Qeios

Measuring the efficacy of a vaccine during an epidemic

Antonio Scala[1,2,3,4☯*], Pierpaolo Cavallo[5☯]

1 CNR-ISC, Applico Lab, Roma, 00185, Italy
2 Centro Ricerche Enrico Fermi, Roma, Italy
3 Big Data in Health Society, Roma, Italy
4 Global Health Security Agenda - GHSA Italy
5 Department of Physics, University of Salerno, 84084 Fisciano, Italy
3 Affiliation Dept/Program/Center, Institution Name, City, State, Country

☯These authors contributed equally to this work.
Current Address: Dept/Program/Center, Institution Name, City, State, Country
* antonio.scala@cnr.it

Abstract

The vaccine rush caused by the current pandemic has led to performing fast clinical trials; in particular, we have observed a wide range of reported efficacy for the different vaccines from phase III cohort studies. We show that we show that when performing large cohorts phase III clinical trials near the epidemic peak, the measured effectiveness represents a strong under-estimate of the vaccine efficacy even in absence of confounding factors. In particular, we show that the underestimation grows with the fraction of infectious individuals present in the population during the experiment and with the severity of the epidemic measured by its basic reproduction number.

Introduction

The vaccine rush caused by the COVID19 pandemic has led to perform clinical trials with procedures that reflect the exceptional circumstances [1] and to the establishment of unprecedented public-private partnership [2]. In particular, we have observed the case of vaccines that have reported widely different efficacies [3], varying from the $\sim 95\%$ of Pfizer and Moderna (mRNA based), to the $\sim 70\%$ of Astra-Zeneca or the $\sim 66\%$ of Johnson & Johnson (viral vector based). While mRNA and viral vectors vaccines use different mechanisms to interact with the cells, both vaccines induce immunity by "instructing" our cells to produce spike proteins. Thus, it is reasonable to ask whether the heterogeneity of the results could be also influenced from the differences in the experimental environments, both in terms of fraction of infectious and in terms of the presence of variants. In this manuscript, we show how performing efficacy measurements at different times of the evolution of an epidemic can lead to serious underestimates of a vaccine's efficacy.

Vaccine efficacy are defined as one minus some measure of relative risk; according to the risk considered, several measures can be defined: efficacy for susceptibility to disease, for colonization, for progression, pathogenicity, infectiousness, indirect effects, population-level effects etc [4]. These measures require specialised and accurate datasets, sometimes with detailed information on the single contact experienced by the experimental cohorts. We will employ a simpler characterization of the vaccine efficacy ϵ defined in terms of the transmission rate β of the epidemic. A vaccine of efficacy ϵ

decreases the transmission rate by a factor $\delta = 1 - \epsilon$, i.e. a vaccinated person has a probability δ times lower of getting infected when coming into contact with an infectious individual; δ can be also indicated as the *relative risk* of vaccinated individuals [4]. Thus, the transmission rate for vaccinated persons lowers from β to $\beta^V = \delta\beta$ [5].

On the other hand, vaccine effectiveness measures the real-world performance of a vaccine [6, 7], in contrast with efficacy that can be defined as the performance of an intervention under ideal and controlled circumstances. Factors concurring in a deviation of effectiveness from efficacy are multifaceted, and the implementation of effectiveness studies (especially troublesome low- and middle-income countries) is affected by several confounding factors like age, socio-demographic factors (ethnicity/religion), geographical location, chronic disease and/or comorbidities and socio-economic status [8]. However, a factor that has been mostly disregarded in large cohort (i.e phase III) studies is the impact of the fraction of infectious individuals during the trial.

A key metric for the impact of an epidemic is the basic reproduction number $\mathcal{R}_0$, measuring the potential number of people an individual can infect; $\mathcal{R}_0$ can be calculated in terms of the transmission probability β and of the average lifetime τ of the infectious state as $\mathcal{R}_0 = \tau\beta$ [9]. The basic reproduction number allows to estimate the herd immunity threshold (HIT), i.e. the fraction $\rho^* = 1 - 1/\mathcal{R}_0$ of immune individuals beyond which no epidemic overburst can happen [9]. The efficacy ϵ is paramount for estimating the *effective* fraction of people $\rho^V = \rho^*/\epsilon$ to reach the HIT: the lower the efficacy, the higher the fraction of individuals to vaccinate.

The efficacy is not known a priori, but must be estimated through an experimental procedure. Overestimating ϵ would underestimate ρ^V, with the danger of not reaching the HIT at the end of the vaccination campaign. Underestimating ϵ ensures that the fraction of vaccinated people is beyond the HIT; however, it expands both the costs and the duration of a vaccination campaign and – in extreme cases – it can lead to an estimate of the number of individuals to vaccinate beyond any practical possibility. As an example, if the fraction of kids in a population is ρ^{kids} and the vaccine that cannot be administered to kids, ρ^V cannot be higher than $1 - \rho^{kids}$.

Reported efficacies are a measure of the reduction in disease incidence in a vaccinated group compared to an unvaccinated group *under optimal conditions in a clinical trial*. However, what happens if clinical trials are performed on large cohorts and during an epidemic, so that it is possible that "optimal conditions" cannot be strictly enforced? As noted by Hallorane et al [10], to avoid that equivalent populations with the same transmission conditions could yield different efficacy estimates, the amount of exposure to infection should be taken into account either by study design or by mathematical modeling.

To isolate the effect of pursuing clinical trials during an ongoing epidemic, we will consider the theoretical case where no confounding factors [6, 7] intervene in the effectiveness – measured as the experimental ratio of infected individuals in a vaccinated and a placebo cohort – showing in long trials performed during an ongoing epidemic the *effectiveness* underestimates the vaccine *efficacy*; such underestimation grows both with the fraction of infectious individuals i present in the population during the experiment and with the severity $\mathcal{R}_0$ of the infection.

Results

The effectiveness η of the vaccine is measured by confronting the infections occurring into two observed groups (also called *cohorts* in the medical language), one that has been vaccinated (cohort V of size N_V) and one that got a placebo (cohort P of size

N_P) [5]. The distribution of the traits (age, census, medical history, etc) of both cohorts must be representative of the whole population; moreover, (i) the individuals of both cohorts should be distributed in the population so not to have contacts among themselves (to avoid spurious correlations), (ii) the observation period T should be long enough to have a statistically significant number of observed cases of infections. Thus, the effectiveness η is estimated as "the proportion of persons in the placebo group of a vaccine trial who would not have become ill if they had received the vaccine" [5]. Let us indicate with A^V and A^P the number of cases in the vaccinated and placebo cohorts, respectively, at the end of the study. Suppose that we are analysing a large trial (like a phase III study) where, given the number of participants, it is not possible to have detailed information about their contacts. If we indicate with $c_P = A_P/N_P$ and $c_V = A_V/N_V$ the attack rates (or cumulative incidence), i.e. the fraction of individuals that get infectious during the trial, the vaccine efficacy can be expressed as [4]

$$\eta = 1 - \frac{c_V}{c_P} \tag{1}$$

; however, we must bear in mind that interpreting efficacy estimates is a multifaceted issue [4].

Notice that eq.1 could be a good estimate of the vaccine efficacy if the observed cohorts are under controlled clinical trial conditions [4]; in the case of large phase III studies, this is not the case: in particular, if phase III studies are performed *during* an epidemic, the measure of η could depend on the fraction of infectious individuals in the population.

The effectiveness η aims to be a proxy (in statistical language - an estimator) of the *real efficacy* ϵ of the vaccine. In the following, due to the nature of a phase III experiment (large numbers of individuals not subject to clinical trial conditions), we will assume that cohort individuals are in contact with the infected population. Since the number of individuals in the cohorts is much smaller than the population, we will also assume that the individuals in the V, P cohorts are uncorrelated as required from the experimental protocol. Finally, we will assume that the dynamics of the cohorts do not influence significantly the ongoing epidemics: this is true if the size of the group is much smaller than the population and if the observation time is much lower than the total duration of the epidemic. Under such assumption, it is possible to derive an explicit formula (eq. 4) for the final values of the effectiveness η in terms of the relative attack rate c, of the transmission rate β and of the relative risk δ. Since the attack rate can be expressed as $c = \bar{i}\,T$, we can see that the key drivers are the length T of the experiment and the average fraction $\bar{i}$ of infectious during the period T .

In the following, we will show the results for SIR models with parameters in the range of COVID19 estimates; in particular, we will assume that the infectious period is $\tau = 15$ days and $\mathcal{R}_0$ is in the range $[2.5, 6.0]$ [11]; however, since eq. 4 does not depend on the details of the dynamics, the results are expected to be robust in respect of the epidemic model employed.

In Fig 1, we show how the expected estimates of η (eq. 4) for a real efficacy $\epsilon = 0.90$ decrease as a function of the attack rate c. The basic reproduction number is fixed to be $\mathcal{R}_0 = 3$, while the duration of the trial is fixed to be $T = 4\tau$, i.e. a period of ≈ 2 months. Lower values of c correspond to initial and final phases of the epidemics where $\bar{i} \ll 1$, while high values of c correspond to experiments performed near the peak of the epidemic. We observe that η tends to underestimate ϵ more when the fraction of infectious individuals is high, while $\eta \approx \epsilon$ in the initial phases where $\bar{i} \ll 1$; in particular, in this regimes the corrections to η (i.e. the systematic errors introduced by using eq. 1) are small and proportional to the attack rate during the trial (see eq. 5).

When the number of individuals is small, the process of getting infected is better described by a stochastic process. We thus perform stochastic simulations of the

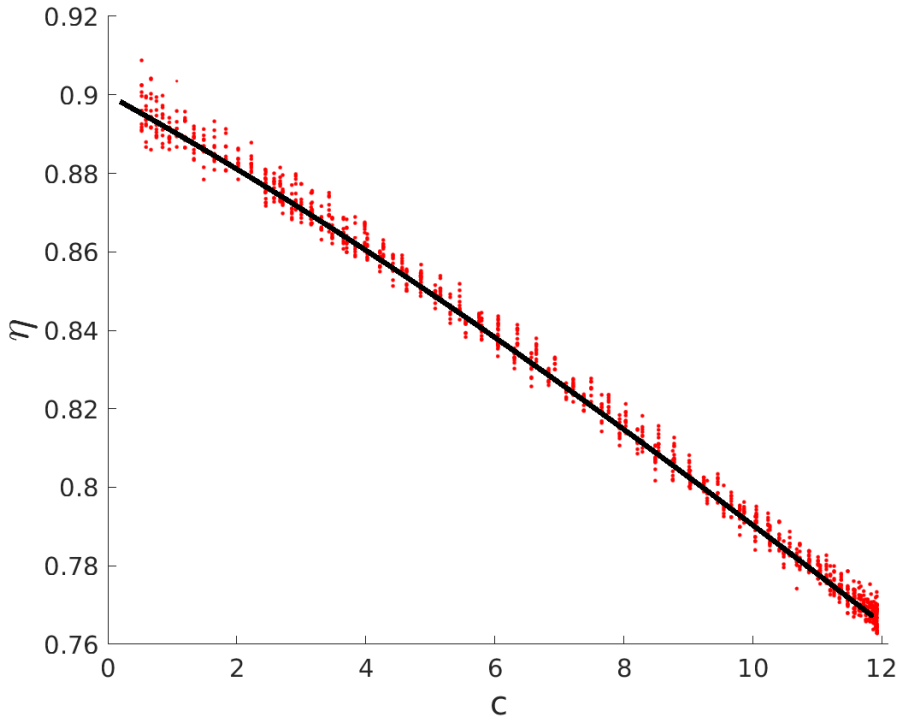

Fig 1. Measured effectiveness η (eq. 1) versus the attack rate c. The epidemic is modelled with a SIR model with basic reproduction number $\mathcal{R}_0 = 3$ and mean infectious period $\tau = 15$ days corresponding to a transmission rate $\beta = 0.2$ days^{-1}. The continuous black line corresponds to the expected values of η (eq. 4) for trials of a duration $T = 2$ months and real efficacy $\epsilon = 0.90$. Curves are obtained by varying the initial time t of the trial; thus, each c corresponds to a period $[t, t + T]$. Lower values of c correspond to initial and final phases of the epidemics where the fraction of infectious individuals is low, while high values of c corresponds to experiments performed near the peak of the epidemic. We observe that η is affected by a systematic error (i.e. $\eta < \epsilon$) that makes it underestimate the real efficacy ϵ; when the fraction of infectious individuals is high, the error is larger, while when it is low, $\eta \approx \epsilon$ and the error is proportional to $\beta\, c$ (see eq. 5). To evaluate the statistical errors, we model the process of getting infected by a stochastic process (eq. 9) and simulate possible values of η for cohorts of $n = 4 \times 10^4$ individuals, i.e. of a size of the same order of the Pfizer trial [12]). As expected, the results of the stochastic simulations (red dots in the figure) fall in a region with a distance of order 10^{-2} around the theoretical curve of eq. 4, i.e a region of order $\pm 1/\sqrt{n}$ as expected for a trial with cohorts of independent, non-interacting individuals.

experiments (see Methods) for cohorts of $n = 4 \times 10^4$ individuals, i.e. of a size of the same order of the Pfizer trial [12]). In Fig 1 we show as red dots the results of stochastic simulations of the trial; the results of the stochastic simulation fall in a region with a relative distance of order 10^{-2} around the theoretical curve of eq. 4, i.e in a region of order $\pm 1/\sqrt{n}$ as expected for a trial with cohorts of independent, non-interacting individuals.

Since the maximum number of infectious individuals is an increasing function of $\mathcal{R}_0$ (see eq. 8), the maximum attainable value of c also increases with $\mathcal{R}_0$; thus, the

worst estimated effectiveness η^{min} (i.e. at the infectious peak) must also be a decreasing function of $\mathcal{R}_0$. In Fig 2, we show that this is the case by plotting η^{min} as a function of $\mathcal{R}_0$ for $\epsilon = 0.90, 0.93, 0.96$ and for a duration of the experiment of 4τ, i.e. $T \approx 2$ months. Notice that, if the vaccine has an higher efficacy ϵ, then η better estimates it: as an example, for $\mathcal{R}_0 = 3$ (a value that has been estimated for COVID19 in France [13]), to an efficacy $\epsilon = 0.96$ corresponds an effectiveness as low as $\eta^{min} \approx 0.90$ (i.e. eq. 1 introduces a systematic error up to $\sim 6\%$), while to an efficacy $\epsilon = 0.90$ corresponds an effectiveness as low as $\eta^{min} \approx 0.77$ (i.e. eq. 1 introduces a systematic error up to $\sim 15\%$).

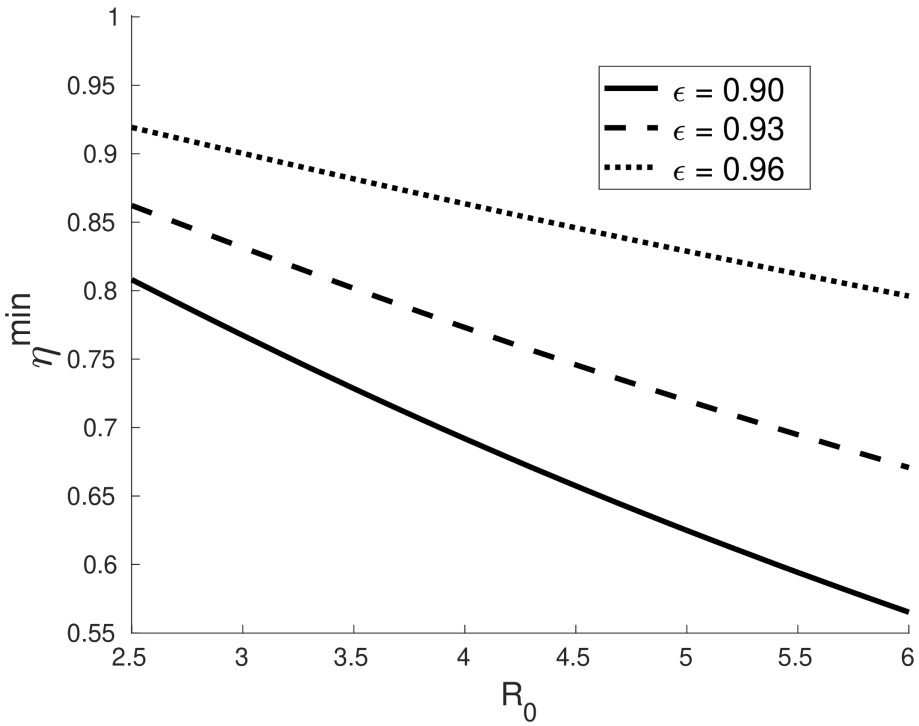

Fig 2. Minimum efficacy vs basic reproduction number. According to eq. 4, the measured effectiveness η (eq. 1) reaches a minimum η^{min} for clinical trials near the epidemic peak. The figure reports the theoretical values of the worst effectiveness estimate η^{min} versus the basic reproduction number $\mathcal{R}_0$ when modelling the epidemic with a SIR of mean infectious period $\tau = 15$ days and considering clinical trials of length $T = 2$ months. The three curves correspond to a true efficacy of $\epsilon = 0.90$ (continuous line), $\epsilon = 0.93$ (dashed line) and $\epsilon = 0.96$ (dotted line). The curves show that the lower a vaccine's efficacy, the worse is its underestimate by the effectiveness (eq. 1).

Discussion

The COVID-19 vaccine rush is pushing governments and developers to set new standards for valid clinical trials in humans, showing that methodological issues in clinical trials can lead to unrepresentative data and communication errors can even fuel vaccine hesitancy [14].

Evaluating vaccine efficacy requires a deep knowledge of the state of the individuals

in the trial and of their contacts; however, this is seldom the case, apart in preliminary phase I and phase II studies where, working with smaller groups in more controlled environments, it is sometimes possible to gain deeper information on the vaccine, both separating the indirect effects of a growing number of infectious in the overall population [15], and eventually discriminating the efficacy in protecting from the disease from the efficacy in dampening the transmission from an infectious vaccinated individual [4]. However, even with a detailed knowledge of the dynamics of the individuals involved in the trial, the statistics of phase I and II is often too low to get accurate enough results. On the other hand, during phase III, large numbers of individuals are enrolled in the trials: thus, statistics is sufficient to reach better accuracy in the estimates, at the price of not being able to have enough detailed information for discriminating various kinds of efficacy, and of being unable to assess the importance of indirect effects. To such an aim, digital contact tracing [16] together with the medical records of the national health service could be paramount to ease all the phases of a pandemic crisis: not only for early detection and isolation of impending outbreaks [17] and for the calibration of pharmaceutical and non pharmaceutical measures [18,19], but also for better phase III estimates for the efficacy of vaccines produced "on the fly" and for the following phase IV evaluation of side effects after they have started to be distributed in the population [14]. However, we have shown that starting to collect more reliable coarser data – like the fraction of infectious individuals in a population – would greatly help interpreting the results of medical trials.

In fact, we have shown how even a coarse knowledge of epidemic data (like estimates of the fraction of infectious versus time) could help to correct phase III (and eventually phase IV) estimates of efficacy from the measured effectiveness. Obviously, more detailed data can help to understand the impact of heterogeneity in contacts [18,20,21] not only on the epidemic dynamics, but also on the estimates of vaccine efficacy and herd immunity thresholds. In particular, digital tracing data would be especially useful for considering the variations of contact patterns due to behavioral data. As an example, vaccinated people may alter their habits if they believe the vaccine is protective; thus, without detailed contact information, the behavioral changes could introduce systematic biases in efficacy trials [10]. In general, human behaviour plays a key role in epidemic spreading, and investigating and quantifying its effects is paramount to effective policies for non-pharmaceutical interventions and vaccination policies [22,23].

Conclusions

Many factors impact the efficacy of a vaccine: from population specific genetic characteristics to partial immunity acquired from previous infections, or even the development of variants during the epidemic: something that, given the duration of the still ongoing pandemic period, has occurred for COVID19. However, our study concentrates on the systematic decrease on the estimated vaccine effectiveness in large cohort studies due to the presence of an high number of infectious individuals in the population. Since vaccines have never been produced, tested and experimented in such exceptional circumstances like the one recently occurred during COVID19, such an issue has not been fully addressed before.

For the sake of simplicity, we have employed a classical epidemiological model with realistic parameters to understand the order of magnitude of the systematic error in efficacy estimates; however, most models of epidemics do not differentiate between infection and disease, while there are cases where that the relation between the biological efficacy of the vaccine and its efficacy as measured by clinical trials is

complex and multi-factorial [4, 24, 25]. Since disease (i.e. observable traits) is what drives behaviour, this is an issue that should be pursued further when modelling epidemics.

In the case of COVID19, the age is a key factor respect to the incidence of severe cases and/or mortality; in fact, efficacy estimates in phase III studies consider the effects of age [12, 26]. In such a case, we will have that effectiveness will be a larger underestimate of efficacy in the subgroups where the latter is lower. Moreover, since we have experienced that the COVID19 vaccine efficacy is time dependent and depends on the number of the doses [27], longitudinal studies should be planned in advance to detail the history-dependence of vaccine efficacy.

Finally, we notice that in the vaccine trials that have occurred it has been observed that antibodies in the vaccinated individuals take time to develop [12, 26]; thus, if also the efficacy grows with time, an extra bias could be introduced in efficacy measurements. In particular, if the trial occurs when the number of infectious is growing, the protection is low at the beginning of the trial, when the probability of getting infected is lower; on the contrary, if the epidemic is decreasing, the vaccine protects less at the beginning, i.e. just when the probability of getting infected is higher. Thus, for two trials – one before the epidemic peak, the other after – with identical time-spans and attack rates, we expect a lower estimate of the vaccine efficacy (i.e. a larger systematic error) for the trial in the decreasing phase.

Methods

Vaccine efficacy

When the frequency of infective events in the susceptible individuals depends on the number of already affected individuals [28], the interpretation of the estimates of a vaccine efficacy can vary depending on the assumptions about the underlying dynamics [4]. Let's assume that, to perform a double-blind evaluation of a vaccine's efficacy, individuals have been divided into two cohorts V (the ones that have received the vaccines) and P (the ones that have received the placebo). Let's also assume that the experimental protocol ensures that: (i) the individual in the cohorts are not in reciprocal contact (the ideal case would be that infectious individuals in the cohorts would remain reciprocally uncorrelated during the experiment); (ii) the infectious dynamics of the cohorts does not influence significantly the ongoing epidemics, i.e the size of the group is much smaller than the population and the observation time is much lower than the total time for the epidemic to evolve. Under these assumptions, infections come only from contacts with infectious individuals outside the cohorts. Thus, assuming full mixing, the probability of meeting an infectious individual is proportional to the fraction i of infectious individuals in the whole population, and the evolution of the fraction of susceptible individuals (i.e. not yet infected) in the P, V cohorts can be written as:

$$\partial_t s_P = -\beta^P \cdot i \cdot s_P \quad , \quad \partial_t s_V = -\beta^V \cdot i \cdot s_V \tag{2}$$

where the transmission probabilities are $\beta^P = \beta$ for the placebo cohort and $\beta^V = \delta\,\beta$ for the vaccine cohort, where $\delta = 1 - \epsilon$. Both equations can be solved yielding the solutions

$$s_P = e^{-\beta^P \cdot c} \quad , \quad s_V = e^{-\beta^V \cdot c} \tag{3}$$

where we indicate with $c = \bar{i}\,T = \int_t^{t+T} i\,d\tau$ the attack rate of the infection for the period $[t, t+T]$, i.e the cumulative fraction of infectious [9] during the trial. The

corresponding attack rates for the cohorts will be $c_P = 1 - s_P$ and $c_V = 1 - s_V$; thus, we can rewrite eq.1 as

$$\eta = \frac{s_V - s_P}{1 - s_P} = \frac{e^{-\delta\beta c} - e^{-\beta c}}{1 - e^{-\beta c}} \tag{4}$$

that tells us that the observed effectiveness η will depend on the attack rate relative to the observation period; notice that such expression is in accordance with the results of [29]. For small values of βc, it is possible to expand eq. 4:

$$\eta = \epsilon \left[1 - \frac{\delta}{2}\beta c + \mathcal{O}(\beta^2 c^2)\right] \tag{5}$$

that tells us that, even for small values of c (i.e. when the average number of infectious $\bar{i}$ is small), there is already a negative correction to the estimate of ϵ by η that is proportional to δ. Also the expansion to the second order

$$\eta = \epsilon \left[1 - \frac{\delta}{2}\left(\beta c + \frac{1 - 2\epsilon}{6}\beta^2 c^2\right) + \mathcal{O}(\beta^3 c^3)\right] \tag{6}$$

retains the same behaviour, since the quadratic term is still negative up to very low efficiency $\epsilon = 0.5$ and the corrections decrease proportionally to δ: the higher the efficiency, the better the estimate.

SIR model

To estimate c, it would be necessary to have accurate data on the fraction of infectious individuals during an epidemic, like the one obtained by testing campaigns. In cases like the COVID19 pandemics where data are scarce and the understanding of the epidemic is still an ongoing process, it is useful to rely on mathematical models whose parameters are tuned on the epidemic's dynamics. For its simplicity and for the few parameters needed, we will use the basic SIR model. In the SIR model the population is divided into three groups S,I,R corresponding to different stages of an infection: S corresponds to susceptible individuals, I to infectious and R to recovered individuals. Indicating with lowercase letters (i.e. s, i, r) the fractions of individuals in a given class, the epidemic is described by the equations

$$\partial_t s = -\beta s i \quad , \quad \partial_t i = \beta s i - i/\tau \quad , \quad \partial_t r = i/\tau \tag{7}$$

where β is the infection rate and τ is the average duration of the infectious period.

For the SIR model, since $\partial \ln s = -\mathcal{R}_0 r$ [19], it is possible to derive the closed solution $\ln s = -\mathcal{R}_0\, r$ for a free epidemic starting from $s(t = 0) = 1$, $r(t = 0) = 0$; thus, since i is maximum when $s = 1/\mathcal{R}_0$ and at this value $r = \ln \mathcal{R}_0/\mathcal{R}_0$, we can explicitly calculate the value of i^{max} from $i + s + r = 1$

$$i^{max} = 1 - \frac{1 + \ln \mathcal{R}_0}{\mathcal{R}_0} \tag{8}$$

showing that in SIR models the maximum fraction of infectious grows as expected with the basic reproduction number following a simple relation with $\mathcal{R}_0$.

Stochastic estimates of the efficiency

While deterministic equations for epidemic dynamics can be a good approximation when the population is large and as soon as there is an extensive number (even if the fraction is small) of infectious [9], in the case of medical experiments cohorts are seldom large enough to disregard statistical fluctuations in the observations. Apart

from particular cases like systems with critical points [30], relative fluctuations for a system of n individuals are of the order $1/\sqrt{n}$. Thus, while for equations like SIR's – describing populations of a size of the order of the inhabitants of a nation – we can disregard fluctuations and we can thus consider i as a good proxy for the evolution of the fraction of infectious, eq.2 does not allow to check for the importance of fluctuations in the experimental setting when the number of cohorts' patients n is not so large. As an example, cohorts of size $n \sim 10000$ are expected to yield relative errors of order $\sim 1\%$.

To estimate such statistical fluctuations, we employ a simplified stochastic approach. Since the V, P experimental cohorts consist of independent and uncorrelated individuals, the infection rates $\beta^X\, i$, $X \in \{V, P\}$ of eq. 2 can be interpreted as independent Poisson rates where each individual in the cohorts has a probability $-\beta^X i$ per unit time to become infected. In a time interval Δt small enough that i can considered constant, the number of infections suffered by a population of S^X individuals will thus follow a Poisson distribution of mean $\beta^X\, i\, S^X\, \Delta t$; thus, the infectious dynamics for the experimental cohorts can be simulated as

$$S^X(t + \Delta t) = S^X(t) - \mathbf{PoissRand}\left(\beta^X \cdot i(t) \cdot S^X(t) \cdot \Delta t\right) \tag{9}$$

where $X \in \{V, P\}$ and $\mathbf{PoissRand}(x)$ generates a random integer number Poisson distributed with rate parameter x. Such an approach has been applied to estimate the fluctuations reported in Fig 1.

Acknowledgments

AS thanks dr. Sandro Bonfigli for his valuable advice on EU and national health service issues, and dr. Emanuele Perugini for his precious advice on data sources for vaccine efficacy.

References

1. Shah A, Marks PW, Hahn SM. Unwavering Regulatory Safeguards for COVID-19 Vaccines. JAMA. 2020;324(10):931. doi:10.1001/jama.2020.15725.

2. Collins FS, Stoffels P. Accelerating COVID-19 Therapeutic Interventions and Vaccines (ACTIV): An Unprecedented Partnership for Unprecedented Times. JAMA. 2020;323(24):2455. doi:10.1001/jama.2020.8920.

3. WHO. Draft landscape and tracker of COVID-19 candidate vaccines; 2021. Available from: `https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines`.

4. Halloran ME, Longini IM, Struchiner CJ. Design and analysis of vaccine studies. Statistics for biology and health. New York: Springer; 2010.

5. Last JM, Association IE, editors. A dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001.

6. Revicki DA, Frank L. Pharmacoeconomic Evaluation in the Real World: Effectiveness Versus Efficacy Studies. PharmacoEconomics. 1999;15(5):423–434. doi:10.2165/00019053-199915050-00001.

7. Singal AG, Higgins PDR, Waljee AK. A Primer on Effectiveness and Efficacy Trials. Clinical and Translational Gastroenterology. 2014;5(1):e45. doi:10.1038/ctg.2013.13.

8. Teerawattananon Y, Anothaisintawee T, Pheerapanyawaranun C, Botwright S, Akksilp K, Sirichumroonwit N, et al. A systematic review of methodological approaches for evaluating real-world effectiveness of COVID-19 vaccines: Advising resource-constrained settings. PLOS ONE. 2022;17(1):e0261930. doi:10.1371/journal.pone.0261930.

9. Diekmann O, Heesterbeek H, Britton T. Mathematical tools for understanding infectious diseases dynamics. Princeton series in theoretical and computational biology. Princeton: Princeton University Press; 2013. Available from: https://press.princeton.edu/books/hardcover/9780691155395/mathematical-tools-for-understanding-infectious-disease-dynamics.

10. Halloran ME, Longini IM, Haber MJ, Struchiner CJ, Brunet RC. Exposure efficacy and change in contact rates in evaluating prophylactic HIV vaccines in the field. Statistics in Medicine. 1994;13(4):357–377. doi:10.1002/sim.4780130404.

11. Alimohamadi Y, Taghdir M, Sepandi M. Estimate of the Basic Reproduction Number for COVID-19: A Systematic Review and Meta-analysis. Journal of Preventive Medicine and Public Health. 2020;53(3):151–157. doi:10.3961/jpmph.20.076.

12. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. New England Journal of Medicine. 2020;383(27):2603–2615. doi:10.1056/NEJMoa2034577.

13. Salje H, Kiem CT, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. Science. 2020;369(6500):208–211. doi:10.1126/science.abc3517.

14. Forman R, Shah S, Jeurissen P, Jit M, Mossialos E. COVID-19 vaccine challenges: What have we learned so far and what remains to be done? Health Policy. 2021; p. S0168851021000853. doi:10.1016/j.healthpol.2021.03.013.

15. Ross R. An application of the theory of probabilities to the study of a priori pathometry.Part I. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character. 1916;92(638):204–230. doi:10.1098/rspa.1916.0007.

16. Moreno Lpez JA, Arregui Garca B, Bentkowski P, Bioglio L, Pinotti F, Bolle PY, et al. Anatomy of digital contact tracing: Role of age, transmission setting, adoption, and case detection. Science Advances. 2021;7(15):eabd8750. doi:10.1126/sciadv.abd8750.

17. Kogan NE, Clemente L, Liautaud P, Kaashoek J, Link NB, Nguyen AT, et al. An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time. Science Advances. 2021;7(10):eabd6989. doi:10.1126/sciadv.abd6989.

18. Scala A, Flori A, Spelta A, Brugnoli E, Cinelli M, Quattrociocchi W, et al. Time, space and social interactions: exit mechanisms for the Covid-19 epidemics. Scientific Reports. 2020;10(1):13764. doi:10.1038/s41598-020-70631-9.

19. Scala A. The mathematics of multiple lockdowns. Scientific Reports. 2021;11(1):8078. doi:10.1038/s41598-021-87556-6.

20. Mistry D, Litvinova M, Pastore y Piontti A, Chinazzi M, Fumanelli L, Gomes MFC, et al. Inferring high-resolution human mixing patterns for disease modeling. Nature Communications. 2021;12(1):323. doi:10.1038/s41467-020-20544-y.

21. Sun K, Wang W, Gao L, Wang Y, Luo K, Ren L, et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. Science. 2021;371(6526):eabe2424. doi:10.1126/science.abe2424.

22. Read JM, Eames KTD, Edmunds WJ. Dynamic social networks and the implications for the spread of infectious disease. Journal of The Royal Society Interface. 2008;5(26):1001–1007. doi:10.1098/rsif.2008.0013.

23. Funk S, Salath M, Jansen VAA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. Journal of The Royal Society Interface. 2010;7(50):1247–1256. doi:10.1098/rsif.2010.0142.

24. Struchiner CJ, Halloran ME, Brunet RC, Ribeiro JMC, Massad E. Malaria vaccines: lessons from field trials. Cadernos de Saude Publica. 1994;10(suppl 2):S310–S326. doi:10.1590/S0102-311X1994000800009.

25. Breslow NE, Storer BE. GENERAL RELATIVE RISK FUNCTIONS FOR CASE-CONTROL STUDIES. American Journal of Epidemiology. 1985;122(1):149–162. doi:10.1093/oxfordjournals.aje.a114074.

26. Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. The Lancet. 2021;397(10269):99–111. doi:10.1016/S0140-6736(20)32661-1.

27. Mller L, Andre M, Moskorz W, Drexler I, Walotka L, Grothmann R, et al. Age-dependent Immune Response to the Biontech/Pfizer BNT162b2 Coronavirus Disease 2019 Vaccination. Clinical Infectious Diseases. 2021;73(11):2065–2072. doi:10.1093/cid/ciab381.

28. Rossi D, Romano E, Guyennon N, Rainaldi M, Ghergo S, Mecali A, et al. The present state of Lake Bracciano: hope and despair. Rendiconti Lincei Scienze Fisiche e Naturali. 2018;doi:10.1007/s12210-018-0733-4.

29. Halloran ME, Struchiner CJ, Longini J Ira M. Study Designs for Evaluating Different Efficacy and Effectiveness Aspects of Vaccines. American Journal of Epidemiology. 1997;146(10):789–803. doi:10.1093/oxfordjournals.aje.a009196.

30. Stanley HE. Introduction to phase transitions and critical phenomena. The International series of monographs on physics. Oxford: Clarendon Press; 1971.