Supplementary material: Assessing the impact of mass testing during a novel infectious disease outbreak: The case of COVID-19

Authors

Yang Ge, Department of Epidemiology and Biostatistics, The University of Georgia, Athens, GA, USA, https://orcid.org/0000-0001-5100-0703

Brian Kenneth McKay, Department of Epidemiology and Biostatistics, The University of Georgia, Athens, GA, USA, https://orcid.org/0000-0003-3859-7634

Shengzhi Sun, Department of Environmental Health, Boston University School of Public Health, Boston, MA, USA.

Feng Zhang, Department of Biostatistics and Data Science, University of Texas Health Science Center at Houston, School of Public Health, Houston, TX, USA.

Andreas Handel, Department of Epidemiology and Biostatistics, The University of Georgia, Athens, GA, USA, http://orcid.org/0000-0002-4622-1146

Corresponding author: yang.ge@uga.edu, ahandel@uga.edu
Estimating F and P for a few pathogens

The MSTI strategy we discuss here depends on values for the proportion, \( P \), of individuals that show a set of symptoms who will be infected with the novel pathogen, and the risk of infection due to being at a testing site, \( F \). Precise estimates for either of those quantities are unlikely to be available for any setting. However, if the right data is collected, it is possible to get decent estimates. Unfortunately, since MSTI has not been formally evaluated previously, the right kind of data is not available. As an alternative, we searched the literature and tried to find reports that had some data that would allow us to estimate \( F \) and \( P \) for a few pathogens. The following sections explain how we obtained those estimates. Due to the strong limitations with the data, these estimates should be considered very approximate and uncertain.

Ebola

One study [1] reports a scenario where 464 of 724 patients (64%) suspected to have Ebola virus disease (EVD) were found to have EVD. A similar fraction was reported for a different setting in another study [2], with 276 of 433 suspected patients (64%) having EVD. In [3], the fraction is \( 108/154 = 70\% \). while [4] report \( 822/2362=35\% \). Based on these reports, we obtain an estimate for \( P \) in the range of 0.35-0.7.

To estimate \( F \), we use data from [5], which reports that among a cohort of 543 patients who initially did not have EVD and were discharged, 18 later likely developed EVD. It is unclear if they got infected in the facility, but an upper bound would be \( 18/543 = 0.03 \). In a similar study [6] found \( 9/138=0.07 \) as an upper bound. This leads to an estimate for \( F \) in the range of 0.03-0.07.
Measles

In [7], the authors analyzed four sets of data to determine confirmed measles cases among those having a clinical diagnosis and found those to be $77/128=0.60$, $1/26=0.04$, $1/93=0.01$, $74/213=0.35$. This leads to an estimate for $P$ covering a wide range of 0.01-0.6.

To estimate $F$, we use data from [8], which reports a case where a single visit to an emergency department resulted in 12 cases. They do not report the number of exposed persons. If we make a guess of around 100, we obtain a risk of infection of $12/100=0.12$. Similarly, [9] report 4 cases caused by an ED visit of an infected patient. Again using a total size of around 100 for those exposed we get $4/100=0.04$. This leads to an estimate for $F$ in the range of 0.04-0.12.

COVID-19

The least information is available for the current COVID-19 outbreak. Notes of a press conference held by National Health Commission of the People’s Republic of China show there were 10261 total fever clinic visits in Wuhan on Jan 27, 2020 [10], with 315 confirmed cases reported next day [11]. Similarly, 12512 and 12568 total fever clinic visits occurred on Feb 3 and 4 [12], with 1967 and 1766 confirmed cases reported the next day [13,14]. If we use the reported fever visits as indication of the total number of individuals showing respiratory symptoms and the next day reported cases as the number among those individuals who are infected with COVID-19, we obtain as estimates for $P$ the values $315/10261$, $1967/12512$ and $1766/12568$, which leads to an estimated range for $P$ of 0.03-0.16.

To estimate $F$, we use data from [15], which reports that one patient infected 4 hospitalized patients in the same ward. We do not know the number of exposed persons in this ward. If we
take a range of 10-100, this would lead to risk of infection from 4/10 - 4/100 and thus $F$ in the range of 0.04 - 0.4.

**Scenarios with non-perfect isolation**

In the main text, we considered the optimistic case where isolation following testing led to no further transmission and no mortality. Here, we explore how results change if instead isolation only reduced those quantities by some amount. As a reminder, the general expressions for transmission risk ratio (TRR) and mortality risk ratio (MRR) are given by $TRR = T_S/T_H = R_I/R + (1 - P)F/P$ and $MRR = M_S/M_H = M_I/M_0 + (1 - P)F/P$. The figure explores scenarios ranging from isolation having very little effect on transmission and mortality risk and only reducing them by 10% (i.e. $R_I = 0.9R$ and $M_I = 0.9M$) up to a reduction of 90% (i.e. $R_I = 0.1R$ and $M_I = 0.1M$). As expected, a smaller reduction in infection and mortality risk shrinks the range for which MSTI is a beneficial strategy.
Figure 1: Simulation of TRR and MRR in scenarios; A: $R_I = 0.9R$, B: $R_I = 0.5R$, C: $R_I = 0.1R$,
D: $M_I = 0.9M$, E: $M_I = 0.5M$, F: $M_I = 0.1M$
An explicit equation for the probability of infection

The probability of infection, $F$, is expected to increase as the proportion of individuals infected with the novel pathogen increases and will saturate at some maximum value. This can be described by a simple function $F = P/(K + P)$, where $K$ describes the strength of the infection control at the testing site. For small $K$, $F$ will eventually approach 1. For infection control strength of $K = 1$, the maximum risk of transmission, (if prevalence were to be $P = 1$) would be $F = 0.5$.

**Figure 2**: MSTI associated infection risk, $F$, as a function of novel pathogen prevalence among symptomatic, $P$, given by $F = P/(K + P)$ for different values of $K$.

If we use $F = P/(K + P)$ in our TRR and MRR equations, we find $TRR = T_S/T_H = R_I/R + (1 - P)/(K + P)$ and $MRR = M_S/M_H = M_I/M_0 + (1 - P)/(K + P)$. With the assumption of perfect isolation ($R_I = M_I = 0$), this becomes $TRR = MRR = (1 - P)/(K + P)$.
For MSTI to be useful, we need to have $TRR = MRR < 1$, and thus $(K + P) > (1 - P)$ or $K > 1 - 2P$. This shows that infection control needs to be strongest at $K \geq 1$ for low prevalence, and once $P \geq 0.5$, even zero infection control will make MSTI beneficial. This corresponds to the main text discussion of keeping $F < P$ for low prevalence, while the value of $F$ does not matter once prevalence reaches 0.5.

Note that the infection control parameter $K$ in our model is in arbitrary units without a clear way to connect and measure it for any specific setting. If any 2 of $K$, $F$ and $P$ are known, the third quantity can be estimated. It is more likely that data can be obtained for infection risk $F$, instead of $K$, which is why we focused on $F$ in the main text.


