SFI TRANSMISSION

COMPLEXITY SCIENCE FOR COVID-19

STRATEGIC INSIGHT: **The disease models used to guide policy for the COVID-19 pandemic must capture key complexities of transmission.**

FROM: **Andrew Dobson, Princeton University; Santa Fe Institute**

DATE: 6 April 2020 **NO:** 005.1

The study of infectious disease dynamics can be divided into two areas. The first is pathology, which focuses on the changes in the hosts due to the presence of the pathogen (e.g., dry cough, high fever, and acute respiratory syndrome). The second is transmission, how the pathogen moves from an infected individual to an uninfected one in order to initiate a new infection. In sharp contrast to pathology, we almost never see transmission occurring — we can only infer that it has occurred when a healthy individual develops signs of infection. We then need to trace back their activities in the hopes of identifying an individual with symptoms that led to the transmission event.

This post will supply an overview of the problems that beset epidemiologists when we try to measure transmission, while familiarizing the reader with some key models used to measure transmission, and prevent it.

Transmission need not be confined to infectious diseases, culture, and ideas are transmitted and even franchised — advertising is an industry focused on transmitting ideas that convince people a commodity is essential to their quality of life. Teaching may be the ultimate form of cultural transmission; from kindergarten right on down to graduate school and beyond, ideas are developed into presentations that transmit knowledge from one individual to others. Similar mathematical frameworks are used to describe the ways in which ideas and pathogens are transmitted. Here I will focus on pathogens, and only occasionally indicate applications to cultural transmission.

Models that are used to describe the dynamics of infectious diseases fall into four broad categories: a major split is into microparasites and macroparasites; a second split is into VBDs, vector-borne diseases, STDs, sexually transmitted diseases, and OIDs, other infectious diseases. The first division acknowledges that some pathogens are too small to be accurately counted and their dynamics are best described by a mathematical framework that divides the host population into Susceptible, Infected, and Recovered/

Removed or Resistant hosts, depending on whether they died, recovered, and rejoined the susceptible hosts, or developed immunological resistance that protects them against future infection. The macroparasites are physically large enough to count: ticks, fleas, and worms. The pathology they create is a function of their abundance; having one or two ticks or fleas is an annoyance, but carrying around a large number of worms in your gut is a major constraint on your ability to grow physically and intellectually. Models for these pathogens need to consider the frequency distribution of parasites across the host population; these are inevitably aggregated, with a large parasite population relative to the host population, and most hosts have only minor infections. The macroparasites produce complex but short-lived immunological responses that only weakly protect them against future infection. However, they can modulate the efficiency of immune responses produced in hosts concomitantly infected with microparasites, and may also significantly reduce the efficacy of vaccines against these pathogens.

The second division, into VBDs, STDs, and OIDs, takes us back to transmission, with examples of each form known for both macro- and micro-parasites. Malaria is the poster child for a vector-borne diseases. The pathogen sequentially uses mosquitoes and humans (or other vertebrates) as hosts, converting each of them from susceptible to infectious individuals. Technically it may be possible to identify which mosquito bite gave rise to an infection, but only a very small proportion of mosquitoes are infected. As a result, we often experience many mosquito bites before getting infected, and we are again stuck with the problem of who the mosquito bit to acquire its infection.

Sexually transmitted diseases have many similarities with vector-borne diseases, although transmission mainly occurs alternatively between each sex within the same species. The problem of when someone was infected and by whom again arises, and it is never apparent that one has been infected until after the incubation period, when symptoms appear. STDs can also be vectored. The smut fungi transmitted between the flowers of many plants are effectively STDs transmitted by pollinating insects. Heterogeneity in transmission is driven by the frequency distributions of numbers of new sexual partners; this variability may be fairly low in the case of insects transmitting smut fungi, but could be significant for human STDs in individuals with many partners. These heterogeneities hugely increase the rate of spread of the pathogen; a big early success in the AIDS epidemic was changing people's behavior by encouraging them to have few new partners. This reduced the variance and significantly reduced net transmission rates.

The OIDs (Other Infectious Diseases) are transmitted by free-living particles, expelled from the infectious hosts, that directly infect the susceptible hosts. The duration of time these infective stages last in the free-living stage is crucial — in the case of influenza and SARS-COV-2, coughing and sneezing releases a cloud of infective particles into the air that may only be infectious for a couple of minutes. In contrast, bacteria such as anthrax

produce spores that may survive in the soil for decades. The parasitic worms that live in the guts of most vertebrates produce eggs that can survive for weeks to months. This creates a bifurcation in transmission mode, when infectious stages are short lived, then transmission from one infected individual is a function of the density of susceptible hosts in their vicinity. You infect many more people when you sneeze on a crowded subway car, than when you sneeze in a nearly empty one. This form of transmission, usually called density-dependent transmission, is assumed to be linear, with each infected individual infected a fixed proportion of the susceptible individuals in their vicinity. In contrast, pathogens with long lived free-living stages, as well as vector-borne and STDs, tend to have saturating transmission functions: mosquitoes have to digest between blood meals, and even Casanova had to rest or dine occasionally. This form of transmission is modeled by frequency-dependent functions with the product of susceptible and infected hosts appearing in the numerator, and the total population in the denominator.

How do pathogens get transmitted between different host species?

Most animal and plant species harbor a significant diversity of parasites and pathogens — it may be there are ten times as many parasitic and inquilinistic species as free-living species, possibly more. Although we focus significant energy on conserving charismatic free-living species, each time one of these goes extinct, it takes with it a significant number of species that depended upon it for their existence. Some of these pathogens will be specialists that are usually only found on that host species, while others are generalists that can also use closely related species, or ones with similar feeding habits. In the majority of cases the parasites only have a minor effect on their hosts and their pathology is relatively low unless the host becomes stressed due to other changes in its environment, such as a reduction in food supply or change in local climate.

Occasionally, the pathogen infects the wrong host. This can occur when a vector feeds on a novel host that is more abundant than its normal host (as occurs when domestic livestock or humans are more abundant than native species). Animals captured for food or the pet trade often become stressed, leading to them releasing significant numbers of infective stages that can contaminate humans involved in the trade. Usually, the pathogen is unable to survive in the new hosts, as the cells it needs to infect in order to replicate are absent. However, when a pathogen does manage to infect novel cells and start replicating, this can then lead to the emergence of a new disease. This seems to be what is happening with the COVID-19. Genetic evidence suggests its natural host is a bat species, or possibly a pangolin. These species have very different physiology from humans (and most of our domestic livestock species); we rarely see any overt pathology in bats infected with these pathogens. This can change dramatically when the pathogen finds itself in the wrong host.

Bats have very different immune systems from other mammals, likely as a consequence of their ability to fly. Humans and other non-volant mammals produce the B-cells of their immune system in their bone marrow. Because bats fly, they have hollow bones; the only place they have bone marrow in their pelvises, so they produce B-cells at much lower rates than other mammals. Similarly, active flight raises their body temperature to levels akin to fever in non-volant mammals, possibly constraining viral growth. Bats also do not store fat as it compromises their aerodynamic ability. Instead, they can enter torpor to get through periods when food resources are low. These all act as constraints on viral pathogens that disappear when the pathogen finds itself in a novel host whose immune response may interact with that pathogen in ways that are detrimental to both the host and the pathogen.

The dynamics of generalist pathogens provide important insights into the transmission dynamics of pathogens in structured human populations. Consider the dynamic of a pathogen that can infect multiple host species, each of which has a different body size and thus different birth and death rates and population densities. The species with the smallest body size will have the highest birth and death rates and population density. The largest will have the opposite. If the pathogen follows simple dynamics, with within-species transmission far exceeding between-species transmission, then each host will interact independently with the pathogen and each will exhibit its own epidemic cycles: large and frequent outbreaks in hosts with low body mass, and slow, less dramatic cycles in larger hosts. As we increase the relative rates of between-species transmission, these cycles will die out. Additionally, any tendency for epidemic outbreaks to occur is buffered by the pathogen's constant jumps between host species, preventing any one species from becoming too abundant. If we increase between-species transmission to levels where it matches within-species transmission, then the small species can use the pathogen to drive the larger species extinct; small species are abundant and recover quickly from outbreaks, while rarer large species cannot recover from frequent epidemics. Ultimately, only the smallest species survive, and they revert to the epidemic behavior they exhibited when between-species transmission was rare.

This exercise suggests that understanding rates of between-species transmission is an additional, vital component of disease dynamics. To that end, Who Acquires Infection From Whom (WAIFW) matrices provide a framework to examine how the pathogen moves between different groups of hosts and allows us to identify which section of the population acts as a reservoir to maintain the infection and which are subject to spillover events.

The matrices were originally developed to study the transmission of pathogens such as measles and rubella between different age-classes in human populations. The population in any area can be divided into pre-school children, kindergarteners, middle school, high school, college, etc. In some ways, it will resemble the matrix for our hypothetical multi-species example; most of the transmission is within the same age class, because most of our interactions are with people are own age. These interactions may become more

diffuse as we get older, and off-diagonal elements will become important as children interact with their parents and grandparents. If we could quantify the structure of these matrices, we would know a lot about how pathogens spread in populations. However, a problem instantly appears. The data on disease exposure we have for each age class increases linearly with the number of age classes into which we divide the population. The number of transmission elements in the matrix increases as the square of this number, so we have created an unsolvable problem. Unsolvable, unless we can assume similar rates of transmission in different age classes, or if we can gather independent social behavior on rates of human interaction. Both approaches have been used. The latter was particularly instructive, as it gathered data across multiple European cultures with different approaches. There was a curious similarity in the structure of the interactions; it would be helpful to expand these surveys to nations where different sex and age-dependent interactions might modify the structure of the matrices.

These matrices are central to the social distancing now being put in place for the COVID-19 epidemic. Essentially, we are trying to massively reduce the strength of interactions within each age class and completely remove the interactions between age classes in order to protect older people who seem to be more susceptible once infected. Quantifying the structure of these transmission matrices is crucial for understanding the size of the epidemic and how to control its spread.

Read more posts in the Transmission series, dedicated to sharing SFI insights on the coronavirus pandemic: santafe.edu/covid19