The Coronavirus Pandemic Working Group
for Models of Total Harm Minimization

COVID-19 Risk
Modeling Options, Conclusions & Concerns to Date

David L. Katz, MD, MPH; Roger M. Stein, PhD; Wesley Pegdan, PhD;
Maria Chikina, PhD; Dina Aronson, MS, RD

April 7, 2020

Table of Contents

THE PURPOSE OF THIS UPDATE 2
- WHERE WE ARE TODAY 2
- WHAT WE NEED TO HELP YOU 4

KEY INTERIM OBSERVATIONS 5
THE DATA CRISIS 6
- MOST COMMON DATA NEEDS IDENTIFIED 7

KEY POLICY IMPLICATIONS 8
Appendix A: Selected COVID-19 live data sources 9
Appendix B: Estimated proportion COVID-19 infections leading to hospitalization (fr. Verity, et al. (2020)*) 9

THE PURPOSE OF THIS UPDATE

The COVID-19 Working Group on Models of Total Harm Minimization has been discussing, in multiple ongoing workstreams, a number of practical approaches for using quantitative tools to inform decisions about public health and economic policy. This document summarizes several of those models and projects that are being undertaken by this group.

This update is made up of three parts:

- This document, summarizing the state of this working group's work and observations;
- A companion Model Atlas that provides a quick reference to the various models that members of the working group are developing; and
- An archive file containing papers describing the various models, as referenced in the Atlas.

WHERE WE ARE TODAY

Accompanying this summary is an atlas of pointers to the various efforts that our working group has underway, along with underlying source documents, where available. We have attempted to make this atlas both self-contained and easy to read, without requiring you to fully digest the underlying research to determine whether it may be of use to your current work. We are available to discuss these in more detail if helpful to you.

Importantly, the scope of this summary, and the accompanying atlas, only describes the current work of this group. We strongly encourage you to also seek advice from other of the many talented groups also working on these problems.

To date, our discussions and research have led us to five preliminary broad conclusions:

1. Even within our own group, there is a richness of models and modeling approaches that are under development. The methodologies for these models are heterogenous as are the units of analysis (e.g., individual vs. population), level of detail and objectives. **However, there does not appear to be any single model that captures all of the facets of the epidemiological and economic implications of current or prospective policy decisions.**

2. The data required to parameterize these models is similarly heterogenous. More challenging, from a practical perspective, is that there is in general a lack of good quality empirical information about almost every aspect of COVID-19 in the US. **Aggressive data collection and integration will be critical to bringing the current models on-line. This data collection must include randomized testing of a subset of the general population.** However, even with such efforts, it may be some time before

this data can be used to inform current and future models. We estimate this time to be on the order of months or quarters.[1]

3. There do not currently appear to be any specific measures of "acceptability" for COVID-19 models for policy use. While it is not the case that these models can be validated in the same way that, e.g., medical diagnostic tests may be, **there is still the need to establish a set of baseline criteria for model performance if they are to be used as a central basis for policy decisions.** This is necessary both to size the potential magnitude of model errors (and their implications), but also to determine how to integrate information from differing models when they disagree. This also requires the data described in (2), above.

4. Although the objective of the group is to provide policymakers with tools to help determine how and when to bring portions of the workforce back on-line while also managing public health, **at this time there is no obvious mechanism for a coherent integration of economic outcomes and health outcomes** using a consistent or comparable framework and assumptions.

5. Given the current state of the modeling work in progress on the one hand, and the lack of reliable data (or, in many cases, *any* data) on the other, we recognize that **policy actions that must be taken in the short term will likely need to either accept substantial model risk and uncertainty, or, alternatively, choose approaches that reduce the dependence of policy on model outputs, until the models and data can be refined**.

While these challenges are a reality for modelers and policy makers, these challenges do not imply either that we cannot build and use these models, or that such models must be somehow perfect in every detail. We can and they don't. However, given the stakes, we favor framing policy to accommodate the current limitations modelers face. Policy design needs should inform modeling exercises, with models used for support, wherever possible.

[1] In addition to actually collecting data, there are a number of other critical activities that we must undertake to prepare the data for modeling. Some of these include: evaluating accuracy and quality of the data; rationalizing data from different sources (e.g., definitions of codes, disambiguation of terms, etc.); adjusting some data for sampling bias; and, most fundamentally, triaging the data needs and collection costs.

WHAT WE NEED TO HELP YOU

We are seeking information about how to make our work most useful to you. Guidance on the following questions, if you can provide it, would be invaluable in achieving this end. Three key perspectives will allow us to focus our efforts where you most need us to:

1. For which specific policy objectives, questions and planning exercises do you need answers most urgently at this time? Are there key trade-offs that you are considering with respect to these?
2. What outputs from models would you find most helpful to inform these policy decisions?
3. For which of these outputs would model errors be most concerning with respect to executing these objectives? Said differently, are there outputs for which simple directional indications would be sufficient versus others where detailed estimates and prediction intervals are required?

At the highest level, this work presupposes that there is more than one way for COVID-19 to devastate health and cost lives: via the direct path of infection, and via the indirect path, mediated via societal upheaval and degradations in social determinants of health.

The explicit aim of this effort is to minimize the total harms incurred along both of these pathways. The models assembled thus far suggest advantages in this regard with an approach to interdictions of contagion that vary with risk tier. Those models are subject to change with more and better information about risk tiers in a specific population.

Further, the harms of infection are immediate, and thus so too are the benefits of interrupting viral transmission. The harms mediated by social determinants of health may extend over a longer and more variable time horizon, and thus so, too, do the benefits mediated via restoration of livelihoods and normalcy.

The economic benefits, apart from the health benefits, of a phased return to normalcy will of course be more immediate, but the preferential focus in this effort is on the net gains achievable for public health.

KEY INTERIM OBSERVATIONS

1. Social distancing, quarantining, and proper hygiene appear to be effective in slowing spread in the short term. However, a total halt to all commercial and social activity (horizontal interdiction) is almost certain to have severe negative impacts on the economy.

2. Current methods of quarantining and social distancing mitigations will likely successfully delay, but not significantly decrease, infection and death rates.

3. The effectiveness of a quarantine measure is only as successful as societal compliance.

4. Widespread, accurate, and efficient **testing** is crucial, including data on the recovered and asymptomatic individuals.

5. To date the populations that have reported the most success in containing the virus are those that have implemented an expansive and responsive testing program.

    a. **Iceland**
    b. China
    c. **South Korea**
    d. Singapore
    e. United Arab Emirates

6. To disrupt spread, healthcare workers should be **screened regularly**.

7. A robust modeling and planning strategy must take a long-term perspective in order to inform policy and avoid a reoccurrence of the current *ad hoc* response.

8. Risk stratification focuses mainly on population separation.

    a. Generational distancing is a first step, but identifying those at risk is equally important

    b. Need robust data on health risk and COVID-19 outcomes

9. Risk stratification approaches must contemplate more than simply the risk of infection, hospitalization and ventilation of an individual patient, but also the risk of that patient's transmission of COVID-19 to other otherwise healthy individuals. This is complex because the risk of a "False Negative" (a cleared individual who is actually infected)

puts not only that individual at risk but also potentially many others who would not have become ill without contact with the infected individual.

10. Current limitations due to incomplete data must be factored into models and recommendations; these include data on partial immunity/re-infection/virus mutation rate, vaccine and treatment development status.

11. Cohabitation remains a major challenge.

12. The immediate goal of preventing overwhelming demands on medical care systems may be addressed effectively, at least in part, if interdictions of viral spread are preferentially directed to the higher-risk populations most likely to impose those demands, namely, those older and with prior, major health conditions; while those with conferred immunity return to the workforce in a measured fashion.

THE DATA CRISIS

It is crucial to have an accurate number of current and past positive cases of COVID-19 (“denominator data”) as well as information about the infected population and various components of the healthcare infrastructure.

Central to these needs is the demands of almost every model to be able to assess prevalence, morbidity and mortality in a realistic – and representative – fashion. To date, most testing in the US has focused either on individuals who have presented as ill at medical facilities, or individuals who are known to have come into contact with infected persons. This data, while informative anecdotally, does not allow modelers either to make estimates of the true risk of infection for the general population or to understand the morbidity and mortality associated with COVID-19 infection. Thus it cannot be used, in its current form, to provide guidance on the impact of proposed policy measures.

Random sampling, testing, and surveys on the other hand, will accomplish this. Data collected from a representative sample of citizens, along with details on the individuals’ health history and demographics, can then be used to develop informative models that will allow us to better understand issues such as:

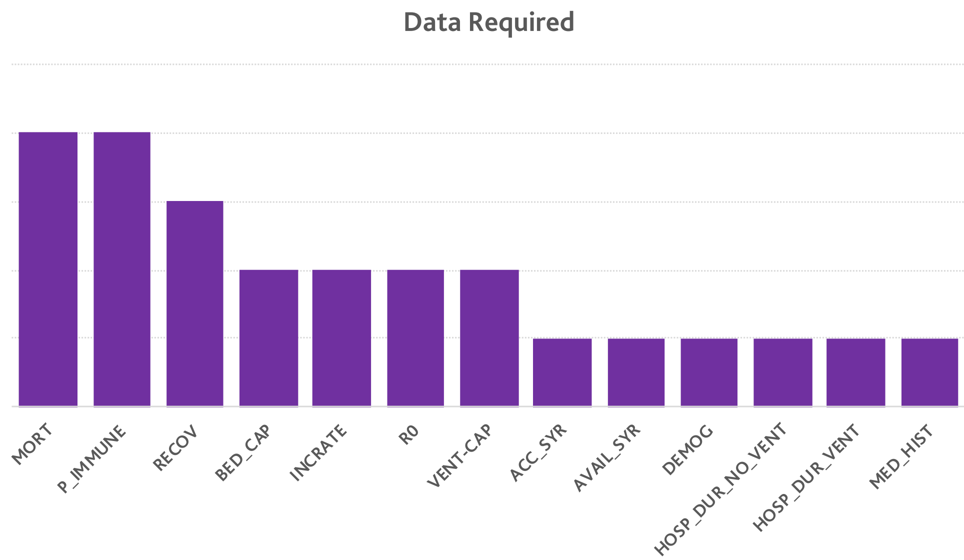
- who is highest at risk and why;
- the patterns and drivers of COVID-19 transmission; and
- whether prevention measures are working.

Such tools can then inform public policy recommendations regarding vertical interdiction, especially with regard to social distancing based on population stratification. In effect, a successful public health platform will simultaneously:

- minimize damage from the disease,
- maximize herd immunity,
- shorten the duration of the disease, and
- minimize the economic impact of social interventions.

MOST COMMON DATA NEEDS IDENTIFIED

This graph shows the distribution of requested data elements across all projects. Note that the number of requests does not necessarily correspond to the importance of the data, particularly with respect to a specific model application.



For a detailed listing of the data required for various models, please see the accompanying Atlas.

KEY POLICY IMPLICATIONS

1. Current social distancing measures should remain in place during the data collection period, except in cases of individuals with conferred immunity or tested presence of COVID-19 antibodies.

2. Vertical interdiction must be implemented in a manner that minimizes harm, particularly within high-risk populations.

3. Pivot date will be informed by a combination of active monitoring, robust data collection, and, where possible, validated models informed by this robust data.

4. Policy makers should adjust policy proposals and legislation as robust data becomes available and these bodies should implement formal monitoring systems and procedures for this purpose.

5. Any federal or state level task forces must prioritize evidence-based policy.

6. Total Harm Minimization (THM) requires establishing the components of, path to, and timeline for an “all clear:” the point at which *all* population segments may once again resume a semblance of “normal” life.

7. Absent the introduction of an effective vaccine, the most expedient route to an all-clear at this point appears to be via herd immunity, i.e., allowing those with confirmed immunity (ie, Selective Early Returners to Life As We Knew It: SERLAWKI) to resume a version of normal activity while other, of the lowest-risk segments of the population to gradually return as well, but with mandatory, perhaps daily, screening before entering public places, mass transportation or work environments.

Appendix A: Selected COVID-19 live data sources

Daily data summary of NYC

COVID-19 Projections Data from healthdata.org (forecasting model used as a planning tool):

COVID-19 Data Hub

Our World in Data visual charts and graphs

Infographics updated in realtime

COVID-19 Hospital Impact Model for Epidemics (CHIME)

Appendix B: Estimated proportion COVID-19 infections leading to hospitalization (fr. Verity, et al. (2020)*)

Table 3 Estimates of the proportion of all infections that would lead to hospitalisation, obtained from a subset of cases reported in mainland China[22]

	Severe cases	All cases	Proportion of infected individuals hospitalised
0–9 years	0	13	0·00% (0·00–0·00)
10–19 years	1	50	0·0408% (0·0243–0·0832)
20–29 years	49	437	1·04% (0·622–2·13)
30–39 years	124	733	3·43% (2·04–7·00)
40–49 years	154	743	4·25% (2·53–8·68)
50–59 years	222	790	8·16% (4·86–16·7)
60–69 years	201	560	11·8% (7·01–24·0)
70–79 years	133	263	16·6% (9·87–33·8)
≥80 years	51	76	18·4% (11·0–37·6)

* *The Lancet*: Estimates of the severity of coronavirus disease 2019: a model-based analysis, March. 2020.