APS COVID Webinar: September 23, 2020

**Epidemiology of COVID-19**

**Marc Lipsitch**, professor of epidemiology and director of the Center for Communicable Disease Dynamics at Harvard University

- The data is terrible, measurements are conflicting, tools are ill-equipped, and thus the data is often not comparable across places and time
  - Data quality also hindered by unmeasured data points and changing delays
  - Delay distribution is not consistent over time
    - Showed plots which demonstrate that delay between symptom onset and diagnosis has been changing with each week since the pandemic started
- Part of the solution = Nowcasting
  - Given how many cases we know about today and in recent past, how many cases will we eventually know about that occurred today and in the recent past?
  - Nowcast ‘learns’ the delay process
  - NYC department of health has been using a method of Nowcasting to fill in estimates of current cases
- It is now clear that travelers initially spread the virus, often undetected
  - Rough linear relationship between number of cases and daily air travel volume
- Spread has depended on:
  - Time of introduction (and some bad luck)
    - Branching process simulation
      - Large uncertainty in the inputs to the simulation
      - Initial studies found to not be able to narrow the uncertainty at all
        - BUT turned out we really needed to expand uncertainty and broaden hypotheses in order to best advice communities
      - Results of this work contributed to the urgency of the NYC response
  - Mobility
    - Mobility study with Facebook data examined counties in Mass, NY, Florida
    - Found heterogeneity of mobility across time between the states
  - Individual responses and risk factors
  - Seasonality (likely mild effect)
    - Compared to 2 common cold coronaviruses [Kissler et al. *Science* 2020]
      - ~20% reduction in summer from winter in these common colds
    - COVID-19 hasn’t looked as seasonal because it has so many hosts to infect
      - Others have a depletion of hosts in summer months
  - A huge number of other possible factors
    - Prior exposures?
    - Demography?
    - Use of the BCG vaccine for TB?
    - T cell cross-immunity
  - Pre-symptomatic transmission is common
    - ~40% of transmission occurs prior to symptom onset (pre/asymptomatic)
      - Was a rough estimate but seems to have stood the test of time thus far
      - This is a challenge for contact tracing
• Plot that shows strong evidence that individual quarantine is much more effective than just active monitoring, but that neither is really effective unless you have a large number of contacts traced (~75%) (Peak et al. 2020 *Lancet Inf Diseases*)

• Contract tracing data from NYC
  o Only about ~25% of contacts are likely found (told about and able to be contacted), and usually already in the midst of their disease
  o It has worked well for other diseases, but it is likely only making a modest contribution to control in this case

• Infection fatality rate: ~0.07% (not a constant!)
  o IFR increases dramatically with age
    ▪ Log-linear increase up to ~10% between old and young
  o Age, race, and socio-economic status have been shown to have significant influence on IFR

• Most of the world has experienced far less than 20% of population infected, already with catastrophic consequences in many places

• Vaccines:
  o SEIR model
    ▪ Susceptible, exposed, infected,
    ▪ Age-stratified SEIR models allow us to ask more targeted questions
  o Lots of unknowns, but know that vaccine will be scarce
  o Modeled different prioritizations for vaccine distribution:
    ▪ For mortality:
      ▪ Vaccinating the oldest is the best strategy
      ▪ Nearly imperturbable.. no matter the assumption, vaccinating the oldest is the most effective at reduction of deaths (robust across countries, etc)
    ▪ For infections:
      ▪ Vaccination the young adults is most effective (modest difference, can change order of best strategies using different assumptions)
  o Might be a benefit from pairing serology tests with a vaccination strategy
    ▪ Test for antibodies (assuming these are effective) and only vaccinate seronegatives

• **Big open questions:**
  o Where does transmission occur? How many activities can we do safely?
  o What are the predictors of superspreading and how can we prevent them?
  o Who is at risk for complications?
  o Will vaccines work at all? If so, for what outcome?

• **A few questions for physicists (areas our expertise could be useful):**
  o Important to have collaboration with relevant epidemiologists/virologists/etc who understand the data very well
  o Role of airborne transmission and ventilation
  o Biophysical mechanisms of seasonality
  o Better ways to account for uncertainty in data
  o Better understanding of the mapping of mobility to transmission
  o Structural biology etc

**Question and Answer:**
• Can you recommend some review articles with relevant model parameters such as the fatality rate as a function of age, asymptomatic population, incubation time, etc? (anonymous)
• Can you discuss the role of uncertainty in case modeling? What are you doing to model uncertainties in your current approaches?
  o Uncertainty about both past and future!
  o Past: Have an observation model on top of virus model, fit to observed data
    ▪ Difficult, and field has not converged on a universal solution to the possibility that we are not detecting a large portion of cases
    ▪ Deconvolution approaches seem to be the best option for many models
  o Future:
    ▪ Need to include a lot of prediction of human behavior! This is a human model, not just a natural model
    ▪ Must be careful to phrase everything as a conditional prediction
• How do you determine and measure the mobility rate and its reduction?
  o Several data sources, most depend on cell phones
    ▪ Login location for facebook
    ▪ Movement between cell towers your phone is pinging
    ▪ Google dataset (not sure how they generate)
    ▪ Third party vendors making graphs based on where you move with a mobile device
• Where have you seen an effective use of the kinds of models you talked about? Is anyone actually using the information we/you are providing?
  o Marc has advised a few countries, usually more general advice, not specific discussions of his models
  o UK is the epicenter of this type of modeling, so they have plenty of good advice, but don’t seem to always be following it
  o Marc doesn’t claim to have a whole lot of insight into the policy-making process and who is using what models – we should follow up here!
  o There is a pretty strong divide between countries that have listened to the science and those that haven’t (many strong actors)
    ▪ It’s rarely the details of the models that matter, but the broad conclusions
• What is role of big data, machine learning for driving epidemiological picture?
  o Really believes in understanding the mechanisms; thus machine learning may have more limited function in this area
  o BUT! Really great work has been done to do more timely predictions of cases and other scattered examples
• Incidence and mortality rates have changed with time...thoughts on the impact of modeling on treatments, improving these numbers?
  o Not a lot
  o Changes in who is getting infected are a big driver of these changing numbers
  o Treatment improvements are beneficial in moving people out of hospitals faster, then improving treatment numbers
• Some reporting that you can get re-infected with the virus...
  o 3 of these cases the last time he checked
  o Might end up being uncommon to get re-infected, but very uncommon to get a severe second case
The discussion on data quality was really interesting... Do you think COVID will be an impetus to try and improve data capture policy for the future?

Can/will epidemiology tell us anything about possible long term side effects of COVID?