Brief Note

How Frequently to Test for Infection?

Yakov Ben-Haim[1]

Contents

> "I've looked at all the models. I've spent a lot of time on the models. They don't tell you anything. You can't really rely upon models."

Dr. Anthony S. Fauci, director of the US National Institute of Allergy and Infectious Diseases (Washington Post, 3 April 2020).

Abstract In some diseases, such as caused by the novel coronavirus, infected individuals are infectious to others long before they themselves develop pathological symptoms of disease. In such cases, testing non-symptomatic individuals for infection, and quarantining those infected, is essential for managing the epidemic. New diseases are often poorly understood and surrounded by deep uncertainty. We use the info-gap concept of robustness to manage uncertainty and to address the question: at what frequency is it necessary to clinically test a target population, in order to have confidence in mastering the epidemic?

Keywords Novel coronavirus, infectiousness, clinical test frequency, deep uncertainty, info-gap theory.

1 Introduction

We consider a medical epidemic such as the coronavirus covid-19 pandemic in 2019 and 2020, with the goal of supporting public health strategy for overcoming the disease. In some diseases, infected individuals are infectious to others long before they themselves develop pathological symptoms of disease. In the case of covid-19, this infectious incubation period can be as long as 14 days, or perhaps even longer. In such cases, testing non-symptomatic individuals for infection, and quarantining those infected, is an essential strategy for managing the epidemic. The basic question we address

\papers\corona2020\hmt004.tex. 8.4.2020.

[1]Yitzhak Moda'i Chair in Technology and Economics, Technion—Israel Institute of Technology, Haifa, Israel. yakov@technion.ac.il

is: at what frequency is it necessary to clinically test a target population, in order to have confidence in mastering the epidemic?

In order to address this question we must manage the vast uncertainty surrounding our understanding of any new disease such as the coronavirus covid-19 pandemic. We develop a response to this challenge, and present a response to Dr. Fauci's concern expressed in the epigraph.

In section 2 we briefly review a few basic epidemiological terms that underlie the subsequent development. In section 3 we develop and discuss an info-gap analysis of robustness to uncertainty, and in section 4 we discuss the implications for answering the question we have posed.

2 Epidemiological Fundamentals

Most of the ideas in this section are based on Grassly and Fraser (2008).

Infectiousness is a characteristic of an infected individual that determines the rate of infection of susceptible individuals. In our analysis we will assume that all uninfected people are susceptible to the disease.

The infectiousness function, $\beta(\tau)$, is the infectiousness of an infected individual over time, in units of 1/time. We assume that the same infectiousness function describes all infected individuals. More precisely, $\beta(\tau)\mathrm{d}\tau$ is the probability that an infected person, A, will infect another person during $[\tau,\, \tau + \mathrm{d}\tau]$ after A's infection. Note that $\beta(\tau)$ is not the probability density for an infected person to infect a susceptible person; it may be constant and it need not be normalized over time.

Y is the number of secondary infections generated by a single infected person, called the offspring of that infected person. $\mathrm{E}(Y)$ is the average number of offspring of an infected person, and equals $\int_0^\infty \beta(\tau)\mathrm{d}\tau$. Recall that we assume all uninfected people are susceptible.

The *reproduction number,* R, equals the average number of secondary infections, $\mathrm{E}(Y)$. In our analysis this is the same as the *basic reproduction number,* R_0, because we are assuming all uninfected people are susceptible. Thus we can write:[2]

$$R = \int_0^\infty \beta(\tau)\mathrm{d}\tau \tag{1}$$

Epidemic can occur only if $R > 1$. Thus reducing R below 1 prevents or ends an epidemic (though not immediately).

3 Formulation and Analysis of Robustness to Uncertainty

Profound uncertainties surround the epidemiology of a new disease. This motivates the use of a very simple model of the epidemiological process. Simple models are wrong in simpler ways than sophisticated models. More sophisticated models, such as discussed by Feller (1941), can and should be studied, but their uncertainties are compounded and more difficult to manage. Furthermore, because we are dealing with highly uncertain models in any case, the approach we adopt is to maximize the robustness to uncertainty while satisfying critical outcome requirements. This

[2] See Diekmann et al (1990) for a much more general version of this relation. Note however that Diekmann makes various approximations, such as only obtaining upper limits for the reproduction number, due to the great mathematical complexity. See also Mattheij and Molenaar (2002).

is different from using the best available model in attempting to obtain the best possible outcome. The robust-satisficing approach developed here is a response to Dr. Fauci's criticism of all available models.

Eq.(1) suggests that one way to reduce the reproduction number to an acceptably low value, $R_{\rm c} < 1$, is to reduce the duration of infectiousness. This can be done by identifying and isolating infected individuals within a time $t_{\rm c}$ of their infection. If that is done, then the infectiousness function becomes zero for $t > t_{\rm c}$. From eq.(1) we choose the testing interval $t_{\rm c}$ to satisfy:

$$\int_0^{t_{\rm c}} \beta(\tau){\rm d}\tau \leq R_{\rm c} \tag{2}$$

where $R_{\rm c}$ is less than one. In other words, each member of the target population must be tested for infection once every $t_{\rm c}$ days, thereby reducing the reproduction number to $R_{\rm c}$.

One problem in choosing the testing interval, $t_{\rm c}$, is that the infectiousness function, $\beta(\tau)$, is poorly known and highly uncertain. We know neither the shape nor magnitude of this function. Furthermore, susceptibility may well vary with age group, medical condition, sex, and other parameters. We are ignoring the different susceptibilities of various sectors of the population due to deficient understanding both of the disease and of how people will behave in response. Stated differently, the deep uncertainty that we are attempting to manage includes the population heterogeneity as well.

As a very rough guess, we adopt the following expression for the infectiousness function, recognizing that it is probably wrong to an unknown degree. Our main task is to address the uncertainty in this function, which we will do shortly.

$$\beta(\tau) = \beta_0, \quad 0 \leq \tau \leq t_{\rm r} \tag{3}$$

where β_0 is a positive constant, and $\beta(\tau) = 0$ for all values of τ outside of the interval $[0,\ t_{\rm r}]$. Recall that τ measures the time since that person was infected, at which time infectiousness of that person begins, though pathological symptoms may not occur until later. Infectiousness may also continue after recovery of that person. We address the deep uncertainty in this infectiousness function shortly.

Consider, for instance, the coronavirus Covid-19 pandemic that started in Wuhan, China, at the end of 2019. The incubation period, between infection and appearance of symptoms, is thought to vary from 1 to 14 days, most commonly around 5 days (WHO, 9.3.2020; Lauer et al, 2020), while Lauer, et al (2020) also write that "97.5% of those who develop symptoms will do so within 11.5 days ([confidence interval], 8.2 to 15.6 days) of infection." The reproduction number, R, is thought to be about 2.2, with a 95% confidence interval from 1.4 to 3.9, (Wu et al, 2020).

Thus, if we adopt $t_{\rm r} = 5$ and $R = 2$, we would combine eqs.(1) and (3) to estimate β_0 as $\widetilde{\beta}_0 = 2/5 = 0.4$ per day. But other plausible estimates could be $1.5/10 = 0.15$ or $2.5/3 = 0.83$ per day. A wider range of $\widetilde{\beta}_0$ estimates is also clearly possible. In short, the value of the infectiousness, β_0, is highly uncertain, and could deviate from the nominal estimate of 0.4 per day by a factor of 2 or more. Furthermore the actual shape and magnitude of the true infectiousness function, $\beta(\tau)$, is unknown.

We have no probabilistic information about the uncertainty in the infectiousness function $\beta(\tau)$, and the expression in eq.(3) is not at all reliable (recall Dr. Fauci's plaint regarding *all* models). Hence we will describe this deep uncertainty with the following fractional-error info-gap model (Ben-

Haim, 2006, 2018):

$$\mathcal{U}(h) = \left\{ \beta(\tau) : \ \beta(\tau) \geq 0, \ \left| \frac{\beta(\tau) - \widetilde{\beta}_0}{\widetilde{\beta}_0} \right| \leq h \right\}, \quad h \geq 0 \tag{4}$$

The set $\mathcal{U}(h)$ contains all positive infectiousness functions, $\beta(\tau)$, that differ from the estimated value, $\widetilde{\beta}_0$, by no more than the fraction h. The value of h is unknown, so the info-gap model is an unbounded family of nested sets of possible infectiousness functions. We are making no assumptions about the shape of these functions, other than that they are non-negative.

The robustness is the greatest horizon of uncertainty, h, up to which the requirement in eq.(2) is satisfied by all infectiousness functions in $\mathcal{U}(h)$. The mathematical definition of robustness is:

$$\widehat{h}(R_\mathrm{c}; t_\mathrm{c}) = \max \left\{ h : \ \left(\max_{\beta(\tau) \in \mathcal{U}(h)} \int_0^{t_\mathrm{c}} \beta(\tau) \mathrm{d}\tau \right) \leq R_\mathrm{c} \right\} \tag{5}$$

The inner maximum occurs for $\beta(\tau) = (1+h)\widetilde{\beta}_0$, from which we find the robustness function:

$$\widehat{h}(R_\mathrm{c}; t_\mathrm{c}) = \frac{R_\mathrm{c}}{\widetilde{\beta}_0 t_\mathrm{c}} - 1 \tag{6}$$

or zero if this is negative.

We see from eq.(6), unsurprisingly, that a shorter interval between tests (smaller t_c) is more robust to uncertainty than a longer testing interval (larger t_c). Furthermore, we see that the robustness trades off against the reproduction number: robustness goes down (gets worse) as the required reproduction number gets smaller and thus more demanding. Finally, the predicted reproduction number, $\widetilde{\beta}_0 t_\mathrm{c}$, has zero robustness to uncertainty.

We will use the robustness function, $\widehat{h}(R_\mathrm{c}; t_\mathrm{c})$, to choose a testing interval, t_c, that can confidently achieve a reproduction number no greater than R_c. We are *not* using an epidemiological model to predict R and thereby choose t_c. We are acknowledging and managing the deep uncertainty — vast ignorance — about the infectiousness function and all models from which it could be derived. We have identified a specific function, $\beta(\tau)$ in eq.(3), but we are not predicating our choice of t_c on the veracity of eq.(3). The function in eq.(3) is no more than the center point of the unbounded family of nested sets, $\mathcal{U}(h)$, of possible infectiousness functions. What we are doing is choosing t_c so that the reproduction number is acceptably small for the widest range of possible infectiousness functions.

4 Discussion

In the discussion preceding the info-gap model of eq.(4) we concluded that, for the covid-19 pandemic, the estimated infectiousness of $\widetilde{\beta}_0 = 0.4$ per day could vary by a factor of 2 or more. Furthermore, the uncertainty is not only in the value of the coefficient $\widetilde{\beta}_0$, but also in the shape and magnitude of the true infectiousness function $\beta(\tau)$. Thus a robustness of at least 2 is needed[3] in order that the testing and isolation will reduce the reproduction number to the value R_c with some confidence. If the required reproduction number is $R_\mathrm{c} = 0.9$, then $\widehat{h} = 2$ in eq.(6) implies $t_\mathrm{c} = 0.75$ days as the interval between testing. That means testing the target population once a day.

[3]From the info-gap model of uncertainty in eq.(4) we see that, if the robustness takes the value $\widehat{h}$, then acceptable outcome is guaranteed for all non-negative infectiousness functions, $\beta(\tau)$, in the interval $[(1-\widehat{h})\widetilde{\beta}_0, \ (1+\widehat{h})\widetilde{\beta}_0]$. Thus $\widehat{h} = 2$ implies acceptable outcome for all infectiousness functions, $\beta(\tau)$, in the interval $[0, \ 3\widetilde{\beta}_0]$. That's moderate robustness.

Eq.(6) can be inverted to determine the testing interval, $t_{\rm c}$, as a function of the required reproduction number $R_{\rm c}$, the estimated infectiousness $\widetilde{\beta}_0$, and the demanded robustness $\widehat{h}_{\rm d}$:

$$t_{\rm c} = \frac{R_{\rm c}}{(1+\widehat{h}_{\rm d})\widetilde{\beta}_0} \tag{7}$$

	Testing Interval, $t_{\rm c}$ [days]			
	$\widetilde{\beta}_0 = 0.05$	$\widetilde{\beta}_0 = 0.1$	$\widetilde{\beta}_0 = 0.2$	$\widetilde{\beta}_0 = 0.4$
$R_{\rm c} = 1$	6.7	3.37	1.7	0.83
$R_{\rm c} = 0.9$	6.0	3.0	1.5	0.75
$R_{\rm c} = 0.7$	4.7	2.3	1.2	0.58
$R_{\rm c} = 0.5$	3.3	1.7	0.83	0.42

Table 1: Testing intervals (days) with $\widehat{h}_{\rm d} = 2$.

Table 1 shows testing intervals, based on eq.(7), for a selection of estimated infectiousness values, $\widetilde{\beta}_0$, and required reproduction numbers $R_{\rm c}$, with demanded robustness $\widehat{h}_{\rm d} = 2$. The value discussed above, $t_{\rm c} = 0.75$, appears in the column for $\widetilde{\beta}_0 = 0.4$ and the row for $R_{\rm c} = 0.9$. Longer intervals between testing are obtained if the estimated infectiousness is lower, which may sometimes be achieved by behavioral modification such as social distancing, as we now explain.

Recall that the infectiousness, $\beta(\tau)$, is the probability per unit time that an infected person will infect another person. In the case of the new coronavirus, face-to-face interaction of 10 or 20 minutes is perhaps sufficient to transmit the disease (though infection is a random process and a single interaction has low probability of infection). Consider a typical person who interacts in this way with 20 other people in the course of a regular day. Social distancing can reduce this number of infectious interactions by physical distancing, by masking, or perhaps by other means. If social distancing reduces infectious interactions to 5 per day, then the probability of infection per day will fall roughly[4] by a factor of 4. This would reduce the estimated infectiousness from $\widetilde{\beta}_0 = 0.4$ to about $\widetilde{\beta}_0 = 0.1$, and increase the required interval between tests from 0.75 to 3 days as seen in table 1 when requiring $R_{\rm c} = 0.9$.

Eq.(7) and table 1 have implications for workforce planning during an epidemic such as Covid-19.

It is not feasible to test the entire population of a country every day, as required if $\widetilde{\beta}_0 = 0.4$ and $R_{\rm c} = 0.9$. Israel, for instance, has a population of 9 million, Germany a population of 83 million, and the U.S. a population of 330 million. However, daily testing of 1 out of 300 citizen could be feasible. For Israel, Germany and the U.S. this comes to 30,000, 277,000 and 1.1 million daily tests. Israel has tested 8,000 people per day, and Germany 50,000 per day, so greater effort would be needed, but this could be achieved. Alternatively, the testing interval could be extended by social distancing in the exposed population as discussed earlier. The point of the analysis summarized in table 1 is to identify testing intervals that reduce the reproduction number below 1, and doing so reliably despite the deep uncertainty in our understanding of the disease.

[4] For example, suppose the probability of infection from a single interaction is $p = 0.025$. Then the probability of no infection from $n = 20$ independent interactions in 1 day is $(1-p)^n = 0.60$. This implies that the probability of at least one infection is 0.40 per day. If $n = 5$, then $(1-p)^n = 0.88$, implying a probability of 0.12 of at least one infection per day.

The tested target population would work in environments that do not constitute isolation. Nonetheless, exposure of the target population to infection would be reduced due to reduction in the size of the workforce, and social distancing measures such as use of masks.

The rest of the population would remain isolated in small family-sized units. A substantial portion of the isolated population can — and must — still work, but from isolation. Many office jobs, construed broadly, will continue. Government offices, banks, universities, insurance companies, IT ventures, and more, will continue to operate from complete isolation. Other sectors, like farming, can operate in near isolation by working in small closed units, and by bringing produce to collection points in a time-staggered manner to avoid contact with distributors. Similar solutions are available in other sectors.

Many questions remain open and require attention.

We have not addressed the question of what diagnostic tests should be used: reverse transcription polymerase chain reaction (rRT-PCR) (Food and Drug Administration), serological tests (Johns Hopkins), etc., how they are to be implemented (individually or by batching?), how reliable they are, and so on.

We have assumed that each member of the on-site workforce who tests positive is immediately replaced with a non-infected person from the remainder of the population. In other words, there must be a reserve tested population from whom replacements can come. What frequency and number of tests, of a targeted subset of the isolated population, are needed in order to maintain a reserve from which to replace infected members of the on-site workforce?

It is necessary to develop a general economic plan for the epidemic transition period. Among other things, this plan will identify the size and composition of the on-site workforce (is 1 in 300 really sufficient?). The plan must also identify critical functions that can be filled remotely, but sometimes with adjustments or adaptations. The plan will develop ways and means to manage myriad logistical problems, such as young parents working from home with school-age children also at home.

In addition to an economic and institutional plan, it is necessary to establish the legal framework for managing the transition period. Presumably it will be necessary to define a centralized planning and management agency, but this must be defined precisely. Who has what powers, for how long and through what means of enforcement?

Forecasting the duration of the epidemic transition period is critical for planning and implementation. This requires careful methodological attention to the deep uncertainties surrounding the epidemiological dynamics.

Other questions requiring attention will also arise.

Acknowledgements The author is indebted to the following people for their constructive critical comments: Anne C. Bader, Yael Ben-Haim, Yohay Carmel and Shai Linn.

5 References

1. Ben-Haim, Yakov, 2018, *Dilemmas of* Wonderland: *Decisions in the Age of Innovation,* Oxford University Press.
2. Ben-Haim, Yakov, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd edition, Academic Press, London.

3. Diekmann, O.; J. A. P. Heesterbeek and J. A. J. Metz, 1990, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, *J. Math. Biol.,* 28: 365–382.
4. Food and Drug Administration, Accelerated Emergency Use Authorization (EUA) Summary, https://www.fda.gov/media/136151/download.
5. Feller, Willy, 1941, On the integral equation of renewal theory, *The Annals of Mathematical Statistics,* 12(3): pp.243–267.
6. Grassly, Nicholas C. and Christophe Fraser, 2008, Mathematical models of infectious disease transmission, *Nature Reviews: Microbiology,* vol. 6, pp.477–487.
7. Johns Hopkins, Serology-based tests for COVID-19,

   http://www.centerforhealthsecurity.org/resources/COVID-19/Serology-based-tests-for-COVID-19.html
8. Lauer, Stephen A.; Kyra H. Grantz; Qifang Bi; Forrest K. Jones; Qulu Zheng; Hannah R. Meredith; Andrew S. Azman; Nicholas G. Reich; Justin Lessler, 2020, *Annals of Internal Medicine,* The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application, doi:10.7326/M20-0504, 10 March 2020, downloaded 2.4.2020.
9. Mattheij, Robert and Jaap Molenaar, 2002, *Ordinary Differential Equations in Theory and Practice,* SIAM Classics in Applied Mathematics.
10. World Health Organization website, https://www.who.int/news-room/q-a-detail/q-a-coronaviruses, dated 9 March 2020.
11. Wu, Di; Tiantian Wu; Qun Liu; Zhicong Yang; 2020, The SARS-CoV-2 outbreak: What we know, *International Journal of Infectious Diseases,* accepted for publication, 5 March 2020.