

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


Enhancing Situational Awareness to Prevent Infectious Disease Outbreaks from Becoming Catastrophic

Citation

Lipsitch M., Santillana M. (2019) Enhancing Situational Awareness to Prevent Infectious Disease Outbreaks from Becoming Catastrophic. In: . Current Topics in Microbiology and Immunology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/82_2019_172

Published version

https://doi.org/10.1007/82_2019_172

Link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:41011214

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles (OAP), as set forth at

https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

Enhancing Situational Awareness to Prevent Infectious Disease Outbreaks from Becoming Catastrophic

Marc Lipsitch (1) and Mauricio Santillana (2)

(1) **Center for Communicable Disease Dynamics, Department of Epidemiology and Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02130 USA**
mlipsitc@hsph.harvard.edu
(2) **Computational Health Informatics Program, Boston Children's Hospital, 1 Autumn st, Boston, MA 02215 USA**
Department of Pediatrics, Harvard Medical School, 25 Shattuck St, Boston, MA 02115 USA
msantill@g.harvard.edu

Abstract

Catastrophic epidemics, if they occur, will very likely start from localized and far smaller (noncatastrophic) outbreaks that grow into much greater threats. One key bulwark against this outcome is the ability of governments and the health sector more generally to make informed decisions about control measures based on accurate understanding of the current and future extent of the outbreak. Situation reporting is the activity of periodically summarizing the state of the outbreak in a (usually) public way. We delineate key classes of decisions whose quality depends on high-quality situation reporting, key quantities for which estimates are needed to inform these decisions, and the traditional and novel sources of data that can aid in estimating these quantities. We emphasize the important role of situation reports as providing public, shared planning assumptions that allow decision makers to harmonize the response while making explicit the uncertainties that underlie the scenarios outlined for planning. In this era of multiple data sources and complex factors informing the interpretation of these data sources, we describe four principles for situation reporting:1. Situation reporting should be thematic, concentrating on essential areas of evidence needed for decisions. 2. Situation reports should adduce evidence from multiple sources to address each area of evidence, along with expert assessments of key parameters. 3. Situation reports should acknowledge uncertainty and attempt to estimate its magnitude for each assessment.4. Situation reports should contain carefully curated visualizations along with text and tables.

1. Introduction

Short of a massive, distributed bioterrorist attack, nearly any imaginable scenario for a globally catastrophic infectious disease outbreak would involve an initially small outbreak spreading from a few infections in a limited geographic area to infect many more people in many more places. It follows that efforts to prevent such catastrophic scenarios from

materializing must include successful measures to stop or limit the spread of severe, but initially subcatastrophic events (Lipsitch 2017). Such measures are more likely to succeed if key decision makers -- and those charged with implementing their decisions -- have access to reliable, timely information on key parameters of an outbreak in progress.

It is a characteristic of infectious disease outbreaks that information available at the early stages is incomplete, uncertain, and often biased in the sense that observations are initially made on unrepresentative samples of the population (for example those reporting to hospitals) that are easily observed, and only later on more representative populations(Lipsitch et al. 2009b, 2011). Knowledgeable public health professionals have a wealth of heuristics for *filtering* and *integrating* data to form early assessments of key quantities that are inputs to decisions – for example, current incidence and prevalence; forecasts of incidence and prevalence; geographic and demographic extent; and severity measures (Fig. 1, "Evidence"). For example, experts in influenza epidemiology know that viral testing and influenza-like illnesses (ILI) are both incomplete measures of incidence with particular biases that vary over time. They have many, often unspoken *filtering* heuristics about how to infer "true" estimates of incidence (absolute or relative) from each type of system. Likewise, they have related sets of heuristics for *integrating* this information to account for biases of individual systems and to assess consistency of different indicators.

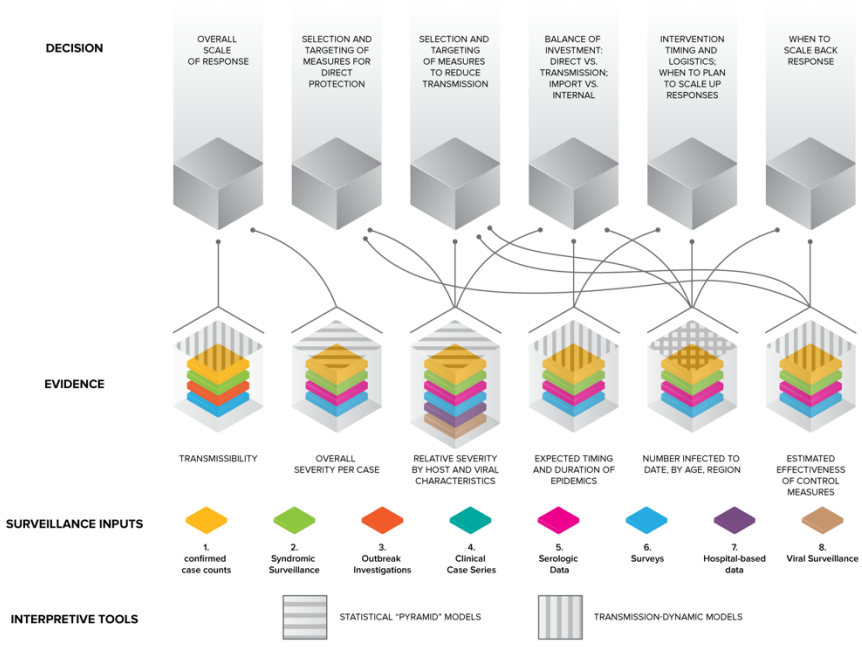


Figure 1*: Key decisions on pandemic response and the evidence base on which they ideally rest; this evidence base is built up from surveillance inputs using interpretive tools such as transmission-dynamic models and "pyramid" severity models. Image adapted from Lipsitch et al. (2011) by Lucia Ricci.*

In contrast to subject-matter experts, senior decision makers are typically generalists with less detailed knowledge of these aspects of any new disease. Their heuristics for interpreting raw data from surveillance systems, epidemiologic investigations, and novel data sources will be less nuanced, less informed by experience, and more variable from person to person. This may cause them to reach faulty conclusions about the magnitude of the threat, the options for and likely effects of potential responses, and the level of certainty surrounding each of these. We suggest in this chapter that information presented to decision makers, commonly known as Situation Reporting, should be tailored to give them not only raw data, but also synthetic expert judgment on the key characteristics of the outbreak and associated uncertainties. Such situation reporting would use a mixture of text and carefully chosen graphical presentations to convey expert estimates of key quantities and levels of certainty for each. In very data-poor settings (such as the early days of an epidemic of a novel disease) these syntheses would reflect expert judgment in interpreting the data that exist. For more familiar diseases, or as an epidemic of a new disease progresses, the presentation would include, in addition to expert interpretation, formal syntheses of emerging data using methods from statistics and machine learning. In all cases, modern situation reporting would incorporate not only traditional public health sources of data (mainly gathered from health systems), but also novel Internet-based data streams that can enhance the context, geographic extent, accuracy and timeliness of traditional sources. Carefully designed visualizations should be used to display spatial and temporal evolution of the current epidemic outbreak, geographic risk predictions, and other high-dimensional information. They may even include historical reconstructions of the spatio-temporal dynamics of previously observed outbreaks that may help contextualize the gravity of the ongoing public health threat.

In this chapter we begin by describing the decisions that rely on good situation reporting and the major topics on which assessments are crucial to good decisions in nearly all outbreaks. Next, we review traditional and novel Internet-based sources of data that can inform these assessments. Motivated by these uses and the available data, we then propose and discuss four criteria for high-quality situation reporting in outbreaks.

2. Decisions that rely on situation reporting

Among the many decisions facing policy makers throughout the course of an infectious disease outbreak (Lipsitch et al. 2011), arguably the most important fall into two broad categories: first, how big should the overall response be at each place and time, and second, how should the response be targeted to maximize effectiveness and limit costs? Specifically:

- **Overall scale of the response.** How many personnel, supplies, and how much money should be allocated to the response, given the opportunity costs of reassigning these personnel from other health-enhancing activities within the health sector (e.g. routine vaccination) and, in the case of large outbreaks, possibly opportunity costs from outside the health sector (eg extra health spending occasioned by the outbreak). As with many of the allocation decisions below, this is a question that will be reassessed repeatedly

throughout the course of the outbreak, up to the decision to terminate the outbreak response after the outbreak is over.

- **Targeting of countermeasures.** If countermeasures to treat or prevent infection are available during an outbreak, they will likely be in short supply. Such countermeasures may include supportive or specific anti-infective medications, personal protective equipment, or vaccines. A key decision for public health officials is to make recommendations or policies for who should receive these treatments for their own direct protection, based on criteria of highest effectiveness, greatest need, greatest social value, or population preferences, among other possible criteria. For the subset of these countermeasures that can prevent transmission of the infection, such as vaccines, the timing and the choice of recipients for the supply of countermeasures might be chosen to optimally reduce the transmission rate of the infection.

3. Assessments that are crucial to making these decisions well

Specific inputs of evidence about the nature of the disease and the state of the outbreak should inform each of these classes of decisions. Crucial assessments, and sources of uncertainty for each, include:

- **Disease severity.** Often measured as a case-fatality proportion or case-hospitalization proportion, severity of the novel infection informs the magnitude and immediacy of the response that should be undertaken, while relative severity measures in different groups inform the appropriate targeting of prevention and treatment interventions. Severity measures may change as the natural history of the disease becomes better understood, as in the case of Zika virus, for which the risk of congenital malformations in the offspring of infected pregnant women came to be appreciated as the primary severity measure. Comparative severity measures in different demographic groups, defined by age or comorbidities for example, can enhance targeting of scarce countermeasures.
  Sources of uncertainty. Especially in the early phases of an outbreak, cases with a known outcome are likely unrepresentative of all cases, thus complicating the effort to estimate the typical severity of infection. On one hand, observed cases early on will typically be more severe than average, as severe cases are more likely to come to medical attention and be diagnosed. Typically at the start of an epidemic, it is unclear what fraction of cases are asymptomatic or subclinical, as these are rarely observed. This factor tends to cause severity to appear higher early on than it is (Lipsitch et al. 2015). On the other hand, in a growing epidemic, it is now recognized that severity may be underestimated when total reported severe outcomes (eg deaths) are divided by total reported cases, because cases may be reported before their outcome is known and reported, so the denominator will include many people who have not yet entered the numerator, but will enter it in the future (Garske et al. 2009). The unknown balance between these opposing biases creates uncertainty in severity estimation. When calculating subgroup-specific severity measures, possible variation between subgroups in the probability a case is detected or reported at varying severity levels (eg symptomatic, hospitalized, fatal) can produce uncertainty in severity comparisons

between subgroups (Jain et al. 2009; Lipsitch et al. 2015; Rudolf et al. 2017; Wolkewitz and Schumacher 2017).

- **Epidemic size and geographic extent.** The total number of cases informs the number of persons affected, the number still at risk, and the resource requirements for treating patients and containing the outbreak. The trend in this total can be used to estimate the rate of spread and measures of contagiousness (such as the basic reproductive number R_0), and to project resource needs into the future. The geographic extent of cases, and its trend, allow similar estimates on smaller spatial scales and may inform efforts to understand the routes of transmission.
  Sources of uncertainty: Not every disease case may be reported, due to limited capacity for surveillance. At the earliest phases of an epidemic, surveillance capacity may not be in place and may miss cases, while later on, there may be too many cases to count, and surveillance methods may need to be modified (Lipsitch et al. 2009a). These constraints may change over time, producing artifactual trends, and may vary from place to place, causing apparent differences between places that are due to surveillance capacity variation rather than only to variation in case numbers(White et al. 2009). In addition to all these factors, nearly all traditional surveillance systems have a delay between the occurrence of a case and its reporting, producing an artifactual decline in the epidemic curve as it approaches the present due to underreporting of recent cases. “Nowcasting” algorithms, often involving nontraditional disease surveillance data sources, can be particularly helpful in addressing these limitations (Höhle and an der Heiden 2014; Bastos et al. 2017; McGough et al. 2017; van de Kasteele et al. 2019) (see below).

- **Transmissibility.** Crucial for any effort to predict how an epidemic will spread are two numbers: how many secondary cases each infected person causes, and how long it takes them to do so. These are known technically as the reproductive number and the serial interval or generation interval. In reality, each of these varies from case to case, so they are more accurately described as distributions, each with its own mean and variation around the mean (Wallinga and Lipsitch 2007). Using these quantities and various types of mathematical models, spread of the infection can be projected over time, and the potential impact of seasonal variation in transmission, depletion of susceptibles by an immunizing infection, and various countermeasures (eg treatment or vaccination) can be estimated.
  Sources of uncertainty: At the very beginning of an outbreak, these quantities may be measured directly by contact tracing, so that secondary cases are traced back to primary cases, and generation intervals are estimated as the time between symptom onset in successive cases in a chain (an approximate measure of the serial interval). As the epidemic expands (and sometimes from the beginning) this is impractical given limited resources, and these quantities must be estimated from the daily number of new cases (epidemic curve) (Wallinga and Teunis 2004; White and Pagano 2008). As a consequence, all the sources of uncertainty in case counting noted above become sources of uncertainty in estimating transmissibility, though there are ways to address these (White et al. 2009), including the use of pathogen genome sequences to provide

independent estimates of the dynamics of the outbreak (Fraser et al. 2009). Methodological errors can mask uncertainty in transmissibility estimates, but these are easily avoided (Magpantay and Rohani 2015).

- **Countermeasure availability, status, and effectiveness.** Central to response planning and implementation is an accurate inventory of what countermeasures are available, in what quantities and locations, and how effective they are projected to be. Countermeasures include supplies to prevent transmission (vaccines, personal protective equipment, prophylactic antiinfective medications) and treat cases (medications for treatment, medical devices such as ventilators, health care disposables and supplies such as IV fluids (Voelker 2018)). The effectiveness of many of these countermeasures will be unknown at the start of an outbreak and may change over time (eg through the development of resistance by the pathogen causing the outbreak). Countermeasures also include behavioral, social and economic interventions such as movement restrictions (Peak et al. 2018), closing of public gatherings and venues (Hatchett et al. 2007), and regionally-varying factors such as opening and closing of schools (Chao et al. 2010; Huang et al. 2014).
  Sources of uncertainty: For novel diseases, countermeasures will have uncertain effectiveness because they have not been tested and may be available, if at all, in short supply (Lipsitch and Eyal 2017). Timetables for producing such countermeasures (eg vaccines) depend on logistical factors that may be independent of, or even exacerbated by, the outbreak itself (Voelker 2018). The situation will change rapidly, as stockpiles are developed, depleted and replenished (Dimitrov et al. 2011). Even for known diseases, such as influenza, vaccine effectiveness varies from year to year (Osterholm et al. 2012).

A number of traditional and novel data sources can inform the real-time estimation of these quantities and the level of uncertainty of each estimate. We next review these data sources.

4. Data sources

To provide evidence in the four key areas noted above, a range of traditional and novel Internet-based data sources are available. We highlight some of the key ones in this section.

4.1. Traditional data sources

Early in an outbreak, the full data on the state of the outbreak may be contained in an epidemiologic *line list*, ideally containing demographic and geographic data on cases, clinical data on the diagnosis, course of their illness and treatment, as well as key dates such as the date on which they were infected (if known), became symptomatic, were reported to public health authorities, and, as applicable, were hospitalized, admitted to intensive care, recovered/discharged, or died. Many of these elements may be unavailable, at least temporarily, for some cases, so several efforts have been made to define minimal data sets needed for basic analyses early in outbreaks (Van Kerkhove et al. 2010; Cori et al.

2017). On the other hand, tools have been developed recently to implement more complex data structures that may include different elements for different cases and can incorporate novel types of data, including pathogen sequences when available (Grad and Lipsitch 2014; Jombart et al. 2014; Finnie et al. 2016).

As an outbreak grows, it will likely become impossible for some jurisdictions to continue testing all suspected cases and/or reporting detailed data on suspected or confirmed cases of disease. Alternative approaches, such as reporting clinical events (emergency department or primary care visits meeting syndromic criteria, for example), combined with diagnostic test results on a fraction of these clinical cases, to maintain a quantitative picture of the progress of the outbreak while using fewer resources (Lipsitch et al. 2009a). In locations with limited resources, this strategy may be employed from the start.

These epidemiologic data will be central to the first three evidence needs outlined above. The fourth need, to estimate countermeasure availability and effectiveness, will primarily require *logistical and supply chain* information about the production and distribution of vaccines, pharmaceutical treatments, and personal protective equipment. For antiinfective treatments, real-time *data on the susceptibility of cases* will be required to assess the likely impact of these treatments, estimate trends in resistance, and inform the optimal use of these and other countermeasures (Leung et al. 2017). To improve estimates of their past and potential effectiveness, data on the *timing and geographic scope of nonpharmaceutical interventions*, such as movement restrictions, safe burial practices (Tiffany et al. 2017) or school closings and openings (Chao et al. 2010; Huang et al. 2014) may be gathered by traditional means (surveys or administrative data) or by some of the novel means described below (Peak et al. 2018).

4.2 Novel data sources

The availability of big data sets, generated and recorded constantly due to the activities of millions of Internet and mobile phone users, has increased significantly and has opened up new ways to understand changes in human behavior. Of particular interest is the availability of Internet-based data that may help us detect changes in human behavior that may signal the emergence of a public health threat in real-time. These data may include unusual surges of symptom-related search activity on Internet search engines, an increase of symptom-related posts on social media, increased sales in over the counter medications to combat fever or other symptoms.

In fact, in the past decade, many research teams have been able to identify historical relationships between information contained in healthcare-based disease surveillance systems-- such as the number of hospitalizations and/or patients seeking medical attention with an array of symptoms-- and symptom-related *Internet search* behavior (Yang et al. 2015), *Wikipedia* article views (Generous et al. 2014; McIver and Brownstein 2014), *clinicians' search* behavior (Santillana et al. 2014), crowd-sourced symptom self-reporting apps (Smolinski et al. 2015; Koppeschaar et al. 2017), symptom-related *Twitter posts*

(Signorini et al. 2011; Paul et al. 2014), prescription changes contained in cloud-based *electronic health records (Santillana et al. 2016; Yang et al. 2017; Lu et al. 2019)* and historical synchronicities in disease activity in neighboring regions (Lu et al. 2019) , *weather* patterns, etc. These studies have shown that behavior changes in human populations, often a consequence of (or correlated with) increased disease activity, have detectable signatures in systems that were not originally designed as public health surveillance systems. These findings suggest that monitoring Internet search and/or social media activity related to symptoms or specific diseases, may help confirm the presence of public health threats.

Once a local disease outbreak has been identified, current and future *weather* patterns that may be conducive for such outbreak to further disseminate may be identified and may allow the creation of risk maps in real-time. For example, it is now well known that changes in ambient air moisture (relative humidity) influence the mechanistic human-to-human transmission of respiratory diseases such as influenza (Lowen et al. 2007; Shaman and Kohn 2009; Shaman et al. 2011). Drier months, such as those that happen during the colder seasons --in mid-latitudes-- enhance disease-transmission. Vector-borne diseases such as Dengue, Malaria, and Yellow Fever, can only be spread if local conditions are suitable for mosquitos to exist and reproduce (Kyle and Harris 2008). Thus, maps of the presence of vectors could be used to product risk maps in real-time (Messina et al. 2015). *Mobile phone* information can be used at the local level to map human mobility, whereas bus, train, or airline logs can be used to assess the likelihood of a given disease to be transmitted from point A to point B. Models incorporating these data have demonstrated the potential to predict outbreaks in new geographic locations, for example with dengue in Pakistan (Wesolowski et al. 2015).

While many of these data sources may be helpful for disease surveillance they have clear limitations. For example, people with mobile phones and/or Internet access do not necessarily reflect the underlying demographics of the locations where they live. This fact introduces biases that need to be considered when using in these data sources as indicators of the presence of a disease. Another limitation stems from the fact that people are susceptible to “panic searching” when news outlets alert them of unusual flu, or dengue, or Ebola disease outbreaks. As a consequence, peaks of search activity and increased social media microblogs discussing symptoms or diseases may only signal a population’s surge of interest in a disease-related topic but may not reflect actual infections. One of us (MS) is actively developing approaches to address these limitations (Santillana et al. 2015).

Finally, it has been shown that some of the uncertainties and limitations inherent to each individual data source may be mitigated by combining multiple data sources in order to assess the gravity of a disease outbreak (Santillana et al. 2015; McGough et al. 2017; Lu et al. 2019).

5. Situation Reporting as a Source of Common Planning Assumptions

A key aim of situation reporting, sometimes underappreciated, is to provide analysts and decision makers with a common set of facts (even if these are uncertain) so that decisions can be made using shared assumptions rather than unstated ones which may vary from person to person and thus cause confusion or error. Publicly stating working interpretations of existing data in a situation report is not intended to suppress disagreements in interpretation but rather to make these explicit, and to note which facts can be known with confidence and what are the key sources of uncertainty. Two examples from the 2009 influenza pandemic may help to illustrate the potential ofsituation reports centered on the four areas of critical evidence stated above to alleviate confusion and improve decisions.

- In the 2009 influenza pandemic, perhaps the most important quantity on which evidence was needed for decision making was the severity – as measured by case-fatality and case-hospitalization rates. Early estimates varied by a factor of 10,000, from a raw estimate in Mexico of 4% based on case and death numbers on May 4, to an adjusted estimate of 0.0004% published in July (Wilson and Baker 2009), spanning the full range of the severity scale established by the US Government for pandemic planning ~~(of Health et al. 2007)~~. The first official US government publication (to our knowledge) that contained a specific scenario at a particular severity level was the August 2009 PCAST report (Executive Office of the President and 's Council of Advisors on Science and Technology 2009), despite the fact that CDC investigations and surveillance had been producing relevant data in the US as early as April-May (Iuliano et al. 2009; Reed et al. 2009). *The act of assembling a thematic situation report that brought together diverse sources of evidence on severity could have helped to narrow this range of uncertainty by bringing together data that had been siloed in individual investigations.*

- Vaccine planning in the United States proceeded in the 2009 pandemic on the assumption of a mid- to late-winter peak of influenza incidence, allowing time for the production of enough doses (160 million) [6] to cover "initial target groups" in a timely fashion. This view was not supported by historical evidence from pandemics [7] cited by NIH authors. As they predicted based on historical experience, the major wave came in the autumn and was largely complete in most places in the US by the time many doses were available. *Making an explicit projection about the likely timing of the peak of cases and its uncertainty -- and specifically, the incorporation of historical data to provide context -- could have improved the quality of assumptions used to plan vaccine rollout and targeting.*

6. Projecting the future

Situation reporting is intimately connected with making projections about how an epidemic may unfold in the future. Indeed, some readers of a situation report may be primarily concerned not with how big or widespread the epidemic is now, but with how big or widespread it could become. Early situation reports will typically contain little in the way of projections, but as an epidemic develops, it may become appropriate to begin including some projections of its likely trajectory under various scenarios. Indeed, to achieve the goal

of creating common planning assumptions described in the previous section, some such scenarios must be developed and include a forward-looking component. Planning scenarios may be developed even without accurate forecasts, but they will be more useful if they are based on the best possible forecasts that can be achieved at a particular stage in the epidemic. Empirically, it should be noted that even when a planning scenario is explicitly and repeatedly annotated as being purely that, and not a prediction, it may be reported in the lay press as if it were a forecast. The 2009 PCAST working group report on the US Government's pandemic response repeatedly characterized its planning scenarios as "not a prediction" in three separate places [ref] but was reported by major news outlets as a prediction (for example, http://www.cnn.com/2009/HEALTH/08/24/us.swine.flu.projections/index.html accessed May 2, 2019).

The technical aspects of how to project disease incidence could fill an entire book, but for the purposes of situation reporting some crucial information should accompany any such projections and should be demanded by decision makers if not explicit in the situation report. The key question for any projection is what assumptions underlie it. In particular, many projections of disease cases indicate that if current trends continue, there may be x cases by a certain time. For infectious diseases, current trends cannot continue indefinitely. The simplest models may assume that the epidemic continues growing exponentially at the same rate as in its earliest phases. For any growing epidemic, such models can project arbitrarily high numbers of cases because exponential growth never ends (Meltzer et al. 2001) -- the only question is how long the epidemic will take to reach a given number of cases (Meltzer et al. 2014). Such projections usually provide a near-worst case scenario, because typically the factors that change during an epidemic tend to moderate transmission rather than increase it. That said, there are important exceptions such as changing weather or vector density for arthropod-borne infections, which can move cyclically with the seasons.

More refined projections -- not assuming "current trends continue" will incorporate factors that modify transmission, including behavior change induced by a desire to control the infection, behavior change for unrelated reasons (eg the beginning and end of school terms that affect directly transmitted diseases), seasonal changes that affect the suitability of transmission through the biology of the infectious agent or its vectors, and depletion of susceptible hosts as individuals previously infected become immune and thus reduce the opportunities for transmission. A projection should clearly state which of these factors it takes account of, what it assumes, and to what evidence exists (or is needed) to support these assumptions. Finally, efforts should be made to include uncertainty estimates (eg confidence intervals) around scenario-based projections that may be displayed on visualizations as uncertainty cones, similar to those used to monitor the likely trajectory of a hurricane in weather prediction systems.

7. Principles for high-quality situation reporting.

The goals of providing evidence to decision makers on key quantities relevant to responding to outbreaks, providing common scenarios for the purposes of planning, and highlighting areas of uncertainty, suggest four principles to enhance the quality of situation reporting in outbreaks.

- **Situation reporting should be thematic, concentrating on essential areas of evidence needed for decisions.** Situation reports should be designed for clarity and value to top-level decision makers, as well as for technical scrutiny by subject-matter experts. Decision makers may lack the time or skills or specialized knowledge to interpret raw data such as case counts, Google search trends, or the like. They may not immediately see the relevance of each data source to the key quantities about which they need information. Thus, maximal value to these consumers of the report will be achieved by organizing data outputs by the quantity of interest they inform, rather than in a simple list. This leads to the second principle:
- **Situation reports should adduce evidence from multiple sources to address each area of evidence, along with expert assessments of key parameters.** Text describing the expert judgment about severity, numbers and geographic extent, and other assessments should be combined with data in the forms of tables and graphs. Notwithstanding the wealth of potential data sources for tracking an outbreak and the response to it, data alone are not sufficient to support evidence-based decisions reflecting a clear picture of the four areas noted above. Key data may be unavailable, especially in the areas hardest hit by an outbreak, and even when available they may be limited, confusing or even misleading. Subject-matter experts -- epidemiologists, clinicians, data managers, and those involved in delivering the public health response -- will typically have knowledge that is vital to sound interpretation of the data. A crucial feature of situation reporting is to make the data, as well as this expert knowledge, widely available to enhance the quality of evidence for decisions, and also to allow scrutiny and critique of the interpretation. Crucial to this presentation is the next principle:
- **Situation reports should acknowledge uncertainty and attempt to estimate its magnitude for each assessment.** This prevents provisional assessments from becoming accepted as unchangeable facts, while acknowledging the possibility that estimates may change as data improve. Finally:
- **Situation reports should contain carefully curated visualizations along with text and tables.** These visualizations should clearly demarcate existing data from projections, visually represent uncertainty bounds, and be presented in intuitive ways that have been tested for clarity with an audience of decision makers before a crisis hits.

8. Conclusion

Accurate, informative, and clear situation reporting is essential for evidence-based decision making and planning in the midst of an outbreak that may be chaotic and full of confusing and contradictory information. In this chapter we have advocated for augmenting raw data with expert interpretation and planning scenarios to aid the decision makers by providing open discussion of what is and is not known and a set of shared assumptions for planning

purposes. New data sources provide unprecedented opportunities to improve our understanding of epidemic dynamics as a new outbreak unfolds, and these must be integrated with more traditional data sources to aid decision makers in understanding the big picture, not only the raw data. Achieving these goals is a crucial part of minimizing the probability that an initially small and local outbreak grows to regionally or globally catastrophic proportions.

9. Acknowledgments

ML was partially supported by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number U54GM088558. MS was partially supported by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number R01GM130668. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Bastos L, Economou T, Gomes M, Villela D, Bailey T, Codeço C. Modelling reporting delays for disease surveillance data [Internet]. arXiv [stat.AP]. 2017. Available from: http://arxiv.org/abs/1709.09150

Chao DL, Halloran ME, Longini IM. School opening dates predict pandemic influenza A (H1N1) outbreaks in the United States. J Infect Dis. 2010;202(6):877–80.

Cori A, Donnelly CA, Dorigatti I, Ferguson NM, Fraser C, Garske T, et al. Key data for outbreak evaluation: building on the Ebola experience. Philos Trans R Soc Lond B Biol Sci [Internet]. 2017 May 26;372(1721). Available from: http://dx.doi.org/10.1098/rstb.2016.0371

Dimitrov NB, Goll S, Hupert N, Pourbohloul B, Meyers LA. Optimizing tactics for use of the U.S. antiviral strategic national stockpile for pandemic influenza. PLoS One. 2011 Jan 19;6(1):e16094.

Executive Office of the President, 's Council of Advisors on Science and Technology. Report to the President on US Preparations for 2009-H1N1 Influenza. 2009 Aug.

Finnie TJR, South A, Bento A, Sherrard-Smith E, Jombart T. EpiJSON: A unified data-format for epidemiology. Epidemics. 2016 Jun;15:20–6.

Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. Science. 2009 Jun 19;324(5934):1557–61.

Garske T, Legrand J, Donnelly CA, Ward H, Cauchemez S, Fraser C, et al. Assessing the severity of the novel influenza A/H1N1 pandemic. BMJ. 2009 Jul 14;339:b2840.

Generous N, Fairchild G, Deshpande A, Del Valle SY, Priedhorsky R. Global disease monitoring and forecasting with Wikipedia. PLoS Comput Biol. 2014 Nov;10(11):e1003892.

Grad YH, Lipsitch M. Epidemiologic data and pathogen genome sequences: a powerful synergy for public health. Genome Biol. 2014 Nov 18;15(11):538.

Hatchett RJ, Mecher CE, Lipsitch M. Public health interventions and epidemic intensity during the 1918 influenza pandemic. Proc Natl Acad Sci U S A. 2007 May 1;104(18):7582–7.

of Health USD, Services H, Others. Community strategy for pandemic influenza mitigation. 2007;

Höhle M, an der Heiden M. Bayesian nowcasting during the STEC O104:H4 outbreak in Germany, 2011. Biometrics. 2014 Dec;70(4):993–1002.

Huang KE, Lipsitch M, Shaman J, Goldstein E. The US 2009 A(H1N1) influenza epidemic: quantifying the impact of school openings on the reproductive number. Epidemiology. 2014 Mar;25(2):203–6.

Iuliano AD, Reed C, Guh A, Desai M, Dee DL, Kutty P, et al. Notes from the field: outbreak of 2009 pandemic influenza A (H1N1) virus at a large public university in Delaware, April-May 2009. Clin Infect Dis. 2009 Dec 15;49(12):1811–20.

Jain S, Kamimoto L, Bramley AM, Schmitz AM, Benoit SR, Louie J, et al. Hospitalized Patients with 2009 H1N1 Influenza in the United States, April–June 2009. N Engl J Med. 2009 Nov 12;361(20):1935–44.

Jombart T, Aanensen DM, Baguelin M, Birrell P, Cauchemez S, Camacho A, et al. OutbreakTools: a new platform for disease outbreak analysis using the R software. Epidemics. 2014 Jun;7:28–34.

van de Kasteele J, Elers P, Wallinga J. Nowcasting the number of new symptomatic cases during infectious disease outbreaks using constrained P-spline smoothing. Epidemiology. 2019;in press.

Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, et al. Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe. JMIR Public Health Surveill. 2017 Sep 19;3(3):e66.

Kyle JL, Harris E. Global spread and persistence of dengue. Annu Rev Microbiol. 2008;62:71–92.

Leung K, Lipsitch M, Yuen KY, Wu JT. Monitoring the fitness of antiviral-resistant influenza strains during an epidemic: a mathematical modelling study. Lancet Infect Dis. 2017 Mar;17(3):339–47.

Lipsitch M. If a Global Catastrophic Biological Risk Materializes, at What Stage Will We Recognize It? Health Secur. 2017 Jul 26;15(4):331–4.

Lipsitch M, Donnelly CA, Fraser C, Blake IM, Cori A, Dorigatti I, et al. Potential Biases in Estimating Absolute and Relative Case-Fatality Risks during Outbreaks. PLoS Negl

Trop Dis. 2015 Jul 16;9(7):e0003846.

Lipsitch M, Eyal N. Improving vaccine trials in infectious disease emergencies. Science. 2017 Jul 14;357(6347):153–6.

Lipsitch M, Finelli L, Heffernan RT, Leung GM, Redd SC, 2009 H1n1 Surveillance Group. Improving the evidence base for decision making during a pandemic: the example of 2009 influenza A/H1N1. Biosecur Bioterror. 2011 Jun;9(2):89–115.

Lipsitch M, Hayden FG, Cowling BJ, Leung GM. How to maintain surveillance for novel influenza A H1N1 when there are too many cases to count. Lancet. 2009a Oct 3;374(9696):1209–11.

Lipsitch M, Riley S, Cauchemez S, Ghani AC, Ferguson NM. Managing and Reducing Uncertainty in an Emerging Influenza Pandemic [Internet]. Vol. 361, New England Journal of Medicine. 2009b. p. 112–5. Available from: http://dx.doi.org/10.1056/nejmp0904380

Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog. 2007 Oct 19;3(10):1470–6.

Lu FS, Hattab MW, Clemente CL, Biggerstaff M, Santillana M. Improved state-level influenza nowcasting in the United States leveraging Internet-based data and network approaches [Internet]. Vol. 10, Nature Communications. 2019. Available from: http://dx.doi.org/10.1038/s41467-018-08082-0

Magpantay FMG, Rohani P. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. R Soc B [Internet]. 2015; Available from: http://rspb.royalsocietypublishing.org/content/282/1806/20150347.short

McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting Zika Incidence in the 2016 Latin America Outbreak Combining Traditional Disease Surveillance with Search, Social Media, and News Report Data. PLoS Negl Trop Dis. 2017 Jan;11(1):e0005295.

McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. PLoS Comput Biol. 2014 Apr;10(4):e1003581.

Meltzer MI, Atkins CY, Santibanez S, Knust B, Petersen BW, Ervin ED, et al. Estimating the future number of cases in the Ebola epidemic--Liberia and Sierra Leone, 2014--2015. 2014; Available from: https://stacks.cdc.gov/view/cdc/24901

Meltzer MI, Damon I, LeDuc JW, Millar JD. Modeling potential responses to smallpox as a bioterrorist weapon. Emerg Infect Dis. 2001 Nov;7(6):959–69.

Messina JP, Brady OJ, Pigott DM, Golding N, Kraemer MUG, Scott TW, et al. The many projected futures of dengue. Nat Rev Microbiol. 2015 Apr;13(4):230–9.

Osterholm MT, Kelley NS, Sommer A, Belongia EA. Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. Lancet Infect Dis. 2012 Jan;12(1):36–44.

Paul MJ, Dredze M, Broniatowski D. Twitter improves influenza forecasting. PLoS Curr [Internet]. 2014 Oct 28;6. Available from:

http://dx.doi.org/10.1371/currents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117

Peak CM, Wesolowski A, Zu Erbach-Schoenberg E, Tatem AJ, Wetter E, Lu X, et al. Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data. Int J Epidemiol. 2018 Oct 1;47(5):1562–70.

Reed C, Angulo F, Swerdlow D, Lipsitch M, Meltzer M, Others. Estimating the burden of pandemic influenza A/H1N1--United States, April-July 2009. Emerg Infect Dis. 2009;

Rudolf F, Damkjær M, Lunding S, Dornonville de la Cour K, Young A, Brooks T, et al. Influence of Referral Pathway on Ebola Virus Disease Case-Fatality Rate and Effect of Survival Selection Bias. Emerg Infect Dis. 2017 Apr;23(4):597–600.

Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance. PLoS Comput Biol. 2015 Oct;11(10):e1004513.

Santillana M, Nguyen AT, Louie T, Zink A, Gray J, Sung I, et al. Cloud-based Electronic Health Records for Real-time, Region-specific Influenza Surveillance. Sci Rep. 2016 May 11;6:25732.

Santillana M, Nsoesie EO, Mekaru SR, Scales D, Brownstein JS. Using clinicians' search query data to monitor influenza epidemics. Clin Infect Dis. 2014 Nov 15;59(10):1446–50.

Shaman J, Goldstein E, Lipsitch M. Absolute humidity and pandemic versus epidemic influenza. Am J Epidemiol. 2011 Jan 15;173(2):127–35.

Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. Proc Natl Acad Sci U S A. 2009 Mar 3;106(9):3243–8.

Signorini A, Segre AM, Polgreen PM. The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic. PLoS One. 2011 May 4;6(5):e19467.

Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons. Am J Public Health. 2015 Oct;105(10):2124–30.

Tiffany A, Dalziel BD, Kagume Njenge H, Johnson G, Nugba Ballah R, James D, et al. Estimating the number of secondary Ebola cases resulting from an unsafe burial and risk factors for transmission during the West Africa Ebola epidemic. PLoS Negl Trop Dis. 2017 Jun;11(6):e0005491.

Van Kerkhove MD, Asikainen T, Becker NG, Bjorge S, Desenclos J-C, dos Santos T, et al. Studies needed to address public health challenges of the 2009 H1N1 influenza pandemic: insights from modeling. PLoS Med. 2010 Jun 1;7(6):e1000275.

Voelker R. Vulnerability to Pandemic Flu Could Be Greater Today Than a Century Ago. JAMA. 2018 Oct 16;320(15):1523–5.

Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth

rates and reproductive numbers. Proc Biol Sci. 2007 Feb 22;274(1609):599–604.

Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol. 2004 Sep 15;160(6):509–16.

Wesolowski A, Qureshi T, Boni MF, Sundsøy PR, Johansson MA, Rasheed SB, et al. Impact of human mobility on the emergence of dengue epidemics in Pakistan. Proceedings of the National Academy of Sciences. 2015;112(38):11887–92.

White LF, Pagano M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. Stat Med. 2008 Jul 20;27(16):2999–3016.

White LF, Wallinga J, Finelli L, Reed C, Riley S, Lipsitch M, et al. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. Influenza Other Respi Viruses. 2009 Nov;3(6):267–76.

Wilson N, Baker MG. The emerging influenza pandemic: estimating the case fatality ratio. Euro Surveill [Internet]. 2009 Jul 2;14(26). Available from: https://www.ncbi.nlm.nih.gov/pubmed/19573509

Wolkewitz M, Schumacher M. Survival biases lead to flawed conclusions in observational treatment studies of influenza patients. J Clin Epidemiol. 2017 Apr;84:121–9.

Yang S, Santillana M, Brownstein JS, Gray J, Richardson S, Kou SC. Using electronic health records and Internet search information for accurate influenza forecasting. BMC Infect Dis. 2017 May 8;17(1):332.

Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. Proc Natl Acad Sci U S A. 2015 Nov 24;112(47):14473–8.