

A physically plausible incidence rate for compartmental epidemiological models

Thomas Pitschel¹, 11th June 2021

Abstract

Motivated by the recent trajectory of SARS-Cov-2 new infection incidences in Germany and other European countries, this note reconsiders the need to use a *non-linear* incidence rate function in deterministic compartmental models for current SARS-Cov-2 epidemic modelling. Employing a homogenous contact model, it derives such function systematically using stochastic arguments. The presented result, which is relevant to modelling of proliferation of arbitrary infectious diseases, integrates well with previous analyses, in particular closes an analytical "gap" mentioned in London and Yorke (1973) and complements the stability related work on incidence rate functions of the form $\beta I^p S^q$ seen for example in Liu, Hethcote and Levin (1987).

Keywords: compartmental model, SIR-model, non-linear incidence rate function, non-linear force of infection

1 Introduction

The massive vaccination campaigns currently underway in the European countries, and nearly completed in some non-European countries such as Canada, the United States, United Kingdom and Israel, have coincided with a rapid decline in new SARS-Cov-2-infection incidences in those countries which is remarkable both in degree and consistency. Given the virus- and contact specific parameter estimates of epidemic models derived just a few months ago, and given the vaccine effectiveness parameters, one may wonder how the currently observed incidence decline could possibly be explained with previously used models.

In Germany, based on data from [1]-[5], the expected overall reduction in infection probability achieved with the current state of vaccination is by a factor of about 0.6391866.² Given that the "unconstrained" (i.e. preventive measures-free) R_0 was estimated at about 3, and the infection

¹Correspondence address: th.pitschel (at) uni (dot) de

²In Germany, according to [1], as of 29th May 2021 there had been administered about 45.3 mio anti-SARS-Cov-2 vaccine doses, and about 14.20 mio persons (17.1% of the population) had been completely vaccinated. In about 42.2% of the population at least one vaccine shot had been administered. According to [2], about 50.5 mio doses had been delivered until 23th May, and of those 97.5% had been administered until 28th May; The distribution of vaccine types among those delivered doses are [2]: 70.5% of BioNTech/Pfizer (BNT162b2), 19.2% of AstraZeneca (AZD1222), 9.4% of Moderna (mRNA-1273), 0.9% of Johnson&Johnson (JNJ-78436735). Except from the JNJ-78436735, which requires one shot to be administered to completion, the three other vaccines require two shots. The reported effectiveness of the vaccines in preventing (symptomatic³) infection was stated as: 82%/95% after first/second shot of BNT162b2 [3], and 76%/81.3% after first/second shot of AZD1222 [4]. For Moderna's mRNA-1273 we assume the same effectiveness numbers as for BNT162b2. If we regard the effectiveness denoting an equivalent reduction of *probability* of infection, and assume vaccine type distribution in the administered doses equals those in the delivered doses, and assume equal vaccination progress in each vaccine type, and neglect J&J vaccines from the calculation, then one must tend to conclude that the overall reduction of infection probability achieved by the current state of vaccination, in Germany, is by an aggregate factor of

$$0.578 \cdot 1 + [(0.705 + 0.094) \cdot (0.251 \cdot 0.18 + 0.171 \cdot 0.05) + 0.192 \cdot (0.251 \cdot 0.24 + 0.171 \cdot 0.187)]/0.991 \approx 0.6391866.$$

Here, the 0.578 represents the share of the not-yet-vaccinated population in which no reduction takes place. The 0.991 adjusts for Johnson&Johnson numbers being omitted.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

probability enters linearly into R_0 in the proliferation models employing a bilinear incidence rate function⁴, it is unreasonable that the current vaccination state should suffice to effect an incidence decline robust even under lifting of strong preventive measures such as contact restrictions and mask-wearing.

It is therefore hypothesized that the better-than-predicted effect of vaccination is due to another mechanism. The mechanism here supposed is that single-shot ("1s") vaccinations *raise the tolerance level* for virus exposure in these 1s-vaccinated susceptibles with the effect that *multiple* exposures with virus particles would be required to trigger an infection. (One may regard as suitable criterion the total virus dose an individual is being exposed to over one day, say, and a threshold on this value.) This in turn is deemed to have the consequence that instead of mere reduction of infection probability, an infection is assuredly prevented in specific contact instances in which virus density remains below a certain threshold level ("threshold effect"). As will be seen in the proposed quantitative analysis, the mathematical description of this mechanism leads to a non-linear incidence function in the on-average (= "deterministic") epidemic modelling. The resulting trajectories then display the larger-than-predicted decline in incidences. As a secondary cause, besides the direct effect of vaccination, the consequential shift of the majority of proliferation events into younger population segments can be stated, which has the same effect of raising the average tolerance level.

The here suggested mechanism has been mentioned as possibly relevant in epidemic modelling, and analytically (nearly) been provisioned for, in much earlier work: In [5], the accounting for multiple exposures is named as a reason to use non-linear incidence functions, and consequently incidence rate functions of form $\sim I^p S^q$ are examined. In [6] the saturation effect, namely that the new infections cannot scale unboundedly linearly with existing number of infectives, is stated to motivate analysis involving non-linear incidence functions, concretely of form $g(I) \cdot S$, with non-negative bounded differentiable g with additional conditions. Most of this previous work however does not focus on a precise translation from physical mechanism of the transmission to a corresponding appropriate mathematical representation of the incidence rate, but often appears to use functional forms guided by tractability or data-matching. What is new in this present note is to give a *systematic* derivation of the non-linearity based on stochastic arguments which properly account for both the saturation and threshold effect. The arguments assume a stylized (in particular: homogenous) contact setting which can be realistic for a substantial part of the contacts occurring in an urbanly situated population ($N \approx 100000 \dots 1000000$).⁵ Focus is on the overall effect in quantitative terms *under assumption* that the stated mechanism is present.

2 Framework epidemic model and derivation of non-linearity g

The result to be derived below is supposed to be a building block in the well-known compartmental epidemic models describing the proliferation in terms of the number of susceptible (S), of "exposed"⁶ (E), infectious (I) and removed (R) individuals, for example the one stated as eqn.

⁴Have $R_0 = \mu \cdot T_I = d \cdot p \cdot T_I$, where μ is a transmission coefficient per time unit, T_I is the average infectiousness period, d is a contact degree and p is an aggregate transmission and reception probability.

⁵In the author's opinion the arguments stated in [7] for focussing merely on concave incidence rates (on biological grounds) are not plausible: The threshold effect will be relevant for any communicable disease in which the counts of transmitted disease carriers can be reduced, say by physical preventive measures (e.g. using masks), or the tolerance level can be raised, as long as a dose-response relationship is at work which has sufficient smoothness. This dose-dependency of the response (and ultimately: of secondary infectiousness) is present e.g. for COVID-19 at least in the younger (i.e. usually more healthy) population ([8,9] with [10] (pg 1043)).

⁶As will become clear later, this group in the present text would rather be called "the individuals in which infection is triggered but who are not yet infectious".

(2.2) in [11]:

$$\frac{dE}{dt} = \beta \cdot g(I)(1 - E - I - R) - (\epsilon + \mu)E, \quad \frac{dI}{dt} = \epsilon E - (\gamma + \mu)I, \quad \frac{dR}{dt} = \gamma I - (\delta + \mu)R \quad (1)$$

with $S(t) + E(t) + I(t) + R(t) = 1$. The time-dependent quantities are understood as *shares* of the respective compartment size relative to the (constant) total population size, and are deemed expected values. The β above is a transmission probability and contact mechanism related coefficient, but for the analysis below it will be assumed equal to one. With $S(t)$ being the share of susceptibles, the $g(I) \cdot S(t)$ then takes the meaning of the number of new infections triggered by all infectious among the total population in a unit time interval, divided by N .

Historically, the analysis began with the incidence function g chosen as linear in I , based on the idea that infections caused by an infective are occurring stochastically independently of the other infections (which is approx. true as long as the infectives density is low) and a single contact suffices for infection. At low infectives counts, i.e. $I \ll 1$, the other members of the population encountered by a specific infective will be nearly disjoint from the ones encountered by any other infective. (I.e. there is no coverage overlap.) Then $E(\#\{\text{"exposed" individuals by any infective}\}) \approx \sum_{i=1}^{N \cdot I(t)} E(\#\{\text{"exposed" individuals by infective } i\}) = I(t) \cdot \text{const.}$ For larger I however, the left-hand side expectation must be bounded by N , and the linearity in I disappears.

To derive $g(I)$ systematically for any value of I , regard the observation time interval be divided into multiple time slots, and allow any infective to make multiple contacts over the observation interval to other members of the population. Concretely, let $N \cdot I(t)$ be the number of infectives being active at the beginning t of the observation time interval, and let each of them make on average d contacts with other members during that interval. The contacts are deemed to uniformly randomly covering the population of N members, one at a time slot. This is analogous to dialing out a deck of $d \cdot N \cdot I$ cards onto N places, for each card choosing one of the places according to the uniform distribution, such that each place may receive multiple cards. We want to compute the expectation of the number of places $K_T = K_T(d \cdot N \cdot I, N, \kappa)$ among the N potential receivers that receive equal or more than $\kappa \in \mathbb{N}$ cards, corresponding to the number of individuals exposed to an infective at least κ times during the observation interval, which is to constitute the condition upon which infection is triggered. (It is assumed here that any contact constitutes an exposure in the sense of contributing to the triggering of an infection; any transmission probability that may exist to describe events of non-exposure inspite of contact is assumed to be absorbed into the value of d . Part of the original numerical value of β is to appear here.) Since the susceptibles are in turn independently distributed among the members of the population, it then follows the expected number of infections triggered in the interval is $E(K_I) = E(K_T) \cdot S(t)$.

The probability p_0 of a specific member of the population being "hit" by zero infectives is given by $((N - 1)/N)^{dNI}$. We will write $d \cdot I =: \tilde{I}$ and focus on large N . Then $p_0 \approx \exp(-\tilde{I})$. The probability p_1 of a specific member being hit once by an infective is

$$p_1 = \left(\frac{N-1}{N}\right)^{dNI-1} \cdot \frac{1}{N} \cdot \binom{dNI}{1} = \left(\frac{N-1}{N}\right)^{dNI} \cdot \frac{1}{N-1} \cdot \frac{NdI}{1} \approx \exp(-\tilde{I}) \cdot \tilde{I} \quad (2)$$

Generally, the probability p_k of a population member being exposed exactly k times is found to be

$$p_k = \left(\frac{N-1}{N}\right)^{dNI-k} \cdot \frac{1}{N^k} \cdot \binom{dNI}{k} = \left(1 - \frac{1}{N}\right)^{dNI} \cdot \frac{1}{(N-1)^k} \cdot \frac{(dNI)!}{(dNI-k)! \cdot k!} \quad (3)$$

For k small compared to N and compared to $\tilde{I} \cdot N$, it is $p_k \approx \exp(-\tilde{I}) \cdot \frac{\tilde{I}^k}{k!}$. (The random variable counting the exposures at a specific population member approximately is found to be Poisson distributed with intensity \tilde{I} .) The approximate values are normalization-consistent. The probability of a specific member being exposed at least once consequently is $(1 - p_0) \approx 1 - \exp(-\tilde{I})$; of being exposed at least κ times it is $(1 - p_0 - p_1 - \dots - p_{\kappa-1}) \approx 1 - \exp(-\tilde{I}) \cdot (\sum_{k=0}^{\kappa-1} \tilde{I}^k/k!)$.

Based on these elementary arguments, one derives the expected values as follows: Since all members of the population are here deemed equal in their role as potential virus receivers, it is

$$E(K_T) \approx N \cdot \sum_{k=\kappa}^{\infty} p_k = N \cdot \left(1 - \exp(-\tilde{I}) \cdot \left(\sum_{k=0}^{\kappa-1} \tilde{I}^k/k! \right) \right). \quad (4)$$

As argued before, $E(K_I) = E(K_T) \cdot S(t)$. The incidence function g derived by above arguments thus is

$$g(I) = \left(1 - \exp(-\tilde{I}) \cdot \left(\sum_{k=0}^{\kappa-1} \tilde{I}^k/k! \right) \right) \Big|_{\tilde{I}=d \cdot I}. \quad (5)$$

2.1 Discussion

As intended, the incidence rate function here is bounded by 1 even as $d \cdot I$ increases to infinity. Further, the historical bilinear form βIS is recovered as special case by observing that for $\kappa = 1$, i.e. a single exposure is deemed sufficient for triggering an infection, the first order approximation of the exponential at \tilde{I} small yields

$$E(K_T) = N \cdot (1 - \exp(-\tilde{I})) \approx N \cdot (1 - (1 - \tilde{I})) = N \cdot d \cdot I \quad (6)$$

$$E(K_I) = N \cdot d \cdot I \cdot S \quad (7)$$

$$g(I) = d \cdot I. \quad (8)$$

The next higher order approximation can be verified to equal $E(K_T) = N \cdot (\tilde{I} - \tilde{I}^2/2) = N \cdot d \cdot I(1 - dI/2)$, yielding the first component accounting for coverage overlap leading to the saturation phenomenon. The stochastic analysis of the here employed homogenous contact mechanism is found to motivate well the non-linear modification of the historical form introduced in [12] (pg 471) based on empirical grounds ("to eliminate systematic differences [to observed data]"), with additional benefit of attaching a meaning to the arising coefficient.

The case of multiple required exposures ($\kappa > 1$):

It is easy to see that the leading monomial in g , for small values of \tilde{I} , will always have degree κ . Exemplarily, one has as approximations for small \tilde{I} :

$$g(I) = 1 - \exp(-\tilde{I}) \cdot (1 + \tilde{I}) \approx \tilde{I}^2/2 - \tilde{I}^3/3 \pm \dots \quad \text{for } \kappa = 2, \text{ and}$$

$$g(I) = 1 - \exp(-\tilde{I}) \cdot (1 + \tilde{I} + \tilde{I}^2/2) = \exp(-\tilde{I}) \cdot (\exp(\tilde{I}) - (1 + \tilde{I} + \tilde{I}^2/2)) \approx \tilde{I}^3/6 - \tilde{I}^4/8 \pm \dots$$

for $\kappa = 3$. This leads, at small \tilde{I} , to a polynomial behaviour of the incidence function, namely $g(I) \cdot S \sim I^\kappa \cdot S$, in line with the form previously analysed in [5] and subsequent work. Based on the host response characteristics determined for COVID-19 in medical literature and shortly described in footnote 5, one concludes that for COVID-19 such super-linear incidence rate behaviour will be relevant rather in younger population segments. These are exactly the segments that constitute the currently remaining "theatre of proliferation" after the advancement of the vaccination campaigns.

References

- [1] Impfquotenmonitoring. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Impfquotenmonitoring.html (retrieved 2021-05-30).
- [2] Impfdashboard. <https://impfdashboard.de/> (retrieved 2021-05-30).
- [3] Vaccines and related biological products advisory committee december 10, 2020 meeting briefing document - FDA, Dec 2020. <https://www.fda.gov/advisory-committees/advisory-committee-calendar/vaccines-and-related-biological-products-advisory-committee-december-10-2020-meeting-announcement>.
- [4] Voysey M, Costa Clemens SA, Madhi SA, Weckx LY, Folegatti PM, Aley PK, and et al. Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials. *Lancet*, 397(10277):881–891, Feb 2021.
- [5] Wei-min Liu, H.W. Hethcote, and S.A. Levin. Dynamical behavior of epidemiological models with nonlinear incidence rates. *J. Math. Biol.*, 25:359–380, 1987.
- [6] Vincenzo Capasso and Gabriella Serio. A generalization of the Kermack-McKendrick deterministic epidemic model. *Mathematical Biosciences*, 42(1):43–61, 1978.
- [7] Andrei Korobeinikov and Philip K. Maini. Non-linear incidence and stability of infectious disease models. *Mathematical Medicine and Biology*, 22:113–128, 2005.
- [8] Monica Gandhi, Chris Beyrer, and Eric Goosby. Masks do more than protect others during COVID-19: Reducing the inoculum of SARS-CoV-2 to protect the wearer. *Journal of General Internal Medicine*, 35(10):3063–3066, Oct 2020. PMC7393808.
- [9] Alok Bhargava. Masks reduce viral inoculum of SARS-CoV-2. *Journal of General Internal Medicine*, 36(4):1123, April 2021. PMC7799398.
- [10] Daniel Blanco-Melo, Benjamin E. Nilsson-Payant, Wen-Chun Liu, and et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell*, 181(5):1036–1045, May 2020. PMC7227586.
- [11] H. W. Hethcote and P. van den Driessche. Some epidemiological models with nonlinear incidence. *J. Math. Biology*, 29(3):271–287, 1991.
- [12] James A. Yorke and Wayne P. London. Recurrent outbreaks of measles, chickenpox and mumps, II. systematic differences in contact rates and stochastic effects. *Am. J. of Epidemiology*, 98(6):469–482, 1973.