ID News: Big Data and predicting, preventing and controlling outbreaks

Christaki E. **New technologies in predicting, preventing and controlling emerging infectious diseases.** Virulence. 2015 Jun 11:1-8. (*Summary*)

Surveillance of emerging infectious diseases is vital for the early identification of public health threats. Emergence of novel infections is linked to human factors such as population density, travel and trade and ecological factors like climate change and agricultural practices. A wealth of new technologies is becoming increasingly available for the rapid molecular identification of pathogens but also for the more accurate monitoring of infectious disease activity. Web-based surveillance tools and epidemic intelligence methods, used by all major public health institutions, are intended to facilitate risk assessment and timely outbreak detection. This review presents new methods for regional and global infectious disease surveillance and advances in epidemic modeling aimed to predict and prevent future infectious diseases threats.

Semenza JC. **Prototype early warning systems for vector-borne diseases in Europe.** Int J Environ Res Public Health. 2015 Jun 2;12(6):6333-51. doi:10.3390/ijerph120606333. (*Summary*)

Globalization and environmental change, social and demographic determinants and health system capacity are significant drivers of infectious diseases which can also act as epidemic precursors. Thus, monitoring changes in these drivers can help anticipate, or even forecast, an upsurge of infectious diseases. The European Environment and Epidemiology (E3) Network has been built for this purpose and applied to three early warning case studies: 1) The environmental suitability of malaria transmission in Greece was mapped in order to target epidemiological and entomological surveillance and vector control activities. Malaria transmission in these areas was interrupted in 2013 through such integrated preparedness and response activities. 2) Since 2010, recurrent West Nile fever outbreaks have ensued in South/eastern Europe. Temperature deviations from a thirty year average proved to be associated with the 2010 outbreak. Drivers of subsequent outbreaks were computed through multivariate logistic regression models and included monthly temperature anomalies for July and a normalized water index. 3) Dengue is a tropical disease but sustained transmission has recently emerged in Madeira. Autochthonous transmission has also occurred repeatedly in France and in Croatia mainly due to travel importation. The risk of dengue importation into Europe in 2010 was computed with the volume of international travelers from dengue-affected areas worldwide.These prototype early warning systems indicate that monitoring drivers of infectious diseases can help predict vector-borne disease threats.

Hay SI, George DB, Moyes CL, Brownstein JS. **Big Data opportunities for global infectious disease surveillance.** PLoS Med 2013;10(4):e1001413. doi:10.1371/journal.pmed.1001413. (*Summary*)

Systems to provide static spatially continuous maps of infectious disease risk and continually updated reports of infectious disease occurrence exist but to date the two have never been combined. Novel online data sources, such as social media, combined with epidemiologically relevant environmental information are valuable new data sources that can assist the “real-time” updating of spatial maps. Advances in machine learning and the use of crowd sourcing open up the possibility of developing a continually updated atlas of infectious diseases. Freely-available dynamic infectious disease risk maps would be valuable to a wide range of health professionals from policy-makers prioritizing limited resources to individual clinicians.

ID News: Big Data and ethics

Ploug T, Holm S. **Meta consent: A flexible and autonomous way of obtaining informed consent for secondary research.** BMJ. 2015 May 7;350:h2146. doi:10.1136/bmj.h2146. (*Summary*)

A rapidly increasing capability for storing, linking and analyzing health data has led to new opportunities for research. However, it also raises new ethical and regulatory concerns. Central among these is the question of the conditions under which secondary research can use data that were collected as part of routine healthcare practice or for a specific research project. Does secondary use require renewed informed consent from the original participants? Consent to date has included: dynamic (when information about specific secondary use of health data or tissue is requested each time to each individual through a web-based platform), broad (when consent is given to future research of a particular type in addition to the current specific research project) or blanket (data could be used without further consent). We propose meta consent which means individuals can choose how they wish to provide consent for future secondary research of data collected in the past or of data that will be stored in the future, thus meta consent is both retrospective and prospective. Meta consent is a truly individual consent procedure that takes into account the differences in personal interests and levels of trust in researchers among the population. The risk of routinisation is reduced because individuals can limit the requests they receive to only those categories of research that really matter to them. Its implementation online makes meta consent easy to revoke or change.

Mittelstadt BD, Floridi L. **The ethics of Big Data: Current and foreseeable issues in biomedical contexts.** Sci Eng Ethics. 2015 May 23. [Epub ahead of print]. (*Summary*)

The capacity to collect and analyze data is growing exponentially. Referred to as 'Big Data', this scientific, social and technological trend has helped create destabilising amounts of information, which can challenge accepted social and ethical norms. As is often the case with the cutting edge of scientific and technological progress, understanding of the ethical implications of Big Data lags behind. By means of a meta-analysis of the literature, a thematic narrative is provided to guide ethicists, data scientists, regulators and other stakeholders through what is already known or hypothesised about the ethical risks of this emerging and innovative phenomenon. Five key areas of concern are identified: 1) informed consent, 2) privacy (including anonymisation and data protection), 3) ownership, 4) epistemology and objectivity and 5) 'Big Data Divides' created between those who have or lack the necessary resources to analyze increasingly large datasets. Six additional areas of concern are then suggested which, although related have not yet attracted extensive debate in the existing literature: 6) the dangers of ignoring group-level ethical harms; 7) the importance of epistemology in assessing the ethics of Big Data; 8) the changing nature of fiduciary relationships that become increasingly data saturated; 9) the need to distinguish between 'academic' and 'commercial' Big Data practices in terms of potential harm to data subjects; 10) future problems with ownership of intellectual property generated from analysis of aggregated datasets; and 11) the difficulty of providing meaningful access rights to individual data subjects that lack necessary resources. Considered together, these eleven themes provide a thorough critical framework to guide ethical assessment and governance of emerging Big Data practices.