Benefit of random testing

Markus Müller,[1, *] Peter M. Derlet,[1, 2] Christopher Mudry,[1, 3] and Gabriel Aeppli[1, 3, 4]

[1] *Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland*
[2] *Department of Materials, ETH Zurich, CH-8093 Zurich, Switzerland*
[3] *Institut de Physique, EPF Lausanne, Lausanne, CH-1015, Switzerland*
[4] *Department of Physics, ETH Zurich, Zurich, CH-8093, Switzerland*
(Dated: May 6, 2020)

I. INTRODUCTION

With the imminent relaxation of socio-economic restrictions, it becomes vital to assess its effect on the prevalence of acute infections within the population, as rapidly as possible. Currently available monitoring instruments for the COVID-19 pandemic have an inherent time delay of about 14 days, as they rely on confirmed infections, hospitalizations, and death numbers. These methods give $R_{\mathrm{eff}}(t)$ (the number of infections caused by a single infected person), but their delay is a significant disadvantage when restrictions are released. If after relaxation, $R_{\mathrm{eff}}(t)$ rises above 1, one will not be able to react adequately before two weeks have passed during which time the prevalence could significantly rise.

Here, we propose the use of random testing to shorten this reaction time, by obtaining direct and model-independent information on $R_{\mathrm{eff}}(t)$. Through random testing of between $2'500$ and $20'000$ people per day, we find that over periods significantly shorter than two weeks, it becomes possible to detect a dangerous increase in R_{eff} with reasonable confidence. When compared to the delay of nonrandom symptomatic testing currently being performed, this shorter response time can save tens of lives per week at the national level, cf. Fig. 1, and reduces costs for the health care system as well as for the economy. Moreover such monitoring provides greater stability and diminishes the probability of a second wave of pandemic.

II. ESTIMATING THE FIRST INTERVENTION TIME

The following is a direct application of the ideas presented in Ref.[1]. Here we focus on the question of how many lives can be saved and how much damage can be prevented by the more rapid intervention possible with random testing.

Assume that in the first half $(T/2)$ of the measuring period T, we detect N_1 infected persons. If we denote with r the number of people tested daily, i_0 the prevalence (infected fraction in the population being randomly sampled) just before the measuring period T, and

$$n_d := i_0\, r \tag{2.1a}$$

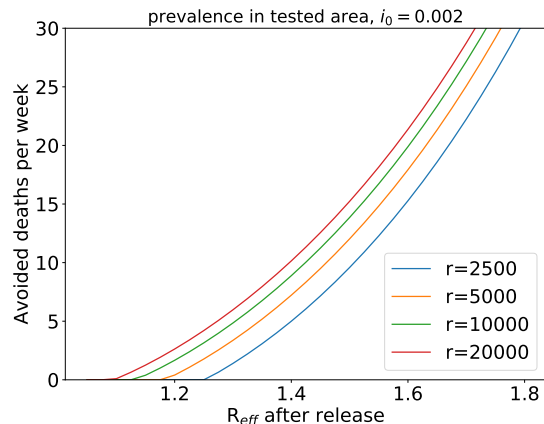the expected number of people detected positive per day,

FIG. 1. Saved lives per week due to the avoided increase of infection numbers as a function of the reproduction rate R_{eff} prevailing right after some restrictive measures are released. The four colored curves correspond to different numbers r of people tested per day. The numbers used for the estimate are given in the main text. We assume a prevalence of $i_0 = 0.002$ in the tested area (as we currently estimate for Geneva) and an average prevalence in Switzerland of $i_{\mathrm{CH}} = 0.0005$.

then we expect

$$N_1 = \frac{T}{2}\, i_0\, r \equiv \frac{T}{2}\, n_d. \tag{2.1b}$$

In the second half of the measuring period T, we would expect

$$N_2 = N_1\, \exp(|k|\, T/2) \tag{2.2}$$

if the growth rate of infections is $|k|$. Now, k is related to the reproduction number R (the expected number of persons that will be infected by one sick person) through

$$k \approx \mu \times (R - 1), \tag{2.3a}$$

where the proportionality constant $\mu > 0$ can be obtained from unmitigated growth data through

$$\mu \approx \frac{k_0}{R_0 - 1}, \qquad R_0 > 1. \tag{2.3b}$$

Here, R_0 is the basic reproduction rate, i.e., the expected number of cases directly generated by one case in a population where all individuals are susceptible to infection,

while k_0 is the ensuing exponential rate of change. We take unmitigated growth for COVID19 to be doubling in 3 days, i.e.,

$$2 = e^{k_0 \times 3} \iff k_0 = \frac{\ln 2}{3} \approx 0.23, \qquad (2.4a)$$

while we choose among the many published estimates[2] of R_0,

$$R_0 \approx 2.9. \qquad (2.4b)$$

From this we get

$$\mu \approx 0.12. \qquad (2.4c)$$

We can tell with reasonable certainty that the growth rate k is positive if the difference

$$N_2 - N_1 = N_1 \left[\exp(|k| T/2) - 1\right] \qquad (2.5a)$$

is larger than its statistical uncertainty $\alpha \sqrt{2 N_1}$, i.e.,

$$N_1^2 \left[\exp(|k| T/2) - 1\right]^2 \geq 2 \alpha^2 N_1. \qquad (2.5b)$$

Here, the typical fluctuation $\sqrt{N_1}$ occurring upon sampling a random number with expectation N_1 is derived from the law of large numbers for independent random variables. The dimensionless parameter α selects the accuracy (confidence) that we seek. We require

$$\alpha \geq 1, \qquad (2.5c)$$

and will assume equality ($\alpha = 1$) below.

Taking the equality in (2.5b), i.e., satisfying the minimal condition on N_1, we get the condition

$$N_1 = 2 \left(\frac{\alpha}{\exp(|k| T/2) - 1}\right)^2 \qquad (2.6a)$$

or, if we use Eq. (2.1b),

$$n_d = \frac{4}{T} \left(\frac{\alpha}{\exp(\mu |R - 1| T/2) - 1}\right)^2. \qquad (2.6b)$$

Given n_d and the growth rate k (or, equivalently, the reproduction number R) after a sudden intervention, we obtain the required sampling time T from solving Eq. (2.6b). Expanding the denominator under the assumption that

$$0 < \mu |R - 1| \frac{T}{2} \ll 1, \qquad (2.7a)$$

one finds

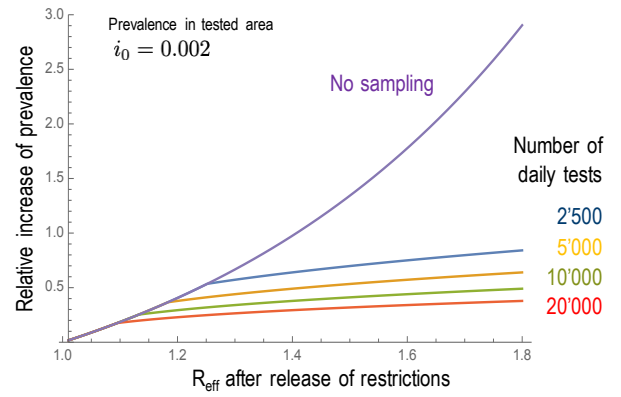$$T = \left(\frac{4 \alpha}{\mu |R - 1|}\right)^{2/3} \frac{1}{n_d^{1/3}}. \qquad (2.7b)$$

FIG. 2. Relative increase of prevalence until an intervention is taken, plotted as a function of the reproduction rate R_{eff} that prevails right after releasing some restrictive measures. The topmost curve shows the result in the absence of sampling, assuming a delay of T^* until a first intervention. The four colored curves correspond to different numbers r of people tested per day. The numbers used for the estimate are given in the main text. We assume a prevalence of $i_0 = 0.002$ in the tested area (as we currently estimate for Geneva).

III. INCREASE OF INFECTION NUMBERS - THE BENEFIT OF RANDOM TESTING

By the time T the prevalence, i.e., the infection numbers, will have increased by the fraction

$$\frac{i(T) - i_0}{i_0} = \exp(\mu |R - 1| T) - 1 \approx \mu |R - 1| T$$
$$= \left(16 \alpha^2 \frac{\mu |R - 1|}{n_d}\right)^{1/3}, \qquad (3.1)$$

which is to be compared to the fraction

$$\frac{i(T^*) - i_0}{i_0} = \exp(\mu |R - 1| T^*) - 1 \approx \mu |R - 1| T^* \quad (3.2)$$

with $T^* \approx 14$ days for methods based on fitting case numbers. These two relative increases are shown in Fig. ?? as a function of the effective reproduction number R_{eff} which prevails after a release of restrictive measures.

Random testing results in a smaller relative increase of the prevalence when enough information has been gathered to intervene. Its decrease amounts to

$$\frac{\Delta i}{i_0} \equiv \frac{i(T^*) - i(T)}{i_0} \approx \mu |R - 1| T^* - \left(16 \alpha^2 \frac{\mu(R - 1)}{n_d}\right)^{1/3}. \qquad (3.3)$$

Assuming an average prevalence in Switzerland of $i_{\text{CH}} = 0.0005$ and a total population of $N_{\text{CH}} \approx 8.4 \cdot 10^6$, we estimate that the number of infected persons is currently $i N_{\text{CH}} \approx 4'200$.

This number multiplied by $\Delta i / i_0$ given in Eq. (3.3) yields the number of additional infections avoided due to

random testing. With a moderate estimate of the mortality rate of $m = 0.5\%$, and assuming that people are infectious for $T_{\text{inf}} = 10$ days this means that the **number of saved lives per day in Switzerland** can be estimated to be

$$4200 \, \frac{m}{T_{\text{inf}}} \, \frac{\Delta i}{i_0} = 2.1 \, \frac{\Delta i}{i_0}. \tag{3.4}$$

The final result is plotted in Fig. 1, which clearly demonstrates that such an approach can save a significant number of lives. Unless the prevalence decreases again due to further lockdown conditions, there is a steady rate of lost lives associated with the increase of the prevalence, which would be avoidable through faster feedback from random sampling.

IV. COST ESTIMATE

We roughly estimate 50 CHF per tested person: 20 CHF for PCR (estimate by Fabian Rudolf BSSE/ETHZ, based on new high throughput tests) and 30 CHF for logistics. The latter could be significantly reduced if individuals were allowed to self-test and send in their probes by mail, as is currently done in a large scale random testing study in the UK (Imperial College, REACT1 study).

The necessary number of tests to assure a certain reaction time to correct for an overshoot in R_{eff}, and thus the costs, are inversely proportional to the prevalence i_0. With $i_0 \approx 0.002$ (our estimate for Geneva), and $n_d = 10'000$ one needs $5'000$ tests per day. This costs $250'000$ CHF per day, or 5 million CHF for 3 weeks, which is a natural time to accompany a strong release measure. The number of lives one is likely able to save thanks to random testing is considerable.

V. FURTHER BENEFITS OF RANDOM TESTING

Random sampling of the prevalence will finally inform us about the actual prevalence, which so far we can only estimate from epidemiological modelling, and which is still subject to substantial uncertainty.

Measuring the prevalence and its time evolution will help to better estimate unknown parameters for epidemiological modelling and thus improves the future predictions of these models.

The fast feedback on the reproduction rate allows to quantify the effect of individual policy measures, possibly with higher accuracy than by fitting with epidemiological models. This knowledge is valuable to optimize future interventions to keep the disease under control.

---

[*] Markus.Mueller@psi.ch

[1] Using random testing to manage a safe exit from the COVID-19 lockdown. M. Müller, P. M. Derlet, Ch. Mudry, and G. Aeppli, *arXiv:2004.04614*.

[2] Sebastian Bonhoeffer, Julien Riou, Christian Althaus, Melissa Penny, National Task Force Report: Response to questions from FOPH 17th April 2020.