

Large-Scale Evaluation of Informal Online Reporting for Outbreak Detection

Emily H. Chan, MSc, Mikaela Keller, PhD, Amy L. Sonricker, MPH,
Clark C. Freifeld, BS, John S. Brownstein, PhD

Children's Hospital Informatics Program, Children's Hospital Boston, Boston, United States

OBJECTIVE

We conducted a large-scale evaluation of informal Internet-based reporting of outbreaks across a variety of diseases and geographic locations, through a comparison to official World Health Organization (WHO) reports.

BACKGROUND

While disease surveillance traditionally relies on a structured public health infrastructure, new approaches take advantage of unofficial sources of information (1). The idea is that, for economic, political and public health reasons, local observers may report outbreaks earlier than local public health authorities. Two such information systems that monitor informal disease reporting are ProMED-mail and HealthMap. ProMED, an expert-monitored global electronic mailing list established in 1994, collects and summarizes information from local media, regional observers and official reports (2). Launched in 2006, HealthMap is a largely automated system that collects disease outbreak reports on the Internet, filters them using natural language processing tools by their geographic location and infectious agent and displays them on a publicly available map at <http://healthmap.org> (3). It has been shown that information collected by such systems provided potentially early awareness of historic outbreaks (eg SARS in China in late 2002 (4)). However, it is unknown how timely and sensitive these datasets are across a wide spectrum of diseases and locations.

METHODS

We first assessed the sensitivity and scope of these datasets with a descriptive analysis of ProMED reports. Next, analyzing the WHO's Global Alert and Response (GAR) reports between 1996 and 2008, we selected human outbreaks of infectious, non-food-borne diseases that were not considered seasonal or endemic to the region, and were not isolated, imported cases, then extracted the corresponding ProMED and HealthMap reports. We subsequently created timelines of the progression and reporting for each outbreak, compared the timing in reporting by official and informal sources, and attempted to identify factors that may contribute to differences in the timing of reporting.

RESULTS

For 355 WHO-confirmed outbreaks, ProMED reported on average 18 days (95% C.I.: 12.2-23.8)

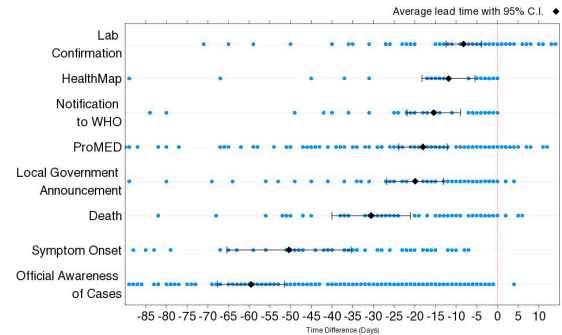


Figure 1 – Timeline of time difference between official WHO reports, informal reports and various “outbreak milestones”. The line at 0 days represents no lead/lag over the WHO report.

earlier than WHO's Global Alert and Response (GAR) reports, while HealthMap reported 12 days (5.4-18.2) earlier (Fig. 1). A further analysis revealed country and disease-dependent differences in reporting. Sensitivity was 0.946 (0.923-0.970) for ProMED (n=355), and 1.000 (1-1) for HealthMap (n=39).

CONCLUSIONS

This preliminary work shows that informal online disease reporting can facilitate both sensitive and timely detection of disease outbreaks. An examination of the finer-grained differences in reporting depending on the disease and location reveals the most informatively valuable areas in which efforts for monitoring the vast amount of informal online reports should be targeted. Early and accurate recognition of outbreaks is crucial for expediting the initiation of appropriate interventions. Evaluation of specificity and signal to noise is an important next research goal.

REFERENCES

1. Keller M, Blench M, Tolentino H, Freifeld CC, Mandl KD, Mawudeku A, et al. Use of unstructured event-based reports for global infectious disease surveillance. *Emerg Infect Dis*. 2009 May;15(5):689-95.
2. 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 Nov-Dec;36(6):724-30.
3. Brownstein JS, Freifeld CC, Reis BY, Mandl KD. Surveillance Sans Frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS Med*. 2008 Jul 8;5(7):e151.
4. Mawudeku A, Lemay R, Werker D, Andraghetti R, St. John R. The Global Public Health Intelligence Network. In: M'ikanatha NM LR, Van Beneden CA, de Valk H, editor. *Infectious disease Surveillance*. Lynn, MA: Blackwell Publishing; 2007. p. 304-17.

Further Information:

Emily Chan, emily.chan2@childrens.harvard.edu