Public Health Demography QE Memo

Ethan Roubenoff

September 2020

This memo includes a description of the main topics, readings, and areas of research pertaining to my Qualifying Exam in Public Health Demography. This field exam focuses on two aspects of public health, as it relates to population: infectious disease modeling and social epidemiology. Infectious disease modeling in populations is concerned with how pathogens spread between hosts (humans usually, for my purposes), which is related to host contact networks; and pathogen-specific likelihood of transmission and mortality. Social epidemiology looks at how health and vulnerability is not equally distributed in a population; social inequality in access and participation in healthcare, employment, income, and infrastructure drives population-level inequality in disease and mortality. Contextualized within demography, Omran's (1971) theory of the Epidemiological Transition enhances demographic transition theory by describing how treating infectious disease is one phase of decreasing total mortality; it is preceded by better farming practices decreasing food scarcity and succeeded by treating degenerative and non-communicable diseases, like Alzheimer's and cardiovascular disease, respectively.

Infectious diseases spread through all populations— human, animal, plant— from one infected individual to the next. It is common for pathogens to be spread directly between individuals, but this process may be mediated by another species (called a vector) or a natural reservoir, like drinking water or sewage. While each pathogen has a unique course of action, infections can be grouped crudely by pathogen type (bacteria, virus, or parasite). Emerging (new) diseases develop from random genetic mutations and by cross-species contact. In order for there to be an outbreak, enough of the population must be susceptible to infection (and therefore, transmitting the disease); vaccination can help lower this proportion. Children and elderly are often the most vulnerable to infections, so age structure of a population is critical in evaluating mortality rates; conversely, understanding disease mortality and vulnerability is critical for mortality studies that consider population growth and structure. Population heterogeneity in baseline health, contact patterns, and other aspects of compositional demography can change the population-level course of a disease.

While much of the epidemiology literature is focused on infectious disease, many diseases are not spread through pathogens; smoking, obesity, and other 'lifestyle' related factors are responsible for cardiovascular disease, cancer, and other non-infectious diseases. The social epidemiology literature considers these risk factors as having exogenous causes: discrimination, segregation, food availability, and healthcare access. The social epidemiology literature reviewed here considers these causes also as risk factors for infectious disease, that would otherwise be seen as population heterogeneity. The social risk factors of infectious disease additionally include infrastructure, climate change, and chronic stress, among others.

Contents

1 Infectious Disease Modeling

- Key papers/books:
  - Keeling and Rohani, 2008: Modeling infectious diseases in humans and animals
    - Chapter 2: Introduction to Simple Epidemic Models
    - Chapter 3: Host Heterogeneities
    - Chapter 6: Stochastic Dynamics
    - Chapter 7: Spatial Models
    - Chapter 8: Controlling Infectious Disease
  - Chowell 2019: A novel sub-epidemic modeling framework for short-term forecasting epidemic waves
  - Bjornstad 2002: Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates Using a Time Series SIR Model
  - Camacho, 2011: Explaining rapid reinfections in multiple-wave influenza outbreaks: Tristan da Cunha 1971 epidemic as a case study
  - Fine 1993: Herd Immunity: History, Theory, Practice
  - Grenfell 2004: Unifying the epidemiological and evolutionary dynamics of pathogens
- Suseceptible-Infected-Recovered (SIR) models (compartmental models) are a frequent first-line model for infectious disease modeling. SIR models assume that individuals fall into one compartment at a time— they are either susceptible to infection, currently infected, or recovered/removed from the population— and use differential equations to model the total count in each category at each time. This model then considers how individuals flow from one

compartment to the next. A simple SIR model in a closed population (no birth or migration) could take the form:

$$\frac{dS}{dt} = -\beta S \cdot I \tag{1}$$

$$\frac{dI}{dt} = \beta \cdot S \cdot I - \gamma \cdot I \tag{2}$$

$$\frac{dR}{dt} = \gamma \cdot I \tag{3}$$

Where S, I, R are Susceptible, Infected, and Recovered proportions of the population, t is the time domain, and β and $1/\gamma$ is the transmission parameter and duration of infectiousness, respectively. Here, quantity βSI represents the number of new infections, parameterized as the number of susceptibles times the number of infecteds scaled by the transmission parameter. The basic reproductive ratio R_0 is β/γ (contact rate * duration of infectiousness), which quantifies the ratio of secondary (new) infecteds to primary (current) infecteds in a fully susceptible population— how many people one sick person will infect, on average. This parameter is roughly equivalent to the NRR in Demography.

- The SIR model is very flexible and can be easily extended to include pre-infectious periods, mortality and birth rates, age structure and contact patterns, and stochastic parameter distributions. SIR models can be adapted to include simple population heterogeneity, including age structure and risk groups, but it can be difficult to incorporate complex, dynamic groupings (see: STI models that divide populations into risk groups). One alternative is agent-based models, which simulate individual members of a population instead of aggregate tallies.

- The outbreak threshold is given by $1/R_0$, and indicates the proportion of the population that must be susceptible in order for an outbreak to occur. $1 - 1/R_0$ is also referred to as the herd immunity threshold (HIT). Maintaining vaccination rates above the HIT is critical in outbreak management (Fine, 1993)

- Grenfell et al (2004), in *Unifying the Epidemiological and Evolutionary Dynamics of Pathogens*, describe a set of pathogen behaviors in RNA viruses they refer to as 'phylodynamics.' RNA viruses are characterized by high mutation rates, large population sizes, and short generation times, which ultimately lead to pathogen population genetic evolutionary processes occurring on roughly the same time scale as epidemics. Pathogenic strains are created through mutation, and their proliferation in host populations determined by the host's immunological response. The high viral mutation rates are modulated by two immunological processes: the selective immunological pressure on the pathogen and the transmission bottleneck between hosts. Host population heterogeneity can further modulate transmission.

  In general, the most pathogenic strain of a virus will proliferate in the population— that is, the one that is most readily spread to other hosts will dominate. Low immune pressure will result in low selection, such that many strains can be equally pathogenic as determined by random mutations. High immune pressure results in high selection forces, and fewer, but potentially more severe strains will circulate, as these strains evolve specifically to survive the host's immune response. Weaker random mutations will be out-selected by the immune response and stronger mutations will proliferate. Advantageous transmission traits include longer duration of infectiousness and increased shedding volume (including coughing severity, among others).

2 Infectious Disease Modeling Case Studies

A number of case studies are considered in these articles, including Measles, Mumps, Rubella, Smallpox, SARS, Covid-19, Dengue, Malaria, and others. A few are highlighted below.

2.1 Measles

- Key papers:
  - Bjornstad, et al 2002: Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates using a Time Series SIR Model
  - Keeling and Grenfell, 1997: Disease Extinction and Community Size: Modeling the Persistence of Measles
  - Smith et al 2018: Agent-based models of malaria transmission: a systematic review
  - Mina et al 2019: Measles virus infection diminishes preexisting antibodies that offer protection from other pathogens
  - He et al 2010: Plug-and-play inference for disease dynamics: measles in large and small populations as a case study
  - Fine 1993: Herd Immunity: History, Theory, Practice
- Measles is one of the most frequently modeled diseases. It is a viral infection with one of the highest R_0 parameters (number of new infections one sick individual is likely to cause) of any commonly modeled disease— estimates range from 12-18 (Fine, 1993)— meaning that it is incredibly easy to transmit. Measles is aerosolized and is spread through droplets expelled during coughing, sneezing, and even breathing. Airborne droplets that carry measles can remain active in an environment for hours. Measles is notable in that outbreak cycles are driven by the population age structure as the proportion of uninfected children govern outbreak timing. A key assumption in measles modeling is that since measles is so transmittable, by adolescence, almost every child will be immune to measles either through exposure or vaccination. Having enough susceptibles to sustain an outbreak— known as the Critical Community Size— is then primarily driven by new births. Since an R_0 of 12-18 corresponds to a herd immunity threshold of 83-94% (ibid), outbreaks can occur when susceptibles make up this proportion of the population; since all individuals will be immune to measles by adolescence, outbreaks can occur once unexposed children make up this proportion of the population. At reasonable birth and death rates, this corresponds roughly to 2-3 year outbreak cycles in unvaccinated populations. Agent-based models (Smith et al 2019) can incorporate population heterogeneity and Partially Observed Markov Process models (POMP, He et al 2010; also known as state-space models) can capture additional variability on parameters not possible in a traditional SIR model.

2.2 COVID-19

- Key papers:
  - Davies et al 2020: Age-dependent effects in the transmission and control of COVID-19 epidemics

  - Miller et al 2020: Disease and healthcare burden of COVID-19 in the United States
  - Kucharski et al 2020: Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study
  - Kissler et al 2020: Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period

- Over 25 million cases of COVID-19, also known as SARS-CoV-2, have been confirmed globally as of August 31st, 2020; 850,000 deaths have been caused by the virus. A large proportion of these deaths have been confirmed in the elderly, who seem to have more severe infections and subsequently higher testing rates (Davies et al, 2020). COVID-19 is spread primarily by droplets expelled through coughing, meaning that more serious cases, where infecteds are coughing more frequently and require more care, are also more contagious (ibid). Miller et al 2020 estimate that the healthcare burden is highest among rural areas, many of which are without a major hospital despite lower population density. Ensuring equitable distribution of care and resources is critical in order to not exceed the care capacity of hospitals. Proper self-isolation and full contact tracing can possibly reduce transmission by 64% (Kucharski et al, 2020). Estimates of R_0 are between 2-2.6 (Kissler et al 2020; Kucharski et al 2020) without intervention; post-intervention R_0 varies dramatically depending on control measures. A major concern with COVID-19, as with many highly pathogenic diseases, is the possibility of 'superspreading' events: activities where people in close proximity have a high potential to cause a large outbreak from only a single or few infected individuals. Intervention strategies like event capacity limits, outdoor/ventilation requirements, and enforcing physical distancing can limit the likelihood of superspreading events. Kissler et al 2020 assess a number of situations based on the SARS-1 epidemic in 2002 and MERS in 2012 (both SARS and MERS are coronaviruses with similar dynamics), including social distancing, immunity, and seasonality in the possibility of endemic cycles.

3 Disease Seasonality

- Key readings:
  - Grassly and Fraser, 2006: Seasonal infectious disease epidemiology
  - Alitzer et al, 2006: Seasonality and the dynamics of infectious diseases
  - Metcalf et al, 2009: Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen

- Seasonal dynamics of disease are complex and highly specific to both the infection and the population of study. Grassly and Fraser (2006) enumerate a series of causes: (a) survival of the pathogen outside the host, (b) host behavior, (c) host immune function, and (d) vector (non-human hosts, like mosquito or rat) abundance. Alitzer et al discuss weather as a primary driver of seasonal forcing in host social behavior and aggregation— including holidays spent in proximity indoors during bad weather— in addition to vector (mosquitos, ticks, and other non-human hosts) abundance during warm, wet weather. High humidity can affect droplet behavior, for example increasing the transmissibility of the flu. Simple seasonal forcing can

be modeled using a sinusoidal function or, in more complex models, using a Fourier decomposition (Bramness 2015). Metcalf et al 2009 measure the relative magnitude of seasonality as the variance of the transmission parameter, and estimate mumps, pertussis, and measles as substantially more seasonal and than diptheria and scarlet fever.

4 Population Heterogeneity

- Key readings:
  - Pitzer et al 2009: Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics
  - Funk et al 2010: Modelling the influence of human behaviour on the spread of infectious diseases: a review
  - Mossong et al 2008: Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases
  - Ferrari et al 2013: Think globally, act locally: the role of local demographics and vaccination coverage in the dynamic response of measles infection to control
  - Bansal 2007: When individual behaviour matters: homogeneous and network models in epidemiology
- Heterogeneity models consider cases when susceptibility to and severity of infection is not universal. Relevant to infection, populations can be heterogeneous in age distributions, contact patterns, and behavior, which can introduce uncertainty in transmission and mortality parameters. In modeling heterogeneity, populations in SIR models can be split into risk groups or stratified by age; transmission and mortality rates can be estimated as distributions with uncertainty instead of point estimates. Bansal et al find that considering the network connectivity of individuals improves modeling accuracy notably, but not as substantially as previously speculated. For a series of different populations, they find that an exponential degree distribution of contact networks fits the observed dynamics best. Mossong et al conduct a large-scale network survey of disease-relevant contact patterns (skin-to-skin contact or two-way conversation), and estimate that 5-19 year-olds are most susceptible in the case of an emerging infection. Funk et al create a taxonomy of prevalence-elastic behavior, whereby individuals modify their behavior when they become aware of a disease. Information can be either belief or prevalence-based and be locally or globally available; behavior can either change the disease course in individuals, change the model parameters, or change the contact structure.

5 Spatial Epidemiology

- Key readings:
  - Elliot and Warternberg, 2004: Spatial Epidemiology: Current Approaches and Future Challenges
  - Robinson, 2012: Spatial dynamics of airborne infectious diseases

- Spatial epidemiology concerns how infectious diseases play out over a two-dimensional surface. Populations spread out over geographic space have potentially heterogeneous environmental factors, affecting seasonal forces between regions. Healthcare and other resources may not be evenly distributed. Finally, disease spread is a fundamentally spatial process: infecteds must come within proximity to susceptibles in order to transmit disease, naturally leading to disease clusters. In my Spatial Methods exam, I will give a more in-depth description of the technical aspects of disease mapping.

6 Social Epidemiology

- Key readings:
  - Link and Phelan, 1995: Social Conditions As Fundamental Causes of Disease
  - Dockery et al, 2006: An Association between Air Pollution and Mortality in Six U.S. Cities
  - Galea et al, 2011: Estimated Deaths Attributable to Social Factors in the United States
  - Glymour et al 2014: Socioeconomic Status and Health, in: Social Epidemiology
  - House et al 1994: The Social Stratification of Aging and Health
  - Case and Deaton, 2017: Mortality and Morbidity in the 21st Century
  - Weiss and McMichael, 2004: Social and evironmental risk factors in the emergence of infectious diseases
- Social epidemiology concerns the social factors behind health disparities. Link and Phelan (1995) and later Phelan et al (2010) contextualize lifestyle health factors, like diet, exercise, and smoking within the larger societal framework that causes individuals to make health-related decisions. Intervention strategies that neglect to consider 'the causes of the causes,' as the authors write, may be largely ineffective. Galea (2011) estimates that:

  > Approximately 245000 deaths in the United States in 2000 were attributable to low education, 176000 to racial segregation, 162000 to low social support, 133000 to individual-level poverty, 119000 to income inequality, and 39000 to area-level poverty.

- Poverty-linked inequities in healthcare access and systemic discrimination by healthcare professionals leads to worse treatment outcomes for people of color and low-income people. Further, lack of sick days for workers and their children can exacerbate the spread of disease in working communities.
- Direct effects of poverty on susceptibility to infectious disease occur either through exposure, poor home construction, or decreased social investment. Weiss and McMichael (2004) recognize poor sanitation in slum housing and the effect of climate change/forest clearance as a potential source of emerging diseases.
- Case and Deaton's seminal work on the opioid crisis (a summary chapter by the authors is included) shows that all-cause mortality among some non-hispanic Whites has been rising since 1997 in relation to cumulative, intergenerational disadvantage— termed 'deaths of despair.'

The primary driving causes of death for this group are drug overdoses, suicides, and alcohol-related liver mortality. Note that Black Americans and nonwhite Hispanics have actually seen an increase in life expectancy over this period.

- Janet Currie's work looks at how health is socially inherited from mother to child— not through a genetic process, but by processes around pollution, educational disadvantage, and environment.

7 Social and environmental risk factors in the emergence of infectious disease (Weiss and McMichael, 2004)

Weiss and McMichael write on how new emerging diseases are exacerbated by human practices, historically from farming practices and trade travel but more recently linked to economic pressures from rapid industrialization, climate change, and high population density. Emerging diseases often first jump the 'species barrier' from animals to humans during periods of excessive contact. The agricultural revolution about 10,000 years ago enabled humans to domesticate livestock and other animals, leading to the first 'crowd' infections due to animal proximity—measles and smallpox, about 7,000 and 4,000 years ago, respectively. As global trade and military networks proliferated, widespread outbreaks became possible; the first outbreak of typhus may have been spread by the Peloponnesian war in 430 BCE. Colonization and exploration beginning in the 16th century spread new diseases to the Americas.

Per the authors, we are currently in the fourth historical transition of globalization as it relates to epidemiology: "Urbanization, dense and usually impoverished peri-urban settlements, social upheaval, air travel, long-distance trade, technological developments, land clearance and climate change all influence the risks of infectious disease emergence and spread" (Weiss and McMichael, 2004). As humans encroach further upon wild habitats, the possibility for inter-species transmission increases; bushmeat practices are suspected of the Ebola outbreak and HIV's transmission to humans. Wet markets were the source of the SARS outbreak in 2002, as many live animals were kept in close contact to humans with poor ventilation and sanitation. Poor home construction without adequate sanitation can lead to outbreaks of dengue and cholera in large, high-density urban slums. Anthropogenic climate change can intensify warm, wet conditions ideal for bacterial and vector growth; natural disasters that displace people put them at risk for outbreaks in temporary housing.

As Braveman and Gottlieb write in the Social Epidemiology literature, it's necessary to "consider the causes of the causes:" interventions to stop emerging diseases need to account for the socio-economic and political forces that create the conditions in the first place in order for them to be successful. Favorable conditions for emerging and endemic diseases—multi-species proximity, poor sanitation, habitat destruction—must be considered as products of social conditions. Intervening for adequate sanitation and safe food practices can limit the spread of emerging and endemic diseases in combination with effective medicine and vaccination dissemination. Political and social apathy and stigma, especially in the case of HIV, must addressed in order for successful treatment delivery, safe practice awareness, and harm reduction.

The authors are focusing primarily on emerging diseases in the developing world, but in the United States inequities in healthcare access, treatment outcomes, employment, and other social factors can influence disease transmission and mortality. Osypuk and Avecedo discuss how geography of opportunity—as a result of racial residential and economic segregation—leads to worse treatment outcomes for communities of color and low-income communities (Osypuk and Avecedo, 2010). Chronic stress, possibly resulting from poverty and discrimination, can inhibit immune response (Sergerstrom and Miller, 2004). Insuring that sick employees are able to take time off (as well as when their children fall sick) can limit the effective contact rate of transmission. Affordability of basic medicines and primary care physicians can reduce duration of infectiousness and ultimately mortality. Although the social epidemiology literature is focused primarily on lifestyle-related health—diet, smoking, exercise— the same risk factors for lifestyle diseases can make populations have more frequent and serious outbreaks.