

The importance of networks in the field of epidemiology

Epidemiology is one of the fields within medicine where network theory has had a great impact, and been of great importance, in the past decades. One event where networks was of great importance was during the H1N1 outbreak in 2009. Due to great improvement in network technology, the time evolution of the disease was predicted before it reached its peak (Barabási). With the aid of networks, transmission and sources of infection can potentially be tracked down, and, knowing the characteristics of certain diseases, there is a possibility to prevent spread at an early stage.

In previous years, epidemiologic data was collected by means of public health agencies. The incorporation of networks as made this distinctively more effective, as data now more easily can be traced down to the individual through digital communication. Information on treatment, symptoms and disease progression can be stored online and used in worldwide research and analysis of disease dynamics.

GIDEON (the Global Infectious Diseases and Epidemiology Network) is an example of a modern day, online program used in diagnosis and informatics in infectious diseases. It can be accessed by everyone and provides up-to-date information on the latest disease outbreaks. It consists of four modules: Diagnosis, Epidemiology, Therapy and Microbiology.

Network epidemics

Infectious diseases are responsible for approximately 43% of the global diseases burden. In other words, getting a better understanding of how pathogens spread can theoretically be of great importance to reduce this number. The framework used in epidemiology to model the spread of pathogens is composed of two main frameworks – compartmentalization and homogenous mixing. Compartmentalization categorizes people based on their stage of disease, being either susceptible, infectious or recovered. Homogenous mixing is a hypothesis that assumes that all individuals have the same chance of being infected by a certain disease.

The two hypotheses mentioned in the former paragraph are applied to different epidemic modeling frameworks. One such framework is the *susceptible-infected model* (SI), where an individual is categorized as being in either of the two states – susceptible or infected. This model starts with an initial exponential growth in the number of infected individuals, and reaches a saturation point when most people have become infected.

Another framework is the *susceptible-infected-susceptible model* (SIS). Compared to the first model mentioned, this one also brings attention to people recovering from the disease, bringing them back to a state of being susceptible without being infected. Everyone is eventually infected in the SI model, whereas in the SIS model there are two possible outcomes – an endemic state and a disease-free state. In the endemic state, the fraction of infected individuals reaches a limit, a stationary state, where the "number of newly infected individuals equals the number of individuals who recover from the disease". In the disease-free state, the infectiosity decreases as more people are cured than people being infected, and eventually the pathogen will be gone. With these two states, the SIS models relies on the possibility that a pathogen will either remain in the population, the endemic state, or disappear, the disease-free state. The characteristics of the pathogen can be described with the metric *basic reproduction number*, R_0 , which is the amount of new cases one case may generate over the course of its infectious period. This metric is the first to be applied when epidemiologists want to estimate the spread of a new pathogen. If R_0 is less than 1, the pathogen will eventually be gone, whereas if it is greater than 1, it will spread and remain in the population.

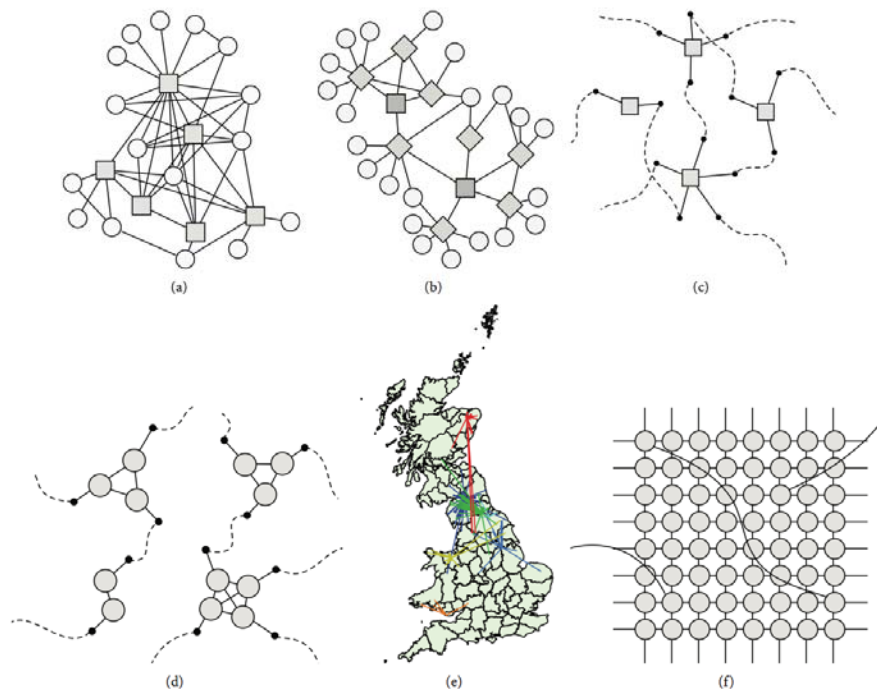
The last model to mention is the *susceptible-infected-recovered model* (SIR). This brings attention to a last possible state of having recovered from the disease, where the individuals are no longer susceptible to a disease as they have developed immunity. The three models mentioned assume homogenous mixing, and are limited in that they do not take into consideration how close contacts and networks have an effect on disease spread. To get a more accurate prediction of disease spread, we must assume that individuals in certain contact networks are more likely to get a certain disease, which can be achieved by further advancement in the model set-up.

Network types

Numerous types of networks can be implemented in epidemiologic study, some will here be summarized with relevant examples.

Contact tracing networks

Contact tracing is elemental in the epidemiology of sexually transmitted infections (STIs). STIs are a disease group where many potential transmission routes exist, at the same time as these are quite obvious. The method is similar to that applied in contact tracing, called snowball sampling ((b) in the diagram below – squares, diamonds and circles refer to primary, secondary and tertiary contacts, respectively), where the individuals are asked to name all their sexual partners over a given time period.



Examples of networks in epidemiology. (a) contacts between IV drug users. One circle represents one node, an individual, an edge represents an interaction, and a node's degree is the number of edges attached to it. (b) mentioned above. (c) example of a configuration model network. (d) a household configuration network. (e) map showing cattle movement.

Contact tracing was also important in the aforementioned H1N1 pandemic. However, impacting factor here was that contacts generally were traced and treated before they generated any secondary cases. Air-borne diseases are also a bigger challenge compared

to STIs due to greater uncertainties in who potential contacts are. Contact tracing is advantageous because it provides information on the transmission of infection, and the modes of contact are obvious. Unfortunately, this network theory relies on the concurrent infectious process, and has limitations in predicting future spread.

Mathematical models have been implemented in epidemiologic research and pattern analysis regarding STIs. These usually target the stage of infection and correlations between infection and sexual activity.

Sexual partnership networks can be studied by simulation models. An example is a study of the role of partnership networks in the epidemiology of gonorrhoea, where heterosexual partnerships in a close population were simulated. In addition to the usual idea that the risk of infection is mainly linked to the number of partners, this study also assessed the link between infection risk and distance between people.

Movement networks

Movement networks are, as the name implies, based on the movement of individuals. Instead of nodes being represented as individuals, they represent locations. Movement forms can roughly be broken down into four types: airline transportation, movement to and from work, movement of dollar bills, and movement of livestock. A significant data collection was done on the latter, the so-called Cattle Tracing Scheme, which gathers data on movement of thousands of farmers in Great Britain. Compared to other movement networks, this is updated frequently, on a daily basis, and is of great importance because movement of cattle is one of the major contributors to disease spread. For instance, the spread of foot and mouth disease in 2001 was due to livestock movement.

Network dynamics

The tracing of livestock movement mentioned in the previous section conquered a big challenge in network epidemiology, which is to create a dynamic network. Many network models simply assume that contacts are static, but more realistic models should also aim at being more dynamic. When setting up an epidemic model, it must first be considered whether if the transmission of disease can be explained through a static

model, or if only a complex, dynamic model can be used to trace it. Static networks are mainly accurate if predicting disease spread in a stable population, for instance in a rapidly spreading spreading pathogen leading to acute infection. Static networks have for instance been of importance in study of animal health problems related to hand foot mouth disease and public health problems related to SARS, walking pneumonia, influenza, gonorrhoea etc.

One way to obtain a more dynamic model is to assume change over time by constructing "time-integrated networks". Referring back to the cattle movement in the UK, Vernon and Keeling compared several dynamic networks and integrated these into a more complete dynamic network.

Dynamic contact behaviour has also been found in research on the transmission of HIV and other STDs. Sexually transmitted diseases usually correlated with a higher number of contacts, causing the outbreaks to grow faster and reaching a larger epidemic size. Physical contacts greatly influences the spread of disease, and must be taken into consideration when we analyse population dynamics. On one hand, a population may be "fluid", where the contacts are random and not limited to certain contacts. On the other hand, a population may be static, where the contacts do not change over time. The dynamics between these two predictabilities are of variable duration and occurrences.

Several factors affect how and why a contact network is dynamic and simply not static. Extrinsic forces, for instance demographic changes, social changes, migration and seasonal changes affect the susceptible host population. The relative importance of these individual factors is again determined by time. An example is how the introduction of disease into a population depletes the number of susceptible people by reducing the number of new births.

Public health interventions is another factor affecting network dynamics.

Implementation of vaccinations or other public health policies acting down to an individual level, greatly reduces the number of susceptible individuals. Another important factor is pathogen-mediated changes. Pathogens may for instance affect changes in host behaviour. Rabies on one hand, cause a more aggressive behaviour in

the most, possibly increasing the rate of transmission. Common cold, on the other hand, causes the host to abstain from social interactions, thereby reducing the rate of transmission. There is still much research to be done on the correlation between symptoms and network changes, but a better understanding of these will allow us to more efficiently shape public health measures based on network dynamics.

MalariaGEN

Networks can be useful both in small-scale studies of STIs outbreak in a smaller population, as well as in global spread of troublesome diseases such as malaria. The Malaria Genomic Epidemiology Network (MalariaGEN) is a large-scale initiative aiming to eliminate malaria, involving researchers from 21 countries. Several factors contribute to the difficulties in completing this study, particularly the fact that it involves developing countries where the lack of technology is obvious. One target area in this network is the research on malaria resistance. Using genome-wide association (GWA) analysis, the genetic basis of resistance can be analysed at the level of the whole genome. The GWA studies use SNP genotyping, one of which difficulties is the large genomic variation on Africa. The ethnic diversity in Africa is also a challenging factor. On the contrary, these variations can be helpful in establishing a correlation between genes, the environment, infection and resistance.

Ethics is a key issue when creating a network, and many ethical and social issues come to light when personalized data is shared online. MalariaGEN set up a separate ethics team to deal with such issues. This team deals with setting up consultation groups, ensuring consent from participants and set up publication guidelines.

The point of mentioning this network was to highlight the importance of networks in global health issues. Large-scale epidemiology networks such as MalariaGEN was hopefully an important contributing factor in a take on the Global Malaria Action Plan.

Conclusion

Setting up a network to be used in epidemic modelling is a complex process, involving complex mathematics and behavioral analytics. A particular challenge lies with the implementation of network dynamics, instead of assuming that populations and disease

spread are static. After more research on this, in order to understand how extrinsic factors and individual behaviour affect the spread of disease, there will be an advancement in epidemiologic networks. Networks have undoubtedly been of great significance in epidemiology until this day, as mentioned with the MalariaGEN, and have an important role in targeting epidemiologic challenges in the future.

Sources

- A. A. (2015). Network Science by Albert-László Barabási. Retrieved from <http://networksciencebook.com/10#epidemic>
- The Malaria Genomic Epidemiology Network. A global network for investigating the genomic epidemiology of malaria. *Nature*. 2008;456(7223):732-737. doi:10.1038/nature07632.
- Leon Danon, Ashley P. Ford, Thomas House, et al., "Networks and the Epidemiology of Infectious Disease," *Interdisciplinary Perspectives on Infectious Diseases*, vol. 2011, Article ID 284909, 28 pages, 2011.
Hiba! A hiperhivatkozás érvénytelen.
- Shweta Bansal, Jonathan Read, Babak Pourbohloul & Lauren Ancel Meyers (2010) The dynamic nature of contact networks in infectious disease epidemiology, *Journal of Biological Dynamics*, 4:5,478-489, DOI: [10.1080/17513758.2010.503376](https://doi.org/10.1080/17513758.2010.503376)
- Victor L. Yu, Stephen C. Edberg; Global Infectious Diseases and Epidemiology Network (GIDEON): A World Wide Web-Based Program for Diagnosis and Informatics in Infectious Diseases, *Clinical Infectious Diseases*, Volume 40, Issue 1, 1 January 2005, Pages 123–126, <https://doi.org/10.1086/426549>