Incubation and latency time estimation for SARS-CoV-2

Arntzen, V.H.

Future directions

Contents

In incubation and latency time estimation, several biases arise from the misfit between gathered data and its analyses. This thesis contributes to the awareness of such biases and to approaches for unbiased estimation. In this chapter, we outline unexplored future directions that arose during our research. These concerning the manifestation of different phenomena in real data, the data collection process and models to inform quarantine length.

6.1 Manifestation in real data

In contrast to simulation studies, in reality it is hard to determine which phenomena are present in the data that we face, let alone to which extent these may bias the estimates when unaddressed in the analysis.

Develop understanding of the presence of different phenomena in real data

There would be merit in developing methods to test or quantify the extent to which each phenomenon is present in real data sets. A straightforward way to do this is to perform

extensive simulations including a broad range of scenarios to which real data may be compared. Particularly relevant are differential recall of exposures and the time-varying risk of infection, that we will discuss in the next two sections.

Study differential recall interdisciplinary

To gain a deeper understanding of the way differential recall, as we conceptualized this in Chapter 3, manifests in contact tracing data, collaboration with researchers in psychology would be beneficial. Earlier studies [Heuch et al., 2018; Salehabadi et al., 2014; Moshiro, 2005; Yoo et al., 2017] on memory decay concerned other event types and examined recall months or years after the event, a time lag that is longer than would be relevant to SARS-CoV-2, where notified cases were typically interviewed days or weeks after exposure.

One could conduct an experiment in which individuals are asked about their contacts in the past few weeks, while these individuals also report their contacts on a daily basis via a simple questionnaire form. This way, one can study the discrepancy that arises between the true exposure(s) that participants report on the day itself and how participants memorize their different exposure(s) a few weeks later. Another interesting direction related to how recall can be improved. The experimental setup as described before allows to study the effect of different suggestions for aided recall provided in Chapter 3.

Improve the goodness-of-fit to the tail of the distribution

As we discussed in Chapter 2 it is common practice to model incubation and latency time with a gamma, lognormal or Weibull distribution and then choose the model that provides the best fit based on a criterion like AIC or LOO-CV. Ideally, the choice of the family of distributions would be supported by a true understanding of the underlying disease mechanisms. Studying the latter using observations of within-host dynamics would be helpful [Nishiura, 2007]. Moreover, it would be valuable to gain insight in the determinants of exceptionally long incubation and latency times: whether these are extreme observations from the same underlying distribution or characterised by an essentially different response to infection.

Instead of selecting a parametric distribution from the abovementioned triplet, we proposed to use a more flexible distribution. We modelled incubation time using a penalized Gaussian mixture (Chapter 2) that offers an adequate fit to the tail of the distribution.

Modelling incubation or latency time using the generalized gamma distribution (Chapter 4) makes the relatively arbitrary choice of parametric distribution from the three distributions redundant because it includes these as special cases. However, the generalized gamma distribution may be less suitable for small data sets due to the additional parameter in comparison to gamma, lognormal and Weibull distribution.

A valuable contribution would involve developing a routine to select one of the more commonly used distributions based on a goodness-of-fit measure that is specific for the tail, that is crucial for informing quarantine length. One can think of an adapted version of the bootstrap-based GPD (generalized Pareto distribution) test [Villaseñor-Alva and González-Estrada, 2009] or similar, that can be borrowed from extreme value theory, which is the field of statistics that focuses on the tail of distributions [Charras-Garrido and Lezaud, 2013]. The generalized Pareto distribution is a family that includes amongst others heavy-tailed distributions [Villaseñor-Alva and González-Estrada, 2009].

Develop a framework to distinguish outliers from erronous observations

A caveat related to estimation of the tail of the distribution is the absence of a statistically sound method to handle outliers in incubation and latency time data. Determining whether extreme time-to-event can be considered biologically plausible or should be considered outliers can be challenging. Currently, infected individuals with an exceptionally large minimum incubation time (from end of exposure to symptom onset) are labeled as recall biased and excluded from analysis [Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021], despite that coronaviruses are known to have a heavier right-hand tail in their distribution [WHO, 2003]. Hence, it is worthwhile to develop a framework to distinguish outliers and extreme time-to-event values. The rule of thumb proposed by Tukey (1977) - considering a possible outlier as an observation larger than 1.5 interquantile range (IQR) from the corresponding quartile Q1 or Q3 - cannot be reliably applied to distributions with tails heavier than that of the normal distribution.

An easy-to-implement outlier detection method for skewed distributions was suggested by Junsawang et al. [2021], but it cannot be used for single or doubly interval censored observations. To address this limitation one can employ multiple imputations of the infection day while keeping track of the outlier classification for each imputed data set. Those values that are often classified as outliers may be treated accordingly.

6.2 Data collection

An important lesson I learned from a rowing coach is that any problem we encounter is often a 'symptom' of something that occurred earlier. For efficient problem-solving, it is worth looking at the root cause, rather than focusing solely on the symptom itself. Upon closer examination, we see that the actual limiting factor in incubation and latency time estimation may be the data *collection* process. Ideally, we would obtain a representative, informative and prospectively collected sample.

Collect detailed exposure information

As discussed in Chapter 2, during the exponential growth phase, the assumption of a constant risk within the exposure window leads to overestimation of the incubation (or latency) time. We addressed this bias in Chapter 4 where we assume an increasing infection risk in line with the exponential growth of the incidence of new cases in the population. However, it is uncertain whether this risk holds on an individual level. For example, when an individual did not make any contacts during a part of their exposure window, the infection risk within the individual's exposure window is unlikely congruent with the incidence in the population. A valuable contribution would be to explore more fine-grained assumptions for the infection risk within the exposure window, necessitating to retrieve more detailed information on exposure history of each individual. We first discuss collection of this data and discuss the corresponding alternative assumption in the next section.

As observed during SARS-CoV-2 pandemic, contact tracing capacities are often limited. Thus, it is unfeasible to collect detailed information about an individual's numerous potential contacts including their duration and the risk of transmission through interviews that are taken upon case notification. The consequence of this lack of information, is that the researcher needs to make post-hoc choices regarding the exposure window that may be relatively arbitrary. A possible alternative would be to ask individuals to rank the different exposure days themselves. This task can be performed timely, for example in the waiting room of a test facility, and consists of ranking their recent exposures based on their perceived risk of infection (e.g. the contact was coughing) and proximity (e.g. in- or outside, type of contact), from highest to lowest risk.

Prevent post-hoc choices to obtain objective estimates

The rationale behind collecting the described, subjective data is that it may eventually lead to less biased estimates than the post-hoc choices of the researcher, that can be even more subjective. Currently, when researchers suspect that the true infection risk increases or decreases strongly over time, the analysis is restricted to observations with a narrow exposure window. As discussed before, this limits the bias imposed by assuming a constant risk of infection within the exposure window. Our sensitivity analyses using data from Vietnam (Chapter 4) indicate that indeed, this procedure limits the bias due to violation of the constant risk assumption as our results assuming a constant risk or exponential growth were comparable making such a selection. However, in Chapter 3 we found that in the presence of differential recall this practice yields underestimation.

Weighting the different partitions of the exposure window, where within each partition a constant risk can be assumed, may be more realistic. Hence, we can assume a piecewise constant risk of infection within an individual's exposure window, parameterised using the weights as described above. When we regard the latter assumption valid, this enables the use of all rather than only a selection of observations with narrow exposure windows for analysis. However, the effectiveness and validity of this approach must be confirmed through a simulation study in which the accuracy of the 'weights' reported by infected individuals is set to imperfect levels. In order to realistically mimic the quality of the exposure rankings, one may seek advice from experts in the field of human memory research. Additionally, an extension of the current software is required to incorporate weights for the different partitions of the exposure window. The `R` package described in Chapter 4 can be easily extended to suit the latter requirement.

Consider prospective data collection

A prospective cohort study is complicated as cases can usually only be confirmed upon symptom onset or positive test for a pathogen. However, we can define the cohort differently: *potentially* infected individuals. For example, all visitors to a certain event with elevated risk of transmission can be a cohort, or individuals who start quarantine around the same calendar time. Note that in the example of quarantine, individuals may have been infected before. However, when the time origin coincides with the start of follow-up, the cohort is prospective. We come to that later.

Explore whether cure models can be employed in the infectious disease context

Suppose that one would monitor such a cohort by means of regular checkups for the onset of symptoms and routine (PCR-)testing. Then, by the end of follow-up, it is possible to distinguish two types of observations regarding the endpoint, that can be (i) observed as the individual developed symptoms or tested positive for the infection or (ii) unobserved as the the individual was still event free at the end of follow up. For an observation of type (ii), it is not possible to distinguish whether the individuals was infected or not. In Chapter 4 we addressed right truncation in the analysis as these individuals were unnoticed. However, individuals that are quarantined and neither develop symptoms, nor test positive during follow-up may carry useful information. This merits to examine another approach that includes all quarantined individuals.

We can model the fraction of uninfected individuals and distribution of the time-to-event jointly by employing a cure model. Cure models [Amico and Van Keilegom, 2018] are useful in different contexts, for example in childhood cancer research. Traditional time-to-event models assume that all individuals remain at risk for the event (for example: death). However, it may occur that a fraction of the individuals will never experience the event of interest. For example when children recover completely from childhood cancer, from a statistical perspective they have infinite survival times and are 'cured'. Cure models are a special type of joint models that consist of two parts that are fitted simultaneously: (i) a model that describes the distribution of time-to-event, for example incubation or latency time; (ii) a model for the probability of cure. Applying this model to the infectious disease context, part (i) can model incubation or latency time, whereas part (ii) would model the probability to be uninfected during follow-up.

Include covariates for personalized estimates of incubation and latency time

Including covariates for both submodels is necessary for identifiability. Therefore, a caveat of the cure model approach is that it requires information about determinants of developing the infection, such as the exposure type, the vaccination status et cetera, and factors determining the time-to-event. Such information may not be part of the standard questionnaire forms that are typically used to collect contact tracing data. However, when these covariates are included and the cohort is similar to individuals that enter quarantine, it may be useful to inform a personalized quarantine length.

When the effect of different individual characteristics is known, the probability to be uninfected after a certain number of days in quarantine as well as the probability to develop symptoms or experience start-of-infectiousness afterwards can be estimated on an individual level. To this end, it can be worthwhile to consider the time of quarantine entry as the time origin, and thus to model the time from quarantine entry to symptom onset or start-of-infectiousness. Note that such an estimate needs to be updated repetitively as calendar time elapses, because the time lag between infection and quarantine entry varies per place and time. Other future directions concerning models to inform quarantine length will be given in the next section.

The above described approach potentially solves several issues that we discussed before: (i) extreme value handling as the assumption that every individual experiences the event eventually is not needed (the probability to be uninfected increases during an individual's follow-up, for example during quarantine); (ii) milder differential recall as data is collected prospectively (Chapter 3); (iii) right truncation (Chapter 4) is naturally addressed as we include the full initial cohort for analysis and the model takes into account that unobserved, long time-to-event cannot be distinguished from uninfected individuals.

6.3 Informing quarantine length

Incubation and latency time are important quantities that inform quarantine length. Quarantine is cumbersome, and therefore, we must broaden our scientific knowledge to optimally inform effective quarantine policies.

The optimal quarantine length depends on many factors

The optimal quarantine policy strongly depends on the spatio-temporal and cultural context. During the SARS-CoV-2 pandemic policy objectives varied worldwide, from aiming for zero transmission (Chapter 4) to mitigating spread, depending on the location and time frame.

Looking solely at infectious disease transmission, quarantine 'fails' when an individual leaves quarantine and becomes infectious afterwards. When an individual enters quarantine relatively late, the chance of 'failure' is typically smaller [Li, Yuan, Chen, Song and Ma, 2021]; however, transmission may have occurred before entry of quarantine. Besides the time elapsed between infection and quarantine and quarantine length. The effectiveness

of quarantine depends on many more factors: e.g. how infectiousness varies during the course of the disease; how many contacts an individual has after quarantine; how convenient it is for individuals to adhere to the policy. Research from Germany indicates that public acceptance of lockdown depends more on the duration than on its intensity or flexibility [Gollwitzer et al., 2020]. The same probably holds for quarantine length.

Develop comprehensive models to optimally inform quarantine length

Ashcroft et al. mathematically modelled the effectiveness of different quarantine strategies where the length of quarantine was varied [Ashcroft et al., 2021]. The researchers distinguished a standard scheme and a test-and-release scheme. Following the latter scheme, individuals were dismissed earlier when testing negative on a specific test day. Additionally, the probability of adherence was varied in their simulations.

A fruitful direction would be to examine how the optimal quarantine length can be informed in real-time, by means of a comprehensive model that takes into account the disease characteristics as well as other factors relevant to decision making, for example from the social domain, economics and logistics. Note that given the different type of data sources that would feed into such a model, its formulation is not straightforward. Obviously, unbiased estimates of the incubation and latency time distributions are crucial input to such models.