Understanding and predicting the global spread of emergent infectious diseases

Dirk Brockmann

The rapid spread of SARS in 2003, the 2009 H1N1 influenza pandemic and the emergence of MERS-CoV in 2013 in the Middle East illustrate the increasing threat of emergent infectious diseases and their global spread to public health. Massive worldwide mobility in combination with exponentially growing, dense populations increase the rate at which novel pathogens emerge and amplify the speed at which they can proliferate worldwide. Compared, for instance, to the 14th century Black Death pandemic that advanced at a speed of only a 3-5 km/day (Noble, 1974) killing 20-50% of the European population in its wake, the 2009 H1N1 pandemic progressed a few hundred km/day on average (Brockmann and Helbing, 2013). Given that more than 10 million passengers travel a total approx. 14 billion km (three times the radius of our solar system!) every day on the worldwide air-transportation network (WAN), this is not surprising. Global mobility has not only increased in magnitude, but also in complexity (Brockmann et al., 2006). The WAN connects more than 4000 airports by more than 25000 links yielding a complex, intricate and multi-scale network (see Fig. 1b).

During the last decade the development and application of high-performance computational models has become an important strategic component in understanding and mitigating disease spread on a global scale (Brockmann and Helbing, 2013, Hufnagel et al., 2004, Colizza et al., 2007, van den Broeck et al., 2011). The majority of these computational approaches incorporate comprehensive data on human mobility and transportation networks. Although the development of parsimonious mathematical models in the context of disease dynamics has a long history (going back to Daniel Bernoulli who developed the first mathematical model to resolve the
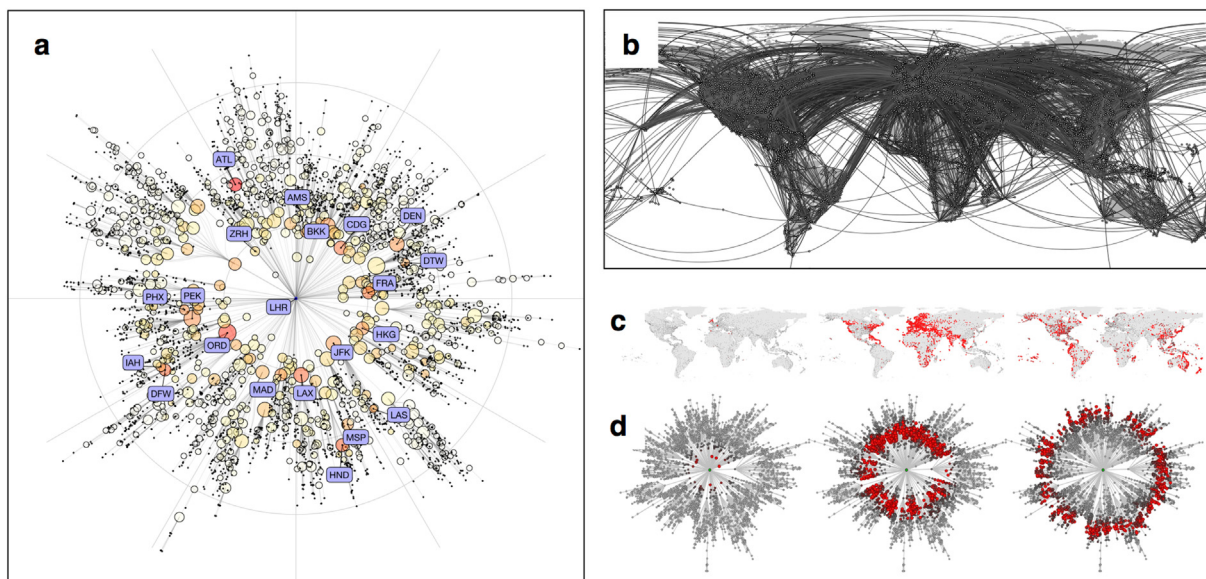

Figure 1. **Understanding global disease dynamics using effective distance. a**: Effective distance from a reference node (London Heathrow) is quantified by radial distance from the origin. Each node in the diagram represents one of approx. 4000 airports in the worldwide air-transportation network. The tree structure represents the most probable spreading path of a potential pandemic. Effective distances and most probable paths are computed from traffic flows in the complete network. **b**: The entire worldwide air-transportation network. More than 4000 airports are connected by more than 25000 direct connections amounting to a total of more than 3 billlion passengers per year. **c**: Temporal snapshots of a simulated pandemic with initial outbreak in London. In the conventional geographic representation spatio-temporal patterns are incoherent and complex. **d**: the identical simulation in the effective distance representation exhibits simple, concentric wave patterns which allows immediate calculations of expected arrival times at any node in the network.

17th century cow-pox inoculation debate (Bernoulli, 1760)), modern computational models account for unprecedented detail and a multitude of factors that are believed to play a role in global disease dynamics. Relying on the increasing performance capabilities of modern supercomputers, agent-based simulations and complex meta-population models have been developed that are able to model the behavior of nearly one billion individuals. While early models in this context were able to reproduce the global dynamics of e.g. the SARS epidemic (Hufnagel et al., 2004), state-of-the art models are now able to produce reliable quantitative forecasts and projections that can help policy makers develop more efficient mitigation strategies and optimize the distribution of e.g. anti-viral drugs and vaccines (van den Broeck et al., 2011). One of the key features that nearly all modern computational models predict is that, unlike historic pandemics that advanced in regular wave like patterns, modern diseases spread in spatially incoherent ways due to the complexity of underlying mobility networks (see Fig. 1c). This makes it impossible to predict e.g. epidemic arrival times based on spreading speed and geographic distance to the outbreak location. In fact, the lack of correlation of arrival times and geographic distance requires running high performance computer simulation in the first place. However, despite their increasing sophistication and accuracy, high-performance computer simulation approaches face a number of fundamental problems. First, the complex nature of spatio-temporal patterns of disease dynamics and computer generated pandemic scenarios are still poorly understood at a fundamental level. Secondly, highly detailed computer simulations that account for all possible factors that

may impact disease dynamics typically come with multiple unknown parameters that must be estimated or even guessed. Because disease dynamical models are typically nonlinear it is often unclear how robust predictions are with respect to errors in these parameter estimates. Because detailed computer simulations are typically computationally resource demanding, systematic parameter scans cannot be performed efficiently. Furthermore, when all possible factors are incorporated into these models from the start, it is generically difficult to assess in simulations which parameters are central and which ones are peripheral for the dynamic outcome. These drawbacks are particularly severe in situations when predictive computer simulations are needed most: during the initial outbreak of novel pathogens with uncertain properties. In this situation, quantitative forecasts come with significant error bars because simulation parameters that define the new pathogen are not known. This dilemma can only be solved if the dynamics of global disease spread is better understood at a fundamental level, i.e. if we identify which model ingredients dominate the patterns of global spread, which ones can be neglected and what features of global spread are universal and valid for the majority of emergent pathogens.

One step in this direction was the recent discovery that the patterns of global spread of SARS, H1N1 and related diseases can be understood and substantially simplified by introducing a novel type of effective distance that is based on methods borrowed from complex network- and random walk theory (Brockmann and Helbing, 2013). Conventional computer simulations that incorporated for example the WAN typically use conventional

geographic map representations in which the spatio-temporal dynamics of epidemic spreads exhibit the aforementioned complex, spatially incoherent structure. When conventional geographic distances are replaced by effective distances between airports, these complex patterns are mapped onto regular wave-like patterns that are reminiscent of historic spreading phenomena such as the historic Black Death pandemic. This allows the measurement of effective spreading speeds and the prediction of relative arrival times with a much higher fidelity and independently of disease specific parameters. This is possible because the effective distance between any two locations in the network can be computed by mobility network properties alone (see Fig. 1a/d).

One of the key implications of this approach is that even in situations when little is known about an emergent pathogen and the likelihood of global spread the most probable spreading pathways and relative arrival times at distant location can be computed. This can help during the design of mitigation strategies that are based e.g. on the implementation of travel restrictions and airport based containment actions. The network theoretic, effective distance approach provides another important feature that is derived from the observation that only from the actual outbreak location a regular wave-front is observed in the effective distance representation. One can make use of this property to solve the so-called inverse problem: Given a complex spatial incidence patterns one can reconstruct the most likely initial outbreak location. This technique has been shown to work efficiently when applied to the 2011 EHEC/HUS outbreak in Germany using a proxy of food transportation in Germany as the

underlying network (Manitz et al., 2014).

Given the increasing availability of data on human mobility and human interactions modern computational and network-theoretic models for disease dynamics will become a central tool for understanding and predicting disease dynamics on local, intermediate and global scales and will aid policy makers and public health research in mitigating their negative effects on society.

Der korrespondierende Autor erklärt, dass kein Interessenkonflikt vorliegt.

Prof. Dr. Dirk Brockmann
Robert Koch-Institute, Seestr. 10,
13353 Berlin
Institute for Theoretical Biology,
Department of Biology
Humboldt University Berlin
Invalidenstraße 43
10115 Berlin
dirk.brockmann@hu-berlin.de

References

Bernoulli D. Reflexions sur les avantages de l'inculation. Mém Math Phys Acad R Sci Paris 1760;1–45.

Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel. Nature 2006;439:462.

Brockmann D, Helbing D. The hidden geometry of complex, network-driven contagion phenomena. Science 2013;342:1337.

Colizza V, et al. Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions. PLoS Med 2007;4:95–110.

Hufnagel L, Brockmann D, Geisel T. Forecast and control of epidemics in a globalized world. Proc Natl Acad Sci 2004;101:15124.

Manitz J, et al. Origin Detection During Foodborne Disease Outbreaks – A Case Study of the 2011 EHEC/HUS Outbreak in Germany. PLoS Currents: Outbreaks 2014; Apr 1. Edition 1.

Noble J. Geographic and temporal development of plagues. Nature 1974;250:726–8.

van den Broeck W, et al. The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. BMC Infect Dis 2011;11:37.

Abstract

The emergence and global spread of human infectious diseases has become one of the most serious public health threats of the 21st century. Sophisticated computer simulations have become a key tool for understanding and predicting disease spread on a global scale. Combining theoretical insights from nonlinear dynamics, stochastic processes and complex network theory these computational models are becoming increasingly important in the design of efficient mitigation and control strategies and for public health in general.

Einleitung

Die Entstehung und globale Ausbreitung neuartiger, von Mensch zu Mensch übertragbarer Infektionskrankheiten ist eines der größten Public-Health-Probleme des 21. Jahrhunderts. Detaillierte, umfangreiche Computersimulationen haben sich in diesem Kontext zu einem wichtigen Werkzeug für das bessere Verständnis der Ausbreitungswege und deren Vorhersage entwickelt. In modernen Ausbreitungssimulationen werden Methoden aus der theoretischen und rechnergestützten Physik, der nichtlinearen Dynamik mit Erkenntnissen aus der Theorie komplexer Netzwerke in numerischen Modellen vereint um genauerer Vorhersagen und potenzielle Ausbreitungsszenarien „in silico" zu untersuchen und damit das Design effizienterer Mitigations- und Kontrollstrategien zu fördern und zu beschleunigen.

Keywords: