ère de la sécurité sanitaire mondiale. La première transition vers l'accès et l'intégration de sources de données ouvertes et non officielles dans la surveillance de menaces pandémiques émergentes est apparue lors de la flambée épidémique de SRAS survenue en 2002-03 dans la province de Guangdong, en Chine.

Le programme : Technologies de mégadonnées pour le suivi des risques de pandémie

Contrairement aux précédentes flambées épidémiques, l'apparition de l'épidémie de SRAS est survenue parallèlement à l'essor de nouvelles sources de données

numériques et d'Internet. Par opposition aux urgences antérieures en matière de santé publique, le premier signalement initial d'un cas inhabituel de pneumonie atypique circulant dans la province de Guangdong ne provenait pas de l'OMS ni des autorités sanitaires chinoises, mais plutôt d'une nouvelle technologie de surveillance numérique, le Réseau mondial d'information en santé publique (RMISP), un prototype lancé par le gouvernement du Canada quelques années auparavant. En tant que nouvelle technologie de surveillance en ligne, le RMISP cherchait à surveiller le contenu numérique en ligne afin d'identifier des indicateurs potentiels de menaces de maladies. En utilisant un logiciel d'indexation web de pointe et des algorithmes numériques pour collecter, trier, et classer les ensembles de données entrants, le RMISP a enregistré les premières alertes d'une urgence sanitaire de santé publique en Chine, plusieurs mois avant les autorités de l'OMS, et malgré les efforts mis en œuvre par le gouvernement chinois pour contrôler étroitement la divulgation de renseignements sur un événement sanitaire inhabituel.

L'utilisation du RMISP pour surveiller les potentielles épidémies au moyen de sources de mégadonnées a marqué des transformations critiques dans l'intégration de données probantes non spécialisées dans la surveillance des urgences de santé publique. Il s'agissait de la première démonstration de la façon dont les ensembles de données numériques non officielles et non médicales saisis en ligne pouvaient fournir des avertissements en temps opportun de flambées épidémiques potentielles à des stades pré-épidémiques. Ces données pouvaient être vérifiées parallèlement à des preuves médicales et scientifiques, et mises à jour, téléchargées et divulguées en temps réel, permettant aux autorités sanitaires nationales et aux agents de santé intergouvernementaux de disposer de capacités d'intervention plus souples. Le développement et l'utilisation de nouvelles techniques de surveillance en matière de santé publique, qui intégraient des sources de données probantes non spécialisées et de mégadonnées et opéraient à travers des paysages de données et Internet, ont signifié que les conventions établies en matière de signalement de maladies à l'ère des mégadonnées n'étaient plus contrôlées bilatéralement par l'OMS et ses états membres. Suite au succès largement cité du RMISP dans la détection de flambées épidémiques à partir de nouvelles sources de données, une série de systèmes de surveillance sanitaire fondés sur des mégadonnées sont entrés en vigueur à partir de 2003.

« Au cours des deux dernières décennies, les nouvelles pratiques en matière de surveillance de maladie fondées sur l'accumulation et le traitement de grands ensembles de données... ont permis de prévoir avec succès les flambées épidémiques de maladies mortelles. »

En 2006, une nouvelle technologie de surveillance des mégadonnées appelée HealthMap a été lancée en ligne. Automatisée et alimentée par des algorithmes pour ses opérations de collecte des données, de filtrage et d'évaluation, le système HealthMap a continué à développer de nouvelles techniques de surveillance de maladies à source ouverte, en termes HealthMap, des « sources de données disparates, y compris des agrégateurs d'informations en ligne, des témoignages, des débats organisés par des experts et des rapports officiels validés, afin d'obtenir une vue unifiée et globale de l'état actuel des maladies infectieuses à l'échelle mondiale et de leurs conséquences sur la santé humaine et animale » (HealthMap 2020).

Dix ans après la flambée épidémique de SRAS en Chine, le 14 mars 2014 HealthMap a téléchargé et diffusé une alerte sanitaire via son site web qui indiquait

l'émergence d'une étrange fièvre hémorragique apparue à Macenta, en Guinée. La prévision précoce de la crise sanitaire avait été détectée en ligne à partir d'un reportage guinéen par un algorithme contenu dans le système d'exploitation de HealthMap. Neuf jours plus tard, le ministère de la Santé et de l'Hygiène publique de Guinée a transmis à l'OMS la notification officielle d'une

flambée épidémique de virus Ébola en évolution rapide au sein des régions fortement boisées du pays. Peu après, un article publié dans la revue *Lancet Infectious Disease* affirmait que la priorité en matière de santé publique à l'échelle mondiale devrait être d'encourager les pays à adopter, intégrer et appliquer ces pratiques évolutives de surveillance numériques afin de suivre et de réguler le mouvement des risques de pandémie mondiale.

Au cours des cinq dernières années, les approches évolutives visant à exploiter diverses sources de mégadonnées numériques pour accélérer la déclaration des maladies infectieuses se sont intensifiées dans les pratiques de sécurité sanitaire. Lors de la flambée épidémique localisée de MERS-CoV survenue en 2016 en Corée du Sud, le gouvernement a utilisé des processus d'apprentissage automatique basés sur des algorithmes pour identifier des cas

présusés d'infections à partir d'enregistrements de téléphones portables fournis par les opérateurs de télécommunications, et pour placer par la suite 17 000 citoyens en quarantaine. Ailleurs, dans le cadre de la pandémie mondiale de COVID-19 en cours survenue à Wuhan en 2019-20, l'immense intérêt mondial pour la recherche a culminé dans la recherche visant à appliquer des techniques d'intelligence artificielle (IA) – y compris l'application de l'apprentissage automatique aux flux de médias sociaux, aux contenus Internet et à une série d'ensembles de données provenant de sources librement accessibles – afin de chercher à dépister et à identifier de nouveaux cas du virus. Au cours des premiers jours critiques de l'émergence de la COVID-19, une start-up basée à Toronto appelée BlueDot, qui fournit une surveillance personnalisée et automatisée des maladies, a affirmé avoir détecté les premiers cas initiaux de COVID-19 par le biais de l'analyse quotidienne automatisée de la société de plus de 100 000 articles de presse dans 65 langues. Le 31 décembre 2019, BlueDot a identifié l'émergence d'une « pneumonie de cause connue » en Chine qui avait déjà entraîné 27 infections. Grâce à des processus d'IA automatisés appliqués à des données provenant de sources librement accessibles en ligne, BlueDot a identifié l'émergence de la COVID-19 environ une semaine avant l'OMS et les Centres américains pour le contrôle et la prévention des maladies (CDC).

L'impact : enseignements tirés de l'intégration de données probantes non spécialisées parmi les pratiques contre les maladies infectieuses

L'intégration de plus en plus fréquente des données probantes non spécialisées mentionnées ci-dessus, principalement sous la forme de sources de mégadonnées, a généré de nombreux avantages pour l'amélioration des pratiques de surveillance de la santé publique dans un contexte de surveillance de pandémie mondiale, y compris : une identification plus rapide d'événements sanitaires inhabituels ; une diversité des sources de données permettant de suivre en temps réel l'émergence des risques pour la santé publique ; et un échange plus rapide de données et d'information stratégiques pour faciliter les ripostes contre les flambées épidémiques.

Toutefois, tandis que la numérisation et les progrès technologiques ont été au cœur de la production d'un plus grand nombre de données et de rapports plus rapides, les pratiques en matière de surveillance

« Tandis que les vagues de numérisation et de datafication dans le cadre de la surveillance d'urgences de santé publique continuent de progresser à un rythme rapide, il est impératif que les interrogations éthiques, juridiques et politiques les concernant... en fassent de même. »

numérique ne sont pas une panacée pour traiter et régler les urgences de santé publique de plus en plus répandues. Ces pratiques en matière de surveillance numérique doivent continuer à opérer en association avec des systèmes de santé solides et des investissements dans le renforcement des systèmes et les capacités de riposte. Ceci est particulièrement important en ce qui concerne l'intégration et l'interopérabilité de ces systèmes dans des systèmes de

santé sentinelles mondiales plus vastes. Plus important encore, alors que les vagues de numérisation et de datafication dans le cadre de la surveillance des urgences de santé publique continuent de progresser à un rythme rapide, il est impératif que les interrogations éthiques, juridiques et politiques inhérentes à ces nouvelles sources de données et de pratiques de production de connaissances continuent d'en faire de même dans le cadre du suivi de la prochaine flambée épidémique.

Lectures complémentaires

Bansal S., Chowell G., Simonsen L., Vespigiani A. and Viboud C. (2016) '[Big Data for Infectious Disease Surveillance and Modeling](#)', *The Journal of Infectious Diseases*, 214.Suppl. 4: S375–S379

Roberts S.L. (2019) '[Big Data, Algorithmic Governmentality and the Regulation of Pandemic Risk](#)', *European Journal of Risk Regulation*, 10.1: 94-115

Roberts S.L. and Elbe S. (2017) '[Catching the flu: Syndromic surveillance, algorithmic governmentality and global health security](#)', *Security Dialogue*, 48.1: 46-62

Vayena E., Salathé M., Madoff L.C. and Brownstein J.S. (2015) '[Ethical challenges of Big Data in Public Health](#)', *PLOS Computational Biology*, 11.2 :1-7

Weir L. and Mykhalovskiy E. (2010) *Global public health vigilance: Creating a world on alert*, New York: Routledge

Nous connaître

L'objectif de la plateforme Social Science in Humanitarian Action (SSHAP) est d'établir des réseaux de spécialistes en sciences sociales possédant une expertise régionale et les connaissances nécessaires pour fournir rapidement des renseignements, des analyses et des conseils adaptés à la demande et sous des formes accessibles, afin de mieux concevoir et mettre en œuvre des interventions d'urgence. La SSHAP est un partenariat entre l'Institute of Development Studies (IDS), la London School of Hygiene and Tropical Medicine (LSHTM), Anthrologica et le programme Communication for Development (C4D) de l'UNICEF.



Social Science in Humanitarian Action

E info@socialscienceinaction.org W www.socialscienceinaction.org



Cette étude de cas fait partie d'un portefeuille inhérent aux enseignements tirés des épidémies en matière de sciences sociales et a reçu le soutien du bureau des États-Unis pour l'assistance à l'étranger en cas de catastrophes naturelles (OFDA) et de l'UNICEF.

Mentions

Cette étude de cas SSHAP a été élaborée par le **Dr Stephen L. Roberts** du Department of Health Policy de la London School of Economics (LSE).

Référence : Roberts, S.L. (2020) *Intégration de données probantes non spécialisées dans la surveillance et la détection précoce d'urgences de santé publique*, Étude de cas SSHAP Numéro 2, UNICEF, IDS et Anthrologica.

Ce document a été financé par l'UNICEF et le bureau de l'agence des États-Unis pour le développement international (USAID) du bureau des États-Unis pour l'assistance à l'étranger en cas de catastrophes naturelles (OFDA). Toutefois, les opinions exprimées ci-après sont celles de l'(des) auteur(s) et ne reflètent pas nécessairement celles de l'USAID, de l'OFDA, ou des partenaires de la plateforme Social Science for Humanitarian Action (SSHAP).



Ceci est un document en libre accès distribué selon les modalités de la version 4.0 de la licence internationale Creative Commons Attribution (CC BY), qui autorise l'utilisation, la distribution et la reproduction sans restriction sur tout support, à condition que les auteurs d'origine et la source soient crédités et que toute modification ou adaptation soit indiquée. <https://creativecommons.org/licenses/by/4.0/legalcode.fr>.

© SSHAP 2020