Incubation and latency time estimation for SARS-CoV-2

Arntzen, V.H.

Incubation and latency time estimation for SARS-CoV-2

Vera Arntzen

Incubation and latency time estimation for SARS-CoV-2

Vera Arntzen

Incubation and latency time estimation for SARV-CoV-2

Vera Hermina Arntzen

The research presented in this thesis was performed at the Mathematical Institute, Leiden University, Leiden, the Netherlands and Oxford Clinical Research Unit (OUCRU), Ho Chi Minh City, Vietnam.

Cover Laetitia Koning & Vera Arntzen
Layout Vera Arntzen
Printed by PrintSupport4U, Steenwijk
ISBN 978-94-93289-57-4

Incubation and latency time estimation for SARV-CoV-2

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 16 oktober 2024
klokke 16:00 uur

door

Vera Hermina Arntzen

geboren te Amsterdam
in 1994

Promotor

Prof. dr. M. Fiocco

Copromotor

Dr. R.B. Geskus (University of Oxford)

Promotiecommissie

Dr. M. Choisy (University of Oxford)

Prof. dr. ir. G.L.A. Derks

Prof. dr. G. Gómez Melis (Universitat Politècnica de Catalunya - Barcelona Tech)

Prof. dr. H. Putter

Prof. dr. F. Spieksma

to curiosity

Preface

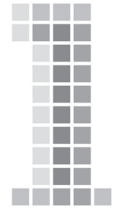
When the work for the thesis in front of you started, the SARS-CoV-2 pandemic had been ongoing for a couple of months already. Suddenly, the novel coronavirus confronted many of us with incredible losses, whether these were lives, health, the proximity to our loved ones, the fearlessness of living in a world free of pandemics or trust in science. At the same time, the world saw how researchers developed the first mRNA-based vaccines against SARS-CoV-2 and rapid antigen tests, and how many countries reduced transmission levels with public health measures.

Humans are a creative and resilient species, yet viruses are so, stressing the importance of continuously extending our knowledge of pathogens. This thesis focuses on estimating the distribution of two critical quantities of an infectious disease to inform public health measures: incubation and latency time.

Contents

List of Abbreviations

AIC . . . . . Akaike information criterion (goodness-of-fit, frequentist paradigm)
AIDS . . . . . acquired immunodeficiency syndrome
BRCA . . . . . breast (and ovarian) cancer associated (genes)
CI . . . . . Confidence Interval (uncertainty bounds, frequentist paradigm)
Covid-19 . . . . . coronavirus disease
CRC . . . . . colorectal cancer
CrI . . . . . Credible Interval (uncertainty bounds, Bayesian paradigm)
DAG . . . . . directed acyclic graph
GWAS . . . . . Genome Wide Association Studies
HIV . . . . . human immunodeficiency virus
LOO IC . . . . . leave-one-out information criterion (goodness-of-fit, Bayesian paradigm)
MLE . . . . . Maximum Likelihood Estimator
MSE . . . . . mean squared error
NP(MLE) . . . . . nonparametric Maximum Likelihood Estimator
PCR . . . . . Polymerase Chain Reaction (laboratory method)
PG(M) . . . . . penalized Gaussian mixture
PRS . . . . . polygenic risk score
reBias . . . . . relative Bias
RNA . . . . . ribonucleic acid (genetic material)
SARS . . . . . severe acute respiratory syndrome (lung infection due to SARS-CoV)
SARS-CoV-2 . . . severe acute respiratory syndrome coronavirus 2
SE . . . . . Standard Error
SNP . . . . . Single Nucleotide Polymorphism (genetic variant)



Introduction

Contents

1.1 Incubation time and latency time distribution

Incubation and latency time of an infectious disease are crucial quantities to understand and manage infectious disease outbreaks. The incubation time is the time interval from infection to symptom onset, and the latency time refers to the period between infection and the start of infectiousness. The corresponding distributions are typically estimated at the beginning of an outbreak with a novel pathogen. Table 1.1 gives some examples of different ways in which incubation time is used [Nishiura, 2007].

Table 1.1: Common uses of the incubation period distribution of infectious diseases [Nishiura, 2007]. The author distinguishes the major field of use and the various functionalities.

Major field of use	Explanation and example
Clinical practice	Rough estimates of the time of exposure of bedside cases (e.g., to determine the causes and/or sources of infection)
	Development of a treatment strategy that extends the incubation period (e.g., antiretroviral therapy for HIV/AIDS)
	Early projection of disease prognosis when the incubation period is clearly associated with clinical severity (e.g., diseases caused by exotoxin)
	Clinical investigations of the impact of infecting dose on the clinical appearance of a disease (i.e., the dose-response mechanism)
Public health practice	Determination of the length of quarantine required for a potentially exposed individual (e.g., limiting the movement of those exposed to SARS within a household)
Epidemiologic study	Determination of the eradicability of a disease (e.g., determination of the effectiveness of isolation measures)
	Estimation of the time of exposure during a point source outbreak (e.g., in identification of the source of infection during large-scale food poisoning)
	Determination of the end of a point source outbreak (i.e., statistical tests that determine if case onset is over)
	Reconstruction of epidemic curves and short-term predictions of slowly progressing diseases (e.g., backcalculation of HIV/AIDS and prion diseases)
	Estimation of the transmission potential and infectiousness relative to disease-age (e.g., estimation of the relative infectiousness of smallpox)
Ecological study	Determination of the adaptation strategy of a parasite (e.g., evolution of vivax malaria owing to seasonal selection pressure)

Estimates of the incubation time and latency time are not necessarily the same because when an infected individual starts to be infectious may not coincide with symptom onset. For SARS-CoV-2, presymptomatic transmission was observed [Tindale et al., 2020].

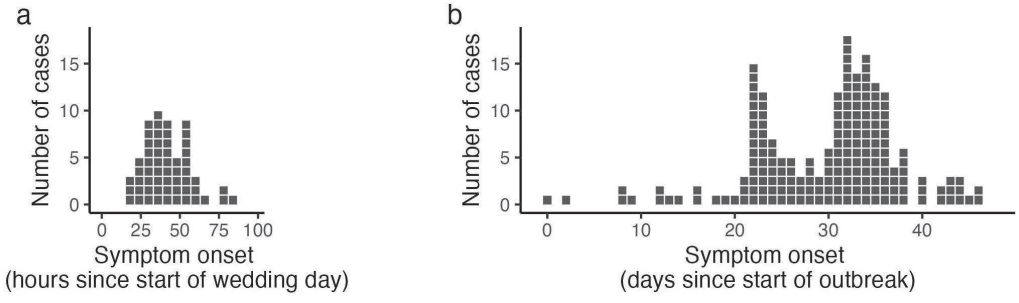
The time elapsed between the start of infectiousness and symptom onset has been

important. Namely, a long time lag between start of infectiousness and symptom onset indicates that isolation of symptomatics is most likely ineffective [Nishiura, 2007], i.e. when latency time is considerably shorter than incubation time.

While all individuals infected with SARS-CoV-2 experience a latency period, the same cannot be said for the incubation time. Among SARS-CoV-2 infected individuals, 40.5% did not develop symptoms until recovery [Ma et al., 2021]. Individuals with asymptomatic infection are typically more challenging to notify. The latency time is one of the factors determining the efforts required to control the spread of an infectious disease [Demers et al., 2023]. Despite its relevance, estimates of the latency time are rare and therefore, public health measures are typically informed by incubation time instead.

The incubation time differs considerably by type of infection, from few hours for toxic food poisoning to sometimes a few decades for tuberculosis, AIDS and variant Creutzfeldt-Jakob disease [Nishiura, 2007]. The incubation time for a specific infectious disease varies as well and may, amongst others, depend on age, transmission route, received pathogen dose, vaccination status and natural immunity [Held et al., 2019].

Figure 1.1: Epidemic curve related to two infectious disease outbreaks: (**a**) point source outbreak of salmonellosis among 57 cases visiting a wedding reception in Malahide, Ireland (1996) [ECDC, 2018]; (**b**) propagated (person-to-person transmission) outbreak of measles in Hagelloch, Germany (1861) in 187 cases [Groendyke et al., 2010]. Each square represents one case.



In some infectious disease outbreaks all infected individuals are thought to be infected at the same calendar time by a shared source and the infection is not transmitted further. These so called point source outbreaks provide the most straightforward scenario for

estimating the incubation time. Figure 1.1a is the *epidemic curve*, a specific type of histogram, corresponding to an outbreak of salmonellosis related to a wedding reception in Ireland [ECDC, 2018]. Each box represents an infected individual. The notified cases (y-axis) are organized by the six-hour time window in which each respective case exhibits the first symptom(s) (x-axis). Contaminated turkey served at the wedding was identified as the most likely vehicle of infection. Since the infection probably occurred at the wedding reception, the distribution of individual symptom onset days can be directly observed, as it approximately equals the incubation time of salmonellosis. The median incubation time of salmonellosis is known to be 45 hours [Eikmeier et al., 2018].

Many outbreaks propagate as direct transmission takes place between individuals or indirect transmissions mitigated by vectors like mosquitos. The epidemic curve curve of such a propagated outbreak is different from a point-source related curve. Figure 1.1b visualizes the daily number of individuals with measles [Groendyke et al., 2010], a highly infectious disease transmitted from person to person. The epidemic curve no longer resembles the incubation time distribution as individuals acquire the infection on different calendar dates. Often, the moment of infection is not precisely observed but it is, at best, known to occur within a specific time window. The information needed to estimate the incubation time distribution typically includes the exposure window along with the symptom onset day. Additional assumptions are needed to estimate the distribution of incubation time.

1.2 Statistical assumptions: a short history

The first estimate of the incubation time for influenza dates back to 1919 [McKendrick, 1925; Nishiura, 2007]. McKendrick, who gained recognition primarily for his infectious disease transmission models, estimated the incubation time on data from 92 maritime ships that left different harbours in Australia. He used the counts of individuals with symptom onset each day t since departure, i.e. $I(t) = 64, 17, 5, 2$ cases on the 1st, 2nd, 3rd and 4th day ($t = 1, 2, 3, 4$), respectively. Assuming that infection took place on shore and no transmission took place on board, he used the idea that individuals that developed symptoms on the second day since departure ($t = 2$) were exposed *at least* two days before when he estimated the daily probability Z_r of the incubation time r days after exposure as

$$Z_r = p(1-p)^{r-1}. \tag{1.1}$$

where $r \geq t$. Until today, the uncertainty of the infection moment that McKendrick acknowledged remains a major challenge in incubation time estimation. Nowadays, it is common to assume that infection is equally likely to occur on all exposure days (in the example: 1, 2, 3 or 4 days before departure). We will revisit the validity of the latter assumption later in the thesis.

The first attempt to model the incubation time distribution for infectious diseases using a continuous parametric distribution was attributed to John Miner in 1916 [Miner, 1922; Nishiura, 2007]. Miner suggested employing the right skewed Pearson I distribution while Philip E. Sartwell later proposed the lognormal distribution as an alternative [Sartwell, 1950; Nishiura, 2007]. The rationale for choosing a lognormal distribution was that pathogens were thought to grow exponentially within a host. While there is little evidence to support this reasoning for all infectious diseases, the lognormal distribution remains part of the commonly assumed triplet of right-tailed distributions today, alongside the gamma and Weibull distribution.

Coronaviruses are known to have a relatively long tailed incubation time distribution. The WHO has expressed concern about the validity of the commonly assumed parametric distributions, as they may not adequately capture the tail behaviour of the incubation time distribution of corona viruses [WHO, 2003]. The mismatch is particularly problematic since the percentiles are of particular interest; for instance, the 95^{th} percentile is typically used to choose the minimum duration of quarantine for potential cases.

In Chapter 2, we investigate the impact of using parametric distributions through a simulation study and assess the performance of a more flexible alternative. In Chapter 4, we assume another flexible distribution for the SARS-CoV-2 latency time distribution that includes the gamma and Weibull distributions as a special case.

Two common statistical concepts for observations of time-to-event complicate estimation of incubation and latency time. A time-to-event or survival time is the time interval between an initial event and the occurrence of an event of interest such as death, disease-progression, relapse, et cetera.

Often, the start- or endpoint of such an interval of interest cannot be observed precisely, which is referred to as *censoring* (Section 1.4). Moreover, in observational data it is

common that certain individuals are observed whereas others go unnoticed, which may lead to *truncation* (Section 1.5). Survival analysis is the statistical discipline devoted to studying time-to-event data, which can be incubation time, latency time, the age of breast cancer diagnosis et cetera.

Several concepts from survival analysis are relevant to the infectious disease context. Estimation of the time from infection to a certain event, such as initial multiplication of gametocytes, a stage of malaria parasites, in the human body, is complicated when some individuals recover from the infection before the endpoint occurred [Andolina et al., 2023; Ramjith et al., 2022]. Clearance of the infection is referred to as a competing risk. There are parallels between survival analysis models and those for spread of infectious disease as well, in specific with the stochastic SIR model [Putter et al., 2024] that models how individuals migrate through the susceptible, infectious and recovery stages. However, the data available early onwards in an infectious disease outbreak is typically fuzzy, stressing the need of tailored approaches for the infectious disease context in specific.

In the applications of this thesis, the estimates rely on observational data. This type of data contrasts clinical trial data in which individuals are assigned specific treatments at known time points and are monitored during follow-up. Before the statistical concepts relevant to our estimation problem are discussed in more detail, the spatiotemporal context and the corresponding data that inspired this thesis are introduced.

1.3 Contact tracing data: SARS-CoV-2 in Vietnam

Acknowledging its limited intensive care capacity and the long-stretched, 1297 km long border with China, the policy of Vietnam was characterised by stringent and early policy measures such as complete border closure. The country initially strived to prevent any introduction and local transmission of SARS-CoV-2. The main pillars of the elimination policy were extensive contact tracing of infected individuals and quarantining of potential infecteds. The quarantine policy for each potential case depended on the closeness to an infected individual, which is referred to as the 'F-system' and is unique to the pandemic response of Vietnam [Hardy et al., 2020]. For direct contacts of an infected individual, quarantine typically took place in designated quarantine facilities. Further details are provided in Chapter 4.

During contact tracing, notified cases were typically asked to recall their potential risk exposures and if so, when they first exhibited symptoms. In the designated quarantine facilities in Vietnam, swabs were taken regularly to test individuals for SARS-CoV-2. This context provides a unique data set that allowed to estimate the latency time for the SARS-CoV-2 Delta variant, which to the best of the author's knowledge is the first estimate based on data from outside of China.

1.4 What is observed and what is not: censoring

Whereas symptom onset is typically observed up to a day precise, the knowledge of the moment of infection and the start of infectiousness is generally limited to the time interval during which these start- and endpoints occurred. Typically, RNA shedding is used as a proxy for infectiousness. Common practice is to assume that the start of infectiousness occurs between the last negative and first positive test for SARS-CoV-2, such that instead of the exact start of infectiousness, a time window containing the endpoint of latency time is observed. Hence, an observation of incubation time consists of an exposure window and the symptom onset day, while an observation of latency time consists of an exposure window and a start-of-shedding window. Observations of incubation and latency time are *single interval censored* (time origin) and *doubly interval censored*, respectively.

Standard methodology is available when the endpoint is interval censored rather than the time origin. It is common to assume a constant risk of infection within the exposure window. As discussed in more detail in Chapter 2, this assumption is convenient because the likelihood can be rewritten with a reversed time axis and yields an interval censored endpoint. Therefore the incubation time can be estimated using available software. Unfortunately, the validity of this assumption is doubtful in the context of an evolving outbreak. At the beginning of an outbreak of a novel pathogen we are confronted with an exponential growth of new infections and this implies that the constant risk assumption is unrealistic. We performed simulation studies to investigate the impact of the constant risk assumption on the estimates of the percentiles of the SARS-CoV-2 incubation time distribution. Real data from the beginning of the pandemic are used as illustration (Chapter 2).

Another bias that may occur is related to the imperfection of our memory. Recalling

when risk exposure took place becomes more challenging when it occurred a long time ago. We refer to this phenomenon as 'differential recall'. Due to uncertainty in recall, exposure windows of less recent exposure may become relatively wide, increasing the risk of bias due to violation of the constant risk assumption. To limit the latter bias, the analysis is often restricted to observations with a narrow exposure window. Even though the term recall bias is frequently mentioned in papers estimating incubation time, to the best of the author's knowledge, differential recall of exposure, where recent exposures are memorised more precise than exposures longer ago, has never been explicitly studied in this context. In Chapter 3 we show that in the presence of differential recall selecting observations with narrow exposure windows leads to an additional bias.

While censoring concerns incomplete information and can be seen as a specific type of missing data, another statistical challenge in incubation and latency time estimation consists of observations that remain unobserved. In the following section we will elaborate on this.

1.5 Who is included and who is not: biases related to ascertainment

Random sampling is at the core of unbiased estimation. Every individual in the population of interest should have the same probability of being included in the sample. The latter can be challenging, especially when data is collected retrospectively. This thesis discusses three examples of observational data in which certain individuals from a study population were included with a higher probability than others. We briefly introduce the three concepts and refer to the respective chapters for further details.

The earliest estimates of the SARS-CoV-2 incubation time distribution were based on data from individuals who left Wuhan around the Lunar New Year and developed symptoms on or after their travel day [Backer et al., 2020]. Individuals who developed symptoms before travelling, i.e. those with short incubation times, were less likely to be included in the analyzed data. What in survival analysis is referred to as *late entry* or *left truncation* was not addressed in the estimates for SARS-CoV-2 in literature. We examined the impact on the estimates in a simulation study (Chapter 3).

Observations are *right truncated* when those with a relatively long time-to-event are less likely to be included in the data set. This phenomenon has been described for observations

of the SARS-CoV-2 latency time from China [Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021]. Right truncation occurred in the data from Vietnam, as individuals were only included in the data when tested positive for SARS-CoV-2 before the end of quarantine or the last day of sampling which in our data was marked by the start of a decline in case incidence due to reporting delay: contact tracing system became overwhelmed with large numbers of cases that could no longer be investigated as thoroughly as before. Right truncation is addressed in our analysis in Chapter 4, a consideration that is not immediately apparent for doubly interval censored observations.

Facing non-random samples is not unique to the infectious disease context and may also occur in other settings. For example, in breast cancer research, individuals are likely to attend a clinic for genetic risk when several family members developed breast cancer. When breast cancer is not frequent in the family, the presence of a genetic component is less likely to be observed. To examine the high risk due to specific genetic variants, researchers include data available from genetic clinics which typically concerns high-risk families with multiple affected individuals. This data is a non-random sample of the population, leading to biased estimates of the increased risk associated with a genetic variant. By means of a tailored weighting method, we restore the data composition such that the results can be extrapolated to the population of interest (Chapter 5). The robustness of our method is investigated by simulations and a two real data applications are provided.

1.6 This thesis

This thesis is a collection of four papers concerning different themes in survival analysis and (infectious disease) epidemiology. Table 1.2 presents an overview of the estimation problems we addressed in each chapter and the field of application. In Chapter 6, we place our work in a broader context by discussing future directions.

Table 1.2: Overview of the estimation problems discussed in this thesis.

Application	Cause	Effect	Remedy
Estimation of incubation and latency time	Assuming a constant risk of infection is not realistic	Overestimation in exponential growth phase (**Chapter 2**)	Assuming that the risk of infection within the exposure window increases congruently with the infection incidence during the exponential growth phase (**Chapter 4**)
	Assuming a gamma, lognormal and/or Weibull distribution for the time-to-event; subsequent choice based on AIC or LOO IC	Biased estimates of the tail percentiles (**Chapter 2**); potential misfit between true and chosen distribution (**Chapter 2**)	A flexible modelling choice, such as Penalized Gaussian Mixture (**Chapter 2**); fitting a generalized gamma distribution that includes gamma, lognormal and Weibull as special cases (**Chapter 4**)
	Differential recall	Underestimation when observations with narrow exposure windows are selected (**Chapter 3**)	Analyze all observations, including also wider exposure windows (**Chapter 4**)
	Delayed entry (left truncation)	Overestimation (**Chapter 3**)	Not straightforward as the late entry time is not observed exactly due to the interval censored infection time
	Right truncation	Underestimation (**Chapter 4**)	Addressed in the analysis; available as a ready-to-use `R` package (`doublIn`) (**Chapter 4**)
The genetic risk of breast, ovarian or prostate cancer	Family-based sampling	Ascertainment bias: underestimation of the risk associated with a genetic variant or polygenic risk score (PRS, **Chapter 5**)	A weighting approach that generalizes the state-of-the-art method; available as a ready-to-use `R` package (`wcox`) (**Chapter 5**)

This chapter is published as Vera H. Arntzen, Marta Fiocco, Nils Leitzinger and Ronald B. Geskus (2023). Towards robust and accurate estimates of the incubation time distribution, with focus on upper tail probabilities and SARS-CoV-2 infection. Statistics in Medicine, 42, 2341–2360 [Arntzen et al., 2023].

Biases in estimation of the incubation time distribution, with focus on upper tail probabilities

Contents

Abstract

Quarantine length for individuals who have been at risk for infection with SARS-CoV-2 has been based on estimates of the incubation time distribution. The time of infection is often not known exactly, yielding data with an interval censored time origin. We give a detailed account of the data structure, likelihood formulation and assumptions usually made in the literature: (i) the risk of infection is assumed constant on the exposure window and (ii) the incubation time follows a specific parametric distribution. The impact of these assumptions remains unclear, especially for the right tail of the distribution which informs quarantine policy. We quantified bias in percentiles by means of simulation studies that mimic reality as close as possible. If assumption (i) is not correct, then median and upper percentiles are affected similarly, whereas misspecification of the parametric approach (ii) mainly affects upper percentiles. The latter may yield considerable bias. We suggest a semiparametric method that provides more robust estimates without the need of a parametric choice. Additionally, we used a simulation study to evaluate a method that has been suggested if all infection times are left censored. It assumes that the width of the interval from infection to latest possible exposure follows a uniform distribution. This assumption gave biased results in the exponential phase of an outbreak. Our application to open source data suggests that focus should be on the level of information in the observations, as expressed by the width of exposure windows, rather than the number of observations.

Keywords SARS-CoV-2 □ incubation time □ interval censored data □ semiparametric □ quarantine period □ uniform infection risk

2.1 Introduction

Isolation of individuals with established SARS-CoV-2 infection and quarantining individuals with higher risk of infection (risk contacts) were two of the widely adopted policy measures to slow down the spread of the virus upon its emergence in 2020. Quarantine length for risk contacts is commonly based on the incubation time distribution, i.e. the time from infection to symptom onset [Cowling et al., 2007]. At the beginning of the SARS-CoV-2 pandemic, the WHO recommended a quarantine length of 14 days from the last time that a risk contact was exposed to an infected individual [WHO, 2020]. Country-specific quarantine lengths deviate from the WHO recommendation, depending on the policy aim, such as 'flattening the curve' in the Netherlands and - until July 2021 - 'zero spread' in Vietnam. A stricter policy requires a longer quarantine period.

Because of its relevance for policy makers, empirical estimates of the incubation time distribution were made soon after the start of the SARS-CoV-2 pandemic, mostly based on early data from Wuhan, China. Xin *et al.* performed a systematic review and meta-analysis of published studies until September 25, 2020 that reported mean, median and/or 95^{th} percentile of the incubation time along with 95% confidence intervals (CIs) [Xin, Wong, Murphy, Yeung, Ali, Wu and Cowling, 2021]. Individual estimates of the 95^{th} percentile ranged from 3.2 to 18.3 days. The pooled estimates of the 95^{th} percentile were dependent on the chosen parametric distribution: 12.6 days (7 studies; 95% CI, 11.2–14.0) and 14.1 days (5 studies; 95% CI, 12.3–15.8) for estimates based on lognormal distribution and Weibull distribution, respectively. These estimates concern the 'wild' type; in comparison, new variants seem to have a shorter incubation time [Wu et al., 2022].

While the event time (time of symptom onset) is generally observed, the big challenge is knowing the time origin (time of infection). Mostly, we only know the start and end of the exposure window, and often just the end. This makes the infection time interval censored or left censored. Most studies assumed a uniform distribution of infection risk within the exposure period. This has the advantage that the time scale can be reversed and traditional methods for interval and right censored time-to-event data can be used. In case all infection times are left censored, this approach cannot be used because all data would be right censored on the reversed time scale. For such data, another approach has been suggested which makes assumptions similar to those in renewal

process theory [Deng et al., 2020; Qin et al., 2020].

While the systematic review by Xin *et al.* focused on the estimates, the aim of our study is to review the methods used to estimate the SARS-CoV-2 incubation time distribution. We give a detailed account of the data structure, the likelihood formulation and the assumptions commonly made in the literature (Section 2.2). By means of simulations we studied the robustness of these assumptions (Section 2.3 and 2.4). The focus of this paper is on the impact of these assumptions on the estimates of the median and upper tail percentiles, as the quarantine length is based on these quantities. As an illustration, percentiles of the SARS-CoV-2 incubation time distribution are estimated using openly available data from the first months of the pandemic (Section 2.5).

2.2 Data collection and methods

2.2.1 Data on infection time

For most infectious diseases it is difficult to obtain information on time of infection. Estimates of the HIV incubation time distribution were mostly based on data from cohort studies, where participants were tested for the presence of HIV antibodies once every three to six months. Since antibodies can be detected within the first months after infection and the median incubation time to AIDS is about ten years, using the midpoint of the seroconversion window as infection time gives negligible bias [Law and Brookmeyer, 1992].

The situation with SARS-CoV-2 is very different. Time from infection to symptom onset varies from a few days to a few weeks. Many symptoms for Covid-19 are not very specific, may have another cause and many individuals remain asymptomatic. Since antibodies develop several weeks after infection, diagnosis of acute infection is based on the RT-PCR test for the presence of RNA. Information on infection time is obtained from four possible sources:

(i) time of direct contact with one or more infected individuals;

(ii) a time period during which an individual was at risk of infection, without having information on specific contacts;

(iii) time of a first positive RT-PCR test;

(iv) time of symptom onset.

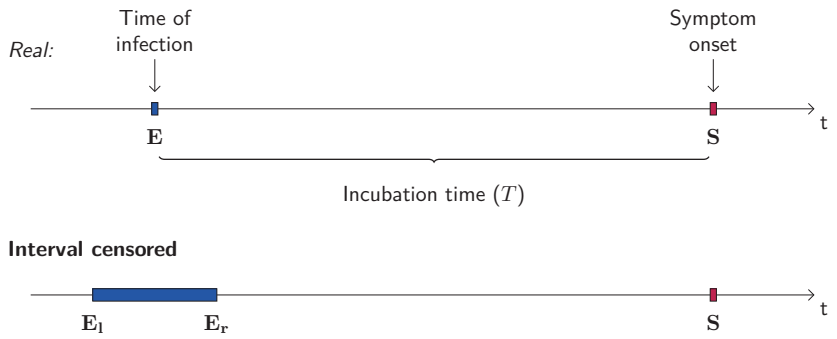
Intensive tracing of all contacts of diagnosed individuals can provide a rich source of information if the incidence of infection is low. However, data quality is hampered by recall bias of the time of contact, the presence of several possible infectors, and the true source of infection may even be missed. Also, in infection clusters where the possible contacts are known, it may be unknown who infected whom.

If no specific contacts are known, there may be general information on the window of exposure. Infected but still undiagnosed individuals can end a period of infection risk if they are quarantined and have no further contact with others. As an example, in the first 1.5 years of the pandemic, the Vietnamese government quarantined all direct contacts (F1) of an infected case (F0) in an allocated facility with active monitoring, and the second line (F2) of contacts at home. The earliest estimates [Backer et al., 2020; Lauer et al., 2020] of the SARS-CoV-2 incubation time were based on individuals who left Wuhan before they developed symptoms. Assuming that the virus was absent outside Wuhan at that time, departure from Wuhan ends the exposure window. These individuals had a minimum incubation time which is the time span between departure and disease onset. Most studies additionally included individuals who arrived in Wuhan during the first outbreak in January 2020. This defines their start of the infection risk period and gives a maximum incubation time. If for nobody the start of the exposure window is known, additional assumptions as discussed in Section 2.2.3 are required.

A positive test result only provides information on the infection time if the person tested positive before symptom onset. An earlier negative test does not provide any information because the person may already be infected at that time. Although symptom onset is the end point of the incubation time, it provides information on an individual's maximum incubation time if the start of his exposure window is known.

Data on infection and symptom onset have primarily been collected for public health purposes to contain further spread of the virus and monitor individuals with symptoms. They have not been collected in a rigorous scientific way. Many studies are based on data from government websites, which lack detailed information on the data collection process, on the choices made with respect to allocation of infection source and on the definition of symptom onset used.

Figure 2.1: Timeline for interval censored observations of incubation time (infection to symptom onset).



2.2.2 Likelihood and assumptions for interval censored infection times

We start by describing the approach used when some individuals have an interval censored exposure window. For individual i ($i = 1, \ldots, N$), let E_{il} and E_{ir} be the calendar times that denote the start and end of the exposure window respectively (Figure 2.1). E_{il} may be missing or set at a value before the start of the outbreak. Let S_i be the calendar time of symptom onset.

Denote by $g_i(\cdot|e_{il}, e_{ir})$ the density of the infection time, given the individual's exposure window. We allow g_i to depend on the individual. Let $f(\cdot)$ and $F(\cdot)$ be the density and cumulative distribution function of the incubation time and $h(\cdot,\cdot)$ the density of the exposure interval.

We assume incubation (e to s) and infection (e) time distributions to be independent. The contribution to the likelihood from individual i is

$$l(e_{il}, e_{ir}, s_i) = h(e_{il}, e_{ir}) \int_{e_{il}}^{e_{ir}} g_i(t|e_{il}, e_{ir}) f(s_i - t) dt. \tag{2.1}$$

Note that commonly E_{il}, E_{ir} and S_i are observed up to a day precise, but this discretisation is not taken into account in the likelihood.

The infection time distribution can be defined at the population level or at the individual level. For the earliest studies on the SARS-CoV-2 incubation time, no individual contact data was used, while the pandemic was in its exponential phase. This suggests using a single population wide distribution g, similar to what has been used for studies on the HIV incubation time where left, right and interval censored infection time are present [Geskus, 2001].

For SARS-CoV-2 there are several reasons to assume an individual exposure distribution $g_i(\cdot|(e_{il}, e_{ir}))$ instead. Infection rates can be very local and depend on the setting in which transmission occurred (at home, at work, in a public place). Also, contact rates fluctuate with an individual's willingness to comply with preventive and lockdown measures. And if contact tracing, precautionary quarantining and testing are related to suspected infection, then the time of infection is more likely to be close to the end of one's exposure window.

Most studies on the SARS-CoV-2 incubation time assume that the risk of infection is constant on the exposure window. Then, the contribution of $g_i(\cdot|e_{il}, e_{ir})$ to the likelihood reduces to a constant $(\frac{1}{e_{ir}-e_{il}})$ that can be left out, yielding

$$l(e_{il}, e_{ir}, s_i) \propto \int_{e_{il}}^{e_{ir}} f(s_i - t) dt = F(s_i - e_{il}) - F(s_i - e_{ir}). \tag{2.2}$$

and we end up maximizing

$$\sum_{i=1}^{N} log\left[F(s_i - e_{il}) - F(s_i - e_{ir})\right]. \tag{2.3}$$

Hence, by assuming a constant risk of infection on the exposure window, the time axis can be reversed and standard methodology for interval censored data can be applied. If

the infection time is left censored, it is possible to treat it as interval censored by choosing the start of the exposure window far before the start of the outbreak. When the time axis is reversed, this transforms into a right censored observation.

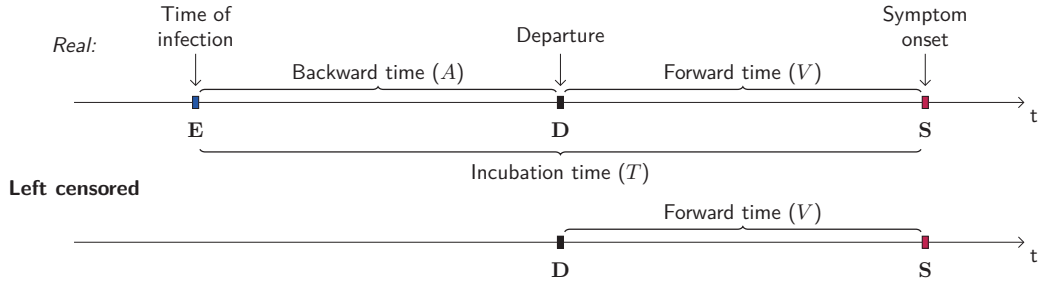
The validity of assuming a constant infection risk can be questioned, as the risk within the exposure window may vary by calendar time (if the outbreak is in the exponential growth phase), location and by type of contact (e.g. infection in a household or at work). We performed a simulation study to quantify the bias when the actual infection time is monotonically increasing or decreasing whereas a constant risk is assumed (Section 2.3).

Common practice is to describe the incubation time using parametric models such as lognormal, gamma and Weibull and choose the one that provides the most conservative tail percentile [Cowling et al., 2007; Nishiura, 2007] or the best-fitting distribution based on AIC [Held et al., 2019]. As there are few observations in the tail, the best fitting distribution and the parameter estimates will mainly be based on the larger amount of information in the middle of the distribution. As a consequence, the estimates of the tail percentiles will strongly depend on the assumed parametric distribution and the form of the incubation time distribution in the middle part. However, there is little biological evidence to support the assumption that the incubation time follows one single parametric distribution over the whole domain. A systematic review showed that the estimates of the 95^{th} percentile of the SARS-CoV-2 incubation time distribution vary according to the choice of the parametric model [Xin, Wong, Murphy, Yeung, Ali, Wu and Cowling, 2021]. Also, confidence intervals for the tail percentiles will be too narrow if assumptions are made that are uncertain to hold. We investigated the presence of bias under the assumption of an incorrect parametric distribution in a simulation study (Section 2.3).

2.2.3 Likelihood and assumptions for the approach of Qin, Deng *et al.*

An analysis by Qin *et al.* [Qin et al., 2020], later refined by Deng *et al.* [Deng et al., 2020], used data from 1211 individuals who developed symptoms after they had left Wuhan during the first exponential phase of the outbreak. The authors only included individuals travelling between January 19 and 23 2020, which is the time from the first public awareness of the severity of the outbreak until the city's lockdown right before Chinese New Year. To further ensure that all individuals were infected in Wuhan, the authors excluded cases who left Wuhan with their infected relatives and friends.

Figure 2.2: Timeline for observations of forward time (departure from Wuhan, China, to symptom onset) to estimate incubation time (infection to symptom onset). As departure in this context marks the end of the exposure window and the start of the exposure window is unknown, the time of infection is left censored.



For each subject $i \in \{1, \ldots, N\}$, date of travel D_i and symptom onset S_i were available (Figure 2.2). The incubation time T_i can be written as the sum of the forward $V_i = S_i - D_i$ and backward time $A_i = D_i - E_i$. Since only the forward time V_i is observed, additional assumptions need to be made to estimate the incubation time. The authors assumed that the time of leaving Wuhan can be seen as an observation time from a truncated renewal process that has reached the equilibrium state, where time of infection and symptom onset are renewal times. However, strictly speaking, this is not a renewal process, as there is no sequence of events of similar type. Rather, each individual has only two events of different type and their timelines overlap in calendar time.

More precisely, they assumed that travel is independent of infection and symptom onset and occurred randomly after infection according to $A_i \sim U(0, \tau)$ with τ set at 30 days. Left truncation arose because individuals that developed symptoms while still in Wuhan were excluded from analysis. Denote by f, F and μ the probability density function, cumulative distribution function and mean value of the incubation time distribution, respectively. Qin *et al.* derived that the density of the forward time $h(v)$ for the individuals in the data set is

$$h(v) = \frac{1 - F(v)}{\mu} \quad 0 \leq v \leq \tau. \tag{2.4}$$

This implies that the density of the included forward time is monotonically decreasing. Since the backward time A is assumed to follow a uniform distribution, it can be shown that the backward and forward time of the included data have the same distribution, denoted by $h(.)$.

The authors found that the forward times were not monotonically decreasing and therefore, they allowed for an additional risk of infection during travel, yielding the following

mixture distribution for the observed forward times

$$q(v, \pi) = \pi f(v) + (1 - \pi)h(v), \qquad v > 0 \tag{2.5}$$

where π is the probability to get infected at the departure time from Wuhan. Note that the forward time V_i equals the incubation time T_i if infection occurs on the day of travel.

The estimation method proposed by Deng *et al.* [Deng et al., 2020] improved Qin *et al.* 's method [Qin et al., 2020] because it takes into account that the symptom onset day is essentially an interval of 24 hours, due to the fact that the daily reports round information by day. Deng *et al.* add and subtract 0.5 to each forward time, i.e. $v_i^+ = v_i + 0.5$ and $v_i^- = v_i - 0.5$, yielding the following contribution of individual i to the likelihood:

$$l(v_i) = \pi\{F(v_i^+) - F(v_i^-)\} + (1 - \pi)\{H(v_i^+) - H(v_i^-)\}. \tag{2.6}$$

In the remainder of this paper, only the method of Deng is discussed. Results from the method of Qin *et al.* were very similar.

The validity of the assumption that travel occurs randomly between infection and day 30, i.e. $A \sim U(0, 30)$, is uncertain. Since the outbreak started in a fully susceptible population without any prevention measures, the incidence was likely to increase exponentially. Also, many people left Wuhan in the few days before the lockdown and Chinese New Year. To investigate the robustness of the method in this particular context, we performed a simulation study with exponential growth of infection incidence and varying rates of leaving Wuhan (Section 2.4).

2.2.4 Software

All analyses were performed in `R` version 4.1.1 [R Core Team, 2021] and `R Studio` version 2021.09.20 ("Ghost Orchid") [RStudio Team, 2021] software environment. `R` code is available from `github.com/vharntzen/simstudy_incubationtime`. This work was performed using the computing resources from the Academic Leiden Interdisciplinary Cluster Environment (ALICE) provided by Leiden University.

2.3 Simulation study I - Interval censored observations

2.3.1 Setup

Individual exposure window widths were sampled randomly from the observed exposure windows in five open source data sets [Backer et al., 2020; Lauer et al., 2020; Linton et al., 2020; Tindale et al., 2020] which we refer to as empirical widths.

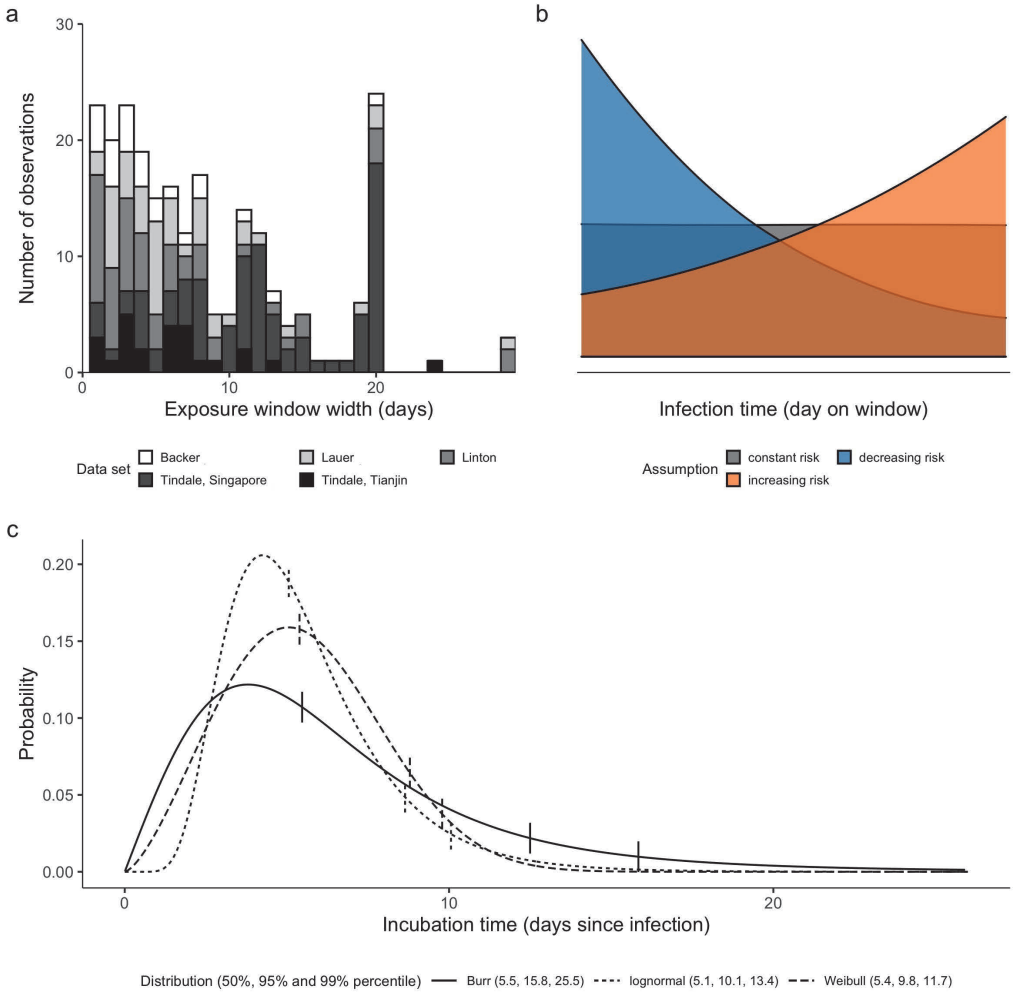
To examine the impact of the assumption of constant risk of infection in the exposure window, three different infection risk distributions were simulated on the individual exposure windows: i) constant risk ($g(t) \sim U(E_l, E_r)$), ii) exponential growth with five-day doubling time of the incidence ($g(t) \propto e^{0.14t}$) which reflects the initial phase of the outbreak in Wuhan [Dorigatti et al., 2020], and iii) a declining risk of transmission ($g(t) \propto p(1-p)^{t-1}$ where $p = 0.2$ on $[E_l, E_r]$), which may reflect household transmission. Figure 2.3b shows the risk functions for an exposure window of 10 days. The inverse cumulative distribution function (CDF) method was used. Moreover, to study how the impact of this assumption is affected by the exposure window width, more extreme scenarios were considered in which the widths were sampled after doubling or squaring the empirical widths.

To examine the impact of assuming an incorrect parametric distribution, incubation times were generated from a lognormal and Weibull distribution with parameters from Lauer *et al.* [2020]. We also generated incubation times from a more heavy-tailed Burr distribution, chosen such that the median was comparable to the two other distributions but with a considerably larger 95^{th} percentile (Figure 2.3c). Note that whereas the exposure window is discrete, no discretisation was applied to the time of infection and symptom onset.

Data generation

For each observation, an exposure window width ($E_r - E_l$), infection time E and incubation time T were sampled. Three different distributions of width were considered: the observed ('empirical') widths, all widths doubled and all widths squared (e.g. an observed width of 4 would become 8 and 16 respectively). Time of symptom onset S was set to $S = E + T$. If $S < E_r$, we set $E_r = S$: Covid-19 related symptom onset determines the end of the exposure window as infection certainly took place before. For each scenario, 1000 data sets with size $N = 100$ and 500 were generated. Details about the algorithm can be found in Supplement 2.7.1 (Algorithm 3).

Figure 2.3: Distributions used in simulation study I. **a** Distribution of exposure window widths in five openly available data sets from early in the pandemic. Shades of grey refer to the corresponding paper by author (and location) [Backer et al., 2020; Lauer et al., 2020; Linton et al., 2020; Tindale et al., 2020]. **b** Infection risk distribution in the exposure window in three different scenarios, indicated by colour. **c** Distributions of incubation time used for simulation study and their median, 95^{th} and 99^{th} percentiles (represented by vertical bars). The parameterizations of Weibull (shape = 2.453, scale = 6.258) and lognormal (meanlog = 1.621, sdlog = 0.418) were based on SARS-CoV-2 specific estimates by Lauer *et al.* [Lauer et al., 2020]; the choice of Burr distribution is such, that its median is comparable to the other distributions, but the tail is more heavy (m = 8.5, s = 2, f = 2).



Estimation

For each of the generated data sets, the 50^{th}, 90^{th}, 95^{th}, 97.5^{th}, and 99^{th} percentiles of the incubation time distribution were estimated using three parametric approaches, a semiparametric and a nonparametric approach. Maximum likelihood estimators (MLEs) assuming parametric distributions (gamma, lognormal and Weibull) were obtained by using the `flexsurv` package [Jackson, 2016] while for the nonparametric maximum likelihood estimator (NPMLE) the `survival` package was used. For the semiparametric approach, a penalized Gaussian mixture (PGM) was employed using the `smoothSurv` package [Komárek et al., 2005]. The smoothing factor λ was chosen based on the maximum AIC in a sequence of values (0.1 to 5.6, or 0.5 to 5.5 with step size 0.5 for data sets of size 100 or 500, respectively) . More details are given in Supplement 2.7.5.

For the percentiles in the parametric approaches, 95% CIs were obtained using parametric bootstrap as implemented in the `flexsurv` package. A method to obtain CIs for the NPMLE was explored (M out of N (M<N) bootstrap [Lee and Pun, 2006]). However, due to the size of the data set in this study, this technique was inadequate for the upper percentiles without smoothing, and CIs are not shown. Often, the estimate was equal to the upper limit of the CI. This is because the confidence interval disappears once the estimate of the cumulative distribution function reaches the value 1. More details can be found in Supplement 2.7.4. For PGM, 95% CIs are obtained by basic bootstrap based on 1000 replications. To limit the computation time, instead of finding the optimal λ for each bootstrapped data set, the λ as obtained for the estimator itself was used for each bootstrap replication. In addition, for each single run (estimate including its bootstrapped confidence interval) we specified an upper time limit of three hours.

To assess the performance of the five estimation approaches, we report the mean deviation of the estimate from the true value (as estimate of bias), as well as the 25^{th} and 75^{th} percentiles of the deviations over all runs. We also report mean width of the 95% CI and its coverage.

2.3.2 Results

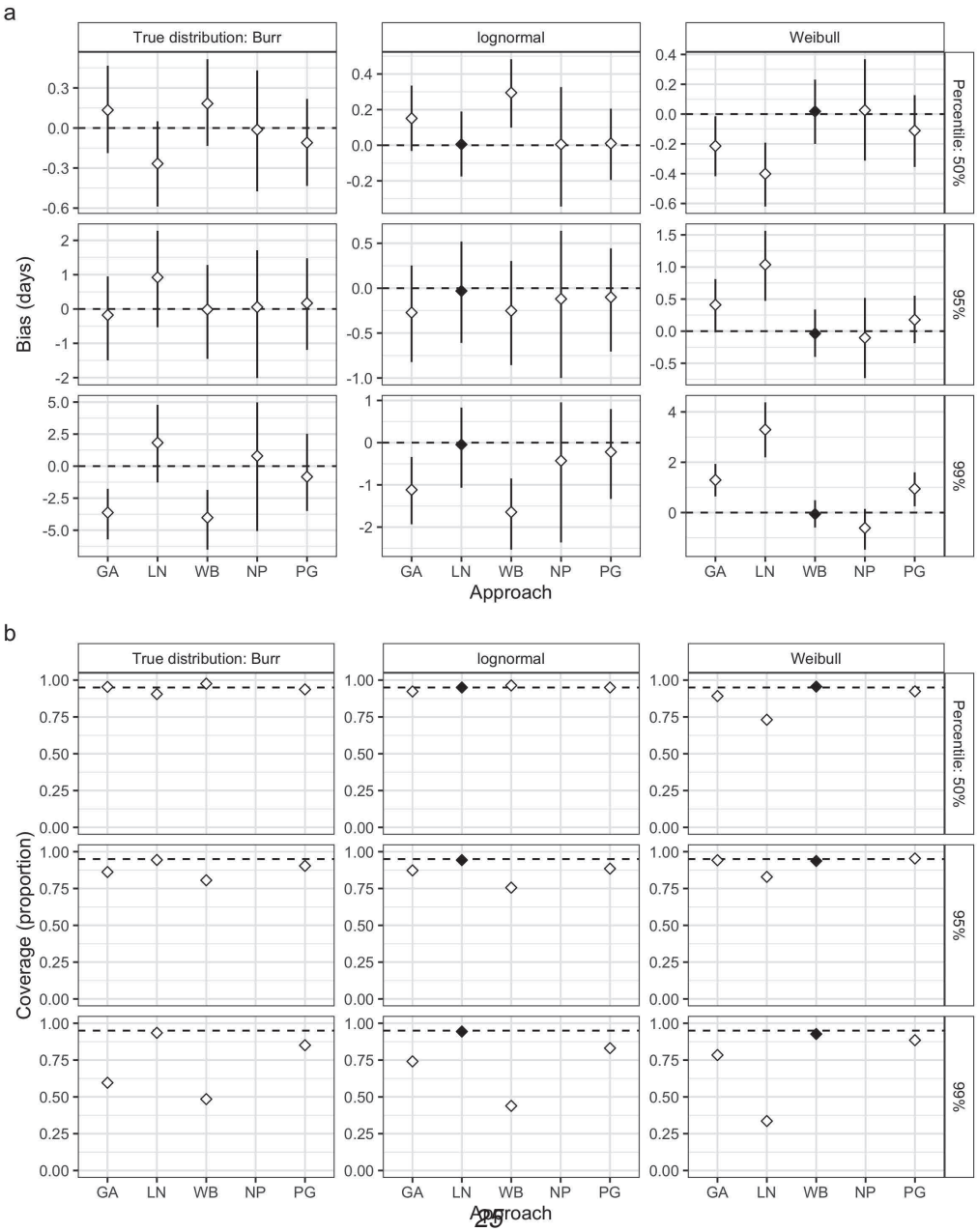
In this section results are shown for sample size N = 100 and three percentiles. Results for N = 500 and other percentiles are provided in Supplement 2.7.2

Incorrect parametric assumption introduces considerable bias in the tail estimates.

Figure 2.4 displays bias (a) and coverage probability (b) for the three chosen incubation time distributions and the five estimation approaches. The infection risk in the exposure window had a uniform distribution, as assumed in all five approaches. Estimates of the median and tail percentiles did not show any bias when the assumed distribution was the same as the true underlying distribution (see Figure 2.4a and Supplemental Figure 2.10a, middle and right panel, closed diamonds). Among all other combinations, NPMLE and PGM showed the smallest bias (cf. open diamonds). The variation (represented by vertical bars connecting quartiles of the estimates) in PGM was less than with the NPMLE, due to smoothing, and comparable to the parametric approaches. The incorrect parametric assumption led to a bias ranging from less than half a day for the median to more than three days in the 99^{th} percentile. The direction of the bias differed between the median and the tail percentiles. For example, for data generated from a Burr distribution (left panel), the bias was slightly upward for the median and downward for the tail percentiles when a gamma or Weibull distribution was assumed, and in the opposite directions when assuming a lognormal. As a consequence, there is a percentile between the median and the 99^{th} percentile where the estimate happens to be unbiased.

For most scenarios, the coverage deviated from 95% when the assumed distribution was different from the true one (Figure 2.4b and Supplemental Figure 2.10b, open diamonds). Coverage for the median dropped to lower levels when sample size increased (cf. Figure 2.4b and Supplemental Figure 2.10b, open diamonds). When the true distribution was Burr, different patterns were seen depending on the assumed distribution. Assuming a gamma or Weibull distribution yielded poor coverage when the percentile of interest was further towards the end of the tail. The lognormal distribution showed better coverage, especially in the tail for N = 100, where it was close to 95%. The latter may be related to the relatively wide confidence intervals (Supplemental Table 2.2). PGM showed fairly good coverage for all scenarios. For PGM, the coverage was based on a considerably smaller number of Monte Carlo replications but with at least 500 for each scenario (Supplemental Tables 2.1-2.6, rightmost column).

Figure 2.4: Results of simulation study investigating the impact of assuming an incorrect parametric distribution of incubation time: bias (**a**) of estimated percentiles and (**b**) coverage of 95% confidence intervals. Vertical bars represent the inter quartile range of the deviation between estimate and true value. Five different estimation methods were used (x-axis): maximum likelihood estimator (MLE) assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP) and penalized Gaussian mixture (PG), respectively. Incubation times were generated from Burr, lognormal and Weibull distribution and a constant infection risk on the exposure window was assumed. Data set size: N = 100.



If the true infection risk in the exposure interval strongly deviates from uniform, assuming a constant risk of infection on the exposure window is only appropriate when intervals are relatively narrow.

Figure 2.5 shows (a) bias in estimates of median and tail percentiles and (b) coverage probability when the infection risk distribution on the individual's exposure window is monotonically increasing (exponential growth) or decreasing (household transmission), respectively. The first resulted in consistent overestimation (see Figure 2.5a and Supplemental Figure 2.11a, middle panel), whereas the latter showed consistent underestimation (cf. Figure 2.5a, right panel), when the incubation time distribution was either chosen correctly (diamonds, Weibull) or modeled flexibly (circles, PGM). Note that for some parametric distributions, the two different biases discussed in this paper cancel each other out (Supplemental Tables B1-6). Violation of the uniform assumption led to similar bias in the median as compared to the tail percentiles, when the parametric assumption was correct (diamonds in Figure 2.5a, middle and right panel). This contrasts what was seen for the incorrect parametric assumption (cf. Figure 2.4a), namely that tail percentiles were more heavily affected than the median. Bias differed by exposure window width. Assuming a constant risk of infection in situations where this assumption is violated led to bias up to three days with a monotonically increasing infection risk on exposure windows of squared width (cf. Figure 2.5a, middle panel, right column).

Under a constant risk of infection (Figure 2.5a, left panel), even though the uniform assumption holds, the bias in the PGM estimate of the tail percentiles was around one day or larger. This is a consequence of the penalty that is imposed to guarantee smoothness of the distribution. For PGM, likewise parametric approaches, the tail behaviour is partly extrapolated from the part of the distribution where there are more observations, but for PGM this extrapolation is more local. Moreover, some of the bias may be due to the limitation discussed in Supplement E. However, in most scenarios this residual bias was smaller than with the incorrect parametric choice (Supplemental Tables B1-6).

For both approaches, coverage deviated from 95% when risk of infection was not constant on the exposure window (see Figure 2.5b and Supplemental Figure B2b). Under exponential growth (middle panel), by using the empirical exposure window size the coverage was good assuming Weibull and poor using PGM, due to differences in bias. In case of declining risk (right panel), the coverage for Weibull was poor and for PGM was

good. Coverage was considerably lower for the median than for the far right tail (cf. first vs. third rows) even though bias in the estimates using the correct distribution (represented by diamonds) is similar across percentiles. This is because the coverage proportion for the percentiles does not solely depend on the bias, but on the length of the confidence intervals as well. In fact, the CIs for the median were much smaller (Supplemental Tables B1-6).

2.4 Simulation study II - Renewal process

2.4.1 Setup

Scenarios

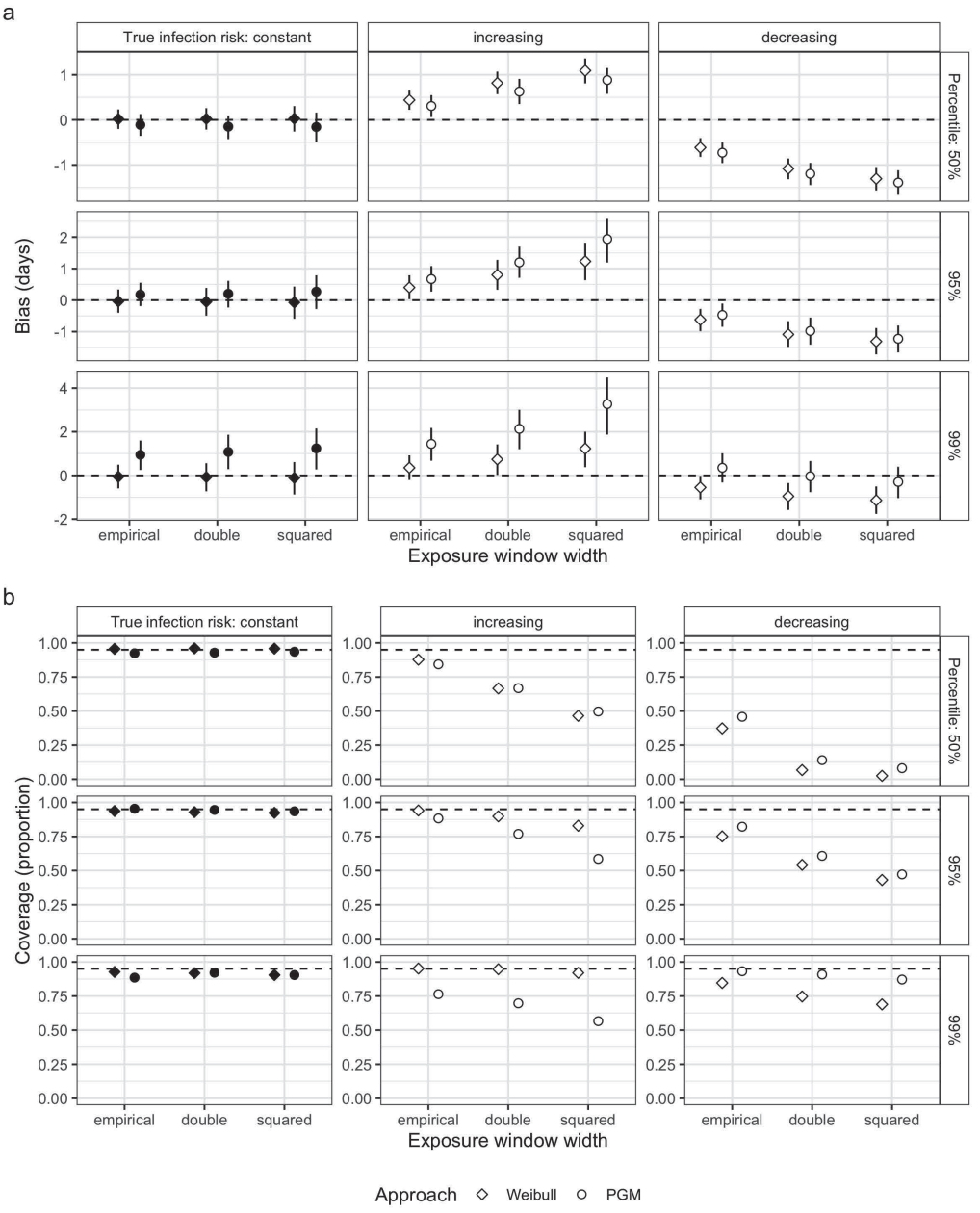
In this simulation study, we investigated the validity of the approach of Deng [Deng et al., 2020; Qin et al., 2020], and in particular the impact of the assumption that leaving Wuhan occurred randomly after infection. We simulated the early weeks of the outbreak in Wuhan, assuming that the incidence of SARS-CoV-2 infections grew exponentially and that the rate of individuals leaving Wuhan sharply increased as the lockdown approached [Gibbs et al., 2020]. We also repeated the data generation process as in the simulation study by Deng *et al.* [2020], in which no epidemic curve was assumed.

Data generation

For each scenario, we chose sample sizes of N = 500 and N= 1200 and generated 1000 data sets. Incubation times were drawn from lognormal and Weibull distributions with parameters as estimated by Lauer *et al.* [2020] and a more heavy-tailed Burr distribution, chosen such that the median was comparable to the two other distributions but with a considerably larger 95^{th} percentile (Figure 2.3c).

`R` code for generating data according to the method proposed by Deng is available on `github.com/naiiife/wuhan` (accessed 06/30/2020). Travel days were drawn from a uniform distribution on domain [0,30], with infection as time origin 0. Denote by π (with $\pi = 0, 0.1, 0.2$) the additional infection probability due to the travel. If an individual was infected before departure, then incubation time T_i and travel time c_i were drawn repeatedly until $T_i > c$, This implies that only individuals with symptom onset after travel were included. More details are provided in Algorithm 1.

Figure 2.5: Results of simulation study investigating the impact of assuming a constant risk of infection: (**a**) bias and (**b**) coverage proportion of percentiles (rows) estimated by MLE assuming Weibull and PGM model (shapes). Vertical bars represent the inter quartile range of the deviation between estimate and true value. Data was generated using different infection risk distributions (panels) and exposure window widths (x-axis). Incubation times were generated from the Weibull distribution. Data set size: N = 100.



Algorithm 1: Algorithm to generate observations of forward time V of SARS-CoV-2, as proposed by Deng.

Result: Data set with $i = 1, 2, ...N$ observations of forward time V, where N = 500 or 1200.

for $i \leftarrow 1$ ***to*** N **do**
  draw $D \sim Bernoulli(\pi)$;
  if $D = 1$ *(infected during travel)* **then**
    return $V_i \sim lognormal(..., ...)*$ **or** $Weibull(..., ...)*$ **or** $Burr(..., ..., ...)$;
  else
    repeat
      draw $T_i \sim lognormal(..., ...)*$ **or** $T_i \sim Weibull(..., ...)*$ **or** $T_i \sim Burr(..., ..., ...)$;
      draw $c_i \sim U(0, 30)$;
    until $T_i > c_i$;
    $V_i \leftarrow T_i - c_i$;
    return V_i;
end

Our alternative generation method (Algorithm 2) mimicked the infection and travel processes in the eighteen days between January 5^{th} and the lockdown of Wuhan on January 23^{rd}, 2020. Resembling the population of Wuhan, we assumed 10 million susceptibles as initial population. As in the real data from the study of Deng, only those who travelled between January 19^{th} and 23^{rd} and developed symptoms afterwards were included. We took January 5^{th}, 2020 as a starting date as those infected before were not likely to meet the criteria. Each day from January 5^{th} to 18^{th}, the same number of people (150,0000) entered and left Wuhan. From January 19^{th} to $23th$, no individuals entered Wuhan, but outbound travelling rate increased to 300,000 per day. The number of new infections on January 5^{th} was chosen to be 125. The daily incidence of SARS-CoV-2 increased according to a five-day doubling time [Dorigatti et al., 2020]. For the infecteds, incubation times were drawn and discretised using `R` function `round()`. This can be interpreted as all events (infection, symptom onset, travel) happening at noon. We selected individuals who left Wuhan between January 19^{th} and 23^{th}, and developed symptoms during or after their day of travel. The travelling rates and initial number of new infections were chosen such that this yielded approximately 1200 observations, comparable to the real data used by Deng (1211 observations). Additionally, from each data set a smaller data set was obtained by randomly sampling 500 observations. Note that the probability

of travelling was unrelated to infection status.

Algorithm 2: Algorithm to generate observations of forward time V, taking into account the exponential growth of SARS-CoV-2 incidence and the sharp increase in people leaving Wuhan, China, before the lockdown. The rate of people entering and leaving Wuhan and initial number of new infections were chosen such that this yielded approximately 1200 observations.

Result: Data set with $\approx$ 1200 observations of forward time V.
1. Initialize;
$P \leftarrow 1 : 10,000,000$ (population of Wuhan);
$S \leftarrow 1 : 10,000,000$ (susceptible population of Wuhan);
E, D, V infection day, travel day and forward time of infecteds;
$I_0 \leftarrow 125$ (number of newly infecteds on January 5^{th});
2. Infection and travel process;
for $t \leftarrow 1$ ***to*** 19 *(January 5^{th} to 23^{rd}, 2020)* **do**
  infect $I_0 e^{0.14(t-1)}$ *with indices* $K \subset S$;
  $E[K] \leftarrow t$;
  $S \leftarrow S[-K]$;
  if $t < 15$ *(before January 19^{th})* **then**
    add $150,000$ *(people entering Wuhan)*;
    $S \leftarrow S[+150,000]$, $P \leftarrow P[+150,000]$;
    remove $150,000$ *(people leaving Wuhan with indices* $M \subset P$*)*;
    $P \leftarrow P[-M]$;
    $S \leftarrow S[-(M \cap S)]$;
  else if $t \geq 15$ *(January 19^{th} to 23^{rd})* **then**
    remove $300,000$ *people leaving Wuhan, with indices* $M \subset P$;
    $P \leftarrow P[-M]$;
    $S \leftarrow S[-(M \cap S)]$;
    $R_t \leftarrow$ indices $M \cap (P \setminus S)$;
    $D[M] \leftarrow t$;
    for $\forall i$ *in* R_t **do**
      $T[i] \leftarrow lognormal(..., ...)$ **or** $Weibull(..., ...)$ **or** $Burr(..., ..., ...)$;
      if $(E[i] + T[i]) \geq D[i]$ **select** *(i.e. left Wuhan between Jan 19^{th} and 23^{rd} and symptom onset on or after travel)*;
      $V[i] \leftarrow E[i] + T[i]) - D[i]$;
    end
end
If N = 500 **then** $V \leftarrow$ sample 500 from V;
return (V)

Estimation

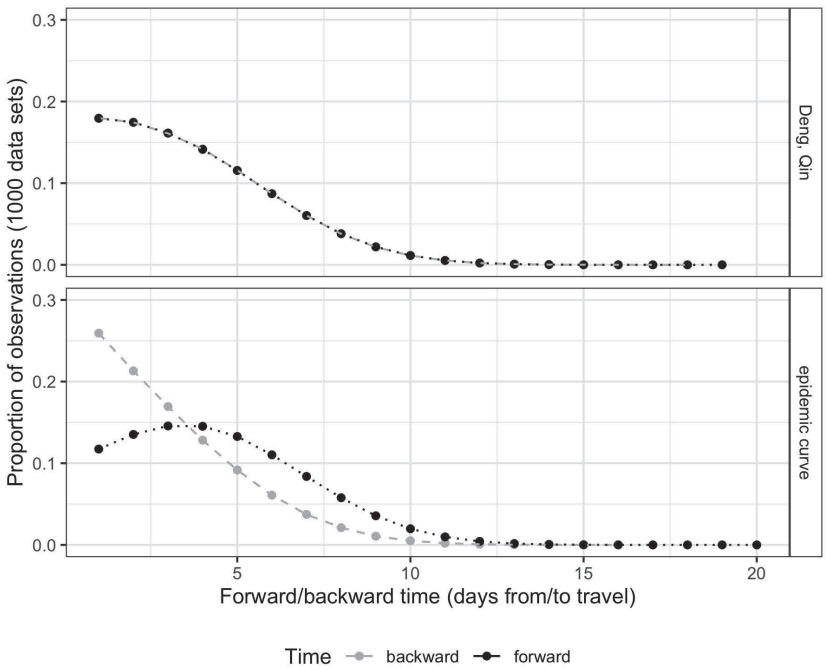
For each of the generated data sets, the 50^{th}, 90^{th}, 95^{th}, 97.5^{th}, and 99^{th} percentiles of the incubation time distribution were estimated using maximum likelihood estimation assuming gamma, lognormal and Weibull distributions.

For each percentile, we report the mean deviation of the estimate from the true value (as estimate of bias), as well as the 25^{th} and 75^{th} percentiles of the deviations over all runs (Figure 2.7). For the mixture approach, we report the average estimate of π and its 95% CI based on a normal approximation, conform Deng. Coverage probabilities for the percentiles are not reported as our main interest is in the bias, bootstrapping as proposed by Deng is computationally demanding, and the authors did not provide the coverage in their work either.

2.4.2 Results

In this section results from the estimation method of Deng and data set size N $\approx$ 1200 (range: 1064 to 1272) are discussed. Smaller sample size (N = 500) showed similar trends. See Supplement 2.7.3 for the additional results. With the assumptions as made by Deng and $\pi = 0$, the forward and backward time densities were equal and monotonically decreasing. This is no longer true when data were generated as in our new approach (Figure 2.6).

Figure 2.6: Summary of forward and backward times (black and grey) in 1000 combined data sets generated by two different approaches (panels) for the following settings: sample size approximately 1200, $\pi = 0$, incubation time follows a Weibull distribution. Note that in the upper panel the two curves overlay.



Little to no bias in estimates when travel day is sampled from uniform distribution and (mixture) model is correctly specified.

Results based on the data generation approach of Deng are shown in Figure 2.7. Figures 2.7a and 2.7c visualize the bias and mean of the estimates of π in the mixture approach. Figure 2.7e shows the bias in the approach without mixture component. Colours refer to the distribution chosen to generate the incubation times given on the x-axis. If for data generated with $\pi = 0.2$ the correct parametric distribution was chosen, median, tail percentiles and π show little to no bias (see Figures 2.7a and 2.7c, right panel, filled diamonds). Similarly, for the correctly chosen parametric distribution, the estimates didn't show any bias when data was generated with $\pi = 0$ and the model did not include a mixture component (cf. Figure 2.7e, left panel, filled diamonds). When data was generated with $\pi = 0$ and a mixture model was fitted little to no bias was observed either. However, when data was generated with $\pi = 0.2$ but analyzed without a mixture component, even the correctly chosen parametric model was strongly biased (Figure 2.7e, right panel).

When data generation incorporates epidemic and travel trends, estimates of median and tail percentiles were heavily biased, even when the correct parametric distribution was chosen.

When no mixture component was included in the likelihood, the bias in the percentiles was considerable (cf. Figure 2.7f). Per contra, when data was generated with $\pi = 0$ and a mixture model was fitted, this bias was reduced (cf. Figure 2.7b, filled diamonds). Our alternative data generation process makes infection close to travel more likely whereas a uniform distribution is assumed in the model. Hence, the model gives an upward bias in the incubation time distribution. Allowing for additional infections on the day of travel via the mixture approach can capture some of this model misspecification. It even gives a downward bias because the true infection date is mostly before the day of travel. It yields large estimates of π even though the data was generated without an excess risk while travelling (cf. Figure 2.7d).

When the choice of the parametric distribution is incorrect, bias in tail estimates can be as high as four days.

Figure 2.7a, b, e, f show that for almost all scenarios, bias in the percentiles was larger when the wrong parametric distribution was chosen (represented by open diamonds). In

particular when the incubation times were generated from a Burr distribution, estimates of median and tail percentiles were strongly biased and showed large variability. The parameter π was strongly overestimated in the mixture approach for incorrect parametric distributions (see Figure 2.7c-d).

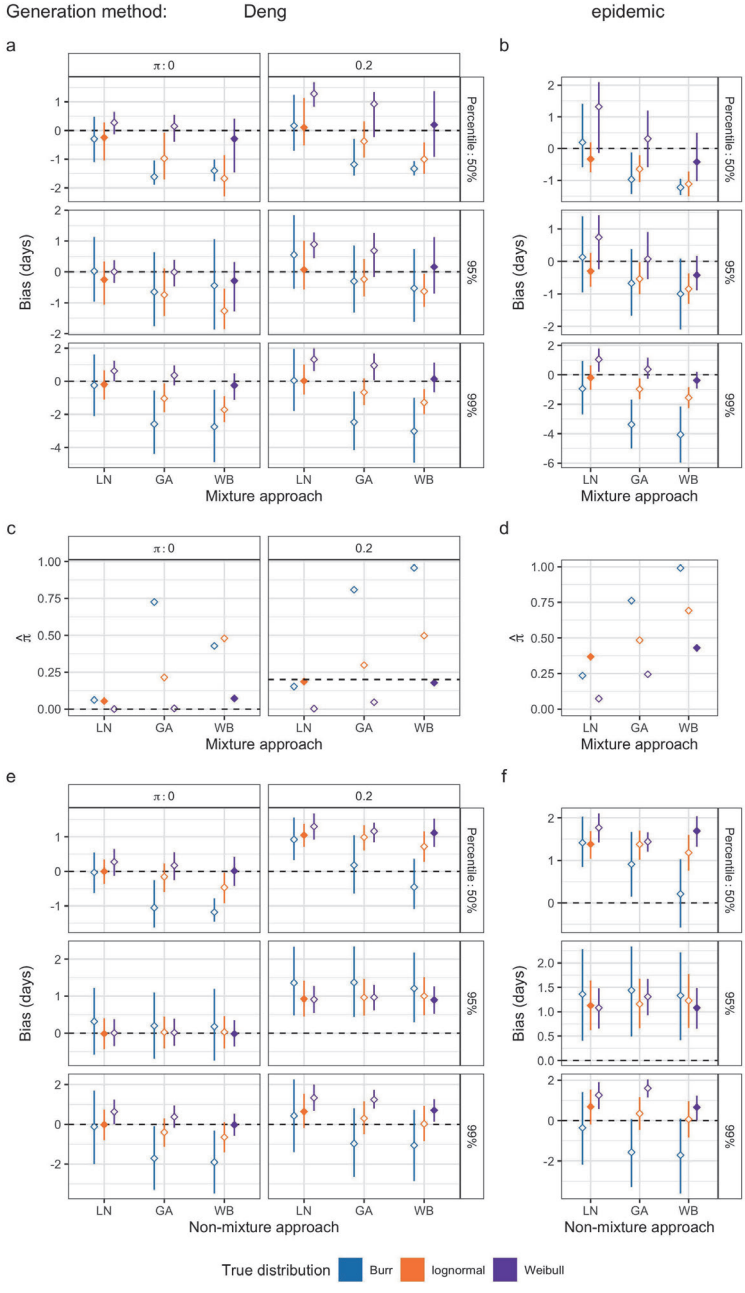


Figure 2.7: (Caption next page.)

Figure 2.7: Results from a simulation study investigating estimation of incubation time distribution using an method inspired by renewal process theory. Sample size is (approximately) 1200. Incubation times were generated from three different distributions: Burr (blue); lognormal (orange); Weibull (purple). Data generation: (**a, c, e**) Dengs method; (**b, d, f**) new method, i.e. epidemic outbreak with $\pi = 0$. Estimation approach: mixture including π (**a**, **b**, **c**, **d**) or excluding π (**e,f**). Fig. **a,b,e,f** and **c-d** show the bias in the estimate of the percentiles and the average estimate of π, respectively. Vertical bars represent the inter quartile range of the deviation between estimate and true value. Dashed lines represent either zero bias, or the π with which the data was generated.

2.5 Data illustration

2.5.1 Open source data

Six publicly available data sets [Backer et al., 2020; Lauer et al., 2020; Linton et al., 2020; Tindale et al., 2020; Yang et al., 2020] with observations collected between 2020/01/31 and 2020/02/29 were combined. Five data sets consisted of individuals infected in China; one data set concerned individuals with local transmission in Singapore as well. The sample size ranged from 52 to 178. Fifteen individuals with interval censored time of symptom onset from one data set [Lauer et al., 2020] were excluded from the analysis. Excluding another 70 asymptomatic individuals, 836 individuals were used. We divided the observations into five groups: exactly observed day of infection, interval censored day of infection with exposure window size smaller than or equal to the median width of four days, interval censored with wider exposure window, and left censored without information on the incubation time (end of exposure window is before, or coincides with symptom onset, respectively). Figure 2.8a visualizes all included observations and Figure 2.8b shows the frequency of each observation type per data set. Following common practice in these studies, missing information on the start of exposure window (left censored observations) was replaced by December 31^{st}, 2019. Note that this is actually not needed, as the observations can be analyzed as left censored as well.

There are two important characteristics of the data that are beyond the scope of this paper but worth mentioning. First, individuals may have been included in multiple data sets. Second, as observations were collected while spread was ongoing, right truncation may occur: individuals who got infected shortly before the end of the follow-up of the study are only included if they have a short incubation period. This phenomenon leads to underestimation of quantiles of the incubation time distribution, that is stronger during an exponential growth phase. Linton *et al.* [2020] and Xin *et al.* [2021] accounted for right truncation, incorporating exponential growth in the number of infections over calendar time in the likelihood.

2.5.2 Results

Figure 2.9 shows the estimates of the median and tail percentiles and their 95% CIs based on different partitions of the data, using the approach discussed in Section 2.2.2 (PGM

was run with an AIC-based choice of λ chosen from 0.5 to 5.5 with step size 0.5 and the same was done for the bootstrap runs to obtain 95% CI). For the NPMLE, confidence intervals were obtained as included in the function `survfit` from `R` package `survival`. We used the log-log transformation which made some of the upper bounds undefined. For more details, see Supplement 2.7.4. The confidence intervals using only the 32 exact observations were quite wide and became much narrower after adding the 216 observations with narrow exposure windows as the amount of information increased. Estimates of the percentiles became smaller. This is likely due to the change in relative contributions of the data sets, which may have different characteristics, resulting in differences in corresponding estimates. The data set from Tianjin, China, by Tindale *et al.* contains the majority of exact observations while the one by Yang *et al.* contains most of the (narrow) interval censored observations. These authors estimated median incubation periods of 8.06 (95% CI 3.35; 5.72) and 5.4 (95% CI 4.8; 6.0), respectively. Next, we added the 185 individuals with wider exposure windows and 238 individuals for whom the start of exposure was unknown but the end of exposure was before symptom onset. This changed the estimates slightly in the upward direction, which may be explained by recall bias: more recent exposure (and infection) tends to be memorised better than exposure longer ago. Hence, individuals with a short incubation time are more likely to have a narrow exposure window than those with a longer incubation time. Therefore, estimates based on narrow exposure windows may be biased in the downward direction. The CI width didn't change much. Lastly, adding the 165 with only the end of exposure known changed the estimates and the CI widths very little. Note that they would not have changed at all if the infection times of individuals with unknown start of exposure window had been treated as left censored. This behaviour was seen in all estimated percentiles. For all methods, the width of the confidence intervals increased with the shift towards the (further) tail percentiles. Weibull and gamma approaches estimated the median higher than lognormal. For the tail percentiles this was reversed. Note that the estimate using the semiparametric PGM method was always in-between the two most extreme parametric estimates. This corresponds with our findings from simulation study I (Section 2.3).

Figure 2.8: Open source data to estimate the incubation time of SARS-CoV-2: (**a**) Visualization of individual timelines. Time is given as time from last possible exposure in days. Grey bars and dots indicate the exposure window and its midpoint (if any). Black dots indicate symptom onset. Cases are ordered by exposure window width and time between end of exposure and symptom onset. (**b**) Type of observations per data set. Data sets are indicated by first author. Bar height represents the number of observations in each data set.

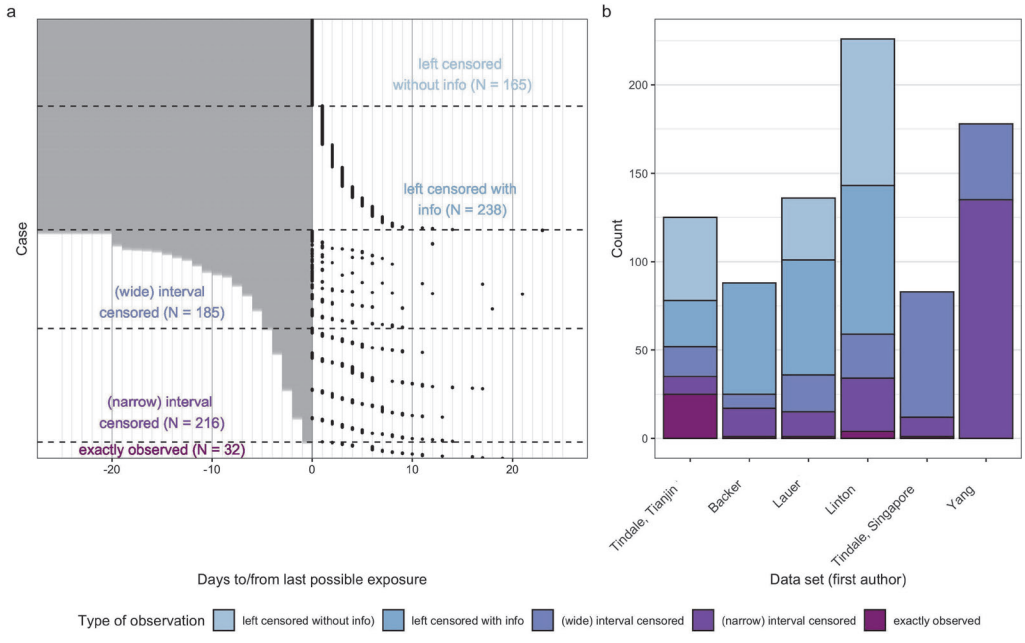
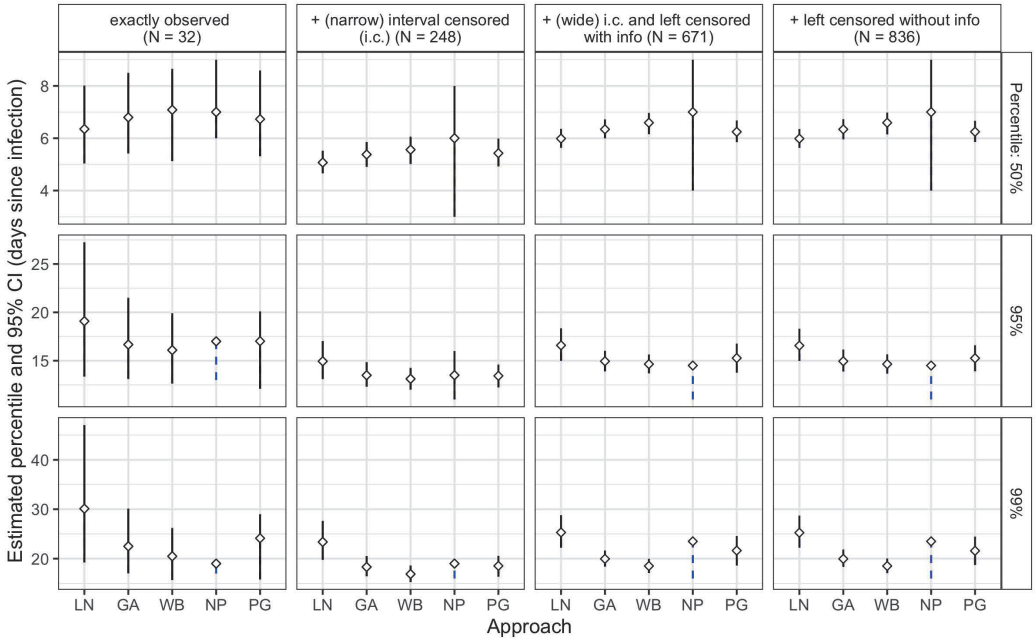


Figure 2.9: Estimates of percentiles of the incubation time distribution based on different partitions of the data (panels) and estimation approaches (x-axis). Five different estimation methods were used: maximum likelihood estimator (MLE) assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP) and penalized Gaussian mixture (PG), respectively. Point estimates along with 95% CI are provided. When the upper bound of the CI was undefined, the lower bound is connected to the estimate instead (blue dashed line, NP).



2.6 Discussion

Estimates of the incubation time distribution have been and will be essential to inform policy makers at the start of an outbreak of a new pathogen like SARS-CoV-1 or SARS-CoV-2. The most challenging part is to obtain accurate data on the infection time. Most infection times are left or interval censored. We discussed and evaluated methods to estimate the incubation time of SARS-CoV-2. Our focus was on the commonly made assumptions and the resulting bias in estimates of the median and upper tail percentiles.

Most estimates are based on data sets in which not all infection times are left censored but at least some individuals have an interval censored infection time. To simplify estimation, standard practice has been to assume (i) a parametric distribution and (ii) a constant risk of infection within the exposure window. We examined the impact of both assumptions (i and ii) in a simulation study. Different parametric and nonparametric approaches (MLE assuming gamma, lognormal, Weibull; NPMLE) were considered. In addition, we proposed a semiparametric approach, that avoids the arbitrary choice of a parametric family yet preserves the smoothness of a parametric curve. We investigated the bias if the true infection risk in the exposure window is exponentially increasing (e.g. an evolving outbreak) or declining (e.g. household transmission). While an incorrect parametric choice mainly affected the upper percentiles, incorrectly assuming a constant risk affected the median and upper percentiles equally. We discuss the impact of each assumption in more detail.

Parameters are estimated based on all observations. The majority of observations is located in the middle. Accordingly, tail behaviour is forced to follow the behaviour in the middle of the distribution. For this reason, estimates of the tail percentiles were not robust to an incorrect parametric assumption of the incubation time distribution. In contrast to the pooled estimate of the 95^{th} percentile based on earlier estimates (13.1 days [Xin, Wong, Murphy, Yeung, Ali, Wu and Cowling, 2021]), a recent study by Zhang *et al.* [2021] reported that more than 10% of individuals have an incubation time of more than 14 days. With our heavy-tailed Burr distribution, assuming a gamma or Weibull distribution estimated the 99^{th} percentile almost 4 days too small. The semiparametric approach proposed here - penalized Gaussian mixture- provides a good alternative. In smaller data sets it outperforms NPMLE, that often has the last jump to the value 1 before the upper percentiles of the true distribution. See Supplement 2.7.4 for details. However, the smoothing parameter λ needs

to be chosen carefully and the default procedure in the `smoothSurv` package did not always give satisfactory results (Supplement 2.7.5). Moreover, for an incorrect parametric choice the confidence intervals tend to be too small for estimates of the tail percentiles. The confidence interval length for PGM and MLE assuming lognormal were fairly similar if the true distribution was Burr or lognormal (both heavy-tailed in our parameterisation). When the true distribution was Weibull, confidence interval length of PGM was most similar to the length obtained by the correct parametric choice.

The bias in the tail percentiles, introduced by falsely assuming a constant risk, tended to be smaller than the bias that can be attributed to the incorrect parametric choice. We saw that the bias increased with increasing average widths. If there is no recall bias, restricting to narrow exposure windows in the data gives the smallest bias, but it throws away information.

For many infectious diseases, like SARS-CoV-1 and Ebola, start of infectiousness coincides with or occurs after symptom onset. However, for SARS-CoV-2, 47.3% (95% CI: 34.0 to 61.0) of individuals remained asymptomatic throughout the course of infection [Sah et al., 2021], while presymptomatic and asymptomatic transmission can occur [Chau et al., 2020; Tindale et al., 2020]. Ideally, the distribution of time from infection to having detectable infection rather than incubation time should inform quarantine length for potentially infected individuals. For many infectious diseases, this will be almost similar to the time from infection to start of infectiousness (latency time). The standard procedure to detect SARS-CoV-2 infection is to perform a PCR-test, giving rise to interval censored event times. Estimation requires both a last negative and first positive PCR-test for at least part of the individuals. As both the start- and endpoint are interval censored (doubly interval censored data), estimation of these distributions is more complicated [Reich et al., 2009]. As a consequence, such estimates are rare and only became available later in the pandemic [Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021].

Another way to avoid handling interval censored observations as such is restricting to exact observations or imputing or backtracing the infection moment. From literature it is known that restricting to exactly observed infection day leads to underestimation, due to recall bias: as individuals tend to recall more recent exposure more easily, observations of short incubation periods are more likely to be exact observations than those of long incubation periods. Midpoint imputation of the infection moment on the exposure window leads to bias as well [Cowling et al., 2007]. Besides, it is only applicable when the start

of the exposure window is known. A recent study [Ejima et al., 2021] utilized viral load measurements to backtrace the actual infection moment but such data is usually lacking. Moreover, estimates may be biased if the strong assumptions that are made are incorrect.

Two earlier studies investigated the validity of assuming a parametric distribution and a constant risk on the exposure window [Reich et al., 2009; Cowling et al., 2007]. Cowling *et al.* noted discrepancies in the tails for different parametric models and named the nonparametric estimate as the gold standard [Cowling et al., 2007]. In line with our results, Reich *et al.* observed that for incorrect parametric choice, coverage for the median was lower as sample size increased [Reich et al., 2009]. The impact of assuming a constant risk on the exposure window was explored for a limited number of deviating risk distributions (piecewise uniform and spiked distribution) and two most extreme scenarios (all infected at beginning of the exposure window, all at the end) [Reich et al., 2009; Cowling et al., 2007]. The authors noted that the tail estimates were more sensitive to the choice of parametric distribution (i) than to the uniform assumption (ii), although performance was poor for the spiked distribution. This is similar to our findings. We contribute to their work by quantifying the bias for several scenarios inspired by SARS-CoV-2.

When data consist of only left censored observations of infection time, the above method based on the likelihood of interval censored data cannot be used. Alternative methods are needed, based on additional assumptions. Our second simulation study shows that the method proposed by Qin is not valid for the setting of an emerging outbreak [Qin et al., 2020]. Nevertheless, their method may be useful with data from individuals arriving in countries with a strict quarantine policy, if infection rates in the country of departure and arrival rates are relatively stable. Examples of such countries are Vietnam, China, New Zealand, Australia and Taiwan during part of the pandemic.

Qin *et al.* performed a sensitivity analysis where they assumed an exponential density for the time from infection to departure. This led to an extra parameter in the likelihood of the forward times, which was estimated to be almost equal to zero. They concluded that the likelihood of forward times is approximately valid, even if the assumption of a uniform distribution of time from infection to departure does not hold. This is uncertain for two reasons. First, it was based on one real data set only. Second, it is difficult to assess how an exponential distribution for time from infection to departure relates to an exponential increase in the number of infections over calendar time which is much

closer to what happened in reality. The authors compared their approach to the approach for interval censored data, but this comparison doesn't make sense. Their generation method for interval censored data leads to a different data set and it additionally has a conceptual mistake (see Supplemental Algorithm 4).

So far we discussed the marginal incubation time distribution, neglecting its dependence on covariables like age and comorbidities which may explain part of the differences between studies. Regression of incubation time on such covariables is needed to come up with a more personalized quarantine length, for example depending on age [Pak, Langohr, Ning, Martínez, Melis and Shen, 2020]. It is important to stress that also other factors are involved in choice of quarantine length. One example is the expectation of how people will adhere to it, which may be improved if quarantine period is shorter. Moreover, in countries that implement quarantining in allocated facilities, the capacity and economic costs may play a role. When policy makers aim for zero SARS-CoV-2 infections, more extreme percentiles than included in this paper might be needed to determine the length of the quarantine period. To accurately estimate percentiles in the far right tail, approaches based on extreme value theory may be more appropriate.

Quarantine length is based on the right tail of the incubation time distribution. Due to the interval censored time of infection, parametric assumptions are commonly made. We show that this can introduce only mild up to rather severe bias, mainly in the tail percentiles. Especially to inform quarantine length, a semiparametric method is a better option and can be used with available `R` software. Whether the bias of the parametric methods is of clinical relevance depends on the aim of the policy. This cannot be seen separate from its societal context (ethics and resources) and disease characteristics (risk of severe outcome and transmissibility). Quarantine is a powerful measure for disease control, but cumbersome, and requires accurate and congruent estimates in the literature to optimally inform decision makers. The penalized Gaussian mixture approach avoids the arbitrary choice between parametric distributions and reduces bias in tail percentiles.

The amount of smoothness can be chosen automatically using the standard procedure in the function `smoothSurvReg` [Komárek, 2020]. If the resulting density shows a too wiggly pattern, the level of smoothness may be increased by increasing the value of the penalty term. We recommend to report the number of observations stratified by exposure window width, which gives additional information on the uncertainty in the

estimate on top of the sample size.

Acknowledgements

The authors thank the Covid-19 modeling team at the Oxford University Clinical Research Unit, Vietnam: Duc Du Hong, Trang Duong Thuy, Lam Phung Khanh, Leigh Jones, Lieu Tran Thi Bich, Maia Rabaa, Manh Nguyen Duc, Marc Choisy, Nguyet Nguyen Thi Minh, Nhat Le Thanh Hoang, Sonia Lewycka, Thomas Kesteman, Trinh Dong Huu Khanh, Tung Trinh Son.

2.7 Supplementary material

2.7.1 Generation algorithms

Algorithm 3 describes the data generation process in Section 2.3.1. The output is a data set with observations of incubation time, where the origin is interval censored.

Algorithm 3: Algorithm to simulate data with interval censored time origin.

Result: Data set with N observations of incubation time.
for $i \leftarrow 1$ ***to*** *N* **do**
  sample $_{,i}$ from set of exposure window widths;
  draw $_i$ on $(0,_{,i})$ according to constant risk ($_i \sim U(_{,i},_{,i})$) **or** exponential growth ($_i \propto e^{0.14t}$) **or** household transmission ($_i \propto p(1-p)^{t-1}$ where $p = 0.2$);
  draw $_i \sim lognormal(..., ...)$ **or** $\sim Weibull(..., ...)$ **or** $\sim Burr(..., ..., ...)$;
  $_i \leftarrow_i +_i$;
  if $_i \leq_{,i}$ **then**
    $_{,i} \leftarrow_i$;
  end
end

Algorithm 4 describes an alternative generation method by Deng *et al.* [2020], that aimed to provide observations with interval censored time origin. However, we noticed an error in this procedure, and as such the resulting observations are not truly interval censored.

The width of the exposure window C was drawn from a uniform distribution on $(0, 30)$ and added to the forward time V, yielding observations (V, $V+C$), see highlighted line. However, this does not guarantee that the actual incubation time is contained within this interval. For example, let us assume $D = 0$, $c = 8$, $C = 5$ and $T = 10$. Then, the true incubation time (T) is not contained in the interval $(V, V + C) = (2, 7)$. Hence, this approach is not valid.

Algorithm 4: Algorithm to generate interval censored observations using forward time V of SARS-CoV-2, as proposed by Deng. There is a conceptual incorrectness in the code, see second to last line.

Result: Data set with N observations of forward time V, where N = 500 or 1200.
for $i \leftarrow 1$ ***to*** *N* **do**
  draw $D \sim Bin(\pi)$;
  if $D = 1$ *(infected during travel)* **then**
    draw $V_i \sim lognormal(..., ...)$ **or** $\sim Weibull(..., ...)$ **or** $\sim Burr(..., ..., ...)$;
    draw $C_i \sim U(0, 30)$;
  else if $D = 0$ *(infected before travel)* **then**
    draw $T_i \sim lognormal(..., ...)$ **or** $\sim Weibull(..., ...)$ **or** $\sim Burr(..., ..., ...)$;
    draw $c_i \sim U(0, 30)$;
    while $T_i < c$ **do**
      draw $T_i \sim lognormal(..., ...) * *$ **or** $\sim Weibull(..., ...)$ **or** $\sim Burr(..., ..., ...)$;
      draw $c_i \sim U(0, 30)$;
    end
    $V_i \leftarrow T_i - c_i$;
    draw $C_i \sim U(0, 30)$ (incorrect);
  return V_i, $V_i + C_i$
end

2.7.2 Tables and figures Simulation study I

Due to feasibility constraints, a time limit was imposed on each of the 1000 Monte Carlo replications per scenario (3 hours for PGM; 2.78 hours for other approaches).

Figure 2.10: Results of simulation study investigating the impact of assuming incorrect parametric distribution of incubation time: bias (**a**) of estimated percentiles and (**b**) coverage of 95% confidence intervals. Five different estimation methods were used (x-axis): maximum likelihood estimator (MLE) assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP) and penalized Gaussian mixture (PG), respectively. Incubation times were generated from Burr, lognormal and Weibull distribution and a constant infection risk on the exposure window was assumed. Data set size: N = 500.

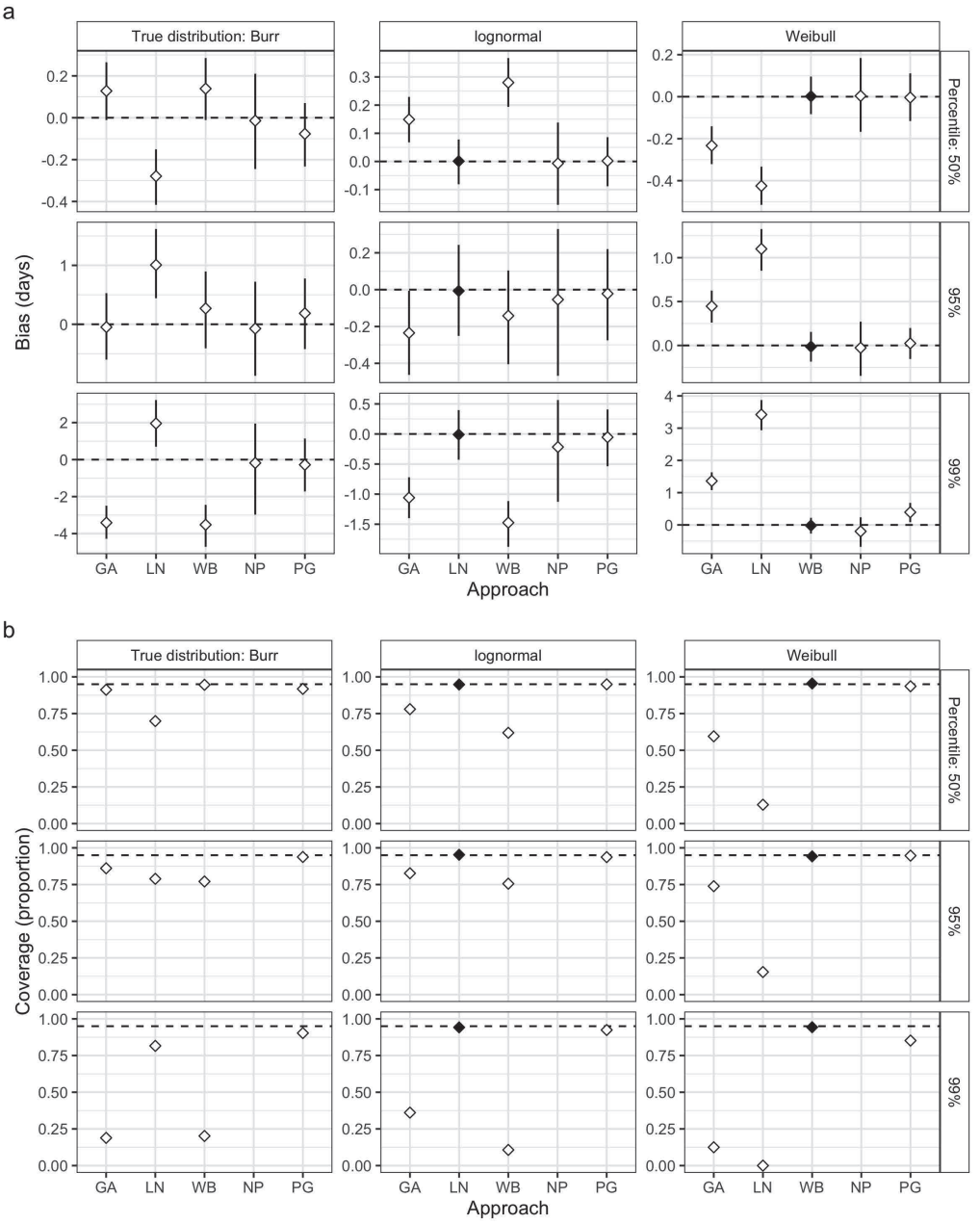


Figure 2.11: Results of simulation study investigating the impact of assuming a constant risk of infection: (**a**) bias and (**b**) coverage proportion of percentiles (rows) estimated by MLE assuming Weibull and PGM model (shapes). Data was generated using different infection risk distributions (panels) and exposure window widths (x-axis). Incubation times were generated from the Weibull distribution. Data set size: N = 500.

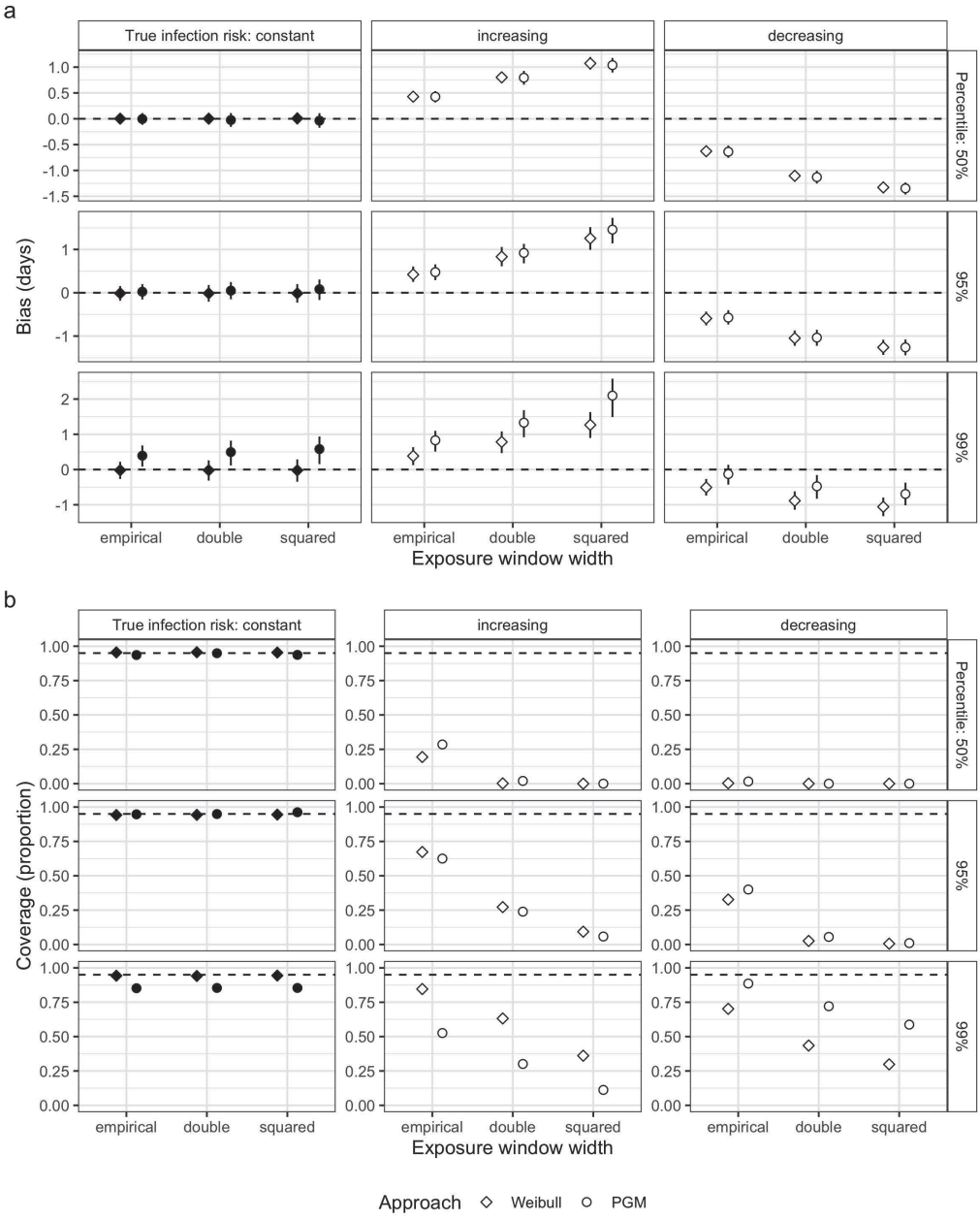


Table 2.1: Results for data generated from Burr distribution, data set size N = 100, for five percentiles. Each row summarizes 1000 runs. *Abbreviations* True infection risk (Inf. risk): constant, increasing (exponential growth) or decreasing (household transmission). Exposure window widths (Width): empirical set (emp.) or its doubled (db.) or squared (sq.) version. Estimation method (Mod.): MLE assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP), penalized Gaussian mixture (PG). Summary measures: bias with (normal approximated) 95% CI, the 25^{th} and 75^{th} percentiles of the deviations over all runs (p25 and p75), Mean Squared Error (MSE), mean length of the CI (LCI), coverage proportion. Percentage of (1000) runs that timed out (TO). ≈ 0 means a value between -0.005 and 0.005.

Mod.	Inf. risk	Width	50%							90%						95%						97.5%						99%						TO
			Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	%
GA			0.13	(0.10; 0.16)	-0.19	0.47	0.23	1.86	.95	0.36	(0.28; 0.44)	-0.60	1.21	1.96	4.23	.88	-0.18	(-0.29; -0.06)	-1.50	0.95	3.36	5.56	.86	-1.24	(-1.38; -1.09)	-2.92	0.21	6.88	6.97	.79	-3.63	(-3.81; -3.44)	-5.72	-1.77	21.97	8.89	.60	0
LN	con-		-0.27	(-0.30; -0.24)	-0.59	0.05	0.27	1.68	.90	0.43	(0.35; 0.51)	-0.48	1.27	1.89	5.38	.95	0.92	(0.80; 1.05)	-0.53	2.28	4.83	8.12	.94	1.39	(1.21; 1.57)	-0.69	3.37	10.21	11.59	.94	1.83	(1.56; 2.10)	-1.27	4.78	22.36	17.44	.94	0
WB	stant		0.18	(0.15; 0.21)	-0.13	0.52	0.26	2.10	.98	0.64	(0.55; 0.73)	-0.38	1.53	2.53	4.14	.84	-0.01	(-0.14; 0.11)	-1.45	1.28	4.17	5.35	.81	-1.27	(-1.43; -1.10)	-3.15	0.36	8.66	6.68	.70	-4.02	(-4.24; -3.80)	-6.52	-1.86	28.34	8.55	.48	0
NP			-0.01	(-0.06; 0.03)	-0.48	0.43	0.50	-	-	-0.02	(-0.13; 0.09)	-1.31	1.13	3.20	-	-	0.06	(-0.12; 0.23)	-2.01	1.71	7.66	-	-	-0.33	(-0.60; -0.07)	-3.31	1.96	18.46	-	-	0.79	(0.29; 1.30)	-5.07	4.98	66.74	-	-	0.1
PG			-0.11	(-0.14; -0.08)	-0.43	0.22	0.23	1.90	.94	0.13	(0.05; 0.21)	-0.76	0.98	1.65	4.90	.93	0.17	(0.05; 0.29)	-1.19	1.48	3.75	7.18	.90	0.06	(-0.11; 0.23)	-1.91	1.90	7.90	9.71	.88	-0.83	(-1.06; -0.60)	-3.50	2.53	14.34	10.50	.85	8.5
GA			0.72	(0.69; 0.75)	0.38	1.07	0.75	1.93	.67	1.22	(1.14; 1.31)	0.23	2.09	3.45	4.40	.76	0.76	(0.64; 0.87)	-0.60	1.87	4.06	5.74	.85	-0.24	(-0.39; -0.10)	-1.96	1.17	5.58	7.14	.87	-2.57	(-2.75; -2.38)	-4.74	-0.77	15.55	9.06	.71	0
LN	exp.		0.27	(0.24; 0.30)	-0.07	0.59	0.28	1.77	.89	1.41	(1.33; 1.50)	0.43	2.28	3.86	5.66	.84	2.07	(1.94; 2.19)	0.57	3.40	8.59	8.45	.84	2.68	(2.50; 2.87)	0.57	4.64	16.08	11.95	.85	3.30	(3.02; 3.58)	0.02	6.18	30.96	17.76	.88	0
WB	growth	emp.	0.84	(0.80; 0.86)	0.50	1.18	0.94	2.19	.78	1.51	(1.41; 1.60)	0.40	2.43	4.64	4.25	.68	0.87	(0.74; 1.00)	-0.63	2.20	5.25	5.43	.79	-0.40	(-0.57; -0.23)	-2.34	1.30	7.57	6.71	.76	-3.19	(-3.41; -2.97)	-5.73	-1.00	22.66	8.49	.56	0
NP			0.58	(0.54; 0.63)	0.05	1.06	0.90	-	-	0.85	(0.73; 0.97)	-0.50	2.01	4.43	-	-	0.91	(0.72; 1.10)	-1.28	2.61	9.95	-	-	0.53	(0.26; 0.81)	-2.50	2.80	19.73	-	-	1.56	(1.05; 2.08)	-4.57	6.17	72.18	-	-	0
PG			0.47	(0.44; 0.50)	0.13	0.81	0.46	1.99	.83	1.03	(0.95; 1.11)	0.08	1.89	2.81	5.09	.90	1.15	(1.02; 1.27)	-0.29	2.46	5.27	7.42	.92	1.09	(0.91; 1.27)	-0.99	3.00	9.47	9.73	.92	0.03	(-0.19; 0.25)	-2.58	3.47	12.38	9.94	.88	11
GA			-0.71	(-0.74; -0.68)	-1.02	-0.42	0.69	1.79	.68	-0.78	(-0.86; -0.70)	-1.70	0.02	2.27	4.05	.79	-1.37	(-1.48; -1.26)	-2.66	-0.27	5.01	5.41	.73	-2.46	(-2.60; -2.32)	-4.12	-1.08	11.24	6.86	.63	-4.87	(-5.05; -4.68)	-7.06	-3.05	32.50	8.86	.47	0
LN	house-		-1.02	(-1.05; -1.00)	-1.31	-0.73	1.22	1.56	.36	-0.84	(-0.91; -0.76)	-1.72	-0.05	2.20	5.10	.90	-0.48	(-0.60; -0.36)	-1.88	0.74	3.86	7.87	.94	-0.10	(-0.28; 0.07)	-2.17	1.65	7.82	11.49	.95	0.29	(0.02; 0.56)	-3.03	3.00	18.79	17.69	.95	0
WB	hold		-0.73	(-0.76; -0.70)	-1.03	-0.43	0.74	1.96	.63	-0.56	(-0.64; -0.47)	-1.52	0.27	2.12	4.03	.79	-1.19	(-1.31; -1.07)	-2.57	0.04	5.13	5.31	.70	-2.38	(-2.54; -2.22)	-4.24	-0.79	12.24	6.77	.59	-5.01	(-5.22; -4.79)	-7.53	-2.92	36.94	8.87	.42	0
NP			-0.80	(-0.84; -0.76)	-1.23	-0.36	1.08	-	-	-1.15	(-1.25; -1.04)	-2.32	≈ 0	4.14	-	-	-1.19	(-1.35; -1.03)	-3.08	0.42	7.98	-	-	-1.41	(-1.66; -1.16)	-4.28	0.82	18.33	-	-	-0.41	(-0.90; 0.08)	-5.88	3.36	62.65	-	-	0.1
PG			-0.91	(-0.94; -0.88)	-1.21	-0.61	1.02	1.77	.52	-1.04	(-1.12; -0.97)	-1.91	-0.24	2.52	4.67	.77	-1.06	(-1.18; -0.95)	-2.37	0.11	4.49	6.96	.82	-1.19	(-1.36; -1.02)	-3.20	0.45	8.79	9.70	.81	-1.88	(-2.12; -1.65)	-4.76	0.97	18.29	11.16	.80	3.9
GA			0.13	(0.10; 0.16)	-0.21	0.49	0.29	2.13	.95	0.42	(0.33; 0.51)	-0.65	1.31	2.44	4.70	.88	-0.09	(-0.21; 0.04)	-1.54	1.17	4.20	6.22	.87	-1.11	(-1.27; -0.95)	-2.93	0.50	8.04	7.84	.82	-3.45	(-3.66; -3.24)	-5.82	-1.43	23.28	10.06	.65	0
LN	con-		-0.26	(-0.29; -0.23)	-0.61	0.10	0.32	1.89	.91	0.32	(0.24; 0.41)	-0.67	1.22	2.08	5.76	.96	0.75	(0.61; 0.88)	-0.79	2.11	5.15	8.72	.95	1.13	(0.94; 1.32)	-1.06	3.12	10.85	12.51	.94	1.43	(1.14; 1.72)	-1.88	4.58	24.13	18.92	.94	0
WB	stant		0.17	(0.14; 0.20)	-0.19	0.55	0.33	2.39	.97	0.68	(0.58; 0.78)	-0.48	1.72	3.02	4.64	.86	0.07	(-0.07; 0.21)	-1.52	1.44	5.10	6.02	.81	-1.15	(-1.33; -0.96)	-3.33	0.59	10.07	7.56	.72	-3.84	(-4.08; -3.60)	-6.69	-1.59	30.11	9.74	.54	0
NP			0.01	(-0.04; 0.06)	-0.59	0.56	0.69	-	-	-0.06	(-0.20; 0.07)	-1.59	1.19	4.57	-	-	-0.10	(-0.30; 0.09)	-2.38	1.74	10.12	-	-	-0.52	(-0.81; -0.22)	-3.80	1.82	23.43	-	-	0.23	(-0.36; 0.81)	-6.81	4.97	90.13	-	-	0
PG			-0.12	(-0.16; -0.09)	-0.49	0.24	0.30	2.15	.94	0.11	(0.03; 0.20)	-0.88	1.03	1.95	5.34	.92	0.17	(0.04; 0.30)	-1.30	1.55	4.43	7.80	.91	0.08	(-0.11; 0.27)	-2.13	2.05	9.33	10.32	.88	-0.88	(-1.12; -0.64)	-3.85	2.62	15.82	11.04	.86	5.8
GA			1.39	(1.35; 1.42)	0.98	1.78	2.27	2.31	.30	2.49	(2.38; 2.60)	1.29	3.54	9.06	5.38	.51	2.24	(2.10; 2.38)	0.65	3.64	10.11	7.04	.71	1.45	(1.27; 1.62)	-0.56	3.22	10.13	8.77	.84	-0.61	(-0.84; -0.39)	-3.18	1.70	13.40	11.11	.86	0
LN	exp.		0.87	(0.84; 0.91)	0.48	1.26	1.07	2.12	.59	2.67	(2.57; 2.77)	1.52	3.69	9.78	6.76	.58	3.62	(3.47; 3.77)	1.86	5.24	19.17	10.09	.64	4.55	(4.33; 4.77)	2.03	6.80	33.08	14.25	.70	5.61	(5.28; 5.93)	1.88	9.06	59.23	21.15	.77	0
WB	growth	db.	1.56	(1.52; 1.60)	1.15	1.98	2.80	2.63	.42	2.75	(2.64; 2.87)	1.42	3.88	11.04	5.15	.44	2.26	(2.10; 2.42)	0.44	3.87	11.63	6.59	.66	1.12	(0.92; 1.32)	-1.21	3.17	11.89	8.14	.79	-1.52	(-1.78; -1.26)	-4.48	1.11	20.04	10.27	.71	0
NP			1.25	(1.19; 1.31)	0.60	1.85	2.46	-	-	2.07	(1.92; 2.22)	0.30	3.59	10.27	-	-	2.27	(2.05; 2.50)	-0.32	4.46	18.74	-	-	2.12	(1.77; 2.47)	-1.78	5.13	36.35	-	-	1.76	(1.15; 2.37)	-5.34	6.22	100.1	-	-	0
PG			1.08	(1.04; 1.12)	0.66	1.48	1.52	2.40	.55	2.29	(2.19; 2.39)	1.15	3.38	7.74	6.07	.72	2.68	(2.53; 2.83)	0.99	4.31	12.78	8.65	.82	2.86	(2.65; 3.07)	0.48	5.11	19.20	10.40	.88	1.39	(1.18; 1.59)	-0.93	4.51	12.63	9.49	.90	10.7
GA			-1.45	(-1.47; -1.42)	-1.74	-1.14	2.30	1.93	.18	-1.93	(-2.01; -1.85)	-2.83	-1.10	5.40	4.07	.55	-2.62	(-2.74; -2.51)	-3.92	-1.52	10.26	5.54	.53	-3.80	(-3.95; -3.65)	-5.53	-2.38	20.28	7.14	.48	-6.30	(-6.50; -6.10)	-8.58	-4.40	50.03	9.39	.37	0
LN	house-		-1.63	(-1.66; -1.61)	-1.94	-1.35	2.86	1.64	.06	-2.23	(-2.30; -2.16)	-3.04	-1.43	6.36	4.84	.61	-2.23	(-2.35; -2.12)	-3.57	-1.04	8.38	7.60	.79	-2.22	(-2.39; -2.05)	-4.19	-0.38	12.51	11.33	.86	-2.35	(-2.62; -2.08)	-5.51	0.50	24.32	17.89	.90	0
WB	hold		-1.51	(-1.54; -1.48)	-1.84	-1.18	2.52	2.03	.12	-1.77	(-1.86; -1.69)	-2.69	-0.97	4.85	4.13	.55	-2.46	(-2.58; -2.34)	-3.86	-1.30	9.78	5.54	.52	-3.66	(-3.82; -3.49)	-5.55	-2.08	20.39	7.21	.46	-6.24	(-6.46; -6.01)	-8.88	-4.10	52.43	9.71	.36	0
NP			-1.50	(-1.54; -1.45)	-1.99	-0.99	2.78	-	-	-2.46	(-2.57; -2.36)	-3.63	-1.37	8.88	-	-	-2.73	(-2.89; -2.57)	-4.58	-1.17	14.02	-	-	-3.21	(-3.46; -2.96)	-6.03	-1.12	26.49	-	-	-2.29	(-2.79; -1.79)	-7.94	1.73	70.73	-	-	0.1
PG			-1.56	(-1.59; -1.53)	-1.86	-1.24	2.64	1.83	.12	-2.32	(-2.39; -2.25)	-3.11	-1.54	6.75	4.59	.46	-2.54	(-2.65; -2.43)	-3.79	-1.41	9.74	6.86	.60	-2.84	(-3.01; -2.67)	-4.78	-1.18	15.50	9.73	.66	-3.62	(-3.87; -3.36)	-6.65	-0.74	29.75	12.08	.68	1.9
GA			0.16	(0.12; 0.20)	-0.26	0.58	0.40	2.47	.95	0.32	(0.21; 0.43)	-0.88	1.37	3.33	5.57	.89	-0.25	(-0.40; -0.10)	-1.90	1.13	5.92	7.32	.87	-1.34	(-1.53; -1.15)	-3.42	0.44	11.20	9.17	.82	-3.76	(-4.01; -3.52)	-6.47	-1.51	29.68	11.70	.68	0
LN	con-		-0.28	(-0.32; -0.24)	-0.69	0.14	0.43	2.22	.91	0.41	(0.30; 0.51)	-0.75	1.53	3.05	7.07	.96	0.92	(0.76; 1.08)	-0.88	2.55	7.68	10.73	.95	1.42	(1.19; 1.66)	-1.26	3.65	16.48	15.40	.94	1.94	(1.57; 2.30)	-2.19	5.27	37.71	23.38	.94	0
WB	stant		0.24	(0.20; 0.28)	-0.18	0.68	0.45	2.80	.98	0.54	(0.42; 0.66)	-0.80	1.70	3.96	5.46	.86	-0.17	(-0.34; -0.01)	-2.09	1.36	7.18	7.02	.81	-1.49	(-1.71; -1.27)	-3.99	0.45	14.27	8.73	.71	-4.32	(-4.60; -4.03)	-7.49	-1.98	39.42	11.16	.54	0
NP			0.04	(-0.01; 0.10)	-0.50	0.57	0.69	-	-	0.01	(-0.14; 0.15)	-1.64	1.38	5.21	-	-	≈ 0	(-0.22; 0.22)	-2.55	2.09	12.85	-	-	-0.24	(-0.61; 0.12)	-4.21	2.26	34.76	-	-	-1.32	(-1.99; -0.66)	-8.06	2.06	115.6	-	-	0
PG			-0.12	(-0.16; -0.08)	-0.54	0.31	0.39	2.54	.95	0.09	(-0.01; 0.20)	-1.07	1.17	2.75	6.25	.90	0.15	(-0.01; 0.30)	-1.64	1.65	6.26	8.97	.89	0.05	(-0.17; 0.28)	-2.51	2.25	12.89	11.38	.87	-1.17	(-1.43; -0.91)	-4.38	2.57	19.10	12.01	.83	10.4
GA			2.34	(2.29; 2.39)	1.77	2.84	6.12	3.03	.07	4.92	(4.76; 5.07)	3.06	6.45	30.61	8.06	.24	5.25	(5.05; 5.46)	2.76	7.26	38.89	10.57	.41	5.03	(4.77; 5.29)	1.90	7.55	42.92	13.15	.63	3.72	(3.39; 4.05)	-0.21	6.90	41.97	16.62	.80	0
LN	exp.		1.71	(1.67; 1.75)	1.22	2.19	3.45	2.89	.26	6.20	(6.03; 6.38)	4.12	7.89	46.23	11.41	.21	8.76	(8.49; 9.03)	5.64	11.34	95.62	17.53	.24	11.62	(11.23; 12.01)	7.11	15.15	175.10	25.34	.29	15.78	(15.18; 16.38)	8.87	21.19	342.8	38.60	.39	0
WB	growth	sq.	2.60	(2.55; 2.65)	2.02	3.13	7.45	3.39	.10	5.00	(4.83; 5.18)	2.95	6.75	32.85	7.61	.25	4.93	(4.70; 5.17)	2.15	7.34	38.51	9.84	.48	4.19	(3.89; 4.48)	0.77	7.14	39.99	12.17	.67	2.03	(1.65; 2.40)	-2.26	5.70	40.55	15.32	.81	0
NP			2.11	(2.04; 2.18)	1.34	2.78	5.68	-	-	5.36	(5.10; 5.62)	2.37	7.54	45.84	-	-	6.63	(6.22; 7.05)	1.92	9.80	89.04	-	-	6.45	(5.92; 6.98)	0.04	10.78	114.96	-	-	2.22	(1.49; 2.96)	-5.56	6.62	145.2	-	-	0.1
PG			1.89	(1.84; 1.94)	1.36	2.37	4.14	3.14	.25	5.14	(4.98; 5.29)	3.23	6.75	32.61	9.52	.40	6.60	(6.37; 6.82)	3.77	9.11	56.88	11.66	.51	6.84	(6.62; 7.07)	4.11	10.40	59.96	10.75	.63	3.46	(3.32; 3.60)	3.82	4.51	17.07	7.64	.76	17.9
GA			-1.91	(-1.94; -1.88)	-2.21	-1.58	3.86	1.97	.03	-3.04	(-3.11; -2.96)	-3.85	-2.34	10.64	3.81	.24	-3.98	(-4.08; -3.87)	-5.13	-2.98	18.74	5.15	.28	-5.39	(-5.54; -5.26)	-6.91	-4.09	34.21	6.67	.28	-8.20	(-8.39; -8.02)	-10.2	-6.43	76.49	8.83	.20	0
LN	house-		-2.04	(-2.07; -2.02)	-2.34	-1.74	4.37	1.64	.01	-3.38	(-3.45; -3.31)	-4.12	-2.66	12.59	4.43	.28	-3.77	(-3.87; -3.66)	-4.93	-2.71	17.03	6.93	.53	-4.20	(-4.35; -4.04)	-5.96	-2.60	23.96	10.32	.68	-4.99	(-5.24; -4.74)	-7.70	-2.46	40.84	16.40	.78	0
WB	hold		-1.99	(-2.02; -1.96)	-2.33	-1.66	4.20	2.02	.01	-2.92	(-3.00; -2.85)	-3.72	-2.21	9.94	3.93	.25	-3.84	(-3.95; -3.73)	-5.05	-2.84	17.86	5.22	.29	-5.25	(-5.40; -5.10)	-6.92	-3.90	33.58	6.80	.28	-8.10	(-8.31; -7.88)	-10.5	-6.30	77.38	9.20	.24	0
NP			-1.92	(-1.96; -1.87)	-2.38	-1.44	4.14	-	-	-3.54	(-3.62; -3.45)	-4.48	-2.61	14.46	-	-	-4.10	(-4.24; -3.95)	-5.59	-2.81	21.91	-	-	-4.96	(-5.17; -4.75)	-7.24	-3.29	36.36	-	-	-3.88	(-4.35; -3.42)	-9.17	-0.19	71.36	-	-	0.1
PG			-1.98	(-2.00; -1.95)	-2.28	-1.66	4.11	1.85	.03	-3.45	(-3.52; -3.39)	-4.18	-2.75	13.08	4.25	.18	-4.04	(-4.14; -3.93)	-5.14	-3.02	19.01	6.30	.33	-4.74	(-4.89; -4.58)	-6.41	-3.27	28.53	9.02	.44	-6.00	(-6.25; -5.76)	-8.75	-3.67	51.62	12.15	.52	2.9

Table 2.2: Results for data generated from lognormal distribution, data set size N = 100, for five percentiles. Each row summarizes 1000 runs. *Abbreviations* True infection risk (Inf. risk): constant, increasing (exponential growth) or decreasing (household transmission). Exposure window widths (Width): empirical set (emp.) or its doubled (db.) or squared (sq.) version. Estimation method (Mod.): MLE assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP), penalized Gaussian mixture (PG). Summary measures: bias with (normal approximated) 95% CI, the 25^{th} and 75^{th} percentiles of the deviations over all runs (p25 and p75), Mean Squared Error (MSE), mean length of the CI (LCI), coverage proportion. Percentage of (1000) runs that timed out (TO). ≈ 0 means a value between -0.005 and 0.005.

Mod.	Inf. risk	Width	50%							90%							95%							97.5%							99%							TO
			Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	%
GA			0.15	(0.14; 0.17)	-0.03	0.34	0.10	1.09	.92	-0.03	(-0.07; 0.01)	-0.47	0.39	0.38	2.03	.90	-0.27	(-0.32; -0.22)	-0.82	0.25	0.68	2.56	.87	-0.59	(-0.65; -0.53)	-1.26	0.02	1.23	3.09	.82	-1.12	(-1.19; -1.05)	-1.93	-0.34	2.58	3.81	.74	0
LN	con-		≈ 0	(-0.01; 0.02)	-0.17	0.19	0.07	1.04	.95	-0.02	(-0.06; 0.01)	-0.44	0.41	0.37	2.32	.94	-0.03	(-0.08; 0.02)	-0.61	0.52	0.67	3.13	.94	-0.04	(-0.10; 0.03)	-0.78	0.63	1.12	4.05	.94	-0.05	(-0.13; 0.04)	-1.07	0.83	1.97	5.42	.94	0
WB	stant		0.30	(0.28; 0.31)	0.10	0.48	0.17	1.32	.96	0.15	(0.11; 0.19)	-0.32	0.58	0.48	1.88	.84	-0.25	(-0.30; -0.20)	-0.86	0.30	0.80	2.24	.76	-0.78	(-0.84; -0.72)	-1.50	-0.12	1.67	2.62	.64	-1.64	(-1.72; -1.57)	-2.53	-0.85	4.26	3.12	.44	0
NP			≈ 0	(-0.02; 0.03)	-0.34	0.33	0.22	-	-	-0.02	(-0.08; 0.04)	-0.71	0.57	0.93	-	-	-0.12	(-0.20; -0.04)	-1.00	0.64	1.68	-	-	-0.29	(-0.4; -0.18)	-1.49	0.71	3.19	-	-	-0.43	(-0.60; -0.26)	-2.36	0.96	7.70	-	-	0.1
PG			0.01	(-0.01; 0.03)	-0.20	0.20	0.08	1.17	.95	-0.06	(-0.10; -0.02)	-0.50	0.36	0.38	2.33	.90	-0.10	(-0.15; -0.05)	-0.70	0.44	0.71	3.07	.88	-0.15	(-0.22; -0.08)	-0.93	0.61	1.24	3.93	.87	-0.22	(-0.32; -0.12)	-1.33	0.80	2.38	5.26	.83	1.5
GA			0.52	(0.50; 0.54)	0.32	0.72	0.36	1.13	.54	0.49	(0.45; 0.53)	0.04	0.90	0.66	2.18	.81	0.30	(0.24; 0.35)	-0.27	0.81	0.76	2.73	.88	0.02	(-0.04; 0.08)	-0.68	0.65	0.97	3.29	.90	-0.46	(-0.53; -0.38)	-1.32	0.34	1.65	4.04	.88	0
LN	exp.		0.36	(0.34; 0.37)	0.17	0.55	0.21	1.09	.72	0.54	(0.50; 0.58)	0.09	0.98	0.72	2.51	.84	0.62	(0.56; 0.67)	≈ 0	1.19	1.15	3.38	.86	0.69	(0.62; 0.76)	-0.12	1.42	1.73	4.36	.88	0.79	(0.70; 0.88)	-0.26	1.75	2.83	5.81	.90	0
WB	growth	emp.	0.72	(0.70; 0.73)	0.51	0.92	0.60	1.35	.61	0.64	(0.60; 0.69)	0.15	1.12	0.93	1.97	.72	0.25	(0.19; 0.30)	-0.36	0.86	0.88	2.35	.80	-0.28	(-0.34; -0.21)	-0.99	0.44	1.23	2.73	.76	-1.14	(-1.22; -1.06)	-2.01	-0.28	2.96	3.23	.58	0
NP			0.38	(0.35; 0.41)	0.03	0.72	0.40	-	-	0.58	(0.51; 0.65)	-0.18	1.29	1.54	-	-	0.46	(0.37; 0.54)	-0.54	1.22	2.25	-	-	0.18	(0.07; 0.30)	-1.11	1.27	3.41	-	-	-0.18	(-0.35; -0.01)	-2.11	1.26	7.37	-	-	0
PG			0.38	(0.36; 0.40)	0.17	0.59	0.24	1.22	.75	0.50	(0.46; 0.54)	0.04	0.91	0.68	2.48	.89	0.51	(0.46; 0.57)	-0.10	1.10	1.05	3.25	.91	0.51	(0.44; 0.58)	-0.27	1.29	1.61	4.15	.91	0.50	(0.40; 0.60)	-0.58	1.57	2.80	5.53	.90	1.8
GA			-0.38	(-0.40; -0.36)	-0.56	-0.21	0.21	1.05	.73	-0.72	(-0.76; -0.69)	-1.11	-0.35	0.84	1.85	.64	-1.01	(-1.06; -0.97)	-1.52	-0.54	1.55	2.34	.60	-1.37	(-1.43; -1.32)	-1.98	-0.81	2.66	2.86	.56	-1.94	(-2.01; -1.88)	-2.71	-1.26	4.98	3.55	.50	0
LN	house-		-0.50	(-0.51; -0.48)	-0.68	-0.32	0.31	0.99	.53	-0.77	(-0.80; -0.73)	-1.16	-0.40	0.90	2.08	.70	-0.86	(-0.91; -0.81)	-1.37	-0.37	1.31	2.83	.77	-0.95	(-1.01; -0.89)	-1.60	-0.32	1.85	3.70	.82	-1.06	(-1.14; -0.98)	-1.96	-0.26	2.83	5.01	.85	0
WB	hold		-0.31	(-0.33; -0.3)	-0.51	-0.12	0.17	1.28	.76	-0.53	(-0.56; -0.49)	-0.96	-0.14	0.65	1.76	.66	-0.92	(-0.97; -0.88)	-1.49	-0.42	1.48	2.12	.54	-1.44	(-1.50; -1.38)	-2.14	-0.83	3.02	2.50	.42	-2.29	(-2.36; -2.21)	-3.16	-1.52	6.65	3.01	.28	0
NP			-0.52	(-0.55; -0.49)	-0.83	-0.21	0.48	-	-	-0.76	(-0.82; -0.71)	-1.39	-0.19	1.34	-	-	-0.81	(-0.88; -0.74)	-1.65	-0.07	2.06	-	-	-1.02	(-1.12; -0.92)	-2.15	-0.13	3.60	-	-	-0.86	(-1.01; -0.70)	-2.71	0.64	7.29	-	-	0
PG			-0.51	(-0.53; -0.49)	-0.69	-0.32	0.33	1.08	.56	-0.80	(-0.83; -0.76)	-1.17	-0.42	0.95	2.18	.62	-0.89	(-0.94; -0.84)	-1.42	-0.40	1.38	2.87	.66	-0.98	(-1.04; -0.92)	-1.70	-0.35	2.00	3.70	.69	-1.09	(-1.18; -1.00)	-2.13	-0.18	3.26	5.00	.70	0.8
GA			0.16	(0.14; 0.18)	-0.06	0.37	0.13	1.27	.92	-0.04	(-0.09; <0.01)	-0.51	0.42	0.51	2.35	.89	-0.29	(-0.35; -0.23)	-0.92	0.29	0.90	2.96	.87	-0.61	(-0.68; -0.55)	-1.37	0.10	1.58	3.59	.84	-1.15	(-1.23; -1.06)	-2.09	-0.28	3.14	4.44	.77	0
LN	con-		0.01	(-0.01; 0.03)	-0.20	0.22	0.10	1.21	.94	-0.05	(-0.10; -0.01)	-0.53	0.40	0.49	2.66	.93	-0.07	(-0.13; -0.02)	-0.73	0.52	0.90	3.60	.93	-0.10	(-0.17; -0.02)	-0.94	0.66	1.50	4.67	.93	-0.13	(-0.23; -0.03)	-1.28	0.87	2.66	6.28	.93	0
WB	stant		0.30	(0.28; 0.33)	0.08	0.53	0.21	1.56	.97	0.12	(0.08; 0.17)	-0.42	0.63	0.62	2.19	.83	-0.29	(-0.35; -0.23)	-1.00	0.36	1.06	2.61	.75	-0.83	(-0.90; -0.75)	-1.67	-0.06	2.11	3.05	.65	-1.70	(-1.79; -1.61)	-2.72	-0.76	4.99	3.63	.49	0
NP			0.02	(-0.01; 0.06)	-0.39	0.39	0.34	-	-	-0.09	(-0.16; -0.02)	-0.91	0.60	1.30	-	-	-0.24	(-0.33; -0.16)	-1.25	0.66	2.13	-	-	-0.51	(-0.63; -0.39)	-1.88	0.59	4.20	-	-	-1.19	(-1.36; -1.02)	-3.14	0.14	8.96	-	-	0
PG			0.02	(<0.01; 0.04)	-0.21	0.24	0.11	1.36	.95	-0.09	(-0.14; -0.05)	-0.57	0.36	0.50	2.68	.90	-0.15	(-0.21; -0.09)	-0.81	0.45	0.92	3.49	.87	-0.21	(-0.29; -0.13)	-1.08	0.59	1.59	4.44	.84	-0.29	(-0.40; -0.18)	-1.53	0.79	3.01	5.87	.82	1.9
GA			0.86	(0.84; 0.89)	0.62	1.09	0.88	1.36	.26	1.01	(0.96; 1.06)	0.44	1.56	1.69	2.74	.64	0.88	(0.82; 0.94)	0.18	1.53	1.82	3.44	.78	0.66	(0.59; 0.74)	-0.18	1.47	1.95	4.15	.86	0.26	(0.17; 0.35)	-0.78	1.26	2.31	5.09	.88	0
LN	exp.		0.69	(0.66; 0.71)	0.45	0.91	0.59	1.32	.42	1.11	(1.06; 1.16)	0.54	1.67	1.92	3.18	.66	1.28	(1.21; 1.35)	0.53	1.97	2.88	4.30	.74	1.46	(1.37; 1.54)	0.46	2.34	4.15	5.57	.78	1.69	(1.57; 1.81)	0.34	2.85	6.43	7.44	.81	0
WB	growth	db.	1.10	(1.08; 1.13)	0.85	1.35	1.36	1.63	.34	1.11	(1.05; 1.16)	0.50	1.71	2.02	2.46	.60	0.73	(0.66; 0.80)	-0.01	1.49	1.76	2.92	.74	0.21	(0.13; 0.29)	-0.70	1.08	1.77	3.40	.80	-0.65	(-0.74; -0.55)	-1.73	0.40	2.86	4.01	.71	0
NP			0.74	(0.70; 0.78)	0.28	1.16	0.96	-	-	1.05	(0.97; 1.14)	0.13	1.86	2.86	-	-	0.88	(0.76; 0.99)	-0.36	1.90	4.01	-	-	0.58	(0.45; 0.72)	-1.01	1.75	5.01	-	-	-0.62	(-0.79; -0.45)	-2.60	0.73	7.89	-	-	0
PG			0.72	(0.69; 0.74)	0.46	0.96	0.65	1.49	.49	1.05	(1.00; 1.10)	0.48	1.59	1.79	3.07	.79	1.14	(1.07; 1.21)	0.34	1.90	2.55	4.01	.84	1.22	(1.13; 1.31)	0.20	2.19	3.59	5.08	.87	1.32	(1.20; 1.44)	-0.11	2.65	5.61	6.69	.89	4.1
GA			-0.76	(-0.78; -0.74)	-0.94	-0.57	0.65	1.20	.31	-1.25	(-1.28; -1.21)	-1.66	-0.85	1.90	1.93	.39	-1.58	(-1.63; -1.53)	-2.14	-1.07	3.08	2.47	.40	-1.98	(-2.04; -1.92)	-2.66	-1.33	4.80	3.04	.40	-2.60	(-2.67; -2.52)	-3.47	-1.78	8.15	3.82	.38	0
LN	house-		-0.84	(-0.86; -0.83)	-1.02	-0.66	0.79	1.11	.18	-1.34	(-1.38; -1.30)	-1.73	-0.97	2.12	2.13	.42	-1.52	(-1.56; -1.47)	-2.07	-1.00	2.90	2.92	.56	-1.69	(-1.75; -1.62)	-2.40	-1.01	3.88	3.86	.64	-1.90	(-1.99; -1.81)	-2.88	-1.01	5.52	5.29	.72	0
WB	hold		-0.75	(-0.76; -0.73)	-0.95	-0.54	0.66	1.45	.28	-1.06	(-1.10; -1.02)	-1.51	-0.64	1.51	1.91	.42	-1.47	(-1.52; -1.42)	-2.08	-0.91	2.83	2.29	.37	-1.99	(-2.05; -1.93)	-2.77	-1.28	5.01	2.72	.31	-2.83	(-2.91; -2.75)	-3.83	-1.93	9.68	3.32	.24	0
NP			-0.88	(-0.91; -0.85)	-1.23	-0.53	1.04	-	-	-1.38	(-1.44; -1.33)	-2.00	-0.80	2.66	-	-	-1.57	(-1.64; -1.49)	-2.44	-0.80	3.94	-	-	-1.75	(-1.85; -1.65)	-2.87	-0.82	5.62	-	-	-1.76	(-1.93; -1.60)	-3.71	-0.40	10.21	-	-	0
PG			-0.86	(-0.87; -0.84)	-1.06	-0.66	0.82	1.17	.22	-1.38	(-1.42; -1.35)	-1.78	-1.00	2.25	2.26	.32	-1.55	(-1.60; -1.50)	-2.11	-1.02	3.03	2.97	.42	-1.71	(-1.77; -1.64)	-2.49	-1.01	4.03	3.81	.50	-1.89	(-1.98; -1.80)	-3.01	-0.90	5.80	5.10	.56	0.8
GA			0.16	(0.13; 0.18)	-0.09	0.40	0.16	1.43	.92	-0.10	(-0.15; -0.05)	-0.68	0.45	0.66	2.67	.89	-0.37	(-0.43; -0.30)	-1.10	0.32	1.17	3.35	.86	-0.71	(-0.79; -0.63)	-1.60	0.13	2.01	4.04	.82	-1.27	(-1.36; -1.18)	-2.33	-0.25	3.87	4.97	.76	0
LN	con-		0.01	(-0.02; 0.03)	-0.23	0.24	0.13	1.37	.94	-0.07	(-0.12; -0.02)	-0.66	0.47	0.65	3.09	.94	-0.10	(-0.17; -0.04)	-0.86	0.58	1.16	4.16	.93	-0.13	(-0.22; -0.04)	-1.09	0.76	1.91	5.38	.93	-0.17	(-0.28; -0.05)	-1.44	0.98	3.36	7.21	.94	0
WB	stant		0.31	(0.29; 0.34)	0.04	0.56	0.24	1.79	.98	0.06	(0.01; 0.12)	-0.60	0.68	0.79	2.49	.83	-0.37	(-0.44; -0.30)	-1.24	0.38	1.39	2.93	.72	-0.93	(-1.02; -0.85)	-1.94	-0.04	2.66	3.39	.62	-1.83	(-1.93; -1.73)	-3.04	-0.78	5.96	4.01	.49	0
NP			0.02	(-0.01; 0.05)	-0.34	0.35	0.29	-	-	0.02	(-0.05; 0.10)	-0.76	0.64	1.33	-	-	-0.06	(-0.17; 0.04)	-1.17	0.79	2.82	-	-	-0.22	(-0.35; -0.09)	-1.64	0.97	4.57	-	-	-1.37	(-1.53; -1.20)	-3.25	0.03	9.10	-	-	0
PG			0.01	(-0.01; 0.04)	-0.25	0.26	0.14	1.54	.95	-0.13	(-0.18; -0.08)	-0.72	0.42	0.67	3.05	.88	-0.20	(-0.27; -0.13)	-0.99	0.53	1.22	3.94	.85	-0.28	(-0.36; -0.19)	-1.25	0.65	2.07	4.97	.83	-0.38	(-0.50; -0.26)	-1.75	0.86	3.85	6.50	.81	3.4
GA			1.14	(1.11; 1.16)	0.81	1.42	1.48	1.60	.16	1.56	(1.49; 1.62)	0.85	2.23	3.52	3.48	.53	1.52	(1.44; 1.60)	0.64	2.37	4.02	4.36	.67	1.40	(1.30; 1.50)	0.33	2.42	4.38	5.24	.77	1.11	(1.00; 1.23)	-0.16	2.35	4.79	6.40	.85	0
LN	exp.		0.95	(0.92; 0.98)	0.65	1.23	1.08	1.58	.28	1.85	(1.78; 1.92)	1.09	2.56	4.69	4.30	.52	2.24	(2.15; 2.34)	1.22	3.19	7.33	5.86	.60	2.65	(2.53; 2.77)	1.33	3.87	10.84	7.60	.64	3.22	(3.06; 3.38)	1.44	4.81	17.10	10.19	.69	0
WB	growth	sq.	1.41	(1.38; 1.43)	1.07	1.72	2.19	1.86	.19	1.55	(1.48; 1.62)	0.76	2.28	3.65	3.02	.52	1.21	(1.12; 1.29)	0.26	2.11	3.34	3.60	.70	0.73	(0.63; 0.83)	-0.38	1.76	3.10	4.17	.78	-0.09	(-0.21; 0.03)	-1.41	1.13	3.57	4.90	.77	0
NP			1.02	(0.98; 1.06)	0.54	1.45	1.46	-	-	1.91	(1.80; 2.02)	0.72	2.81	6.63	-	-	1.87	(1.73; 2.00)	0.32	3.04	8.54	-	-	1.35	(1.19; 1.52)	-0.54	2.76	9.15	-	-	-0.49	(-0.66; -0.32)	-2.40	0.91	7.82	-	-	0.1
PG			0.98	(0.95; 1.00)	0.65	1.28	1.15	1.77	.36	1.74	(1.67; 1.81)	1.00	2.49	4.29	4.04	.65	2.03	(1.94; 2.13)	0.98	3.00	6.44	5.38	.73	2.32	(2.20; 2.45)	0.91	3.57	9.31	6.90	.78	2.72	(2.55; 2.89)	0.78	4.44	14.60	9.06	.83	5
GA			-0.90	(-0.92; -0.89)	-1.09	-0.71	0.90	1.26	.19	-1.51	(-1.54; -1.47)	-1.91	-1.15	2.60	1.90	.23	-1.88	(-1.92; -1.83)	-2.41	-1.41	4.08	2.42	.27	-2.31	(-2.37; -2.25)	-2.98	-1.73	6.20	2.98	.29	-2.97	(-3.05; -2.90)	-3.83	-2.26	10.22	3.75	.28	0
LN	house-		-0.97	(-0.99; -0.96)	-1.15	-0.78	1.03	1.15	.11	-1.61	(-1.64; -1.57)	-2.00	-1.27	2.88	2.07	.25	-1.84	(-1.89; -1.80)	-2.37	-1.37	3.96	2.82	.41	-2.08	(-2.14; -2.01)	-2.79	-1.46	5.27	3.71	.51	-2.38	(-2.46; -2.29)	-3.29	-1.56	7.43	5.09	.62	0
WB	hold		-0.93	(-0.95; -0.91)	-1.16	-0.70	0.98	1.52	.12	-1.32	(-1.36; -1.29)	-1.74	-0.94	2.11	1.93	.28	-1.75	(-1.80; -1.70)	-2.34	-1.23	3.70	2.29	.25	-2.29	(-2.35; -2.22)	-3.05	-1.63	6.24	2.70	.22	-3.14	(-3.22; -3.06)	-4.08	-2.31	11.49	3.29	.17	0
NP			-1.01	(-1.03; -0.98)	-1.31	-0.71	1.22	-	-	-1.64	(-1.69; -1.59)	-2.15	-1.19	3.28	-	-	-1.82	(-1.89; -1.75)	-2.58	-1.16	4.53	-	-	-1.98	(-2.07; -1.89)	-3.02	-1.13	6.06	-	-	-1.77	(-1.93; -1.61)	-3.63	-0.29	9.95	-	-	0
PG			-0.98	(-1.00; -0.96)	-1.16	-0.77	1.05	1.20	.13	-1.65	(-1.69; -1.62)	-2.05	-1.28	3.04	2.23	.18	-1.88	(-1.93; -1.83)	-2.43	-1.39	4.11	2.91	.29	-2.10	(-2.16; -2.04)	-2.81	-1.45	5.43	3.71	.38	-2.38	(-2.46; -2.29)	-3.40	-1.49	7.69	4.94	.47	1.1

Table 2.3: Results for data generated from Weibull distribution, data set size N = 100, for five percentiles. Each row summarizes 1000 runs. *Abbreviations* True infection risk (Inf. risk): constant, increasing (exponential growth) or decreasing (household transmission). Exposure window widths (Width): empirical set (emp.) or its doubled (db.) or squared (sq.) version. Estimation method (Mod.): MLE assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP), penalized Gaussian mixture (PG). Summary measures: bias with (normal approximated) 95% CI, the 25^{th} and 75^{th} percentiles of the deviations over all runs (p25 and p75), Mean Squared Error (MSE), mean length of the CI (LCI), coverage proportion. Percentage of (1000) runs that timed out (TO). ≈ 0 means a value between -0.005 and 0.005.

Mod.	Inf. risk	Width	50%							90%							95%							97.5%							99%							TO
			Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	%
GA			-0.21	(-0.23; -0.20)	-0.42	-0.01	0.14	1.16	.89	0.10	(0.07; 0.13)	-0.22	0.42	0.25	2.22	.97	0.41	(0.37; 0.45)	-0.02	0.81	0.56	2.81	.94	0.77	(0.72; 0.81)	0.25	1.26	1.17	3.41	.89	1.29	(1.23; 1.35)	0.64	1.93	2.59	4.22	.78	0
LN	con-		-0.40	(-0.42; -0.38)	-0.62	-0.19	0.26	1.14	.73	0.32	(0.29; 0.36)	-0.08	0.69	0.42	2.72	.96	1.04	(0.98; 1.09)	0.47	1.56	1.75	3.76	.83	1.91	(1.84; 1.98)	1.14	2.66	4.93	4.95	.58	3.30	(3.19; 3.39)	2.20	4.38	13.45	6.77	.34	0
WB	stant		0.02	(<0.01; 0.04)	-0.20	0.23	0.10	1.32	.96	-0.03	(-0.05; <0.01)	-0.32	0.27	0.21	1.87	.94	-0.04	(-0.07; <0.01)	-0.40	0.34	0.31	2.25	.94	-0.05	(-0.09; -0.01)	-0.48	0.40	0.43	2.64	.93	-0.06	(-0.11; -0.01)	-0.59	0.49	0.63	3.16	.93	0
NP			0.03	(-0.01; 0.06)	-0.31	0.37	0.28	-	-	-0.03	(-0.08; 0.02)	-0.55	0.46	0.57	-	-	-0.10	(-0.16; -0.05)	-0.73	0.52	0.79	-	-	-0.25	(-0.31; -0.18)	-1.01	0.37	1.18	-	-	-0.61	(-0.69; -0.53)	-1.47	0.14	2.03	-	-	0
PG			-0.11	(-0.13; -0.09)	-0.36	0.13	0.14	1.38	.92	-0.04	(-0.06; -0.01)	-0.32	0.26	0.21	1.95	.93	0.18	(0.14; 0.21)	-0.19	0.55	0.37	2.42	.95	0.47	(0.42; 0.51)	-0.02	0.93	0.77	2.98	.94	0.94	(0.88; 1.01)	0.25	1.60	1.92	3.81	.88	26.4
GA			0.18	(0.16; 0.20)	-0.02	0.39	0.13	1.19	.90	0.64	(0.61; 0.67)	0.30	0.98	0.67	2.36	.86	0.99	(0.95; 1.03)	0.53	1.42	1.41	2.98	.76	1.38	(1.33; 1.43)	0.82	1.93	2.56	3.60	.64	1.95	(1.89; 2.01)	1.24	2.64	4.83	4.43	.52	0
LN	exp.		-0.01	(-0.03; 0.01)	-0.23	0.20	0.10	1.19	.94	0.94	(0.90; 0.98)	0.50	1.33	1.25	2.92	.78	1.73	(1.68; 1.79)	1.12	2.31	3.80	4.01	.50	2.69	(2.61; 2.77)	1.83	3.47	8.73	5.26	.31	4.17	(4.06; 4.28)	2.96	5.32	20.42	7.15	.16	0
WB	growth	emp.	0.44	(0.42; 0.46)	0.22	0.65	0.29	1.34	.88	0.42	(0.39; 0.45)	0.11	0.74	0.40	1.96	.92	0.40	(0.37; 0.44)	0.03	0.79	0.49	2.34	.94	0.38	(0.34; 0.42)	-0.07	0.85	0.60	2.74	.95	0.35	(0.30; 0.40)	-0.20	0.92	0.78	3.25	.95	0
NP			0.45	(0.42; 0.49)	0.09	0.78	0.48	-	-	0.40	(0.35; 0.44)	-0.13	0.91	0.79	-	-	0.31	(0.25; 0.36)	-0.38	0.90	0.93	-	-	0.12	(0.05; 0.18)	-0.62	0.75	1.20	-	-	-0.46	(-0.54; -0.38)	-1.30	0.25	1.81	-	-	0
PG			0.31	(0.28; 0.33)	0.06	0.55	0.22	1.40	.84	0.45	(0.42; 0.48)	0.15	0.76	0.42	2.03	.91	0.67	(0.63; 0.71)	0.27	1.08	0.82	2.54	.88	0.96	(0.91; 1.01)	0.39	1.52	1.56	3.13	.84	1.44	(1.38; 1.51)	0.68	2.17	3.26	4.00	.76	30.1
GA			-0.80	(-0.81; -0.78)	-1.00	-0.60	0.72	1.13	.23	-0.61	(-0.63; -0.58)	-0.91	-0.30	0.57	2.06	.83	-0.33	(-0.36; -0.29)	-0.72	0.06	0.44	2.63	.94	0.01	(-0.04; 0.05)	-0.46	0.47	0.50	3.22	.97	0.51	(0.46; 0.57)	-0.08	1.10	1.06	4.03	.96	0
LN	house-		-0.97	(-0.99; -0.95)	-1.17	-0.76	1.03	1.09	.09	-0.48	(-0.51; -0.44)	-0.83	-0.13	0.49	2.51	.92	0.17	(0.12; 0.22)	-0.32	0.62	0.56	3.50	.98	0.99	(0.93; 1.05)	0.31	1.62	2.00	4.68	.90	2.31	(2.22; 2.40)	1.33	3.21	7.46	6.50	.64	0
WB	hold		-0.61	(-0.63; -0.60)	-0.82	-0.41	0.47	1.29	.37	-0.64	(-0.67; -0.62)	-0.93	-0.36	0.60	1.79	.66	-0.62	(-0.65; -0.59)	-0.98	-0.28	0.66	2.17	.75	-0.59	(-0.63; -0.55)	-1.04	-0.19	0.75	2.58	.81	-0.55	(-0.60; -0.50)	-1.09	-0.03	0.9	3.13	.85	0
NP			-0.66	(-0.69; -0.62)	-1.00	-0.29	0.71	-	-	-0.66	(-0.71; -0.62)	-1.17	-0.16	0.97	-	-	-0.65	(-0.70; -0.60)	-1.21	-0.12	1.12	-	-	-0.69	(-0.76; -0.63)	-1.45	-0.03	1.52	-	-	-0.78	(-0.86; -0.70)	-1.68	-0.01	2.33	-	-	0
PG			-0.73	(-0.75; -0.71)	-0.96	-0.50	0.65	1.35	.46	-0.69	(-0.72; -0.67)	-0.98	-0.40	0.67	1.86	.64	-0.47	(-0.50; -0.44)	-0.85	-0.11	0.52	2.32	.82	-0.16	(-0.21; -0.12)	-0.64	0.31	0.52	2.87	.90	0.35	(0.29; 0.41)	-0.32	1.01	1.05	3.72	.93	15.2
GA			-0.21	(-0.23; -0.19)	-0.43	0.02	0.17	1.34	.91	0.07	(0.04; 0.11)	-0.30	0.43	0.31	2.55	.97	0.37	(0.32; 0.41)	-0.11	0.83	0.64	3.23	.97	0.71	(0.66; 0.77)	0.11	1.29	1.27	3.94	.93	1.23	(1.16; 1.29)	0.44	1.94	2.71	4.89	.85	0
LN	con-		-0.40	(-0.42; -0.37)	-0.64	-0.15	0.29	1.30	.79	0.23	(0.20; 0.27)	-0.18	0.63	0.44	3.04	.98	0.90	(0.84; 0.95)	0.30	1.48	1.60	4.20	.90	1.73	(1.65; 1.80)	0.87	2.51	4.48	5.55	.74	3.03	(2.92; 3.14)	1.83	4.14	12.26	7.63	.52	0
WB	stant		0.02	(<0.01; 0.04)	-0.22	0.26	0.13	1.57	.96	-0.03	(-0.07; <0.01)	-0.39	0.31	0.28	2.20	.94	-0.05	(-0.09; -0.01)	-0.49	0.39	0.42	2.64	.93	-0.06	(-0.11; -0.01)	-0.59	0.46	0.59	3.10	.92	-0.08	(-0.14; -0.02)	-0.73	0.56	0.87	3.71	.92	0
NP			0.03	(-0.01; 0.07)	-0.40	0.45	0.42	-	-	-0.05	(-0.10; 0.01)	-0.69	0.57	0.83	-	-	-0.20	(-0.26; -0.13)	-0.97	0.49	1.11	-	-	-0.41	(-0.48; -0.33)	-1.26	0.33	1.61	-	-	-1.06	(-1.14; -0.98)	-1.95	-0.28	2.87	-	-	0
PG			-0.15	(-0.17; -0.12)	-0.42	0.09	0.18	1.59	.93	-0.05	(-0.08; -0.01)	-0.39	0.29	0.28	2.26	.93	0.20	(0.16; 0.24)	-0.23	0.61	0.48	2.84	.94	0.53	(0.48; 0.59)	-0.02	1.10	1.01	3.53	.95	1.08	(1.00; 1.15)	0.29	1.86	2.47	4.56	.92	22.5
GA			0.55	(0.52; 0.57)	0.31	0.79	0.42	1.41	.65	1.13	(1.09; 1.17)	0.70	1.54	1.67	2.93	.66	1.52	(1.47; 1.57)	0.93	2.05	2.95	3.69	.58	1.94	(1.88; 2.01)	1.23	2.61	4.77	4.47	.50	2.56	(2.48; 2.63)	1.68	3.36	8.08	5.50	.42	0
LN	exp.		0.35	(0.33; 0.37)	0.11	0.60	0.26	1.40	.80	1.46	(1.42; 1.51)	0.94	1.92	2.69	3.58	.57	2.31	(2.24; 2.38)	1.55	3.00	6.51	4.92	.37	3.32	(3.23; 3.41)	2.30	4.24	13.22	6.46	.26	4.87	(4.74; 5.00)	3.37	6.18	28.08	8.77	.16	0
WB	growth	db.	0.82	(0.79; 0.84)	0.57	1.07	0.80	1.61	.67	0.82	(0.79; 0.86)	0.45	1.21	1.01	2.42	.83	0.80	(0.76; 0.84)	0.33	1.28	1.13	2.90	.90	0.78	(0.72; 0.83)	0.21	1.33	1.28	3.37	.93	0.74	(0.68; 0.80)	0.02	1.42	1.52	3.97	.95	0
NP			0.85	(0.81; 0.89)	0.36	1.29	1.16	-	-	0.74	(0.68; 0.80)	0.05	1.42	1.50	-	-	0.50	(0.43; 0.57)	-0.30	1.18	1.47	-	-	0.20	(0.12; 0.28)	-0.65	0.92	1.54	-	-	-0.74	(-0.82; -0.66)	-1.63	0.01	2.20	-	-	0
PG			0.63	(0.60; 0.65)	0.35	0.91	0.56	1.62	.67	0.91	(0.87; 0.94)	0.54	1.31	1.16	2.50	.79	1.20	(1.15; 1.25)	0.71	1.70	2.00	3.16	.77	1.56	(1.50; 1.62)	0.87	2.21	3.39	3.93	.75	2.13	(2.05; 2.21)	1.20	3.00	6.29	5.04	.70	34.6
GA			-1.22	(-1.24; -1.20)	-1.45	-1.00	1.60	1.28	.05	-1.16	(-1.20; -1.13)	-1.51	-0.83	1.60	2.16	.51	-0.92	(-0.96; -0.88)	-1.34	-0.51	1.24	2.79	.78	-0.62	(-0.66; -0.57)	-1.13	-0.11	0.98	3.46	.90	-0.14	(-0.21; -0.08)	-0.80	0.51	0.99	4.37	.97	0
LN	house-		-1.36	(-1.39; -1.34)	-1.60	-1.15	1.99	1.21	.02	-1.15	(-1.19; -1.12)	-1.52	-0.79	1.61	2.53	.65	-0.62	(-0.66; -0.57)	-1.13	-0.13	0.93	3.57	.93	0.10	(0.04; 0.16)	-0.60	0.74	1.05	4.82	.98	1.28	(1.19; 1.38)	0.25	2.19	3.83	6.82	.91	0
WB	hold		-1.08	(-1.10; -1.06)	-1.31	-0.86	1.30	1.50	.07	-1.13	(-1.16; -1.10)	-1.45	-0.80	1.51	1.99	.38	-1.09	(-1.12; -1.05)	-1.48	-0.67	1.53	2.41	.54	-1.03	(-1.08; -0.99)	-1.54	-0.55	1.58	2.88	.66	-0.95	(-1.01; -0.90)	-1.58	-0.35	1.69	3.55	.75	0
NP			-1.13	(-1.17; -1.09)	-1.57	-0.70	1.68	-	-	-1.17	(-1.22; -1.12)	-1.72	-0.60	2.05	-	-	-1.15	(-1.21; -1.09)	-1.83	-0.49	2.24	-	-	-1.21	(-1.27; -1.14)	-1.97	-0.50	2.67	-	-	-1.37	(-1.46; -1.29)	-2.34	-0.52	3.69	-	-	0
PG			-1.20	(-1.22; -1.17)	-1.45	-0.96	1.57	1.49	.14	-1.23	(-1.26; -1.20)	-1.57	-0.90	1.75	2.04	.31	-0.98	(-1.01; -0.94)	-1.41	-0.55	1.32	2.54	.61	-0.63	(-0.68; -0.58)	-1.19	-0.11	0.99	3.18	.78	-0.04	(-0.10; 0.03)	-0.76	0.66	1.10	4.18	.91	9
GA			-0.21	(-0.24; -0.19)	-0.49	0.05	0.21	1.55	.92	0.11	(0.07; 0.15)	-0.34	0.55	0.43	2.98	.97	0.42	(0.37; 0.47)	-0.14	0.98	0.86	3.76	.96	0.78	(0.72; 0.85)	0.11	1.47	1.64	4.57	.94	1.32	(1.24; 1.40)	0.45	2.19	3.36	5.65	.89	0
LN	con-		-0.41	(-0.44; -0.38)	-0.70	-0.12	0.35	1.52	.81	0.38	(0.33; 0.43)	-0.12	0.84	0.72	3.73	.97	1.14	(1.07; 1.21)	0.40	1.80	2.53	5.17	.90	2.07	(1.97; 2.16)	1.02	3.00	6.68	6.85	.76	3.54	(3.40; 3.68)	1.98	4.83	17.53	9.44	.55	0
WB	stant		0.03	(<0.01; 0.05)	-0.26	0.30	0.18	1.81	.96	-0.05	(-0.09; -0.01)	-0.48	0.37	0.37	2.54	.94	-0.07	(-0.12; -0.03)	-0.59	0.43	0.55	3.01	.92	-0.09	(-0.14; -0.04)	-0.70	0.51	0.77	3.50	.91	-0.11	(-0.18; -0.05)	-0.87	0.61	1.11	4.16	.90	0
NP			0.03	(-0.01; 0.07)	-0.37	0.43	0.37	-	-	-0.01	(-0.07; 0.05)	-0.63	0.56	0.86	-	-	-0.08	(-0.15; -0.01)	-0.82	0.60	1.18	-	-	-0.15	(-0.22; -0.07)	-1.03	0.57	1.53	-	-	-0.96	(-1.05; -0.88)	-1.94	-0.16	2.69	-	-	0
PG			-0.16	(-0.19; -0.13)	-0.48	0.16	0.23	1.83	.94	-0.01	(-0.05; 0.03)	-0.44	0.40	0.38	2.62	.93	0.27	(0.22; 0.32)	-0.28	0.79	0.69	3.29	.94	0.64	(0.58; 0.70)	-0.07	1.30	1.44	4.08	.93	1.24	(1.15; 1.32)	0.28	2.15	3.43	5.26	.90	24.9
GA			0.84	(0.82; 0.87)	0.57	1.11	0.87	1.69	.44	1.79	(1.74; 1.84)	1.20	2.31	3.88	3.78	.43	2.31	(2.24; 2.38)	1.54	2.98	6.49	4.77	.38	2.87	(2.78; 2.95)	1.96	3.70	9.96	5.76	.34	3.64	(3.54; 3.74)	2.50	4.69	15.99	7.06	.30	0
LN	exp.		0.66	(0.64; 0.69)	0.38	0.94	0.60	1.74	.65	2.53	(2.46; 2.60)	1.74	3.20	7.74	5.12	.29	3.76	(3.66; 3.87)	2.60	4.72	17.08	7.13	.20	5.19	(5.04; 5.33)	3.55	6.53	32.48	9.46	.14	7.35	(7.14; 7.56)	4.98	9.25	65.36	13.02	.08	0
WB	growth	sq.	1.09	(1.07; 1.12)	0.81	1.36	1.36	1.82	.47	1.22	(1.18; 1.26)	0.75	1.68	1.99	2.96	.73	1.23	(1.18; 1.28)	0.64	1.82	2.26	3.55	.83	1.23	(1.17; 1.30)	0.53	1.90	2.55	4.11	.89	1.23	(1.15; 1.30)	0.38	2.00	2.96	4.83	.92	0
NP			1.11	(1.07; 1.15)	0.67	1.56	1.62	-	-	1.24	(1.18; 1.31)	0.44	1.90	2.73	-	-	0.98	(0.91; 1.06)	0.13	1.70	2.46	-	-	0.44	(0.36; 0.52)	-0.50	1.20	1.92	-	-	-0.57	(-0.65; -0.49)	-1.50	0.19	2.05	-	-	0
PG			0.88	(0.86; 0.91)	0.58	1.15	0.97	1.82	.50	1.49	(1.45; 1.54)	0.99	2.02	2.83	3.32	.58	1.94	(1.87; 2.00)	1.19	2.61	4.88	4.33	.59	2.46	(2.38; 2.55)	1.48	3.38	8.01	5.46	.59	3.27	(3.15; 3.38)	1.87	4.48	14.25	7.12	.57	36.2
GA			-1.42	(-1.44; -1.40)	-1.67	-1.17	2.15	1.38	.02	-1.42	(-1.45; -1.39)	-1.76	-1.07	2.26	2.18	.35	-1.19	(-1.22; -1.15)	-1.61	-0.75	1.81	2.79	.68	-0.89	(-0.94; -0.84)	-1.41	-0.34	1.40	3.47	.86	-0.42	(-0.48; -0.36)	-1.08	0.28	1.17	4.41	.94	0
LN	house-		-1.56	(-1.59; -1.54)	-1.82	-1.29	2.59	1.27	.01	-1.43	(-1.46; -1.40)	-1.81	-1.06	2.33	2.54	.49	-0.91	(-0.95; -0.86)	-1.43	-0.42	1.39	3.57	.87	-0.20	(-0.27; -0.14)	-0.94	0.46	1.11	4.85	.97	0.98	(0.88; 1.07)	-0.08	1.92	3.24	6.90	.93	0
WB	hold		-1.30	(-1.33; -1.28)	-1.56	-1.05	1.85	1.59	.03	-1.36	(-1.39; -1.33)	-1.69	-1.04	2.10	2.05	.24	-1.31	(-1.35; -1.27)	-1.72	-0.89	2.08	2.46	.43	-1.24	(-1.29; -1.20)	-1.74	-0.72	2.07	2.93	.58	-1.14	(-1.20; -1.08)	-1.76	-0.50	2.11	3.62	.69	0
NP			-1.37	(-1.40; -1.33)	-1.71	-1.00	2.20	-	-	-1.35	(-1.39; -1.30)	-1.84	-0.86	2.32	-	-	-1.29	(-1.34; -1.23)	-1.88	-0.72	2.44	-	-	-1.23	(-1.29; -1.16)	-1.94	-0.52	2.59	-	-	-1.18	(-1.26; -1.09)	-2.16	-0.34	3.25	-	-	0.1
PG			-1.39	(-1.42; -1.37)	-1.66	-1.12	2.09	1.61	.08	-1.47	(-1.50; -1.44)	-1.82	-1.11	2.41	2.08	.19	-1.23	(-1.26; -1.19)	-1.66	-0.81	1.89	2.56	.47	-0.88	(-0.93; -0.83)	-1.45	-0.36	1.39	3.18	.70	-0.30	(-0.36; -0.23)	-1.04	0.40	1.21	4.17	.87	7.3

Table 2.4: Results for data generated from Burr distribution, data set size N = 500, for five percentiles. Each row summarizes 1000 runs. *Abbreviations* True infection risk (Inf. risk): constant, increasing (exponential growth) or decreasing (household transmission). Exposure window widths (Width): empirical set (emp.) or its doubled (db.) or squared (sq.) version. Estimation method (Mod.): MLE assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP), penalized Gaussian mixture (PG). Summary measures: bias with (normal approximated) 95% CI, the 25^{th} and 75^{th} percentiles of the deviations over all runs (p25 and p75), Mean Squared Error (MSE), mean length of the CI (LCI), coverage proportion. Percentage of (1000) runs that timed out (TO). ≈ 0 means a value between -0.005 and 0.005.

Mod.	Inf. risk	Width	50%							90%							95%							97.5%							99%							TO
			Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	%
GA			0.13	(0.12; 0.14)	-0.01	0.26	0.06	0.84	.91	0.45	(0.41; 0.49)	0.04	0.88	0.58	1.90	.77	-0.05	(-0.10; <0.01)	-0.60	0.53	0.70	2.49	.86	-1.07	(-1.14; -1.01)	-1.76	-0.35	2.27	3.12	.68	-3.41	(-3.50; -3.33)	-4.28	-2.49	13.49	3.97	.19	0.3
LN	con-		-0.28	(-0.29; -0.27)	-0.42	-0.15	0.12	0.75	.70	0.49	(0.45; 0.52)	0.11	0.88	0.56	2.38	.88	1.01	(0.95; 1.06)	0.44	1.62	1.76	3.57	.79	1.50	(1.42; 1.58)	0.67	2.34	3.80	5.07	.77	1.96	(1.84; 2.07)	0.70	3.22	7.37	7.56	.82	0
WB	stant		0.14	(0.12; 0.15)	-0.01	0.28	0.07	0.94	.95	0.82	(0.78; 0.86)	0.35	1.31	1.17	1.88	.57	0.27	(0.21; 0.33)	-0.41	0.90	1.08	2.44	.77	-0.89	(-0.97; -0.81)	-1.78	-0.04	2.55	3.05	.63	-3.52	(-3.63; -3.41)	-4.71	-2.44	15.49	3.90	.20	0.4
NP			-0.01	(-0.04; 0.01)	-0.24	0.21	0.11	-	-	-0.04	(-0.09; 0.02)	-0.61	0.48	0.69	-	-	-0.07	(-0.15; <0.01)	-0.88	0.72	1.51	-	-	-0.11	(-0.24; 0.01)	-1.59	1.17	4.14	-	-	-0.18	(-0.41; 0.05)	-2.97	1.95	13.90	-	-	0
PG			-0.08	(-0.09; -0.06)	-0.23	0.07	0.06	0.86	.92	0.13	(0.10; 0.17)	-0.23	0.52	0.35	2.26	.94	0.19	(0.13; 0.24)	-0.42	0.78	0.80	3.44	.94	0.12	(0.04; 0.20)	-0.76	0.99	1.73	5.12	.93	-0.28	(-0.41; -0.15)	-1.72	1.14	4.62	7.10	.90	12.3
GA			0.72	(0.71; 0.74)	0.57	0.87	0.57	0.87	.08	1.32	(1.28; 1.36)	0.91	1.76	2.15	1.98	.26	0.89	(0.84; 0.94)	0.34	1.47	1.52	2.58	.67	-0.07	(-0.14; -0.01)	-0.77	0.64	1.15	3.20	.87	-2.35	(-2.43; -2.26)	-3.23	-1.41	7.37	4.05	.42	0.4
LN	exp.		0.26	(0.24; 0.27)	0.11	0.40	0.11	0.80	.74	1.47	(1.43; 1.51)	1.07	1.89	2.52	2.51	.29	2.15	(2.10; 2.21)	1.54	2.76	5.45	3.72	.28	2.79	(2.71; 2.87)	1.94	3.63	9.46	5.22	.36	3.43	(3.31; 3.55)	2.14	4.62	15.46	7.69	.51	0
WB	growth	emp.	0.80	(0.78; 0.81)	0.65	0.96	0.69	0.98	.09	1.72	(1.68; 1.77)	1.22	2.21	3.52	1.93	.12	1.19	(1.12; 1.25)	0.48	1.84	2.52	2.48	.52	0.02	(-0.07; 0.10)	-0.88	0.87	1.88	3.06	.75	-2.65	(-2.76; -2.54)	-3.84	-1.57	10.24	3.88	.32	0.4
NP			0.57	(0.54; 0.59)	0.32	0.81	0.45	-	-	0.81	(0.75; 0.86)	0.22	1.39	1.41	-	-	0.87	(0.78; 0.95)	-0.07	1.74	2.69	-	-	0.88	(0.75; 1.01)	-0.61	2.14	5.04	-	-	0.66	(0.43; 0.90)	-2.10	3.04	14.56	-	-	0
PG			0.50	(0.49; 0.52)	0.34	0.66	0.31	0.90	.42	1.03	(0.99; 1.06)	0.62	1.41	1.41	2.35	.60	1.16	(1.10; 1.21)	0.50	1.76	2.15	3.54	.79	1.15	(1.07; 1.24)	0.23	2.01	3.10	5.24	.88	0.75	(0.62; 0.88)	-0.72	2.21	4.83	6.64	.92	15.1
GA			-0.72	(-0.73; -0.71)	-0.85	-0.59	0.56	0.80	.07	-0.70	(-0.73; -0.66)	-1.08	-0.26	0.83	1.82	.65	-1.25	(-1.30; -1.20)	-1.77	-0.67	2.22	2.42	.48	-2.30	(-2.37; -2.24)	-2.96	-1.58	6.40	3.06	.28	-4.67	(-4.75; -4.58)	-5.55	-3.72	23.62	3.94	.05	0.4
LN	house-		-1.04	(-1.05; -1.03)	-1.17	-0.92	1.12	0.70	< .01	-0.79	(-0.82; -0.75)	-1.14	-0.41	0.91	2.25	.75	-0.41	(-0.46; -0.36)	-0.94	0.22	0.86	3.45	.92	-0.01	(-0.08; 0.07)	-0.75	0.90	1.47	4.98	.96	0.40	(0.29; 0.52)	-0.77	1.74	3.66	7.59	.95	0
WB	hold		-0.78	(-0.80; -0.77)	-0.93	-0.64	0.66	0.88	.06	-0.39	(-0.43; -0.35)	-0.82	0.05	0.58	1.83	.76	-0.93	(-0.99; -0.87)	-1.58	-0.31	1.78	2.42	.59	-2.03	(-2.11; -1.95)	-2.88	-1.24	5.78	3.08	.34	-4.53	(-4.64; -4.42)	-5.66	-3.48	23.63	4.03	.11	0.1
NP			-0.81	(-0.83; -0.79)	-1.01	-0.59	0.75	-	-	-1.19	(-1.23; -1.14)	-1.72	-0.68	1.98	-	-	-1.25	(-1.32; -1.18)	-2.02	-0.49	2.88	-	-	-1.37	(-1.49; -1.25)	-2.66	-0.27	5.37	-	-	-1.34	(-1.57; -1.11)	-3.82	0.76	15.29	-	-	0
PG			-0.88	(-0.90; -0.87)	-1.02	-0.76	0.82	0.80	.02	-1.02	(-1.05; -0.98)	-1.36	-0.67	1.32	2.16	.55	-1.03	(-1.08; -0.98)	-1.58	-0.43	1.75	3.33	.74	-1.12	(-1.20; -1.04)	-1.98	-0.28	2.87	5.00	.82	-1.48	(-1.61; -1.35)	-2.90	-0.12	6.79	7.47	.83	8.2
GA			0.12	(0.11; 0.14)	-0.03	0.28	0.07	0.96	.93	0.53	(0.48; 0.57)	0.07	1.02	0.75	2.12	.77	0.07	(0.01; 0.12)	-0.56	0.70	0.88	2.79	.86	-0.92	(-0.99; -0.84)	-1.69	-0.12	2.27	3.51	.75	-3.20	(-3.30; -3.11)	-4.22	-2.19	12.65	4.49	.32	0.4
LN	con-		-0.28	(-0.29; -0.26)	-0.43	-0.12	0.13	0.85	.75	0.39	(0.35; 0.42)	-0.01	0.81	0.53	2.55	.92	0.84	(0.78; 0.90)	0.22	1.44	1.58	3.82	.85	1.25	(1.17; 1.33)	0.35	2.13	3.36	5.43	.84	1.57	(1.44; 1.69)	0.24	2.93	6.58	8.11	.88	0
WB	stant		0.11	(0.09; 0.12)	-0.05	0.29	0.08	1.07	.94	0.90	(0.86; 0.95)	0.38	1.39	1.43	2.10	.59	0.41	(0.34; 0.48)	-0.34	1.09	1.46	2.74	.78	-0.69	(-0.78; -0.59)	-1.68	0.19	2.76	3.44	.70	-3.23	(-3.35; -3.10)	-4.56	-2.11	14.55	4.44	.30	0.1
NP			-0.01	(-0.03; 0.02)	-0.28	0.27	0.17	-	-	-0.06	(-0.12; <0.01)	-0.78	0.63	1.00	-	-	-0.09	(-0.18; <0.01)	-1.05	0.84	2.07	-	-	-0.21	(-0.36; -0.07)	-1.88	1.19	5.36	-	-	-0.40	(-0.66; -0.14)	-3.50	2.11	17.56	-	-	0
PG			-0.09	(-0.11; -0.08)	-0.26	0.06	0.07	0.96	.92	0.14	(0.11; 0.18)	-0.28	0.56	0.41	2.44	.93	0.22	(0.16; 0.27)	-0.42	0.86	0.95	3.68	.94	0.17	(0.08; 0.26)	-0.81	1.06	2.05	5.44	.92	-0.21	(-0.35; -0.07)	-1.74	1.27	5.23	7.33	.90	10.4
GA			1.39	(1.37; 1.41)	1.22	1.57	2.00	1.04	< .01	2.63	(2.58; 2.67)	2.12	3.12	7.50	2.42	.01	2.43	(2.37; 2.50)	1.72	3.11	6.99	3.15	.16	1.69	(1.61; 1.77)	0.83	2.55	4.56	3.92	.57	-0.30	(-0.41; -0.20)	-1.41	0.78	2.86	4.96	.86	0.3
LN	exp.		0.86	(0.84; 0.87)	0.69	1.03	0.80	0.95	.05	2.76	(2.71; 2.80)	2.28	3.24	8.11	3.00	.01	3.75	(3.68; 3.82)	3.03	4.44	15.19	4.44	.03	4.72	(4.62; 4.81)	3.73	5.67	24.54	6.22	.07	5.81	(5.67; 5.95)	4.36	7.25	38.88	9.14	.18	0
WB	growth	db.	1.52	(1.50; 1.54)	1.34	1.71	2.40	1.18	< .01	3.05	(2.99; 3.11)	2.42	3.60	10.17	2.34	.01	2.69	(2.62; 2.78)	1.82	3.44	8.96	3.00	.13	1.68	(1.58; 1.79)	0.60	2.61	5.69	3.72	.56	-0.80	(-0.93; -0.66)	-2.23	0.42	5.50	4.70	.68	0.2
NP			1.22	(1.19; 1.25)	0.90	1.53	1.71	-	-	2.07	(1.99; 2.14)	1.21	2.86	5.64	-	-	2.32	(2.21; 2.43)	1.10	3.43	8.53	-	-	2.46	(2.29; 2.62)	0.53	4.15	13.20	-	-	2.50	(2.22; 2.79)	-0.83	5.37	27.34	-	-	0
PG			1.12	(1.10; 1.14)	0.94	1.30	1.33	1.08	.01	2.31	(2.26; 2.35)	1.83	2.79	5.84	2.81	.10	2.70	(2.63; 2.77)	1.91	3.45	8.47	4.18	.27	2.94	(2.84; 3.04)	1.87	3.96	11.21	6.00	.48	2.46	(2.34; 2.58)	1.12	4.51	9.70	5.72	.78	14.3
GA			-1.46	(-1.48; -1.45)	-1.60	-1.32	2.18	0.86	< .01	-1.82	(-1.86; -1.79)	-2.23	-1.42	3.68	1.83	.10	-2.47	(-2.52; -2.42)	-3.04	-1.90	6.82	2.48	.11	-3.60	(-3.67; -3.53)	-4.36	-2.87	14.20	3.19	.07	-6.04	(-6.13; -5.95)	-7.05	-5.12	38.67	4.18	.02	0.3
LN	house-		-1.66	(-1.67; -1.64)	-1.79	-1.53	2.78	0.73	< .01	-2.17	(-2.20; -2.14)	-2.54	-1.80	4.98	2.13	.04	-2.13	(-2.18; -2.08)	-2.71	-1.53	5.21	3.31	.37	-2.09	(-2.16; -2.02)	-2.96	-1.20	5.83	4.86	.64	-2.18	(-2.30; -2.06)	-3.50	-0.85	8.33	7.55	.81	0
WB	hold		-1.58	(-1.60; -1.57)	-1.75	-1.43	2.56	0.92	< .01	-1.61	(-1.65; -1.57)	-2.04	-1.21	2.97	1.88	.18	-2.18	(-2.24; -2.12)	-2.81	-1.59	5.64	2.53	.19	-3.27	(-3.35; -3.19)	-4.15	-2.51	12.40	3.30	.15	-5.71	(-5.82; -5.59)	-6.94	-4.68	35.95	4.44	.07	0
NP			-1.51	(-1.53; -1.49)	-1.77	-1.24	2.42	-	-	-2.36	(-2.40; -2.31)	-2.91	-1.87	6.14	-	-	-2.69	(-2.77; -2.62)	-3.50	-1.92	8.67	-	-	-3.05	(-3.16; -2.93)	-4.33	-1.91	12.74	-	-	-3.33	(-3.56; -3.10)	-6.06	-1.25	25.14	-	-	0
PG			-1.57	(-1.58; -1.55)	-1.71	-1.43	2.50	0.83	< .01	-2.25	(-2.28; -2.21)	-2.62	-1.89	5.32	2.10	.03	-2.44	(-2.49; -2.39)	-3.00	-1.88	6.61	3.24	.23	-2.69	(-2.77; -2.62)	-3.51	-1.84	8.79	4.88	.46	-3.24	(-3.37; -3.11)	-4.62	-1.93	14.85	7.76	.61	1.7
GA			0.16	(0.14; 0.17)	-0.05	0.36	0.10	1.12	.92	0.48	(0.43; 0.54)	-0.06	1.03	0.91	2.52	.81	-0.01	(-0.08; 0.06)	-0.76	0.70	1.23	3.30	.87	-1.03	(-1.12; -0.95)	-1.97	-0.14	3.05	4.12	.75	-3.37	(-3.48; -3.26)	-4.62	-2.23	14.64	5.25	.35	0.3
LN	con-		-0.31	(-0.32; -0.29)	-0.50	-0.11	0.17	1.00	.75	0.53	(0.48; 0.57)	0.02	1.03	0.83	3.13	.92	1.10	(1.03; 1.17)	0.32	1.84	2.51	4.68	.85	1.67	(1.57; 1.77)	0.57	2.73	5.51	6.65	.83	2.26	(2.10; 2.41)	0.57	3.84	11.40	9.92	.84	0
WB	stant		0.18	(0.16; 0.19)	-0.04	0.38	0.12	1.26	.95	0.86	(0.80; 0.92)	0.22	1.43	1.61	2.50	.67	0.30	(0.22; 0.39)	-0.59	1.06	1.91	3.23	.79	-0.86	(-0.97; -0.75)	-2.02	0.10	3.91	4.04	.68	-3.49	(-3.63; -3.34)	-5.01	-2.21	17.81	5.17	.33	0
NP			≈ 0	(-0.03; 0.02)	-0.26	0.25	0.15	-	-	-0.03	(-0.09; 0.04)	-0.74	0.66	1.08	-	-	-0.11	(-0.20; -0.01)	-1.20	0.92	2.52	-	-	-0.06	(-0.22; 0.10)	-1.87	1.45	6.80	-	-	-0.06	(-0.37; 0.25)	-3.69	2.56	25.21	-	-	0
PG			-0.09	(-0.10; -0.07)	-0.29	0.11	0.09	1.13	.93	0.16	(0.11; 0.20)	-0.33	0.65	0.58	2.92	.93	0.22	(0.15; 0.29)	-0.53	0.98	1.34	4.38	.93	0.16	(0.05; 0.26)	-1.00	1.23	2.84	6.39	.92	-0.30	(-0.46; -0.14)	-2.20	1.39	6.73	8.09	.90	9.7
GA			2.34	(2.32; 2.37)	2.10	2.58	5.62	1.36	< .01	5.10	(5.03; 5.17)	4.32	5.81	27.34	3.62	< .01	5.51	(5.41; 5.60)	4.47	6.46	32.68	4.74	< .01	5.36	(5.24; 5.47)	4.06	6.52	32.33	5.88	.05	4.13	(3.98; 4.28)	2.51	5.58	22.88	7.42	.38	0
LN	exp.		1.69	(1.67; 1.71)	1.48	1.92	2.96	1.30	< .01	6.33	(6.25; 6.40)	5.52	7.09	41.47	5.00	< .01	8.94	(8.82; 9.05)	7.66	10.1	83.29	7.58	< .01	11.83	(11.66; 11.99)	10.01	13.51	147.15	10.82	< .01	15.99	(15.74; 16.24)	13.24	18.61	272.49	16.23	< .01	0
WB	growth	sq.	2.59	(2.57; 2.61)	2.35	2.84	6.85	1.52	< .01	5.43	(5.34; 5.52)	4.52	6.23	31.52	3.48	< .01	5.54	(5.42; 5.66)	4.26	6.58	34.45	4.51	.01	4.96	(4.80; 5.11)	3.36	6.26	30.69	5.59	.12	3.00	(2.80; 3.20)	0.95	4.62	19.18	7.04	.58	0
NP			2.02	(1.99; 2.05)	1.67	2.34	4.33	-	-	5.06	(4.94; 5.18)	3.80	6.07	29.06	-	-	6.83	(6.63; 7.02)	4.58	8.73	56.56	-	-	8.39	(8.06; 8.72)	4.52	11.15	98.80	-	-	9.84	(9.28; 10.40)	3.42	14.24	177.88	-	-	0
PG			1.92	(1.89; 1.94)	1.69	2.13	3.79	1.38	< .01	4.96	(4.90; 5.03)	4.29	5.68	25.64	4.49	< .01	6.35	(6.25; 6.44)	5.29	7.43	42.72	6.82	.01	7.56	(7.43; 7.69)	6.17	9.21	61.48	6.96	.04	4.34	(4.29; 4.38)	4.51	4.51	19.40	2.40	.17	26.4
GA			-1.93	(-1.94; -1.92)	-2.08	-1.79	3.77	0.89	< .01	-2.93	(-2.96; -2.90)	-3.29	-2.57	8.88	1.71	< .01	-3.83	(-3.87; -3.78)	-4.34	-3.34	15.22	2.32	< .01	-5.20	(-5.26; -5.13)	-5.87	-4.57	28.01	2.99	< .01	-7.94	(-8.03; -7.86)	-8.86	-7.10	64.94	3.94	< .01	0.3
LN	house-		-2.07	(-2.08; -2.06)	-2.21	-1.94	4.32	0.73	< .01	-3.31	(-3.34; -3.28)	-3.64	-2.98	11.19	1.95	< .01	-3.65	(-3.70; -3.61)	-4.15	-3.16	13.90	3.01	.01	-4.05	(-4.11; -3.98)	-4.77	-3.33	17.56	4.42	.13	-4.79	(-4.90; -4.69)	-5.94	-3.72	25.95	6.89	.31	0
WB	hold		-2.07	(-2.09; -2.06)	-2.22	-1.92	4.35	0.92	< .01	-2.77	(-2.81; -2.74)	-3.13	-2.42	7.99	1.77	< .01	-3.58	(-3.63; -3.53)	-4.15	-3.07	13.52	2.37	.02	-4.89	(-4.97; -4.82)	-5.67	-4.18	25.30	3.09	.02	-7.60	(-7.70; -7.50)	-8.73	-6.62	60.51	4.19	.01	0
NP			-1.91	(-1.93; -1.89)	-2.12	-1.72	3.75	-	-	-3.46	(-3.50; -3.43)	-3.88	-3.05	12.40	-	-	-4.13	(-4.19; -4.07)	-4.82	-3.45	18.02	-	-	-4.84	(-4.93; -4.74)	-5.88	-3.91	25.59	-	-	-5.58	(-5.78; -5.39)	-7.78	-3.79	41.06	-	-	0
PG			-1.99	(-2.00; -1.97)	-2.12	-1.86	3.98	0.84	< .01	-3.36	(-3.39; -3.33)	-3.68	-3.04	11.51	1.93	< .01	-3.90	(-3.95; -3.86)	-4.38	-3.39	15.78	2.94	.01	-4.56	(-4.63; -4.49)	-5.31	-3.83	22.03	4.41	.08	-5.73	(-5.84; -5.62)	-6.98	-4.59	36.24	7.26	.24	2.1

Table 2.5: Results for data generated from lognormal distribution, data set size N = 500, for five percentiles. Each row summarizes 1000 runs. *Abbreviations* True infection risk (Inf. risk): constant, increasing (exponential growth) or decreasing (household transmission). Exposure window widths (Width): empirical set (emp.) or its doubled (db.) or squared (sq.) version. Estimation method (Mod.): MLE assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP), penalized Gaussian mixture (PG). Summary measures: bias with (normal approximated) 95% CI, the 25^{th} and 75^{th} percentiles of the deviations over all runs (p25 and p75), Mean Squared Error (MSE), mean length of the CI (LCI), coverage proportion. Percentage of (1000) runs that timed out (TO). ≈ 0 means a value between -0.005 and 0.005.

Mod.	Inf. risk	Width	50%							90%						95%						97.5%						99%						TO
			Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	Bias	(95% CI)	p25	p75	MSE	LCI	Cov	%
GA			0.15	(0.14; 0.16)	0.07	0.23	0.04	0.49	.78	≈ 0	(-0.02; 0.01)	-0.19	0.17	0.07	0.91	.91	-0.23	(-0.26; -0.21)	-0.46	-0.01	0.17	1.14	.83	-0.54	(-0.57; -0.52)	-0.82	-0.26	0.46	1.38	.64	-1.06	(-1.09; -1.03)	-1.40	-0.72	1.38	1.69	.36	0
LN	con-		≈ 0	(-0.01; 0.01)	-0.08	0.08	0.01	0.47	.95	≈ 0	(-0.02; 0.01)	-0.18	0.18	0.07	1.03	.95	-0.01	(-0.03; 0.01)	-0.25	0.24	0.13	1.38	.95	-0.01	(-0.04; 0.02)	-0.33	0.30	0.21	1.78	.95	-0.01	(-0.05; 0.03)	-0.43	0.40	0.37	2.37	.94	0
WB	stant		0.28	(0.27; 0.29)	0.19	0.37	0.10	0.58	.62	0.23	(0.21; 0.25)	0.02	0.42	0.14	0.84	.75	-0.14	(-0.17; -0.12)	-0.41	0.10	0.17	1.00	.76	-0.64	(-0.67; -0.61)	-0.97	-0.35	0.64	1.18	.45	-1.48	(-1.51; -1.44)	-1.88	-1.11	2.51	1.41	.11	0
NP			-0.01	(-0.02; 0.01)	-0.15	0.14	0.05	-	-	-0.01	(-0.04; 0.02)	-0.33	0.29	0.20	-	-	-0.05	(-0.09; -0.02)	-0.47	0.33	0.35	-	-	-0.07	(-0.12; -0.02)	-0.71	0.50	0.70	-	-	-0.22	(-0.30; -0.14)	-1.13	0.56	1.74	-	-	0
PG			≈ 0	(-0.01; 0.01)	-0.09	0.09	0.02	0.52	.95	-0.01	(-0.03; 0.01)	-0.19	0.17	0.07	1.07	.94	-0.02	(-0.04; <0.01)	-0.28	0.22	0.13	1.45	.94	-0.04	(-0.06; <0.01)	-0.38	0.29	0.23	1.91	.93	-0.05	(-0.10; -0.01)	-0.54	0.41	0.47	2.68	.92	0.3
GA			0.52	(0.51; 0.53)	0.43	0.60	0.28	0.51	.01	0.52	(0.50; 0.54)	0.33	0.7	0.35	0.98	.43	0.33	(0.31; 0.36)	0.10	0.56	0.24	1.22	.78	0.07	(0.04; 0.10)	-0.22	0.35	0.19	1.47	.91	-0.40	(-0.43; -0.36)	-0.73	-0.06	0.43	1.80	.83	0
LN	exp.		0.35	(0.34; 0.36)	0.27	0.44	0.14	0.49	.17	0.56	(0.55; 0.58)	0.37	0.75	0.40	1.11	.45	0.64	(0.62; 0.67)	0.39	0.89	0.56	1.49	.59	0.72	(0.69; 0.75)	0.40	1.04	0.76	1.92	.67	0.83	(0.79; 0.87)	0.40	1.23	1.10	2.54	.74	0
WB	growth	emp.	0.70	(0.70; 0.71)	0.61	0.79	0.51	0.59	< .01	0.73	(0.71; 0.75)	0.52	0.94	0.64	0.88	.17	0.37	(0.34; 0.39)	0.09	0.63	0.30	1.05	.67	-0.13	(-0.16; -0.10)	-0.47	0.18	0.26	1.23	.76	-0.96	(-1.00; -0.92)	-1.37	-0.60	1.28	1.47	.34	0
NP			0.37	(0.35; 0.38)	0.20	0.52	0.20	-	-	0.54	(0.51; 0.58)	0.20	0.88	0.56	-	-	0.55	(0.50; 0.59)	0.09	0.96	0.75	-	-	0.56	(0.50; 0.62)	-0.14	1.14	1.22	-	-	0.44	(0.36; 0.53)	-0.54	1.28	2.08	-	-	0
PG			0.37	(0.36; 0.38)	0.27	0.46	0.15	0.54	.23	0.55	(0.53; 0.56)	0.35	0.73	0.38	1.14	.55	0.59	(0.57; 0.62)	0.33	0.84	0.50	1.54	.70	0.62	(0.59; 0.66)	0.28	0.94	0.66	2.03	.81	0.66	(0.62; 0.70)	0.16	1.11	0.96	2.82	.88	0.3
GA			-0.38	(-0.39; -0.37)	-0.46	-0.31	0.16	0.47	.12	-0.69	(-0.71; -0.68)	-0.86	-0.52	0.54	0.83	.16	-0.97	(-0.99; -0.95)	-1.18	-0.76	1.04	1.05	.12	-1.31	(-1.34; -1.29)	-1.58	-1.06	1.88	1.28	.07	-1.88	(-1.91; -1.85)	-2.21	-1.57	3.75	1.59	.03	0
LN	house-		-0.50	(-0.51; -0.49)	-0.58	-0.43	0.26	0.44	.02	-0.74	(-0.76; -0.73)	-0.90	-0.58	0.61	0.93	.16	-0.83	(-0.85; -0.81)	-1.05	-0.62	0.79	1.26	.32	-0.91	(-0.93; -0.88)	-1.20	-0.64	1.00	1.63	.45	-1.01	(-1.04; -0.97)	-1.40	-0.64	1.33	2.19	.60	0
WB	hold		-0.33	(-0.34; -0.32)	-0.42	-0.25	0.13	0.56	.26	-0.45	(-0.47; -0.44)	-0.64	-0.27	0.28	0.79	.42	-0.82	(-0.84; -0.80)	-1.06	-0.59	0.80	0.95	.17	-1.31	(-1.34; -1.28)	-1.62	-1.03	1.91	1.13	.07	-2.12	(-2.15; -2.09)	-2.51	-1.78	4.80	1.37	.01	0
NP			-0.52	(-0.54; -0.51)	-0.67	-0.36	0.32	-	-	-0.73	(-0.76; -0.71)	-1.02	-0.47	0.69	-	-	-0.80	(-0.83; -0.76)	-1.18	-0.46	0.94	-	-	-0.85	(-0.90; -0.81)	-1.37	-0.38	1.27	-	-	-0.98	(-1.06; -0.91)	-1.80	-0.30	2.29	-	-	0
PG			-0.52	(-0.52; -0.51)	-0.60	-0.43	0.28	0.48	.02	-0.74	(-0.76; -0.73)	-0.91	-0.58	0.62	0.99	.18	-0.81	(-0.83; -0.79)	-1.04	-0.59	0.77	1.34	.35	-0.87	(-0.90; -0.84)	-1.18	-0.57	0.96	1.80	.52	-0.93	(-0.97; -0.89)	-1.40	-0.52	1.29	2.56	.65	0.1
GA			0.15	(0.14; 0.16)	0.06	0.24	0.04	0.57	.82	0.01	(-0.01; 0.03)	-0.21	0.21	0.10	1.05	.91	-0.22	(-0.25; -0.20)	-0.50	0.04	0.21	1.32	.85	-0.53	(-0.56; -0.50)	-0.86	-0.21	0.51	1.60	.72	-1.04	(-1.08; -1.01)	-1.45	-0.65	1.44	1.98	.48	0
LN	con-		≈ 0	(<0.01; 0.01)	-0.09	0.09	0.02	0.54	.94	-0.01	(-0.03; <0.01)	-0.23	0.19	0.09	1.18	.95	-0.02	(-0.05; <0.01)	-0.30	0.26	0.17	1.58	.94	-0.03	(-0.06; <0.01)	-0.39	0.33	0.28	2.04	.94	-0.04	(-0.08; <0.01)	-0.52	0.44	0.50	2.73	.94	0
WB	stant		0.28	(0.28; 0.29)	0.18	0.38	0.10	0.68	.72	0.24	(0.22; 0.26)	≈ 0	0.47	0.18	0.97	.77	-0.13	(-0.16; -0.10)	-0.44	0.15	0.22	1.17	.78	-0.63	(-0.67; -0.60)	-1.00	-0.31	0.70	1.38	.52	-1.46	(-1.50; -1.42)	-1.91	-1.07	2.58	1.66	.18	0
NP			≈ 0	(-0.02; 0.02)	-0.21	0.20	0.09	-	-	-0.05	(-0.08; -0.01)	-0.45	0.29	0.32	-	-	-0.06	(-0.11; -0.01)	-0.58	0.40	0.58	-	-	-0.14	(-0.20; -0.08)	-0.88	0.52	1.08	-	-	-0.36	(-0.46; -0.27)	-1.41	0.48	2.38	-	-	0
PG			≈ 0	(-0.01; 0.01)	-0.10	0.10	0.02	0.59	.94	-0.02	(-0.04; <0.01)	-0.24	0.19	0.10	1.23	.94	-0.04	(-0.07; -0.01)	-0.34	0.25	0.18	1.65	.93	-0.06	(-0.10; -0.03)	-0.46	0.30	0.31	2.16	.92	-0.10	(-0.14; -0.05)	-0.66	0.42	0.62	2.99	.91	0.4
GA			0.86	(0.85; 0.87)	0.75	0.96	0.76	0.61	< .01	1.06	(1.03; 1.08)	0.82	1.28	1.24	1.23	.07	0.94	(0.91; 0.97)	0.64	1.23	1.08	1.54	.31	0.74	(0.71; 0.77)	0.37	1.09	0.83	1.85	.63	0.36	(0.31; 0.40)	-0.08	0.78	0.55	2.26	.88	0
LN	exp.		0.68	(0.67; 0.69)	0.58	0.77	0.48	0.59	< .01	1.14	(1.12; 1.17)	0.90	1.38	1.44	1.41	.07	1.33	(1.30; 1.36)	1.01	1.65	2.00	1.90	.15	1.52	(1.48; 1.55)	1.12	1.92	2.68	2.44	.22	1.77	(1.72; 1.82)	1.23	2.28	3.78	3.24	.34	0
WB	growth	db.	1.09	(1.08; 1.10)	0.98	1.20	1.23	0.70	< .01	1.23	(1.20; 1.25)	0.96	1.47	1.67	1.09	.03	0.89	(0.86; 0.92)	0.54	1.20	1.05	1.31	.31	0.41	(0.37; 0.45)	-0.01	0.78	0.54	1.53	.73	-0.41	(-0.45; -0.36)	-0.90	0.04	0.70	1.82	.69	0
NP			0.71	(0.69; 0.73)	0.47	0.93	0.61	-	-	1.10	(1.06; 1.14)	0.65	1.50	1.63	-	-	1.16	(1.11; 1.22)	0.50	1.74	2.20	-	-	1.14	(1.07; 1.22)	0.32	1.84	2.83	-	-	0.97	(0.86; 1.08)	-0.24	1.95	4.14	-	-	0
PG			0.70	(0.69; 0.71)	0.59	0.80	0.52	0.66	.01	1.11	(1.09; 1.14)	0.87	1.35	1.37	1.41	.10	1.24	(1.21; 1.27)	0.88	1.56	1.77	1.90	.26	1.35	(1.31; 1.39)	0.91	1.77	2.24	2.49	.44	1.48	(1.42; 1.54)	0.88	2.07	3.01	3.41	.61	0.3
GA			-0.76	(-0.77; -0.75)	-0.85	-0.68	0.60	0.54	< .01	-1.20	(-1.21; -1.18)	-1.38	-1.03	1.50	0.87	< .01	-1.51	(-1.53; -1.49)	-1.74	-1.28	2.40	1.10	< .01	-1.89	(-1.92; -1.86)	-2.18	-1.60	3.75	1.36	< .01	-2.49	(-2.52; -2.46)	-2.85	-2.12	6.47	1.70	< .01	0
LN	house-		-0.85	(-0.86; -0.84)	-0.94	-0.77	0.74	0.50	< .01	-1.30	(-1.31; -1.28)	-1.47	-1.14	1.75	0.94	< .01	-1.46	(-1.48; -1.44)	-1.70	-1.24	2.25	1.28	.02	-1.61	(-1.64; -1.59)	-1.92	-1.32	2.80	1.68	.09	-1.81	(-1.85; -1.77)	-2.23	-1.41	3.64	2.29	.22	0
WB	hold		-0.78	(-0.78; -0.77)	-0.86	-0.68	0.62	0.64	< .01	-0.97	(-0.99; -0.95)	-1.16	-0.78	1.01	0.84	.04	-1.34	(-1.36; -1.31)	-1.58	-1.09	1.93	1.03	.02	-1.82	(-1.85; -1.79)	-2.12	-1.50	3.54	1.23	.01	-2.62	(-2.66; -2.58)	-3.02	-2.23	7.22	1.51	.01	0
NP			-0.88	(-0.90; -0.86)	-1.05	-0.69	0.84	-	-	-1.32	(-1.35; -1.29)	-1.64	-1.02	1.94	-	-	-1.46	(-1.49; -1.42)	-1.87	-1.06	2.47	-	-	-1.56	(-1.62; -1.51)	-2.13	-1.08	3.13	-	-	-1.70	(-1.78; -1.63)	-2.63	-0.88	4.38	-	-	0
PG			-0.87	(-0.88; -0.86)	-0.96	-0.78	0.78	0.52	< .01	-1.31	(-1.33; -1.30)	-1.50	-1.15	1.79	1.03	< .01	-1.44	(-1.47; -1.42)	-1.69	-1.22	2.21	1.38	.04	-1.56	(-1.59; -1.53)	-1.89	-1.24	2.66	1.84	.13	-1.70	(-1.74; -1.65)	-2.17	-1.24	3.35	2.62	.32	0.1
GA			0.16	(0.15; 0.17)	0.06	0.26	0.05	0.65	.86	-0.03	(-0.05; -0.01)	-0.28	0.20	0.13	1.20	.90	-0.28	(-0.30; -0.25)	-0.60	0.02	0.28	1.50	.85	-0.60	(-0.63; -0.56)	-0.99	-0.24	0.65	1.81	.70	-1.13	(-1.17; -1.09)	-1.61	-0.69	1.71	2.22	.52	0
LN	con-		≈ 0	(<0.01; 0.01)	-0.10	0.10	0.02	0.62	.94	-0.01	(-0.03; 0.01)	-0.26	0.22	0.12	1.37	.95	-0.02	(-0.05; 0.01)	-0.37	0.29	0.22	1.83	.94	-0.03	(-0.06; 0.01)	-0.47	0.38	0.36	2.35	.95	-0.04	(-0.09; 0.01)	-0.60	0.51	0.64	3.12	.94	0
WB	stant		0.30	(0.29; 0.31)	0.19	0.40	0.12	0.77	.79	0.21	(0.18; 0.23)	-0.06	0.48	0.20	1.11	.78	-0.18	(-0.21; -0.15)	-0.52	0.16	0.29	1.32	.75	-0.69	(-0.73; -0.65)	-1.12	-0.30	0.86	1.55	.53	-1.54	(-1.59; -1.49)	-2.06	-1.08	2.93	1.84	.22	0
NP			≈ 0	(-0.02; 0.01)	-0.17	0.16	0.06	-	-	≈ 0	(-0.04; 0.03)	-0.38	0.35	0.29	-	-	-0.05	(-0.10; -0.01)	-0.56	0.40	0.55	-	-	-0.02	(-0.09; 0.05)	-0.85	0.69	1.27	-	-	-0.12	(-0.22; -0.02)	-1.28	0.80	2.77	-	-	0
PG			0.01	(<0.01; 0.02)	-0.10	0.11	0.03	0.68	.94	-0.03	(-0.05; <0.01)	-0.29	0.22	0.13	1.40	.93	-0.05	(-0.08; -0.02)	-0.40	0.26	0.24	1.88	.93	-0.07	(-0.11; -0.03)	-0.54	0.34	0.42	2.46	.91	-0.10	(-0.16; -0.05)	-0.72	0.46	0.83	3.38	.90	0.3
GA			1.12	(1.11; 1.13)	1.00	1.25	1.30	0.72	< .01	1.59	(1.56; 1.62)	1.28	1.87	2.73	1.55	.01	1.57	(1.54; 1.61)	1.19	1.93	2.78	1.93	.09	1.46	(1.42; 1.50)	1.01	1.90	2.58	2.32	.27	1.20	(1.15; 1.25)	0.66	1.73	2.08	2.83	.59	0
LN	exp.		0.93	(0.92; 0.94)	0.82	1.05	0.90	0.71	< .01	1.87	(1.84; 1.90)	1.54	2.18	3.72	1.88	.01	2.27	(2.23; 2.31)	1.83	2.69	5.56	2.54	.01	2.68	(2.63; 2.73)	2.12	3.22	7.88	3.27	.04	3.25	(3.19; 3.32)	2.52	3.96	11.8	4.35	.07	0
WB	growth	sq.	1.40	(1.39; 1.41)	1.26	1.53	2.00	0.80	< .01	1.69	(1.65; 1.72)	1.34	1.99	3.09	1.35	.01	1.39	(1.35; 1.43)	0.96	1.76	2.32	1.62	.14	0.95	(0.90; 0.99)	0.43	1.38	1.44	1.89	.50	0.17	(0.12; 0.23)	-0.44	0.69	0.80	2.23	.80	0
NP			0.96	(0.94; 0.98)	0.75	1.17	1.01	-	-	1.88	(1.83; 1.93)	1.32	2.35	4.14	-	-	2.19	(2.12; 2.27)	1.36	2.85	6.29	-	-	2.38	(2.28; 2.49)	1.18	3.28	8.63	-	-	2.27	(2.12; 2.43)	0.51	3.65	11.45	-	-	0
PG			0.95	(0.94; 0.97)	0.82	1.07	0.95	0.78	< .01	1.82	(1.79; 1.85)	1.49	2.13	3.54	1.91	.02	2.15	(2.11; 2.19)	1.69	2.58	5.07	2.60	.06	2.48	(2.43; 2.54)	1.89	3.02	6.98	3.42	.13	2.94	(2.85; 3.02)	2.12	3.70	10.29	4.68	.25	1
GA			-0.90	(-0.91; -0.89)	-0.99	-0.82	0.83	0.56	< .01	-1.45	(-1.47; -1.43)	-1.63	-1.28	2.16	0.86	< .01	-1.80	(-1.82; -1.78)	-2.04	-1.58	3.35	1.08	< .01	-2.22	(-2.24; -2.19)	-2.51	-1.94	5.07	1.33	< .01	-2.86	(-2.89; -2.83)	-3.24	-2.50	8.44	1.67	< .01	0
LN	house-		-0.97	(-0.98; -0.96)	-1.06	-0.89	0.96	0.51	< .01	-1.56	(-1.57; -1.54)	-1.73	-1.40	2.49	0.92	< .01	-1.78	(-1.80; -1.76)	-2.02	-1.56	3.27	1.24	< .01	-2.00	(-2.02; -1.97)	-2.31	-1.69	4.16	1.62	.01	-2.28	(-2.31; -2.24)	-2.73	-1.87	5.51	2.20	.07	0
WB	hold		-0.96	(-0.97; -0.96)	-1.06	-0.86	0.95	0.67	< .01	-1.23	(-1.25; -1.21)	-1.41	-1.05	1.58	0.85	< .01	-1.61	(-1.64; -1.59)	-1.86	-1.38	2.73	1.02	< .01	-2.11	(-2.14; -2.09)	-2.44	-1.81	4.66	1.22	< .01	-2.92	(-2.96; -2.89)	-3.33	-2.55	8.87	1.50	< .01	0
NP			-1.00	(-1.01; -0.98)	-1.13	-0.86	1.04	-	-	-1.57	(-1.60; -1.55)	-1.82	-1.34	2.61	-	-	-1.75	(-1.78; -1.72)	-2.11	-1.42	3.31	-	-	-1.92	(-1.96; -1.88)	-2.38	-1.48	4.12	-	-	-2.04	(-2.11; -1.97)	-2.89	-1.31	5.38	-	-	0
PG			-0.99	(-1.00; -0.98)	-1.07	-0.90	0.99	0.54	< .01	-1.58	(-1.59; -1.56)	-1.75	-1.41	2.56	0.99	< .01	-1.77	(-1.79; -1.75)	-2.01	-1.55	3.26	1.33	< .01	-1.96	(-1.98; -1.93)	-2.28	-1.64	4.02	1.77	.03	-2.18	(-2.22; -2.14)	-2.64	-1.75	5.15	2.50	.14	0.1

Table 2.6: Results for data generated from Weibull distribution, data set size N = 500, for five percentiles. Each row summarizes 1000 runs. *Abbreviations* True infection risk (Inf. risk): constant, increasing (exponential growth) or decreasing (household transmission). Exposure window widths (Width): empirical set (emp.) or its doubled (db.) or squared (sq.) version. Estimation method (Mod.): MLE assuming a gamma (GA), lognormal (LN) or Weibull (WB) distribution, nonparametric MLE (NP), penalized Gaussian mixture (PG). Summary measures: bias with (normal approximated) 95% CI, the 25^{th} and 75^{th} percentiles of the deviations over all runs (p25 and p75), Mean Squared Error (MSE), mean length of the CI (LCI), coverage proportion. Percentage of (1000) runs that timed out (TO). ≈ 0 means a value between -0.005 and 0.005.

Mod.	Inf. risk	Width	50%							90%							95%							97.5%							99%							TO
			Bias	(95% CI)		p25	p75	MSE	LCI	Cov	Bias	(95% CI)		p25	p75	MSE	LCI	Cov	Bias	(95% CI)		p25	p75	MSE	LCI	Cov	Bias	(95% CI)		p25	p75	MSE	LCI	Cov	Bias	(95% CI)		p25	p75	MSE	LCI	Cov	%
GA			-0.23	(-0.24; -0.22)		-0.32	-0.14	0.07	0.52	.60	0.12	(0.11; 0.14)		-0.02	0.26	0.06	1.00	.96	0.45	(0.43; 0.46)		0.26	0.62	0.28	1.26	.74	0.82	(0.80; 0.84)		0.59	1.03	0.79	1.53	.40	1.36	(1.33; 1.39)		1.08	1.63	2.04	1.89	.12	0
LN	con-		-0.42	(-0.43; -0.42)		-0.52	-0.33	0.20	0.51	.13	0.36	(0.34; 0.38)		0.18	0.53	0.20	1.22	.83	1.10	(1.08; 1.12)		0.85	1.33	1.35	1.67	.15	2.00	(1.97; 2.04)		1.67	2.32	4.27	2.19	.01	3.42	(3.38; 3.46)		2.94	3.87	12.21	2.99	< .01	0
WB	stant		≈ 0	(-0.01; 0.01)		-0.08	0.10	0.02	0.57	.96	-0.01	(-0.02; <0.01)		-0.14	0.13	0.04	0.82	.94	-0.01	(-0.03; <0.01)		-0.18	0.15	0.07	1.00	.94	-0.02	(-0.04; <0.01)		-0.22	0.19	0.09	1.18	.95	-0.02	(-0.04; <0.01)		-0.27	0.22	0.14	1.42	.94	0
NP			≈ 0	(-0.01; 0.02)		-0.17	0.18	0.06	-	-	-0.01	(-0.03; 0.01)		-0.26	0.23	0.14	-	-	-0.03	(-0.06; <0.01)		-0.35	0.27	0.21	-	-	-0.07	(-0.11; -0.04)		-0.50	0.30	0.34	-	-	-0.20	(-0.25; -0.15)		-0.68	0.23	0.59	-	-	0
PG			≈ 0	(-0.01; 0.01)		-0.12	0.11	0.03	0.65	.94	-0.05	(-0.07; -0.04)		-0.19	0.08	0.05	0.86	.94	0.02	(0.01; 0.04)		-0.15	0.20	0.07	1.06	.95	0.15	(0.13; 0.17)		-0.07	0.37	0.14	1.29	.93	0.39	(0.36; 0.42)		0.09	0.68	0.36	1.65	.85	32.5
GA			0.17	(0.16; 0.18)		0.08	0.26	0.05	0.54	.76	0.66	(0.65; 0.68)		0.50	0.81	0.49	1.06	.24	1.02	(1.01; 1.04)		0.81	1.23	1.14	1.33	.08	1.43	(1.41; 1.45)		1.19	1.67	2.18	1.61	.02	2.02	(1.99; 2.04)		1.71	2.31	4.28	1.98	< .01	0
LN	exp.		-0.03	(-0.04; -0.02)		-0.12	0.06	0.02	0.53	.93	0.98	(0.96; 1.00)		0.79	1.17	1.04	1.31	.07	1.80	(1.78; 1.83)		1.54	2.06	3.41	1.79	< .01	2.78	(2.75; 2.82)		2.42	3.14	8.05	2.34	< .01	4.30	(4.25; 4.35)		3.77	4.81	19.08	3.16	< .01	0
WB	growth	emp.	0.43	(0.42; 0.44)		0.34	0.52	0.20	0.57	.19	0.44	(0.42; 0.45)		0.29	0.58	0.24	0.86	.51	0.42	(0.41; 0.44)		0.25	0.60	0.25	1.04	.67	0.41	(0.39; 0.43)		0.19	0.62	0.27	1.23	.78	0.38	(0.36; 0.41)		0.12	0.63	0.29	1.46	.85	0
NP			0.43	(0.42; 0.45)		0.26	0.59	0.25	-	-	0.43	(0.41; 0.45)		0.16	0.70	0.34	-	-	0.37	(0.34; 0.40)		0.02	0.68	0.38	-	-	0.33	(0.29; 0.37)		-0.12	0.73	0.52	-	-	0.18	(0.13; 0.23)		-0.39	0.70	0.69	-	-	0
PG			0.43	(0.42; 0.44)		0.32	0.54	0.21	0.66	.28	0.40	(0.39; 0.41)		0.26	0.54	0.21	0.90	.61	0.48	(0.46; 0.49)		0.29	0.66	0.30	1.11	.62	0.60	(0.58; 0.62)		0.36	0.81	0.48	1.34	.59	0.83	(0.80; 0.86)		0.51	1.10	0.92	1.70	.53	41.8
GA			-0.81	(-0.82; -0.81)		-0.91	-0.73	0.68	0.51	< .01	-0.58	(-0.60; -0.57)		-0.73	-0.46	0.38	0.93	.32	-0.29	(-0.31; -0.28)		-0.47	-0.12	0.15	1.18	.87	0.06	(0.04; 0.08)		-0.16	0.27	0.10	1.44	.97	0.58	(0.56; 0.61)		0.32	0.84	0.50	1.80	.78	0
LN	house-		-0.99	(-1.00; -0.98)		-1.08	-0.90	1.00	0.49	< .01	-0.44	(-0.46; -0.43)		-0.60	-0.29	0.25	1.12	.72	0.23	(0.21; 0.25)		0.01	0.45	0.16	1.56	.95	1.08	(1.05; 1.11)		0.77	1.37	1.37	2.07	.39	2.44	(2.39; 2.48)		1.98	2.85	6.37	2.87	.02	0
WB	hold		-0.63	(-0.64; -0.62)		-0.72	-0.53	0.41	0.56	< .01	-0.62	(-0.64; -0.61)		-0.76	-0.50	0.43	0.79	.13	-0.59	(-0.61; -0.58)		-0.75	-0.44	0.41	0.97	.33	-0.56	(-0.58; -0.54)		-0.75	-0.37	0.39	1.16	.52	-0.51	(-0.53; -0.48)		-0.74	-0.27	0.38	1.41	.70	0
NP			-0.65	(-0.66; -0.63)		-0.83	-0.46	0.49	-	-	-0.60	(-0.62; -0.58)		-0.83	-0.38	0.47	-	-	-0.59	(-0.62; -0.57)		-0.87	-0.31	0.53	-	-	-0.60	(-0.64; -0.57)		-0.97	-0.27	0.63	-	-	-0.63	(-0.67; -0.59)		-1.11	-0.22	0.86	-	-	0
PG			-0.64	(-0.65; -0.63)		-0.76	-0.52	0.43	0.64	.02	-0.68	(-0.69; -0.66)		-0.81	-0.55	0.50	0.84	.11	-0.57	(-0.59; -0.56)		-0.73	-0.40	0.39	1.02	.40	-0.41	(-0.43; -0.39)		-0.62	-0.21	0.27	1.25	.69	-0.13	(-0.16; -0.10)		-0.43	0.13	0.21	1.60	.89	19.3
GA			-0.23	(-0.24; -0.22)		-0.34	-0.13	0.08	0.60	.68	0.10	(0.08; 0.12)		-0.06	0.27	0.07	1.15	.97	0.41	(0.39; 0.43)		0.20	0.63	0.27	1.45	.82	0.78	(0.75; 0.80)		0.51	1.03	0.76	1.76	.58	1.31	(1.28; 1.34)		0.97	1.63	1.96	2.18	.26	0
LN	con-		-0.43	(-0.44; -0.42)		-0.53	-0.32	0.21	0.59	.21	0.28	(0.26; 0.29)		0.09	0.46	0.16	1.36	.92	0.97	(0.95; 1.00)		0.70	1.23	1.11	1.87	.38	1.83	(1.79; 1.86)		1.45	2.17	3.66	2.45	.07	3.17	(3.13; 3.22)		2.62	3.66	10.71	3.35	< .01	0
WB	stant		≈ 0	(-0.01; 0.01)		-0.10	0.10	0.03	0.66	.96	-0.01	(-0.03; <0.01)		-0.16	0.15	0.06	0.96	.95	-0.01	(-0.03; <0.01)		-0.20	0.18	0.09	1.17	.94	-0.02	(-0.04; <0.01)		-0.26	0.22	0.12	1.39	.95	-0.02	(-0.05; 0.01)		-0.31	0.26	0.18	1.67	.94	0
NP			-0.01	(-0.03; 0.01)		-0.24	0.20	0.11	-	-	-0.03	(-0.06; <0.01)		-0.35	0.28	0.22	-	-	-0.06	(-0.09; -0.02)		-0.46	0.28	0.31	-	-	-0.14	(-0.18; -0.10)		-0.66	0.31	0.47	-	-	-0.26	(-0.31; -0.21)		-0.83	0.29	0.76	-	-	0
PG			-0.03	(-0.04; -0.01)		-0.16	0.11	0.04	0.75	.95	-0.05	(-0.07; -0.04)		-0.21	0.10	0.06	1.01	.95	0.05	(0.03; 0.07)		-0.15	0.25	0.09	1.25	.95	0.21	(0.18; 0.23)		-0.06	0.46	0.20	1.53	.93	0.50	(0.46; 0.53)		0.12	0.82	0.53	1.94	.85	31.5
GA			0.53	(0.52; 0.54)		0.42	0.63	0.30	0.63	.08	1.16	(1.15; 1.18)		0.97	1.36	1.44	1.31	.02	1.57	(1.54; 1.59)		1.31	1.82	2.60	1.65	.01	2.01	(1.99; 2.04)		1.70	2.32	4.27	2.00	< .01	2.65	(2.61; 2.68)		2.26	3.02	7.34	2.45	< .01	0
LN	exp.		0.33	(0.32; 0.34)		0.22	0.43	0.13	0.63	.46	1.52	(1.50; 1.54)		1.28	1.75	2.44	1.60	< .01	2.41	(2.38; 2.44)		2.05	2.73	6.04	2.19	< .01	3.45	(3.40; 3.49)		2.96	3.90	12.35	2.85	< .01	5.04	(4.98; 5.10)		4.35	5.66	26.28	3.85	< .01	0
WB	growth	db.	0.80	(0.79; 0.81)		0.69	0.91	0.67	0.67	< .01	0.84	(0.83; 0.86)		0.66	1.03	0.79	1.06	.12	0.83	(0.81; 0.85)		0.61	1.06	0.80	1.29	.27	0.81	(0.79; 0.84)		0.56	1.07	0.81	1.51	.47	0.78	(0.75; 0.81)		0.47	1.08	0.83	1.80	.63	0
NP			0.82	(0.80; 0.84)		0.60	1.05	0.78	-	-	0.82	(0.79; 0.85)		0.46	1.16	0.94	-	-	0.71	(0.67; 0.75)		0.26	1.12	0.90	-	-	0.59	(0.54; 0.63)		0.08	1.06	0.89	-	-	0.30	(0.24; 0.36)		-0.40	0.90	0.96	-	-	0
PG			0.80	(0.78; 0.81)		0.66	0.93	0.67	0.76	.02	0.82	(0.80; 0.84)		0.65	0.99	0.74	1.10	.15	0.92	(0.90; 0.94)		0.68	1.13	0.97	1.36	.24	1.07	(1.04; 1.10)		0.76	1.34	1.35	1.67	.29	1.33	(1.29; 1.37)		0.92	1.68	2.16	2.10	.30	45.6
GA			-1.24	(-1.25; -1.23)		-1.35	-1.15	1.57	0.58	< .01	-1.14	(-1.15; -1.13)		-1.28	-1.00	1.35	0.97	< .01	-0.88	(-0.90; -0.86)		-1.06	-0.69	0.85	1.25	.21	-0.56	(-0.58; -0.54)		-0.78	-0.32	0.42	1.54	.76	-0.06	(-0.09; -0.04)		-0.35	0.23	0.19	1.95	.96	0
LN	house-		-1.40	(-1.41; -1.39)		-1.51	-1.30	1.98	0.54	< .01	-1.12	(-1.14; -1.11)		-1.28	-0.97	1.31	1.13	.02	-0.56	(-0.58; -0.54)		-0.77	-0.34	0.42	1.58	.78	0.19	(0.16; 0.22)		-0.12	0.48	0.24	2.12	.97	1.42	(1.38; 1.46)		0.96	1.86	2.43	2.96	.45	0
WB	hold		-1.10	(-1.11; -1.09)		-1.22	-1.00	1.24	0.65	< .01	-1.10	(-1.12; -1.09)		-1.25	-0.97	1.26	0.86	< .01	-1.05	(-1.06; -1.03)		-1.23	-0.87	1.17	1.06	.03	-0.98	(-1.00; -0.96)		-1.20	-0.77	1.06	1.29	.16	-0.89	(-0.91; -0.86)		-1.14	-0.62	0.94	1.60	.44	0
NP			-1.14	(-1.16; -1.12)		-1.37	-0.92	1.41	-	-	-1.11	(-1.14; -1.08)		-1.39	-0.86	1.40	-	-	-1.02	(-1.05; -0.99)		-1.36	-0.70	1.28	-	-	-1.02	(-1.05; -0.98)		-1.44	-0.63	1.37	-	-	-1.05	(-1.10; -1.00)		-1.57	-0.56	1.67	-	-	0.1
PG			-1.13	(-1.14; -1.12)		-1.25	-1.01	1.31	0.72	< .01	-1.17	(-1.19; -1.16)		-1.31	-1.03	1.42	0.91	< .01	-1.04	(-1.05; -1.02)		-1.22	-0.86	1.15	1.11	.06	-0.83	(-0.86; -0.81)		-1.09	-0.60	0.81	1.37	.35	-0.48	(-0.51; -0.45)		-0.83	-0.16	0.47	1.77	.72	9.7
GA			-0.24	(-0.25; -0.23)		-0.35	-0.13	0.09	0.70	.76	0.16	(0.14; 0.17)		-0.02	0.35	0.11	1.34	.97	0.50	(0.47; 0.52)		0.26	0.74	0.38	1.69	.83	0.88	(0.85; 0.91)		0.59	1.18	0.98	2.05	.58	1.45	(1.41; 1.48)		1.07	1.81	2.41	2.53	.30	0
LN	con-		-0.44	(-0.46; -0.43)		-0.56	-0.33	0.23	0.69	.33	0.46	(0.44; 0.48)		0.22	0.69	0.33	1.67	.87	1.27	(1.24; 1.30)		0.90	1.60	1.86	2.30	.31	2.25	(2.20; 2.29)		1.73	2.69	5.54	3.02	.05	3.78	(3.71; 3.84)		3.03	4.44	15.30	4.12	.01	0
WB	stant		0.01	(<0.01; 0.02)		-0.11	0.12	0.03	0.76	.95	-0.01	(-0.03; 0.01)		-0.18	0.17	0.07	1.10	.95	-0.01	(-0.03; 0.01)		-0.23	0.20	0.11	1.33	.94	-0.02	(-0.04; 0.01)		-0.28	0.24	0.15	1.57	.95	-0.02	(-0.05; 0.01)		-0.35	0.29	0.22	1.88	.94	0
NP			≈ 0	(-0.01; 0.02)		-0.18	0.19	0.08	-	-	0.01	(-0.02; 0.03)		-0.25	0.26	0.17	-	-	-0.01	(-0.04; 0.02)		-0.36	0.33	0.25	-	-	-0.06	(-0.10; -0.02)		-0.51	0.36	0.43	-	-	-0.16	(-0.21; -0.10)		-0.76	0.39	0.76	-	-	0
PG			-0.04	(-0.05; -0.02)		-0.17	0.11	0.05	0.85	.94	-0.04	(-0.06; -0.03)		-0.23	0.14	0.07	1.15	.96	0.08	(0.06; 0.10)		-0.17	0.30	0.12	1.43	.96	0.26	(0.23; 0.29)		-0.05	0.55	0.26	1.74	.93	0.58	(0.54; 0.62)		0.16	0.94	0.71	2.20	.85	33.7
GA			0.82	(0.81; 0.83)		0.71	0.94	0.71	0.76	.01	1.83	(1.80; 1.85)		1.57	2.08	3.48	1.69	< .01	2.37	(2.34; 2.40)		2.03	2.69	5.86	2.13	< .01	2.95	(2.91; 2.99)		2.53	3.35	9.06	2.56	< .01	3.75	(3.70; 3.80)		3.23	4.23	14.62	3.13	< .01	0
LN	exp.		0.64	(0.63; 0.65)		0.52	0.76	0.44	0.79	.07	2.62	(2.59; 2.66)		2.26	2.97	7.18	2.27	< .01	3.90	(3.85; 3.95)		3.33	4.40	15.82	3.14	< .01	5.36	(5.30; 5.43)		4.57	6.04	29.95	4.13	< .01	7.57	(7.48; 7.67)		6.43	8.50	59.74	5.62	< .01	0
WB	growth	sq.	1.07	(1.06; 1.08)		0.95	1.20	1.18	0.75	< .01	1.24	(1.22; 1.26)		1.01	1.46	1.64	1.30	.02	1.26	(1.23; 1.28)		0.99	1.51	1.74	1.58	.09	1.26	(1.24; 1.29)		0.95	1.57	1.82	1.84	.20	1.27	(1.23; 1.30)		0.90	1.63	1.91	2.17	.36	0
NP			1.10	(1.08; 1.11)		0.89	1.29	1.29	-	-	1.21	(1.18; 1.24)		0.84	1.53	1.72	-	-	1.16	(1.11; 1.20)		0.67	1.57	1.83	-	-	1.06	(1.00; 1.11)		0.48	1.61	1.89	-	-	0.78	(0.71; 0.84)		0.01	1.34	1.74	-	-	0
PG			1.04	(1.02; 1.05)		0.90	1.18	1.13	0.84	< .01	1.27	(1.25; 1.29)		1.01	1.48	1.72	1.37	.02	1.46	(1.43; 1.49)		1.14	1.73	2.35	1.74	.06	1.70	(1.67; 1.74)		1.27	2.05	3.28	2.15	.09	2.10	(2.04; 2.15)		1.49	2.58	5.11	2.73	.11	46.3
GA			-1.45	(-1.46; -1.44)		-1.56	-1.34	2.12	0.62	< .01	-1.39	(-1.40; -1.37)		-1.53	-1.24	1.97	0.98	< .01	-1.13	(-1.15; -1.12)		-1.32	-0.94	1.36	1.25	.06	-0.81	(-0.84; -0.79)		-1.05	-0.58	0.78	1.55	.50	-0.32	(-0.35; -0.30)		-0.62	-0.02	0.29	1.97	.93	0
LN	house-		-1.60	(-1.61; -1.59)		-1.72	-1.49	2.60	0.57	< .01	-1.39	(-1.40; -1.37)		-1.55	-1.23	1.98	1.14	< .01	-0.83	(-0.85; -0.81)		-1.06	-0.60	0.80	1.59	.52	-0.09	(-0.12; -0.06)		-0.41	0.24	0.22	2.14	.98	1.14	(1.10; 1.19)		0.66	1.61	1.77	3.02	.64	0
WB	hold		-1.33	(-1.34; -1.32)		-1.44	-1.22	1.79	0.70	< .01	-1.33	(-1.34; -1.32)		-1.47	-1.19	1.82	0.88	< .01	-1.26	(-1.28; -1.25)		-1.44	-1.08	1.66	1.08	.01	-1.18	(-1.20; -1.16)		-1.39	-0.96	1.49	1.31	.06	-1.06	(-1.08; -1.03)		-1.32	-0.80	1.27	1.63	.30	0
NP			-1.36	(-1.38; -1.34)		-1.54	-1.18	1.92	-	-	-1.31	(-1.33; -1.29)		-1.53	-1.11	1.82	-	-	-1.23	(-1.25; -1.21)		-1.51	-0.96	1.66	-	-	-1.18	(-1.21; -1.15)		-1.49	-0.88	1.60	-	-	-1.15	(-1.19; -1.11)		-1.60	-0.72	1.72	-	-	0
PG			-1.35	(-1.36; -1.33)		-1.47	-1.23	1.85	0.77	< .01	-1.40	(-1.41; -1.39)		-1.54	-1.24	2.01	0.92	< .01	-1.26	(-1.28; -1.25)		-1.45	-1.08	1.66	1.12	.01	-1.06	(-1.08; -1.03)		-1.29	-0.82	1.23	1.37	.16	-0.69	(-0.72; -0.66)		-1.01	-0.37	0.71	1.77	.59	7.9

2.7.3 Tables and figure Simulation study II

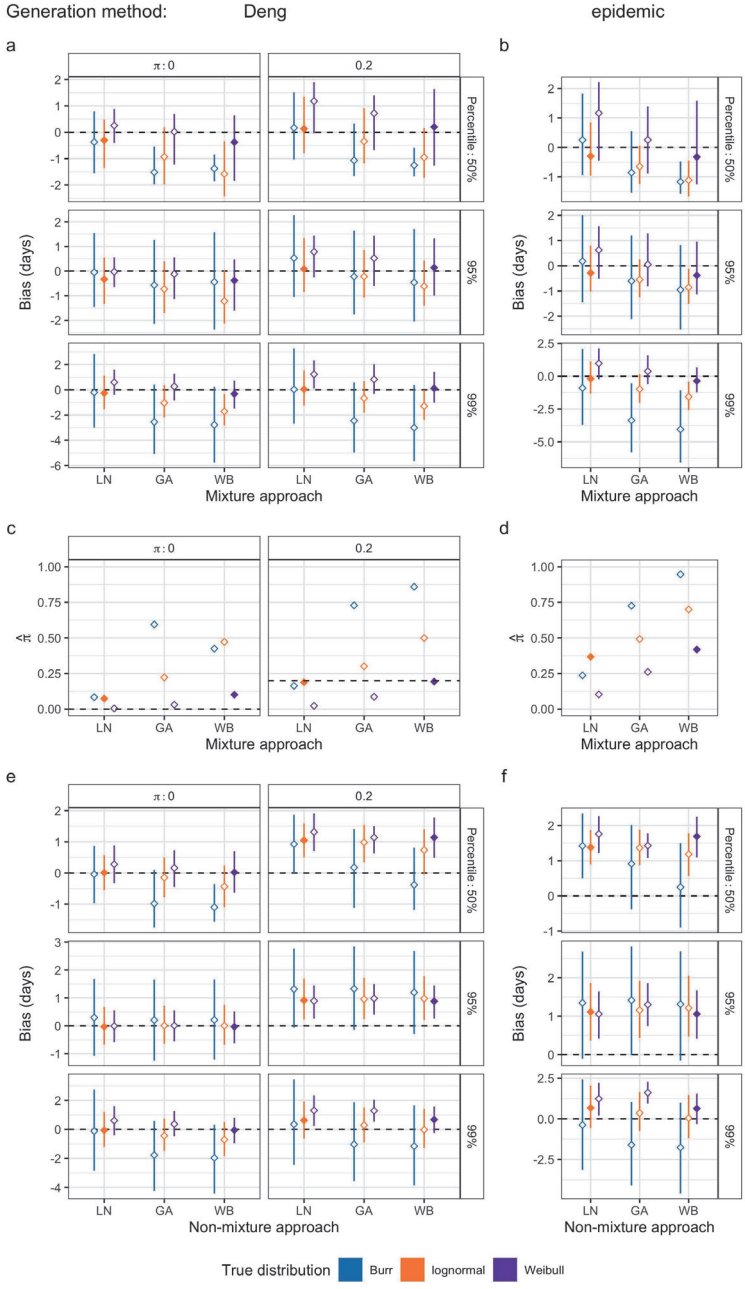


Figure 2.12: (Caption next page.)

Figure 2.12: Results of simulation study investigating estimation of incubation time distribution using an method inspired by renewal process theory. Sample size is 500. Incubation times were generated from three different distributions: Burr (blue); lognormal (orange); Weibull (purple). Data generation: left panel **a, c, e** Dengs method; right panel **b, d, f** new method: epidemic outbreak with $\pi = 0$. Estimation approach: mixture including π (Fig. **a**, **b**, **c**, **d**) or excluding π (e,f). Fig. **a,b,e,f** and **c-d** show the bias in the estimate of the percentiles and the average estimate of π respectively. The lines represent either zero bias, or the π with which the data was generated.

Table 2.7: Results for data generated according to Deng, data set size N = 500. Incubation times were generated from three different distributions. A mixture approach including π or a non-mixture approach was taken, assuming different parametric distributions. The following summary measures are given per estimated percentile: the bias, the 25% and 75% percentiles of the deviations and the average estimate of π where applicable, along with a (normal approximated) 95% confidence interval. ≈ 0 means a value between -0.005 and 0.005.

True distribution	Assumed distribution	True π	(Mixture) model	50%			90%			95%			97.5%			99%			π		
				Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Mean	95% CI	
Burr	gamma	0	no π	-0.98	-1.75	0.10	0.18	-1.01	1.41	0.21	-1.24	1.65	-0.24	-2.05	1.53	-1.77	-4.23	0.58	-	-	-
	lognormal			-0.04	-0.97	0.86	0.18	-0.97	1.35	0.30	-1.08	1.68	0.28	-1.53	2.11	-0.12	-2.84	2.75	-	-	-
	Weibull			-1.10	-1.56	-0.36	0.20	-0.92	1.33	0.21	-1.20	1.66	-0.30	-2.12	1.49	-1.97	-4.41	0.32	-	-	-
	gamma		with π	-1.52	-1.97	-0.55	-0.57	-1.79	0.87	-0.57	-2.14	1.26	-1.03	-3.02	1.27	-2.55	-5.07	0.39	.59	.57	.62
	lognormal			-0.37	-1.55	0.79	-0.21	-1.56	1.22	-0.05	-1.46	1.54	0.01	-1.82	1.85	-0.22	-2.99	2.84	.08	.08	.09
	Weibull			-1.38	-1.85	-0.85	-0.37	-1.84	1.21	-0.44	-2.37	1.58	-1.02	-3.39	1.43	-2.78	-5.76	0.24	.42	.40	.45
	gamma	0.1	no π	-0.45	-1.47	0.82	0.79	-0.37	2.09	0.79	-0.62	2.33	0.29	-1.54	2.08	-1.32	-3.86	1.16	-	-	-
	lognormal			0.43	-0.47	1.40	0.76	-0.39	2.00	0.84	-0.62	2.31	0.75	-1.20	2.60	0.19	-2.62	2.98	-	-	-
	Weibull			-0.80	-1.39	0.23	0.68	-0.49	1.90	0.71	-0.74	2.20	0.20	-1.67	2.08	-1.49	-4.14	0.95	-	-	-
	gamma		with π	-1.30	-1.83	0.09	-0.34	-1.68	1.22	-0.37	-2.13	1.59	-0.86	-3.02	1.56	-2.44	-5.24	0.67	.70	.68	.73
	lognormal			-0.07	-1.22	1.22	0.17	-1.30	1.59	0.31	-1.36	1.91	0.33	-1.69	2.31	-0.01	-2.95	2.96	.12	.11	.13
	Weibull			-1.32	-1.80	-0.63	-0.33	-1.77	1.37	-0.43	-2.32	1.78	-1.04	-3.37	1.67	-2.83	-5.73	0.54	.67	.64	.69
	gamma	0.2	no π	0.18	-1.11	1.40	1.40	0.16	2.57	1.33	-0.14	2.85	0.73	-1.02	2.69	-1.03	-3.55	1.86	-	-	-
	lognormal			0.93	-0.02	1.87	1.29	0.12	2.41	1.32	-0.07	2.77	1.12	-0.73	3.11	0.35	-2.44	3.44	-	-	-
	Weibull			-0.38	-1.18	0.81	1.20	-0.02	2.38	1.20	-0.29	2.68	0.63	-1.17	2.63	-1.16	-3.86	1.66	-	-	-
	gamma		with π	-1.06	-1.66	0.33	-0.15	-1.36	1.30	-0.22	-1.76	1.64	-0.78	-2.72	1.50	-2.45	-4.95	0.58	.73	.71	.75
	lognormal			0.17	-1.03	1.51	0.43	-0.99	2.03	0.53	-1.05	2.27	0.49	-1.40	2.69	0.03	-2.69	3.28	.16	.15	.17
	Weibull			-1.25	-1.67	-0.59	-0.32	-1.49	1.32	-0.46	-2.04	1.71	-1.13	-3.15	1.53	-3.00	-5.66	0.37	.86	.84	.88
lognormal	gamma	0	no π	-0.15	-0.78	0.49	0.08	-0.5	0.65	0.01	-0.64	0.72	-0.14	-0.96	0.77	-0.44	-1.49	0.74	-	-	-
	lognormal			0.01	-0.55	0.57	-0.02	-0.59	0.55	-0.03	-0.67	0.67	-0.05	-0.85	0.89	-0.06	-1.20	1.18	-	-	-
	Weibull			-0.43	-1.09	0.24	0.10	-0.49	0.71	≈ 0	-0.68	0.75	-0.23	-1.08	0.70	-0.71	-1.84	0.50	-	-	-
	gamma		with π	-0.93	-1.98	0.19	-0.70	-1.69	0.35	-0.73	-1.70	0.38	-0.82	-1.86	0.35	-1.04	-2.17	0.36	.22	.21	.23
	lognormal			-0.30	-1.35	0.47	-0.34	-1.35	0.44	-0.33	-1.33	0.55	-0.31	-1.40	0.72	-0.27	-1.53	1.11	.07	.07	.08
	Weibull			-1.58	-2.43	-0.35	-1.17	-2.04	0.13	-1.22	-2.13	0.04	-1.37	-2.29	-0.15	-1.72	-2.82	-0.35	.47	.46	.49
	gamma	0.1	no π	0.42	-0.19	1.06	0.62	0.03	1.21	0.52	-0.22	1.26	0.34	-0.59	1.30	-0.01	-1.22	1.31	-	-	-
	lognormal			0.54	-0.01	1.12	0.52	-0.07	1.09	0.48	-0.25	1.18	0.42	-0.53	1.37	0.33	-0.96	1.69	-	-	-
	Weibull			0.14	-0.59	0.86	0.67	0.08	1.24	0.53	-0.24	1.26	0.24	-0.74	1.26	-0.32	-1.63	1.08	-	-	-
	gamma		with π	-0.55	-1.55	0.69	-0.31	-1.25	0.84	-0.36	-1.34	0.87	-0.47	-1.55	0.84	-0.73	-2.04	0.72	.24	.23	.25
	lognormal			0.04	-1.00	0.98	0.02	-1.07	0.95	0.01	-1.07	1.06	0.01	-1.18	1.19	0.01	-1.38	1.46	.11	.10	.11
	Weibull			-1.21	-2.05	0.34	-0.76	-1.60	0.57	-0.83	-1.73	0.38	-1.01	-2.02	0.20	-1.41	-2.62	-0.04	.45	.44	.47
	gamma	0.2	no π	0.98	0.35	1.54	1.10	0.53	1.73	0.95	0.24	1.72	0.72	-0.19	1.66	0.29	-0.89	1.51	-	-	-
	lognormal			1.05	0.50	1.59	1.00	0.44	1.63	0.91	0.23	1.69	0.80	-0.12	1.77	0.62	-0.63	1.92	-	-	-
	Weibull			0.74	-0.05	1.39	1.18	0.60	1.82	0.98	0.21	1.78	0.62	-0.35	1.68	-0.03	-1.28	1.41	-	-	-
	gamma		with π	-0.34	-1.17	0.91	-0.15	-0.93	0.95	-0.22	-1.07	0.85	-0.36	-1.31	0.76	-0.66	-1.80	0.68	.30	.29	.31
	lognormal			0.13	-0.78	1.35	0.11	-0.82	1.37	0.09	-0.84	1.35	0.07	-0.98	1.43	0.04	-1.23	1.52	.19	.18	.20
	Weibull			-0.95	-1.73	0.16	-0.52	-1.24	0.58	-0.61	-1.40	0.42	-0.83	-1.73	0.24	-1.29	-2.38	≈ 0	.50	.49	.51
Weibull	gamma	0	no π	0.16	-0.45	0.73	-0.07	-0.57	0.41	0.01	-0.55	0.56	0.14	-0.56	0.83	0.37	-0.46	1.26	-	-	-
	lognormal			0.28	-0.32	0.88	-0.12	-0.62	0.38	-0.01	-0.58	0.56	0.21	-0.52	0.92	0.61	-0.39	1.60	-	-	-
	Weibull			0.02	-0.63	0.70	-0.02	-0.53	0.47	-0.03	-0.62	0.52	-0.04	-0.75	0.62	-0.06	-0.95	0.80	-	-	-
	gamma		with π	0.02	-1.22	0.69	-0.21	-1.31	0.41	-0.12	-1.14	0.56	0.02	-0.97	0.83	0.26	-0.84	1.26	.03	.03	.04
	lognormal			0.26	-0.40	0.88	-0.15	-0.70	0.37	-0.03	-0.65	0.56	0.19	-0.55	0.92	0.59	-0.41	1.60	$< .01$	$< .01$	.01
	Weibull			-0.38	-1.84	0.64	-0.40	-1.71	0.42	-0.38	-1.60	0.47	-0.36	-1.52	0.59	-0.32	-1.48	0.72	.10	.09	.11
	gamma	0.1	no π	0.66	0.05	1.14	0.43	-0.02	0.94	0.50	-0.03	1.06	0.62	0.02	1.26	0.83	0.09	1.65	-	-	-
	lognormal			0.78	0.15	1.40	0.38	-0.08	0.89	0.47	-0.08	1.04	0.66	-0.08	1.35	1.01	-0.03	2.00	-	-	-
	Weibull			0.56	-0.13	1.24	0.49	0.04	0.99	0.45	-0.11	1.01	0.41	-0.29	1.08	0.36	-0.55	1.19	-	-	-
	gamma		with π	0.44	-0.86	1.09	0.22	-0.92	0.88	0.29	-0.83	1.01	0.42	-0.73	1.25	0.64	-0.55	1.65	.05	.04	.05
	lognormal			0.73	-0.16	1.38	0.33	-0.53	0.88	0.43	-0.35	1.02	0.62	-0.23	1.35	0.98	-0.11	1.98	.01	.01	.01
	Weibull			≈ 0	-1.52	1.14	-0.02	-1.35	0.92	-0.02	-1.25	0.91	-0.01	-1.22	1.00	-0.01	-1.19	1.09	.13	.12	.14
	gamma	0.2	no π	1.14	0.63	1.50	0.92	0.42	1.36	0.98	0.40	1.50	1.09	0.43	1.69	1.28	0.54	2.04	-	-	-
	lognormal			1.31	0.71	1.91	0.84	0.33	1.29	0.89	0.26	1.45	1.03	0.23	1.77	1.30	0.24	2.35	-	-	-
	Weibull			1.14	0.49	1.78	0.97	0.46	1.45	0.88	0.26	1.45	0.79	0.04	1.48	0.67	-0.24	1.58	-	-	-
	gamma		with π	0.72	-0.67	1.40	0.47	-0.71	1.29	0.52	-0.60	1.44	0.63	-0.44	1.65	0.83	-0.33	2.02	.09	.08	.09
	lognormal			1.18	-0.04	1.89	0.72	-0.37	1.29	0.79	-0.25	1.45	0.93	-0.07	1.76	1.23	0.12	2.33	.02	.02	.03
	Weibull			0.20	-1.27	1.64	0.16	-1.07	1.32	0.14	-0.99	1.33	0.13	-0.96	1.34	0.11	-1.02	1.42	.19	.18	.20

Table 2.8: Results for data generated according to Deng, data set size N = 1200. Incubation times were generated from three different distributions. A mixture approach including π or a non-mixture approach was taken, assuming different parametric distributions. The following summary measures are given per estimated percentile: the bias, the 25% and 75% percentiles of the deviations and the average estimate of π where applicable, along with a (normal approximated) 95% confidence interval. ≈ 0 means a value between -0.005 and 0.005.

True distribution	Assumed distribution	True π	(Mixture) model	50%			90%			95%			97.5%			99%			π		
				Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Mean	95% CI	
Burr	gamma	0	no π	-1.05	-1.62	-0.25	0.14	-0.63	0.95	0.20	-0.69	1.10	-0.22	-1.36	0.95	-1.72	-3.30	-0.09	-	-	-
	lognormal			-0.03	-0.63	0.55	0.20	-0.49	0.96	0.32	-0.58	1.22	0.29	-0.95	1.54	-0.12	-2.00	1.69	-	-	-
	Weibull			-1.18	-1.46	-0.78	0.14	-0.59	0.91	0.18	-0.74	1.19	-0.30	-1.47	0.96	-1.91	-3.48	-0.32	-	-	-
	gamma		with π	-1.61	-1.89	-1.04	-0.67	-1.51	0.36	-0.65	-1.76	0.63	-1.09	-2.49	0.50	-2.58	-4.38	-0.56	.73	.70	.75
	lognormal			-0.29	-1.10	0.48	-0.12	-1.06	0.82	0.02	-0.96	1.13	0.06	-1.19	1.37	-0.24	-2.10	1.62	.06	.06	.07
	Weibull			-1.40	-1.76	-1.00	-0.38	-1.49	0.80	-0.44	-1.86	1.06	-1.01	-2.74	0.81	-2.75	-4.87	-0.52	.43	.41	.45
	gamma	0.1	no π	-0.46	-1.36	0.37	0.81	0.03	1.63	0.82	-0.08	1.73	0.34	-0.89	1.57	-1.26	-2.93	0.47	-	-	-
	lognormal			0.44	-0.16	1.11	0.79	0.04	1.55	0.88	-0.06	1.79	0.79	-0.49	2.01	0.23	-1.62	2.00	-	-	-
	Weibull			-0.89	-1.29	-0.24	0.62	-0.13	1.43	0.69	-0.24	1.66	0.23	-1.01	1.51	-1.38	-3.12	0.33	-	-	-
	gamma		with π	-1.42	-1.74	-0.54	-0.45	-1.31	0.60	-0.45	-1.56	0.76	-0.92	-2.32	0.55	-2.47	-4.28	-0.56	.83	.81	.84
	lognormal			-0.02	-0.91	0.88	0.24	-0.84	1.30	0.37	-0.77	1.53	0.38	-0.99	1.72	0.01	-1.91	1.84	.10	.10	.11
	Weibull			-1.40	-1.69	-0.98	-0.44	-1.36	0.79	-0.53	-1.74	1.04	-1.13	-2.64	0.78	-2.90	-4.81	-0.55	.78	.76	.79
	gamma	0.2	no π	0.18	-0.64	1.04	1.44	0.67	2.23	1.37	0.44	2.34	0.78	-0.41	2.03	-0.97	-2.65	0.81	-	-	-
	lognormal			0.92	0.33	1.55	1.32	0.57	2.09	1.36	0.48	2.34	1.18	-0.04	2.41	0.43	-1.4	2.26	-	-	-
	Weibull			-0.45	-1.09	0.36	1.18	0.42	1.95	1.21	0.30	2.17	0.68	-0.54	1.91	-1.06	-2.85	0.72	-	-	-
	gamma		with π	-1.18	-1.57	-0.30	-0.25	-1.08	0.82	-0.30	-1.31	0.86	-0.83	-2.12	0.58	-2.47	-4.16	-0.62	.81	.79	.82
	lognormal			0.17	-0.70	1.25	0.45	-0.55	1.60	0.55	-0.55	1.83	0.51	-0.83	1.99	0.05	-1.79	1.97	.15	.15	.16
	Weibull			-1.33	-1.57	-1.06	-0.40	-1.18	0.55	-0.53	-1.61	0.74	-1.18	-2.58	0.42	-3.02	-4.90	-0.99	.96	.95	.96
lognormal	gamma	0	no π	-0.16	-0.60	0.23	0.09	-0.28	0.42	0.03	-0.41	0.45	-0.11	-0.67	0.40	-0.40	-1.14	0.29	-	-	-
	lognormal			≈ 0	-0.37	0.35	-0.01	-0.36	0.33	-0.01	-0.42	0.40	-0.02	-0.57	0.51	-0.02	-0.80	0.73	-	-	-
	Weibull			-0.47	-0.92	-0.04	0.12	-0.25	0.47	0.03	-0.41	0.46	-0.19	-0.75	0.36	-0.65	-1.41	0.10	-	-	-
	gamma		with π	-0.97	-1.70	-0.07	-0.72	-1.38	0.13	-0.74	-1.43	0.11	-0.83	-1.55	0.06	-1.04	-1.87	-0.13	.21	.21	.22
	lognormal			-0.25	-1.04	0.28	-0.26	-1.10	0.29	-0.25	-1.06	0.33	-0.23	-1.04	0.45	-0.19	-1.10	0.67	.05	.05	.06
	Weibull			-1.67	-2.30	-0.86	-1.24	-1.83	-0.50	-1.26	-1.85	-0.55	-1.40	-2.02	-0.64	-1.72	-2.47	-0.89	.48	.47	.49
	gamma	0.1	no π	0.42	0.04	0.81	0.63	0.27	1.00	0.53	0.08	1.01	0.35	-0.23	0.99	≈ 0	-0.79	0.86	-	-	-
	lognormal			0.53	0.19	0.90	0.53	0.17	0.89	0.49	0.04	0.96	0.44	-0.14	1.07	0.36	-0.47	1.25	-	-	-
	Weibull			0.12	-0.33	0.56	0.69	0.32	1.06	0.55	0.09	1.04	0.28	-0.31	0.94	-0.27	-1.07	0.67	-	-	-
	gamma		with π	-0.59	-1.25	0.38	-0.35	-0.97	0.54	-0.39	-1.02	0.46	-0.50	-1.18	0.39	-0.76	-1.59	0.20	.24	.23	.24
	lognormal			0.06	-0.70	0.78	0.05	-0.71	0.76	0.04	-0.72	0.81	0.04	-0.78	0.93	0.03	-0.93	1.11	.10	.09	.10
	Weibull			-1.29	-1.84	-0.59	-0.82	-1.37	-0.11	-0.87	-1.43	-0.19	-1.04	-1.65	-0.31	-1.43	-2.17	-0.52	.46	.45	.47
	gamma	0.2	no π	0.99	0.61	1.34	1.11	0.73	1.48	0.96	0.48	1.46	0.73	0.12	1.36	0.30	-0.50	1.15	-	-	-
	lognormal			1.05	0.71	1.38	1.00	0.65	1.36	0.92	0.45	1.41	0.82	0.20	1.47	0.64	-0.19	1.53	-	-	-
	Weibull			0.72	0.28	1.16	1.20	0.81	1.58	1.00	0.49	1.51	0.66	≈ 0	1.36	0.02	-0.85	0.93	-	-	-
	gamma		with π	-0.37	-0.94	0.32	-0.17	-0.70	0.46	-0.24	-0.79	0.41	-0.38	-0.99	0.34	-0.67	-1.43	0.18	.30	.29	.30
	lognormal			0.10	-0.52	1.13	0.08	-0.56	1.08	0.07	-0.57	1.01	0.05	-0.65	0.96	0.03	-0.80	1.01	.19	.18	.19
	Weibull			-1.00	-1.51	-0.41	-0.54	-1.06	≈ 0	-0.63	-1.13	-0.06	-0.84	-1.42	-0.16	-1.28	-1.99	-0.47	.50	.49	.50
Weibull	gamma	0	no π	0.17	-0.25	0.55	-0.06	-0.37	0.26	0.01	-0.34	0.39	0.15	-0.28	0.59	0.37	-0.17	0.95	-	-	-
	lognormal			0.28	-0.13	0.65	-0.11	-0.42	0.21	0.01	-0.35	0.38	0.22	-0.21	0.68	0.63	0.01	1.24	-	-	-
	Weibull			0.01	-0.42	0.42	-0.01	-0.33	0.32	-0.02	-0.36	0.35	-0.02	-0.43	0.42	-0.03	-0.57	0.52	-	-	-
	gamma		with π	0.15	-0.39	0.55	-0.08	-0.50	0.26	-0.01	-0.47	0.39	0.13	-0.40	0.59	0.36	-0.25	0.95	< .01	< .01	.01
	lognormal			0.28	-0.13	0.65	-0.11	-0.42	0.21	0.01	-0.35	0.38	0.22	-0.21	0.68	0.63	0.01	1.24	< .01	-	-
	Weibull			-0.29	-1.46	0.41	-0.30	-1.36	0.30	-0.29	-1.28	0.32	-0.27	-1.21	0.38	-0.25	-1.12	0.47	.07	.06	.08
	gamma	0.1	no π	0.69	0.32	1.03	0.45	0.14	0.76	0.51	0.14	0.88	0.62	0.18	1.08	0.83	0.27	1.42	-	-	-
	lognormal			0.79	0.43	1.17	0.40	0.09	0.71	0.50	0.11	0.87	0.69	0.18	1.17	1.05	0.36	1.73	-	-	-
	Weibull			0.55	0.16	0.96	0.51	0.20	0.82	0.48	0.10	0.84	0.44	-0.04	0.90	0.40	-0.20	0.99	-	-	-
	gamma		with π	0.61	-0.30	1.00	0.37	-0.48	0.76	0.44	-0.39	0.87	0.55	-0.25	1.08	0.76	-0.06	1.42	.01	.01	.02
	lognormal			0.78	0.41	1.17	0.39	0.07	0.71	0.50	0.09	0.87	0.69	0.18	1.17	1.05	0.36	1.73	< .01	< .01	< .01
	Weibull			0.08	-1.20	0.91	0.08	-1.01	0.79	0.08	-0.96	0.79	0.09	-0.91	0.83	0.09	-0.85	0.90	.10	.09	.11
	gamma	0.2	no π	1.16	0.85	1.41	0.91	0.61	1.23	0.96	0.62	1.31	1.06	0.68	1.46	1.24	0.79	1.73	-	-	-
	lognormal			1.30	0.92	1.68	0.85	0.56	1.16	0.91	0.55	1.27	1.05	0.56	1.53	1.34	0.67	1.99	-	-	-
	Weibull			1.11	0.71	1.52	0.98	0.67	1.29	0.90	0.52	1.26	0.81	0.36	1.26	0.70	0.13	1.26	-	-	-
	gamma		with π	0.93	-0.23	1.33	0.64	-0.27	1.15	0.69	-0.16	1.26	0.78	-0.09	1.42	0.95	0.07	1.68	.05	.04	.05
	lognormal			1.28	0.83	1.68	0.83	0.41	1.16	0.89	0.45	1.27	1.04	0.50	1.53	1.33	0.62	1.99	< .01	< .01	< .01
	Weibull			0.20	-0.92	1.36	0.17	-0.73	1.18	0.16	-0.70	1.13	0.15	-0.66	1.13	0.14	-0.66	1.13	.18	.17	.19

Table 2.9: Results for data generated according to the epidemic curve, data set size N = 500. Incubation times were generated from three different distributions. A mixture approach including π or a non-mixture approach was taken, assuming different parametric distributions. The following summary measures are given per estimated percentile: the bias, the 25% and 75% percentiles of the deviations and the average estimate of π where applicable, along with a (normal approximated) 95% confidence interval.

True distribution	Assumed distribution	(Mixture) model	50%			90%			95%			97.5%			99%			π		
			Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Mean	95% CI	
Burr	gamma	no π	0.91	-0.38	2.01	1.75	0.59	2.85	1.42	-0.02	2.82	0.55	-1.33	2.41	-1.60	-4.09	1.03	-	-	-
	lognormal		1.42	0.50	2.35	1.54	0.43	2.57	1.35	-0.10	2.68	0.87	-1.04	2.78	-0.37	-3.14	2.42	-	-	-
	Weibull		0.25	-0.90	1.50	1.57	0.41	2.65	1.31	-0.15	2.68	0.46	-1.44	2.30	-1.76	-4.59	0.98	-	-	-
	gamma	with π	-0.86	-1.54	0.56	-0.32	-1.53	1.28	-0.60	-2.12	1.20	-1.38	-3.26	0.75	-3.37	-5.79	-0.55	.73	.71	.74
	lognormal		0.25	-0.94	1.83	0.25	-1.17	2.02	0.18	-1.44	2.00	-0.09	-2.03	2.06	-0.90	-3.70	2.09	.24	.23	.25
	Weibull		-1.17	-1.57	-0.48	-0.59	-1.77	0.67	-0.95	-2.52	0.82	-1.85	-3.79	0.41	-4.05	-6.57	-1.08	.95	.94	.96
lognormal	gamma	no π	1.37	0.87	1.88	1.36	0.79	1.96	1.16	0.43	1.92	0.86	-0.03	1.86	0.36	-0.75	1.64	-	-	-
	lognormal		1.38	0.90	1.87	1.25	0.69	1.82	1.11	0.37	1.86	0.94	-0.02	1.93	0.67	-0.56	2.04	-	-	-
	Weibull		1.19	0.56	1.78	1.49	0.92	2.09	1.21	0.47	2.05	0.79	-0.15	1.90	0.04	-1.18	1.46	-	-	-
	gamma	with π	-0.65	-1.24	0.05	-0.48	-1.10	0.25	-0.55	-1.25	0.25	-0.69	-1.53	0.18	-0.98	-2.03	0.15	.49	.49	.50
	lognormal		-0.30	-0.96	0.85	-0.31	-0.98	0.78	-0.29	-1.02	0.80	-0.25	-1.15	0.80	-0.19	-1.33	1.13	.37	.36	.37
	Weibull		-1.11	-1.66	-0.45	-0.74	-1.34	-0.08	-0.86	-1.51	-0.13	-1.10	-1.91	-0.23	-1.57	-2.58	-0.43	.70	.69	.71
Weibull	gamma	no π	1.43	1.07	1.78	1.24	0.73	1.74	1.30	0.74	1.86	1.41	0.80	2.02	1.60	0.93	2.27	-	-	-
	lognormal		1.76	1.22	2.27	1.10	0.58	1.60	1.06	0.42	1.65	1.09	0.33	1.81	1.23	0.21	2.20	-	-	-
	Weibull		1.69	1.09	2.25	1.25	0.72	1.77	1.06	0.41	1.67	0.87	0.08	1.59	0.63	-0.31	1.54	-	-	-
	gamma	with π	0.26	-0.88	1.39	-0.01	-0.90	1.20	0.05	-0.81	1.28	0.17	-0.73	1.39	0.38	-0.61	1.58	.26	.25	.27
	lognormal		1.16	-0.44	2.21	0.60	-0.68	1.52	0.62	-0.52	1.57	0.73	-0.38	1.75	0.97	-0.23	2.12	.10	.09	.11
	Weibull		-0.33	-1.25	1.58	-0.38	-1.15	1.12	-0.38	-1.13	0.95	-0.37	-1.14	0.75	-0.36	-1.23	0.68	.42	.41	.43

Table 2.10: Results for data generated according to the epidemic curve, data set size N = 1200. Incubation times were generated from three different distributions. A mixture approach including π or a non-mixture approach was taken, assuming different parametric distributions. The following summary measures are given per estimated percentile: the bias, the 25% and 75% percentiles of the deviations and the average estimate of π where applicable, along with a (normal approximated) 95% confidence interval.

True distribution	Assumed distribution	(Mixture) model	50%			90%			95%			97.5%			99%			π		
			Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Bias	25^{th} p.	75^{th} p.	Mean	95% CI	
Burr	gamma	no π	0.91	0.15	1.67	1.76	1.03	2.52	1.44	0.49	2.34	0.58	-0.69	1.74	-1.57	-3.29	0.07	-	-	-
	lognormal		1.42	0.84	2.03	1.55	0.83	2.29	1.36	0.41	2.28	0.89	-0.37	2.09	-0.36	-2.18	1.41	-	-	-
	Weibull		0.21	-0.57	1.03	1.57	0.85	2.35	1.33	0.42	2.21	0.49	-0.80	1.68	-1.71	-3.61	0.10	-	-	-
	gamma	with π	-0.97	-1.43	-0.12	-0.41	-1.22	0.49	-0.67	-1.67	0.38	-1.43	-2.66	-0.11	-3.38	-5.01	-1.69	.76	.75	.78
	lognormal		0.20	-0.59	1.41	0.20	-0.74	1.49	0.13	-0.95	1.39	-0.13	-1.45	1.22	-0.94	-2.68	0.94	.24	.23	.24
	Weibull		-1.22	-1.47	-0.95	-0.65	-1.43	0.09	-1.00	-2.09	0.09	-1.88	-3.32	-0.46	-4.06	-5.95	-2.15	.99	.99	.99
lognormal	gamma	no π	1.38	1.02	1.70	1.37	0.98	1.78	1.16	0.66	1.68	0.86	0.23	1.49	0.35	-0.46	1.17	-	-	-
	lognormal		1.38	1.04	1.69	1.26	0.87	1.65	1.13	0.62	1.64	0.96	0.31	1.59	0.69	-0.18	1.54	-	-	-
	Weibull		1.18	0.76	1.60	1.50	1.09	1.92	1.23	0.67	1.77	0.80	0.11	1.50	0.06	-0.83	0.96	-	-	-
	gamma	with π	-0.64	-1.05	-0.21	-0.47	-0.89	-0.02	-0.54	-1.01	-0.03	-0.68	-1.22	-0.07	-0.97	-1.65	-0.23	.48	.48	.49
	lognormal		-0.33	-0.74	0.20	-0.33	-0.75	0.19	-0.30	-0.78	0.26	-0.26	-0.86	0.38	-0.20	-1.01	0.63	.37	.36	.37
	Weibull		-1.11	-1.50	-0.72	-0.74	-1.14	-0.33	-0.85	-1.30	-0.37	-1.08	-1.61	-0.50	-1.55	-2.26	-0.83	.69	.69	.70
Weibull	gamma	no π	1.44	1.21	1.66	1.24	0.90	1.57	1.31	0.93	1.67	1.42	1.00	1.81	1.61	1.15	2.04	-	-	-
	lognormal		1.77	1.42	2.10	1.12	0.77	1.44	1.08	0.65	1.48	1.12	0.60	1.61	1.26	0.58	1.90	-	-	-
	Weibull		1.69	1.32	2.04	1.27	0.91	1.61	1.08	0.65	1.49	0.90	0.39	1.37	0.66	0.01	1.24	-	-	-
	gamma	with π	0.31	-0.59	1.19	0.02	-0.65	0.88	0.07	-0.55	0.90	0.18	-0.44	1.01	0.38	-0.26	1.16	.24	.24	.25
	lognormal		1.32	-0.14	2.09	0.74	-0.38	1.41	0.74	-0.24	1.42	0.83	-0.07	1.53	1.05	0.21	1.80	.07	.07	.08
	Weibull		-0.42	-1.02	0.50	-0.44	-0.95	0.23	-0.42	-0.89	0.17	-0.40	-0.89	0.15	-0.38	-0.93	0.20	.43	.42	.44

2.7.4 Note on NPMLE

There are some issues when using the NPMLE for estimation of percentiles of the incubation time distribution. This especially holds with data on the SARS-CoV-2 incubation time that are small in size and discrete in nature. We explain the issues using the open source data set provided by Lauer *et al.* [2020] (Figure 2.8).

Figure 2.13 visualizes the uncertainty in the incubation time per individual via grey horizontal lines. Each line starts at the minimum possible incubation time and it ends at the maximum possible incubation time if it is smaller than 15 days; most individuals have an upper limit of 15 days or larger. The solid black curve is the NPMLE of the cumulative distribution function (CDF). The grey shaded area denote intervals in which the NPMLE is not uniquely defined. For example, the last jump in the NPMLE can be at day 9, 10 or 11. This is a consequence of the interval censored nature of the data, with one individual having a minimum incubation time of 9 days, and another individual having a maximum incubation time of 11 days; there is no further information in-between. The dashed horizontal line represents the 95^{th} percentile. The dashed line crosses the black curve in a grey shaded area, which implies that the 95^{th} percentile is between 9 and 11 days.

In simulation study I (Section 2.3) we saw that for small data sets, the NPMLE often had its last jump from a value below 0.95 to the value 1, and the location of the jump was not uniquely defined. In that case, the midpoint of the horizontal segment was returned by the function `quantile.survfit` from the `R survival` package that we used. As a consequence, each of the estimated tail percentiles of 95 and higher has the same value. We used the confidence intervals as provided by the `survfit` function, which are known to be inconsistent. The NPMLE does not have $\sqrt{n}$ convergence rate and the central limit theorem does not hold [Anderson-Bergman, 2017]. It is known that an "n-out-of-n" bootstrap sampling gives inconsistent confidence intervals and that m-out-of-n subsampling is needed.

The large jumps in the tail also impact the confidence intervals. Confidence intervals for the CDF reduce to zero width (using option 'plain' in the `survfit` function) once the estimate reaches the value 1. For the other options ('log' and 'log-log') the confidence interval is missing. Confidence intervals for the percentiles (the horizontal direction) are typically based on the confidence intervals for the CDF (the vertical direction): lower and upper bound of the confidence interval for the percentile are the points where a horizontal

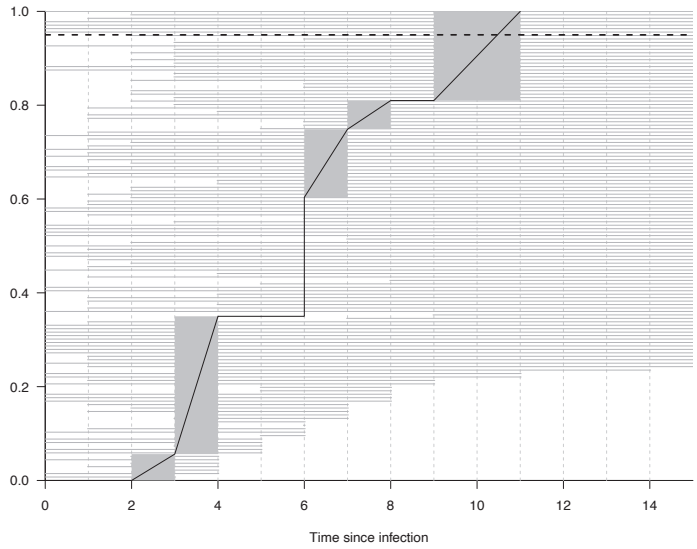


Figure 2.13: Minimum and maximum possible incubation time for each individual along with the NPMLE. Grey horizontal lines represent the interval from the minimum to maximum possible incubation time in days for each observation. Intervals are sorted by maximum possible incubation time. The solid black line represents the NPMLE of the cumulative distribution function (CDF). The grey shaded areas represent the Turnbull intervals where the jump location is not uniquely defined. The jumps in the curve without grey shaded are a represent a point mass on a specific incubation time instead. The dashed line represents the 95^{th} percentile. The figure restricts to the first fifteen days.

line at the percentile crosses the vertical confidence intervals. If the last jump in the NPMLE goes from a value below 0.95 to the value 1, the upper bound of the confidence interval around the 95th percentile (and higher ones) coincides with the estimate (option 'plain') of the percentile or is missing (options 'log' and 'log-log').

2.7.5 Note on PGM

The penalized Gaussian mixture approach (PGM) is implemented in `R` package `smoothSurv` [Komárek, 2020]. As default, the function `smoothSurvReg` chooses the optimal smoothing parameter λ from a grid of values $e^{2:-9}$ based on AIC in a bootstrap routine. We found that sometimes a small λ value was selected that gave a density estimate with multiple modes, while we generated data from a distribution with a unimodal density. We think that with real data a unimodal incubation time distribution is more common. The

Table 2.11: Performance of penalized Gaussian mixture approach with different settings regarding smoothing parameter λ: standard (default grid of λ in `smoothSurvReg` function, i.e. $e^{2:-9}$) and alternative (sequence of λ's 0.1 to 5.6 by steps of 0.5). We generated 1000 data sets with 100 observations from a Weibull distribution with a constant risk of infection on and empirical width of the exposure window. p25 and p75 denote the first and third percentile of the deviations between true value and estimate; TO denotes the percentage of runs that timed out. Note that the percentage TO is lower than for the same scenario in Table 2.3 as for the latter, besides point estimates their Confidence Intervals were calculated as well.

Approach	Percentile (%)	Bias	p25	p75	MSE	TO (%)
Standard	50	-0.04	-0.31	0.20	0.14	0
	90	-0.06	-0.38	0.26	0.23	0
	95	0.07	-0.36	0.48	0.38	0
	97.5	0.26	-0.34	0.83	0.74	0
	99	0.59	-0.25	1.44	1.71	0
Alternative	50	-0.07	-0.32	0.17	0.13	0.1
	90	-0.05	-0.35	0.26	0.21	0.1
	95	0.14	-0.26	0.54	0.36	0.1
	97.5	0.41	-0.11	0.90	0.73	0.1
	99	0.86	0.16	1.54	1.76	0.1

package does not incorporate such a constraint. When fitting PGM to real data, we advise to plot the resulting distribution, and to adjust the grid of λ values only when a too wiggly pattern is observed. However, in our simulations, an automatised procedure was needed, and computation time was limited. Accordingly, we chose a lowest value in our sequence of λ values by plotting densities for some λ values, and chose the smallest one that gave a unimodal distribution. We chose a sequence of 0.5 to 5.5, or 0.1 to 5.6 (step size 0.5) for data sets of size 100 or 500, respectively. Figure 2.14a shows a histogram of the chosen λ for each of 1000 runs for one scenario (Weibull incubation times; constant risk of infection; data set size 100). Using the default range, choosing λs smaller than 0.5 is frequent, yielding a wiggly pattern (Figure 2.14b). The difference in performance between the two procedures is small (Table 2.11). Our choice gave estimates with a slightly smaller mean squared error for all percentiles considered, except for the most extreme one. However, bias tended to be slightly larger. Note that the results for the alternative approach were regenerated such that they randomly deviate from the results in Table 2.3.

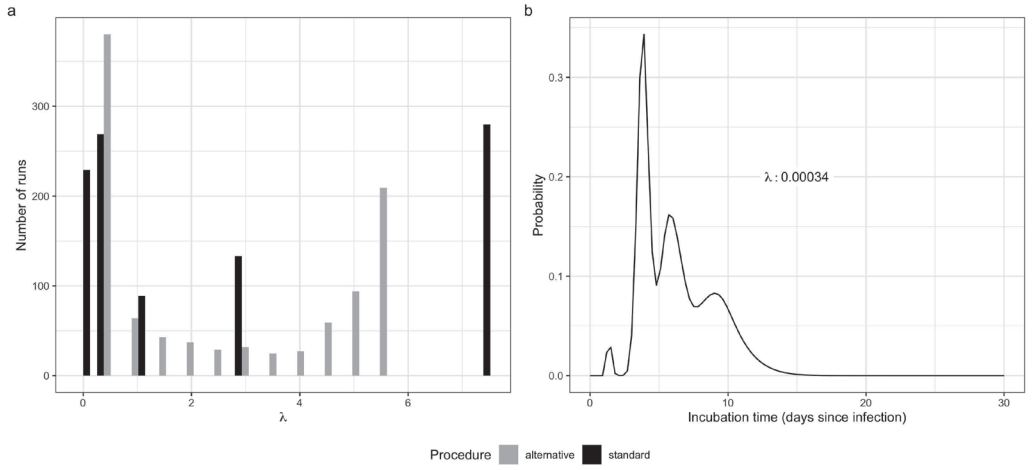


Figure 2.14: Comparison between standard (default grid of λ in `smoothSurvReg` function, i.e. $e^{2:-9}$ and alternative (sequence of λs 0.1 to 5.6 by steps of 0.5) choice of lambda's in penalized Gaussian mixture approach. The alternative approach was employed to automatise the approach and save computation time in the simulation study. We generated 1000 data sets with 100 incubation times from a Weibull distribution with a constant risk of infection on and empirical with of the exposure window. Panel (**a**) summarizes the chosen λ for each run using the standard and alternative procedure (black and grey); panel (**b**) shows the fitted incubation time distribution for one data set. The chosen λ was smaller than 0.5 and a wiggly pattern is observed.

This chapter is published as Vera H. Arntzen, Marta Fiocco and Ronald B. Geskus (2024). Two biases in incubation time estimation related to exposure. BMJ Infectious Diseases, 24, 555 [Arntzen et al., 2024].

Two biases in incubation time estimation related to exposure

Contents

Abstract

Estimation of the SARS-CoV-2 incubation time distribution is hampered by incomplete data about infection. We discuss two biases that may result from incorrect handling of such data.

Notified cases may recall recent exposures more precisely (differential recall). This creates bias if the analysis is restricted to observations with well-defined exposures, as longer incubation times are more likely to be excluded.

Another bias occurred in the initial estimates, that were based on data concerning travellers from Wuhan. Only individuals who developed symptoms after their departure were included, leading to under-representation of cases with shorter incubation times (left truncation). This factor was not considered in the analyses.

We performed simulations and give a literature review to investigate the amount of bias in estimated percentiles of the SARS-CoV-2 incubation time distribution. Depending on the rate of differential recall, restricting the analysis to a subset of narrow exposure windows resulted in underestimation in the median and even more in the 95th percentile. Failing to account for left truncation led to an overestimation of multiple days in both the median and the 95th percentile.

Keywords differential recall □ left truncation □ interval censored data □ incubation time □ SARS-CoV-2 □ interval censoring

3.1 Introduction

Incubation time is the period from infection to symptom onset. Knowing its distribution is relevant to make decisions about public health measures as well as to parameterize mathematical models for disease spread. For the SARS-CoV-2 virus, the right tail of the distribution played a crucial role in determining the appropriate duration of quarantine following infection. Estimation of the incubation time distribution of an infectious disease is hampered by incomplete data about infection. While time of symptom onset is usually known, the time origin is not. Typically, the only information available is a range of potential exposure times, yielding data with interval censored time origins. Insights into transmission are primarily obtained via contact tracing, where individuals with confirmed infection are asked about potential sources of transmission. As such, infectors and infectees can be traced.

While methods to estimate a distribution based on interval censored endpoints are well established, estimation with interval censored time origins is less straightforward. A commonly made assumption in SARS-CoV-2 incubation time estimation is that the infection time is uniformly distributed within the exposure window. Then the likelihood with interval censored time origins can be written as the likelihood for interval censored endpoints. In an earlier study, we quantified the bias introduced when the uniform assumption is violated by means of a simulation study [Arntzen et al., 2023]. We found that the incubation time is overestimated if the infection risk increases rather than remains constant within an exposure window, as happens during the initial outbreak phase of a novel pathogen. To limit bias, analysis is often restricted to observations with narrow and well-defined exposure windows [McAloon et al., 2020]. For instance, from the 255 first PCR confirmed cases of mpox in Italy, only 30 observations were used to estimate incubation time [Guzzetta et al., 2022]. These observations were chosen because both a narrow period of exposure and symptom onset could be identified. This would not be a problem, had the observation been a random subset of the data. However, this selection of observations may introduce another bias, due to the presence of differential recall.

In order to inform policy makers with respect to prevention measures at the start of an outbreak, a rapid assessment of the incubation time distribution is needed. For SARS-CoV-2, these estimates were based on data from individuals who became infected in Wuhan,

travelled from Wuhan right before the lockdown started, and developed symptoms after departure [Backer et al., 2020]. This means that their exposure window ended on the day of travel. Such data may be subject to two forms of length biased sampling. Right truncation occurs when individuals are omitted due to their ongoing incubation at the time of data collection, leading to under-representation of longer incubation times [Linton et al., 2020]. Left truncation occurs because data from Wuhan travellers only included individuals who developed symptoms after departure, leading to an under-representation of shorter incubation times. To the best of our knowledge, occurrence of left truncation in this context has not been described elsewhere.

This paper explores two biases that have been overlooked in the estimation of the incubation time for SARS-CoV-2: differential recall and left truncation. The structure of the paper is as follows. Section 2 introduces the likelihood and commonly made assumptions. Section 3 discusses the literature on differential recall and left truncation in the presence of interval censored time origin. Sections 4 and 5 present the simulation scenarios and results. The paper ends with a discussion where findings and their implications are presented. Practical recommendations for incubation time estimation are provided.

3.2 Likelihood and commonly made assumptions

Denote by E the time of infection. Typically, the knowledge about E is limited to an exposure period within which the infection took place, or only the end of the exposure period is known.

We denote the start and end of the exposure window by E_l (left) and E_r (right) respectively, with E_l possibly missing. The onset of symptoms (S) is usually known up to the precise day. The observed data with respect to the incubation time is (e_{il}, e_{ir}, s_i), all given with respect to calendar time (see Figure 3.1). Let $g_i(\cdot|e_{il}, e_{ir})$ represent the individual-specific density of the infection time, having $[e_{il}, e_{ir}]$ as support. Denote by $f(\cdot)$ and $F(\cdot)$ the density and the cumulative distribution function of the incubation time $T = S - E$, and let $h(\cdot, \cdot)$ denote the density of the observation points that define the start and end of the exposure window.

Three assumptions are commonly made:

a) The start and end of the exposure window are independent of the incubation time, i.e. $(E_{il}, E_{ir}) \perp T_i$.

b) The individual's risk of infection is constant within the exposure window, i.e. $E_i|(e_{il}, e_{ir}) \sim \text{Unif}(e_{il}, e_{ir})$.

c) The distribution of the incubation time follows a parametric distribution, such as Weibull, lognormal and gamma.

Under assumption (a), the contribution to the likelihood for individual i is given by:

$$l(e_{il}, e_{ir}, s_i) = h(e_{il}, e_{ir}) \int_{e_{il}}^{e_{ir}} g_i(u|e_{il}, e_{ir}) f(s_i - u) du. \tag{3.1}$$

Note that although E_{il}, E_{ir} and S_i are commonly observed up to a specific day, this discretization is not accounted for in the likelihood.

It is challenging to verify the validity of assumption (a) since the moment of infection, and hence also the incubation time, are typically not precisely observed. We therefore rely on reasoning why (a) is valid. Observations of incubation time are usually collected retrospectively through interviews with diagnosed individuals. Suppose that at the beginning of an outbreak, individuals who developed symptoms are interviewed on the day of symptom onset (S). Then, a person with a long incubation time needs to recall an exposure that occurred longer ago compared to a person with a short incubation time. Assumption (a) is violated if some individual characteristics that on average increase incubation time and decrease recall ability are present. However, if the ability to recall possible exposure decays over time before symptom onset similarly for all individuals, assumption (a) still holds. We provide further details in Section 3.3.1.

Assumption (b) is convenient because it makes the likelihood proportional to a likelihood for interval censored end points

$$l(e_{il}, e_{ir}, s_i) \propto F(s_i - e_{il}) - F(s_i - e_{ir}). \tag{3.2}$$

Standard estimation approaches and software are available with such interval censored end points. Assumption (b) is violated during the epidemic growth phase, leading to moderate bias in the estimates [Arntzen et al., 2023]. The amount of bias depends on the width of the exposure windows. Wider intervals do not necessarily result in greater bias. Specifically, individuals with very wide exposure windows that end at the time of symptom onset, do not provide any information. Individuals with a narrow exposure window contribute more to the estimate.

Assumption (c) is unrealistic. Historically, a lognormal distribution was commonly assumed, but the validity of the rationale behind this choice is nowadays considered questionable [Nishiura, 2007]. Whether this choice is problematic depends on the quantity of interest. The mean or median value will often be little affected by an incorrect choice of the parametric distribution of the incubation time. If the focus is on the estimation of a tail percentile, it becomes crucial to consider more flexibility in the choice of distribution. Coronaviruses are known to have an incubation time distribution with a long tail [WHO, 2003]. Hence, the gamma, lognormal or Weibull distribution may not adequately capture the true shape of the tail [WHO, 2003]. This issue can be partially overcome by using a semiparametric approach [Arntzen et al., 2023].

3.3 Literature

3.3.1 Differential recall

When collecting exposure information retrospectively through interviews with diagnosed individuals, it is important to keep in mind that our memory is not flawless. Recall bias is a term encompassing all sorts of biases that arise from differences in recall among participants in retrospective studies. A well-known example of recall bias is observed in case-control studies and retrospective cohort studies when estimating the risk associated with an exposure [Neugebauer and Ng, 1990]. Cases tend to remember exposure status more accurately than controls. This misclassification inflates odds ratios and can lead to erroneous associations [Raphael, 1987; Krämer et al., 2010]. However, differential recall is not limited to case-control studies but may occur in all observational data [Neugebauer and Ng, 1990].

Several papers on estimating SARS-CoV-2 incubation time highlight recall bias as a problem [Wu et al., 2022; Bikbov and Bikbov, 2020]. A systematic review and meta-analysis based on 42 studies where the aim was to determine the incubation period of COVID-19, showed that 78.6% (N = 33) of the estimates were potentially affected by recall bias [Dhouib et al., 2021]. Note that recall bias may occur in these studies as these typically rely on data obtained by *backward* tracing of potential infectors [Chen et al., 2022], e.g. tracing potential infectors. Recall bias is less likely to occur in data collected by *forward* tracing, e.g. tracing contacts that a notified case might have infected, but this practice less common.

Memory of an event is worse if it happened longer ago (this phenomenon inspired the game "Match the Memory"). This has been observed for exposures in cases of foodborne Hepatitis A [Petrignani et al., 2014] and prion disease [Ruegger et al., 2009], that tend to have a long incubation time.

It may also be, that the timing of the event is remembered without systematic bias, but with less precision when it occurred further in the past. In the context of estimating the incubation time distribution, individuals with confirmed infection are asked by public health officials to report their potential past exposures at the time of the interview. Typically, individuals recall recent exposures more accurately than those that occurred further in the past, leading to a broader exposure window being reported. We call this differential recall.

The two definitions are provided here:

Recall bias *Umbrella term encompassing various biases that arise from differences in recall ability among participants in retrospective studies.*

Differential recall *The phenomenon that individuals exhibit less precise recollection of the timing of an event if the event occurred further in the past. In the context of incubation time estimation, this event typically refers to potential risk exposure.*

Differential recall does not necessarily introduce bias. It becomes problematic if researchers choose to restrict the analysis only to observations with "well-defined" exposure [McAloon et al., 2020], where well-defined means that the exposure is either observed exactly or it falls within a narrow exposure window. Reasons for this choice are ample, such as: considering these observations to be more reliable; attempting to limit bias if a constant risk of infection over time is assumed [Arntzen et al., 2023]; or simplifying the analysis by treating exposures as exact rather than interval censored observations. When there is differential recall, restricting the analysis to observations with well-defined exposure may introduce bias, because observations with shorter incubation times tend to have shorter exposure windows and therefore are more likely to be included.

There is no differential recall if the exposure windows are based on test results. One example is estimation of the HIV incubation time distribution based on data from cohort studies in which individuals are tested for HIV infection at each visit [Geskus, 2001].

Literature on memory decay and differential recall is scarce, and studies typically do

not concern the infectious disease context. Most studied the strength of memory decay. Literature in experimental psychology suggests that memory decays exponentially with time [Sudman and Bradburn, 1973]. Two studies found that the recall of injuries declined if they happened longer before the interview [Heuch et al., 2018; Moshiro, 2005]. Since these studies did not consider respiratory infection and considered recall aggregated by month, results cannot necessarily be extrapolated to the SARS-CoV-2 setting. In elderly, the recalled fall rate showed a decline of 9% in a one-year compared to a quarterly survey [Yoo et al., 2017]. Two studies describe differential recall of age at menarche [Sukumaran and Dewan, 2018; Salehabadi et al., 2014]. Their data include a combination of observations with exact event times and current status data, where the age at menarche is left- or right-censored. The probability of recall, i.e. exactly observing the age at menarche, is assumed to depend on the time between menarche and the moment of recall, and it is modeled with a piecewise function.

One study focused on the mechanisms of differential recall, and discusses methods to improve the responses [Sudman and Bradburn, 1973]. Note that their conceptualization differs from differential recall as we stated in our definition. In the analysis of the impact of memory decay on responses in surveys, the authors propose a model for the effect of time on memory in survey interviews. This model consists of two components: forgetting an exposure entirely or placing it more recently than it actually occurred, which is known as forward telescoping. The latter was observed to occur more frequently than misplacing the exposure in the opposite time direction (backward telescoping). In survey research, Weber's law [Haigh et al., 2020] describes the error in time perception due to telescoping as a function of the logarithm of the time period.

Other directions to mitigate bias due to differential recall relate to the interview process [Sudman and Bradburn, 1973]. The following techniques may be beneficial for memory responses:

a) Use of records: this involves providing records of event details.

b) Aided control: by providing specific cues, such as using pictures or lists of possible exposure locations or using aided recall questions like "Did you visit a grocery store, and if so, when?"

c) Bounded recall: conducting a series of interviews covering bounded time periods (e.g. biweekly, focusing on the last two weeks).

Additionally, they discussed how interview characteristics can influence recall bias. These factors include whether it is self-administered or face-to-face, the positioning of questions, and the type of questions (open or closed).

McAloon *et al.* warn that the subset of observations with well-characterized exposures for SARS-CoV-2 may be biased toward more severe cases [McAloon et al., 2020], thus violating assumption (a). If severe cases tend to have shorter incubation periods [Lai et al., 2020], the estimates may be biased downward.

Determining the presence of differential recall in the data is challenging due to the unknown exact moment of infection. Ideally, one would assess the correlation between the width of the exposure window and the incubation time to quantify the extent of differential recall. As an approximation, the interval between the end of exposure and the interview date can be used instead of the incubation time. If a strong positive correlation is identified between the exposure window width and this interval, it serves as an indication of the presence of differential recall in the collection of exposure information.

3.3.2 Left truncation

The initial studies using data from Wuhan only included individuals who left Wuhan before the lockdown started (January 23, 2020) and were free of symptoms until the day they left Wuhan. Consequently, individuals with shorter incubation times were more likely to be excluded.

Apart from E and S as the calendar time of infection and onset of symptoms respectively, we additionally denote V as the calendar time of leaving Wuhan. The observed data for individual i are $(e_{il}, e_{ir}, v_i, s_i)$ where individual i is included in the analysis because $v_i < s_i$. For many individuals $e_{ir} = v_i$. In the likelihood specification, this leads to a denominator term that quantifies the probability to be free of symptoms at the time of leaving Wuhan. Let $h'(e_l, e_r, v)$ denote the joint density of the observation points around the moment of infection and the time of leaving Wuhan.

Then the likelihood provided in (3.1) is replaced by

$$l(e_{il}, e_{ir}, v_i, s_i | v_i < s_i) = \frac{h'(e_{il}, e_{ir}, v_i) \int_{e_{il}}^{e_{ir}} g_i(u|e_{il}, e_{ir}) f(s_i - u) du}{\int_{e_{il}}^{e_{ir}} g_i(u|e_{il}, e_{ir})[1 - F(v_i - u)] du}. \tag{3.3}$$

Currently, there is no suitable R package available for this specific type of survival data. Pak *et al.* consider a similar type of data structure. They postulate a distribution for the time from infection to enrollment $V - E$ with density k [Pak, Liu, Ning, Gómez and Shen, 2020]. Assuming that $V - E$ and T are independent, the following likelihood is obtained

$$l(e_{il}, e_{ir}, v_i, s_i | v_i < s_i) = \frac{h(e_{il}, e_{ir}) \int_{v_i - e_{ir}}^{v_i - e_{il}} k(u) f(u + s_i - v_i) du}{\int_{v_i - e_{ir}}^{v_i - e_{il}} k(u)[1 - F(u)] du}.$$

They applied the data to a cohort study on HIV infection, allowing for right censored data with respect to symptom onset.

Qin *et al.* rightly acknowledge that the sampling mechanism of traveler data from Wuhan introduces length biased sampling [Qin et al., 2020]. They treated the incubation period as a renewal time and the duration from departure to symptom onset as forward time in a renewal process. This approach it is not suitable for our specific context [Arntzen et al., 2023].

3.4 Simulation setup

We performed a simulation study to investigate the effects of differential recall and the presence of left truncated data. To examine differential recall, we varied: the strength of differential recall; the implementation of differences in memory; whether the complete data or a subset was used in the analysis. To investigate how the presence of left truncated data affects the results, the following aspects were changed: the width of the exposure window; the distribution of infection risk: constant, increasing, or decreasing; whether or not to account for the presence of left truncation.

3.4.1 Data generation

In the following sections, we provide details about how the data were generated to study the effect of differential recall and the presence of left truncated data, on the estimate of the incubation time distribution. In each scenario, the incubation time (T) was generated from a Weibull distribution with parameter values based on estimates for SARS-CoV-2 during the early stages of the pandemic (median 5.4 days, 95th percentile 9.8 days) [Lauer et al., 2020]. Since we were not interested in the bias due to an incorrect parametric model, a Weibull distribution was assumed for estimation as well. One thousand data sets were generated in each scenario.

Differential recall

The basic idea of the data generation is sketched in Figure 3.1. For each individual i, we first generated a sequence of daily 'checkup' times, spanning from 1 to 20 days before the onset of symptoms (indicated by vertical tick marks). This approach builds upon the work of Dejardin and Lessaffre [2013]. These 'checkup' times act as observation points concerning infection status. They are forgotten with a certain probability (represented by crosses). As we remove observation times, we end up with a new, wider, exposure window. This means that we cannot employ the data generation approach used in our previous work, where we initially generated infection times uniformly within an exposure window and then generated an incubation time (Chapter 2). Instead, we generated times from symptom onset backwards to infection time as indicated by the arrow in Figure 3.1. This process allowed us to directly create interval censored time-to-event data, and we do not need to assume a constant infection risk within the exposure window.

We considered two different scenarios regarding incomplete memory. In scenario A, we assumed that individuals with longer incubation times tend to be more forgetful. Hence, the observation times depend on the incubation time, which violates assumption a) in Section 3.2. In scenario B, we assumed that forgetfulness increases as individuals have to look further back in time. Specifically, under scenario A, the probability of omission, i.e. missing a checkup time, increases with a person's incubation time. The probability of omission varies among individuals but remains the same for each checkup point within an individual. In scenario B, the probability of missing a checkup time increases as the exposure time moved further away from symptom onset. In this case, the probability varies with the timing of the checkup point, but remains the same across individuals. Additionally, we generated a subset of individuals (10%) with perfect recall of the time of infection.

The probability to remember was modeled as $e^{-\lambda d}$, where the parameter λ represents the differential recall rate and d the number of days that elapsed since the checkup time at the interview day. Different values for the strength of λ were used. We explored two estimation approaches: one that uses the complete data set (N = 500), while the other restricts the analysis to exposure windows narrower than 5 days, which we will refer to as the "subset" approach.

We describe how we obtained exposure information in the simulations using Figure 3.1, which illustrates the timelines of two individuals. Both individuals are infected (indicated by

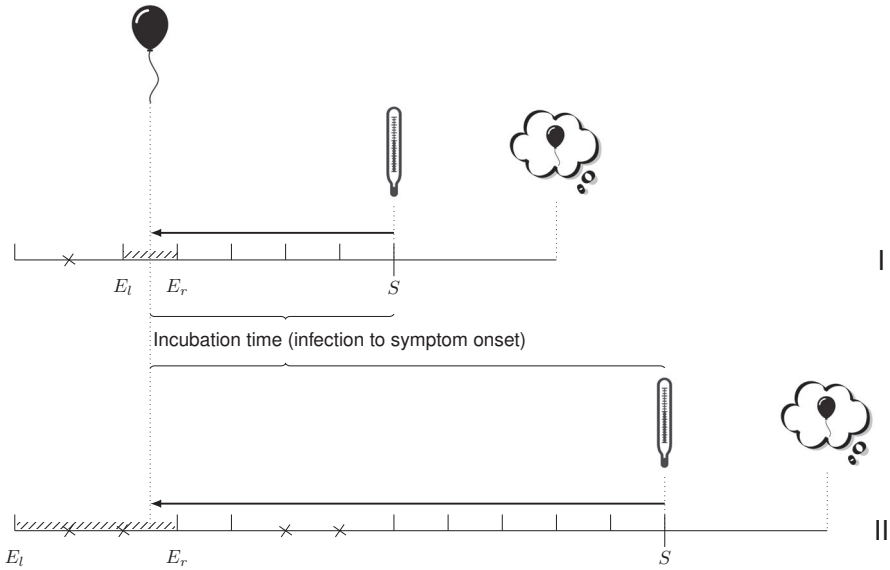


Figure 3.1: Illustration of memory decay. Graphical representation of incubation time and differential recall for two individuals I and II, both infected at the same party. During an interview conducted three days after symptom onset, both individuals were asked to recall their risk exposures. Individual I had a shorter incubation time (infection to symptom onset S) than individual II and therefore was exposed closer to symptom onset. In the simulation setup, decay of memories was mimicked by generating daily 'checkup' times (vertical tick lines) that may be forgotten (crosses) with certain probability as explained in the text. The observed exposure window consists of the last checkup time before infection (E_l) and the first checkup time after infection (E_r) that are not forgotten.

balloon) at the same event but individual I develops symptoms (indicated by thermometer) soon after infection, while individual II has a much longer incubation period. Upon diagnosis, both are asked to recall their risk exposure. Individual I was exposed more recently at the time of interview (indicated by thinking cloud).

We generated daily 'checkup' times represented by vertical tick lines. Memory decay is incorporated by considering a probability of omission (indicated by crosses) that either increases with incubation time (scenario A) or increases as the checkup times are longer ago (scenario B). The observed exposure window consists of the two memorised checkups closest to the moment of infection (E_l and E_r). In the example in Figure 3.1, the exposure window of individual II is wider than of individual I.

Left truncation

We generated data in a similar way as in Chapter 2, but with a different selection of exposure windows. Ten per cent had the moment of infection (E) observed exactly and they travelled on the day of infection, while the remaining 90% all had the same width of the exposure window (0 to E_r where E_r represents the preset width). We varied the width of this window among scenarios. Next, we generated the time of infection (E) within the exposure window, an incubation time (T) and a time of symptom onset ($S = E + T$). This generation process made three different assumptions with respect to the time of infection within the exposure window:

a) A constant risk of infection ($g(t) \sim U(E_l, E_r)$);

b) Exponential growth with a five-day doubling time of the incidence ($g(t) \propto e^{0.14t}$), which reflects the initial phase of the outbreak in Wuhan [Dorigatti et al., 2020];

c) A declining infection risk ($g(t) \propto p(1-p)^{t-1}$ where $p = 0.2$ on the interval $[E_l, E_r]$), which may represent household transmission.

We only included individuals who experienced symptom onset after the end of the exposure window, i.e. after leaving Wuhan ($S > E_r = V$). This leads to left truncated data as illustrated in Figure 3.2. In the figure, individual I had a shorter incubation time than individual II. Individual I developed symptoms (indicated by thermometer) before their scheduled departure from Wuhan (indicated by train), and remained in Wuhan. In contrast, individual II traveled while incubating and developed symptoms later. Individuals with $T < E_r - E$ were discarded. In this example, it means that individual I is excluded from the data, while individual II is included, i.e. observed in the data.

For each data set, we initially generated 50,000 observations, and then a random sample of 500 observations was selected satisfying the condition $S > E_r$.

3.4.2 Estimation

The method used for data with differential recall yields data sets containing exact, interval censored and right censored observations. Observations are right censored when there is no memorised checkup before infection, which may occur because we limited the maximum

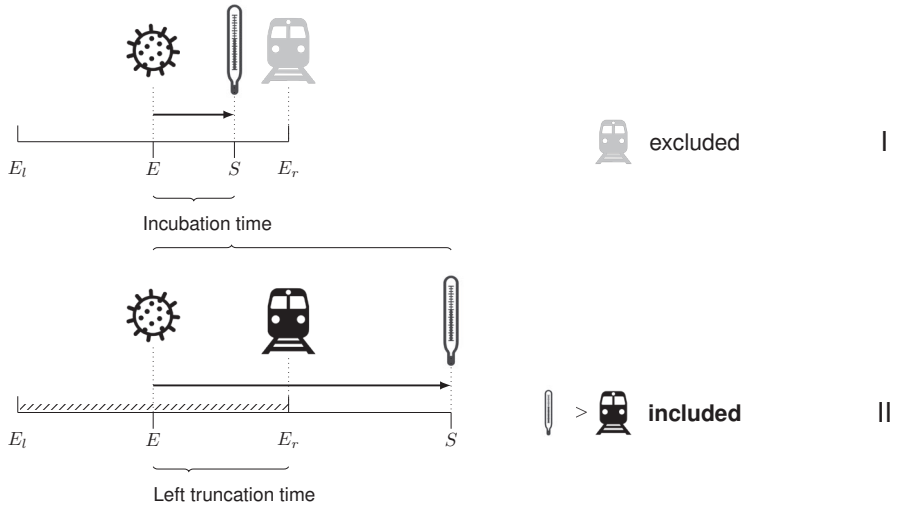


Figure 3.2: Illustration of left truncation. Two individuals were infected on the same day during the outbreak in Wuhan. Individual I had a shorter incubation time (infection to symptom onset, E to S) than individual II. Individual I and II planned to leave Wuhan at the same calendar date. However, individual I developed symptoms before the travel day; individual II developed symptoms after leaving Wuhan. Individual II is included in the data concerning travellers from Wuhan, with a left truncation time (interval) from infection (E) to travel day (E_r). Individual I is excluded from the data.

number of checkup points to 20. The `R` package `survival` was used to fit the appropriate models to these data sets, assuming a Weibull distribution.

For the simulations with left truncated data, we use the reversed time scale, which assumes a constant risk of infection. We also assume that the time of infection is known for the truncated part of the likelihood, which is not the case in practice. Hence instead of Equation (3.3), we maximized the "oracle" likelihood based on

$$l'(e_{il}, e_{ir}, s_i | e_{ir} < s_i) = \frac{F(s_i - e_{il}) - F(s_i - e_{ir})}{1 - F(e_{ir} - e_i)}. \tag{3.4}$$

This is somewhat artificial and merely serves to illustrate the problem with left truncated data, rather than to provide an actual solution. As the `survival` package does not incorporate the combination of interval censoring and left truncation, the `R` package `MixtureRegLTIC` [Chen et al., 2013] was used to fit an AFT to the data. The latter uses the extended generalized gamma (EGG) distribution, which was introduced by Farewell and Prentice [Farewell and Prentice, 1977] and includes the Weibull distribution as a special case.

3.4.3 Performance measures

The performance of the model across 1000 estimates of the median and 95% percentile of the incubation time distribution per scenario is summarized by the bias and the interquartile range (p25 and p75) of the deviations between true and estimated value. Additionally, for the simulations concerning left truncated data, the mean proportion of exact observations in the resulting data sets is provided.

While all runs for the data sets with differential recall provided a model fit, for the scenarios concerning left truncated data, the model did not converge for some of the runs. This issue is due to an artifact inherent in the simulation setup. Specifically, it may occur that an observation has a late entry time that exceeds the lower bound of the interval censored incubation time, i.e. $s_i - e_{ir} < e_{ir} - e_i$ in Equation (3.4). The `MixtureRegLTIC software` package was designed for observations with exactly observed time origin, interval censored endpoints and left truncation with respect to the endpoint. The percentage of invalid runs is shown in Supplement 3.7.2 and indicated by 'Inv.'.

3.4.4 Software

All analyses were performed in R version 4.1.1 [R Core Team, 2021] and R Studio version 2021.09.20 ("GhostOrchid") [RStudio Team, 2021] software environment, using the computing resources from the Academic Leiden Interdisciplinary Cluster Environment (ALICE) provided by Leiden University. The analysis code can be accessed via `www.github.com/vharntzen/TwoBiasesExposure`.

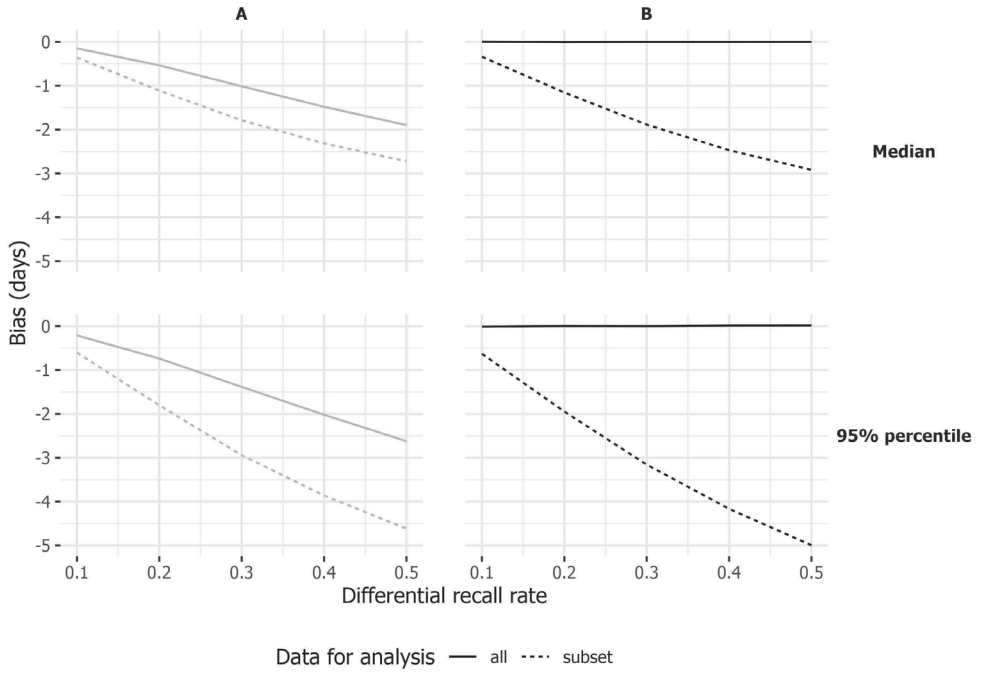
3.5 Results

All performance measures can be found in the tables in Supplement 3.7.1 and 3.7.2.

3.5.1 Differential recall

Memory decay was simulated using two different scenarios (A and B). Figure 3.3a visualizes the resulting bias (y-axis) for different percentiles (upper row: median; lower row: 95^{th} percentile) and approaches. These approaches include analyzing all observations (solid line) and analyzing a subset (dashed line).

Figure 3.3: Results of simulations concerning differential recall. The bias (y-axis) is presented, based on 1000 generated data sets, for the estimated medians (upper panel) and 95^{th} percentiles (lower panel) under different strengths of differential recall (x-axis). Two analysis approaches are considered: one using all 10 000 observations (solid line) and the other using a subset of narrow exposure windows (window width $\leq$ 5 days, dashed line). The recall probability per checkup time, as depicted in Figure 3.1, either depends on incubation time (scenario A, left panels) or backward time from symptom onset (scenario B, right panels).



In scenario A, both estimation approaches give biased estimates (Figure 3.3, left panel), because the distribution of non-omitted checkup times depends on the incubation time. Scenario B is unbiased if all data is analyzed, regardless of the rate of memory decay, since the distribution of non-omitted observation times is independent of the incubation time.

Using only observations with well-defined exposure (window width $\leq$ 5 days, on average 39% of the observations for a differential recall rate of 0.3, for both scenarios) gives a similar downward bias in both scenario A and B (Figure 3.3). In both scenarios, individuals with longer incubation times tend to have wider exposure windows. Therefore, restricting to narrow windows selectively includes those with shorter incubation times. The magnitude of this bias increases with more extreme levels of differential recall.

3.5.2 Left truncation

The bias when we corrected for left truncation in the analysis is shown in the left panel of Figure 3.4a, while the bias resulting from leaving truncation uncorrected is shown in the right panel. The figure visualises the results for different exposure window widths (x-axis), percentiles (upper panel: median; lower panel: 95^{th} percentile), and true infection risk distributions (line type).

When the risk of infection is constant on the exposure window (solid line) and left truncation is accounted for in the analysis (left panel), estimates are unbiased regardless of the exposure window width (x-axis). However, when left truncation is neglected (right panel), estimates exhibit an upward bias. This bias initially increases with exposure window width, followed by a decline until it appears to stabilize.

Under a decreasing risk of infection within the exposure window (represented by the dashed line), the bias approaches zero as the exposure window width increases. This is because only the non-truncated, exact observations remain for the analysis (travel on day of infection). The rationale behind this is as follows: infection is most likely to occur at the beginning of an individual's exposure window. The wider the window, the less likely it is for symptoms to develop after the end of the exposure window rather than within it. Hence, with the left truncation mechanism in place, it is less likely for such an observation to be included. Note that the absolute bias in the right panel of Figure 3.4 is smaller than in the left panel and it operates in the opposite direction. This difference is because the two components of bias in the right panel partially cancel each other out. There is an upward bias when left truncation is not accounted for and a downward bias due to the violation of constant risk of infection (assumption (b) in Section 3.2). In the left panel of Figure 3.4, only the latter component of bias is present (resulting in downward bias).

At the beginning of an outbreak, the cumulative infection incidence and, consequently, the risk of infection grows exponentially. When the risk is increasing, an upward bias is observed (as indicated by the dotted line in Figure 3.4), regardless of whether we corrected for truncation (left panel) or not (right panel). This bias increases with window width as the constant risk assumption is more strongly violated. In contrast to a decreasing risk (dashed line), both components of the bias point in the same upward direction. Without correction for truncation (right panel), the bias is larger than when left truncation is addressed in the analysis (left panel). Since infection is most likely to occur right before symptom onset,

a substantial portion of the data is used for analysis even when the exposure window width is large, preventing the bias from vanishing. Moreover, as exposure windows get wide, the bias plateaus rather than vanishes as exposure windows get even wider. Note that the bias remaining after correction for left truncation (left panel) is the same bias as observed in Chapter 2.

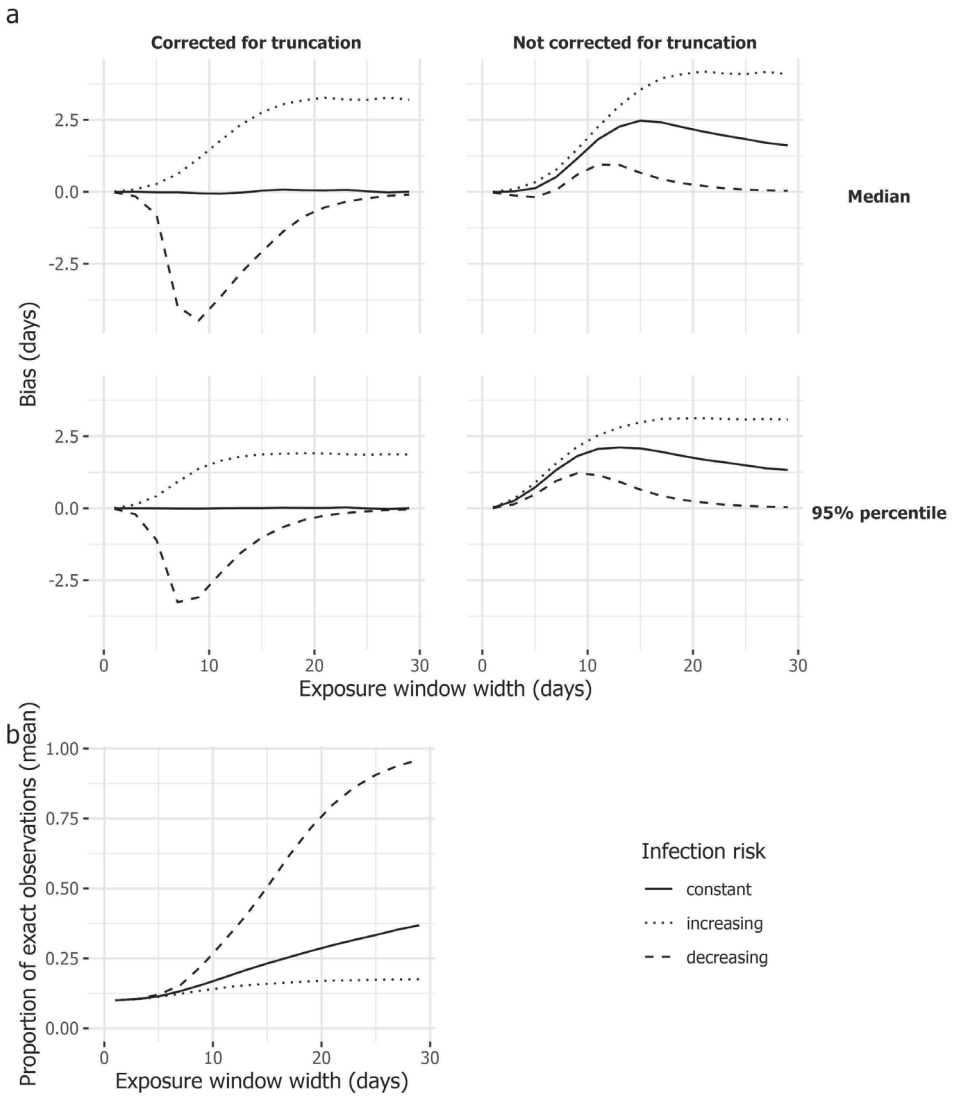
3.6 Discussion

Incubation time plays a critical role in informing policy makers during the early stages of an outbreak. However, accurate estimation is challenging due to limitations in the data, which is typically collected retrospectively through interviews with infected individuals regarding their exposure. In this study, we investigated the impact of two phenomena in SARS-CoV-2 contact tracing data that have been neglected in estimation: differential recall of exposure and left truncation. Our simulations revealed that, under the most plausible scenario, where the start- and endpoints of the exposure windows, as well as the inclusion criteria, are independent of incubation time (scenario B and analysing complete dataset), differential recall does not introduce bias in the estimates. However, when the analysis is restricted to individuals with well-defined exposure, incubation time tends to be underestimated. Neglecting left truncation in the analysis consistently leads to overestimation.

The value of our study lies in recognizing the sources of bias involved in incubation time estimation. The phenomenon of differential recall may also occur in other contexts where time-to-event data is observed, such as environmental or work-related exposure to toxic agents and the subsequent development of health conditions. However, verifying its presence in real-world scenarios can be a challenging task.

Although right truncation has been mentioned in previous papers on SARS-CoV-2 incubation time estimation [Linton et al., 2020], left truncation has mostly been overlooked. Qin and Deng did consider left truncation in their analysis of the Wuhan data [Qin et al., 2020; Deng et al., 2020]. However, we found that the method they proposed was not suitable for this particular context (Chapter 2, second simulation study). Since the exact moment of infection is not observed for most individuals, the same holds for the time from infection to entry into the study (i.c. leaving Wuhan). It is possible to adjust the likelihood to account for this specific problem by integrating over all possible infection moments within

Figure 3.4: Simulation results concerning left truncation. (**a**) The bias (y-axis) is presented, based on 1000 generated data sets, for the estimated medians (upper panel) and 95^{th} percentiles (lower panel) across various exposure window widths (x-axis; excluding the 10% of observations with exactly observed moment of infection). Three different scenarios for the risk of infection within the exposure window are considered: constant (solid lines); increasing (dotted line) or decreasing (dashed line). The analysis is performed with truncation incorporated (left panel) or without (right panel). (**b**) The mean proportion of exact observations (y-axis) in the data set used for analysis, for different exposure window widths (x-axis). The infection risk distributions (constant, increasing, decreasing) on the exposure window domain are represented by different line types.



the exposure window. We explored the method proposed by Pak *et al.* [2020] and the corresponding R software that they provided upon request, but did not include it because it assumes a distribution for the time between infection and travel.

The concept of differential recall of time-to-event data has received little attention in the literature. Our simulations show that neglecting this phenomenon does not introduce any bias when the distribution of observation points (i.e., start and end of exposure) is unrelated to the time-to-event distribution and we use the full data set. While we consider this independence assumption plausible in the context of retrospectively collected contact information, verifying it in reality is difficult since the moment of infection is interval censored at best. Future research is needed to develop an algorithm capable of distinguishing between the two situations in real data. Note that in fact, window width may depend on the number of risk contacts as well, which was beyond the scope of our study. Regardless of (non-)independence of the exposure window boundaries and incubation time, our simulations revealed that restricting the analysis to narrow exposure windows introduces bias. It is important to note that the analysis is usually restricted to observations with narrow exposure windows for a valid reason, specifically to mitigate bias resulting from the violation of the constant risk assumption, particularly during the exponential growth phase of an outbreak. Apart from preventing bias due to differential recall, also including the individuals with wider intervals increases the size of the typically small data set, thereby increasing statistical power and narrowing the width of the confidence intervals.

Another concern is that individuals with narrow exposure windows may not be representative of the entire population, but over-represent a group with shared characteristics such as a certain age, health status or attending a certain event with high transmission rates [WHO, 2003; McAloon et al., 2020]. If the incubation time distribution depends on such a characteristic, the resulting estimate is not representative for the entire population. An example is age, which was found to be related to memory in survey questions [Sudman and Bradburn, 1973], as in differential recall scenario A in our simulations. Software for analyzing data with an interval censored time origin rather than endpoint, where a more realistic distribution of the infection risk within the exposure window can be used using a population-wide estimate of the infection incidence, would circumvent the need to assume a constant risk of infection. This would eliminate the need to restrict the analysis to a well-defined subset.

Our study offers practical recommendations for researchers involved in estimation of incubation time. Firstly, caution is warranted when restricting the analysis to observations with narrow exposure windows. While this reduces bias resulting from the potential violation of the constant risk assumption, it may lead to underestimation of the incubation time distribution due to differential recall. If there is doubt whether differential recall plays a role, a sensitivity analysis comparing results with and without wide exposure windows is recommended. Secondly, researchers need to be aware that left truncation may be present in the data. We gave the specific example of data on SARS-CoV-2 infection based on individuals that left Wuhan. A scenario other than traveller data in which this may occur, is when infected individuals experience a high case fatality rate, and are ascertained by screening. Individuals with a short incubation time may tend to have deceased already, such that exposure information cannot be obtained anymore via retrospective interviews.

In a more general context, obtaining optimal estimates of the incubation time distribution requires comprehensive retrieval of exposure information. Typically, this information is obtained through retrospective interviews with detected cases, and these interviews should cover a sufficiently long period to capture all potential risk exposures. If the period is too short, the true infection may not fall within the given exposure window. Additionally, when the case definition assumes only a narrow range of potential incubation periods, implying a limited exposure period, longer incubation periods may go unnoticed. To prevent the latter problem from occurring, the incubation time could be excluded from the case definition, but this increases the risk of misdiagnoses. In other words, a less specific case definition complicates diagnosis. For example, in the case of influenza and corona viruses, for which the clinical presentation shows strong similarities, including the incubation period in the case definition is useful for distinguishing between these respiratory infections [Nishiura et al., 2012].

Our study discusses two overlooked sources of bias in incubation time estimation, acknowledging that resolving them in practice may not be straightforward. We provide practical recommendations for researchers engaged in estimating incubation time.

Figures

Thermometer https://www.vecteezy.com/vector-art/10405665-vector-illustration-of-mercury-thermometer-icon **Thinking balloon**

https://www.vecteezy.com/vector-art/10651867-collection-set-of-blank-black-and-white-hand-drawing-speech-

bubble-balloon-think-speak-talk-text-box-banner-flat-vector-illustration-design **Balloon** https://www.vecteezy.com/vector-art/577902-flying-vector-festive-balloons-shiny-with-glossy-balloons-for-holiday **Train** https://commons.wikimedia.org/wiki/File:BSicon_TrainCHN.svg **Virus** https://commons.wikimedia.org/wiki/File:Coronavirus_icon.svg

3.7 Supplementary material

3.7.1 Simulation results concerning differential recall

Table 3.1: Results of simulations concerning differential recall. Each row summarizes 1000 runs, each with data set size N = 500. *Abbreviations* Differential recall rate (Recall rate); Percentile estimand (Perc.); Data for analysis (Data): all 10 000 observations or the subset of observations with an exposure window width smaller than 5 days; Scenario to differential recall: A: probability of missing checkup time depends on incubation time; B: probability depends on backward time of infection from symptom onset; p25/p75: percentiles of the distribution of deviations between true and estimated percentile.

Recall rate	Perc.	Data	Scenario A			Scenario B		
			Bias	*p25*	*p75*	*Bias*	*p25*	*p75*
0.1	50	all	-0.15	-0.24	-0.07	0.00	-0.08	0.08
0.2			-0.54	-0.62	-0.45	0.00	-0.09	0.08
0.3			-1.01	-1.11	-0.92	0.00	-0.10	0.10
0.4			-1.48	-1.58	-1.38	0.00	-0.14	0.14
0.5			-1.90	-2.01	-1.78	0.00	-0.18	0.16
0.1		subset	-0.36	-0.45	-0.27	-0.34	-0.42	-0.26
0.2			-1.11	-1.21	-1.01	-1.16	-1.25	-1.06
0.3			-1.79	-1.90	-1.68	-1.89	-1.99	-1.78
0.4			-2.32	-2.43	-2.20	-2.47	-2.58	-2.36
0.5			-2.72	-2.86	-2.59	-2.92	-3.04	-2.79
0.1	95	all	-0.21	-0.37	-0.07	-0.01	-0.16	0.15
0.2			-0.74	-0.92	-0.56	0.00	-0.20	0.21
0.3			-1.38	-1.59	-1.18	0.00	-0.28	0.26
0.4			-2.02	-2.28	-1.78	0.01	-0.39	0.38
0.5			-2.62	-2.91	-2.36	0.02	-0.52	0.49
0.1		subset	-0.60	-0.75	-0.46	-0.63	-0.78	-0.48
0.2			-1.80	-1.97	-1.63	-1.95	-2.12	-1.77
0.3			-2.94	-3.16	-2.72	-3.16	-3.36	-2.95
0.4			-3.86	-4.13	-3.59	-4.16	-4.42	-3.92
0.5			-4.62	-4.93	-4.31	-5.00	-5.29	-4.71

3.7.2 Simulation results concerning left truncation

Table 3.2: Results of simulations concerning left truncation. Each row summarizes 1000 runs, each with data set size N = 500. *Abbreviations* Exposure window width in days (Width); Percentile estimand (Perc.); Infection risk distribution on exposure window (Risk): constant (cons.), exponentially increasing (incr.) or decreasing (decr.); mean proportion of exact observations in the final data sets (PE); p25/p75 quantiles of the distribution of deviations between true and estimated percentile; percentage of invalid runs, i.e. for which the model did not converge (Inv.).

Width	Perc.	Risk	PE	Corrected for truncation				Not corrected for truncation			
				Bias	*p25*	*p75*	*Inv.*	*Bias*	*p25*	*p75*	*Inv.*
3	50	cons.	0.10	0.00	-0.10	0.10	0.0	0.26	0.18	0.33	0.3
7			0.13	-0.01	-0.22	0.24	0.0	1.33	1.24	1.41	0.0
11			0.18	0.00	-0.29	0.32	1.7	2.06	1.94	2.19	0.0
17			0.25	0.02	-0.22	0.28	23.7	1.96	1.81	2.11	0.0
25			0.33	-0.01	-0.20	0.20	18.1	1.49	1.36	1.61	0.0
3		incr.	0.10	0.13	0.03	0.22	0.0	0.33	0.25	0.42	0.0
7			0.12	0.92	0.76	1.07	0.0	1.55	1.46	1.64	0.0
11			0.15	1.64	1.43	1.86	1.1	2.53	2.40	2.66	0.0
17			0.16	1.90	1.68	2.12	10	3.10	2.94	3.27	0.0
25			0.17	1.86	1.59	2.11	8.6	3.08	2.85	3.29	0.0
3		decr.	0.10	-0.21	-0.31	-0.11	0.0	0.14	0.06	0.21	0.2
7			0.15	-3.26	-3.82	-2.74	12.1	0.95	0.86	1.04	0.1
11			0.30	-2.29	-2.63	-1.95	98.3	1.16	1.05	1.26	0.0
17			0.62	-0.65	-0.78	-0.53	44.5	0.43	0.35	0.52	0.0
25			0.91	-0.10	-0.19	-0.02	11.1	0.08	0.00	0.16	0.0
3	95	cons.	0.10	0.00	-0.12	0.12	0.0	0.02	-0.11	0.13	0.3
7			0.13	-0.02	-0.19	0.18	0.0	0.51	0.38	0.64	0.0
11			0.18	-0.07	-0.43	0.38	1.7	1.82	1.68	1.98	0.0
17			0.25	0.08	-0.38	0.58	23.7	2.41	2.11	2.71	0.0
25			0.33	0.02	-0.32	0.38	18.1	1.83	1.57	2.07	0.0
3		incr.	0.10	0.09	-0.03	0.21	0.0	0.09	-0.02	0.22	0.0
7			0.12	0.62	0.49	0.75	0.0	0.76	0.64	0.89	0.0
11			0.15	1.76	1.54	1.99	1.1	2.26	2.10	2.41	0.0
17			0.16	3.04	2.68	3.44	10.0	3.94	3.61	4.28	0.0
25			0.17	3.19	2.62	3.76	8.6	4.09	3.58	4.55	0.0
3		decr.	0.10	-0.16	-0.29	-0.04	0.0	-0.13	-0.25	-0.01	0.2
7			0.15	-3.98	-4.87	-3.02	12.1	0.06	-0.07	0.19	0.1
11			0.30	-3.69	-4.18	-3.33	98.3	0.95	0.81	1.09	0.0
17			0.62	-1.39	-1.62	-1.14	44.5	0.44	0.29	0.59	0.0
25			0.91	-0.23	-0.35	-0.08	11.1	0.07	-0.05	0.2	0.0

This chapter will be submitted soon as Vera H. Arntzen, Manh Nguyen Duc, Marta Fiocco, Lan Truong Thi Thanh, Tam Nguyen Hoai Thao, Buu Mai Thanh, Tu-Anh Nguyen, Nhat Le Thanh Hoang, Marc Choisy, Lam Phung Khanh, Nga Le Hong and Ronald B. Geskus, on behalf of the Covid-19 modelling team Oxford University Clinical Research Unit, Vietnam[+]. The latency time of the SARS-CoV-2 Delta variant in naive individuals from Vietnam.

[+] *Duc Du Hong, Lam Phung Khanh, Leigh Jones, Marc Choisy, Nhat Le Thanh Hoang, Ronald Geskus, Sonia Lewycka, Thomas Kesteman, Trinh Dong Huu Khanh, Tung Trinh Son, Manh Nguyen Duc, Nguyet Nguyen Thi Minh, Thinh Ong Phuc, Trang Duong Thuy, Lieu Tran Thi Bich, Maia Rabaa.*

The latency time of the SARS-CoV-2 Delta variant in naive individuals from Vietnam

Contents

Abstract

The latency time (infection to start-of-infectiousness) is one of the factors that determines the efforts required to control an infectious disease. Yet, estimates of the SARS-CoV-2 latency time remain sparse. Information on the endpoint requires repeated testing for viremia. Moreover, proper estimation is challenging as both the start- and endpoint are typically interval censored and long latency times may be underrepresented.

We collected detailed exposure information from public health reports produced during an outbreak with the SARS-CoV-2 Delta variant in Ho Chi Minh City, Vietnam, from May 2021 onwards. Using a tailor-made digital form and application to make reliable choices, we distilled a comprehensive data set with information on exposure and test results from 1951 individuals. This is the first data set of its kind outside of China, collected in the absence of large-scale vaccination or earlier transmission.

Our analysis is unique as we respect the doubly interval censored nature of the observations and make realistic assumptions regarding the infection risk (exponential growth) and the latency time distribution (generalized gamma), while also addressing truncation due to sampling cutoff and a finite quarantine length. Our implementation using the Bayesian program JAGS is freely available in the R package `doublIn`.

The estimated mean SARS-CoV-2 Delta variant's latency time was 3.22 (95% Credible Interval 2.89; 3.55) days; the median 1.81 (95% CrI 1.44; 2.16); the 95% percentile 10.98 (95% CrI 9.91; 12.41). Sensitivity analyses showed that the estimates depend strongly on the model assumptions.

Our results indicate that the SARS-CoV-2 Delta latency time may be shorter than previously assumed, requiring timely identification of infecteds. This may explain why the combination of contact tracing and quarantine was a more successful strategy for variants characterized by a longer latency time.

Keywords latency time □ SARS-CoV-2 □ contact tracing data □ doubly interval censoring □ quarantine

4.1 Introduction

Understanding the natural history of SARS-CoV-2 variants has been essential to shape public health measures optimally. Viral shedding patterns differ per SARS-CoV-2 variant [Puhach et al., 2023], and it naturally follows that the latency time (infection to start of infectiousness) is variant-specific as well. Quarantine length is one of the public health measures that is ideally informed by latency time. However, given the sparsity of such estimates, the choice of quarantine length is typically based on estimates of the incubation time (infection to symptom onset) distribution, even though 40.5% of the SARS-CoV-2 infected individuals do not develop symptoms [Ma et al., 2021]. Knowledge of the latency time is essential, as quarantining is cumbersome for individuals and demanding in terms of logistics, especially when performed in a designated facility, as was the policy in countries like Vietnam. More generally, estimates of the latency time indicate how extensive a public health response needs to be to control the spread of SARS-CoV-2 [Demers et al., 2023]. The shorter the latency time, the less time there is to identify and isolate new cases before they potentially transmit to others. Together with the basic reproduction number R_0 and the amount of non-symptomatic transmission, the latency time determines the required effort to control spread of an infectious disease.

Literature on the SARS-CoV-2 latency time distribution is scarce [Kang et al., 2022; Ma et al., 2022; Jiang et al., 2023; Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021; Li et al., 2024]. As the roll-out of vaccination campaigns coincided with the emergence of novel variants, one cannot easily compare the estimates between variants. Current estimates of latency time are from China and concern either the Delta variant in the presence of partial vaccination [Kang et al., 2022; Ma et al., 2022; Li et al., 2024], or individuals of which both the vaccination status and the variant are unknown [Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021]. Using a unique data set originating from Vietnam, we present the first estimate of the latency time of the SARS-CoV-2 Delta variant in mostly naive individuals.

Data to estimate latency time is difficult to collect, as both infection and start of infectiousness are not observed directly. Exposure information is typically collected by retrospectively interviewing notified cases as part of a contract tracing policy. When contact tracing is implemented such data is usually available, as we know from the availability of

estimates of the incubation time (infection to symptom onset). However, information about the endpoint of the latency time, i.e. the start of infectiousness, is rare. As we cannot observe the start of infectiousness, one typically chooses the start of RNA shedding as a proxy [Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021; Kang et al., 2022; Li et al., 2024], which we refer to as start-of-shedding in this Chapter. When an individual tests positive by a PCR- or rapid antigen test, we assume that shedding has started. Hence, we assume that shedding started between the last negative test (when available) and the first positive test. Latency time can only be estimated if PCR- or antigen tests are performed repeatedly during follow-up. Since Vietnam adopted the policy of quarantining in facilities allocated by the government, PCR- and antigen tests were performed repeatedly during follow-up; therefore it is possible to estimate the latency time distribution using this data.

Observations of latency time are typically doubly interval censored, which means that both the start- and endpoint are at best known to fall in specific time windows. This complicates the analysis. To simplify estimation, common practice is to assume that the risk of infection within the exposure window is constant and that the latency time follows a gamma, lognormal or Weibull distribution. In previous work, we found that this can introduce bias in the estimates (Chapter 2). In this study, we relax these assumptions by considering alternative infection risk distributions using the observed trends in first positive test results in the population and using the more flexible generalized gamma distribution for the latency time, which includes the commonly used distributions as a special case. Also, we address right truncation to prevent sampling bias from two phenomena. First, longer latency times may be missed due to ongoing, increasing transmission during data collection [Chen et al., 2022; Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021]. Second, as quarantine has a definite duration, some individuals may have started shedding afterwards. As far as we are aware, addressing the latter source of bias is a novelty in SARS-CoV-2 latency time estimation.

This paper is organized as follows. In Section 4.2, we provide a short overview about SARS-CoV-2 latency time estimates from the literature. In Section 4.3 information about data collection, extraction and cleaning is discussed. The methodology used and the results are presented in Sections 4.4 and 4.5 respectively. The article ends with a discussion where the implications of our findings and future directions of research are outlined.

4.2 Literature

Studies estimating the SARS-CoV-2 latency time are scarce. We found five studies that directly estimated the SARS-CoV-2 latency time, i.e. not inferring it from a model. Unfortunately, one paper by Jiang [Jiang et al., 2023], who is shared first author on one of the other papers [Li et al., 2024], was in Chinese and could not be accessed via our university.

All estimates of the SARS-CoV-2 latency time were based on contact tracing data from China, including repeated PCR-test results [Kang et al., 2022; Ma et al., 2022; Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021; Jiang et al., 2023; Li et al., 2024].

The estimates concerned the Delta variant (n = 93 [Kang et al., 2022], n = 40 [Ma et al., 2022] and n = 672 [Li et al., 2024]), the Omicron variant (n = 467 [Jiang et al., 2023] and n = 885 [Li et al., 2024]) or unknown variant(s) (n = 177 [Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021]). In two studies, the majority of the included individuals was vaccinated (70% [Ma et al., 2022], 57% for Delta and 92% for Omicron infected individuals [Li et al., 2024]) and in another at least part of the individuals was vaccinated [Kang et al., 2022]; for the other two studies vaccination status was unclear ([Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021] [Jiang et al., 2023]). One study restricted the analysis to those with a single-day exposure [Ma et al., 2022], which simplifies analysis but may not be a representative selection [Li, Zhang, Peng, Gao, Jing, Wang, Ren, Xu and Wang, 2021].

4.3 Data

4.3.1 Spatiotemporal context

Because of its long-stretched border with China and limited ICU capacities, Vietnam initially strived to prevent any SARS-CoV-2 outbreak [Thai et al., 2020; Cobelens and Harris, 2020]. The country's public health response comprehended early and stringent policy measures, which included almost complete border closure, extensive contact tracing and quarantining. In 2020, the country experienced only one two-week period of national lockdown in April and a couple of local lockdowns. For an overview of the Vietnamese approach to the pandemic during its initial phase, we refer to Tan *et al.* [Tan, 2021].

Extensive contact tracing was employed, using an epidemiological classification system that is referred to as the 'F-system' [Hardy et al., 2020]. Direct (F_1) contacts and secondary (F_2) contacts (contacts of contacts) of confirmed cases (F_0) were actively traced. Contact tracing typically concerned the period from 14 to 21 days before testing positive or symptom onset of the confirmed case, whichever event occurred earlier, until the time of quarantine or the time of investigation. Contact tracing is bidirectional; it aims to find infection chains (infectees) from the F_0 onwards, but it also aims to find the source who infected the F_0. Attendance at events that are known to have led to transmission is explicitly asked for. For F_1 persons, besides finding possible (i.e. in case F_1 is infected) infectees (F_2), the public health staff also ask for the last date of exposure to the original case (F_0) to define the length of their quarantine. Direct contacts (F_1) as well as the rare incoming passengers that tested negative were required to quarantine in a government-allocated quarantine facility. Individuals were transferred to the hospital upon testing positive during quarantine. Secondary contacts (F_2) quarantined at home. Quarantine was for a minimum period of 14 days since the last date of exposure. For F_1 contacts and the incoming passengers it went up to three weeks from May 5^{th} to July 14^{th}, 2021 [Vietnam Center for Disease Control, 2021]. They ended their quarantine if they still tested negative at the end of this period. Once an F_1 tests positive, the person will be moved to the hospital, and all the contacts move from F_2 to F_1 status.

Until May 2021, Ho Chi Minh City in the south of Vietnam had barely reported any community transmission, because the few small-scale outbreaks were effectively contained. Community transmission considerably increased from mid May 2021 onwards. Infections were predominantly, and from June onwards solely, from the Delta variant, specifically the AY.57 lineage [Tam et al., 2023]. Vaccination campaigns still needed to be rolled out on a large scale. Before July 2021, Ho Chi Minh City received 923 050 vaccination doses [City Press Center Ho Chi Minh, 2021], on a city population of more than 10 million. In Vietnam in general, a populous country of 97.47 million, the daily number of administered vaccination doses only started to increase around June 14th, 2021, with total numbers increasing from 1.55 million to 4.06 million doses on July 12th [Our World in Data, 2024]. As such, we can assume that most individuals who were infected with SARS-CoV-2 during this period were naive, i.e. they did not contract SARS-CoV-2 before, and neither received vaccination. The Delta strain outbreak in Vietnam is thought to originate from a single

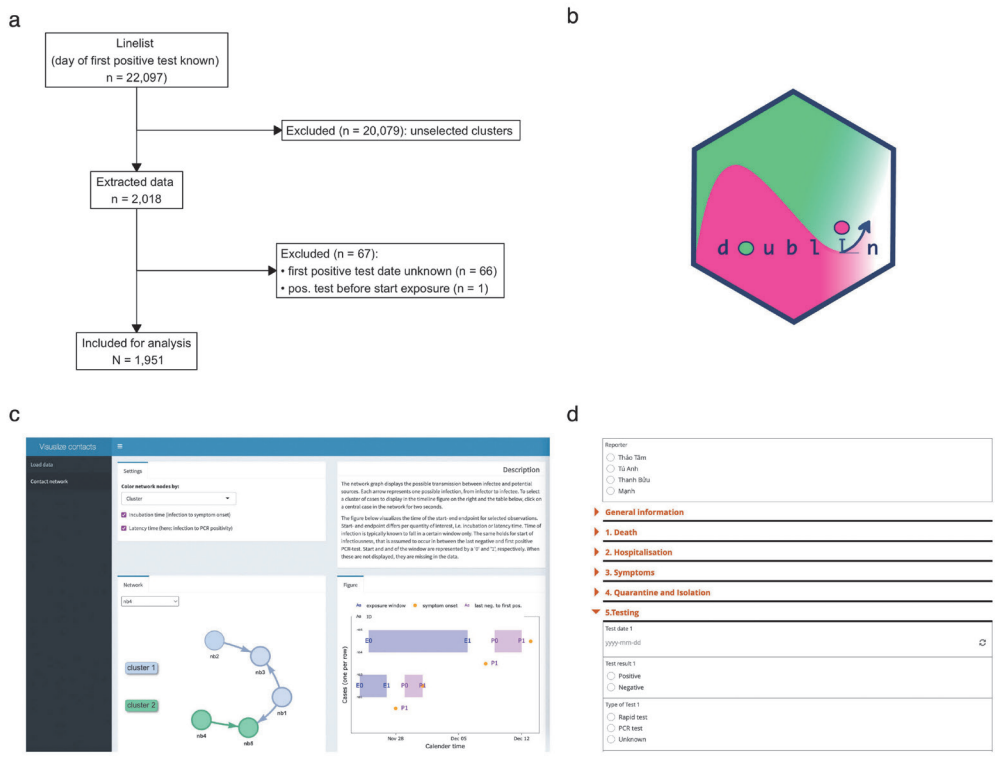


Figure 4.1: Illustration of the data collection process. Panel (**a**) shows the data flow from line list to data. Panel (**b**) shows the icon of the R package 'doublIn'. Panel (**c**) shows some of the R Shiny app functionalities (example data). Panel (**d**) shows the Kobo toolbox form used for data entry.

introduction in the second half of April 2021 [Tam et al., 2023].

The 2021 Ho Chi Minh City outbreak provides a rich data set that allows us to estimate the latency time of individuals for the Delta variant, as detailed exposure information and repeated test results are available. The line listing (data frame with one individual case per row) covers nearly all of the infections in this geographical region. The data used for analysis is representative of the population regarding the latency time due to the cluster-based sampling from the line listing, similar to snowball-sampling.

4.3.2 Data flow

Data in this study originates from several documents, provided by the Ho Chi Minh City Center for Disease Control (HCDC), which is the public health institute in Ho Chi Minh City. Flowchart 4.1a gives an overview. The starting point was a line list provided by HCDC containing basic information on 22097 PCR-confirmed cases, collected in Ho Chi Minh City between April 29th and July 15th, 2021. These collection dates refer to the day on which the case was made public; with the exception of one individual (excluded) the first positive test of any type (PCR or rapid antigen) always occurred before publication of the case, and typically three days before (34.2% of the individuals, n = 7576). The exact day of infection was generally unknown. Additional case information was retrieved from public health reports compiled as part of the contact tracing policy. We used 2827 public health reports in MS Word files. Each report covered one or multiple PCR-confirmed cases and gave more detailed information about exposure to potential sources. This information was obtained from individual case questionnaires that were taken upon confirmation of infection. We received the files in different folders, mostly but not exclusively covering a specific cluster of transmission. For simplicity, in the remainder of this paper we refer to 'cluster' as cases that were in the same folder. In July 2021, as the city saw a dramatic increase in the number of new cases, the contact tracing system got overwhelmed, and as a consequence not all cases could be described (timely) in public health reports.

4.3.3 Data extraction

During four months (from June 20th, 2022 to October 20th, 2022), four researchers (three from HCDC, one from the Oxford University Clinical Research Unit (OUCRU)) extracted the relevant case information. The researchers used a list with all clusters to divide their work. If a researcher found a link between the respective cluster and another cluster, the researcher would work on the related cluster next. For training purposes, at the very beginning, all four researchers worked on the same cluster of roughly 80 individuals. Then, they discussed their findings and aligned the way they retrieved data. Afterwards, the researchers started working independently, while the coordinating researcher from OUCRU performed random cross-checks of the data retrieved by the other researchers in the early stage of their work, again to align their work further. We included 75 of more than 100

clusters described by the public health reports in our data.

To support the data extraction from the public health reports, we developed a simple KoboToolbox form [KoBoToolbox, 2022] that facilitated manual data entry and prevented typos as much as possible. When a researcher read a report, they filled the relevant information into the form for those individuals that satisfied the following criteria: at least one potential source was mentioned in the report or at least one of two dates, namely the test date (PCR- or antigen test) or a day of onset of any symptom. When working on one cluster, the researcher typically drew all the potential transmission trees on paper, to assist with choosing the exposure periods.

If there was uncertainty in exposure, we considered two assumption in case one led to a narrower window than another. For instance, consider cases A and B residing together. They were exposed to a possible source between t_1 and t_2 and separated from each other at a later time t_3. Assuming that A and B contracted the infection from the same source led to the narrower window of t_1 to t_2. In the broader time frame of t_1 to t_3, we considered the possibility that A was initially infected from the source and subsequently transmitted the infection to B or vice versa.

To check and support the choice of these exposure windows, we developed an R Shiny application that visualizes the contact network based on entered data and the individual exposure information. We utilized the `KoboconnectR` R package [Sen, 2023] to directly import data collected from KoboToolbox that contained one row per infectee into R, facilitating visualization within our application.

Lastly, for the included cases, we merged the data obtained using KoboToolbox with the respective information of the initial line list, which had more complete information on testing.

Figure 4.1c and d give a glimpse of the KoboToolbox form and the Shiny app, respectively. The KoboToolbox form (Figure 4.1d) is available via `https://ee.kobotoolbox.org/x/dcXRd59G`. It supports entry of all relevant information from the public health reports, such as risk contacts and test results. Drop-down menus reduce the risk of data entry errors. We used the Shiny app (Figure 4.1c) to check if the extracted data was realistic. The software for both tools is available, with guidance and example data via `https://github.com/manhnguy/Contact-Tracing-for-Respiratory-Transmitted-Diseases`.

4.3.4 Cleaning

Our software suits observations for whom a) start of both windows (exposure and start-of-shedding) are known; b) both window widths (end minus start) are positive and non-zero; c) exposure coincides with, or occurs before the end of the start-of-shedding window.

For the vast majority (N = 2797, 94.8%) of all SARS-CoV-2 tests (N = 2950) performed during our study period on individuals in the extracted data (N = 2018, Figure 4.4b), the type of test was unknown, i.e. either a PCR or a rapid antigen test. We defined the first positive test date as the first day an individual tested positive for SARS-CoV-2 on *either* test. The last negative test date was defined as the day of the previous test of *either* type. For all individuals, the SARS-CoV-2 infection was confirmed by PCR test at some point during follow-up. We selected (Figure 4.1a) those observations for whom we had a positive test day, as well as at least one of the exposure window limits. For 1968 individuals out of 2018 for whom data was extracted, at least one test result was observed (Figure 4.4b). For 1952 individuals, the first positive test date was known (Figure 4.1a). We excluded one case for which the first positive test occurred before any exposure window (type 1 or 2) started. Hence, we included 1951 for analysis.

To comply with the requirements a-c, we prepared the data for analysis in six steps. We strived to include all possible information carried by the observations. We took the following steps in chronological order:

(i) if the end of exposure (E_r) is missing, set it to the day of the first positive test;

(ii) if the start of exposure (E_l) is missing, set it to the probable start of the Delta outbreak (April 30, 2021)

(iii) if the last negative test date (S_l) is missing, set it to the start of the exposure window ($S_l \leftarrow E_l$);

(iv) guarantee doubly interval-censored observations by considering a window width of at least one day in case date of infection or start-of-shedding were known, by adding or extracting half a day, respectively ($E_l - 0.5; E_r + 0.5; S_l - 0.5; S_r + 0.5$);

(v) if the exposure window ends after first positive test ($E_r > S_r$), let the end of exposure equal the first positive test date ($E_r \leftarrow S_r$);

(vi) if the last negative test occurs before start of exposure ($S_l < E_l$), set the last negative test to the start of exposure ($S_l \leftarrow E_l$).

Steps (i) to (vi) provided the subset of observations for analysis. Regarding step (v), note that we can naturally assume the individual to be infected by that time. We applied these rules separately for the two different ways exposure windows were determined.

4.4 Methods

4.4.1 Likelihood for doubly interval censored observations

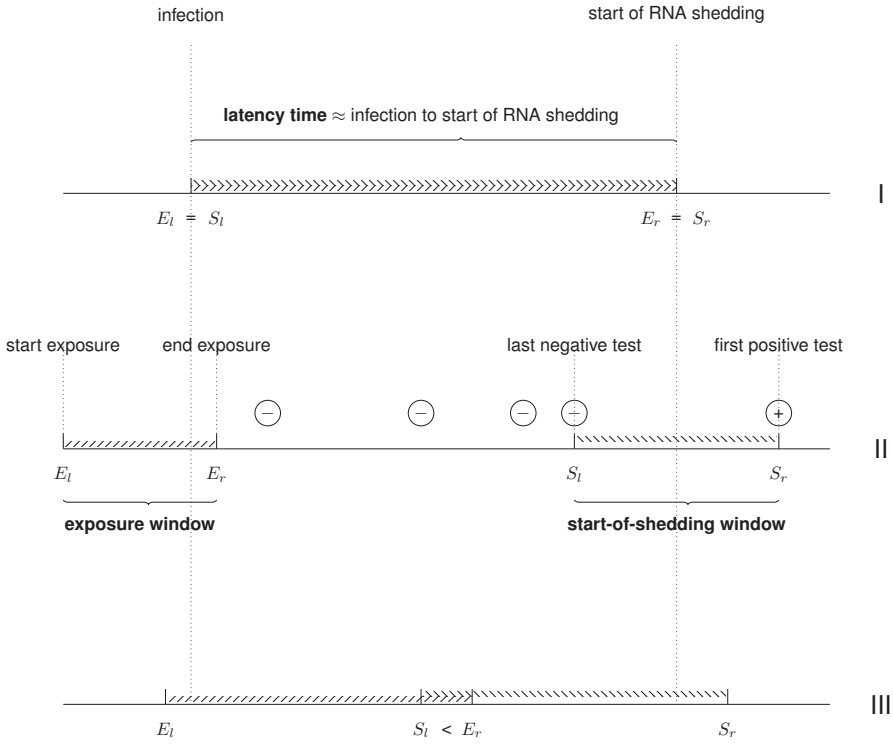


Figure 4.2: Illustration of observations of the latency time. Data representation of three individuals with an equal latency time which are observed differently due to double interval censoring. The window containing the origin is referred to as the exposure window, from the first possible moment of infection (E_l) to the last one (E_r). The boundaries of the window in which RNA shedding starts are the last negative test result and the first positive test result, where 'positive' refers to the detected presence of SARS-CoV-2 RNA (or antigen). The exposure window and start-of-shedding window may completely overlap ($E_l = S_I$ and $E_r = S_r$, indicated by I), not overlap ($E_r < S_l$, indicated by II) or partially overlap ($E_r > S_l$, indicated by III).

Building on previous work [Ramjith et al., 2022; Reich et al., 2009], the data representation of an observation of latency time consists of an exposure window and a start-of-shedding window that may completely coincide (I), not overlap (II), or partially overlap (III), as shown in Figure 4.2.

We first assume the setting without truncation. For an infected individual i $(i = 1, ..., N)$, let E_{il} and E_{ir} be the boundaries of the exposure window in calendar time; likewise, let S_l and S_r denote the start-of-shedding window.

Assume that the distributions of calendar time of infection (E) and latency time (E to S) are independent. When the exposure window ends before or on the same day as the start-of-shedding window begins (i.e. no overlap, Figure 1, type II), the contribution to the likelihood of the latency time for individual i is

$$l^{II} = \int_{e_{il}}^{e_{ir}} \int_{s_{il}}^{s_{ir}} g(e|e_{il}, e_{ir}) f(s-e) ds de \tag{4.1}$$

where $g(e|e_{il}, e_{ir})$ and $f(s-e)$ denote the probability density function of infection given the exposure window and the latency time respectively.

The likelihood contribution for an observation with complete overlap of the exposure and start-of-shedding windows (Figure 4.2, type I) is

$$l^{I} = \int_{e=e_{il}}^{e_{ir}} \int_{s=e}^{e_{ir}} g(e|e_{il}, e_{ir}) f(s-e) de ds. \tag{4.2}$$

When both windows partially overlap (i.e. Figure 4.2, type III) the likelihood $l(e_{il} \leq s_{il} \leq e_{ir} \leq s_{ir}$ can be decomposed in three parts using Equations 4.2 and 4.1. For details, we refer to the figure in Supplement 4.7.5.

4.4.2 Infection risk within the exposure window

The distribution of $g(e|e_{il}, e_{ir})$ depends on the assumptions made for the risk of infection within the exposure window $[e_l, e_r]$. A typical choice is a uniform distribution $g(e|e_{il}, e_{ir}) \sim Unif(e_{il}, e_{ir})$, implying that the infection risk is constant over time. However, the start of an outbreak is characterised by exponential growth and in previous work [Arntzen et al., 2023], we showed that violation of this assumption leads to biased estimates. Therefore, besides a constant risk, we explore an alternative assumption.

We assume an exponential growth for the infection risk at the population level. The incidence of new infections i is assumed to grow as $i(t) = i(t_0)e^{rt}$ where $i(t_0)$ is the

incidence when one started counting cases from the outbreak ($t = t_0$) and r is the epidemic growth rate (per day) with a corresponding doubling time of $\frac{ln(2)}{r}$ [Dorigatti et al., 2020]).

Estimates of growth rate and doubling time are obtained using the `R` package`incidence` [Kamvar et al., 2019]. We log-regressed the case incidence per day ($log[i(t)] = rt + log[i(t_0)]$) where we excluded all days on which the incidence was underreported (from July 12th onwards) to obtain the epidemic growth rate and doubling time, and corresponding Standard Error (SE). In the analysis we assume a normally distributed exponential growth factor with mean r and variance SE^2. Following Xin *et al.* [2021], we performed sensitivity analyses where we consider a less and more extreme scenario by extracting and adding $\approx$ 2 SEs to the estimated r, respectively. For the above-described approach, we assume that few infections will be unobserved in the line list, given the extensiveness of contact tracing.

4.4.3 Flexible distribution of latency time

The common practice in the literature is to fit three parametric distributions (gamma, lognormal and Weibull) and select one model by AIC (frequentist paradigm) or LOO IC (Bayesian paradigm). Especially when the interest is in the right-hand tail of the distribution, relaxing this assumption is recommended (Chapter 2). As the latency time may inform quarantine length, for which one often looks at estimates of the upper percentiles, we additionally fitted a generalized gamma distribution, first introduced by Stacy and Mihram [1965]; gamma, lognormal and Weibull distribution are special cases. For details on the parameterisation, we refer to Supplement 4.7.1. To implement the generalized gamma distribution, we used the parameterisation as proposed by Stacy and Mihram. As far as we know, there is only one earlier study on incubation and latency time estimation that utilized a generalized gamma distribution [Olivera and Muñoz, 2024].

4.4.4 Truncation

Our interest is to estimate the intrinsic latency time distribution [Park et al., 2024], which is representative for all individuals under stable conditions. However, due to several factors the likelihood (Eq. 4.2-4.3) does not directly concern the intrinsic distribution of interest in our data. Specifically, (i) the repeated testing for SARS-CoV-2 infection typically ended upon discharge from quarantine; (ii) cases that were made public after July 15th were not included in the line list from HCDC. Therefore, individuals for whom more time

elapsed between infection and first positive test are more likely to go unnoticed in our data than those for whom this interval is short, which is referred to as right truncation. If this phenomenon is not properly addressed, the latency time distribution is biased downwards in the early phase of an outbreak, which is illustrated with real data by Xin *et al.* [2021] and Chen *et al.* [2022]. Note that the size of the bias induced by (ii) is enlarged as the infection risk increased during the study period. We first discuss this issue in more detail, and then present how we addressed it in our analysis.

Individuals that kept on testing negative were typically dismissed from the quarantine facility after 21 days. Hence, there is a chance that individuals who have an exceptionally long latent period and would test positive for SARS-CoV-2 infection after quarantine have been missed. Figure 4.3a illustrates this phenomenon.

The cut-off criterion utilized by HCDC for the line list we obtained was that the publication date of a case was July 15th or earlier. Note that the first positive test typically occurred three days before publication. Therefore, individuals that were infected close to July 12th are less likely to be included in the data when they have a longer latency time (Figure 4.3b).

Xin *et al.* and Linton *et al.* addressed right truncation due to exponential growth or decay during the sampling phase in their analyses [Xin, Wong, Murphy, Yeung, Ali, Wu and Cowling, 2021; Linton et al., 2020] using their estimated exponential growth or decay factor r. This was formalized by Park *et al.* [2024] (see Eq. 23 in their paper). A difference between the likelihood proposed by Park which was utilized by Xin, and the approach by Linton is that the latter assumes a constant risk of infection within the exposure window, while addressing truncation due to exponential growth whereas Park and Xin assume an increasing risk within the exposure window.

For the truncation date for each individual $t_{i,end}$ in calendar time, we used the start of quarantine plus 21 days or July 12th (end of sampling), whichever occurred earlier in time. The start of quarantine was known for 53% of the included individuals (n = 1038); for the other half of the observations, July 12th was used as a truncation date. For those with unobserved quarantine date (43.8%, n = 913), the exposure window ends at first positive test or symptom onset, typically yielding wide exposure windows. Hence, these observations do not contribute much to the estimate.

Denote by $t_{i,end}$ the calendar time at which the observation was truncated; let the condition for the inclusion of individual i in the data set be $s_{ir} \leq t_{i,end}$, i.e. an individual

tests positive before the end of study. Recall that each individual observation of latency time $(e_{il}, e_{ir}, s_{il}, s_{ir})$ represents two coinciding (type I, Eq. 4.2), distinct (type II, Eq. 4.1) or partially overlapping (type III, Supplement 4.7.5) windows and therefore belongs to one of the three respective partitions of the data denoted by I, II or III (Figure 4.2). The likelihood of the complete set of observations $i = 1, ..., N$ addressing right truncation is

$$\prod_{i=1}^{N} \frac{l^I \cdot l^{II} \cdot l^{III}}{\int_{e_{il}}^{e_{ir}} g(e|e_{il}, e_{ir}) F(t_{i,end} - e) de}. \tag{4.3}$$

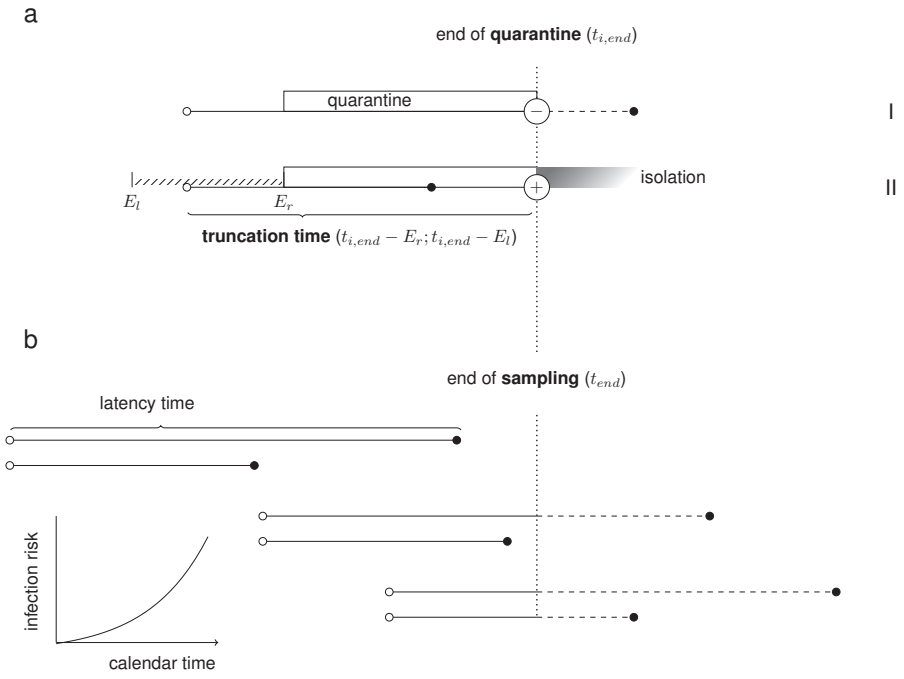


Figure 4.3: Illustration of truncation in the context of latency time estimation. Data representation of two individuals with a latency time of 9 days (upper) and 5 days (lower). In panel (**a**) both individuals were infected (open bullet) on the same calendar day and entered quarantine on the same day. However, individual I is unobserved (left quarantine while testing negative) whereas individual II tested positive by the end of quarantine and therefore appears in the data set with a (calendar) truncation time. Because infection is not observed exactly but known to fall within the exposure window $(E_l; E_r)$, the truncation time is interval censored. Panel (**b**) visualizes the same pair or latency times as before, but with different infection moment for each pair. Towards the end of the sampling period, relatively more individuals get infected due to exponential growth. The effect of truncation is the same as before.

4.4.5 Software

We used a Bayesian approach for estimation, using Markov Chain Monte Carlo to quantify the posterior distribution (Supplement 4.7.4). We extended code by Charniga *et al.*, used to estimate the incubation time for mpox [Charniga et al., 2022]. We adapted their code to allow for doubly interval censored data with all three types of observations of the latency time (example I, II and III in Figure 4.2). Apart from a constant infection risk within the exposure window we also considered exponential growth. For the latency time distribution we assumed gamma, Weibull and the more flexible generalized gamma model. We additionally corrected for right truncation (see Equation 4.3). The JAGS code that implements doubly interval censoring and right truncation is surprisingly concise as the different components can be combined without necessity to write down the complete likelihood (Supplement 4.7.4 for details). We assumed non-negative, flat priors for all parameters (normal distribution with μ = 0 and $\sigma^2 = 1000$, truncated at 0). We assumed that the uncertainty in the estimated epidemic growth rate r is normally distributed, with a standard deviation equal to the Standard Error of the estimate. We used the following settings: 3 chains, each with 500 000 iterations; a burn-in period of 10 000 and thinning factor 10, respectively.

We validated our code against the `coarseDataTools` package using simulated data, using Weibull and gamma for the time-to-event distribution, uniform distribution for infection and no right truncation. We found that the estimates were very similar (results not shown).

All analyses were performed in `R` Studio [RStudio Team, 2021]. Models were fitted via the `rjags` interface to JAGS [Plummer, 2017]. We used the ALICE computing resources provided by Leiden University. All code is available via *github.com/vharntzen/LatencyCovid Vietnam*.

4.4.6 Estimates

We report estimates of the mean, median, and the 95^{th} percentile, along with the corresponding 95% credible intervals. Parameter estimates that characterize the distributions can be found in Supplement 4.7.2.

Sensitivity analyses were performed in which only observations with narrow (≤ 4 days) exposure windows were analysed and in which truncation was not addressed

(Supplement 4.7.3).

4.5 Results

4.5.1 Descriptives

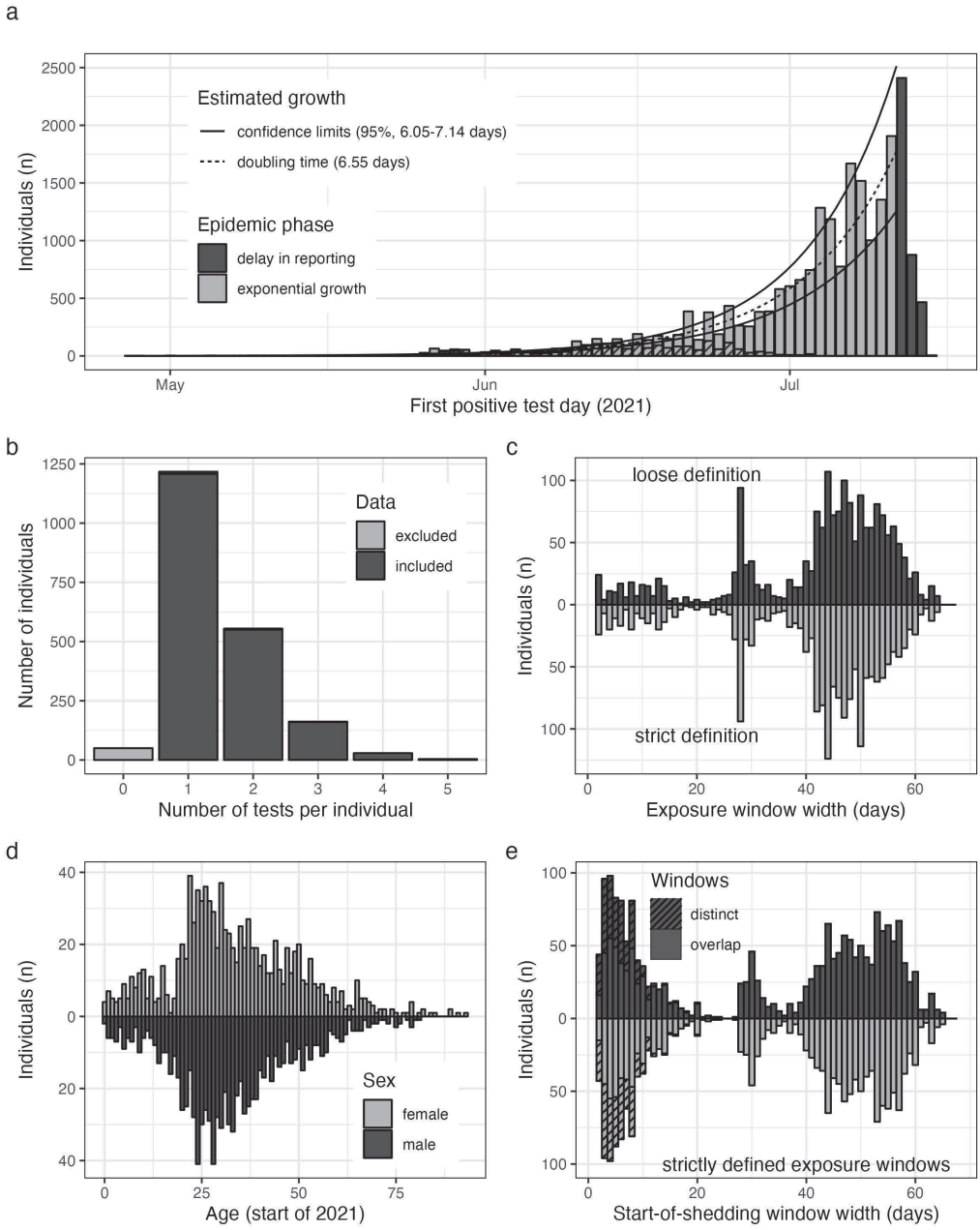


Figure 4.4: (Caption next page.)

Figure 4.4: Data characteristics. Panel (**a**): daily number of individuals from the line list that test positive for SARS-CoV-2 for the first time (per day on which the sample was taken). Greyscale indicates whether the data was used for estimation of the exponential growth factor (light grey) or considered underreporting (dark grey); the dashed part of each bar represents individuals included for analysis. As a first positive test typically occurred three days before a case was made public and the line list concerned all infecteds with publication on July 15th latest, the delay phase starts on July 12th, 2021. The solid line represents the estimated exponential growth curve; the dotted lines represent the confidence bounds around this curve. Panel (**b**): distribution of the total number of tests per individual. Note that the figure displays 2018 individuals for whom data was extracted. Only individuals with at least one test result (positive/negative) and known test day were included for analysis. Panel (**c**): the empirical distribution of the width of exposure windows included for analysis, according to a strict definition (below zero) and more loose definition (above zero). Panel (**d**): age distribution by sex of individuals included for analysis. Panel (**e**): the empirical distribution of start-of-shedding windows using different definitions of the exposure window (see above). Dashed areas indicate observations for whom exposure and start-of-shedding windows (partially) overlap.

Figure 4.4a visualizes the calendar day of first positive test of all individuals in the line list (N = 22,097). Excluding 3759 individuals (17.0%) with a first positive test for which the sample was taken from July 12th onwards (dark grey area) due to reporting delay as the public health system started to experience an overload, we estimated a doubling time of 6.55 days (95% CI 6.05; 7.14) with a corresponding exponential growth factor r of 0.106 (95% CI 0.097; 0.114).

The age distributions of males and females were almost equal (Figure 4.4d).

Figure 4.4c and e visualize the exposure window and start-of-shedding window widths, respectively, obtained by handling a more strict definition (below zero) and a looser definition (above zero) of the exposure window. Around 12% (n = 235) of individuals had an exposure window smaller than ten days (strict definition). The more strict definition did not decrease window widths substantially. The wider the start-of-shedding window, the more likely it is to overlap with the exposure window (Figure 4.4e). The start-of-shedding windows tend to be narrower than the exposure windows.

4.5.2 Estimates

Figure 4.5 visualizes the latency time estimates in our analyses and in the literature.

With the most flexible model (generalized gamma distribution), assuming exponential growth as estimated using the line list (r = 0.106), a strict assumption regarding the exposure window bounds, and addressing truncation, we estimated the latency time for the SARS-CoV-2 Delta variant as follows: mean 3.22 (95% CrI 2.89; 3.55) days; median 1.81 (95% CrI 1.44; 2.16); 95% percentile 10.98 (95% CrI 9.91; 12.41) (Figure 4.5a, dotted line and Figure 4.6).

Most outstanding is that all estimates are considerably larger when a constant infection risk is assumed instead of exponential growth (Figure 4.5a). The weaker and stronger exponential growth scenarios (r $\pm \approx$ 2SE) were comparable to the results assuming the estimated growth factor. The estimates when we considered a loose definition of the exposure window (Figure 4.5a, second row) are slightly smaller compared to the estimates using the strict definition. The differences between assuming generalized gamma, gamma and Weibull distributions were negligible.

The results of the sensitivity analyses are shown in Supplement 4.7.3. The impact of restricting the analysis to observations with a narrow (≤ 4 days) exposure windows was different per assumption for the infection risk. Assuming exponential growth, the estimates were larger, while for an assumed constant risk of infection these were smaller. Not addressing truncation shrunk the credible intervals, specifically obtaining more symmetric credible intervals around the 95^{th} percentile. When exponential growth was assumed, addressing truncation did not change the point estimates much; under the constant risk assumption the point estimates became larger.

4.6 Discussion

The Vietnamese data set of latency time observations is unique in its size and level of detail. Nonetheless, our estimates of the latency time distribution depend strongly on the assumptions that we made, specifically regarding the exposure window and the risk of infection within this window. We estimated a mean latency time of 3.22 (95% CrI 2.89; 3.55) days for mostly unvaccinated individuals predominantly infected with the Delta variant, with a median of 1.81 (95% CrI 1.44; 2.16) and a 95% percentile of 10.98

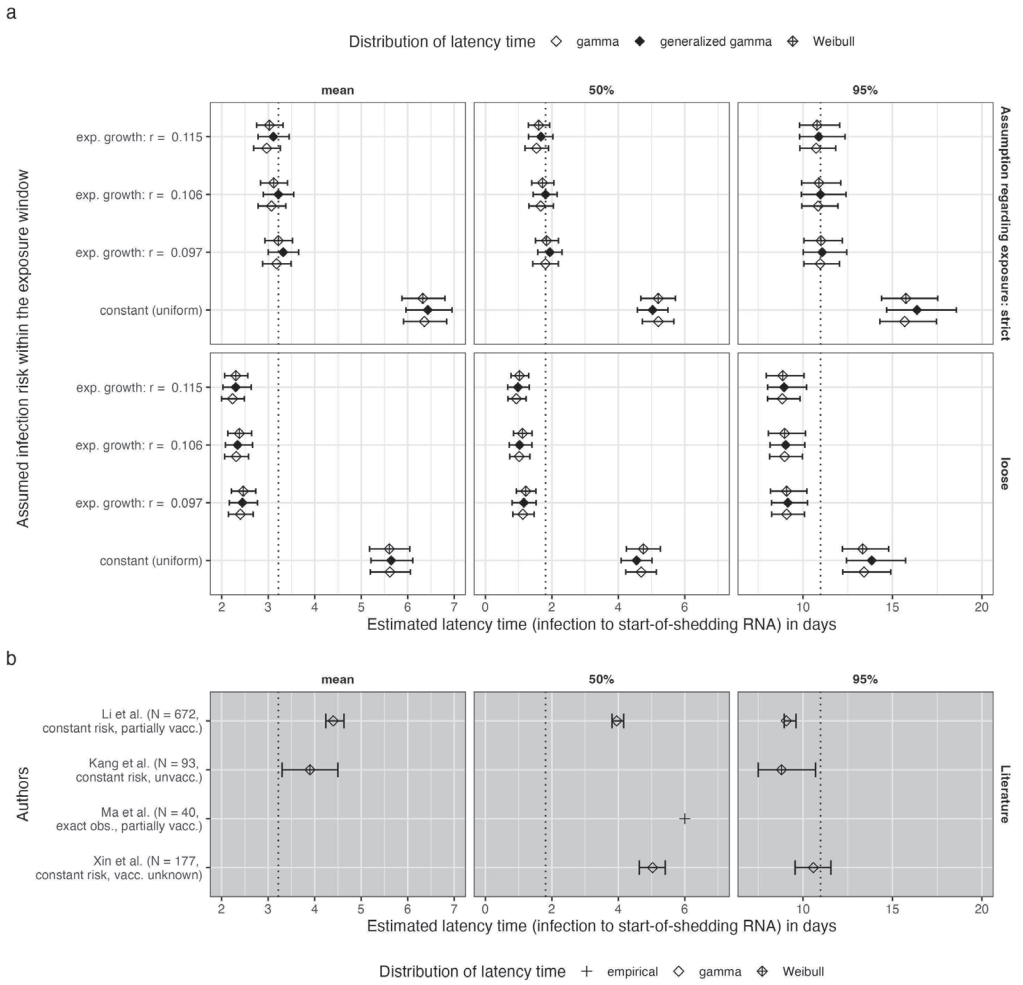


Figure 4.5: Latency time estimates for SARS-CoV-2. Panel (**a**): estimates based on our data concerning mostly the Delta variant; panel (**b**): estimates from the literature, concerning different variants. The mean, median and 95% percentile are presented with corresponding 95% credible intervals (**a**) or credible/confidence intervals (**b**) represented by error bars. Symbols refer to the assumed parametric shape of the distribution of latency time. All estimates are given in days (x-axis). In panel (**a**), estimates are given for different assumptions for the infection risk within the exposure window (y-axis) and exposure window definitions (rows). The estimated exponential growth rate was 0.106, with a standard error (SE) of 0.004; the results are shown assuming the estimated mean plus or minus 2 SEs as well.

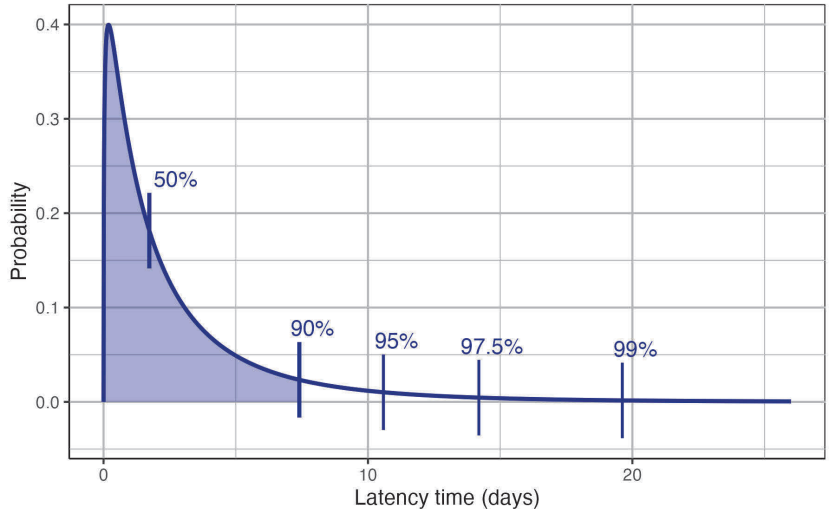


Figure 4.6: Estimated latency time distribution for SARS-CoV-2. Vertical bars correspond to the 50%, 90%, 95%, 97.5% and 99% percentile, respectively.

(95% CrI 9.91; 12.41) with the strict definition of the exposure window bounds. Our estimates assuming exponential growth are smaller than any of the estimates we found in the literature regarding the Delta variant.

As far as we are aware, we present the first estimate of the SARS-CoV-2 latency time for which the infection risk is assumed to increase according to the exponential growth of the case incidence. In similar analyses in the literature, it was assumed to remain constant in calendar time, which is known to be unrealistic during infectious disease outbreaks. Our analyses assuming a constant risk of infection provide estimates that are indeed comparable with most estimates in earlier studies (Figure 4.5b). However, it is unclear whether (part of) their observations concern the Delta variant. The estimates by Kang *et al.* for the Delta variant, assuming a constant risk, are considerably smaller than our estimates under the same assumption [Kang et al., 2022]. This is especially surprising given that they estimated a larger exponential growth rate during their sampling phase (0.33, 95% CrI: 0.18–0.48) than we found using our data (0.106, 95% CI 0.097; 0.114). In contrast to our analyses and the analysis by Xin *et al.* [2021], none of the other papers on SARS-CoV-2 latency time mentioned that they address truncation in their analysis.

Incubation time and latency time are closely related. In a large study, the mean

incubation time for the Delta variant was estimated to be 4.43 days (95% CI 4.36 – 4.49) [Galmiche et al., 2023]. Unfortunately, we could not estimate the incubation time distribution because our data on symptom onset was incomplete. Our estimate of the latency time is shorter, as one would expect given the possibility of presymptomatic spread of the infection [Tindale et al., 2020]. As Galmiche *et al.* noted, comparing the incubation time for variants is complicated due to differences in study design and vaccination status, and the same holds for latency time estimates.

We can safely assume that the vast majority of infections concerns the Delta variant and was unvaccinated at the time of infection. Note that the Vietnamese population is relatively young, e.g. median age is 31 (IQR 23; 43) in our data set and 31.6 countrywide (2020 [United Nations, Department of Economic and Social Affairs, Population Division, 2022]). Our estimate of the latency time distribution is probably generalizable to other low- and middle income countries with a similar age structure. Our study contributes by providing ready-to-use software that extends the current methods by addressing right truncation as we discussed before as well as in two other aspects that we discuss next.

We modelled the infection risk within the exposure window assuming exponential growth. The exponential growth factor r that we used in our analyses was based on the simple relationship where the incidence $i(t)$ can be expressed as $i(t) = i(t_0)e^{rt}$ where t_0 is the start of an outbreak. Note that the exponential growth factor (r) determines the doubling time ($ln(2)/r$) of an infectious disease. One caveat is that strictly speaking t_0 is unknown [Anzai and Nishiura, 2022]. The alternative equation proposed by Anzai *et al.* uses the time elapsed since one started counting cases from the outbreak (t_0), denoted by τ, i.e. $i(t) = i(t_0)e^{rt} \cdot \frac{e^{r\tau}-e^{-r}}{r}$. However, as there was basically no transmission in Vietnam just before the Delta outbreak started, we assumed that $\tau \approx 0$. Another limitation is that for some individuals, the population risk does not equal the individual risk within their exposure window, for example, when an infection is introduced into a household (Chapter 2). When viral sequencing data is available, one can obtain more certainty in the pairing of infector and infectee using the method described by Stockdale *et al.* [2023]. This would yield narrower exposure windows and thus milder bias when there is a discrepancy between the true and assumed infection risk distribution.

We contribute to the literature in the field by providing our code for doubly interval censored observations, which allows to assume an exponentially increasing (or decreasing)

risk within the exposure window, various assumptions regarding the shape of the latency time distribution and to address right truncation, implemented using JAGS [Plummer, 2017]. As WHO stressed after the emergence of SARS-CoV-1 [WHO, 2003], coronaviruses tend to have a long tailed incubation time distribution. In earlier work, we found that the common choices of parametric distributions (gamma, lognormal, Weibull) do not always capture the shape of the tail adequately, which is particularly problematic as the tail of the distribution informs quarantine length. Moreover, with parametric choices that differ by study, pooling estimates is not easy [McAloon et al., 2020]. Therefore, we assumed a generalized gamma distribution to allow for more flexibility regarding the shape of the latency time distribution. In the future, it would be worth exploring even more flexible methods such as splines that allow for simultaneous estimation of infection and latency time. The `R` package `TwoTimeScales` currently only includes exact and right censored observations; however the authors plan to extend this [Carollo et al., 2023].

One can also use our software for incubation time estimation as the symptom onset is typically interval censored with a window of one day. Furthermore, it can be used for estimation of other delay distribution such as the generation interval and serial interval. In both cases, observations are doubly interval censored and a flexible modelling approach using the generalized gamma distribution may be beneficial.

A limitation of our study is that transmission within quarantine facilities cannot be ruled out as some individuals quarantined with infected individuals in one room. However, when transmission within the quarantine facility was observed and mentioned in the public health reports we took this into account in the choice of exposure window: following the loose definition, the end of exposure would in such a case be the moment that an infected room member was transferred to the hospital. This may explain why the estimates with a loose exposure window definition are rather short as individuals may not experience the exponentially growing risk of infection during quarantine. Despite these efforts, some of these transmissions may be unnoticed. Moreover, we did not account for imperfect specificity and sensitivity of the PCR- and rapid antigen tests as this would be challenging, especially as the type of test was mostly unknown. Lastly, one should note that the start of infectiousness may occur earlier than its detection by PCR [Xin, Li, Wu, Li, Lau, Qin, Wang, Cowling, Tsang and Li, 2021].

A fruitful future direction would be to collect data in real-time by using a digital question-

naire form with a scientific purpose besides the typical aim of controlling the outbreak. We have developed a questionnaire form which is openly available via `https://github.com/manhnguy/Contact-Tracing-for-Respiratory-Transmitted-Diseases`. Answers to the questionnaire form can be directed imported into R, and we provide guidance to do so.

In this paper, we provide an estimate of the latency time for the SARS-CoV-2 Delta variant, making novel, realistic assumptions regarding exposure information, and we make our code openly available. To facilitate future estimation of such quantities in a timely manner, more research is needed into both data collection and model assumptions regarding exposure information that match the unruly reality of the infectious disease context.

Acknowledgements

The authors thank Jordache Ramjith for providing his code and Le Van Tan for input on SARS-CoV-2 testing and vaccination practices.

4.7 Supplementary material

4.7.1 Generalized gamma distribution

The probability density function of the generalized gamma distribution as proposed by Stacy and Mihram [1965] is

$$f(t|\theta,\kappa,\delta) = \frac{\frac{\delta}{\theta^{\kappa}}}{\Gamma(\frac{\kappa}{\delta})} t^{\kappa-1} e^{-(\frac{t}{\theta})^{\delta}} \tag{4.4}$$

where θ, κ, $\delta > 0$. The generalized gamma distribution contains the gamma distribution ($\delta = 1$), lognormal ($\frac{\kappa}{\delta} \to \infty$) and Weibull ($\kappa = \delta$) distribution as special cases. The mean is

$$\theta \frac{\Gamma(\frac{\kappa+1}{\delta})}{\Gamma(\frac{\kappa}{\delta})} \tag{4.5}$$

and the variance

$$\theta^2 \left(\frac{\Gamma(\frac{\kappa+2}{\delta})}{\Gamma(\frac{\kappa}{\delta})} - \left(\frac{\Gamma(\frac{\kappa+1}{\delta})}{\Gamma(\frac{\kappa}{\delta})} \right)^2 \right). \tag{4.6}$$

Table 4.1 shows how this parameterization relates to the software implementations of the generalized gamma distribution in the `R` package `flexsurv` [Jackson, 2016] and

in `JAGS` [Plummer, 2017]. We also show the relation between the generalized gamma distribution and the gamma and Weibull distributions in base `R` `dgamma` and `dweibull` distributions, respectively.

For the frequentist approach that we describe in Supplement 4.7.5, we used the relationship between the cumulative distribution functions of the gamma distribution with scale θ and shape κ ($G(t|\theta,\kappa)$) and the generalized gamma distribution ($F(t|\theta,\kappa,\delta)$) [Rubio, 2020], i.e.

$$F(t|\theta,\kappa,\delta) = G(t^{\delta}|\theta^{\delta},\frac{\kappa}{\delta}), \tag{4.7}$$

and the `R` functions provided by Rubio on his webpage.

Table 4.1: Software implementations of generalized gamma distribution as parameterized by Stacy and Mihram [1965] and gamma and Weibull distributions as special cases.

Parameter	**`flexsurv`**	**JAGS, `rjags`**	**base `R`**	
Stacy and Mihram [1965]	`GenGamma.orig(shape, scale, k)`	`dgen.gamma(a, b, c)`	`dgamma(shape, scale)`	`dweibull(shape, scale)`
θ	`scale`	$\frac{1}{b}$	`scale`	`scale`
κ	`shape` $\cdot$ `k`	$a \cdot c$	`shape`	`shape`
δ	`shape`	c	1	`shape`

4.7.2 Parameter estimates of latency time distribution

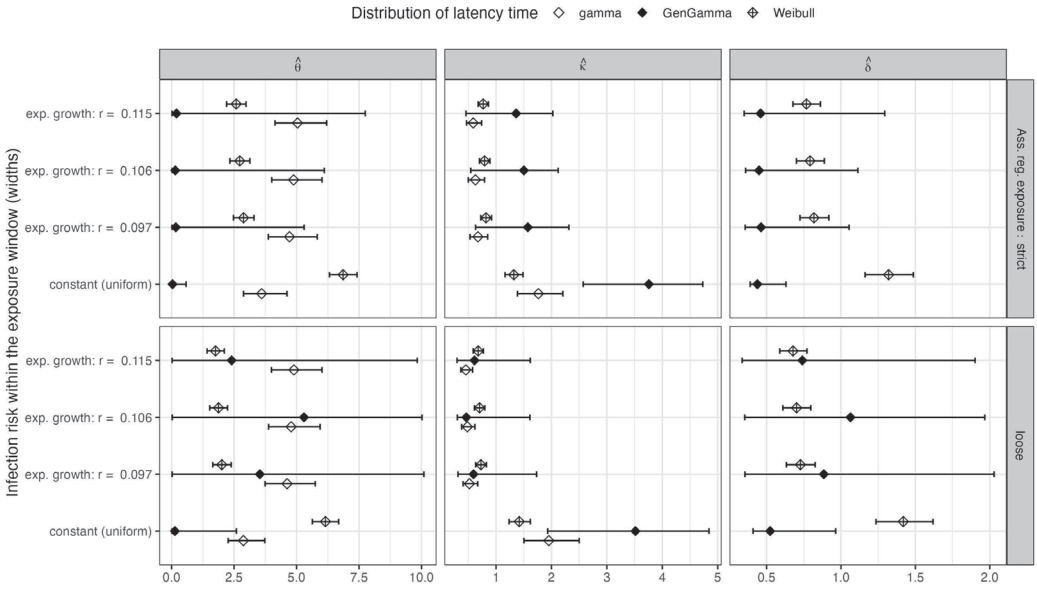


Figure 4.7: Parameter estimates of the latency time distribution for SARS-CoV-2. Parameters refer to the generalized gamma distribution with three parameters θ, κ and δ (columns) [Stacy and Mihram, 1965], of which the gamma (δ = 1, omitted) and Weibull ($\kappa = \theta$) distribution are special cases. Parameter estimates are presented as symbols that refer to the assumed parametric shape of the distribution of latency time and their corresponding 95% credible intervals are represented by error bars. Estimates are given making two different assumptions for the risk of infection within the exposure window (within panel) and two different assumptions regarding the exposure window bounds (rows).

Although the uncertainty in the estimated percentiles based on the generalized gamma distribution is comparable to the gamma and Weibull distributions, the uncertainty in the parameter estimates is considerable (Figure 4.7). This is probably due to strong collinearity between the parameters that define the generalized gamma distribution (Figure 4.8).

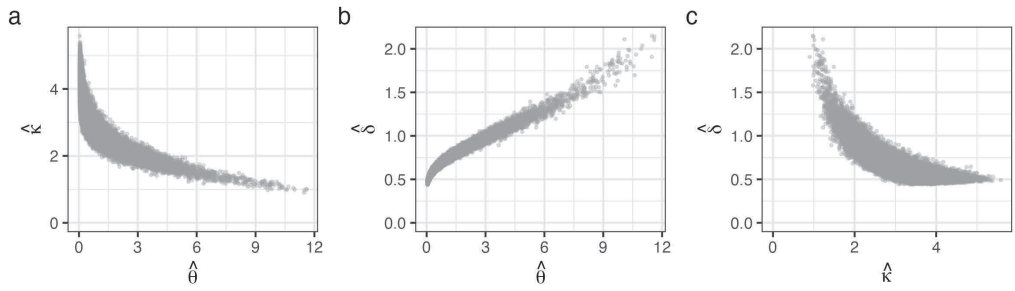


Figure 4.8: Bivariate posterior distribution of parameter pairs of the generalized gamma distribution for the latency time. Settings: handling a strict assumption regarding exposure window and assuming a constant infection risk. Each dot represents one iteration (45 000 in total). Each of the panels (**a**-**c**) visualizes a different combination of parameters. We observe strong collinearity between each of the parameters.

4.7.3 Sensitivity analyses

Figure 6.1 presents the latency time estimates corresponding to sensitivity analyses regarding the observations that were included, whether or not truncation was addressed and the assumption that was made when determining the exposure window bounds. Note that the difference between assuming exponential growth or a constant risk of infection disappears when analysis is restricted to observations with narrow exposure windows (≤ 4 days). The latter estimates are in-between the estimates for the two assumptions regarding the infection risk when all observations are included for analysis.

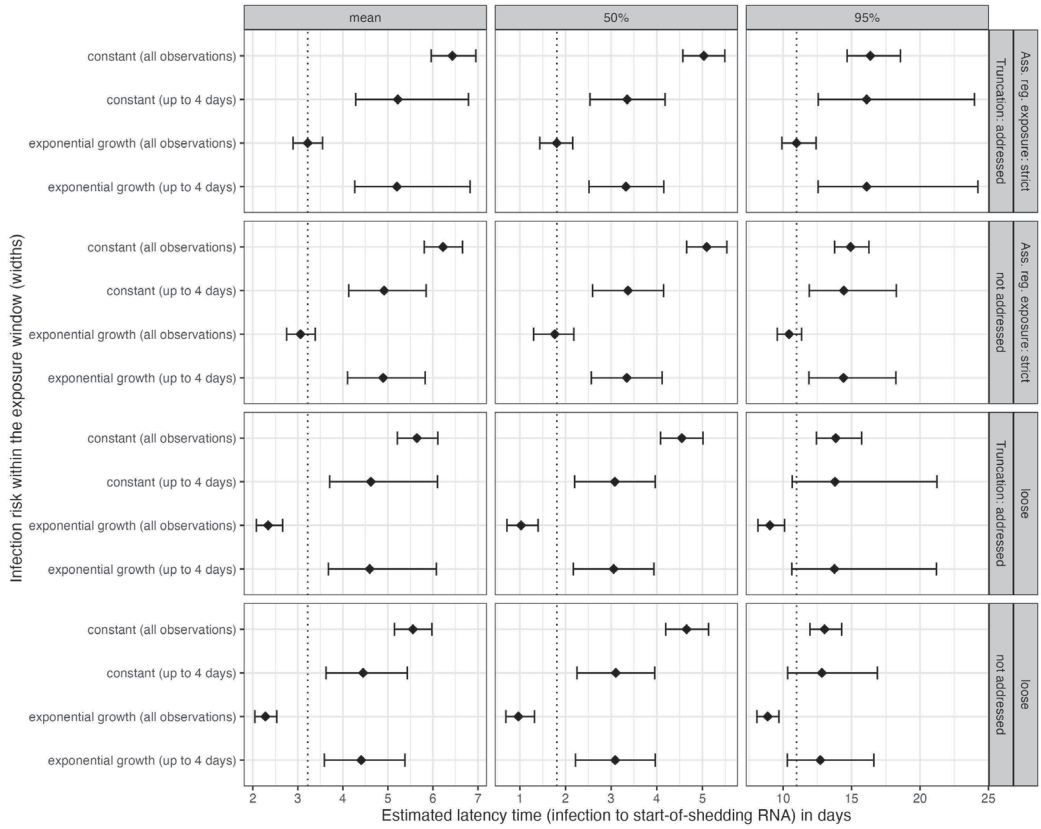


Figure 4.9: Estimated mean and percentiles for different assumptions regarding exposure and right truncation. The y-axis represents the assumed infection risk within the exposure window (r = 0.106 for exponential growth) and whether all observations or observations with narrow ($\leq$ 4 days) were selected for analysis. Panels correspond to the assumption regarding the exposure window bounds and whether truncation was addressed in the analysis or not (rows) and the estimated outcome measure (columns). All analyses assumed a generalized gamma distribution.

4.7.4 Implementation of our model in JAGS

The implementation is surprisingly concise, thanks to the incorporation of both interval censoring and truncation in the JAGS program. The code for a model assuming an increasing risk of infection within the exposure window and a generalized gamma- distributed latency time distribution is:

```
model {
  for(i in 1:N){
```

Table 4.2: Estimates for different assumptions regarding the analysis. All models assume that the latency time distribution follows the shape of generalized gamma. *Abbreviations* Exp. window: (assumption regarding) exposure window. Narrow exp.w.: observations with an exposure window width $\leq$ 4 days. CrI: credible interval.

Truncation	Exp. window	Inf. Risk	Data	Mean	(95% CrI)	50%	(95% CrI)	95%	(95% CrI)	$\hat{\theta}$	(95% CrI)	$\hat{\kappa}$	(95% CrI)	$\hat{\delta}$	(95% CrI)
not addressed	loose	constant (uniform)	all	5.56	(5.15; 5.98)	4.65	(4.19; 5.13)	13.03	(11.96; 14.28)	0.84	(0.07; 7.24)	2.83	(1.33; 4.45)	0.73	(0.5; 1.66)
			narrow exp.w.	4.45	(3.63; 5.43)	3.10	(2.25; 3.95)	12.83	(10.33; 16.88)	0.04	(0.01; 10.17)	2.60	(0.67; 3.93)	0.42	(0.35; 2.00)
		exp. growth: r = 0.097	all	2.36	(2.12; 2.62)	1.05	(0.76; 1.42)	9.02	(8.23; 9.85)	7.74	(2.76; 10.8)	0.39	(0.29; 0.70)	1.49	(0.84; 2.35)
			narrow exp.w.	4.41	(3.59; 5.38)	3.09	(2.21; 3.97)	12.71	(10.31; 16.65)	0.09	(0.02; 12.83)	2.28	(0.52; 3.68)	0.46	(0.37; 3.06)
		exp. growth: r = 0.106	all	2.28	(2.05; 2.53)	0.97	(0.69; 1.32)	8.87	(8.09; 9.69)	7.66	(2.18; 10.63)	0.38	(0.29; 0.72)	1.46	(0.77; 2.26)
			narrow exp.w.	4.41	(3.59; 5.38)	3.09	(2.22; 3.97)	12.71	(10.31; 16.63)	0.12	(0.02; 12.71)	2.23	(0.53; 3.62)	0.47	(0.37; 3.00)
		exp. growth: r = 0.115	all	2.20	(1.98; 2.45)	0.88	(0.64; 1.20)	8.75	(7.97; 9.56)	7.74	(2.87; 10.53)	0.36	(0.28; 0.62)	1.46	(0.83; 2.21)
			narrow exp.w.	4.39	(3.56; 5.35)	3.09	(2.18; 3.99)	12.64	(10.31; 16.44)	0.31	(0.02; 13.64)	1.96	(0.48; 3.51)	0.54	(0.37; 3.54)
	strict	constant (uniform)	all	6.22	(5.81; 6.66)	5.09	(4.65; 5.54)	14.93	(13.76; 16.27)	0.15	(0.02; 1.24)	3.72	(2.49; 4.84)	0.53	(0.43; 0.77)
			narrow exp.w.	4.92	(4.13; 5.85)	3.37	(2.59; 4.15)	14.43	(11.91; 18.27)	0.05	(0.02; 1.50)	2.49	(1.39; 3.55)	0.42	(0.36; 0.71)
		exp. growth: r = 0.097	all	3.21	(2.88; 3.55)	1.98	(1.48; 2.38)	10.49	(9.63; 11.43)	1.60	(0.02; 9.27)	1.08	(0.44; 2.51)	0.72	(0.38; 1.62)
			narrow exp.w.	4.89	(4.11; 5.81)	3.34	(2.57; 4.12)	14.41	(11.91; 18.17)	0.07	(0.02; 2.51)	2.38	(1.21; 3.39)	0.44	(0.37; 0.81)
		exp. growth: r = 0.106	all	3.06	(2.75; 3.39)	1.76	(1.30; 2.18)	10.43	(9.57; 11.35)	4.05	(0.09; 9.96)	0.72	(0.40; 1.92)	0.96	(0.44; 1.72)
			narrow exp.w.	4.89	(4.1; 5.83)	3.34	(2.57; 4.12)	14.41	(11.89; 18.24)	0.05	(0.01; 2.17)	2.47	(1.26; 3.54)	0.42	(0.36; 0.77)
		exp. growth: r = 0.115	all	2.93	(2.64; 3.24)	1.58	(1.15; 2.01)	10.37	(9.52; 11.28)	5.76	(0.17; 10.52)	0.57	(0.36; 1.61)	1.12	(0.47; 1.82)
			narrow exp.w.	4.89	(4.1; 5.81)	3.34	(2.56; 4.11)	14.41	(11.9; 18.20)	0.07	(0.02; 2.27)	2.38	(1.24; 3.41)	0.44	(0.36; 0.78)
addressed	loose	constant (uniform)	all	5.64	(5.21; 6.11)	4.55	(4.08; 5.01)	13.84	(12.43; 15.73)	0.13	(0.02; 2.59)	3.51	(1.93; 4.84)	0.52	(0.41; 0.96)
			narrow exp.w.	4.62	(3.71; 6.10)	3.08	(2.20; 3.96)	13.79	(10.67; 21.24)	0.05	(0.02; 7.26)	2.31	(0.83; 3.58)	0.42	(0.34; 1.40)
		exp. growth: r = 0.097	all	2.44	(2.16; 2.77)	1.16	(0.80; 1.52)	9.15	(8.25; 10.25)	3.53	(0.02; 10.09)	0.59	(0.31; 1.73)	0.88	(0.35; 2.03)
			narrow exp.w.	4.59	(3.67; 6.06)	3.06	(2.18; 3.95)	13.7	(10.61; 21.03)	0.06	(0.02; 10.64)	2.20	(0.64; 3.53)	0.43	(0.34; 2.13)
		exp. growth: r = 0.106	all	2.34	(2.08; 2.66)	1.03	(0.72; 1.40)	9.03	(8.16; 10.10)	5.30	(0.03; 10.02)	0.46	(0.30; 1.61)	1.06	(0.35; 1.97)
			narrow exp.w.	4.60	(3.68; 6.08)	3.05	(2.17; 3.94)	13.75	(10.64; 21.20)	0.05	(0.02; 8.76)	2.24	(0.73; 3.53)	0.42	(0.34; 1.65)
		exp. growth: r = 0.115	all	2.30	(2.03; 2.63)	0.98	(0.67; 1.32)	8.94	(8.03; 10.20)	2.40	(0.02; 9.83)	0.61	(0.30; 1.62)	0.74	(0.34; 1.90)
			narrow exp.w.	4.59	(3.67; 6.04)	3.06	(2.18; 3.94)	13.68	(10.61; 20.98)	0.06	(0.02; 10.41)	2.19	(0.65; 3.51)	0.43	(0.34; 2.04)
	strict	constant (uniform)	all	6.43	(5.96; 6.95)	5.03	(4.57; 5.49)	16.37	(14.68; 18.57)	0.04	(0.01; 0.58)	3.76	(2.57; 4.73)	0.44	(0.39; 0.63)
			narrow exp.w.	5.22	(4.28; 6.79)	3.35	(2.54; 4.18)	16.10	(12.56; 23.98)	0.05	(0.01; 1.17)	2.21	(1.28; 3.26)	0.4	(0.33; 0.64)
		exp. growth: r = 0.097	all	3.32	(3.00; 3.65)	1.94	(1.57; 2.30)	11.07	(10.02; 12.45)	0.17	(0.02; 5.30)	1.57	(0.63; 2.31)	0.46	(0.36; 1.05)
			narrow exp.w.	5.20	(4.26; 6.79)	3.32	(2.51; 4.16)	16.11	(12.56; 24.05)	0.05	(0.01; 1.22)	2.20	(1.25; 3.25)	0.4	(0.33; 0.65)
		exp. growth: r = 0.106	all	3.22	(2.89; 3.55)	1.81	(1.44; 2.16)	10.98	(9.91; 12.41)	0.15	(0.02; 6.11)	1.50	(0.54; 2.12)	0.45	(0.36; 1.11)
			narrow exp.w.	5.20	(4.26; 6.82)	3.32	(2.51; 4.15)	16.10	(12.55; 24.23)	0.05	(0.02; 1.02)	2.21	(1.30; 3.24)	0.4	(0.33; 0.62)
		exp. growth: r = 0.115	all	3.11	(2.78; 3.45)	1.67	(1.30; 2.03)	10.88	(9.81; 12.35)	0.20	(0.02; 7.75)	1.36	(0.46; 2.02)	0.46	(0.35; 1.29)
			narrow exp.w.	5.20	(4.26; 6.80)	3.32	(2.51; 4.15)	16.09	(12.54; 24.12)	0.05	(0.01; 1.35)	2.22	(1.24; 3.27)	0.4	(0.33; 0.66)

```
# Interval censoring of the endpoint
          type_R[i] ~ dinterval( Y[i] , C[i, 1:2] )
           C[i,1] <- max( 0.000000001, (R0[i] - L[i]) )
           C[i,2] <- max( 0.000000001, (R1[i] - L[i]) )

# Exponential growth (or decay)
           L_star[i] ~ dexp(r) T(0, L1[i])
           L[i] <- L1[i] - L_star[i]

# Generalized gamma-distributed time-to-event
           Y[i] ~ dgen.gamma(a, b, c)
  }

# Priors
  a ~ dnorm(0, 1/1000)I(0, )
  b ~ dnorm(0, 1/1000)I(0, )
  c ~ dnorm(0, 1/1000)I(0, )
  r ~ dnorm(r_est, prec_r)
  prec_r <- 1/(r_se^2)
}
```

The interval censoring of the endpoint is incorporated by the JAGS function `dinterval`. The interval censoring of the origin is implemented via truncation on the exposure window `T(0, L1[i])` of the pre-specified distribution (here exponential `dexp(r)`). Exponential decay can be implemented easily by replacing `L_star` directly by `L[i]`, saving one line. The parameterization of the generalized gamma distribution in JAGS differs from the one proposed by Stacy and Mihram [1965]. In our R code we rewrite the parameter estimates accordingly, see Table 4.1. To avoid any confusion, in the above code we use a, b, and c. We use an uninformative truncated normal prior for these parameters, via the term $I(0,)$. For the exponential growth factor r, we assume a normal distribution parameterized according to the input data, i.e. the estimate (`r_est`) and its Standard Error (`r_se`). Note that JAGS works with precision rather than standard deviation.

We verified the validity of this implementation of right truncated data by means of a small simulation study of 500 generated data sets per scenario and 1000 infected individuals per data set, before truncation. We generated data assuming a random exposure window width varying from one to five days and a constant risk of infection within the window. We chose a gamma distributed latency time as estimated by Xin *et al.* [2021] (shape = 4.05, rate = 0.74; median 5.03; 95^{th} percentile 10.57). The endpoint was observed up to the day accurate. Quarantine started at the end of the exposure window. Individuals were tested at the end of quarantine. We mimicked three scenarios with respect to truncation due to end of quarantine: i) individuals left quarantine 5, 10, 15, 20 or 25 days after entering, each occurring with equal probability; ii) quarantine lasted 7 days for all; iii) quarantine lasted 7 or 14 days (with equal probability). Only those individuals that tested positive before exiting quarantine were included for the analyses, meaning that those individuals for whom start-of-infectiousness occurred later remained unobserved. Our simulation code can be accessed via the example section the R help file for function `Estimate_doublIn` in our R package `doublIn`.

Figure 4.10 shows that our JAGS code adequately addresses truncation. Note that the difference between the analyses with and without correction for truncation is small when the level of truncation is very mild (14 or 21 days) whereas the difference is large when the level of truncation varies (5, 10, 15, 20, or 25 days) or is consistently strong (7 days). The coverage when truncation is not addressed is especially poor for the tail percentile. For quarantine scenario ii), 150 of 500 data sets did not provide a model fit

for one of the two models (addressing truncated or naive) which we hypothesize was due to too little information in the data set.

4.7.5 Alternative analyses

We explored several alternative analyses, that were eventually not suitable for our data, but may inspire the reader facing another data set.

We explored an equivalent frequentist approach, where we built upon `R` code by Ramjith *et al.* [2022] suitable for doubly interval censored observations with two non-overlapping or completely overlapping windows. We utilized the idea present in the source code of the `R` package `coarseDataTools` on representing the observation when exposure and start-of-shedding windows partially overlap and adapted the code by Ramjith *et al.* accordingly. Then, the contribution to the likelihood can be written as the sum of three district parts as shown in Figure 4.11. The same expressions are used as in the main text, e.g. Expression 4.2 for type (a) and 4.1 for type (b) and (c), but the start and end of the windows containing the origin (E) and (S) are adapted to match the respective part. With our data set we faced convergence issues that we hypothesize to be related to local maxima in the likelihood.

In an attempt to move away from the assumption of a constant risk of infection, we considered exponential risk as described in the main text but also two other variations. In both we first estimated the risk of infection over calendar time using the nonparametric maximum likelihood estimator for interval censored data (NPMLE; `R` package `interval` [Fay and Shaw, 2010]). In the first variation we considered a piecewise constant infection risk where we assume a constant risk during each day to which the NPMLE, a discrete estimator, assigned probability mass. In the second variation, we determined one or multiple peaks and lows in the risk of infection with the function `find_peaks(span = 3, ignore_treshold = 0.1)` from R package `ggpmisc` [Aphalo, 2023; Dayal, 2021]. For each of the decreasing and increasing phases, we estimated the corresponding exponential growth factor using the R package `incidence` [Kamvar et al., 2019]. In the analysis, we used the 'individual' exponential growth factor that was applicable midway the individual's exposure window. However, we saw that the NPMLE had a very limited amount of jumps, indicating that our data contained too much uncertainty regarding the moment of infection to choose one of these approaches.

We also fitted the infection risk distribution within each of the three largest clusters. Again, we saw that this gave too little information in order to provide a finescaled NPMLE.

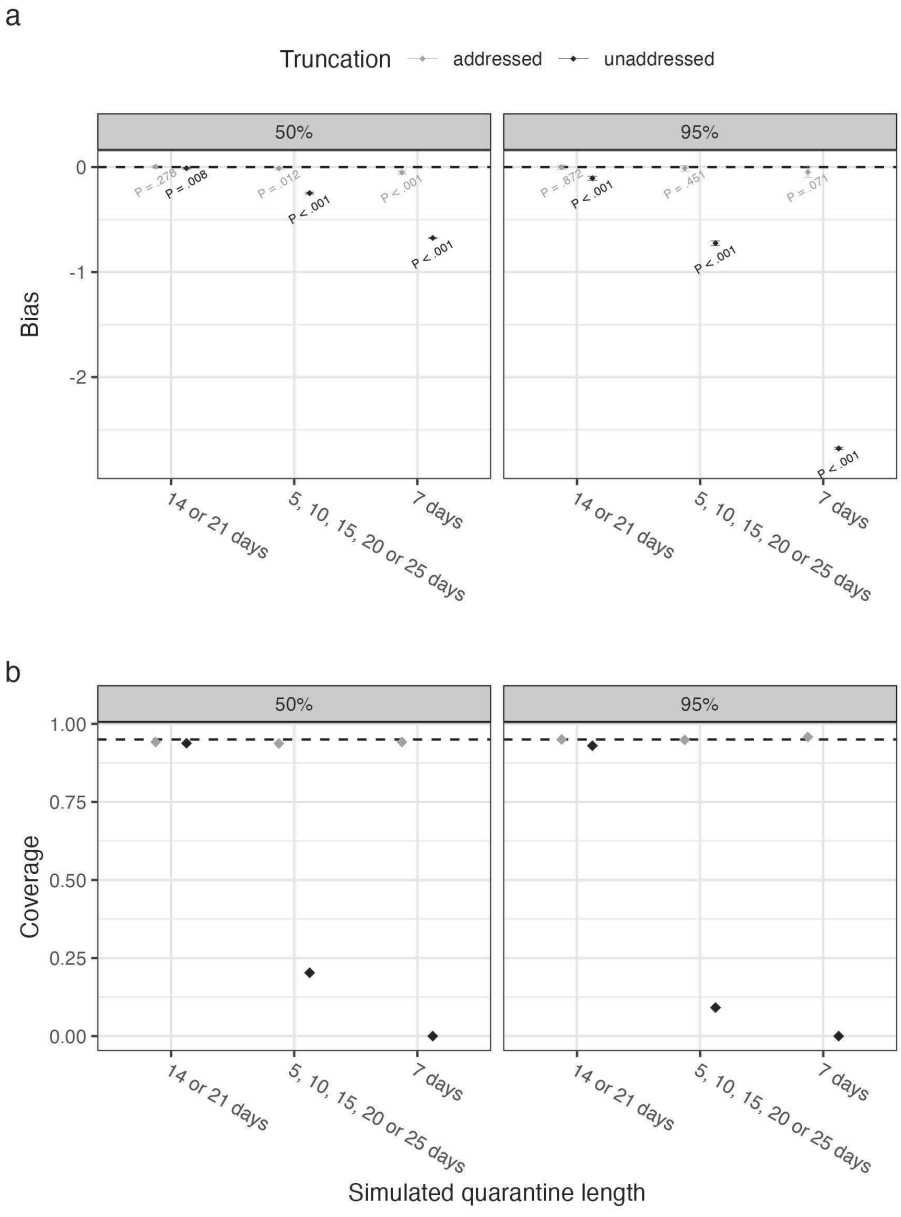


Figure 4.10: Simulation results to verify our JAGS implementation. Colors represent whether or not truncation was addressed in the analysis. Figure (**a**) visualizes the bias, its 95% Confidence Interval and the corresponding P-value (t-test). Figure (**b**) presents the coverage proportion of the true quantiles by the Confidence Intervals.

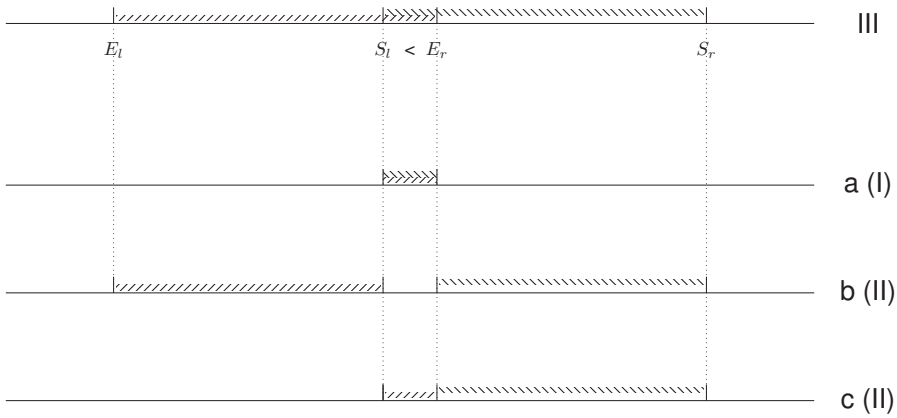


Figure 4.11: Illustration of data representation of partially overlapping exposure and start-of-shedding windows. In our code for the frequentist approach, we decompose partially overlapping windows (type III in Figure 4.2) in three distinct parts, one for the completely overlapping part (a) or two for the non-overlapping parts to the left and right (b and c). Part a is conform type I in Figure 4.2, whereas part b and c are conform type II. Inspired by the source code in the `coarseDataTools` package [Reich et al., 2009], we use this idea to write the likelihood of an observation with partially overlapping windows, by summing the likelihoods of part a, b and c.

This chapter will be published soon as Vera H. Arntzen, Marta Fiocco, Inge M.M. Lakeman, Maartje Nielsen and Mar Rodríguez-Girondo. A new inverse probability of selection weighted Cox model to deal with outcome-dependent sampling in survival analysis. Biometrical Journal.

A new inverse probability of selection weighted Cox model to deal with outcome-dependent sampling in survival analysis

Contents

Abstract

Motivated by the study of genetic effect modifiers of cancer, we examined weighting approaches to correct for ascertainment bias in survival analysis. Outcome-dependent sampling is common in genetic epidemiology leading to study samples with too many events in comparison to the population and an overrepresentation of young, affected subjects. A usual approach to correct for ascertainment bias in this setting is to use an inverse probability-weighted Cox model, using weights based on external available population-based age-specific incidence rates of the type of cancer under investigation. However, the current approach is not general enough leading to invalid weights in relevant practical settings if oversampling of cases is not observed in all age groups. Based on the same principle of weighting observations by their inverse probability of selection, we propose a new, more general approach. We show the advantage of our new method using simulations and two real datasets. In both applications the goal is to assess the association between common susceptibility loci identified in Genome Wide Association Studies (GWAS) and cancer (colorectal and breast) using data collected through genetic testing in clinical genetics centers.

Keywords Cox regression □ genetic epidemiology □ outcome-dependent sampling □ survival analysis □ weighting

5.1 Introduction

Outcome-dependent sampling is common in genetic epidemiology. Since harmful variants in cancer associated high risk genes are typically rare, an efficient sampling strategy to find carriers of these variants is to oversample affected individuals with a family history of a specific disease. For example, carriers of pathogenic variants in the Lynch syndrome associated gene *PMS2* and the breast- and ovarian cancer associated genes *BRCA1* and *BRCA2*, are often detected through genetic screening programs in which testing is targeted to families with multiple cases. Due to this testing strategy, the available study cohorts to investigate modifiers of cancer risk are often non-representative samples of the population of interest: carriers with an early diagnosis of cancer are more frequently included in the sampled population compared to those with delayed cancer diagnoses or individuals who remain disease-free.

In the context of survival analysis, family-based outcome-dependent sampling results in an over-representation of events and short lifetimes, which without adjustment, leads to biased estimates of covariate effects when using, for example, a Cox proportional hazards model. This happens because the sampling mechanism affects the joint distribution of the time-to-event and covariate.

To solve this problem, two main approaches have been proposed in the literature: methods based on retrospective likelihood [Carayol and Bonaïti-Pellié, 2004; Chatterjee et al., 2006; Barnes et al., 2013] and the weighted cohort method [Antoniou et al., 2005] based on weighted Cox regression. The general idea of the methods based on retrospective likelihood is to formulate the likelihood of the observed covariate values conditional on the observed outcomes. These methods typically require to know the familial relations within the sample and the distribution of the covariate of interest, leading to analytically complex and computationally intensive methods. When the overall age-specific incidence rates in the population of interest are known, an alternative approach to estimate the association between a set of covariates and time to cancer diagnosis under outcome-dependent sampling is to use a weighted Cox regression model [Antoniou et al., 2005]. The general idea is to propose a weighting scheme with different weights for affected (observed events) and unaffected (right-censored) individuals according to an external source so that the resulting weighted sample mimics the true target population [Antoniou et al., 2005; Barnes

et al., 2012] in terms of the age-specific proportions of affected and unaffected individuals.

Due to its simplicity, this is an attractive approach. However, the proposed weighted scheme has some limitations: it often leads to invalid weights in relevant practical situations since it is only workable under particular sampling schemes, such as those involving substantial oversampling of cases.

The primary objective of this study is to introduce a novel and more versatile inverse probability of selection weighting scheme, utilizing population-based age-specific incidence rates of the event of interest. This leads to the development of a generalized weighted cohort method capable of accommodating arbitrary levels of outcome-dependent sampling, offering an improved alternative to the existing approach. As a secondary goal, we aim to conduct a sensitivity analysis to assess the performance of the weighted approaches in the presence of unobserved heterogeneity, particularly exploring within-family correlations arising from shared, unobserved factors. Despite the frequent inclusion of multiple family members in studies employing the original weighted cohort method (see Supplemental Table 5.5 for details), the influence of unobserved heterogeneity in this context remains unexplored. This aspect merits thorough investigation.

The rest of the paper is organised as follows. In Section 5.2, the commonly used weighted cohort Cox approach is revisited and its assumptions are discussed. A new alternative weighting scheme is proposed in Section 5.2.2. In Section 5.3, both weighting schemes are compared by means of an intensive Simulation study. In Section 5.4, we present two real data illustrations. In both illustrations, the role of genetic variants as modifiers of cancer risk is studied using datasets of affected individuals and family members ascertained through genetic counseling in a clinical genetic center. In the first application, we focus on colorectal cancer in carriers of the pathogenic variant *PMS2*, and in the second application, we analyse the association between a Polygenic Risk Score (PRS) based on common breast cancer-associated variants and breast cancer risk in multiple case families. Main conclusions, recommendations, and a final discussion follow in Section 5.5.

5.2 Weighted Cox regression to deal with outcome-dependent sampling

Let T be the time to event of interest in the target population of interest. The typical target population in our context comprises individuals who are carriers of a specific rare mutation of interest. Denote by C the right censoring time, assumed to be uninformative. Denote by Z the covariate of interest. Since sampling schemes in genetic epidemiology are typically family-based, denote the observed sample information by $(t_{ij}, \delta_{ij}, z_{ij})$, where $i = 1, \ldots, n_j$ index all included individuals in the sample belonging to family j. It is important to note that even though the sampling is family-based, this does not necessarily imply the inclusion of multiple family members in the resulting sample. Assume that N families are observed, with varying observed size $n_1, \ldots, n_N$, so that $n = \sum_{j=1}^{N} n_j$ individuals are included in the sample. The observed time to event for individual i in family j (denoted from now on by ij), is given by $t_{ij} = \min(T_{ij}, C_{ij})$. Define the non-censoring indicator $\delta_{ij} = I(T_{ij} \leq C_{ij})$ where δ_{ij} is 1 if the event is observed or 0 if observation ij is right censored. z_{ij} denotes covariate value for individual ij.

The observed data is collected through a family-based outcome-dependent sampling scheme. The selection process begins by testing the first individual in a family, focusing on those diagnosed with cancer at a young age and with a family history of cancer. If this initial individual tests positive for the mutation, the rest of the family is invited to participate in genetic testing. This approach may identify additional carriers of the mutation within the family, though this is not always the case. Consequently, due to variations in age and family history criteria across studies, influenced by disease severity and prevalence, the level of outcome-dependent sampling varies across studies. Despite the diversity in the final configurations, samples of carriers of rare genetic variants obtained via genetic testing typically result in an over-representation of young cases in the sample.

A common approach to estimate the effect of covariate Z on T is to use the Cox proportional hazards model with hazard function $h(t|z) = h_0(t)\exp(\beta z)$ where $h_0(t)$ is the baseline hazard. With prospective cohort data, the parameter β can be estimated maximizing the partial likelihood. However, the over-representation of events and short event times in the sample due to outcome-dependent sampling affects the risk set composition along the follow-up time in comparison to the true population, which may

result in biased estimation of the covariate effect. A possible solution to this problem is to consider a weighted Cox model using external information about the distribution of T in the population to construct weights reflecting individuals' selection probabilities.

5.2.1 The weighted cohort approach revisited

When T represents the age at cancer diagnosis, or another common disease, registry data about the marginal distribution of T in the target population is often accessible. In practical scenarios, the available external information is typically aggregated into K distinct age intervals, defined as $I_1 = [a_0, a_1)$, $I_2 = [a_1, a_2)$,..., $I_K = [a_{K-1}, a_K)$. For cancer studies, the commonly available external data comprises the population cancer incidence rate μ_k for each age interval I_k. The seminal work of Antoniou et al. [2005] introduced a weighted Cox regression model with sampling weights derived in such a way that the incidence rates in each interval I_k, in the resulting pseudo-population after weighting, align with the incidence rates μ_k in the target population.

However, before presenting the specific calculation of these weights, it is essential to acknowledge two main assumptions regarding the observed data in this context, given that the externally available data is discrete in time. First, we assume constant hazards within each interval I_k for $k = 1, \ldots, K$. Second, we assume that right-censoring is also discrete and occurs at the specified time points defining the intervals. This implies that if censored observations happen to fall within interval I_k, we assume that the censoring took place at point a_k. When these two prerequisites are met, the marginal distribution of T in the resulting weighted pseudo-population, based on interval-specific incidence rates, will follow the same distribution as in the reference population.

Let r_k denote the number of individuals experiencing the event within the age interval, $I_k = [a_{k-1}, a_k)$, $k = 1, \ldots K$. Similarly, s_k denotes the number of individuals right-censored within the age interval I_k (i.e. follow-up ends between age a_{k-1} and a_k without the event being observed). The term $p_k = \sum_{\{ij: t_{ij} \in I_k, \delta_{ij}=1\}} t_{ij}$ denotes the total follow-up time accumulated by all r_k individuals experiencing the event in age interval I_k; the equivalent total follow-up time accumulated by the s_k right-censored individuals is denoted by $q_k = \sum_{\{ij: t_{ij} \in I_k, \delta_{ij}=0\}} t_{ij}$. Then, all r_k cases in interval I_k are assigned weight w_k and all s_k right-censored individuals in interval I_k are assigned weight v_k such that:

$$\mu_k = \frac{w_k r_k}{w_k p_k + v_k q_k + (a_k - a_{k-1}) \sum_{l>k} (v_l s_l + w_l r_l)}. \tag{5.1}$$

In the right part of Expression (5.1) the weighted total of affected observations is divided by the weighted total of observations at risk. Then, this weighted ratio is imposed to be equal to the population incidence rate μ_k. As a result, after weighting the sample age-specific incidence rates resemble the age-specific incidence rates of the population. However, since equation (5.1) alone does not guarantee unique weights w_k and v_k, the following constraint is incorporated to guarantee unique weights:

$$\frac{w_k r_k + v_k s_k}{r_k + s_k} = 1. \tag{5.2}$$

Combining equations (5.1) and (5.2) provides unique expressions for v_k and w_k:

$$w_k = \frac{\mu_k (q_k (r_k + s_k) + (a_k - a_{k-1}) s_k \sum_{l>k} (r_l + s_l))}{r_k s_k + \mu_k (q_k r_k - p_k s_k)}, \tag{5.3}$$

where $\sum_{l>k}(r_l + s_l)$ are all observations in age groups older than k. The weight equation for censored individuals is given by

$$v_k = \frac{1}{s_k}(r_k + s_k - w_k r_k). \tag{5.4}$$

Once weights v_k, w_k for each age interval I_k, $k = 1, \ldots, K$ are calculated, the regression parameter β can be estimated using the following weighted score equation:

$$U_a(\beta) = \sum_{ij:\delta_{ij}=1} z_{ij} - \sum_{ij:\delta_{ij}=1} \frac{\sum\limits_{l \epsilon R(t_{ij})} W_l z_l \exp[\beta z_l]}{\sum\limits_{l \epsilon R(t_{ij})} W_l \exp[\beta z_l]}, \tag{5.5}$$

where $R(t_{ij})$ is the set of individuals still at risk just before t_{ij}, i.e. $R(t_{ij}) = \{l : t_{ij} \leq t_l\}$, and weight W_{ij} for individual ij ($i = 1, \ldots, n_j$, $j = 1, \ldots, N$) is defined as

$$W_{ij} = \begin{cases} w_k, & \text{if } \delta_{ij} = 1 \text{ and } t_{ij} \in [a_{k-1}, a_k) \\ v_k, & \text{if } \delta_{ij} = 0 \text{ and } t_{ij} \in [a_{k-1}, a_k). \end{cases} \tag{5.6}$$

The derivation and unbiasedness of the estimator resulting from Expression (5.5) are outlined in Supplement 5.6.1. We followed the same reasoning as presented by Mandel *et*

al. [2017], who introduced an inverse probability weighted Cox model to address double truncation. According to the authors, their findings extend beyond the context of double truncation to deal with a broader range of biased sampling scenarios. To elaborate further, Supplement 5.6.1 provides a detailed explanation. Crucially, in our setting, we assume conditional independence between the selection event and covariate values, given the observed event time. This assumption allows us to apply the results of Mandel *et al.* [2017], with the selection event treated as a function of the observed event times.

A number of conditions are required to guarantee finite and positive weights w_k and v_k, namely:

$$r_k > 0 \tag{5.7}$$

$$s_k > 0 \tag{5.8}$$

$$r_k > \mu_k p_k \frac{s_k}{s_k + \mu_k q_k} \tag{5.9}$$

$$w_k < 1 + \frac{s_k}{r_k}. \tag{5.10}$$

Conditions (5.7) and (5.9) are required to get proper w_k weights for the cases, while conditions (5.8) and (5.10) are required to get valid v_k weights for those that are censored. Condition (5.7) implies the observation of events in all the considered intervals while condition (5.8) implies the presence of right-censored observations in all intervals under consideration. Conditions (5.9) and (5.10) are more difficult to interpret and evaluate beforehand, but they are both related to the level of oversampling of events. As discussed by [Antoniou et al., 2005], if oversampling of events occurs in all considered age groups both conditions are typically fulfilled. However, as we will show in our real data application, oversampling of young cancer cases is the norm in genetic epidemiology, but not necessarily the case at older ages, so condition (5.9) and especially (5.10) might not be fulfilled in relevant practical scenarios.

Under oversampling of events at interval I_k, condition (5.9) is verified since $r_k > \mu_k p_k$, i.e., the observed number of events in interval I_k is larger than the expected number of events assuming the population incidence rate (μ_k). Actually, since $\mu_k q_k$ is usually positive, $0 < \frac{s_k}{s_k + \mu_k q_k} < 1$ in general which implies that condition (5.9) is fulfilled even when no oversampling of events is observed in interval I_k. However, condition (5.10) is cumbersome. Since it involves the estimated weight for events w_k together with the ratio of events and

right-censored observations in interval I_k ($\frac{s_k}{r_k}$), this condition is often not satisfied when there is no clear oversampling of cases in interval I_k. In such situations, w_k can still be calculated but it becomes small, which leads to violating condition (5.10).

When any of the conditions (5.7)-(5.10) are not satisfied, the weighted cohort method can still be applied by merging intervals, however then the method becomes less precise and dependent on sample specific characteristics which may hamper comparability among studies using this method.

We therefore propose an alternative, more general weighting scheme which allows to overcome the aforementioned limitations.

5.2.2 The new generalised weighted cohort approach

We propose a new weighting scheme to correct outcome-dependent sampling using external information. In contrast to the previous weighted cohort method, the new approach is more general, as it can be applied with arbitrary levels of over or under-representation of events.

Similar to the original method, the newly proposed weights represent sampling probabilities given the observed time to event of each individual so that the resulting pseudo-population matches the target population of reference in terms of the marginal distribution of T. The same score function (5.5) and justification of its validity applies. However, here we take a different approach to derive the weights. Instead of directly using the incidence rates, we focus on the risk sets at the beginning at each interval I_k and weight the individuals so that the resulting weighted risk set presents the same ratio of events and non-events as one would expect if the sample would have been randomly drawn from the target population.

Let N_k denote the number of individuals at risk (those who did not experience the event yet) at the beginning of the interval I_k in our sample, denoted by $\mathcal{S}_O$, potentially drawn under an outcome-dependent sampling mechanism. Now denote by $\mathcal{S}_P$ a hypothetical random sample of the target population with the same N_k number of individuals at risk at the beginning of the interval I_k. In both cases, N_k can be split into two disjoint parts: (i) the number of individuals that experience the event within the interval I_k and (ii) those experiencing the event in later intervals. However, if $\mathcal{S}_O$ is obtained using outcome-dependent sampling, the expected number of individuals belonging to each of these two parts in $\mathcal{S}_O$ and $\mathcal{S}_P$ will, in general, differ.

For the hypothetical random sample $\mathcal{S}_P$, N_k can be decomposed as follows:

$$N_k = N_k S_k + N_k(1 - S_k) \tag{5.11}$$

where $S_k = P(T > a_k | T > a_{k-1})$ represents the conditional probability of experiencing the event in a later time interval than interval I_k given that the event has not been experienced before interval I_k in the reference population. S_k can be directly calculated from the typically available population cancer incidence rates μ_k for each age interval I_k, since $S_k = e^{-\mu_k(a_k - a_{k-1})}$, $k = 1, \dots K$. Accordingly, $1 - S_k$ is the probability of experiencing the event in the interval I_k given that it has not been experienced before, in the reference population. From Expression (5.11) follows that the ratio between events and non-events in interval I_k in the reference population is given by $\frac{1-S_k}{S_k}$. Under the assumption of constant hazards within each pre-specified interval I_k, the ratio of events to non-events completely determines the incidence rate in interval I_k, thereby entirely characterizing the marginal distribution of T.

The same decomposition of the risk set at the beginning of interval I_k can be made for the observed sample $\mathcal{S}_O$, potentially subject to outcome-dependent sampling:

$$N_k = N_k S_k^o + N_k(1 - S_k^o) \tag{5.12}$$

where S_k^o is the observed proportion of individuals at risk at time a_{k-1} experiencing the event beyond I_k, calculated with the sample data.

In our new approach, we keep those subjects non-experiencing the event at interval I_k unweighted ($v_k = 1$) while we assign specific weights (w_k) to those subjects experiencing the event of interest in interval I_k making use of the decompositions given by Expressions (5.11) and (5.12). Specifically, weights w_k correct the oversampling (or undersampling) of cases, such that the ratio between events and non-events on the interval I_k in the resulting pseudo-population after weighting is the same as in the reference population:

$$w_k = \frac{(1 - S_k)}{S_k} \frac{S_k^o}{(1 - S_k^o)}. \tag{5.13}$$

Equation (5.13) illustrates that the population ratio between events and non-events within interval I_k, denoted as $\frac{1-S_k}{S_k}$, is multiplied by the inverse quantity based on the observed data, $\frac{1-S_k^o}{S_k^o}$. After weighting, the composition of the risk set within interval I_k, in terms of the ratio of events to non-events, resembles the composition of the risk set within

interval I_k in the reference population. As a result, under oversampling of cases, weights for affected individuals in interval I_k are $w_k < 1$, representing the inverse of the probability of being selected. Alternatively, under undersampling of cases, $w_k > 1$. Interestingly, in absence of outcome-dependent sampling, i.e. under random sampling, $w_k = 1$ and the new method coincides with the regular unweighted Cox model.

With our new proposal, two conditions need to be fulfilled in order to get valid weights: $(1 - S_k^o) > 0$, $k = 1, \ldots, K$ and $S_K > 0$. The first condition $(1 - S_k^o) > 0$ is satisfied if events are observed in each interval I_k, so as the original weighted cohort method, observation of events in all group ages is a requirement of our new method. However, the new method does not require the presence of right-censoring which makes it a more general and natural approach. The condition $S_K > 0$ only involves the last interval and implies that the method is suitable for studying events not experienced by a part of the population during the relevant follow-up time. This is a mild condition that is always satisfied when studying defective distributions ($S(\infty) > 0$) such as time to cancer or other diseases since not all population members will develop the event of interest. Even if our interest would be to study time to death or the target population would be a highly susceptible population to a specific cancer with lifetime risk of 1, the new weights could still be applied with an appropriate choice of the upper limit of the last interval K.

Once the weights are calculated, the regression parameter β can be estimated using the weighted score equation (5.5) and robust estimates of the standard errors can be obtained using a sandwich estimator, as proposed by Antoniou *et al.* [2005] in the original weighted cohort approach.

In summary, both the existing weighted cohort and the new generalised weighted cohort approaches generate pseudo-populations by means of inverse probability of selection weighting, but these pseudo-populations are different. The method developed in this study, is more general since it does not require oversampling or undersampling of events in all or at specific intervals and does not make assumptions about the right-censoring distribution.

5.3 Simulation study

A simulation study was conducted to assess the new generalized weighted cohort method's performance and compare it with the existing approach in several scenarios intended to

mimic relevant situations in practice. We consider two main simulation settings. First, we generate data so that the (weighted) Cox approaches are well specified. Second, we study the sensitivity of the weighted methods to model misspecification due to the presence and failure to adjust for unobserved heterogeneity.

5.3.1 Simulation setup I

Simulated data was generated using the following model:

$$\lambda_{ij}(t) = \lambda_0 \exp(\beta z_{ij}), \tag{5.14}$$

where t is the observed event time, $\lambda_0 = \frac{1}{60}$ represents the constant baseline hazard, Z is a continuous covariate assumed to be normally distributed ($Z \sim N(0,1)$) and β represents the associated log-hazard ratio. If the resulting event times were larger than 100, these were set to 100. Censoring times were sampled from an exponential distribution ($C \sim Exp(60)$) and the family size in the population is set to n_j (family size of size $n_j = 2$ and 5 members were considered). In each Monte Carlo trial, we generated N families (N = 250, 500, 750). Family-based outcome-dependent sampling was implemented by including families in the sample if for at least n_A family members the event was observed before the end of follow-up ($n_A = 1, 3$). The different combinations of n_A and n_j lead to three different scenarios with increasing level of outcome-dependent sampling: scenario 1 (A1) with $n_j = 5$ and $n_A = 1$ represents the mildest level of selection, scenario 2 (A2) with $n_j = 5$ and $n_A = 3$ represents a medium level of outcome dependent selection, and scenario 3 (A3) with $n_j = 2$ and $n_A = 1$ represents the strongest level of outcome dependent sampling in the simulation study. Moreover, all included families had at least one 'young affected' defined as having observed event time smaller than the first quartile of simulated T distribution. This mimics the common practice in clinical genetics centers: families are invited to participate in genetic studies when a young family member is diagnosed with the event at a young age. In terms of covariate effect, the null case ($\beta = 0$) and two alternative scenarios ($\beta = 0.3, 1$) were considered.

For each considered value of β, the underlying population was generated by simulating a large data set (N = 200 000, $n_j = 1$) without ascertainment and it was used to approximate the population hazards needed to calculate the weights. In all simulation scenarios, we considered five intervals, each with a width of 20 years. Relative bias, mean square

error and coverage proportions of the 95% confidence intervals are reported in Table 5.1. Moreover, the proportion of invalid weights in the M Monte Carlo trials is reported for the two weighted methods.

5.3.2 Simulation setup II

In the previous simulation setting, we have assumed, as the proposed models in Section 5.2, that differences among individuals in terms of hazards can be fully accounted for by including covariates in the Cox proportional hazards model. However, when samples contain multiple members of the same family (often the case when applying the traditional weighted cohort approach as shown in Supplemental Table 5.5), unmeasured hetereogeneity may arise since members of the same family often share common unmeasured characteristics such as genetic, social, dietary or other factors. In this second simulation setting, in order to introduce such unmeasured heterogeneity in the simulated data, we consider an extension of the data generation model specified in Expression (5.14) by adding a latent (frailty) term, U, shared by all members of the same family:

$$\lambda_{ij}(t) = u_j \lambda_0 \exp(\beta z_{ij}), \tag{5.15}$$

where $u_j \sim \Gamma(1, \theta)$ is a latent term (frailty), shared by the n_j members of a given family j. The larger the value of the variance θ, the more family members are alike and the larger the difference between families, yielding larger unobserved family effects. We consider two different values of within-family correlation: 'low' ($\theta = 0.1$) and 'large' ($\theta = 1$). Note that the latent frailty U and the covariate under investigation Z are independent. We expect that, as in the traditional unweighted Cox regression context [Henderson and Oman, 1999], ignoring the presence of U introduces bias in the hazard ratio estimation due to non-collapsability, even when U is independent of the covariate of interest Z. However, we also expect that such bias diminishes in the presence of outcome-dependent sampling and with the use of inverse probability of selection weighted Cox models.

The rest of the simulation settings were as in the previous Simulation setting I, except N, the number of families was fixed to 500 in this setting. For each scenario, weighted Cox models were estimated using the traditional and the generalized weighting scheme. Results obtained with the standard choices of using an unweighted Cox model or a shared gamma frailty model (unweighted) are also reported. Note that the application of the weighted

Table 5.1: Simulation I. Relative bias (reBias), mean square error (MSE) and coverage probability (Coverage) for $\widehat{\beta}$ along 1000 trials. A1: mild level of ascertainment. A2: medium level of ascertainment A3: strong level of ascertainment. N: number of families. For the weighted approaches, the proportion of invalid (negative) weights along 1000 trials is also reported.

			Unweighted			Weighted cohort				Generalized weighted cohort			
β	Scenario	N	reBias	MSE	Coverage	reBias	MSE	Coverage	Invalid weights	reBias	MSE	Coverage	Invalid weights
$\beta = 0$	A1	250	< 0.001	0.003	0.947	< 0.001	0.003	0.939	0.002	< 0.001	0.003	0.944	0.000
		500	0.002	0.002	0.953	< 0.001	0.011	0.953	0.000	< 0.001	0.002	0.949	0.000
		750	< 0.001	0.001	0.942	-0.002	< 0.001	0.939	0.000	-0.002	0.001	0.947	0.000
	A2	250	-0.003	0.004	0.942	< 0.001	0.006	0.938	0.005	< 0.001	0.005	0.940	0.000
		500	0.001	0.002	0.950	< 0.001	0.003	0.943	0.000	< 0.001	0.003	0.939	0.000
		750	-0.002	0.001	0.953	-0.003	0.002	0.941	0.000	-0.002	0.002	0.942	0.000
	A3	250	0.003	0.012	0.948	0.010	0.029	0.925	0.570	0.001	0.034	0.914	0.000
		500	< 0.001	0.006	0.939	< 0.001	0.012	0.937	0.191	< 0.001	0.014	0.935	0.000
		750	-0.003	0.004	0.957	0.002	0.007	0.947	0.058	0.002	0.008	0.951	0.000
$\beta = 0.3$	A1	250	-0.044	0.003	0.948	-0.091	0.004	0.912	0.000	-0.041	0.004	0.944	0.000
		500	-0.046	0.002	0.932	-0.092	0.002	0.887	0.000	-0.042	0.002	0.931	0.000
		750	-0.036	0.001	0.943	-0.100	0.002	0.841	0.000	-0.049	0.001	0.919	0.000
	A2	250	-0.200	0.008	0.818	-0.107	0.007	0.916	0.006	-0.129	0.007	0.893	0.000
		500	-0.195	0.005	0.733	-0.107	0.004	0.907	0.000	-0.128	0.004	0.877	0.000
		750	-0.210	0.005	0.587	-0.115	0.003	0.871	0.000	-0.138	0.003	0.837	0.000
	A3	250	-0.221	0.017	0.885	-0.012	0.024	0.941	0.562	0.071	0.033	0.933	0.000
		500	-0.229	0.010	0.851	-0.034	0.013	0.933	0.208	0.006	0.015	0.937	0.000
		750	-0.220	0.008	0.795	-0.037	0.008	0.947	0.077	-0.005	0.010	0.938	0.000
$\beta = 1$	A1	250	-0.063	0.007	0.805	-0.036	0.006	0.919	0.000	-0.001	0.005	0.946	0.000
		500	-0.064	0.006	0.655	-0.037	0.003	0.893	0.000	-0.003	0.002	0.956	0.000
		750	-0.067	0.034	0.463	-0.038	0.003	0.822	0.000	-0.004	0.002	0.943	0.000
	A2	250	-0.173	0.035	0.294	-0.042	0.008	0.926	0.000	-0.059	0.010	0.891	0.000
		500	-0.173	0.032	0.053	-0.045	0.005	0.911	0.000	-0.060	0.007	0.813	0.000
		750	-0.175	0.032	0.009	-0.047	0.004	0.837	0.000	-0.062	0.006	0.738	0.000
	A3	250	-0.270	0.083	0.228	0.024	0.030	0.907	0.025	0.070	0.043	0.886	0.000
		500	-0.269	0.078	0.054	< 0.001	0.015	0.929	0.014	0.043	0.020	0.925	0.000
		750	-0.269	0.076	0.011	-0.002	0.010	0.938	0.000	0.041	0.014	0.917	0.000

approaches is not possible in the context of frailty models since the correct estimation of the weights would require knowing the true value of the frailty variance, which cannot be correctly estimated under outcome-dependent sampling.

5.3.3 Simulation results

Simulation I

We first present the results obtained when data is generated under the assumption of fully observed heterogeneity. Both the level of outcome-dependent sampling and the covariate effect size determine the observed differences among the studied methods. If the covariate effect is strong ($\beta = 1$), the naive unweighted method is outperformed by the weighted approaches, even when the level of outcome-dependent selection is low (scenario A1). When the covariate effect is weaker ($\beta = 0.3$), and the level of ascertainment is medium

(scenario A2) or high (scenario A3), both weighted cohort methods perform similarly and clearly outperform the naive, unweighted approach. Under a weak level of ascertainment (scenario A1), the new generalized weighted cohort method performs as well as the naive unweighted approach and they slightly outperform the traditional weighted cohort approach. Importantly, the traditional weighted approach is often not applicable when the assumed covariate effect is weak. Negative weights are often obtained in this setting due to violation of condition (Eq. 5.10). The same problem is observed in the null case scenario when assuming $\beta = 0$. The new generalized weighted cohort method does not suffer from this problem, yielding valid and satisfactory results in all studied scenarios.

Simulation II

Table 5.2: Simulation II. Relative bias (reBias), mean square error (MSE) and coverage probability (Coverage) for $\widehat{\beta}$ along 1000 trials. A1: mild level of ascertainment. A2: medium level of ascertainment A3: strong level of ascertainment. N: number of families. Data is generated according to a shared frailty model with frailty variance θ. For the weighted approaches, the proportion of invalid (negative) weights along 1000 trials is also reported.

				Unweighted			**Weighted cohort**				**Generalized weighted cohort**				**Shared frailty**		
θ	β	Scenario	N	reBias	MSE	Coverage	reBias	MSE	Coverage	Invalid weights	reBias	MSE	Coverage	Invalid weights	reBias	MSE	Coverage
$\theta = 0.1$	$\beta = 0$	A1	500	0.002	0.002	0.937	0.003	0.002	0.949	0.000	0.003	0.002	0.953	0.000	< 0.001	0.002	0.939
		A2	500	< 0.001	0.002	0.953	0.003	0.005	0.952	0.000	0.003	0.004	0.952	0.000	< 0.001	0.002	0.954
		A3	500	< 0.001	0.006	0.945	-0.002	0.012	0.941	0.036	< 0.001	0.015	0.936	0.000	0.002	0.006	0.942
	$\beta = 0.3$	A1	500	-0.065	0.002	0.918	-0.092	0.002	0.895	0.000	-0.060	0.002	0.934	0.000	-0.069	0.002	0.914
		A2	500	-0.208	0.006	0.705	-0.087	0.005	0.926	0.000	-0.119	0.004	0.896	0.000	-0.217	0.006	0.688
		A3	500	-0.239	0.011	0.824	-0.037	0.013	0.951	0.206	-0.017	0.016	0.941	0.000	-0.228	0.011	0.848
	$\beta = 1$	A1	500	-0.094	0.010	0.363	-0.059	0.006	0.766	0.000	-0.032	0.003	0.895	0.000	-0.092	0.010	0.394
		A2	500	-0.195	0.040	0.027	-0.057	0.007	0.865	0.000	-0.080	0.010	0.734	0.000	-0.194	0.049	0.030
		A3	500	-0.278	0.083	0.045	-0.012	0.017	0.921	0.000	0.017	0.022	0.912	0.000	-0.278	0.083	0.063
$\theta = 1$	$\beta = 0$	A1	500	0.001	0.002	0.945	< 0.001	0.003	0.940	0.000	0.001	0.002	0.939	0.000	< 0.001	0.006	0.954
		A2	500	< 0.001	0.002	0.944	-0.004	0.008	0.933	0.000	-0.002	0.006	0.926	0.000	-0.002	0.002	0.957
		A3	500	0.001	0.007	0.947	-0.003	0.025	0.940	0.000	-0.002	0.023	0.912	0.000	-0.001	0.007	0.950
	$\beta = 0.3$	A1	500	-0.238	0.007	0.613	-0.130	0.005	0.885	0.000	-0.164	0.006	0.854	0.000	-0.131	0.004	0.830
		A2	500	-0.279	0.009	0.567	0.070	0.008	0.939	0.070	-0.043	0.006	0.940	0.000	-0.231	0.007	0.667
		A3	500	-0.310	0.015	0.793	0.076	0.044	0.879	0.480	-0.051	0.039	0.899	0.000	-0.323	0.016	0.756
	$\beta = 1$	A1	500	-0.263	0.072	0.001	-0.151	0.028	0.391	0.000	-0.177	0.037	0.262	0.000	-0.110	0.015	0.391
		A2	500	-0.293	0.090	0.001	0.001	0.010	0.908	0.000	-0.081	0.014	0.794	0.000	-0.188	0.040	0.123
		A3	500	-0.355	0.135	0.024	-0.045	0.040	0.858	0.492	-0.100	0.051	0.841	0.000	-0.355	0.134	0.028

Table 5.2 shows the results assuming the presence of unobserved family-shared heterogeneity. When unmeasured within-family correlation is mild ($\theta = 0.1$), we found similar results as in the previous simulation study: weighted methods perform similarly and provide better results than the unweighted model. Also, weighted methods, which deal with outcome-dependent sampling but ignore the presence of unobserved heterogeneity, perform better than a gamma shared frailty model which deals with shared unobserved heterogeneity but ignores outcome-dependent sampling.

For strong within-family correlation ($\theta = 1$) the performance of both weighted methods is, in general, good if the level of ascertainment is moderate or high (scenarios A2 and A3). If the level of ascertainment is mild (scenario A1), weighted methods would still outperform the traditional unweighted Cox approach but a shared frailty model seems a better choice in this setting. Bias is still noticeable with the shared frailty model, but of a smaller magnitude. Finally, the original weighted cohort also provided negative weights in this setting with unobserved family-shared heterogeneity, while the newly proposed generalized weighted cohort method proved to be more robust.

Overall, the simulation results show that the new generalized weighted cohort method is preferred over the original weighted cohort approach proposed by Antoniou *et al.* 2005. The original weighted cohort method performs well, in general, in the presence of a combination of a strong covariate effect and strong outcome-dependent sampling, as expected. However, its applicability is restricted to certain scenarios, and it is not general enough. Our sensitivity analysis, based on assuming the existence of unobserved heterogeneity, shows that inverse probability of selection weighted Cox models can still perform properly in the presence of mild unobserved family-shared heterogeneity, but they lead to biased results when the size of the frailty variance is large. Still, weighted methods seem to be preferred over the alternative approach of ignoring outcome-dependent sampling and fitting a shared frailty model if the level of outcome-dependent sampling is strong. If the level of ascertainment is mild, the results indicate a preference for the shared frailty model.

5.3.4 Software implementation

The generalized weighted cohort method developed in this work was implemented in the user-friendly `R` package `wcox`, which can be downloaded from CRAN and `https://github.com/vharntzen/wcox`.

5.4 Real data applications

We present two applications to illustrate the performance of the new generalized weighted cohort method compared to the traditional approaches on real data. In both applications, the goal is to assess the association between common susceptibility *loci* (gene locations on the chromosome) identified in Genome Wide Association Studies (GWAS) and cancer,

using data collected through genetic testing in clinical genetics units. Specifically, the first application is devoted to study the association between a Single Nucleotide Polymorphism (SNP) and colorectal cancer (CRC) in carriers of a pathogenic variant in the *PMS2* gene while the second one focuses on the association of a 161 SNP-based polygenic risk score with breast cancer. The selection of both datasets was based on family history of cancer with oversampling of cancer cases with the aim of finding carriers of certain genetic variants. As a result, the sample used in the first application is composed of *PMS2* mutation carriers. In the second application, the sample is composed of women with a family history of breast cancer and without BRCA1 or BRCA2 mutations.

5.4.1 Application to colorectal cancer

In this application, we consider a sample of male carriers of the germline PMS2 mutation. Motivated by the previous promising findings reported by Ten Broeke *et al.* [2018], we studied the association between the SNP rs1321311 and colorectal cancer in men. The sample consisted of 191 males belonging to 102 different families collected in eight Dutch clinical genetics centers between 2007 and 2016. Details on the selection criteria can be found in Ten Broeke *et al.* [2015]. The distribution of the number of individuals belonging to the same family was very skewed, the mean number of individuals per family was 1.83 and most of the families (55 %) contributed with one single member (Figure 2, left panel). The last age of follow-up ranged between 25 and 88 years, but given that no events were observed after 75 years old, we censored observations at 75 years. The range of observed ages at CRC diagnosis varied between 25 and 75, and 58 events were observed. From the 191 studied individuals, 116 were homozygotes of the non-risk allele, 65 were heterozygotes and 10 were homozygotes of the risk allele. Because of the limited size of the last category, we evaluated the effect of the indicator of being a carrier of the rs1321311 allele.

We considered four different models: unweighted Cox regression, the state-of-the-art weighted cohort method, our new method based on the new and more general weighting scheme, and a shared gamma frailty model as a sensitivity analysis to measure the potential impact of unobserved family-specific heterogeneity. The two studied weighted methods require the knowledge of incidence rates for CRC in carriers of pathogenic variants in *PMS2*. These were obtained by multiplying the population-based incidence rates of

CRC in the Netherlands in 2011 [Netherlands Cancer Registry, 2021] by the previously published [Ten Broeke et al., 2015] age-dependent hazard ratios of CRC for *PMS2* carriers. The choice of the year 2011 as the reference is justified because it is the middle point of the data collection period (2007-2016). The specific age-specific intervals and incidence rates used in this application can be found in Supplemental Table 5.6.

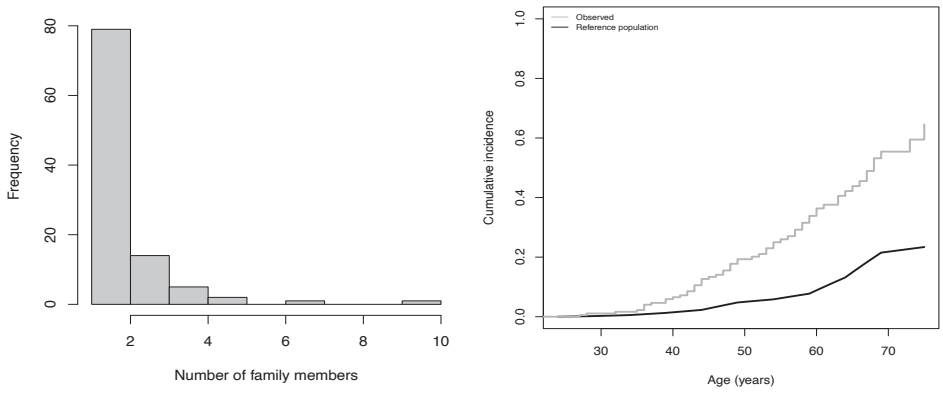


Figure 5.1: Application 1: Study of the association between SNP rs1321311 and CRC cancer in male carriers of a pathogenic variant in the gene *PMS2*. Left panel: Size of the families included in the sample. Right panel: Cumulative incidence of colorectal cancer at different ages. The grey line shows the observed risk in the sample. The black line reflects the expected cumulative colorectal cancer risk for the population of *PMS2* mutation carriers based on previous literature [Ten Broeke et al., 2015]. Specifically, age-specific CRC incidence rates of *PMS2* mutation carriers are obtained multiplying the point estimates of the age-dependent hazard ratios as reported in Table 2 in Ten Broeke *et al.* [2015] by the underlying population-based incidence rates of CRC for males in the Netherlands in 2011 according to the Netherlands Cancer Registry (NCR).

From the results reported in the bottom line of Table 5.3, it is observed that the new generalized weighted cohort method provides slightly larger estimated effects than the well-known (unweighted) Cox regression. In agreement to the result obtained with the unweighted method, the estimated association between the risk allele rs1321311 and CRC was statistically significant at the usual 5% level when using the new method. Importantly, the traditional weighted cohort approach could not be used because negative weights were obtained. Specifically, the oversampling of cases was not strong enough in the age group 65-70 years old and restriction (5.10) discussed in Section 5.2.1 was not met leading to negative weights for unaffected individuals in this age group. The shared frailty model provides the lower estimated covariate effect among the evaluated methods. This is

probably due to the limited sizes of the family clusters and a small underlying unobserved heterogeneity. The estimated frailty variance was 0.15 with a broad confidence interval (0-1), indicating difficulties of the model to give reliable estimates of the level of unobserved heterogeneity. A likely major driving cause for this difficulty is the limited cluster size of this application since most of the families contribute a single individual to the analysis. As a consequence, the shared frailty approach is not recommended in this application and one would rather choose the new generalised weighted cohort approach.

Table 5.3: Application to CRC in male carriers of PMS2. Estimated regression coefficients ($\widehat{\beta}$) and corresponding 95% Confidence Intervals for the effect of the SNP rs1321311 for different Cox models. Case weights are calculated based on incidence rates of CRC for *PMS2* mutation carriers defined as the point estimates of the age-dependent hazard ratios reported in Ten Broeke *et al.* [2018] multiplied by the population-based rates of CRC in Netherlands in 2011.

Model	$\widehat{\beta}$ **(95% CI)**
Unweighted	0.723 (0.182; 1.265)
Frailty	0.671 (0.149; 1.192)
Weighted cohort	- (*negative weights*)
Generalized weighted cohort	0.771 (0.234; 1.308)

5.4.2 Application to breast cancer

In this application, the association between a PRS score and breast cancer was analyzed using a sample of 579 clinically ascertained women belonging to 101 families. On average, six women were included per family (mean family size = 5.73 and standard deviation = 4.66, Figure 2 right panel). The inclusion criterion was two-fold. Per family, one of the women should be tested negative for BRCA1 or BRCA2 pathogenic variants. This was a special feature of this sample and means that family aggregation and early-onset of cancer are not explained by pathogenic variants in these high-risk genes. Furthermore, breast cancer had to occur in at least three female family members or in two females if at least one had bilateral breast cancer before the age of 60. The families were selected between 1990 and 2012 by Clinical Genetic Services in four Dutch cities (Groningen, Leiden, Nijmegen and Rotterdam) and one Hungarian city (Budapest). Given the scarcity of events after 80 years of age (only one observed event at 90), we censored observations at age 80. The PRS was based on 161 SNPs weighted by previously published log-odds ratios (mostly based on population-based case-control studies). Detailed description of the calculation of PRS can

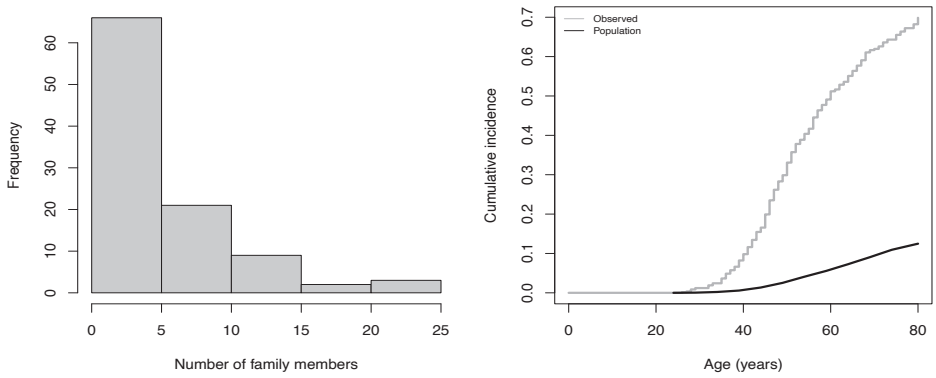


Figure 5.2: Application 2: Study of the association between a polygenic risk score and female breast cancer. Left panel: Size of the families included in the sample. Right panel: Cumulative incidence of breast cancer at different ages. The gray line shows the observed risk in the sample. The black line shows the population-based (the Netherlands, 2001 [Netherlands Cancer Registry, 2021]) cumulative incidence used as reference in the weighted analyses.

be found elsewhere [Lakeman et al., 2019]. As before, to establish the association between the marker of interest, the PRS, and breast cancer, we considered four different models: the traditional unweighted Cox regression, the state of the art weighted cohort method to deal with outcome-dependent sampling, our new weighted method and a shared gamma frailty model. Population-based incidence rates of the Netherlands in 2001 [Netherlands Cancer Registry, 2021] (mid point of the sample selection period) were used as external input to construct the weights. The specific age-specific intervals and incidence rates used in this application can be found in Supplemental Table 5.6.

From the results reported in Table 5.4, we observe that the new method provides a slightly smaller effect than the previously proposed weighted cohort approach and that both provided smaller effects than the unweighted Cox model. None of these three approaches reached statistical significance at the 5% level. In order to estimate the level of heterogeneity due to unmeasured within-family similarity, a shared frailty model was also fitted. The estimated frailty variance was 0.41, indicating that unobserved heterogeneity is not negligible in this application. This, together with the large size of the included families, is probably the reason why the shared frailty model seems to outperform the other methods. The estimated conditional hazard ratio using a shared frailty model is larger

than the ones obtained using unweighted and weighted versions of the Cox model even if statistical significance at the 5% level is also not reached with this approach. According to our simulation results, we infer that the association between PRS and breast cancer is likely obscured by ignoring the strong unobserved heterogeneity and that the frailty approach is preferred in this application.

Table 5.4: Application to female breast cancer in non-BRCA1/2 families. Estimated regression coefficients ($\widehat{\beta}$) and corresponding 95% confidence intervals for the effect of polygenic risk score (PRS) for different Cox models.

Model	$\widehat{\beta}$ **(95% CI)**
Unweighted	0.110 (-0.096; 0.317)
Frailty	0.173 (-0.045; 0.390)
Weighted cohort	0.079 (-0.226; 0.385)
Generalized weighted cohort	0.062 (-0.261; 0.384)

5.5 Discussion

In this paper, we have revisited the analysis of outcome-dependently sampled survival data with weighted Cox regression using external data to construct inverse probability of selection weights. Our research is motivated by the interest in the effect of potential modifying factors on cancer risk using clinically ascertained data. Typically, those data sets are collected through ongoing genetic testing programs, where selection criteria lead to an over-representation of young cases and hence, the resulting samples are not representative of the target population of interest. We proposed a new weighting scheme that restores the expected ratio of events and non-events at each follow-up time using population-based hazard information. Our simulation study has shown that the new method can be applied to a broader set of realistic scenarios. Our real data applications support the same conclusion indicating the broader applicability of the new weighting scheme and it should be the preferred option to analyze data obtained under family-based outcome-dependent sampling when unobserved heterogeneity is negligible or mild.

A strength of the new weighting scheme is that it relies on fewer assumptions to provide valid, non-negative weights. The traditional weighted cohort [Antoniou et al., 2005] approach requires that a number of conditions are fulfilled, which hamper its applicability. Specifically, the original method is problematic if oversampling of cases is not observed in

all age groups. In practice, although overall oversampling of events is expected, it does not necessarily hold for all age groups. Our new method overcomes this restriction and can be applied to a wider set of oversampling schemes, hence it can be regarded as a generalization of the traditional weighted cohort approach. This together with user-friendly implementation makes it an attractive analysis tool for applied researchers in the field.

Likewise the previously proposed weighted cohort method, our approach relies on a number of assumptions. First, a crucial assumption is the existence of a well-established external source of population-based incidence rates. Second, the sampling probabilities of observed individuals depend on the age at onset but they are assumed to be conditionally independent of the risk modifier under investigation. These two assumptions have been previously discussed in the context of the weighted cohort method [Antoniou et al., 2005; Barnes et al., 2012]. Furthermore, the relationship between the hazard and the risk modifier under investigation should approximately follow a proportional hazards specification. We have examined the performance of both the traditional and the new generalised weighted cohort approaches under model misspecification, specifically, under non-collapsability due to the presence of residual familial aggregation. In this case, we have also observed that the use of weighted approaches seems advisable compared to the naive unweighted approach. Additionally, if the number of available individuals per family is limited, which is the most common situation in practice, our new method might be the preferred option, outperforming a shared frailty model and the traditional weighted cohort approach. However, we would like to caution about the interpretation of the estimated effect and point out the systematic downward bias of the regression coefficient in this setting, proposing the systematic inclusion of the results of a shared frailty model as a sensitivity analysis.

The extension of weighting approaches to deal with outcome-dependent sampling to the context of frailty models would be interesting but challenging. Since the estimated incidence rate in the sample depends on the correct estimation of the frailty variance, it would be necessary to know the value of the frailty variance to derive correct weights. However, the frailty variance is latent and hence we anticipate an identifiability problem in such an approach. More sophisticated modeling, using a frailty model with explicit correction for ascertainment is possible but not straightforward and it is left as future research. It is noteworthy that such a complex approach will presumably require large clusters and sample sizes and hence our simpler approach based on borrowing information

from a trustworthy external source will still be preferred in a number of relevant practical situations, such as our application to PMS2 carriers.

In conclusion, for performing regression analysis using survival data obtained under family-based outcome dependently sampling, specialized techniques are required to avoid bias and provide valid inference. We have proposed an accurate and conceptually simple method which generalizes and outperforms existing methods based on weighted Cox regression.

Acknowledgments

The departments of Clinical Genetics and Human Genetics (LUMC, Leiden) are gratefully acknowledged for providing the breast cancer data set.

5.6 Supplementary material

5.6.1 Unbiasedness of the inverse probability of selection weighted Cox approach

Denote by D the selection event, and let $W(t)$ be the probability of being selected given that that the observed event time is t, assume that $W(t) > 0$ on the support of T.

In the absence of covariates, the density of the observed data, obtained under biased (in particular within an outcome-dependent sampling where $D \perp Z|T$) is given by the following weighted density:

$$f_{T|D}(t) = \frac{P(D=1|T=t)f_T(t)}{\int_0^\infty P(D=1|T=s)f_T(s)ds}.$$

Accordingly, the joint density of the sampled information including covariate Z is given by:

$$f_{T,Z|D}(t,z) = \frac{W(t)h(t|z;\beta)exp(-\int_0^t h(y|z;\beta)dy)f_Z(z)}{E(W(t))}.$$

The density of the event times conditional on the covariate is therefore:

$$f_{T|Z,D}(t,z) = \frac{W(t)h(t|z;\beta)exp(-\int_0^t h(y|z;\beta)dy)}{E(W(t)|z)}.$$

Note that $E(W(t)|z) = P(D = 1|t, z)$ depends on β and thus this biased sampling must be accounted for in the estimation of β. However, since the weight $W(t)$ is a function of t alone, based on the key assumption $D \perp Z|T$, it follows that $f_{Z|T,D} = f_{Z|T}$ and the following standard probabilistic result from the Cox model can be used:

$$E(Z|T = t, D) = E(Z|T = t) = \frac{E\left[Ze^{\beta Z}\overline{F}_{T|Z}(t|Z)\right]}{E\left[e^{\beta Z}\overline{F}_{T|Z}(t|Z)\right]},$$

where $\overline{F} = 1 - F$ denotes the survival function.

Let $f_{Z|D} = E(W(t)|z)f_Z(z)/E(W(T))$ denote the marginal weighted density of the covariate, then we can adapt the former general result to our setting and rewrite it as follows:

$$E(Z|T = t, D) = \frac{E\left[Ze^{\beta Z}\overline{F}_{T|Z}(t|Z)/E(W(T)|Z)|D\right]}{E\left[e^{\beta Z}\overline{F}_{T|Z}(t|Z)/E(W(T)|Z)|D\right]}.$$

The former expression still involved functionals of Z and T unconditionally on D. To rewrite the expectation as a function of observed variables we use the expression on the density of the event times conditional on the covariate $f_{T|Z,D}(t, z)$ and the fact that $W(t) > 0$ on the support of T, which implies $\overline{F}_{T|Z}(t|z)/E(W(T)|Z = z) = E\left[W(T)^{-1}I(T \geq t)|Z = z, D)\right]$. As a result:

$$E(Z|T = t, D) = \frac{E\left[Ze^{\beta Z}W(T)^{-1}I(T \geq t)|D\right]}{E\left[e^{\beta Z}W(T)^{-1}I(T \geq t)|D\right]},$$

which implies the unbiasedness of the weighted estimating equation

$$U(\beta) = \sum_{i=1}^{n}\{Z_i - \frac{\sum_{j-1}^{n} Z_j e^{\beta Z_j}(W(T_j)^{-1}I(T_j \geq T_i)}{\sum_{j-1}^{n} e^{\beta Z_j}(W(T_j)^{-1}I(T_j \geq T_i)}\}.$$

5.6.2 Literature review: family size and the use of the weighted cohort approach

Among the 81 citations of Antoniou *et al.*'s 2005 paper [Antoniou et al., 2005], we found 51 papers that used a weighted cohort approach to obtain unbiased Hazard Ratios in a Cox model (list available upon request). The majority (62.7%, n = 32) were studies into breast and ovarian cancer risks. In 48 of the 51 papers applying the weighted cohort approach, the study sample certainly includes multiple members per family, but the exact number of families was only mentioned in 19 papers, shown in Table 5.5. For each study, we calculated the average number of family members included. Note that this may fluctuate: one study [Andrieu et al., 2006] described the exact family composition of the sample, see Table 5.5 footnote 6. The median of the paper-specific, average family cluster sizes was 2.5. Generally, this data was collected at family cancer- or genetics clinics, where relatives of the index case (proband) were invited to be tested. Sometimes this was combined with 'population-based' recruitment [Chau et al., 2016; Ait Ouakrim et al., 2015].

Table 5.5: Family information (when reported) in papers applying weighted cohort approach. This list is a subset of all (81) PubMed citations of the paper of Antoniou *et al.* [2005] on 06/02/2021, with inclusion criteria 1) applying weighted Cox, 2) mentioning the number of families and sample size.

Authors	**Year**	**Sample**	**Average number of relatives per family**	**(Cancer) research area**
Borde *et al.*	2020	578 families; 760 carriers	1.6	Breast
Dashti *et al.*	2018	774 families; 2042 carriers	2.6	Ovarian
Ten Broeke *et al.*	2018	152 families; 521 samples	3.4	Colorectal
Kamiza *et al.*	2016	62 families; 260 carriers	4.2	Endometrial
Dashti *et al.*	2017	761 families; 1925 carriers	2.5	Colorectal
Chau *et al.*	2016	15049 families; 42489 participants[(1)]	5.3	Colorectal
Win *et al.*	2015 (b)	330 families; 1098 carriers	3.3	Colorectal
Win *et al.*	2015 (a)	593 families; 854 individuals	1.4	Colorectal
Dashti *et al.*	2015	548 families; 1128 women	2.1	Breast
Ait Ouakrim et al.	2015	748 families; 1858 carriers[(2)]	2.5	Breast
Pooley et al.	2014	3134 families; 4822 included[(3]	1.5	Breast
Killick et al.	2014	115 families; 158 included[(4)]	1.4	Breast
Win et al.	2013	315 families; 927 carriers	2.9	Breast
Pijpe et al.	2012	930 families; 1122 carriers[(4)]	1.2	Breast
Win et al.	2011 (a)	498 families provided 1324 carriers; 287 families provided 1219 non-carriers	2.7 (carriers); 4.2 (non-carriers)	Colorectal
Win et al.	2011 (b)	286 families provided 601 carriers; 182 families provided 533 non-carriers	2.1 (carriers); 2.9 (non-carriers)	Endometrial
Amos et al.	2010	93 families; 489 included (of which 45 married-in)	5.3	Ovarian
Antoniou et al.	2006	392 families; 810 carriers	2.1	Breast
Andrieu et al.	2006	1074 families; 1601 women	1.5[6]	Breast

(1) Includes some population-based recruitment for which average number of recruited relatives per family was 2.6, vs. 5.3 for clinic-based families.
(2) 25% population-based.
(3) Sampled non-carrying relatives only.
(4) Relatives functioned as controls, i.e. non-carriers.
(5) This concerns only one of the data sets in the paper.
(6) The relative occurrence of relatives per family was 71.1% for size 1, 17.9% for size 2, 6.2% for size 3, 2.8% for size 4 and 2.0% for size 5 up until 11.

5.6.3 Population incidence rates used in real data applications

Table 5.6: Population-based age-specific incidence rates (in cases per 100.000) used for weight construction. For breast cancer, we used the registered data of The Netherlands in 2001 for women [Netherlands Cancer Registry, 2021]. For colorectal cancer, age-specific incidence rates were obtained by multiplying the population-based incidence rates of CRC in the Netherlands in 2011 [Netherlands Cancer Registry, 2021] by the previously published [Ten Broeke et al., 2015] age-dependent hazard ratios of CRC for *PMS2* carriers. The choice of the year 2011 as the reference is justified because it is the middle point of the data collection period (2007-2016).

Age group	Colorectal (men)	Breast (women)
25-30	40.56	9.09
30-34	69.39	33.61
35-39	146.60	72.41
40-44	204.38	156.17
45-49	518.28	241.48
50-54	221.10	323.56
55-59	411.68	310.83
60-64	1214.11	363.63
65-69	2016.21	388.41
70-74	405.12	415.42
75-79	-	293.22

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.

List of Publications

Academic papers

Arntzen, V.H., Feenstra, S.G., Benincà, E., Le, T.T.N., Mascini, E.M., Nabuurs-Franssen, M.H., Voss, A., Marik, A.M., Jong, E. de, Silvis, W., Schijffelen, M.J., Schneeberger, P., Hopman, J., Korthals Altes, H. Wertheim, H.F.L. (contributed equally) [2023, most work in 2017] 'Spatial analysis of methicillin-resistant Staphylococcus aureus carriage (MRSA) at hospital admission in a livestock dense region', *medRXiv (preprint)*, doi: 10.1101/2023.05.01.23289266

Smith A.F., Huss A., Dorevitch, S., Heijnen, L., **Arntzen, V.H.**, Davies, M., Robert-Du Ry van Beest Holle, M., Fujita, Y., Verschoor, A.M., Raterman, B., Oesterholt, F., Heederik, D., and Medema, G. [2019] 'Multiple Sources of the Outbreak of Legionnaires' Disease in Genesee County, Michigan, in 2014 and 2015', *Environmental Health perspectives*, doi: 10.1289/EHP566

Arntzen, V. H., Fiocco, M., Leitzinger, N. and Geskus, R. B. [2023], 'Towards robust and accurate estimates of the incubation time distribution, with focus on upper tail probabilities and SARS-CoV-2 infection', *Statistics in Medicine*, doi: 10.1002/sim.9726

Arntzen, V.H., Fiocco, M. and Geskus, R.B. [2024], 'Two biases in incubation time estimation related to exposure', *BMC Infectious Diseases*, doi: 10.1186/s12879-024-09433-7

Vera H. Arntzen, Manh Nguyen Duc, Marta Fiocco, Lan Truong Thi Thanh, Tam Nguyen Hoai Thao, Buu Mai Thanh, Tu-Anh Nguyen, Nhat Le Thanh Hoang, Marc Choisy, Lam Phung Khanh, Nga Le Hong and Ronald B. Geskus, on behalf of the Covid-19 modelling team Oxford University Clinical Research Unit, Vietnam^{+} [2024] 'The latency time of the SARS-CoV-2 Delta variant in naive individuals from Vietnam', *to be submitted*

[+] Duc Du Hong, Lam Phung Khanh, Leigh Jones, Marc Choisy, Nhat Le Thanh Hoang, Ronald Geskus, Sonia Lewycka, Thomas Kesteman, Trinh Dong Huu Khanh, Tung Trinh Son, Manh Nguyen Duc, Nguyet Nguyen Thi Minh, Thinh Ong Phuc, Trang Duong Thuy, Lieu Tran Thi Bich, Maia Rabaa.

Arntzen, V. H., Fiocco, M., Lakeman, I.M.M., Nielsen, M., Rodríguez-Girondo, M. [2024] 'A new inverse probability of selection weighted Cox model to deal with outcome-dependent sampling in survival analysis', *to be published in Biometrical Journal*

R Software packages

Arntzen, V.H. and Rodríguez-Girondo, M. [2023] '`wcox`: weights to correct for outcome dependent sampling in time to event data', *available from CRAN*

Arntzen, V.H. [2024] '`doublIn`: estimate incubation or latency time using doubly interval censored observations', *available from CRAN*

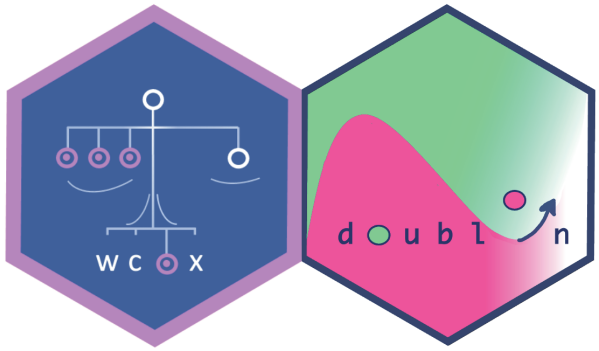


Figure 6.1: Logos corresponding to the two `R` Software packages developed for the work in this thesis: `wcox` and `doublIn`.

Summary

The incubation (infection to symptom onset) and latency time (infection to start-of-infectiousness) are crucial quantities to inform control measures at the beginning of an infectious disease outbreak. An example is the duration of quarantine for potentially infected individuals, a control measure that was frequently imposed after SARS-CoV-2 emerged in 2020. This thesis is inspired by the Vietnamese context at the time. With limited intensive care capacity and a long-stretched border with China, Vietnam initially strived to prevent any transmission of SARS-CoV-2. Government-allocated quarantine facilities in which individuals were regularly tested for presence of the infection provided a unique data set.

What is known about the time of infection is usually limited to the window of exposure, i.e. the interval from the first to the last possible moment of infection, such that the start point of incubation time observations is *interval censored*. Common practice is to assume that (i) the risk of infection within the exposure window is constant and (ii) incubation time follows a *gamma*, *lognormal* or *Weibull* distribution (think of three different baking forms for cake dough). However, during the beginning of an outbreak, the daily number of new cases grows exponentially and coronaviruses are known to have a long, right-tailed incubation time distribution, leaving (i) and (ii) unrealistic. In Chapter 2, we investigated this issue by generating toy data sets. We observed that a model that allows for more flexibility in the shape of the distribution provides a better fit to the right-hand tail. For the estimation of the latent period of SARS-CoV-2 with data from Vietnam we assumed an increasing risk of infection within the exposure window (Chapter 4).

Contract tracing aims to notify cases soon after infection to prevent further transmission. Exposure information is retrieved retrospectively, through interviews with notified cases. Estimates of incubation time typically use such contact tracing data. Differential recall may occur: at the time of the interview, infected individuals may recall recent exposure more precisely than less recent exposure. To mitigate the impact of violating assumption (i), analysis is often restricted to observations with well-defined exposure. However, in the presence of differential recall this restriction leads to underestimation of the incubation time.

In Chapter 3, we discuss this issue and another phenomenon that has been overlooked in early estimates of the incubation time distribution of SARS-CoV-2.

Quarantine length for SARS-CoV-2 was typically informed by estimates of the incubation time, even though two out of five infected individuals will never develop symptoms. Latency time would be a more logical quantity for informing this decision, but its estimates are sparse for two reasons: the required data is rarely available and estimation is complicated further because start-of-infectiousness cannot be exactly observed. Utilizing RNA shedding as a proxy for infectiousness, the start-of-infectiousness likely took place between the last negative and first positive test for SARS-CoV-2. Therefore, observations of latency time consist of two windows such that standard methods cannot be used. We developed an `R` software package (`doublIn`) suitable for this type of observations. Using unique data from Vietnam and realistic model assumptions, we estimated the latency time of the SARS-CoV-2 Delta variant for unvaccinated and non-immune individuals in Vietnam in 2021 (Chapter 4).

The incubation and latency time are examples of time-to-event data, the type of data that the survival analysis field within statistics studies. Another example is the age at which a woman develops breast cancer. Understanding the risk associated with genetic variants like BRCA1/2 is crucial as it informs decisions about precautionary measures such as preventive surgical removal of the breast. Estimates are typically based on data from high-risk families with multiple affected members, requiring a tailored approach to provide estimates applicable to individuals from low-risk families as well. In Chapter 5, we generalized the state-of-the-art approach. Our software is openly available for researchers working on similar estimation problems (`R` package `wcox`).

Samenvatting

De incubatietijd (infectie tot de start van symptomen) en latente periode (infectie tot start-van-infectieusheid) zijn cruciale informatie voor controle maatregelen aan het begin van een infectieziektenuitbraak. Een voorbeeld is de duur van quarantaine van potentieel geïnfecteerde individuen, een controle maatregel die geregeld is opgelegd sinds SARS-CoV-2 in 2020 de kop opstak. Dit proefschrift is geïnspireerd door de Vietnamese context in die tijd. Met een beperkte capaciteit van de *intensive care* en een uitgestrekte grens met China streefde Vietnam er in eerste instantie naar om alle verspreiding van SARS-CoV-2 te voorkomen. Door de overheid aangewezen quarantaine faciliteiten waar individuen herhaaldelijk getest werden op aanwezigheid van de infectie verschaften een unieke dataset.

Wat betreft de tijd van infectie is meestal slechts de periode van blootstelling bekend, dat wil zeggen het interval van het eerst tot het laatst mogelijke moment van infectie. Daarom is het beginpunt van observaties van de incubatietijd *interval gecensureerd*. Normaalgesproken wordt er aangenomen dat (i) het risico op infectie binnen de blootstellingsperiode constant is en dat (ii) de incubatietijd een *gamma*, *lognormaal* of *Weibull* verdeeld is (denk hierbij aan drie verschillende bakvormen voor een cakebeslag). In het begin van een uitbraak groeit het dagelijke aantal nieuwe ziektegevallen echter exponentieel en daarbij komt dat we weten dat de incubatietijdverdeling van corona virussen een lange rechterstaart heeft. Daarom zijn (i) en (ii) niet erg realistisch. In Hoofdstuk 2 onderzochten we dit door oefendata te genereren. We zagen dat een model wat meer flexibiliteit toestaat in the vorm van de verdeling, een betere pasvorm heeft in diens staart. Bij het schatten van de latente periode van SARS-CoV-2 met data uit Vietnam namen we aan dat het risico op infectie toenam tijdens de blootstellingsperiode (Hoofdstuk 4).

Contactonderzoek poogt ziektegevallen vlak na het moment van infectie te notificeren om verdere verspreiding te voorkomen. Blootstellingsgegevens worden retrospectief verkregen door interviews met genotificeerde ziektegevallen. Schattingen van de incubatietijd worden gewoonlijk gebaseerd op dergelijke gegevens die zijn verkregen tijdens

contactonderzoek. Individuen kunnen zich blootstellingen verschillend herinneren: tijdens het interview herinneren geïnfecteerden zich blootstelling preciezer wanneer deze recent plaatsvond dan langer geleden. Om vertekening van de schatting door de ongeldigheid van aanname (i) zoveel mogelijk te beperken, analyseert men geregeld alleen de observaties met een duidelijk gedefinieerde blootstellingsperiode. Wanneer individuen zich blootstelling verschillend herinneren zoals eerder beschreven, wordt de incubatietijd onderschat. In Hoofdstuk 3 bespreken we dit en nog een ander fenemeen waar overheen gekeken is in de eerste schattingen van de incubatietijdverdeling van SARS-CoV-2.

De lengte van de quarantaineperiode voor SARS-CoV-2 werd normaliter gebaseerd op onder andere schattingen van de incubatietijd, terwijl twee van de vijf geinfecteerden nooit symptomen zal ontwikkelen. De latente periode zou logischerwijs veel informatiever zijn voor deze keuze, maar schattingen daarvan zijn helaas schaars om twee redenen: de benodigde data is vrijwel nooit beschikbaar en het schatten wordt verder bemoeilijkt doordat de start-van-infectieusheid niet exact kan worden waargenomen. Gebruikmakende van de aanwezigheid van RNA als maatstaf voor infectieusheid, is infectieusheid naar alle waarschijnlijkheid begonnen na de laatste negatieve en voor de eerste positieve test voor SARS-CoV-2. Observaties van de latente periode bestaan dus uit twee perioden rond het start- en eindpunt, waardoor we geen gebuik kunnen maken van gangbare methoden. We ontwikkelden een `R` software pakket (`doublIn`) wat geschikt is voor dit type observaties. Met unieke data uit Vietnam, en realistische aannames voor ons model, schatten we de latente periode voor de SARS-CoV-2 Delta variant voor ongevaccineerde, niet-immune individuen uit Vietnam in 2021 (Hoofdstuk 4).

De incubatietijd en de latente periode zijn zogenoemde *time-to-event* data, het type data wat bestudeerd wordt in het overlevingsanalyse veld binnen de statistiek. Een ander voorbeeld is de leeftijd waarop een vrouw borstkanker ontwikkelt. Begrip van het risico geassocieerd met genetische varianten als BRCA1/2 is cruciale informatie voor beslissingen over voorzorgsmaatregelen, zoals het preventief chirurgisch verwijderen van de borst. Schattingen zijn gewoonlijk gebaseerd op gegevens afkomstig van hoog-risico families met meerdere familieleden met de aandoening. Daarom is er maatwerk nodig om schattingen te verkrijgen die eveneens van toepassing zijn op individuen uit laag-risico families. In Hoofdstuk 5 veralgemeniseerden we de aanpak die op dit moment geldt als de gouden standaard. Onze software is openbaar beschikbaar voor onderzoekers die

werken aan vergelijkbare schattingsproblemen (R package wcox).

Curriculum Vitae

Vera Arntzen was born in Amsterdam, the Netherlands, on April 10th 1994. During primary education at the Bijlmermontessorischool she was awarded an oversized t-shirt and a science magazine subscription for her result on a national math test. During secondary education at Vossius Gymnasium she combined the natural sciences with art as an elective.

At Radboud University, Nijmegen (the Netherlands), she studied biomedical sciences (bachelor) and epidemiology (master) with electives in infectious diseases, global health, informatics and logic. During her bachelor she was study representative. Her interest in infectious disease research was sparked by lectures on malaria. She went to Copenhagen, Denmark, for an internship on influenza at WHO. Her graduation research concerned MRSA carriage in a livestock-dense region.

Fascinated by modelling the spread of infectious diseases research, she studied statistics (master) at Leiden University (the Netherlands). On the side, she worked on a retrospective investigation of an unresolved Legionnaires disease outbreak in Flint, United States (KWR Water Research Institute) and as a freelance journalist at Leidsch Universitair Weekblad Mare (local university newspaper). As part of an internship at Statistics Netherlands she used web scraping and machine learning techniques to assess the sustainability efforts of companies. Professor dr. Marta Fiocco supervised her master thesis in survival analysis techniques for familial breast cancer data and gave her the opportunity to pursue a PhD combining her interests in analytical problems and infectious diseases. The PhD was co-supervised by dr. Ronald Geskus from Oxford University Clinical Research Unit, a research institute dedicated to mostly infectious disease research in Vietnam, where she visited for at total of eight months. Besides her scientific work she teached introduction courses in statistics at Leiden University College and regularly organized activities for fellow PhDs.

Dankwoord

De weg naar een proefschrift is geen rechte lijn, eerder heuvelachtig. Maar, in de woorden van Veldhuis en Kemper, de allermooiste bloemen groeien vlak langs het ravijn, en het plukken daarvan lukte dankzij de hulp van zovelen, waarvan een aantal in het bijzonder bloemen toekomt.

Als eerste wil ik mijn (co-)promotoren bedanken. Marta, je creëerde een PhD positie op basis van mijn interesses ('statistiek', 'infectieziekten' en 'iets met het buitenland'). Je hield het overzicht en gaf me volop mogelijkheden om cursussen te volgen. Jouw enthousiasme is voor mij onlosmakelijk verbonden met dit hoofdstuk. Ronald, jouw inhoudelijke expertise was van onschatbare waarde. Als ik vastzat op een probleem merkte je luchtig op dat dat nou juist het allerleukste onderdeel was van onderzoek en dacht je mee. Eenmaal opgelost zag ik bij jou iets wat me inspireerde: het geluksmomentje van de onderzoeker in hart en nieren. De online meetings tijdens de pandemie verbreedden mijn horizon, net als de perioden in Vietnam daarna.

Promovendi in Leiden, bedankt voor het plezier wat we hebben gehad, in de kantine, Foo Bar, India, Italië, Portugal of elders. Nandan, je liep steevast langs op kantoor voor een praatje, dikwijls omdat je wist dat ik dat kon gebruiken. Ik ben ontzettend blij dat je naast een vriend ook paranimf bent. Maria, we begonnen dit avontuur samen en ik ben trots op beider paden. Erwin en Tessa, soms zijn sterke bonen en een gesprekje alles wat je nodig hebt en laat er dan net een nieuw koffietentje opstarten. Collega's van OUCRU waaronder Duc, James, Hoang, Hang, Thuong, voor ins en outs over Vietnam en infectieziekten, lunchgesprekken en gezelligheid. Nguyet, Nguyen en Dex voor de spelletjesavonden. Thao Dien Riders en Turtle Squad voor fietsritjes voor zonsopkomst en koffieleuten met uitzicht op skyscrapers. Kien en familie voor het onderdompelen in de Vietnamese cultuur en me thuis laten voelen in Ho Chi Minh City. Petra, voor het smachten naar de samenvatting.

Mama, je hebt altijd benadrukt dat school leuk mag zijn en daar ben ik je dankbaar voor. Papa, jouw onderzoeksverhalen legden de kiem voor dit proefschrift. Zus, vroeger koos je de twee efficiëntste supermarktrijen -een voor elk- om onze boodschappen razendsnel

te verplaatsen zodra eentje sneller bleek; nu promoveer je op wachtrijonderzoek en daar ben ik trots op. Broer, ik waardeer je eigenheid en oprechtheid. Een boom tot proefschrift verwerken of er zachtjes tegenaan tikken; geluk zit in allebei.

Tmnit, in groep twee zetten we samen onze eerste schreden in de schoolbankjes. Ik bewonder je vuur voor onrecht. Claartje en Laetitia, voor heel veel lol op de middelbare en de dierbare vriendschap die mijn tig verhuizingen wonderwel doorstond. Laetitia, dankzij jouw pixeltechniek en avondenlang post-its plakken herrees Jimi Hendrix op de schoolmuur en de skyline van Ho Chi Minh City op dit omslag. Mirte en Karin, de best denkbare (ex-)buuf. Elske en Loes, de rode draad door mijn (studenten)leven. Francis en Wiki, jullie leerden me dat er meer kan dan je zelf denkt. Charlotte, ik waardeer je open blik en onze gesprekken over het werkende leven. Noor, voor een stuk kletsen over hebben zodra de avond op is. Eva en Sylvia, voor het eetclubje wat uitgroeide tot een vriendschap die voor- hoofd en na overstijgt. Dank voor jullie support en de gezelligheid. Sebastiaan, Sofie, Aron en Ruben, de pandemie bracht ons dichter bij elkaar en een gezonde dosis taart. Bedankt dat jullie altijd voor me klaarstaan. Lieve Manon, Floor, Jill, Layla, Jannie, Lotte, Sanne, jullie blijven me inspireren. Bij ons mag alles op tafel komen (maar het liefste lekker eten en wijn). Doris, Esther, Roos (2x), voor het weekend in roeien ofwel geluk pur sang. Roos, alias huisgenootje, wildplukbuddy, roeimaatje - je was er als ik je nodig had, en verder gelukkig ook gewoon heel vaak. Arie, Bas, Mirjam en Vivina, voor de warme ontvangst en de fijne mensen die jullie zijn.

Lieve Myrthe, dit proefschrift was er niet geweest zonder jouw niet aflatende steun en ik ben blij dat je ook tijdens de verdediging naast mij staat. Dat je er het liefste een themafeestje van had gemaakt (de een ‘para’ en de ander ‘nimf’) is tekenend: al sinds het begin van onze studententijd hang je slingers op in mijn leven. Jouw fijne kijk op de mens en wereld, creativiteit en spontaniteit sieren je. Het is een voorrecht om een vriendschap te kennen als de onze.

Wouter, dankjewel voor het maatje wat je bent, je liefde en dat wij zijn. Ik hoor je al zeggen dat ik je daar niet voor hoef te bedanken, en dat is het mooiste en onbegrijpelijkste wat er is. Jouw interesse voor mijn onderzoek was onbegrensd; je bemoedigende woorden en grapjes over academia evenzo. Je had er een rotsvast vertrouwen in dat we mijn periode in Vietnam samen door zouden komen. Lieve Wouter, bloemen hoef je niet ("verleppen te snel") en wat let je: het beste uitzicht is dat op onze toekomstdromen.

Bibliography

Ait Ouakrim, D., Dashti, S. G., Chau, R., Buchanan, D. D., Clendenning, M., Rosty, C., Winship, I. M., Young, J. P., Giles, G. G., Leggett, B., Macrae, F. A., Ahnen, D. J., Casey, G., Gallinger, S., Haile, R. W., Le Marchand, L., Thibodeau, S. N., Lindor, N. M., Newcomb, P. A., Potter, J. D., Baron, J. A., Hopper, J. L., Jenkins, M. A. and Win, A. K. [2015], 'Aspirin, Ibuprofen, and the risk of colorectal cancer in Lynch syndrome.', *Journal of the National Cancer Institute* **107**.

Amico, M. and Van Keilegom, I. [2018], 'Cure models in survival analysis', *Annual Review of Statistics and Its Application* **5**(1), 311–342.

Amos, C. I., Pinney, S. M., Li, Y., Kupert, E., Lee, J., de Andrade, M. A., Yang, P., Schwartz, A. G., Fain, P. R., Gazdar, A., Minna, J., Wiest, J. S., Zeng, D., Rothschild, H., Mandal, D., You, M., Coons, T., Gaba, C., Bailey-Wilson, J. E. and Anderson, M. W. [2010], 'A susceptibility locus on chromosome 6q greatly increases lung cancer risk among light and never smokers.', *Cancer research* **70**, 2359–67.

Anderson-Bergman, C. [2017], 'An efficient implementation of the EMICM algorithm for the interval censored NPMLE', *Journal of Computational and Graphical Statistics* **26**(2), 463–467.

Andolina, C., Ramjith, J., Rek, J., Lanke, K., Okoth, J., Grignard, L., Arinaitwe, E., Briggs, J., Bailey, J., Aydemir, O., Kamya, M. R., Greenhouse, B., Dorsey, G., Staedke, S. G., Drakeley, C., Jonker, M. and Bousema, T. [2023], 'Plasmodium falciparum gametocyte carriage in longitudinally monitored incident infections is associated with duration of infection and human host factors', *Scientific Reports* **13**(1).

Andrieu, N., Goldgar, D. E., Easton, D. F., Rookus, M., Brohet, R., Antoniou, A. C., Peock, S., Evans, G., Eccles, D., Douglas, F., Noguès, C., Gauthier-Villars, M., Chompret, A., Van Leeuwen, F. E., Kluijt, I., Benitez, J., Arver, B., Olah, E. and Chang-Claude, J. [2006], 'Pregnancies, breast-feeding, and breast cancer risk in the international BRCA1/2 carrier cohort study (IBCCS).', *Journal of the National Cancer Institute* **98**, 535–44.

Antoniou, A. C., Goldgar, D. E., Andrieu, N., Chang-Claude, J., Brohet, R., Rookus, M. A. and Easton, D. F. [2005], 'A weighted cohort approach for analysing factors modifying disease risks in carriers of high-risk susceptibility genes.', *Genetic epidemiology* **29**, 1–11.

Antoniou, A. C., Shenton, A., Maher, E. R., Watson, E., Woodward, E., Lalloo, F., Easton, D. F. and Evans, D. G. [2006], 'Parity and breast cancer risk among BRCA1 and BRCA2 mutation carriers.', *Breast cancer research : BCR* **8**, R72.

Anzai, A. and Nishiura, H. [2022], 'Doubling time of infectious diseases', *Journal of Theoretical Biology* **554**, 111278.

Aphalo, Pedro, J. [2023], 'R package ggpmisc: Miscellaneous extensions to 'ggplot", p. 2.
URL: *https://github.com/aphalo/ggpmisc*

Arntzen, V. H., Fiocco, M. and Geskus, R. B. [2024], 'Two biases in incubation time estimation related to exposure', *BMC Infectious Diseases* **24**(1).

Arntzen, V. H., Fiocco, M., Leitzinger, N. and Geskus, R. B. [2023], 'Towards robust and accurate

estimates of the incubation time distribution, with focus on upper tail probabilities and SARS-CoV-2 infection', *Statistics in Medicine* .

Ashcroft, P., Lehtinen, S., Angst, D. C., Low, N. and Bonhoeffer, S. [2021], 'Quantifying the impact of quarantine duration on COVID-19 transmission', *eLife* **10**.

Backer, J. A., Klinkenberg, D. and Wallinga, J. [2020], 'Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020', *Eurosurveillance* **25**(5).

Barnes, D., Barrowdale, D., Beesley, J., Chen, X., kConFab Investigators, Group, A. O. C. S., James, P., Hopper, J., Goldgar, D., Chenevix-Trench, G., Antoniou, A. and Mitchell, G. [2013], 'Estimating single nucleotide polymorphism associations using pedigree data: applications to breast cancer', *British Journal of Cancer* **108**, 2610–2622.

Barnes, D. R., Lee, A., Easton, D. F. and Antoniou, A. C. [2012], 'Evaluation of association methods for analysing modifiers of disease risk in carriers of high-risk mutations', *Genetic epidemiology* **36**(3), 274–291.

Bikbov, B. and Bikbov, A. [2020], 'Maximum incubation period for COVID-19 infection: do we need to rethink the 14-day quarantine policy?'.

Borde, J., Ernst, C., Wappenschmidt, B., Niederacher, D., Weber-Lassalle, K., Schmidt, G., Hauke, J., Quante, A. S., Weber-Lassalle, N., Horváth, J., Pohl-Rescigno, E., Arnold, N., Rump, A., Gehrig, A., Hentschel, J., Faust, U., Dutrannoy, V., Meindl, A., Kuzyakova, M., Wang-Gohrke, S., Weber, B. H. F., Sutter, C., Volk, A. E., Giannakopoulou, O., Lee, A., Engel, C., Schmidt, M. K., Antoniou, A. C., Schmutzler, R. K., Kuchenbaecker, K. and Hahnen, E. [2020], 'Performance of breast cancer polygenic risk scores in 760 FemaleCHEK2Germline mutation carriers', *JNCI: Journal of the National Cancer Institute* **113**(7), 893–899.

Carayol, J. and Bonaïti-Pellié, C. [2004], 'Estimating penetrance from family data using a retrospective likelihood when ascertainment depends on genotype and age of onset.', *Genetic epidemiology* **27**, 109–117.

Carollo, A., Eilers, P. H. C., Putter, H. and Gampe, J. [2023], 'Smooth hazards with multiple time scales', *arXiv* .
URL: *https://arxiv.org/abs/2305.09342*

Charniga, K., Masters, N. B., Slayton, R. B., Gosdin, L., Minhaj, F. S., Philpott, D., Smith, D., Gearhart, S., Alvarado-Ramy, F., Brown, C., Waltenburg, M. A., Hughes, C. M. and Nakazawa, Y. [2022], 'Estimating the incubation period of monkeypox virus during the 2022 multi-national outbreak', *medRxiv* .

Charras-Garrido, M. and Lezaud, P. [2013], 'Extreme value analysis: an introduction', *Journal de laSociete Française de Statistique, Societe Française de Statistique et Societe Mathematique de France* **154**(2), 66–97.

Chatterjee, N., Kalaylioglu, Z., Shih, J. H. and Gail, M. H. [2006], 'Case–control and case-only designs with genotype and family history data: estimating relative risk, residual familial aggregation, and cumulative risk', *Biometrics* **62**, 36–48.

Chau, N. V. V., Lam, V. T., Dung, N. T., Yen, L. M., Minh, N. N. Q., Hung, L. M., Ngoc, N. M., Dung, N. T., Man, D. N. H., Nguyet, L. A., Nhat, L. T. H., Nhu, L. N. T., Ny, N. T. H., Hong, N. T. T., Kestelyn, E., Dung, N. T. P., Xuan, T. C., Hien, T. T., Phong, N. T., Tu, T. N. H., Geskus, R. B., Thanh, T. T., Truong, N. T., Binh, N. T., Thuong, T. C., Thwaites, G., Tan, L. V., Chau, N. V. V., Dung, N. T., Hung, L. M., Loan, H. T., Truong, N. T., Phong, N. T., Man, D. N. H., Hao, N. V.,

Thuy, D. B., Ngoc, N. M., Lan, N. P. H., Thoa, P. T. N., Thao, T. N. P., Phuong, T. T. L., Uyen, L. T. T., Tam, T. T. T., That, B. T. T., Nhung, H. K., Tai, N. T., Tu, T. N. H., Vuong, V. T., Ty, D. T. B., Dung, L. T., Uyen, T. L., Tien, N. T. M., Thao, H. T. T., Thao, N. N., Vuong, H. N. T., Thao, P. N. P., Phuong, P. M., Tam, D. T. H., Kestelyn, E., Joseph, D., Geskus, R., Thwaites, G., van Doorn, H. R., Hien, H. V., Huy, H. L. A., Ha, H. N., Yen, H. X., Nuil, J. V., Day, J., Donovan, J., Lawson, K., Nguyet, L. A., Yen, L. M., Nhu, L. N. T., Nhat, L. T. H., Tan, L. V., Odette, S. L., Thwaites, L., Rabaa, M., Choisy, M., Chambers, M., Rahman, M., Hoa, N. T., Nhien, N. T. T., Ny, N. T. H., Tuyen, N. T. K., Dung, N. T. P., Hong, N. T. T., Truong, N. X., Khanh, P. N. Q., Yen, P. L. K., Yacoub, S., Kesteman, T., Thuong, N. T. T., Thanh, T. T., Hien, T. T., Hang, V. T. T., Dung, N. T. and and, L. H. N. [2020], 'The natural history and transmission potential of asymptomatic Severe Acute Respiratory Syndrome Coronavirus 2 infection', *Clinical Infectious Diseases* **71**(10), 2679–2687.

Chau, R., Dashti, S. G., Ait Ouakrim, D., Buchanan, D. D., Clendenning, M., Rosty, C., Winship, I. M., Young, J. P., Giles, G. G., Macrae, F. A., Boussioutas, A., Parry, S., Figueiredo, J. C., Levine, A. J., Ahnen, D. J., Casey, G., Haile, R. W., Gallinger, S., Le Marchand, L., Thibodeau, S. N., Lindor, N. M., Newcomb, P. A., Potter, J. D., Baron, J. A., Hopper, J. L., Jenkins, M. A. and Win, A. K. [2016], 'Multivitamin, calcium and folic acid supplements and the risk of colorectal cancer in Lynch syndrome.', *International journal of epidemiology* **45**, 940–53.

Chen, C., Tsay, Y., Wu, Y. and Horng, C. [2013], 'Logistic-aft location-scale mixture regression models with nonsusceptibility for left-truncated and general interval-censored data', *Statistics in Medicine* **32**(24), 4285–4305.

Chen, D., Lau, Y.-C., Xu, X.-K., Wang, L., Du, Z., Tsang, T. K., Wu, P., Lau, E. H. Y., Wallinga, J., Cowling, B. J. and Ali, S. T. [2022], 'Inferring time-varying generation time, serial interval, and incubation period distributions for COVID-19', *Nature Communications* **13**(1).

City Press Center Ho Chi Minh [2021], 'More than 3.4 million vaccinations have been administered to people in Ho Chi Minh City'.
URL: *https://covid19.hochiminhcity.gov.vn/-/hon-3-4-trieu-mui-vac-xin-a-uoc-tiem-cho-nguoi-dan-tphcm?inheritRedirect=true (accessed on 22/02/2024)*

Cobelens, F. G. and Harris, V. C. [2020], 'Untangling severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic control—lessons from Vietnam', *Clinical Infectious Diseases* **72**(9), e343–e344.

Cowling, B. J., Muller, M. P., Wong, I. O. L., Ho, L.-M., Louie, M., McGeer, A. and Leung, G. M. [2007], 'Alternative methods of estimating an incubation distribution: examples from severe acute respiratory syndrome.', *Epidemiology (Cambridge, Mass.)* **18**, 253–259.

Dashti, S. G., Buchanan, D. D., Jayasekara, H., Ait Ouakrim, D., Clendenning, M., Rosty, C., Winship, I. M., Macrae, F. A., Giles, G. G., Parry, S., Casey, G., Haile, R. W., Gallinger, S., Le Marchand, L., Thibodeau, S. N., Lindor, N. M., Newcomb, P. A., Potter, J. D., Baron, J. A., Hopper, J. L., Jenkins, M. A. and Win, A. K. [2017], 'Alcohol consumption and the risk of colorectal cancer for mismatch repair gene mutation carriers.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **26**, 366–375.

Dashti, S. G., Chau, R., Ouakrim, D. A., Buchanan, D. D., Clendenning, M., Young, J. P., Winship, I. M., Arnold, J., Ahnen, D. J., Haile, R. W., Casey, G., Gallinger, S., Thibodeau, S. N., Lindor, N. M., Le Marchand, L., Newcomb, P. A., Potter, J. D., Baron, J. A., Hopper, J. L., Jenkins, M. A. and Win, A. K. [2015], 'Female hormonal factors and the risk of endometrial cancer in Lynch syndrome.', *JAMA* **314**, 61–71.

Dashti, S. G., Win, A. K., Hardikar, S. S., Glombicki, S. E., Mallenahalli, S., Thirumurthi, S., Peterson, S. K., You, Y. N., Buchanan, D. D., Figueiredo, J. C., Campbell, P. T., Gallinger, S., Newcomb, P. A., Potter, J. D., Lindor, N. M., Le Marchand, L., Haile, R. W., Hopper, J. L., Jenkins, M. A., Basen-Engquist, K. M., Lynch, P. M. and Pande, M. [2018], 'Physical activity and the risk of colorectal cancer in Lynch syndrome.', *International journal of cancer* **143**, 2250–2260.

Dayal, V. [2021], 'COVID-19: visualizing peaks and waves in recorded cases across the globe'.

Dejardin, D. and Lesaffre, E. [2013], 'Stochastic EM algorithm for doubly interval-censored data', *Biostatistics* **14**(4), 766–778.

Demers, J., Fagan, W. F., Potluri, S. and Calabrese, J. M. [2023], 'Testing-isolation interventions will likely be insufficient to contain future novel disease outbreaks'.

Deng, Y., You, C., Liu, Y., Qin, J. and Zhou, X.-H. [2020], 'Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for COVID-19 outbreak in China', *Biometrics* .

Dhouib, W., Maatoug, J., Ayouni, I., Zammit, N., Ghammem, R., Fredj, S. B. and Ghannem, H. [2021], 'The incubation period during the pandemic of COVID-19: a systematic review and meta-analysis', *Systematic Reviews* **10**(1).

Dorigatti, I., Okell, L., Cori, A., Imai, N., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba Perez, Z., Cuomo-Dannenburg, G., Fitzjohn, R., Fu, H., Gaythorpe, K., Hamlet, A., Hinsley, W., Hong, N., Kwun, M., Laydon, D., Nedjati Gilani, G., Riley, S., Van Elsland, S., Volz, E., Wang, H., Walters, C., Xi, X., Donnelly, C., Ghani, A. and Ferguson, N. [2020], 'Report 4: Severity of 2019-novel coronavirus (nCoV)'.

ECDC [2018], Guidelines for writing outbreak investigation reports, Technical report, ECDC. **URL:** *https://www.ecdc.europa.eu/sites/default/files/documents/Annex 05_Guide for writing outbreak investigation reports_2019.pdf (accessed on 03/2024)*

Eikmeier, D., Medus, C. and Smith, K. [2018], 'Incubation period for outbreak-associated, non-typhoidal salmonellosis cases, Minnesota, 2000–2015', *Epidemiology and Infection* **146**(4), 423–429.

Ejima, K., Kim, K. S., Ludema, C., Bento, A. I., Iwanami, S., Fujita, Y., Ohashi, H., Koizumi, Y., Watashi, K., Aihara, K., Nishiura, H. and Iwami, S. [2021], 'Estimation of the incubation period of COVID-19 using viral load data', *Epidemics* **35**, 100454.

Farewell, V. T. and Prentice, R. L. [1977], 'A study of distributional shape in life testing', *Technometrics* **19**(1), 69–75.

Fay, M. P. and Shaw, P. A. [2010], 'Exact and asymptotic weighted logrank tests for interval censored data: Theintervalrpackage', *Journal of Statistical Software* **36**(2).

Galmiche, S., Cortier, T., Charmet, T., Schaeffer, L., Chény, O., von Platen, C., Lévy, A., Martin, S., Omar, F., David, C., Mailles, A., Carrat, F., Cauchemez, S. and Fontanet, A. [2023], 'SARS-CoV-2 incubation period across variants of concern, individual factors, and circumstances of infection in France: a case series analysis from the ComCor study', *The Lancet Microbe* **4**(6), e409–e417.

Geskus, R. B. [2001], 'Methods for estimating the AIDS incubation time distribution when date of seroconversion is censored', *Statistics in Medicine* **20**(5), 795–812.

Gibbs, H., Liu, Y., Pearson, C. A. B., Jarvis, C. I., Grundy, C., Quilty, B. J., Diamond, C., Simons, D., Gimma, A., Leclerc, Q. J., Auzenbergs, M., Lowe, R., O'Reilly, K., Quaife, M., Hellewell, J.,

Knight, G. M., Jombart, T., Klepac, P., Procter, S. R., Deol, A. K., Rees, E. M., Flasche, S., Kucharski, A. J., Abbott, S., Sun, F. Y., Endo, A., Medley, G., Munday, J. D., Meakin, S. R., Bosse, N. I., Edmunds, W. J., Davies, N. G., Prem, K., Hué, S., Villabona-Arenas, C. J., Nightingale, E. S., Houben, R. M. G. J., Foss, A. M., Tully, D. C., Emery, J. C., van Zandvoort, K., Atkins, K. E., Rosello, A., Funk, S., Jit, M., Clifford, S., Russell, T. W. and and, R. M. E. [2020], 'Changing travel patterns in China during the early stages of the COVID-19 pandemic', *Nature Communications* **11**(1).

Gollwitzer, M., Platzer, C., Zwarg, C. and Göritz, A. S. [2020], 'Public acceptance of Covid-19 lockdown scenarios', *International Journal of Psychology* **56**(4), 551–565.

Groendyke, C., Welch, D. and Hunter, D. R. [2010], 'Bayesian inference for contact networks given epidemic data', *Scandinavian Journal of Statistics* **38**(3), 600–616.

Guzzetta, G., Mammone, A., Ferraro, F., Caraglia, A., Rapiti, A., Marziano, V., Poletti, P., Cereda, D., Vairo, F., Mattei, G., Maraglino, F., Rezza, G. and Merler, S. [2022], 'Early estimates of monkeypox incubation period, generation time, and reproduction number, Italy, May–June 2022', *Emerging Infectious Diseases* **28**(10), 2078–2081.

Haigh, A., Apthorp, D. and Bizo, L. A. [2020], 'The role of Weber's law in human time perception', *Attention, Perception and Psychophysics* **83**(1), 435–447.

Hardy, A., Shum, M. and Quyên, V. N. [2020], 'The 'F-system' of targeted isolation: A key method in Vietnam's suppression of Covid-19.'.
URL: *https://halshs.archives-ouvertes.fr/halshs-03151062/document*

Held, L., Hens, N., O'Neill, P. and Wallinga, J. [2019], *Handbook of Infectious Disease Data Analysis*, Chapman Hall/CRC.

Henderson, R. and Oman, P. [1999], 'Effect of frailty on marginal regression estimates in survival analysis.', *Journal of the Royal Statistics Society. Series B* **61**(2), 367–379.

Heuch, I., Abdalla, S. and Tayeb, S. E. [2018], 'Modelling memory decay after injuries using household survey data from Khartoum State, Sudan', *BMC Medical Research Methodology* **18**(1).

Jackson, C. [2016], 'flexsurv: A platform for parametric survival modeling in R', *Journal of Statistical Software* **70**(8).

Jiang, X. L., Qiu, Y., Zhang, Y. P., Yang, P., Huang, B., Lin, M., Ye, Y., Gao, F., Li, D., Qin, Y., Li, Y. and Li, Z. J. [2023], 'Latent period and incubation period with associated factors of COVID-19 caused by Omicron variant.', *Zhonghua yu fang yi xue za zhi [Chinese journal of preventive medicine]* **57**, 659–666.

Junsawang, P., Promwongsa, M. and Srisodaphol, W. [2021], 'Robust outliers detection method for skewed distribution', *Thailand Statistician* **19**(3).

Kamiza, A. B., Hsieh, L.-L., Tang, R., Chien, H.-T., Lai, C.-H., Chiu, L.-L., Lo, T.-P., Hung, K.-Y., You, J.-F., Wang, W.-C., Hsiung, C. A. and Yeh, C.-C. [2016], 'TP53 polymorphisms and colorectal cancer risk in patients with Lynch syndrome in Taiwan: A retrospective cohort study.', *PloS one* **11**, e0167354.

Kamvar, Z. N., Cai, J., Pulliam, J. R. C., Schumacher, J. and Jombart, T. [2019], 'Epidemic curves made easy using the R package 'incidence'', *F1000Research* **8**, 139.

Kang, M., Xin, H., Yuan, J., Ali, S. T., Liang, Z., Zhang, J., Hu, T., Lau, E. H., Zhang, Y., Zhang, M., Cowling, B. J., Li, Y. and Wu, P. [2022], 'Transmission dynamics and epidemiological

characteristics of SARS-CoV-2 Delta variant infections in Guangdong, China, May to June 2021', *Eurosurveillance* **27**(10).

Killick, E., Tymrakiewicz, M., Cieza-Borrella, C., Smith, P., Thompson, D. J., Pooley, K. A., Easton, D. F., Bancroft, E., Page, E., Leongamornlert, D., Kote-Jarai, Z. and Eeles, R. A. [2014], 'Telomere length shows no association with BRCA1 and BRCA2 mutation status.', *PloS one* **9**, e86659.

KoBoToolbox [2022], 'KoBoToolbox simple, robust and powerful tools for data collection.'. **URL:** *https://www.kobotoolbox.org*

Komárek, A., Lesaffre, E. and Hilton, J. F. [2005], 'Accelerated failure time model for arbitrarily censored data with smoothed error distribution', *Journal of Computational and Graphical Statistics* **14**(3), 726–745.

Komárek, A. [2020], *smoothSurv package*. **URL:** *https://CRAN.R-project.org/package=smoothSurv*

Krämer, A., Kretzschmar, M. and Krickeberg, K., eds [2010], *Modern Infectious Disease Epidemiology*, Springer New York.

Lai, C., Yu, R., Wang, M., Xian, W., Zhao, X., Tang, Q., Chen, R., Zhou, X., Li, X., Li, Z., Li, Z., Deng, G. and Wang, F. [2020], 'Shorter incubation period is associated with severe disease progression in patients with COVID-19', *Virulence* **11**(1), 1443–1452.

Lakeman, I., Hilbers, F., Rodríguez-Girondo, M., Lee, A., Vreeswijk, M., Hollestelle, A., Seynaeve, C., Meijers-Heijboer, H., Oosterwijk, J. and Hoogerbrugge, N. [2019], 'Addition of a 161-SNP polygenic risk score to family history-based risk prediction: impact on clinical management in non-BRCA1/2 breast cancer families', *Journal of Medical Genetics* pp. jmedgenet–2019.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G. and Lessler, J. [2020], 'The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application', *Annals of Internal Medicine* **172**(9), 577–582.

Law, C. G. and Brookmeyer, R. [1992], 'Effects of mid-point imputation on the analysis of doubly censored data', *Statistics in Medicine* **11**(12), 1569–1578.

Lee, S. M. S. and Pun, M. C. [2006], 'On m out of n bootstrapping for nonstandard M-estimation with nuisance parameters', *Journal of the American Statistical Association* **101**(475), 1185–1197.

Li, M., Yuan, Q., Chen, P., Song, B. and Ma, J. [2021], 'Estimating the quarantine failure rate for COVID-19', *Infectious Disease Modelling* **6**, 924–929.

Li, Y., Jiang, X., Qiu, Y., Gao, F., Xin, H., Li, D., Qin, Y. and Li, Z. [2024], 'Latent and incubation periods of Delta, BA.1, and BA.2 variant cases and associated factors: a cross-sectional study in China', *BMC Infectious Diseases* **24**(1).

Li, Z.-Y., Zhang, Y., Peng, L.-Q., Gao, R.-R., Jing, J.-R., Wang, J.-L., Ren, B.-Z., Xu, J.-G. and Wang, T. [2021], 'Demand for longer quarantine period among common and uncommon COVID-19 infections: a scoping review.', *Infectious diseases of poverty* **10**, 56.

Linton, N., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A., Jung, S.-m., Yuan, B., Kinoshita, R. and Nishiura, H. [2020], 'Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: A statistical analysis of publicly available case data', *Journal of Clinical Medicine* **9**(2), 538.

Ma, Q., Liu, J., Liu, Q., Kang, L., Liu, R., Jing, W., Wu, Y. and Liu, M. [2021], 'Global percentage of

asymptomatic SARS-CoV-2 infections among the tested population and individuals with confirmed COVID-19 diagnosis: A systematic review and meta-analysis', *JAMA Network Open* **4**(12), e2137257.

Ma, T., Ding, S., Huang, R., Wang, H., Wang, J., Liu, J., Wang, J., Li, J., Wu, C., Fan, H. and Zhou, N. [2022], 'The latent period of coronavirus disease 2019 with SARS-CoV-2 B.1.617.2 Delta variant of concern in the postvaccination era', *Immunity, Inflammation and Disease* **10**(7).

Mandel, M., de Uña-Álvarez, J., Simon, D. K. and Betensky, R. A. [2017], 'Inverse probability weighted cox regression for doubly truncated data', *Biometrics* **74**(2), 481–487.

McAloon, C., Collins, Á., Hunt, K., Barber, A., Byrne, A. W., Butler, F., Casey, M., Griffin, J., Lane, E., McEvoy, D., Wall, P., Green, M., O'Grady, L. and More, S. J. [2020], 'Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research', *BMJ Open* **10**(8), e039652.

McKendrick, A. G. [1925], 'Applications of mathematics to medical problems', *Proceedings of the Edinburgh Mathematical Society* **44**, 98–130.

Miner, J. [1922], 'The incubation period of typhoid fever.', *J Infect Dis* **31**, 296–301.

Moshiro, C. [2005], 'Effect of recall on estimation of non-fatal injury rates: a community based study in Tanzania', *Injury Prevention* **11**(1), 48–52.

Netherlands Cancer Registry [2021], 'Netherlands comprehensive cancer organisation (iknl).'.
URL: *www.iknl.nl/en/ncr/ncr-data-figures*

Neugebauer, R. and Ng, S. [1990], 'Differential recall as a source of bias in epidemiologic research', *Journal of Clinical Epidemiology* **43**(12), 1337–1341.

Nishiura, H. [2007], 'Early efforts in modeling the incubation period of infectious diseases with an acute course of illness', *Emerging Themes in Epidemiology* **4**(1).

Nishiura, H., Mizumoto, K., Ejima, K., Zhong, Y., Cowling, B. and Omori, R. [2012], 'Incubation period as part of the case definition of severe respiratory illness caused by a novel coronavirus.', *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **17**.

Olivera, M. J. and Muñoz, L. [2024], 'Exploring the latency period in Chagas disease: duration and determinants in a cohort from Colombia', *Transactions of The Royal Society of Tropical Medicine and Hygiene* .

Our World in Data [2024], 'Coronavirus (COVID-19) vaccinations'.
URL: *https://ourworldindata.org/covid-vaccinations (accessed on 19/03/2024)*

Pak, D., Langohr, K., Ning, J., Martínez, J. C., Melis, G. G. and Shen, Y. [2020], 'Modeling the coronavirus disease 2019 incubation period: Impact on quarantine policy', *Mathematics* **8**(9), 1631.

Pak, D., Liu, J., Ning, J., Gómez, G. and Shen, Y. [2020], 'Analyzing left-truncated and right-censored infectious disease cohort data with interval-censored infection onset', *Statistics in Medicine* **40**(2), 287–298.

Park, S. W., Akhmetzhanov, A. R., Charniga, K., Cori, A., Davies, N. G., Dushoff, J., Funk, S., Gostic, K., Grenfell, B., Linton, N., Lipsitch, M., Lison, A., Overton, C. E., Ward, T. and Abbott, S. [2024], 'Estimating epidemiological delay distributions for infectious diseases', *medRXiv* .
URL: *https://www.medrxiv.org/content/10.1101/2024.01.12.24301247v1*

Petrignani, M., Verhoef, L., Vennema, H., van Hunen, R., Baas, D., van Steenbergen, J. E. and Koopmans, M. P. [2014], 'Underdiagnosis of foodborne hepatitis A, the netherlands, 2008–20101', *Emerging Infectious Diseases* **20**(4), 596–602.

Pijpe, A., Andrieu, N., Easton, D. F., Kesminiene, A., Cardis, E., Noguès, C., Gauthier-Villars, M., Lasset, C., Fricker, J.-P., Peock, S., Frost, D., Evans, D. G., Eeles, R. A., Paterson, J., Manders, P., van Asperen, C. J., Ausems, M. G. E. M., Meijers-Heijboer, H., Thierry-Chef, I., Hauptmann, M., Goldgar, D., Rookus, M. A. and van Leeuwen, F. E. [2012], 'Exposure to diagnostic radiation and risk of breast cancer among carriers of BRCA1/2 mutations: retrospective cohort study (GENE-RAD-RISK).', *BMJ (Clinical research ed.)* **345**, e5660.

Plummer, M. [2017], 'JAGS Version 4.3.0 user manual'.
URL: *https://people.stat.sc.edu/hansont/stat740/jags_user_manual.pdf*

Pooley, K. A., McGuffog, L., Barrowdale, D., Frost, D., Ellis, S. D., Fineberg, E., Platte, R., Izatt, L., Adlard, J., Bardwell, J., Brewer, C., Cole, T., Cook, J., Davidson, R., Donaldson, A., Dorkins, H., Douglas, F., Eason, J., Houghton, C., Kennedy, M. J., McCann, E., Miedzybrodzka, Z., Murray, A., Porteous, M. E., Rogers, M. T., Side, L. E., Tischkowitz, M., Walker, L., Hodgson, S., Eccles, D. M., Morrison, P. J., Evans, D. G., Eeles, R. A., Antoniou, A. C., Easton, D. F. and Dunning, A. M. [2014], 'Lymphocyte telomere length is long in BRCA1 and BRCA2 mutation carriers regardless of cancer-affected status.', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **23**, 1018–24.

Puhach, O., Meyer, B. and Eckerle, I. [2023], 'SARS-CoV-2 viral load and shedding kinetics.', *Nature reviews. Microbiology* **21**, 147–161.

Putter, H., Goeman, J. and Wallinga, J. [2024], 'Stochastic epidemic models and their link with methods from survival analysis', *medRxiv* .
URL: *https://www.medrxiv.org/content/early/2024/02/20/2024.02.18.24302991*

Qin, J., You, C., Lin, Q., Hu, T., Yu, S. and Zhou, X.-H. [2020], 'Estimation of incubation period distribution of COVID-19 using disease onset forward time: A novel cross-sectional and forward follow-up study', *Science Advances* **6**(33), eabc1202.

R Core Team [2021], *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

Ramjith, J., Andolina, C., Bousema, T. and Jonker, M. A. [2022], 'Flexible time-to-event models for double-interval-censored infectious disease data with clearance of the infection as a competing risk', *Frontiers in Applied Mathematics and Statistics* **8**.

Raphael, K. [1987], 'Recall bias: A proposal for assessment and control', *International Journal of Epidemiology* **16**(2), 167–170.

Reich, N. G., Lessler, J., Cummings, D. A. T. and Brookmeyer, R. [2009], 'Estimating incubation period distributions with coarse data', *Statistics in Medicine* **28**(22), 2769–2784.

RStudio Team [2021], *RStudio: Integrated Development Environment for R*, RStudio, PBC., Boston, MA.

Rubio, J. F. [2020], 'The Generalised Gamma Distribution'.
URL: *https://rpubs.com/FJRubio/GG (last accessed on 08/04/2024).*

Ruegger, J., Stoeck, K., Amsler, L., Blaettler, T., Zwahlen, M., Aguzzi, A., Glatzel, M., Hess, K. and Eckert, T. [2009], 'A case-control study of sporadic Creutzfeldt-Jakob disease in Switzerland:

analysis of potential risk factors with regard to an increased CJD incidence in the years 2001–2004', *BMC Public Health* **9**(1).

Sah, P., Fitzpatrick, M. C., Zimmer, C. F., Abdollahi, E., Juden-Kelly, L., Moghadas, S. M., Singer, B. H. and Galvani, A. P. [2021], 'Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis', *Proceedings of the National Academy of Sciences* **118**(34), e2109229118.

Salehabadi, S. M., Sengupta, D. and Das, R. [2014], 'Parametric estimation of menarcheal age distribution based on recall data', *Scandinavian Journal of Statistics* **42**(1), 290–305.

Sartwell, P. E. [1950], 'The distribution of incubation periods of infectious disease', *American Journal of Epidemiology* **51**(3), 310–318.

Sen, A. [2023], 'R package 'koboconnectr': Download data from kobotoolbox to r', Webpage: https://github.com/asitav-sen/koboconnectR (accessed in 2023).

Stacy, E. W. and Mihram, G. A. [1965], 'Parameter estimation for a generalized gamma distribution', *Technometrics* **7**(3), 349–358.

Stockdale, J. E., Susvitasari, K., Tupper, P., Sobkowiak, B., Mulberry, N., Gonçalves da Silva, A., Watt, A. E., Sherry, N. L., Minko, C., Howden, B. P., Lane, C. R. and Colijn, C. [2023], 'Genomic epidemiology offers high resolution estimates of serial intervals for COVID-19.', *Nature communications* **14**, 4830.

Sudman, S. and Bradburn, N. M. [1973], 'Effects of time and memory factors on response in surveys', *Journal of the American Statistical Association* **68**(344), 805–815.

Sukumaran, A. and Dewan, I. [2018], 'Modelling and analysis of recall-based competing risks data', *Journal of Applied Statistics* **46**(9), 1621–1635.

Tam, N. T., Anh, N. T., Tung, T. S., Thach, P. N., Dung, N. T., Trang, V. D., Hung, L. M., Dien, T. C., Ngoc, N. M., Van Duyet, L., Cuong, P. M., Phuong, H. V. M., Thai, P. Q., Tung, N. L. N., Man, D. N. H., Phong, N. T., Quang, V. M., Thoa, P. T. N., Truong, N. T., Thao, T. N. P., Linh, D. P., Tai, N. T., Bao, H. T., Vuong, V. T., Nhung, H. T. K., Hong, P. N. D., Hanh, L. T. P., Chung, L. T., Nhan, N. T. T., Thanh, T. T., Hung, D. T., Mai, H. K., Long, T. H., Trang, N. T., Thuong, N. T. H., Hong, N. T. T., Nhu, L. N. T., Ny, N. T. H., Thuy, C. T., Thanh, L. K., Nguyet, L. A., Mai, L. T. Q., Thuong, T. C., Nga, L. H., Thanh, T. T., Thwaites, G., Rogier van Doorn, H., Chau, N. V. V., Kesteman, T. and Van Tan, L. [2023], 'Spatiotemporal evolution of SARS-CoV-2 Alpha and Delta variants during large nationwide outbreak of COVID-19, Vietnam, 2021', *Emerging Infectious Diseases* **29**(5).

Tan, L. V. [2021], 'COVID-19 control in Vietnam', *Nature Immunology* **22**(3), 261–261.

Ten Broeke, S. W., Brohet, R. M., Tops, C. M., van der Klift, H. M. and Velthuizen, M. E. e. a. [2015], 'Lynch syndrome caused by germline PMS2 mutations: delineating the cancer risk.', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **33**, 319–325.

Ten Broeke, S. W., Elsayed, F. A., Pagan, L., Olderode-Berends, M. J. W., Garcia, E. G., Gille, H. J. P., van Hest, L. P., Letteboer, T. G. W., van der Kolk, L. E., Mensenkamp, A. R., van Os, T. A., Spruijt, L., Redeker, B. J. W., Suerink, M., Vos, Y. J., Wagner, A., Wijnen, J. T., Steyerberg, E. W., Tops, C. M. J., van Wezel, T. and Nielsen, M. [2018], 'SNP association study in PMS2-associated Lynch syndrome.', *Familial cancer* **17**, 507–515.

Ten Broeke, S. W., van Bavel, T. C., Jansen, A. M., Gómez-García, E., Hes, F. J., van Hest, L. P., Letteboer, T. G., Olderode-Berends, M. J., Ruano, D., Spruijt, L., Suerink, M., Tops, C. M., van Eijk, R., Morreau, H., van Wezel, T. and Nielsen, M. [2018], 'Molecular background of colorectal tumors from patients with Lynch syndrome associated with germline variants in pms2',

Gastroenterology **155**(3), 844–851.

Thai, P. Q., Rabaa, M. A., Luong, D. H., Tan, D. Q., Quang, T. D., Quach, H.-L., Thi, N.-A. H., Dinh, P. C., Nghia, N. D., Tu, T. A., Quang, L. N., Phuc, T. M., Chau, V., Khanh, N. C., Anh, D. D., Duong, T. N., Thwaites, G., van Doorn, H. R., Choisy, M., Chambers, M., Choisy, M., Day, J., Trinh, D. H. K., Tam, D. T. H., Donovan, J., Duc, D. H., Geskus, R. B., Chanh, H. Q., Van, H. H., Thao, H. D., le Anh Huy, H., Ha, H. N., Trieu, H. T., Yen, H. X., Kestelyn, E., Kesteman, T., Nguyet, L. A., Yen, L. M., Lawson, K., Thanh, L. K., Nhu, L. N. T., Nhat, L. T. H., Lan, L. T. H., Van, T. L., Lewycka, S. O., Tran, N. B., Nguyet, N. M., Quyen, N. T. H., Ngoc, N. T., Ny, N. T. H., Thuong, N. T. H., Trang, N. T. H., Tuyen, N. T. K., Diep, N. T. N., Dung, N. T. P., Tam, N. T., Hong, N. T. T., Trang, N. T., Van, V. C. N., Truong, N. X., Van, N. T. T., Khanh, P. N. Q., Lam, P. K., Yen, P. L. K., Nhat, P. T. H., Rabaa, M., Thuong, T. N. T., Thwaites, G., Thwaites, L., Phuc, T. M., Thanh, T. T., Ngoc, T. T. B., Hien, T. T., van, D. H. R., van, N. J., Chau, V., Bich, V. T. N., Hang, V. T. T. and and, S. Y. [2020], 'The first 100 days of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) control in Vietnam', *Clinical Infectious Diseases* .

Tindale, L. C., Stockdale, J. E., Coombe, M., Garlock, E. S., Lau, W. Y. V., Saraswat, M., Zhang, L., Chen, D., Wallinga, J. and Colijn, C. [2020], 'Evidence for transmission of COVID-19 prior to symptom onset', *eLife* **9**.

United Nations, Department of Economic and Social Affairs, Population Division [2022], 'World population prospects: The 2022 revision'.
URL: *https://population.un.org/wpp/ (accessed in 03/2024).*

Vietnam Center for Disease Control [2021], 'Dispatch no. 600/CD-BCD dated May 5, 2021 on adjusting the centralized quarantine time, management time after the end of centralized quarantine and testing to prevent COVID-19 epidemic.'.
URL: *https://vncdc.gov.vn/cong-dien-so-600cd-bcd-ngay-552021-ve-viec-dieu-chinh-thoi-gian-cach-ly-tap-trung-thoi-gian-quan-ly-sau-khi-ket-thuc-cach-ly-tap-trung-va-xet-nghiem-phong-chong-dich-covid-9-nd16053.html (accessed on 20/03/2024)*

Villaseñor-Alva, J. A. and González-Estrada, E. [2009], 'A bootstrap goodness of fit test for the generalized pareto distribution', *Computational Statistics amp; Data Analysis* **53**(11), 3835–3841.

WHO [2003], *Consensus document on the epidemiology of severe acute respiratory syndrome (SARS), May 2003.*, World Health Organization.
URL: *http://www.who.int/csr/sars/WHOconsensus.pdf (accessed on 14/12/2021)*

WHO [2020], Considerations for quarantine of individuals in the context of containment for coronavirus disease (COVID-19) WHO/2019-nCoV/IHR Quarantine/2020.2, Technical report.

Win, A. K., Clendenning, M., Crawford, W., Rosty, C., Preston, S. G., Southey, M. C., Parry, S., Giles, G. G., Macrae, F. A., Winship, I. M., Baron, J. A., Hopper, J. L., Jenkins, M. A. and Buchanan, D. D. [2015], 'Genetic variants within the hTERT gene and the risk of colorectal cancer in Lynch syndrome.', *Genes & cancer* **6**, 445–51.

Win, A. K., Dowty, J. G., Antill, Y. C., English, D. R., Baron, J. A., Young, J. P., Giles, G. G., Southey, M. C., Winship, I., Lipton, L., Parry, S., Thibodeau, S. N., Haile, R. W., Gallinger, S., Le Marchand, L., Lindor, N. M., Newcomb, P. A., Hopper, J. L. and Jenkins, M. A. [2011], 'Body mass index in early adulthood and endometrial cancer risk for mismatch repair gene mutation carriers.', *Obstetrics and gynecology* **117**, 899–905.

Win, A. K., Dowty, J. G., English, D. R., Campbell, P. T., Young, J. P., Winship, I., Macrae, F. A., Lipton, L., Parry, S., Young, G. P., Buchanan, D. D., Martínez, M. E., Jacobs, E. T., Ahnen, D. J.,

Haile, R. W., Casey, G., Baron, J. A., Lindor, N. M., Thibodeau, S. N., Newcomb, P. A., Potter, J. D., Le Marchand, L., Gallinger, S., Hopper, J. L. and Jenkins, M. A. [2011], 'Body mass index in early adulthood and colorectal cancer risk for carriers and non-carriers of germline mutations in dna mismatch repair genes.', *British journal of cancer* **105**, 162–9.

Win, A. K., Hopper, J. L., Buchanan, D. D., Young, J. P., Tenesa, A., Dowty, J. G., Giles, G. G., Goldblatt, J., Winship, I., Boussioutas, A., Young, G. P., Parry, S., Baron, J. A., Duggan, D., Gallinger, S., Newcomb, P. A., Haile, R. W., Le Marchand, L., Lindor, N. M. and Jenkins, M. A. [2013], 'Are the common genetic variants associated with colorectal cancer risk for DNA mismatch repair gene mutation carriers?', *European journal of cancer (Oxford, England : 1990)* **49**, 1578–87.

Win, A. K., Reece, J. C., Buchanan, D. D., Clendenning, M., Young, J. P., Cleary, S. P., Kim, H., Cotterchio, M., Dowty, J. G., MacInnis, R. J., Tucker, K. M., Winship, I. M., Macrae, F. A., Burnett, T., Le Marchand, L., Casey, G., Haile, R. W., Newcomb, P. A., Thibodeau, S. N., Lindor, N. M., Hopper, J. L., Gallinger, S. and Jenkins, M. A. [2015], 'Risk of colorectal cancer for people with a mutation in both a MUTYH and a DNA mismatch repair gene.', *Familial cancer* **14**, 575–83.

Wu, Y., Kang, L., Guo, Z., Liu, J., Liu, M. and Liang, W. [2022], 'Incubation period of COVID-19 caused by unique SARS-CoV-2 strains', *JAMA Network Open* **5**(8), e2228008.

Xin, H., Li, Y., Wu, P., Li, Z., Lau, E. H. Y., Qin, Y., Wang, L., Cowling, B. J., Tsang, T. K. and Li, Z. [2021], 'Estimating the latent period of coronavirus disease 2019 (COVID-19)', *Clinical Infectious Diseases* .

Xin, H., Wong, J. Y., Murphy, C., Yeung, A., Ali, S. T., Wu, P. and Cowling, B. J. [2021], 'The incubation period distribution of coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis', *Clinical Infectious Diseases* .

Yang, L., Dai, J., Zhao, J., Wang, Y., Deng, P. and Wang, J. [2020], 'Estimation of incubation period and serial interval of COVID-19: analysis of 178 cases and 131 transmission chains in Hubei province, China', *Epidemiology and Infection* **148**.

Yoo, J., Kim, S., Park, W.-C., Kim, B.-S., Choi, H. and Won, C. W. [2017], 'Discrepancy between quarterly recall and annual recall of falls: A survey of older adults', *Annals of Geriatric Medicine and Research* **21**(4), 174–181.

Zhang, Z.-J., Che, T.-L., Wang, T., Zhao, H., Hong, J., Su, Q., Zhang, H.-Y., Zhou, S.-X., Teng, A.-Y., Zhang, Y.-Y., Yang, Y., Fang, L.-Q. and Liu, W. [2021], 'Epidemiological features of COVID-19 patients with prolonged incubation period and its implications for controlling the epidemics in China', *BMC Public Health* **21**(1).

Stellingen

behorende bij het proefschrift
Incubation and latency time estimation for SARS-CoV-2
van
Vera Arntzen

1. *One hopes that investment in resilient public health infrastructures and effective responses to health crises can be legacy and counterpart to the painful collective global memory of loss that COVID-19 has imposed.*

- Frank G. Cobelens and Vanessa C. Harris, (Clinical Infectious Diseases, 2020)

The unruly reality of the infectious disease context requires ongoing development of tailored approaches. (Chapter 4)

2. A semi-parametric approach provides a better fit to the tail of the incubation time distribution than traditional, parametric approaches. (Chapter 2)

3. There would be merit in exploration of an evidence-based personalized quarantine length. (Chapter 2)

4. Our memory is imperfect. Hence, knowledge on *how* we memorise potential risk exposures is as important as the reported exposure information itself. (Chapter 3)

5. Inverse probability weighting turns the statistician into a magician with a trick that keeps its magic even once you understand it. (Chapter 5)

6. The optimal model to inform quarantine length comprehends transmission characteristics as well as social, ethical and logistic factors. (Chapter 6)

7. Negative gossip in the workplace has all the traits of an infectious disease.

8. *Single vision produces worse illusions than double vision or many-headed monsters*.

- Donna Haraway (A Cyborg Manifesto, 1987)

The quality of science benefits from diversity and interdisciplinary collaboration.

9. *There are two kinds of people in the world: those who divide the world into two kinds of people, and those who don't.*

- Robert Benchley (orig.), Lammert Kamphuis (Verslaafd aan ons eigen gelijk / Pleidooi voor perspectivistische lenigheid, 2023)

Er zijn twee soorten onderzoekers in academia: zij die collega's verdelen in alpha/beta, frequentistisch/Bayesiaans, theoretisch/toegepast, wiskundig/onkundig, bondgenoten/tegenstanders, .../... en zij die dat niet doen.

10. *In die periode was ik net een paar jaar afgestudeerd (...) en probeerde wijs te worden uit het labyrintisch complex van menselijke drijfveren.*

- Roxane van Iperen (De Genocidefax, 2022)

Leren werken is mensenwerk.

Vera Arntzen,
Leiden, 16 oktober 2024