Roles of heterogeneity in infectious disease epidemiology: implications on dynamics, inference and control of influenza and COVID-19

Akira Endo

Thesis submitted in accordance with the requirements for the degree of
Doctor of Philosophy
of the
University of London
December 2020

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

I, Akira Endo, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Akira Endo, December 2020

Abstract

Heterogeneity plays a vital role in the epidemiology of infectious diseases. Individual-level variation in susceptibility or infectiousness due to predictable biological, behavioural or random factors can modify the transmission dynamics and cause deviation from what is expected from models assuming that transmission is homogeneous. One of the important sources of heterogeneity are social contact networks; respiratory infectious diseases including influenza and COVID-19 spread over social contact networks that are formed through mixing in multiple social settings. Households and schools are important places of transmission of many respiratory infectious diseases (although the role of schools in the transmission of COVID-19 remain unclear). Schoolchildren are often the main drivers of influenza epidemics, and they further spread the disease to other age groups by introducing infection into households and other settings. However, detailed transmission dynamics in households and schools have not been fully understood; in particular, it is not well known how group sizes and contact patterns affect transmission risks in heterogeneous populations.

The underlying mechanisms of variations in transmission may not necessarily be explained by known factors. Even in such cases, quantifying such variation can be useful in characterising the transmission dynamics. SARS-CoV-2, along with other related coronaviruses, exhibits strong dispersion in the number of secondary transmissions per case. In addition to the basic reproduction number R_0, which represents the mean number of secondary transmissions in fully susceptible population, the degree of variability around the mean highlights the importance of superspreading and potentially informs control policy targeting superspreading events.

This PhD study attempts to further improve the current understanding of how heterogeneity affects transmission dynamics, inference and public health applications. Using datasets and models of high burden respiratory infectious diseases, influenza and COVID-19, this thesis investigated the roles of heterogeneity in various contexts, i.e. transmission settings such as households and schools and important research topics such as vaccine evaluation studies, international dissemination and contact tracing.

Acknowledgements

Contents

List of Figures

List of Tables

Acronyms

COVID-19	Coronavirus disease 2019
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
TND	Test-negative design
CPI	community probability of infection
SAR	Secondary attack rate
SITP	Susceptible-infectious transmission probability
SARS-CoV	severe acute respiratory syndrome coronavirus
MERS-CoV	Middle East respiratory syndrome coronavirus
WHO	World Health Organization
MCMC	Markov-chain Monte Carlo
WBIC	Widely-Applicable Bayesian Information Criterion
ESS	Effective sample size
PCR	Polymerase chain reaction
CrI	Credible interval
VE	Vaccine effectiveness
MO	Multiple overimputation
TD	Target disease
ND	Non-target disease
MLE	Maximum likelihood estimate
EM	Expectation maximisation

1 Introduction

1.1 Motivation

Influenza is a respiratory infectious disease that causes outbreaks almost every year, imposing substantial burden on people, not only infants and the elderly who suffers more severe disease outcomes but also schoolchildren who exhibit high incidence and their parents who may need to stay home to provide care. As influenza is expected to exhibit representative trans(4–7)mission patterns of respiratory infectious diseases in general (e.g., transmission via droplets and short-range aerosols through physical and conversational contacts; on the other hand, diseases involving long-range droplet-nuclei transmission, i.e. measles, tuberculosis and varicella may not necessarily share similar transmission patterns), the epidemiological characteristics estimated from influenza outbreak data has the potential to also inform other less-documented respiratory diseases. Yet, our understanding of the transmission dynamics of influenza at different scales across different social settings is limited partly because most infections cause mild or no symptoms and may not be diagnosed or tested at medical institutions. In this light, Uchida et al. (1–4)(1–5)(Uchida, Kaneko, Hidaka, Yamamoto, & Honda, 2017; Uchida, Kaneko, Hidaka, Yamamoto, Honda, et al., 2017b) collected an extensively detailed, laboratory-confirmed dataset of influenza from a citywide survey of primary school students in Matsumoto city, Japan in 2014/15 season, which provided an opportunity to perform in-depth analyses that were hardly possible in previous studies, if combined with mathematical modelling approaches.

A notable strength of this dataset is that individual-level details of each student (e.g. class/grade, household composition and adherence to precautionary measures) are available, which allows for modelling of transmission dynamics accounting for heterogeneity. Heterogeneity discussed in this thesis refers to variability in transmission patterns between individuals and between populations. For example, characteristics related to transmission, e.g. susceptibility and infectiousness, may vary between individuals. Community structures such as school classes and household compositions may introduce nonuniform patterns into social interactions resulting in transmission. Epidemiological differences between populations, e.g. demography, geography, social systems, culture and public health practices can also constitute variations in transmission patterns. As the real-world transmission dynamics of infectious diseases is essentially heterogenous due to these sources, models need to account for heterogeneity if their aim is to better understand the fine-grained transmission dynamics. While homogeneous models including the compartmental Susceptible-Infectious-Recovered (SIR) model are often used for their convenience and applicability to simple data (e.g. single-series daily case counts), oversimplifying the heterogeneity in data may result in neglecting nontrivial effects in the

1

transmission dynamics not captured by homogeneous models. Applying heterogeneous model to this dataset and exploring the role of heterogeneity and its implications will improve our understanding of desirable practices in infectious disease modelling.

The original aim of this PhD project was to investigate fine-scale heterogeneous transmission patters of influenza in schools and households using mathematical models. The inferred heterogeneous transmission across the school and household settings would then be used to propose designs for optimal school-based studies on intervention including mass vaccination. However, in early 2020, the emergence of coronavirus disease 2019 (COVID-19) rapidly developed into a pandemic and there was an urgent public demand for insights into its epidemiological characteristics and possible control strategies. To accommodate such a critical research agenda and also to further develop the scope of the PhD project, the research plan was slightly modified. Namely, additional studies focusing on the heterogeneity of transmission potential of SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2; the causative agent of COVID-19) were included, and the planned chapters for school-based intervention studies were reorganised to inform pandemic management strategies in school settings, reinstating the school influenza inference studies in the context of the COVID-19 pandemic. COVID-19 and influenza share many epidemiological characteristics, e.g. main routes of transmission and tendency to spread over social interactions, time scale of the course of infection (incubation period and generation time are on the order of days or 1-2 weeks) and the potential role of presymptomatic/asymptomatic infections. On the other hand, some of the observed features of COVID-19 are distinct from those of influenza, e.g. relatively minor occurrences of outbreaks in children and the relative importance of superspreading events (Hébert-Dufresne et al., 2020). These similarities and differences need to be considered when extrapolating understanding of influenza to COVID-19.

1.2 Aims and objectives

1.2.1 Aims

1. To understand the role of heterogeneity in dynamics and control of infectious diseases, specifically influenza and COVID-19.
2. To develop methodologies to handle heterogeneities, including multi-layer mixing in schools and households, individual-level covariates and overdispersion.

1.2.2 Objectives

1. To optimise the complexity of models to analyse the within-household transmission of

influenza and estimate relevant parameters to understand the dynamics among families (Paper 1)

2. To develop a method to correct for bias due to misclassified outcomes in test-negative design vaccine effectiveness studies in the presence of individual heterogeneities (Paper 2)

3. To quantify the degree of overdispersion of SARS-CoV-2 transmission in the initial phase of the 2020 pandemic (Paper 3)

4. To assess the effectiveness of 'backward contact tracing' in the presence of overdispersion in transmission dynamics (Paper 4)

5. To estimate the transmissibility of influenza over intra-class, intra-grade and inter-grade contacts in school and compare intervention strategies for schools during pandemics (Paper 5)

1.3   Outline of the thesis

This *research paper style* thesis consists of five papers, each representing a main research chapter. The current chapter serves as a brief overview of the PhD project. **Chapter 2** provides the background knowledge and contexts of the studies and **Chapter 3** describes the datasets used in the thesis, along with their collection process and limitations. **Chapter 4 (Paper 1)** uses a household transmission model to estimate parameters characterising the introduction and propagation of seasonal influenza in households of primary school students. Multiple models with different levels of complexity were compared to explore the best practice to analyse household outbreaks. **Chapter 5 (Paper 2)** proposes bias correction methods for the test-negative design (TND) studies, a study design that has been frequently used for recent influenza vaccine effectiveness (VE) studies. When imperfect tests are used in VE studies, misclassified disease outcomes can cause biased estimates. This study formularises the degree of bias in TND studies and derives a statistical approach to adjust for the bias, even in the presence of individual covariates usually included in vaccine studies as a standard practice. **Chapter 6 (Paper 3)** and **Chapter 7 (Paper 4)** explores the role of overdispersion in transmission (individual-level variation in the number of secondary transmissions) typically observed in the COVID-19 outbreak. **Chapter6** quantifies the degree of overdispersion of SARS-CoV-2 in a form of an overdispersion parameter of a negative binomial distribution from the earliest international case data. **Chapter 7** assesses how the existence of overdispersion improves the effectiveness of backward contact tracing compared with more traditional forward tracing. **Chapter 8 (Paper 5)** estimates transmission patterns of seasonal influenza within and between classes and grades in primary schools and applies them to simulated school outbreaks of COVID-19 and pandemic influenza to assess optimal management strategies in school settings. **Chapter 9** provides discussion and conclusion of the overall thesis.

2   Background

This chapter provides the background for the thesis and review of relevant literature. The chapter is divided into two sections: topics related to influenza and those related to COVID-19.

2.1   Influenza

2.1.1   Epidemiology and transmission dynamics

Influenza is an acute contagious respiratory disease caused by influenza virus. Three of the four types of influenza virus, A, B and C are known to infect humans (influenza D virus is not known to cause human infections), of which A and B causes seasonal epidemics (Su et al., 2017). Of all, influenza A virus (IAV) is of major public health concern. Frequent mutations in the antibody binding target hemagglutinin (HA) enable the virus to cause recurrent epidemic with limited effects of existing immunity in the population ("antigenic drift"). IAV subtypes are classified by the combination of hemagglutinin and neuraminidase (NA). Of the 18 HA and 11 NA subtypes identified to date, only 3 HA (H1, H2, H3) and 2 NA (N1, N2) subtypes are known to cause seasonal epidemics in human population (T. Watanabe et al., 2014). Natural reservoir of IAV is considered to be waterfowl, and reassortment of viral gene segments ("antigenic shift") from zoonotic and human origins, which may take place in pigs that can be infected both avian and human influenzas, is a potential source of a novel pandemic strain against which immunity in the human population is almost nonexistent (T. Watanabe et al., 2014). Influenza B virus (IBV), on the other hand, is known to be genetically more stable and has a narrower host range (humans and seals) (Osterhaus, 2000). Two IBV lineages, Yamagata and Victoria, are known to be circulating. Influenza C virus is known to cause mild upper respiratory tract infections, mostly in children, and is speculated to be widely distributed and people are immunised in their early life (1)(Matsuzaki et al., 2006; Salez et al., 2014).

In humans, influenza virus infects the upper respiratory tract and patients typically present common cold-like symptoms such as fatigue, sore throat and cough with sudden onset of fever (Punpanich & Chotpitayasunondh, 2012). The severity of the disease varies widely, and it sometimes causes severe illness including pneumonia and encephalitis, which can lead to a fatal outcome (Punpanich & Chotpitayasunondh, 2012; Van Kerkhove et al., 2011). Influenza virus circulates globally every year, estimated to be causing 1 billion annual cases and to be associated with 290,000-650,000 annual respiratory deaths (14–17)(Iuliano et al., 2018; World Health Organization, 2019). Especially, high mortality rates are estimated for sub-Saharan Africa, southeast Asia and elderly population aged 75 years or older (22–28)(Iuliano et al., 2018).

Influenza is easily transmitted from human to human via droplets, contacts and potentially short-range aerosols (Killingley & Nguyen-Van-Tam, 2013). The basic reproduction number, the number of secondary transmissions caused by a typical primary case in a totally susceptible population, is estimated at 1.2-1.4 for seasonal influenza (Biggerstaff et al., 2014). Symptoms appear after an incubation period of 1-2 days, but viral shedding starts before symptom onset and lasts for about a week (Carrat et al., 2008; Paules & Subbarao, 2017). Published estimates suggested that asymptomatic cases account for around 5-30% of total cases (although there is substantial heterogeneity in the estimates across different studies and designs) (Leung et al., 2015). Asymptomatic individuals can be infectious while they may not be captured by symptom-based surveillance. Some studies reported that their viral shedding level is lower than that of symptomatic cases (Carrat et al., 2008; Ip et al., 2017), but a model comparison study suggested infectiousness may not vary between symptomatic and asymptomatic cases (Wardell et al., 2017).

Influenza virus and other directly transmitted pathogens spread over social contact networks (Christakis & Fowler, 2010; Eubank et al., 2004; Keeling & Eames, 2005; Meyers et al., 2003, 2005; Newman, 2002; Volz & Meyers, 2009; Wang et al., 2014). In particular, social scenes which involve frequent conversational or physical contacts play important roles in the transmission dynamics (Ferguson et al., 2006; le Polain de Waroux et al., 2018). Typically, influenza is assumed to be transmitted on three major layers of mixing: households, schools/workplaces and general community (Cauchemez, Donnelly, et al., 2009; Ferguson et al., 2006; Fumanelli et al., 2012; Mossong et al., 2008). General community in this context summarises miscellaneous sources of contacts, including casual contacts with strangers, and therefore the first two layers, households and schools/workplaces, are in particular crucial components in social network studies.

Influenza vaccines are widely recommended to mitigate the disease burden. Due to the antigenic drift, vaccines need to be regularly updated to target the circulating strain, and yearly vaccinations are recommended especially for high risk groups including the elderly, small children aged 6-59 months, individuals with specific medical conditions and health care workers (World Health Organization, 2012). Most seasonal influenza vaccines include 2 influenza A strains and 1 or 2 influenza B strain(s) (trivalent/quadrivalent vaccines). The estimated effectiveness of influenza vaccines varies from year to year and is typically around 40-60% against symptomatic infections (Centers for Disease Control and Prevention, 2020a). Most studies assessing the effectiveness of influenza vaccine use symptomatic infection or hospitalisations as endpoints, and effectiveness against all infections (including asymptomatic infections) or transmission remains unclear.

2.1.2 Within-household transmission of influenza (Paper 1)

Quantification of transmission risk within households

Households are considered as one of the most important settings of transmission, as family members closely contact each other both conversationally and physically on a daily basis (Goeyvaerts et al., 2017; Ibuka et al., 2016; Mossong et al., 2008). Many epidemiological studies have investigated household data to infer the transmission dynamics of influenza within households (Lau et al., 2012; Tsang et al., 2016). By collecting infection status and other information of members from the entire household, researchers have estimated epidemiological properties of influenza, such as within-household transmission risk, serial interval, and risk factors. In particular, estimating within-household transmission risk is often regarded as a major objective in household studies, as it is an important quantity which characterises the importance of household transmission during the epidemic. The secondary attack rate (SAR) is a frequently used estimator for this purpose, which reflects the infectiousness of an index case in a household (Lau et al., 2012). SAR is defined as the number of secondary cases in households divided by the number of household contacts and is easily calculated from household study data. SAR has been historically used as a measure of proportion while the term "rate" usually implies time differential. Some recent studies refer to SAR as "secondary attack proportion" or "secondary infection risk" to avoid the use of "rate" (Klick et al., 2012; Lau et al., 2012; Tsang et al., 2016). The primary objective of SAR is to estimate the probability of infection per infectious case in the household (we hereafter refer to this probability as susceptible-infectious transmission probability; SITP). However, it has been pointed out that the crude SAR can be subject to bias (Tsang et al., 2016); the crude SAR and SITP are equal only if all secondary cases are caused by index cases, while this assumption is usually violated due to tertiary transmission caused by secondary cases and infections acquired from outside the household (community probability of infection; CPI). In such cases, SAR as a proxy for SITP often results in overestimation. This issue can be addressed by accounting for potential coprimary and tertiary cases (e.g. by defining generations of infection based on symptom onset dates), however, it is not necessarily clear whether this principle is endorsed in SAR studies (Madewell et al., 2020). A potentially more powerful approach to estimating SITP is to utilise mathematical models. Longini and Koopman (Longini & Koopman, 1982) developed a model capable of separately estimating SITP and CPI without the bias from final outcome data (i.e. without time-series information). The likelihood function of observing n cases in a household of size N is given by the following recursive equations:

$$\pi(n; N, \lambda, r) = \pi(n; n, \lambda, r) \binom{N}{n} \exp\big(-(\lambda + nr)\big)^{N-n},$$

$$\pi(n; n, \lambda, r) = 1 - \sum_{m=0}^{n-1} \pi(m; n, \lambda, r),$$

where λ and nr are a cumulative hazard of infection from outside and inside the household, respectively. The first equation decomposes the likelihood as a product of the probability that n individuals are fully infected and the probability that N-n individuals remain uninfected. $\exp(-(\lambda + nr))$ represents the probability of a susceptible individual escaping the risk of infection throughout the observation period, and thus the factor $\binom{N}{n} \exp(-(\lambda + nr))^{N-n}$ in the first equation gives the probability that N-n individuals escape the risk of infection. The probability of n individuals being fully infected, $\pi(n; n, \lambda, r)$, is obtained by subtracting from 1 the probability of observing less than n infections. The Longini-Koopman model and other forms of household final outcome models have enabled researchers to perform likelihood-based estimation of SITP and CPI for influenza (Ball & Neal, 2002; Becker & Britton, 1999; Cauchemez et al., 2014; House et al., 2012; O'Neill et al., 2000; Wardell et al., 2017).

Development of Longini-Koopman and related models

One of the oldest mathematical models that deals with transmission process in a discrete and closed population is first introduced in a lecture at Harvard University in 1928. This model, later known as the "Reed-Frost model", however was not documented in publications when it was first developed (Frost, 1976). The Reed-Frost model describes a chain of transmissions in a closed population as a simple stochastic process in discrete time steps representing generations of transmission. Individuals are assumed to follow homogeneous mixing and thus the risk of infection for an individual at specific time step t is given as $1 - q^{I_t}$, where I_t is the number of currently infectious individuals at time step t and q is the probability of not being infected by a specific infectious individual throughout a time step. By appropriately choosing parameters, the stochastic process of the Reed-Frost model can mimic the typical time evolution of an epidemic (Fine, 1977). If time-series data, i.e. the number of cases in each time step, is available, the parameter of interest, q, can be simply estimated by binomial likelihoods. If data is given only in a form of the final size of an outbreak, e.g. by serology, the likelihood for such an observation becomes slightly complicated and requires recursive computation (Bailey, 1975; Longini & Koopman, 1982; Ludwig, 1975): the Longini-Koopman model, defined using such a recursive process, allows for estimation of CPI and SITP separately as shown above. This model allows for relaxation of some of the assumptions of the Reed-Frost model, e.g. discrete time steps and identification of generation of cases because temporal information is not utilized in the likelihood of the final size.

One of the major assumptions of Longini-Koopman model is independence between households, i.e. CPI is treated as a single independent parameter and is not assumed to be affected by the level of infections in households in the community, which clearly does not hold in reality. A model proposed by (Ball et al., 1997) used a similar framework of the Longini-Koopman model but incorporated between-household transmission risks. They also showed that, with a large number of households, the assumption of a global CPI parameter in the Longini-Koopman model gives a sufficient approximation. The assumption of common within-household force of infection between individuals is another major assumption in the Longini-Koopman model. While discrete heterogeneity arising from individual characteristics is addressed by the heterogeneous extension (Longini et al., 1988), other sources of heterogeneity may remain. By assuming that infectious period may vary between individuals, (Addy et al., 1991; Ball, 1986) accounted for individual-level variations in the secondary transmission rates and proposed a random-effect model which requires a Laplace transform of the distribution of infectious period. This assumption of individual-level variation may be especially relevant if the disease exhibits a substantial overdispersion in transmission, e.g. COVID-19. Moreover, it needs to be noted that temporal variation in within-household transmission risks (e.g. during school term vs holiday) can also introduce differential secondary transmission rates, which is however difficult to address in estimation based on final outbreak sizes.

With the development of Bayesian inference approaches, which enabled imputation of latent variables in a relatively straightforward manner, recent household final size studies also attempted to reconstruct transmission trees within households and/or infection times along with parameter inference instead of computing the exact likelihood for the data (Bi et al., 2021; Cauchemez et al., 2014; House et al., 2012; O'Neill & Roberts, 1999). While such imputation approaches can typically be computationally intensive, their ability to directly obtain samples of transmission trees/infection times can be a strength compared to the traditional direct likelihood-based approaches. On the other hand, it should be noted that model selection in the presence of data imputation can be a complex problem.

Modelling heterogeneity and dependency in within-household transmission risk

In the original Longini-Koopman model, it was assumed that SITP is a single value within a household study dataset. However, it is natural to expect that SITP varies across different types of households. It may depend on individual characteristics of the infector and infectee and their contact frequency within the household. Characteristics of households such as household size, family composition or living environment may also be associated with SITP.

Heterogeneity is often incorporated in a model when the direction of transmission can be determined by data components, e.g. onset dates of cases (Cauchemez et al., 2011; Cauchemez, Donnelly, et al., 2009; Tsang et al., 2014). If the data lacks such temporal information, accounting for individual heterogeneity (multiple categories of individuals representing family structure) becomes a challenge (Tsang et al., 2016). Although a theoretical framework for a heterogeneous final outcome model has been proposed earlier (Demiris & O'Neill, 2005; Longini et al., 1988; Van Boven et al., 2007), to the best of our knowledge, only a limited number of household final outcome studies on influenza used it to account for more than two classes (e.g. adults and children) of individuals (Cauchemez et al., 2014; Wardell et al., 2017). These studies successfully performed likelihood-based estimation by imputing transmission chains within households. Although this approach can be powerful, one must be aware of the possible computational burden due to the nature of the imputation algorithm.

Household size is another important factor that may influence SITP. Many studies reported that household size was negatively associated with SITP (or SAR) (Buchholz et al., 2010; Cauchemez et al., 2011; Cauchemez, Donnelly, et al., 2009; House et al., 2012; Thai et al., 2014). This may be because household transmission is well described as frequency dependent. Frequency-dependent models assume that the risk of infection is determined by the proportion infectious in the household, not by the absolute number (Begon et al., 2002). Under this assumption, the contribution of an infected individual is divided by the total number of household members, so that SITP and household size are inversely proportional. In the presence of heterogeneity, the family composition may also be responsible for the negative association; larger households tend to consist of multiple generations and contact rates may vary from one type of individual to another (e.g., children, parents and grandparents). Previous studies reporting the effect of household size only used crude household size, and family compositions have not been accounted for in this context.

Determining optimal model complexity for household final outcome models

Complex models such as those incorporating heterogeneity have rarely been used in previous household final outcome studies, partly because of the difficulty in collecting data of sufficient size. In household studies, one household (not one individual) serves as one sample; therefore, a large number of households has to be included in order to be able to increase model complexity. A typical sample size of a household study is around a few hundreds of households. Simpler models may be able to estimate a set of parameters from such datasets, but when researchers wish to compare the performance of multiple complex models, larger sample sizes (typically thousands) may be required. As only few studies reported goodness-of-fit of multiple models with

given complexity (Cauchemez et al., 2011, 2014; Wardell et al., 2017), a consensus has not yet been established on optimal model complexity for capturing household final outcome data.

2.1.3 Influenza transmission on school network (Paper 5)

Social network of students in the class-grade-school structure

School is recognised as another important layer of influenza transmission (Glezen, 1996). A large number of children, who tend to exhibit a high prevalence of influenza possibly because of being less immune, closely interact with each other in a closed environment. The impact of the school-age population on outbreaks or effects of school-based interventions have been important targets of previous studies (Cauchemez, Ferguson, et al., 2009; Eames, 2014; Gemmetto et al., 2014; Heymann et al., 2009; Nishiura et al., 2014). In order to deliver useful insights into school transmission, it is necessary to understand the social network in school with a hierarchical structure (class-grade-school). Previous studies collected contact pattern data of students in hierarchical structures using self-written reports (Conlan et al., 2011; Leecaster et al., 2016) or wearable sensor devices (Fournet & Barrat, 2014; Guclu et al., 2016; Leecaster et al., 2016; Stehlé et al., 2011). These contact pattern studies suggested strongly assortative contact behaviours within classes and within grades. Strong assortativity was also observed in a contact tracing study in the 2009/10 H1N1 outbreak (Wang et al., 2014). Two previous studies modelled influenza outbreak in primary schools by differentiating intra-class, intra-grade and inter-grade contacts (Cauchemez et al., 2011; Clamer et al., 2016). It was suggested in (Cauchemez et al., 2011) that the inferred sources of infection of pupils were almost equally distributed among intra-class, intra-grade and inter-grade transmission; a similar distribution can also be conjectured from (Clamer et al., 2016) (though not explicitly presented in the paper).

Density- vs. frequency- dependent models for the school transmission network

Although there are a number of school contact pattern studies, their implication on real epidemics has been scarcely investigated due to limited data on school outbreaks which have class/grade information. Therefore, it is not well known how different features of contacts characterise transmission probability, e.g., the number, duration, mode of contacts and class/grade profile.
In modelling studies on school transmission, one has to make an assumption whether contact is density- or frequency-dependent (or the mixture of the two). This is especially important when different sizes of schools and classes are compared. This assumption also influences how parameter estimates should be interpreted. The force of infection that an individual experiences from a part of the population (e.g., individuals in the same class, same grade or same school) is given as

$$\lambda_i = \beta_{ij} c_{ij} \frac{I_j}{N_j}.$$

I_j and N_j are the numbers of infectious and total individuals in population j. β_{ij} is the transmission probability given infectious contact and c_{ij} is the contact rate, respectively, from population j to i (typically, β_{ij} and c_{ij} are simplified so that the model only differentiates between intra-class, intra-grade and inter-grade interactions). Contact is density-dependent if c_{ij} is proportional to N_j, and frequency-dependent if they are independent. The equation can be re-parameterised as $\lambda_i^{\mathrm{D}} = \beta_{ij}^{\mathrm{D}} I_j$ for density-dependent and $\lambda_i^{\mathrm{F}} = \beta_{ij}^{\mathrm{F}} \frac{I_j}{N_j}$ for frequency-dependent mixing. More generally, the dependency of the contact rates can be represented by introducing an exponent coefficient γ: $\lambda_i^{\mathrm{M}} = \beta_{ij}^{\mathrm{M}} \frac{I_j}{N_j^{\gamma}}$. Moreover, if a fine-scale dataset is available, complex modelling (e.g. using both class sizes and the number of classes as explanatory variables for contact rates) or estimation of the dependency (the values of γ) separately for different levels of the school structure (i.e. at intra-class, inter-class and inter-grade levels) may be possible.

Mixing assumptions should be carefully chosen for larger-scale school transmission studies that include various sizes of population units (classes/grades/schools). However, as there are few school outbreak data that have sufficient details to inform granular transmission patterns and their dependencies on the school structure. Two existing school outbreak studies (Cauchemez et al., 2011; Clamer et al., 2016) estimated transmission parameters at different levels of school structure using mathematical models. They classified any possible pair of students in the dataset into intra-class, intra-grade and inter-grade pairs and assigned different parameters for each type of relationships. The results suggested that such relationships are an important determinant of transmission risks of influenza. Intra-class contacts, followed by intra-grade, was estimated to be substantially frequent sources of influenza transmission between students. Meanwhile, a potential caveat in extending these studies is that both studies assumed density-dependent mixing patterns. Although the effect of arbitrarily adopting this assumption may have been minimal in these studies as the number of schools included was very small (*n*=1 and 2, respectively), it may have a non-negligible impact if their approach is directly applied to large-scale datasets. While the previous contact pattern studies in school settings reported that the number of contacts was correlated with larger class sizes (Hens et al., 2009; Melegaro et al., 2017), due to the lack of comparison with actual patterns of disease transmission, such observations from contact studies may be sensitive to how contacts are defined and measured.

2.1.4 Test negative design as a tool for assessment of vaccine effectiveness (Paper 2)

Due to rapid mutation in the antigenic domains (antigenic drift), influenza vaccines typically need to be updated annually in order to target the most circulating strain in the year (Boni, 2008). For this reason, vaccine effectiveness (VE) studies should be routinely conducted to monitor the field effectiveness of vaccines in each season. Test negative design (TND) is a relatively recently developed design for VE studies, often used for influenza vaccines. Unlike the traditional case-control design that recruits a control group from the general population, TND enrols medically-attended patients with specific symptoms (e.g. flu-like illness) and classifies them into case and control groups according to their test results. Participants of TND studies are thus conditioned not only on the presence of symptoms (including those caused by causes other than the disease of interest) but also on medical attendance, which is expected to minimise the ascertainment bias (De Serres et al., 2013; Fukushima & Hirota, 2017). In addition to this methodological strength, TND is preferred in clinical studies because it enables the use of routinely collected clinical records without additional recruitment effort for a control group. VE is estimated as an odds ratio between test-positive ('cases') and test-negative ('controls') patients, and in well-controlled settings, the obtained odds ratio corresponds to the relative risk attributed to vaccination (Haber et al., 2015). However, previous studies found that TND is often more sensitive to the misclassification bias, where the estimated VE is affected due to disease outcomes potentially mislabelled with false results of tests (De Smedt et al., 2018; Jackson & Rothman, 2015; Orenstein et al., 2007). These studies also showed that specificity may have a strong influence on the degree of the misclassification bias, although they were limited to specific ranges of parameter values and lacked comprehensive exploration of possible settings.

Even if misclassifications of disease outcomes are inevitable due to the limited test performance, a statistical adjustment may be possible if the estimates of sensitivity and specificity of the test are available. Such a bias correction method has been proposed for cohort studies, but a similar correction was shown to be impossible for case-control studies without external information such as the baseline prevalence of the disease (Greenland, 1996). While a number of studies assessed the potential degree of bias in different settings (De Smedt et al., 2018; Jackson & Rothman, 2015; Orenstein et al., 2007), bias correction methods have not been developed for TND studies. This may be partially because TND studies are often considered as a special case of case-control studies. However, TND studies have a statistically distinct property from case-control studies which can potentially allow for bias corrections. As TND has become a central approach to VE studies, practical bias correction methods specifically designed for TND will be of paramount epidemiological interest. Such methods need to be able to handle multivariate analysis because it is a standard approach to adjust for multiple covariates such as age and sex in TND studies.

2.2 COVID-19

2.2.1 Epidemiology and transmission dynamics

In late December 2019, China reported a cluster of pneumonia cases of unknown cause in Wuhan city, Hubei Province. In January, a novel coronavirus was isolated as a causative agent of the disease, which was later officially named SARS-CoV-2 (and the disease COVID-19) (World Health Organization, 2020c). By the end of January, nearly 10,000 confirmed cases were confirmed in China and 19 countries across regions including Asia, Europe and the Americas reported importation of cases (World Health Organization, 2020b). WHO declared a pandemic on 11 March 2020, when 114 countries had reported confirmed cases (World Health Organization, 2020d).

Overall, COVID-19 presents a similar spectrum of symptoms to influenza, ranging from common cold-like illness (e.g. cough, fever and fatigue) to serious symptoms (e.g. dyspnoea and pneumonia). Symptoms can also include taste, olfactory or gastrointestinal disorders, which are also reported for influenza (Minodier et al., 2015; Vetter et al., 2020; Welge-Lüssen & Wolfensberger, 2006). While most infections only experience asymptomatic or mild infection, COVID-19 cases can also lead to serious conditions and deaths (especially among those with comorbidities or in old age) at a significantly high risk than seasonal influenza (Petersen et al., 2020; UK Office for National Statistics, 2020).

Transmission of SARS-CoV-2 follows typical routes of respiratory infectious diseases, such as via droplets and fomites (Rahman et al., 2020). In addition, growing evidence suggests that aerosols, especially in a relatively short-range, may also be involved in transmission, as was also suggested for influenza (Banik & Ulrich, 2020; Klompas et al., 2020). The basic reproduction number R_0 of SARS-CoV-2 is estimated to be around 2-3, which is slightly higher than that of influenza. While symptoms usually appear 2-11 days post-infection (McAloon et al., 2020), infectiousness can precede the symptom onset by 1-2 days (He et al., 2020).

2.2.2 Overdispersion in transmission of SARS-CoV-2 (Papers 3, 4)

From the early phase of the pandemic, it has been suggested that the distribution of the number of secondary transmissions (offspring distribution) of COVID-19 cases may be highly heterogeneous. Contact tracing conducted in multiple countries found that most traced cases caused no or only few secondary transmissions, while a small proportion of cases were involved in so-called superspreading events (Bi et al., 2020; Nishiura et al., 2020), thereby raising the reproduction number as an average up to 2-3.

Such highly dispersed offspring distributions are often represented by a negative binomial distribution (Lloyd-Smith et al., 2005). A negative binomial distribution is characterised by two parameters: mean, which corresponds to R_0 in the case of offspring distribution, and overdispersion parameter (often denoted by k), which quantifies the degree of dispersion:

$$\text{NB}(x; R_0, k) = \binom{x + k - 1}{x} \left(\frac{R_0}{R_0 + k}\right)^x \left(\frac{k}{R_0 + k}\right)^k.$$

A negative binomial distribution converges to a Poisson distribution as k approaches infinity. In the previous novel coronavirus outbreaks (SARS-CoV; severe acute respiratory syndrome coronavirus and MERS-CoV; Middle East respiratory syndrome coronavirus), k was estimated to be small (0.16 for SARS-CoV (Lloyd-Smith et al., 2005) and 0.26 for MERS-CoV (Kucharski & Althaus, 2015)), suggesting a substantial degree of dispersion in the offspring distribution. Being closely related to these viruses, SARS-CoV-2 was expected to have a similar overdispersion, while few reliable estimates of k were available in the early phase of the pandemic.

Existence of substantial overdispersion in transmission has several implications on the dynamics and control of an outbreak. The probability of extinction of an outbreak given a small number of initial cases becomes higher in the presence of overdispersion because most initial cases do not contribute to the expansion of the outbreak (Waxman & Nouvellet, 2019). Early observations of a number of imported cases in some countries that did not lead to sustained local outbreaks were in line with this property of overdispersion (World Health Organization, 2020a). If the upper tail of the offspring distribution corresponding to those causing superspreading events are effectively suppressed by limiting the size of social gatherings or identifying factors associated with superspreading events (Leclerc et al., 2020), R_0 could be efficiently reduced. Namely, a high degree overdispersion suggests that transmissions tend to form chains of large clusters that contribute to the sustained spread of the outbreak and many more chains involving few or no secondary transmissions becoming extinct, which suggests that control measures may benefit from focusing on larger clusters.

2.2.3   COVID-19 and school outbreaks (Paper 5)

Children accounted for a relatively small proportion of confirmed cases of COVID-19 and are suggested to present generally mild symptoms, if any (Centers for Disease Control and Prevention, 2020b; European Centre for Disease Prevention and Control, 2020; Götzinger et al., 2020; Lee et al., 2020; Wu & McGoogan, 2020). Empirical evidence on the role of schoolchildren in the transmission of SARS-CoV-2 is still limited. Most studies estimated that children are less

susceptible than adults, while some suggested equal or higher susceptibility of children, especially in older age groups (10-19 years old) (Viner et al., 2020). Estimates of infectiousness of children relative to adults are scarce. However, existing studies suggest they may be slightly higher than adults (Fateh-Moghadam et al., 2020; Park et al., 2020). While viral loads of SARS-CoV-2 infected children showed similar or slightly higher distributions compared with adults (Heald-Sargent et al., 2020; Yonker et al., 2020), how these may translate to infectiousness is unclear.

COVID-19 outbreaks at school settings are considered to be rare (Buonsenso et al., 2020; European Centre for Disease Prevention and Control, 2020; Ismail et al., 2020; Russell et al., 2020). However, the currently observed school outbreaks may be underrepresented due to multiple factors. Children may be less likely to be tested due to mild or asymptomatic infections, which preclude possible school outbreaks from being recognised. Many countries enforced countrywide school closures from spring to summer in 2020 and some countries continued to keep them closed also in fall onward (United Nations Educational Scientific and Cultural Organization, 2020), restricting the chance of transmissions in school settings. A recent study suggested that these interventions may have been effective in reducing transmission (Brauner et al., 2020). Schools were advised to implement intensive prevention measures including physical distancing, environmental cleaning and reduction in class sizes upon reopening (Centers for Disease Control and Prevention, 2020c; Department for Education; UK Government, 2020), which may also have contributed to limiting the risk of sustained transmission at schools. Nonetheless, a number of outbreaks in school settings have been reported including those involving transmissions between a large number of students (Goldstein et al., 2020; Kommuneoverlegen in Lillestrøm municipality et al., 2020; Ministry of Education Culture Sports Science and Techonogy Japan, 2020; Stein-Zamir et al., 2020; Torres et al., 2020), which indicates that school outbreaks could still occur under certain conditions.

2.3 Bayesian inference

2.3.1 Parameter estimation using Markov-chain Monte Carlo (Papers 1, 3, 5)

The major characteristics of Bayesian inference is that parameter estimates are given as conditional distributions given data (posterior distribution) as opposed to point estimates often given in other inference frameworks, e.g. maximum likelihood estimation. Given data D, the posterior distribution for parameter θ is formulated using the Bayes' theorem

$$p(\theta|D) = \frac{p(\theta)p(D|\theta)}{p(D)},$$

where $p(\theta)$ is the prior distribution, $p(D|\theta)$ is the likelihood and $p(D) = \int p(\theta)p(D|\theta)d\theta$

is the marginal likelihood. A choice of the prior distribution needs to be made a priori in Bayesian inference, and this can sometimes be a debatable issue. Existing data, background knowledge and requirement for mathematical convenience are often considered in the choice of the prior. The prior distribution constitutes an important part of the model specification, and therefore reasoning and the resulting model performance should be clarified and assessed.

Even with the likelihood and prior distribution specified, the posterior distribution cannot be directly obtained because the integral involved in the marginal likelihood $p(D) = \int p(\theta)p(D|\theta)d\theta$ is typically intractable. Instead, Monte Carlo approaches are used in practice to efficiently sample from the posterior distribution.

Markov-chain Monte Carlo (MCMC) is one of the most utilised method for sampling from the posterior distribution where the integral of the marginal likelihood is intractable (Kass et al., 2012). MCMC requires an "unnormalised" function of θ proportional to the density distribution to sample from; i.e. the integral of the function over all possible region of θ does not have to be 1. As the values of the likelihood $p(D|\theta)$ and prior distribution $p(\theta)$ for given θ are usually available, their product $f(\theta) = p(\theta)p(D|\theta)$, which is proportional to the target posterior distribution $p(\theta|D)$, can be supplied.

While many specific algorithms for MCMC are available, the most basic algorithm is the Metropolis-Hastings algorithm. Starting from an arbitrary initial value, the algorithm iteratively updates the parameter sample θ by the following update rule and yield a sequence of samples $\boldsymbol{\theta} = (\theta^1, \theta^2, \dots \theta^N)$.

1. Given the current sample θ, propose a new sample θ' from the "proposal distribution" $q(\theta'|\theta)$

2. Accept-reject the proposed sample θ' at a probability $p_{\text{Accept}} = \min\left\{1, \frac{f(\theta')q(\theta|\theta')}{f(\theta)q(\theta'|\theta)}\right\}$.

   If θ' is rejected, use the current sample as a new sample instead.

3. Repeat 1-2 until convergence.

This update rule ensures that after a sufficient number of iterations, the resulting sequence θ converges to the target distribution $p(\theta|D)$. In practice, the first part of samples is under the influence of the arbitrary initial value and therefore discarded as "burn-in" samples. In addition, "thinning" of samples is often performed to reduce the sample length, where only 1 in every few samples are extracted and recorded.

Of multiple criteria proposed to assess whether a convergence of an MCMC chain is achieved,

the effective sample size (ESS) (Kass et al., 2012) is one of the most convenient and intuitive criterion, defined as

$$\text{ESS} = \frac{N}{1 + 2\sum_{k=1}^{\infty} \rho(k)},$$

where N is the raw sample size and $\rho(k)$ is the autocorrelation of samples with lag k. As an MCMC chain derives from a Markov chain, the samples have a certain level of autocorrelation, it contains less information than an independent set of samples with the same size. ESS indicates what independent sample size the MCMC chain with autocorrelation is equivalent to and thus can be used to assess whether the chain is long enough.

2.3.2 Model selection (Papers 1, 3)

When there are multiple candidate models to apply to the same data, one needs a standard quantitative measure to compare the performance of the models. The marginal likelihood $p(D) = \int p(\theta)p(D|\theta)d\theta$ provides one of such measures because the marginal likelihood corresponds to the probability of observing the data given the specific model (i.e. the "likelihood of a model"). Alternatively, the prediction performance of models measured by the Kullback-Leibler divergence between the predictive distribution and true distribution of data can be used as a measure for model comparison, e.g. Akaike information criterion (AIC). However, these prediction-based criteria were not employed here since the main interest in model selection in this thesis is understanding mechanisms behind data generating process rather than prediction.

While the marginal likelihood itself cannot be easily estimated, there are several methods that provide an estimate of the marginal likelihood multiplied with a constant factor. One of the most well-known examples is the Bayesian information criterion (BIC), which uses a Laplace approximation of the likelihood function around the mode. Although BIC is convenient and widely used, it can only be applied to regular models, i.e. models with a positive finite Fisher information matrix. If the model is singular (with a non-positive definite Fisher information matrix) or nearly singular, BIC does not have a theoretical support because the likelihood function around the mode is not asymptotically normal. Widely-applicable Bayesian information criterion (WBIC) is proposed as a more general measure in Bayesian inference that can be used for both regular and singular models (S. Watanabe, 2013). When data D of size N is independently and identically distributed, WBIC is defined as

$$\text{WBIC} = \frac{p(D|\theta)^{1/\log N}p(\theta)}{\int p(D|\theta)^{1/\log N}p(\theta)d\theta} \log p(D|\theta).$$

In practice, this can be computed as the mean log-likelihood over MCMC samples where

$p(D|\theta)^{1/\log N} p(\theta)$ is the target distribution. WBIC as defined here has half the scale of AIC/BIC, and a difference of 1 is considered as significant.

3 Data sources used in the thesis

3.1 Matsumoto city primary school influenza data (Papers 1, 5)

3.1.1 Overview

This dataset was collected as part of a citywide primary school influenza survey for an observational study that looked into the effect of vaccination and non-pharmaceutical preventions against seasonal influenza in Matsumoto city, Nagano prefecture, Japan (Uchida, Kaneko, Hidaka, Yamamoto, Honda, et al., 2017b, 2017a). In 2014/15 flu season, students and their parents in all the 29 public primary schools in Matsumoto city were asked to respond to a questionnaire on influenza during the season. The survey was two-fold: in the first part ("prospective survey") conducted from October 2014 to February 2015, students who reported infection to school for a granted absence were enrolled; in the second part ("retrospective survey") conducted in March 2015, all students were eligible regardless of influenza episode. Students who had influenza during the season may have been included in both of the surveys; however, their data were unlinkable between the surveys as the questionnaires were separate and anonymous. In both of the surveys, questionnaires were distributed to eligible students, and their guardians were asked to respond on behalf of the students. The questionnaire consisted of a variety of questions, including the student's background information, whether the students had influenza during the season, onset date and observed symptoms, vaccination history, the family composition and who in the same household had influenza during the season. Participants were asked to report diagnosed influenza in the questionnaire and most of the student cases (95%) had rapid diagnostic test results of influenza-A positive. As the diagnosis of influenza by physicians is usually required for primary school students to be granted absence in Japan, students with influenza-like illness are strongly encouraged to visit medical institutions. During the study period, the schools reported 2,651 cases to the municipal board of education and 2,548 students (96%) responded to the prospective survey. In the retrospective survey, 13,217 students were eligible and 11,390 (86%) students responded. After removing those with missing values, 10,486 (79%) were available for analysis. A household may have had more than one student eligible for the study; in such cases, questionnaires were collected independently and were not linked with each other due to the anonymity of the questionnaire. Therefore, it is expected that the data contains multiple records on the same household but are handled as different households in the following analysis. The original study found that 48.1% of the students

included had vaccines, 52.0% regularly wore masks and 79.1% engaged themselves in hand washing. Further details of the dataset can also be found in the corresponding research papers (Papers 1 and 5).
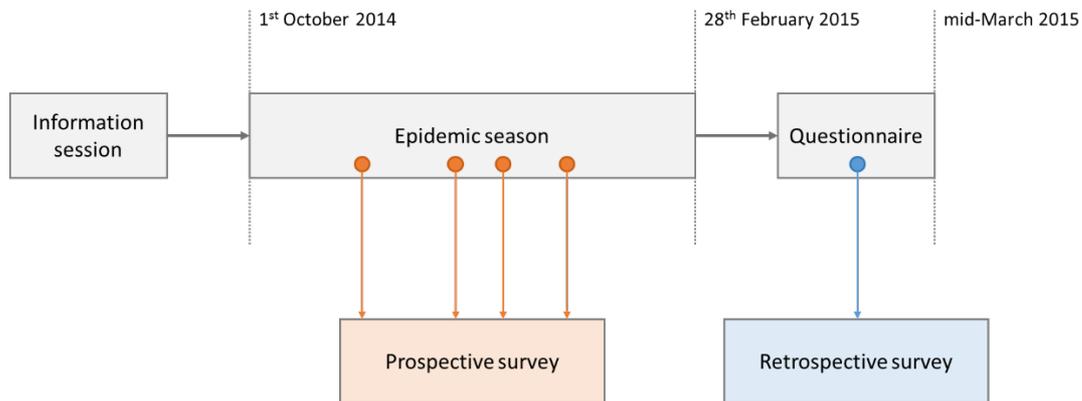

Figure 3.1. A schematic illustration of the data collection timeline.
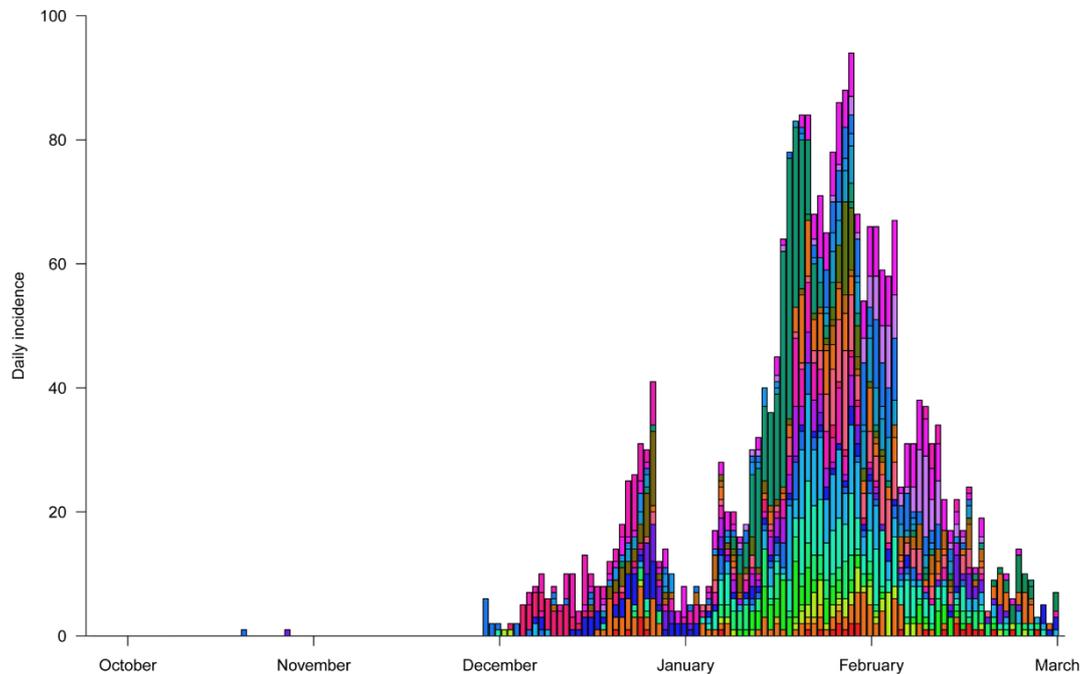
Figure 3.2. Temporal distribution of seasonal influenza cases by symptom onsets in 29 primary schools in Matsumoto city, Japan, 2014-15 epidemic season. The reported cases in different primary schools are denoted by colours.

19

3.1.2 Case finding

Japanese primary schools are legally obliged to report the number of influenza cases in their students to the municipal board of education. Many Japanese primary schools thus require students to report influenza after recovery, along with a letter of diagnosis issued by medical institutions. In Japan, when a physician suspects influenza from the symptoms of a patient, usually a rapid diagnostic test, which detect influenza virus antigen in nasal swab by immunochromatography, is used to confirm diagnosis. Since it was not recorded which kit was used for diagnosis, it was not possible to establish the precise sensitivity and specificity values relevant to the data in this study. Multiple systematic reviews reported that the sensitivity and specificity of rapid diagnostic tests are estimated to be 50-70% and 98-99%, respectively (Bruning et al., 2017; Chartrand et al., 2012). However, we found in these reviews that the sensitivity for studies conducted in Japan was relatively high (range: 72.9-96.4%). Those values are also consistent with earlier studies conducted in Japan (M Hara et al., 2004; Michimaru Hara et al., 2006; Yamazaki et al., 2004) and we assumed that the sensitivity in Japanese setting may be around 80-90%. It should also be noted that in some cases physicians may diagnose influenza without performing test or with a negative result if they conclude it is highly likely from clinical and epidemiological information. However, such cases must have been rare in this dataset, as 96.4% of the cases reported that their diagnosis specified the virus type (type A: 94.9%, B: 1.5%), which is unlikely without positive rapid test results (Uchida, Kaneko, Hidaka, Yamamoto, Honda, et al., 2017b).

3.1.3 Characteristics of study participants

The retrospective survey data provided the distribution of household compositions in the study population (Figure 3.3 and Table 3.1). Households of size four were the most frequently observed, which were typically composed of two parents and two children. The maximum household size in the study population was 10 and the maximum number of siblings was 6, but such large households were very rare. The 20 most frequent household compositions shown in Table A1 accounted for over 95% of the reported household compositions. The profiles of cases in the prospective and retrospective surveys are displayed in Table 3.2. Further details of the study population are documented in the original study (Uchida, Kaneko, Hidaka, Yamamoto, Honda, et al., 2017a).

3.1.4 Limitations

Multiple limitations of this dataset must be noted. First, the available data may have misrepresented the population of interest (issue of representability of participants). Such potential

misrepresentation can be classified as that regarding internal validity and external validity. While the prospective survey included almost every documented influenza case (96%) during the study period, the retrospective survey had a slightly lower respondent rate (86%). Although this respondent rate may not be particularly low compared with other questionnaire-based studies, if refusal to respond to the survey was associated with specific characteristics, e.g. history of influenza or adherence to precautionary measures, the dataset may be subject to selection bias (affecting internal validity). The dataset represents a specific subpopulation in Matsumoto city, i.e. households with at least one primary school student. Therefore, the inferred transmission patterns such as within-household transmission risks may not be generalisable to households with no primary school age child. In addition, careful assessment is required when extrapolating findings from this dataset to another geographical, social and cultural settings as the dataset may reflect characteristics unique to Matsumoto city, Japan. For example, education systems, typical familial roles and interactions, immunological landscape and circulating strains of virus that are specific to Matsumoto city may need to be considered when the study findings are transferred to different settings (issues of external validity). Second, the collected data may be subject to 'measurement errors', where the information provided by each participant may not reflect their true status. In the current dataset, underascertainment and recall bias in particular may be the major sources of bias. As mentioned in Section 3.1.2, it is expected that the dataset captured most symptomatic influenza in children. However, asymptomatic infections or very mild symptoms may have been missed; false negative test results may also have led to loss of 10-20% of symptomatic cases (unless test-negative cases are still diagnosed based on clinical findings). Moreover, the presence of illness, medical attendance or test results were not specifically asked in the questionnaire for influenza episodes of household members, which increased the uncertainty on the infection status of household members. Potential underascertainment arising from these factors, especially differential underascertainment between students and household members can lead to bias in the results, typically underestimation of transmission risks. Sensitivity analysis was conducted in Paper 1 to account for potential underascertainment. However, the same approach was not possible for the time-series analysis conducted in Paper 5 (because the temporal information of unreported cases cannot be reconstructed) and the interpretation of results require caution. Recall bias is another major issue in the dataset, especially in the retrospective survey. The retrospective survey was conducted in March and asked the influenza episodes, precautionary measures, etc. during the whole study period (November 2014-February 2015). Given that the long time frame of the survey, it is not unlikely that some participants might have provided inaccurate information. Students with reported history of influenza episodes were included in the prospective survey where they responded right after recovery, therefore their information may be more reliable as long as the prospective data is used as reference; however,

only the retrospective data is available for those without influenza episodes, which may even emphasize systematic recall bias between case and control groups. This can be an issue particularly in Paper 5 where the prospective and retrospective survey datasets are combined. Covariates on precautionary measures were adjusted for as part of the analysis assuming that both case and control groups have the same likelihood of giving inaccurate responses. However, this approach could not account for possible differences in accuracies between case and control groups and thus recall bias may remain unadjusted for.
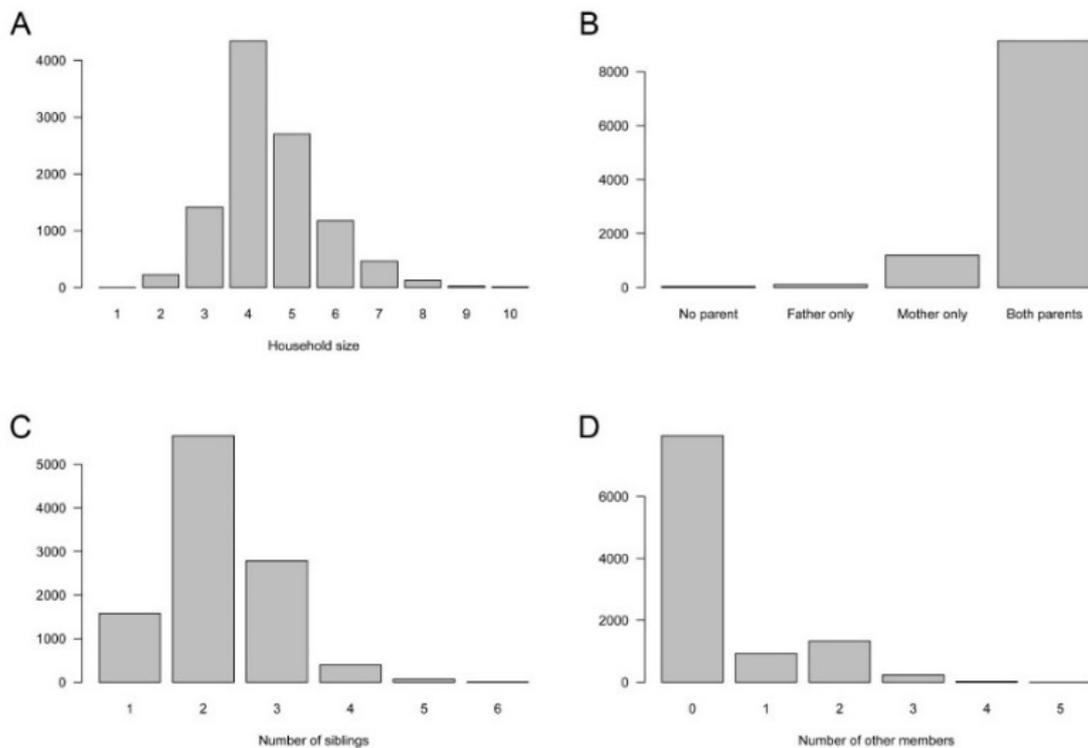
Figure 3.3. The frequency distribution for the household size, the number of parents, siblings and other members. (A) Household size, (B) parents in the households, (C) the number of siblings (including the respondents themselves), and (D) the number of family members other than parents or siblings.

Table 3.1. Frequency distribution table for compositions of households included in the retrospective data

Order	Composition	# of households	Order	Composition	# of households
1	FM-2	3,915	11	M-3	160
2	FM-3	1,971	12	FM-1-2	134
3	FM-1	899	13	FM-1-1	97

4	FM-2-2	606	14	M-1-2	86
5	M-2	429	15	M-2-2	80
6	FM-2-1	415	16	FM-2-3	70
7	FM-3-2	297	17	FM-3-3	57
8	FM-4	250	18	FM-4-2	55
9	FM-3-1	232	19	M-1-1	43
10	M-1	205	20	M-2-1	39
				Subtotal	10,040 (95.7%)

Only 20 most frequent compositions are shown, accounting for 95.7% of the total 10,486 responses. Household compositions are denoted in the following manner.

FM: households with both father and mother; M: households with only mother; The first number: the total number of siblings in the household; The second number (where applicable): the number of other members (mostly grandparents) in the household.

Table 3.2. The number of individuals and influenza cases in each type

| Individual type | | Prospective | Retrospective | | |
		Cases	Counts	Cases	Attack ratio (%)
Student	Overall	2,537	10,410	2,137	20.5
	Male	1,329	5,311	1,132	21.3
	Female	1,208	5,099	1,005	19.7
	Year 1	488	1,831	406	22.2
	2	418	1,773	363	20.5
	3	419	1,731	342	19.8
	4	446	1,717	375	21.8
	5	374	1,674	322	19.2
	6	396	1,684	329	19.5
Father			9,201	629	6.8
Mother			10,260	866	8.4
Sibling			12,632	2,320	18.4
Other			4,356	191	4.4

* The number of respondents and cases for "Father", "Mother", "Sibling" and "Other" is obtained from the response to the questionnaire and may be redundant due to the inclusion of multiple students from the same household.

3.2 WHO situation reports for COVID-19 (Paper 3)

In response to major outbreaks of infectious diseases, World Health Organization (WHO) typically issues situation reports which aim to deliver up-to-date reports of the status and impacts of the outbreak. The first WHO situation reports on COVID-19 was published on 21 January 2020, 3 weeks after WHO was informed of pneumonia cases of unknown cause in China and shortly after imported confirmed cases of COVID-19 (known as 2019-nCoV at the time) were reported in several neighbouring countries (World Health Organization, 2020c). Since then, situation reports were issued every day until 16 August 2020 when the reports became weekly. In the earliest phase of the pandemic, the number of confirmed cases in each country was reported by likely source of infection (i.e. with or without likely exposure in China and foreign countries outside China). However, due to the rise in the overall number of cases, WHO started to aggregate the number of cases in reporting countries and only provided the countries-level status ("Imported cases only", "Local transmission" or "Under investigation"). In Paper 3, we used the latest situation report with the stratified number of cases by likely sources (Situation report 30, 27 April 2020 (World Health Organization, 2020a)).

3.2.1 Limitations

This dataset is a summary data based on reports from member countries to WHO and thus the quality and characteristics of data are not necessarily standardised. In particular, the degree of underascertainment and the accuracy of imported/local cases labels are possible factors that could have affected the estimation results. The degree of underascertainment can be affected by multiple aspects of surveillance, e.g. the intensity of border control, testing capacity and accuracy, public awareness and administrative process. Differential ascertainment where the degree of underascertainment is different between types of cases may also need to be considered. For example, compared with imported cases with travel history, identifying local cases can be difficult especially in the earliest phase of an outbreak. As a sensitivity analysis, the effect of potential underascertainment of cases on the results has been explored including differential ascertainment. The results of the sensitivity analysis suggested that while the estimates were relatively robust to differential ascertainment, underascertainment of imported cases can cause underestimation of the overdispersion parameter k. Mislabelling of imported and local cases can happen when, for example, travel history of a case is unreported or a case with travel history (and thus considered as an imported case) actually acquired infection locally. While such mislabelling cannot be completely excluded, the relative contribution of mislabelling may be limited in the early phase of an outbreak when travel history was a crucial information in epidemiological investigation.

3.3 Ethical approval

Ethical approval for the secondary analysis of Matsumoto city dataset was obtained from the LSHTM ethics committee (reference number: 14599). The original study was approved by Committee for Medical Ethics of Shinshu University (approval number: 2715). WHO situation report is publicly available and thus did not require ethical approval.

4 Paper 1: Fine-scale family structure shapes influenza transmission risk in households: Insights from primary schools in Matsumoto city, 2014/15

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1700902		Title	Dr
First Name(s)	Akira			
Surname/Family Name	Endo			
Thesis Title	Roles of heterogeneity in infectious disease epidemiology: implications on dynamics, inference and control of influenza and COVID-19			
Primary Supervisor	Sebastian Funk			

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	PLoS Computational Biology		
When was the work published?	26 December 2019		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate conceived the study, designed the model, performed the analysis and wrote the original draft of the manuscript.

SECTION E

Student Signature	Akira Endo
Date	10 December 2020

Supervisor Signature	Sebastian Funk
Date	10 December 2020

Improving health worldwide   www.lshtm.ac.uk

Improving health worldwide   Page 2 of 2   www.lshtm.ac.uk

PLOS | COMPUTATIONAL BIOLOGY

Fine-scale family structure shapes influenza transmission risk in households: Insights from primary schools in Matsumoto city, 2014/15

Akira Endo[1]*, **Mitsuo Uchida**[2], **Adam J. Kucharski**[1,3], **Sebastian Funk**[1,3]

1 Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, United Kingdom, **2** Department of Public Health, Graduate School of Medicine, Gunma University, Gunma, Japan, **3** Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

* akira.endo@lshtm.ac.uk

Check for updates

Abstract

Households are important settings for the transmission of seasonal influenza. Previous studies found that the per-person risk of within-household transmission decreases with household size. However, more detailed heterogeneities driven by household composition and contact patterns have not been studied. We employed a mathematical model that accounts for infections both from outside and within the household. The model was applied to citywide primary school seasonal influenza surveillance and household surveys from 10,486 students during the 2014/15 season in Matsumoto city, Japan. We compared a range of models to estimate the structure of household transmission and found that familial relationship and household composition strongly influenced the transmission patterns of seasonal influenza in households. Children had a substantially high risk of infection from outside the household (up to 20%) compared with adults (1–3%). Intense transmission was observed within-generation (between children/parents/grandparents) and also between mother and child, with transmission risks typically ranging from 5–20% depending on the transmission route and household composition. Children were identified as the largest source of secondary transmission, with family structure influencing infection risk.

Author summary

We characterised detailed heterogeneity in household transmission patterns of influenza by applying a mathematical model to citywide primary school influenza survey data from 10,486 students in Matsumoto city, Japan, one of the largest-scale household surveys on seasonal influenza. Children were identified as the largest source of secondary transmission, with family structure influencing infection risk. This suggests that vaccinating children would have stronger secondary effects on transmission than would be assumed without taking into account transmission patterns within the household.

Introduction

Respiratory infectious diseases transmitted by droplets, exemplified by influenza, are known to spread through social contact networks [1,2]. Social settings that involve frequent contacts play important roles in transmission dynamics [3,4]. Households are considered one of the main layers of transmission because individuals come in close contact with each other both conversationally and physically on a daily basis [5–8]. Several epidemiological studies have used household data to investigate the transmission dynamics of influenza within households [9,10], particularly in terms of the secondary attack rate (the number of secondary household cases divided by the number of household members at risk). However, this assumes that an index case (the first case in a household, who is considered to be infected outside the household) is responsible for all subsequent household cases and that all the other household members are equally at the risk of secondary infection.

The possibility of co-primary infections and tertiary transmission is neglected under such assumptions [9], with potentially heterogeneous transmission patterns between household members also radically simplified. The former limitation can be addressed by mathematical models which separately estimate the risk of infection from outside the household and the within-household transmission risk [11]. Many household studies have employed the Longini-Koopman model and other related models to study the within-household transmission dynamics of influenza [12–18].

On the other hand, potentially-heterogeneous transmission patterns have not been fully studied with empirical data. Multiple household modelling studies have incorporated factors including age, vaccination status and antibody titres to account for heterogeneity, but these are usually used to identify individual risk factors that determine relative susceptibility of individuals [15,17,19–21]. Given typical behaviours within the family, it is natural to expect substantial heterogeneity in household contact patterns related to familial relationships and household compositions, on top of those individual factors [7,8]. Addy et al. [22] estimated a within-household transmission matrix consisting of two classes (children and adults), but more classes might be needed to better account for the heterogeneity of household contact patterns. In actual implementation, even such two-class analysis is very rare; in most cases, household size is the only family-related covariate for modelling of within-household transmissions in outbreak data [14,15,18,19,23]. Further, due to the limited sample size of households in these studies, a rationale on the quantitative effect of household size in transmission has not been established. Familial roles/relationships (e.g., father, mother, grandparent, etc.) have been paid far less attention to in household outbreak studies; we found only one field study on influenza that included familial roles as a covariate, a descriptive study that did not quantify the risk by familial roles [24].

Households serve as important units in intervention policies [25,26]. Tailored quantification of the transmission risks from outside and inside the household could help prioritise and promote household-level prevention strategies including vaccination. If specific compositions of households have a higher risk of outbreaks than others, intervention policies may be optimised by targeting such households. Moreover, as vaccine uptake is shown to be influenced by the perceived risk of infection and vaccine effectiveness [27,28], identifying the household-specific risk of infection and the possible reduction by vaccines may support highlight the individual benefit of vaccination.

To investigate the within-household transmission dynamics of seasonal influenza, we applied a highly flexible household transmission model that accounts for heterogeneity to a large influenza dataset. The dataset included more than 10,000 primary school students with the infection status not only of students but also of their household members, which enabled a detailed investigation of within-household transmission dynamics. Focusing on the effect of

28

familial roles and household compositions, we compared multiple models with different levels of complexity to find the best model to describe the transmission patterns.

Methods

Ethics statement

The data analysis in the present study was secondary and was approved by the ethics committee at the London School of Hygiene & Tropical Medicine (reference number: 14599). Consent was not obtained because the data were anonymous upon collection. The original study was approved by the Committee for Medical Ethics of Shinshu University (reference number: 2715).

Data source

We used data from a citywide primary school influenza survey. At the end of the 2014/15 season (early March), parents of students at all 29 public primary schools in Matsumoto city, Nagano prefecture, Japan, were asked to respond to a questionnaire consisting of a variety of questions including whether the students had influenza during the season, onset date and observed symptoms, vaccination history, family composition and who in the same household had influenza episodes during the season. The data was originally collected for an observational study on the effect of prevention measures against seasonal influenza (Uchida et al., 2017) [29]. In the present study, we only considered data on influenza episodes in students, their household composition and influenza episodes in the household members. Participants reported the number of siblings in the household, and also ticked the type of family members (such as "father", "younger sister" or "uncle") with whom they live, as well as whether they acquired influenza in the 2014/15 season. Among 13,217 students eligible, 11,390 (86%) responded to the survey. After removing those with missing values, 10,486 surveys were used in the present study. Characteristics of the population and frequent household compositions are shown in Tables 1 and 2. The influenza types reported for the student cases during the season were mostly A (95% of those tested positive) [30]. The national-level surveillance data suggested that AH3 strain was predominant, accounting for 99% of the type A isolates [31]. The

Table 1. The number of individuals and influenza cases in each type.

Individual type		Counts*	Cases*	Attack ratio (%)
Student	Overall	10,410	2,137	20.5
	Male	5,311	1,132	21.3
	Female	5,099	1,005	19.7
	Grade 1	1,831	406	22.2
	2	1,773	363	20.5
	3	1,731	342	19.8
	4	1,717	375	21.8
	5	1,674	322	19.2
	6	1,684	329	19.5
Father		9,201	629	6.8
Mother		10,260	866	8.4
Sibling		12,632	2,320	18.4
Other		4,356	191	4.4

* The number of respondents and cases for "Father", "Mother", "Sibling" and "Other" is obtained from the response to the questionnaire and may be redundant due to the inclusion of multiple students from the same household.

https://doi.org/10.1371/journal.pcbi.1007589.t001

29

Table 2. Frequency distribution table for compositions of households included in the retrospective data.

Order	Composition	# of households	Order	Composition	# of households
1	FM-2	3,915	11	M-3	160
2	FM-3	1,971	12	FM-1-2	134
3	FM-1	899	13	FM-1-1	97
4	FM-2-2	606	14	M-1-2	86
5	M-2	429	15	M-2-2	80
6	FM-2-1	415	16	FM-2-3	70
7	FM-3-2	297	17	FM-3-3	57
8	FM-4	250	18	FM-4-2	55
9	FM-3-1	232	19	M-1-1	43
10	M-1	205	20	M-2-1	39
				Subtotal	10,040 (95.7%)

Only 20 most frequent compositions are shown, accounting for 95.7% of the total 10,486 responses. Household compositions are denoted in the following manner. FM: households with both father and mother; M: households with a single mother; The first number: the total number of siblings in the household; The second number (where applicable): the number of other members (mostly grandparents) in the household.

https://doi.org/10.1371/journal.pcbi.1007589.t002

vaccination coverage of the students in the dataset was 48%; however, we did not consider vaccination in the present analysis. Further details of the data collection and descriptive epidemiology can be found in the original studies [29,30].

In the present study, we classified each individual in households as one the following type: "father", "mother", "student", "sibling", or "other". "Students" are participants of the survey (i.e., students of primary schools in Matsumoto city), and "siblings" are their elder/younger siblings, who may have also been recruited in the survey if they are primary school students (however, they are not linked in the data and thus unidentifiable as participants). The parameters for "students" and "siblings" were differentiated because "siblings" are not necessarily primary school students, therefore their characteristics may be different from "student". "Father" and "mother" were labelled as "single-parent" if they are only one parent in the family; models were considered in the model selection where their parameter values were differentiated from cohabiting parents (details described in "model selection"). Most individuals classified as "other" were grandparents (90.1%). Uncles/aunts accounted for 6.7%, and the remaining 3.2% was "none of the above categories".

In the survey, all students who were reported to have acquired influenza were also reported to have been diagnosed at a medical institution. For other household members, clinical diagnosis was not clearly required on the question sheet. In Japan, rapid diagnostic tests are usually used for suspected patients. International systematic reviews estimated that the sensitivity and specificity of rapid diagnostic tests are 50–70% and 98–99%, respectively [32,33]. However, the sensitivity for studies conducted in Japan included in these reviews was relatively high (range: 72.9–96.4%), consistent with other earlier studies conducted in Japan [34–36]. Considering that many Japanese primary schools encourage students presenting influenza-like symptoms to consult medical institutions so that they are granted absence, we believe that the reported influenza episodes in the dataset were sufficiently inclusive for our analysis. We also performed a sensitivity analysis to address possible underreporting in the survey (described later).

Heterogeneous chain binomial model

We employed the chain-binomial model presented in [37] which allows for heterogeneous transmission (see Fig 1 for schematic illustration). Let N be a vector representing the number
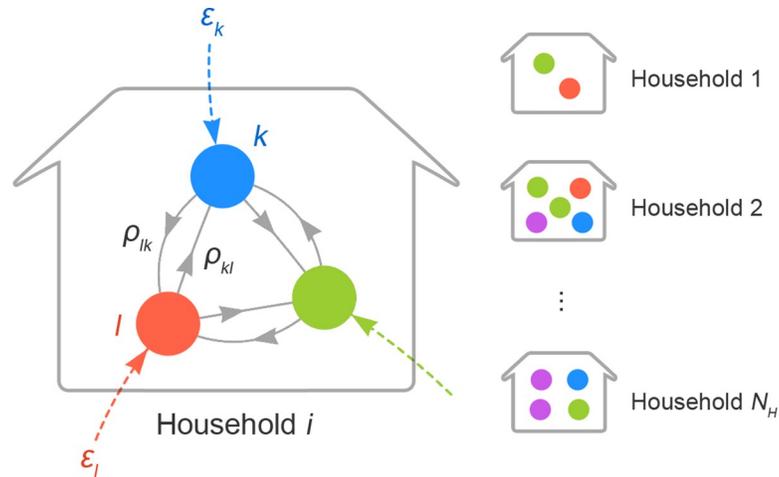
Fig 1. A schematic illustration of household chain-binomial model. Nodes in different colours correspond to different types of individuals (e.g., father, sibling, etc.). Transmission patterns are illustrated taking household i as an example. Coloured dotted edges represent the risk of external infection $\boldsymbol{\varepsilon}$ to each individual. Solid grey edges denote person-to-person transmission risk (PTR) from one type of person to another. PTR from type l to k is given as ρ_{kl}, which refers to the risk of transmission given that the individual of type l is infectious. Households have different compositions and ρ_{kl} may also vary according to the composition. On the other hand, $\boldsymbol{\varepsilon}$ is the risk from outside the household and thus assumed to be identical across households.

of family members stratified by individual type (e.g., father, mother, child, etc.) in a household. The probability that a certain combination of individuals (represented by a vector $\boldsymbol{n}$) in the household are infected by the end of the season is given by the following recursive equations.

$$\pi(\boldsymbol{n}; \boldsymbol{N}, \boldsymbol{\varepsilon}, H) = \pi(\boldsymbol{n}; \boldsymbol{n}, \boldsymbol{\varepsilon}, H) \prod_k \binom{N_k}{n_k} S_k(\boldsymbol{n}, \boldsymbol{\varepsilon}, H)^{N_k - n_k},$$ (1)

$$\pi(\boldsymbol{n}; \boldsymbol{n}, \boldsymbol{\varepsilon}, H) = 1 - \sum_{\boldsymbol{v} < \boldsymbol{n}} \pi(\boldsymbol{v}; \boldsymbol{n}, \boldsymbol{\varepsilon}, H).$$

where N_k and n_k are the k-th component of $\boldsymbol{N}$ and $\boldsymbol{n}$, respectively ($1 \leq k \leq K$). The sum $\sum_{\boldsymbol{v} < \boldsymbol{n}}$ is taken for all vector $\boldsymbol{v}$ satisfying $0 \leq v_k \leq n_k$ ($\forall k$) and $\boldsymbol{v} \neq \boldsymbol{n}$. We denoted by $\boldsymbol{\varepsilon}$ the external risk of infection over the season for each type of individual. The susceptible-infectious transmission probability (SITP) ρ_{kl} is the probability of within-household transmission for a specific infectious-susceptible pair [18] and has been used to quantify within-household transmission. However, it is more convenient to use the effective household contact matrix $H = (\eta_{kl})$ in the model; η_{kl} is defined to satisfy $\rho_{kl} = 1 - \exp(-\eta_{kl})$, and is interpreted as the amount of contact that leads to within-household transmission (effective contact) from type l to k. That is, η_{kl} denotes the amount of exposure that an individual k experiences when another individual of type l in the same household is infectious. $S_k(\boldsymbol{n}, \boldsymbol{\varepsilon})$, the probability that a type k individual escapes infection from both outside and inside the household throughout the season, is given as

$$S_k(\boldsymbol{n}, \boldsymbol{\varepsilon}, H) = (1 - \varepsilon_k) \exp\left(-\sum_l \eta_{kl} n_l\right).$$ (2)

$(1 - \varepsilon_k)$ is the probability that the individual is not infected outside the household, and $\exp\left(-\sum_l \eta_{kl} n_l\right)$ is the probability that the individual is not infected from any of the household infectives. When a dataset $\{\boldsymbol{N}_i, \boldsymbol{n}_i\}$ contains the family composition and infection status in each household i, the pseudo-likelihood function (where interaction between households is

31

neglected) is given as

$$L(\boldsymbol{\varepsilon}, H; \{\boldsymbol{N_i}, \boldsymbol{n_i}\}) = \prod_i \pi(\boldsymbol{n_i}; \boldsymbol{N_i}, \boldsymbol{\varepsilon}, H). \tag{3}$$

The household-wise likelihood $\pi(\boldsymbol{n_i}; \boldsymbol{N_i}, \boldsymbol{\varepsilon}, H)$ is computed by recursively applying Eq (1) starting with $\pi(\boldsymbol{0}; \boldsymbol{0}, \boldsymbol{\varepsilon}, H) = 1$.

Transmission risk in households

We modelled the possible heterogeneity in household transmission by parameterising the effective household contact matrix $H = (\eta_{kl})$. Our basic assumptions are: (i) each pair of individuals have a specific "intensity of contact"; (ii) the relative importance of each household contact may be reduced if an individual experiences a large amount of household contacts in total; (iii) the contact intensity adjusted by the total amount of contact is proportional to the force of infection. That is, we modelled η_{kl} as

$$\eta_{kl} = \beta \frac{c_{kl}}{C_k^{\gamma}}. \tag{4}$$

The contact intensity c_{kl} represents the (hypothetical) number of household contacts between type k and l, and β is the transmissibility coefficient. C_k represents the total number of household contacts experienced by an individual of type k, which we introduced to investigate how η_{kl} differs in households of different sizes and compositions. Noting that the number of individuals in contact is N_k-1 if $k = l$, we get

$$C_k = \sum_l c_{kl}(N_l - \delta_{kl}), \tag{5}$$

where δ_{kl} is the Kronecker delta. The value of the exponent parameter γ determines how strongly η_{kl} is scaled by C_k, which associates our model with density-dependent vs. frequency-dependent mixing assumptions [38]. The value $\gamma = 0$ corresponds to the density-dependent mixing assumption, where the force of infection is proportional to the total number of contacts (weighted by intensity) with infectives, whereas $\gamma = 1$ corresponds to the frequency-dependent mixing assumption, where it is the proportion of infectious contacts among total contacts that matters. In addition to $\gamma = 0$ and $\gamma = 1$, γ was also allowed to be estimated as a free parameter in the model selection, representing a mixture of density-dependent and frequency-dependent mixing.

The contact intensity matrix (c_{kl}) is interpreted as the per-individual version of the contact matrix ($c_{kl} = b_{kl}/N_l$ where b_{kl} is the contact matrix). The parameter c_{kl} generally constitutes a $K \times K$ matrix and contains too many parameters to estimate. We, therefore, reduced the number of parameters by categorising contacts into the following 5 pairs first:

$$c_{kl} = \begin{cases} c_{\text{CC}} \ (\text{Child} - \text{Child}) \\ c_{\text{FC}} \ (\text{Father} - \text{Child}) \\ c_{\text{MC}} \ (\text{Mother} - \text{Child}) \\ c_{\text{OC}} \ (\text{Other} - \text{Child}) \\ c_{\text{AA}} (\text{Adult} - \text{Adult}) \end{cases} \tag{6}$$

Child included both "student" and "sibling", and adult included "father", "mother" and "other". (In models where "single-parent" is a separate type, another parameter c_{SC} (Single parent–Child) was added.) The matrix was assumed to be symmetric, i.e, $c_{kl} = c_{lk}$. We did not vary β in our baseline analysis such that transmission is also symmetric ($\eta_{kl} = \eta_{lk}$), but the possibility of type-specific susceptibility was addressed in our sensitivity analysis. Since we did not have a measurement for the intensity of household contacts in our dataset, we used relative values of

c_{kl} in our analysis where c_{AA} was assumed to be 1. The parameter β is approximately equal to the probability of transmission in a (hypothetical) household composed of only father and mother (since $\frac{c_{kl}}{C_k^i} = 1$ regardless of γ).

Statistical analysis and model selection

We sampled parameter values from a posterior distribution yielded from the pseudo-likelihood function (3) and priors in Table 3 using the Markov-chain Monte Carlo (MCMC) method. An optimal variance-covariance matrix for proposal was explored by the adaptive Metropolis algorithm, and then the random-walk Metropolis algorithm was used to obtain final samples. All MCMC sampling was performed using the R package {LaplacesDemon}. The scripts and dataset to produce MCMC samples for the main results are reposited on GitHub (https://github.com/akira-endo/HHstudy_FluMatsumoto2014-15).

First, we tested various possible combinations of assumptions on the effective contact matrix and the risk of external infection (shown in Table 3) and compared their goodness of fit by Widely-applicable Bayesian Information Criterion (WBIC) [39]. Model variants included (i) homogeneous or heterogeneous mixing in households (c_{kl}), (ii) uniform or heterogeneous risk of external infection (ε_k), (iii) the value of the exponent parameter (γ), and (iv) whether the parameter values for a single parent is differentiated from those of cohabiting parents. Characteristics of compared models are documented in S1 File, Section 1. WBIC for each model was computed from 80,000 MCMC samples which were thinned from 125,000 samples × 8 chains so that the chains had the effective sample size (ESS) ~40,000.

We then used the models selected by WBIC to estimate the parameters. As final samples, 10,000 thinned samples were recorded from 40,000 pre-thinned MCMC samples. It was ensured that the ESS was at least 500 for each parameter.

Using the estimated parameters, we computed the source-stratified risk of infection and the risk attributable to the introduction into the household (see S1 File, Section 2 for further details).

Further model development

When the parameters were estimated with the best model selected, we found that the estimates for c_{FC} and c_{OC} were very similar, which suggested that we might be able to equate these two

Table 3. Parameter estimates by the best model.

Parameter		Prior	Estimate (95% CrI)
External risk (ε_k)	Student	1-LogUnif(0,1)*	0.197 (0.188–0.207)
	Sibling		0.161 (0.153–0.169)
	Mother		0.035 (0.030–0.040)
	Father		0.038 (0.033–0.043)
	Other		0.013 (0.009–0.017)
Contact intensity (c_{kl})	Child-Child	Unif(0,∞)	1.04 (0.88–1.23)
	Mother-Child		1.16 (1.00–1.32)
	Father-Mother		1 (0.748–1.282)
	Other-Other		1.97 (1.10–3.24)
	Cross generational		0.43 (0.35–0.52)
Transmissibility (β)		(derived quantity; not sampled by MCMC)	0.20 (0.16–0.24)
Exponent parameter (γ)		Unif(−∞,∞)	0.51 (0.33–0.69)

* Cumulative force of infection is uniformly distributed.

33

parameters and further stratify the contacts between adults (c_{AA}) with the degree of freedom earned. We tested some other contact intensity matrices, including

$$c_{kl} = \begin{cases} c_{CC} \ (\text{Child} - \text{Child}) \\ c_{MC} \ (\text{Mother} - \text{Child}) \\ c_{FM} \ (\text{Father} - \text{Mother}) \\ c_{OO} \ (\text{Other} - \text{Other}) \\ c_{X} \ (\text{Cross generational}) \end{cases} \qquad (7)$$

which gave the best performance in the end. Explored candidate models and selection results are detailed in S1 File, Section 2.

Sensitivity analysis

We performed a sensitivity analysis to address potential biases in our dataset. We considered in our sensitivity analysis (i) ascertainment bias, (ii) different susceptibility in children, (iii) multiple counting of households and (iv) censoring of sibling cases.

The first two points are related to the assumptions in our models. Influenza can have a low reporting rate due to mild clinical presentation (including asymptomatic infections), and therefore some infectious individuals may not have been included in our dataset. The reporting rate of influenza is considered to be very high in primary school students in Japan, who are often required to report influenza to their schools. On the other hand, the reporting rate of adults can be lower, as they may be less likely to seek medical treatment than children. A serosurvey conducted in Japan after the 2009/10 H1N1 influenza pandemic suggested that while influenza in children was almost fully reported, the reporting rate of adults were relatively low (30–50%) [40].

Another possible difference between adults and children is susceptibility: adults may be less likely to be infected by the same amount of exposure due to the previous history of infections or stronger immune systems than children. Conversely, children may exhibit lower susceptibility if the vaccine uptake for them is higher than adults. The majority of household transmission studies from a systematic review [9] reported a significant association between susceptibility and age (although this becomes the minority when limited to the studies with PCR-confirmed cases). Our baseline model assumes that transmissibility β is identical between individuals, but in reality, transmissibility might depend on the age of the susceptibles.

The remaining points explored in sensitivity analysis are inherent limitations in our dataset. One of the limitations is that, because students in the same household responded to the questionnaire separately, households with multiple siblings may have been counted more than once. As this was an anonymous questionnaire, data obtained from different students were not linked with each other even if they were from the same household. If there was more than one child in a household who was eligible for the study, the same household transmissions can appear multiple times in the dataset, which could modify the results. Lastly, because of the design of the questionnaire, the number of influenza cases in siblings may have been underreported. The questionnaire asked whether each type of individual in the same household had influenza during the season, and the respondents ticked if at least one individual of that type was infected since it was a yes-no question. Therefore, even if there was more than one case in the same type of individuals, the number was not reported and treated as a single case; that is, if a respondent has two older brothers, he/she only reports that "older brother had influenza", and there was no distinction on the dataset whether it was only one or both of them. This issue was addressed by modifying the likelihood function.

34
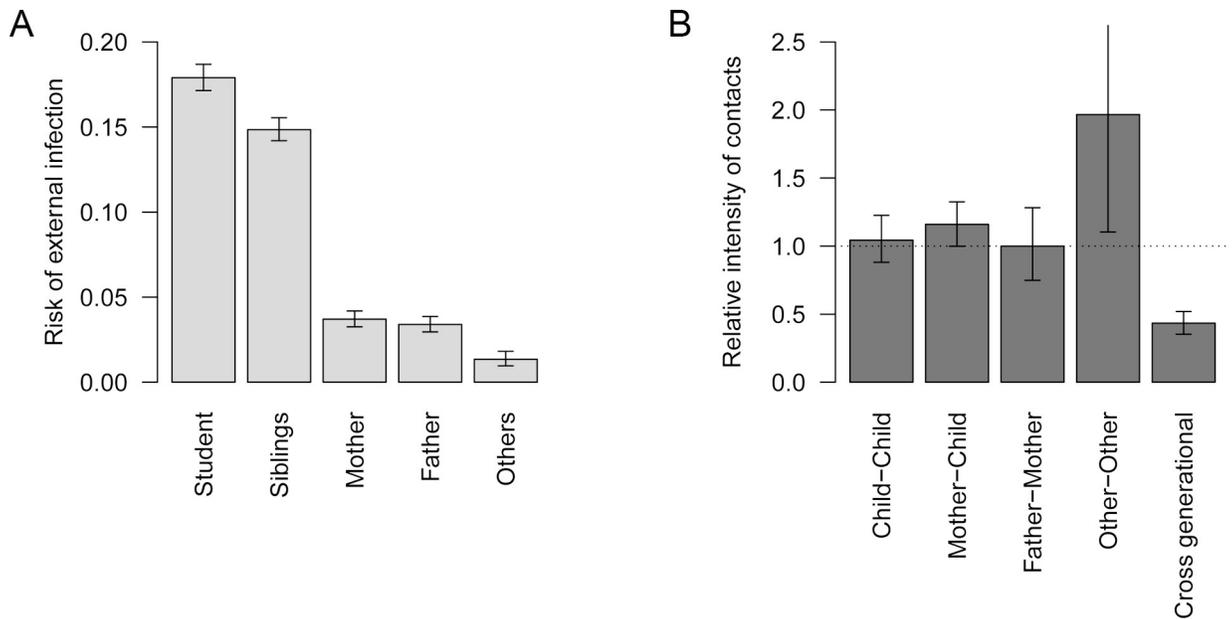
Fig 2. Estimated risk of external infection and relative intensity of within-household contact. (A) Estimated risk of external infection for each type of individual. (B) The estimated relative intensity of within-household contact. Values are scaled so that the median of contact intensity between adults is 1 (horizontal dotted line). Whiskers indicate 95% credible intervals (CrI).

Each potential source of bias was addressed by incorporating the data-generating process causing the bias into the model. Technical details of the sensitivity analysis can be found in S1 File, Section 3.

Results

We found considerable heterogeneity in both the risk of external infection and the risk of within-household transmission (Table 3 and Fig 2). The best performing mathematical model suggested that children had a comparatively high risk of infection outside the household: 20% in the primary school students and 16% in their siblings, compared to only 1–3% in adults. Within-household contact patterns showed strong generational clustering. High contact intensities were observed within the same generation (between siblings, parents and grandparents), and the intensity of cross-generational contacts was less than half the intensity within the same generation. Contact between mothers and children was an exception to this, showing a higher intensity than between parents. The estimated contact intensity relative to that between parents (father-mother) was highest between other-other (1.97; CrI: 1.10–3.24), most of whom were grandparents in our data, followed by mother-child (1.16; CrI: 1.00–1.32) and child-child (1.04; 0.88–1.23), both of which are insignificantly different from father-mother (1; 0.75–1.28). The model did not support a significant difference between parameter estimates for single and cohabiting parents.

The inferred networks of household transmission suggest that various contact patterns between household members exist in different household compositions. The contact intensity between individuals are shown in network graphs (Fig 3A–3C) for three selected characteristic household composition models, "nuclear family": FM-2 (see Table 2 for the notation), (b) "many-siblings family": FM-4, and (c) "three-generation family": FM-2-2. Mothers served to bridge between the generations of children and parents; clusters of grandparents were relatively independent of other household members.
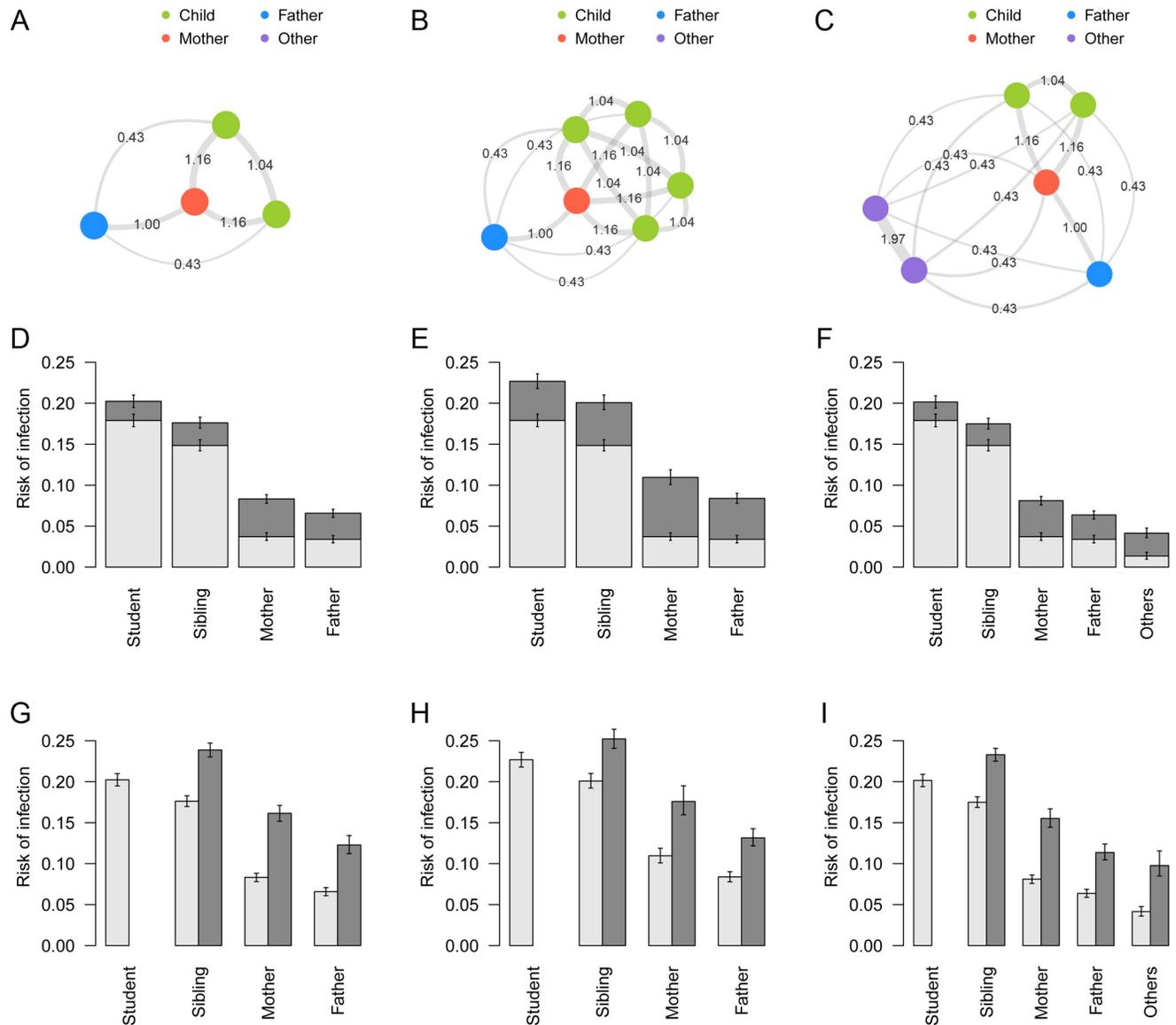
Fig 3. Contact patterns and risk of infection in specific household compositions. (A)-(C) Network graphs showing contact intensity between individuals for different household compositions: (A) "nuclear family", (B) "many-siblings family", (C) "three-generation family". Node colours represent the type of individuals. Edges denote the relative intensity of contact (c_{kl}) between individuals. (D)-(F) Risk of infection in households of different compositions stratified by source. Light grey: risk of infection from outside the household; dark grey: risk of infection from within the household. Whiskers indicate the 95% CrI. (G)-(I) Unconditional risk of infection and conditional risk given an introduction of infection into a household. Light grey: overall risk of infection for each individual in the household; dark grey: risk of overall infection conditional that a student is infected outside and introduces infection into the household. Infection of the student is considered given, and thus the conditional risk for the student is not shown. Whiskers indicate the 95% CrI.

https://doi.org/10.1371/journal.pcbi.1007589.g003

The overall risk of infection and the breakdown of infection source presented in Fig 3D–3F suggests that risk of infection in children was mostly from outside the household, whereas a larger proportion of risk in adults was attributed to within-household transmission. Risk of within-household infection increased when more children were in the household (Fig 3E); however, the influence of additional members categorised as "others" (grandparents in most cases) was minimal, probably due to their low risk of external infection and contact intensity (Fig 3F). On the other hand, for grandparents in a typical three-generation household, the risk of infection from inside the household was twice the risk from outside.

36
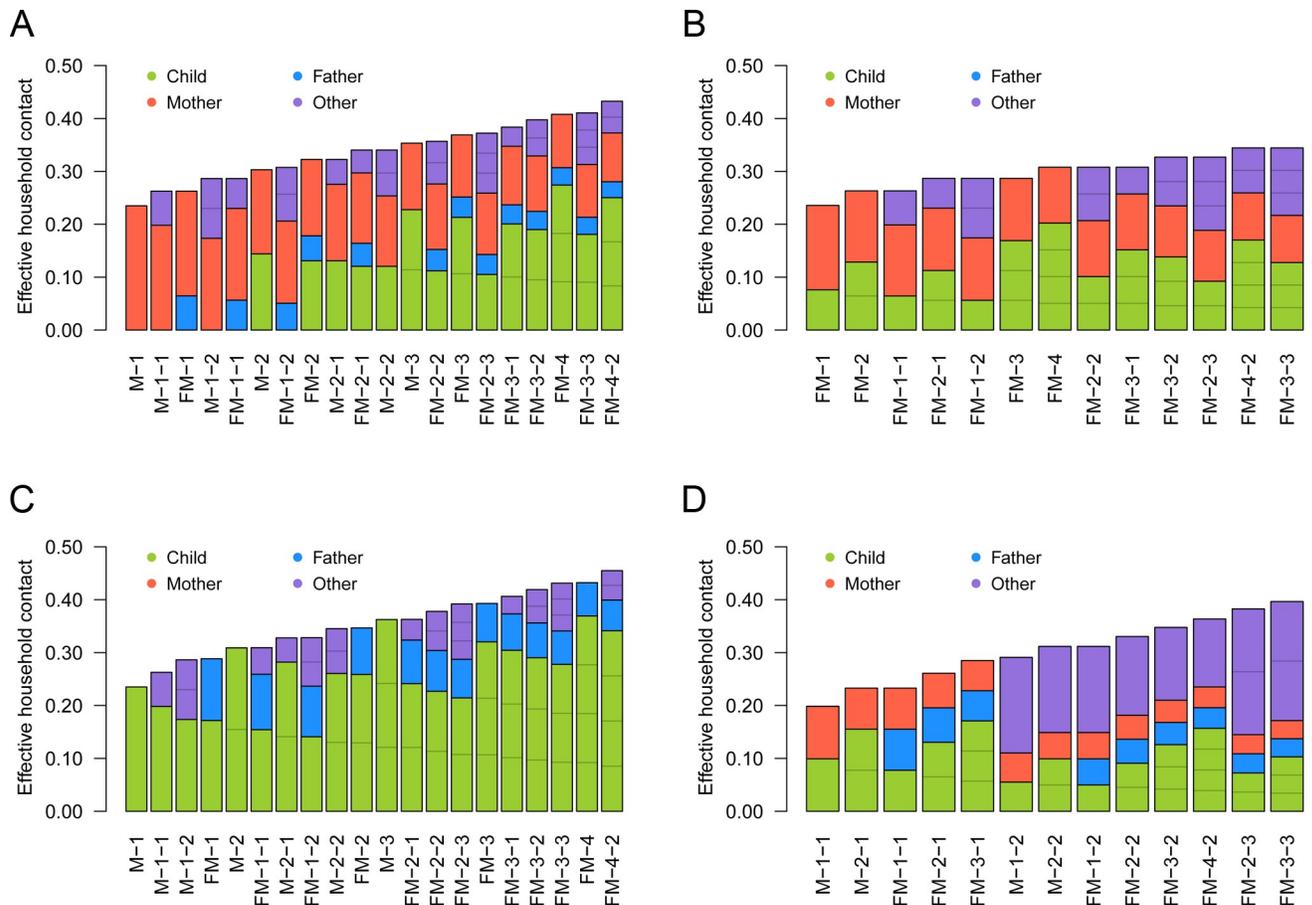
Fig 4. The effective amount of contacts experienced by individuals (η_{kl}) in different household compositions. (A) Child; (B) Father; (C) Mother; (D) Other. The coloured compartments denote the breakdown of effective contacts allocated to each individual in the household, which corresponds to SITP given that individual is infectious.

https://doi.org/10.1371/journal.pcbi.1007589.g004

Once influenza was brought into a household by a student, the conditional risk of infection in other members of the household increased substantially; the implication of disease introduction into households can be seen in the simulated risk of infection after introduction (Fig 3G–3I). In "nuclear family" and "three-generation family" models, the risk in adults increased by a factor of 2–3 if a primary school student in the family was infected.

The effective household contacts that each type of individual experiences are displayed in Fig 4, indicating the substantial variation in household contact patterns between individuals and between households. SITP typically ranged around 5–20%, depending on the contact pair and household composition. Reflecting the estimated value of $\gamma = 0.5$ (CrI: 0.3–0.7), the total amount of effective household contacts was greater in larger households, but the weight of each single contact (the effective contact corresponding to contact with one individual in the household) decreased with household size. This is because the effective household contact η_{kl} that one experiences followed an "inverse square root law", i.e., η_{kl} is inversely proportional to the square root of the total amount of contact C_k ($\eta_{kl} \propto 1/C_k^{0.5}$; see Eq 4).

Although Fig 4 summarises the heterogeneous within-household transmission patterns, one must note that the secondary transmission is conditional to infection in the primary case. When the contacts were weighted by the risk of external infection to visualise the source of primary and secondary infections for each individual, it can be seen that the children were
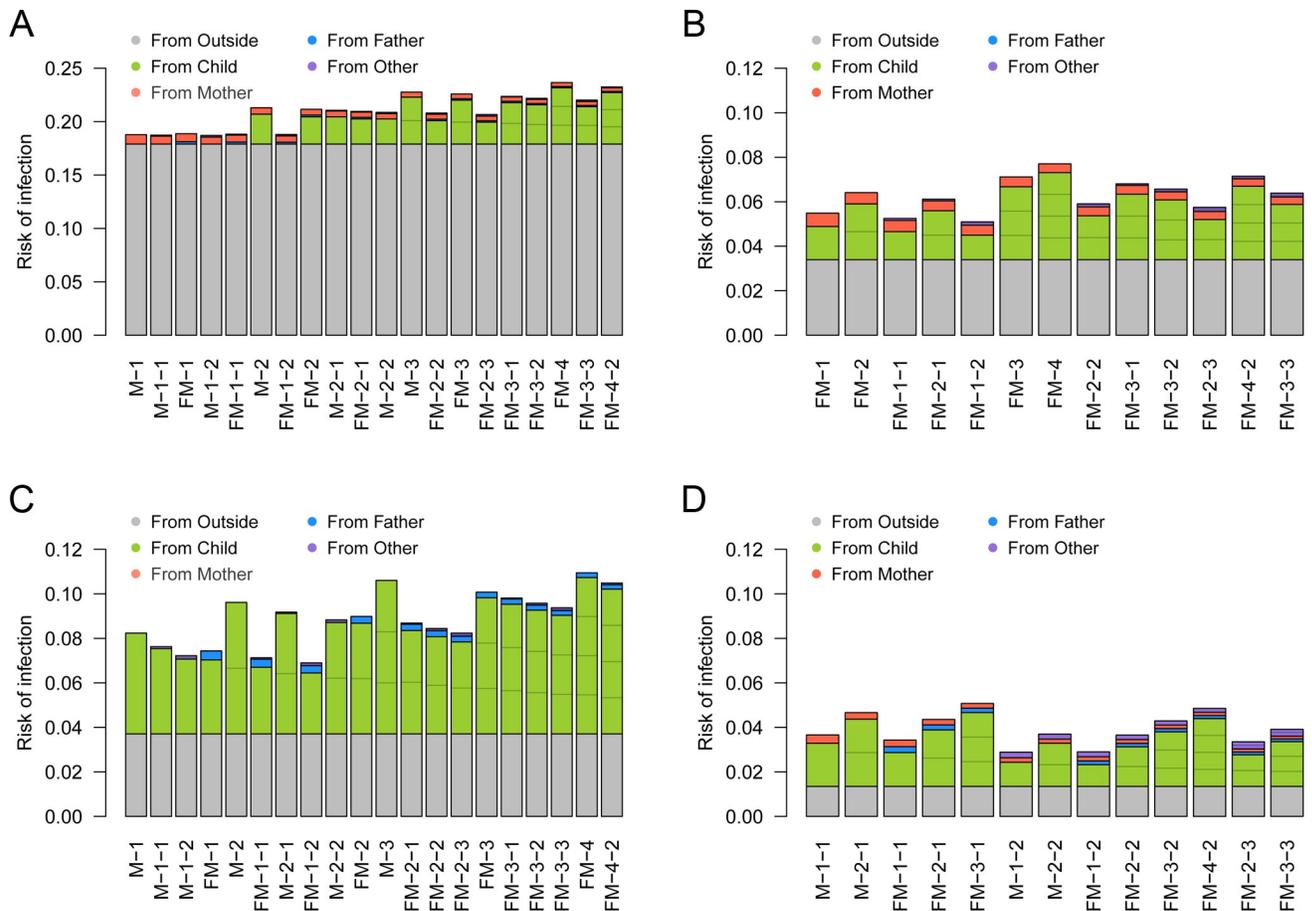
Fig 5. The risk of primary/secondary infection to individuals in different household compositions and its source. (A) Child; (B) Father; (C) Mother; (D) Other. The coloured compartments denote the breakdown of sources. Household compositions are displayed in the same order as Fig 4. The risk of primary infection in children was set to be 16.4%, the average between those of "students" and "siblings". Note that the scale of the y-axis in (A) is different from the other three panels.

responsible for the most of secondary transmissions within households (Fig 5): as children were more than five times likely to acquire influenza from outside the household than adults, they were the most likely source of secondary transmission. As a consequence, the individual risk of infection was mostly determined the number of children in the household. A sensitivity analysis suggested that the effective household contacts between children may have been lower than the baseline estimates under some assumptions (Figure S1 in S1 File). However, the overall trend did not change substantially. The importance of children introducing influenza into household remained unchanged throughout the sensitivity analysis. The model prediction was highly consistent with the observed outcome patterns (Figures S2 and S3 in S1 File), suggesting our model could successfully capture the heterogeneous transmission patterns of influenza in households.

Discussion

We applied a household-based mathematical model to a large-scale influenza survey data including 10,000 primary school students and their families in Matsumoto city, Japan, 2014–15. With the dataset of an extensive sample size on morbidity and familial roles of household

members, the model captured heterogeneous transmission patterns in households in greater detail than previous household studies.

Our results are supportive of the common perception that influenza is brought into households by schoolchildren [41]. With their high probability of contracting influenza outside the household, they were responsible for most secondary transmissions within households. Once they brought the virus from outside the household, their mother and other siblings were exposed to a higher risk of within-household secondary transmission. The estimated breakdown of infection source showed that within-household transmission accounted for a large proportion of the overall risk in adults. The relative importance of within-household transmission was especially highlighted in grandparents in "three-generation" households. In a typical three-generation family composed of two children, two parents and two grandparents, the risk of infection in grandparents was tripled by within-household transmission. Besides, it must be noted that infection of a grandparent is likely to be followed by that of another due to a high transmission risk between grandparents. These emphasise the importance of controlling school epidemic and household contagion, as the symptoms of influenza tend to be more severe in the elderly [41–43].

The results of the present study could have implications for household-level control measures. There are two steps in a household outbreak: introduction and within-household transmission. Due to the different risk patterns between the two steps, the focus of prevention measures should also change accordingly. At the pre-introduction stage when no one in the household is yet infected with influenza, the primary target is to prevent the first infection in the household from happening. Children, with the risk of external infection up to 20%, are most likely to be the first case in the household and thus should be prioritised at this stage. As the high risk of external infection is probably from schools [3], household members are advised to monitor the trend of school outbreaks and guide children to comply with daily precautions [44,45]. Our results suggest that vaccinating children is an effective strategy not only because their risk of infection is high but also because they are responsible for a substantial fraction of within-household secondary infections. Especially for adults living with many children, protecting children from infection is as important as (or even more important in some cases) protecting themselves. If one of the household members contracts influenza despite the pre-introduction control effort, the primary target shifts to preventing further transmissions within the household. Household members are now exposed to an infectious person within the same household, which substantially elevates their risk. At this post-introduction stage, preventing subsequent transmissions is important because every additional infection further increases the exposure. Our findings on household transmission patterns can be used to identify key individuals in the household network. For example, if the primary case is a child, the most probable secondary case is either the mother or another sibling. If the mother gets infected, that may be followed by transmission to either the father or another child. Direct transmissions between children and father/grandparent may be relatively rare. Grandparents are suggested to be at lower risk of infection from other household members. However, their contacts with each other are closer than any other pair of household members, which warrants attention provided the high disease burden of influenza in the elderly.

To our best knowledge, the present study first reported a parametric relationship between within-household influenza transmission and household composition with high precision. With a detailed dataset consisting of up to 10,000 households, the present study was able to employ a highly flexible modelling framework to explore previously used modelling assumptions in great detail. A decrease of the per-person risk of within-household infections with household size has been observed in previous studies [9]; our model selection supported that this reduced effect of household contact is better characterised as a function of the total

39

amount of contact experienced by an individual (C_k) rather than the household size (N) and that the relationship follows an inverse square root law. Previous modelling studies used different frameworks to study the relationship between SITP and household composition. Cauchemez et al. (2004) and (2014) [15,46] selected the frequency-dependent mixing assumption (SITP inversely proportional to N) over the density-dependent mixing (SITP independent of N). Many similar studies were also supportive of the frequency-dependent mixing assumption [14,19,23], while Azman et al. (2013) reported an increased transmission rate in larger households (SITP proportional to $N^{0.7}$; although not conclusive due to the limited sample size). One of the strengths of our results is that not only did we propose a better alternative measure to scale SITP than household size, we also differentiated the model from both density- and frequency-dependent models with sufficient support. The best model suggested that within-household transmission patterns lie half-way between the two extremes of density- and frequency-dependent models (we call this the semi-density-dependent model as the total effective contact experienced by an individual is proportional to the total contact intensity to the power of 0.5). Although a similar approach (without incorporating heterogeneous contact patterns) was employed in [19], where the authors estimated the STIP proportional to $1/N^{1.2}$, their CrI was too wide (0.13–2.3) to be conclusive. The large-scale dataset enabled us to obtain a narrower CrI (0.30–0.72) that distinguished the model with significance from the density- and frequency-dependent models. In the semi-density-dependent model, the total amount of effective contact increases in larger household despite the reduced importance of each contact (Fig 4). Therefore, if the risk of external infection is similar between household members, having many household members is a risk factor (which is not usually the case in the frequency-dependent model) because the effect of reduced SITP is outweighed by the increased number of household members who potentially bring infection into the household. Although such effect was not clearly visible in the present study due to the almost exclusive primary infections in children (Fig 5), more distinct characteristics may be seen in other epidemic settings with the semi-density dependent model.

Multiple limitations in the present study must be acknowledged. Firstly, the case definition in the dataset was not very strict. The data was collected by self-written questionnaires and it was impossible to validate their response. In the dataset, all student cases were reported to be with a clinical diagnosis, and more than 95% of diagnoses were based on rapid diagnostic tests [30]. Considering that primary school students in Japan are highly motivated to visit medical institutions to obtain a leave of absence from school, we believe that our data was able to capture influenza incidence in primary schools at high accuracy. However, it is not clear if the same applies to their household members; diagnoses were not explicitly required for household members on the question sheet, although the term "influenza" rather than "influenza-like illness" was used. Moreover, subclinical infections were probably present both in children and adults. Because of this, we considered underreporting in the sensitivity analysis, leaving the main conclusions unaltered. Secondly, our model formulation is only one possible candidate for parameterising within-household transmission patterns. "Contact" in our model was merely a hypothetical quantity and may not be directly related to actual physical or social contacts. We also had to use a relatively simple contact pattern matrix for successful parameter estimation. Although our model successfully explained the current data incorporating in an interpretable manner, future development may include theoretical frameworks that can explain empirical household contact patterns. A recent study have suggested the possible age-dependency in the contact frequency between siblings [7], but the age of household members were not available in the current dataset. More informative dataset and understanding of age-dependent household contact patterns will yield further clarification on this point. Furthermore, one must be aware that our analysis based on a unique study population, i.e., households

40

with at least one primary school student in Matsumoto city, may not be overgeneralized. Extrapolating our household transmission model to household compositions not included in the dataset, e.g., households with no children, may be unreliable. Thirdly, the present study radically simplified the risk factors of individuals. Covariates other than familial roles and household compositions, e.g., comorbidities, vaccination history, previous exposures or habits of personal hygiene, were not considered. The risk of external infection in children was estimated as a single value, which may potentially vary between classes, grades and schools. Overdispersion in infectiousness as addressed in [14,47,48] was also assumed to be negligible. Nonetheless, it is of note that the model had a fairly good performance despite considerable simplification.

Although more follow-up studies that supplement our findings are to be awaited, we believe that the present study has presented useful insights on the household-level dynamics of influenza. Understanding of the household-specific contact patterns will help us illustrate how influenza spreads across multiple social settings and facilitate individual and political decisions on disease control accounting for household-specific characteristics.

Supporting information

S1 File. Supplementary materials.
(PDF)

Author Contributions

Conceptualization: Akira Endo.

Data curation: Mitsuo Uchida.

Formal analysis: Akira Endo.

Methodology: Akira Endo, Adam J. Kucharski, Sebastian Funk.

Software: Akira Endo.

Supervision: Adam J. Kucharski, Sebastian Funk.

Visualization: Akira Endo.

Writing – original draft: Akira Endo.

Writing – review & editing: Mitsuo Uchida, Adam J. Kucharski, Sebastian Funk.

References

1. le Polain de Waroux O, Flasche S, Kucharski AJ, Langendorf C, Ndazima D, Mwanga-Amumpaire J, et al. Identifying human encounters that shape the transmission of Streptococcus pneumoniae and other acute respiratory infections. Epidemics. 2018. https://doi.org/10.1016/j.epidem.2018.05.008 PMID: 30054196

2. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. PLoS One. 2010. https://doi.org/10.1371/journal.pone.0012948 PMID: 20856792

3. Ackerman E, Elveback LR. Simulation of infectious disease epidemics. Springfield, Ill.: C. C. Thomas; 1984.

4. Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature. 2006. https://doi.org/10.1038/nature04795 PMID: 16642006

5. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008. https://doi.org/10.1371/journal.pmed.0050074 PMID: 18366252

41

6. Ibuka Y, Ohkusa Y, Sugawara T, Chapman GB, Yamin D, Atkins KE, et al. Social contacts, vaccination decisions and influenza in Japan. J Epidemiol Community Health. 2016. https://doi.org/10.1136/jech-2015-205777 PMID: 26424846

7. Goeyvaerts N, Santermans E, Potter G, Torneri A, Kerckhove K Van, Willem L, et al. Household members do not contact each other at random: implications for infectious disease modelling. Proc R Soc B Biol Sci. 2018; 285: 20182201. https://doi.org/10.1098/rspb.2018.2201 PMID: 30963910

8. Ozella L, Gesualdo F, Tizzoni M, Rizzo C, Pandolfi E, Campagna I, et al. Close encounters between infants and household members measured through wearable proximity sensors. PLoS One. 2018. https://doi.org/10.1371/journal.pone.0198733 PMID: 29879196

9. Tsang TK, Lau LLH, Cauchemez S, Cowling BJ. Household Transmission of Influenza Virus. Trends Microbiol. 2016; 24: 123–133. https://doi.org/10.1016/j.tim.2015.10.012 PMID: 26612500

10. Lau LLH, Nishiura H, Kelly H, Ip DKM, Leung GM, Cowling BJ. Household transmission of 2009 pandemic influenza A (H1N1): a systematic review and meta-analysis. Epidemiology. 2012. https://doi.org/10.1097/EDE.0b013e31825588b8 PMID: 22561117

11. Longini IM, Koopman JS. Household and community transmission parameters from final distributions of infections in households. Biometrics. 1982; 38: 115–126. PMID: 7082755

12. Becker NG, Britton T. Statistical studies of infectious disease incidence. J R Stat Soc B. 1999; 61: 287–307. https://doi.org/10.1111/1467-9868.00177

13. Ball F, Neal P. A general model for stochastic SIR epidemics with two levels of mixing. Math Biosci. 2002; 180: 73–102. https://doi.org/10.1016/s0025-5564(02)00125-6 PMID: 12387917

14. House T, Inglis N, Ross J V., Wilson F, Suleman S, Edeghere O, et al. Estimation of outbreak severity and transmissibility: Influenza A(H1N1)pdm09 in households. BMC Med. 2012; 10: 117. https://doi.org/10.1186/1741-7015-10-117 PMID: 23046520

15. Cauchemez S, Ferguson NM, Fox A, Mai LQ, Thanh LT, Thai PQ, et al. Determinants of Influenza Transmission in South East Asia: Insights from a Household Cohort Study in Vietnam. PLoS Pathog. 2014; 10: 2–9. https://doi.org/10.1371/journal.ppat.1004310 PMID: 25144780

16. O'Neill PD, Balding DJ, Becker NG, Eerola M, Mollison D. Analyses of infectious disease data from household outbreaks by Markov chain Monte Carlo methods. J R Stat Soc Ser C-Applied Stat. 2000; 49: 517–542. https://doi.org/10.1111/1467-9876.00210

17. Wardell R, Prem K, Cowling BJ, Cook AR. The role of symptomatic presentation in influenza A transmission risk. Epidemiol Infect. 2017. https://doi.org/10.1017/S0950268816002740 PMID: 27916020

18. Azman AS, Stark JH, Althouse BM, Vukotich CJ, Stebbins S, Burke DS, et al. Household transmission of influenza A and B in a school-based study of non-pharmaceutical interventions. Epidemics. 2013. https://doi.org/10.1016/j.epidem.2013.09.001 PMID: 24267874

19. Cauchemez S, Bhattarai A, Marchbanks TL, Fagan RP, Ostroff S, Ferguson NM, et al. Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. Proc Natl Acad Sci. 2011; 108: 2825–2830. https://doi.org/10.1073/pnas.1008895108 PMID: 21282645

20. Cauchemez S, Donnelly CA, Reed C, Ghani AC, Fraser C, Kent CK, et al. Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. N Engl J Med. 2009. https://doi.org/10.1056/NEJMoa0905498 PMID: 20042753

21. Tsang TK, Cauchemez S, Perera RAPM, Freeman G, Fang VJ, Ip DKM, et al. Association between antibody titers and protection against influenza virus infection within households. J Infect Dis. 2014. https://doi.org/10.1093/infdis/jiu186 PMID: 24676208

22. Addy CL, Longini IM, Haber M. A generalized stochastic model for the analysis of infectious disease final size data. Biometrics. 1991.

23. Cauchemez S, Ferguson NM, Wachtel C, Tegnell A, Saour G, Duncan B, et al. Closure of schools during an influenza pandemic. The Lancet Infectious Diseases. 2009. https://doi.org/10.1016/S1473-3099(09)70176-8

24. Thai PQ, Mai LQ, Welkers MRA, Hang NLK, Thanh LT, Dung VTV, et al. Pandemic H1N1 virus transmission and shedding dynamics in index case households of a prospective Vietnamese cohort. J Infect. 2014. https://doi.org/10.1016/j.jinf.2014.01.008 PMID: 24491598

25. Wu JT, Riley S, Fraser C, Leung GM. Reducing the impact of the next influenza pandemic using household-based public health interventions. PLoS Med. 2006. https://doi.org/10.1371/journal.pmed.0030361 PMID: 16881729

26. Budge PJ, Griffin MR, Edwards KM, Williams J V., Verastegui H, Hartinger SM, et al. Impact of home environment interventions on the risk of influenza-associated ARI in Andean Children: Observations from a prospective household-based cohort study. PLoS One. 2014. https://doi.org/10.1371/journal.pone.0091247 PMID: 24622044

42

27. Yeung MPS, Lam FLY, Coker R. Factors associated with the uptake of seasonal influenza vaccination in adults: A systematic review. Journal of Public Health (United Kingdom). 2016. https://doi.org/10.1093/pubmed/fdv194 PMID: 28158550

28. Wu S, Su J, Yang P, Zhang H, Li H, Chu Y, et al. Factors associated with the uptake of seasonal influenza vaccination in older and younger adults: A large, population-based survey in Beijing, China. BMJ Open. 2017. https://doi.org/10.1136/bmjopen-2017-017459 PMID: 28951412

29. Uchida M, Kaneko M, Hidaka Y, Yamamoto H, Honda T, Takeuchi S, et al. Effectiveness of vaccination and wearing masks on seasonal influenza in Matsumoto City, Japan, in the 2014/2015 season: An observational study among all elementary schoolchildren. Prev Med Reports. 2017; 5: 86–91. https://doi.org/10.1016/j.pmedr.2016.12.002 PMID: 27981021

30. Uchida M, Kaneko M, Hidaka Y, Yamamoto H, Honda T, Takeuchi S, et al. Prospective epidemiological evaluation of seasonal influenza in all elementary schoolchildren in Matsumoto city, Japan, in 2014/2015. Jpn J Infect Dis. 2017. https://doi.org/10.7883/yoken.JJID.2016.037 PMID: 27580571

31. Analysis of influenza virus isolates from the 2014/15 influenza season, Japan. Infectious Agents Surveillance Report, National Institute of Infectious Diseases and Tuberculosis and Infectious Diseases Control Division. 2015. pp. 202–207. Available: https://www0.niid.go.jp/niid/idsc/iasr/36/429e.pdf

32. Chartrand C, Leeflang MMG, Minion J, Brewer T, Pai M. Accuracy of rapid influenza diagnostic tests: A meta-analysis. Annals of Internal Medicine. 2012. https://doi.org/10.7326/0003-4819-156-7-201204030-00403 PMID: 22371850

33. Bruning AHL, Leeflang MMG, Vos JMBW, Spijker R, de Jong MD, Wolthers KC, et al. Rapid Tests for Influenza, Respiratory Syncytial Virus, and Other Respiratory Viruses: A Systematic Review and Meta-analysis. Clin Infect Dis. 2017. https://doi.org/10.1093/cid/cix461 PMID: 28520858

34. Yamazaki M, Mitamura K, Ichikawa M, Kimura K, Komiyama O, Shimizu H, et al. [Evaluation of flow-through immunoassay for rapid detection of influenza A and B viruses]. Kansenshogaku Zasshi. 2004. https://doi.org/10.11150/kansenshogakuzasshi1970.78.865 PMID: 15508721

35. Hara M, Takao S, Fukuda S, Shimazu Y, Miyazaki K. [Comparison of three rapid diagnostic kits using immunochromatography for detection of influenza A virsuses]. Kansenshogaku Zasshi. 2004. https://doi.org/10.11150/kansenshogakuzasshi1970.78.935 PMID: 15628525

36. Hara M, Sadamasu K, Takao S, Shinkai T, Kai A, Fukuda S, et al. [Evaluation of immunochromatography test for rapid detection of influenza A and B viruses using real-time PCR]. Kansenshogaku Zasshi. 2006. https://doi.org/10.11150/kansenshogakuzasshi1970.80.522 PMID: 17073266

37. Longini IM, Koopman JS, Haber M, Cotsonis GA. Statistical inference for infectious diseases: Risk-specific household and community transmission parameters. Am J Epidemiol. 1988. https://doi.org/10.1093/oxfordjournals.aje.a115038 PMID: 3421247

38. Begon M, Bennett M, Bowers RG, French NP, Hazel SM, Turner J. A clarification of transmission terms in host-microparasite models: Numbers, densities and areas. Epidemiol Infect. 2002. https://doi.org/10.1017/S0950268802007148 PMID: 12211582

39. Watanabe S. A Widely Applicable Bayesian Information Criterion. 2013; 14: 867–897. https://doi.org/10.1088/0953-8984/23/18/184115

40. Mizumoto K, Yamamoto T, Nishiura H. Age-dependent estimates of the epidemiological impact of pandemic influenza (H1N1-2009) in Japan. Comput Math Methods Med. 2013. https://doi.org/10.1155/2013/637064 PMID: 23509599

41. Glezen WP. Emerging infections: Pandemic influenza. Epidemiologic Reviews. 1996. https://doi.org/10.1093/oxfordjournals.epirev.a017917 PMID: 8877331

42. Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, et al. Influenza-associated hospitalizations in the United States. J Am Med Assoc. 2004. https://doi.org/10.1001/jama.292.11.1333 PMID: 15367555

43. Schanzer DL, Tam TWS, Langley JM, Winchester BT. Influenza-attributable deaths, Canada 1990–1999. Epidemiol Infect. 2007. https://doi.org/10.1017/S0950268807007923 PMID: 17306052

44. Jefferson T, Foxlee R, Del Mar C, Dooley L, Ferroni E, Hewak B, et al. Physical interventions to interrupt or reduce the spread of respiratory viruses: Systematic review. BMJ. 2008. https://doi.org/10.1136/bmj.39393.510347.BE PMID: 18042961

45. Aiello AE, Murray GF, Perez V, Coulborn RM, Davis BM, Uddin M, et al. Mask Use, Hand Hygiene, and Seasonal Influenza-Like Illness among Young Adults: A Randomized Intervention Trial. J Infect Dis. 2010. https://doi.org/10.1086/650396 PMID: 20088690

46. Cauchemez S, Carrat F, Viboud C, Valleron AJ, Boëlle PY. A Bayesian MCMC approach to study transmission of influenza: Application to household longitudinal data. Stat Med. 2004. https://doi.org/10.1002/sim.1912 PMID: 15505892

43

47. Ball F. A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. Adv Appl Probab. 1986. https://doi.org/10.2307/1427301

48. Van Boven M, Koopmans M, Van Beest Holle MDR, Meijer A, Klinkenberg D, Donnelly CA, et al. Detecting emerging transmissibility of avian influenza virus in human households. PLoS Comput Biol. 2007; 3: 1394–1402. https://doi.org/10.1371/journal.pcbi.0030145 PMID: 17676981

Supplementary materials

Fine-scale family structure shapes influenza transmission risk in households: insights from primary schools in Matsumoto city, 2014/15.

Akira Endo, Mitsuo Uchida, Adam Kucharski, Sebastian Funk

1.   Model selection

Model selection on the complexity

In the first round of model selection, we compared models with different complexity. Our household transmission model was mainly characterised by two components, the effective household contact $\eta_{kl} = \beta \frac{c_{kl}}{C_k{}^\gamma}$ and the risk of external infection ε_k. Models corresponding to all possible combinations of assumptions were compared based on the Widely-applicable Bayesian information criterion (WBIC) (1). WBIC has the same scale as the Bayesian information criterion. A difference of 2 in WBIC is considered as an indication of statistical significance, while a difference greater than 5 is deemed as strong support. Table S1 compares the candidate models and their WBIC.

The parameters c_{kl} and ε_k were estimated as a single value $c_{kl} = c$ and $\varepsilon_k = \varepsilon$ under "Homogeneous"/"Uniform" assumptions, respectively. We fixed γ at 0 in "DD" (density-dependent) models and 1 in "FD" (frequency-dependent), and freely estimated in "IM" (intermediate) models. In "Single parent-Y" models, fathers and mothers who do not live with a spouse were classified as an additional type "single parent" (thus the number of types was 6 in these models). The best model (Model 12) was selected with very strong support: ΔWBIC from the second-best model was 16.9.

Model selection on the contact pattern matrix

After selecting the model complexity, we further tried to explore different contact pattern matrices c_{kl}. Let the rows and columns of c_{kl} correspond to (Student, Sibling, Father, Mother, Other). Five parameters $(c_{CC}, c_{FC}, c_{MC}, c_{OC}, c_{AA})$ being denoted by numbers 1 to 5, the contact pattern matrix c_{kl} in the previous model selection had the following structure:

$$c_{kl} = \begin{bmatrix} 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 & 4 \\ 2 & 2 & 5 & 5 & 5 \\ 3 & 3 & 5 & 5 & 5 \\ 4 & 4 & 5 & 5 & 5 \end{bmatrix} \quad\quad (S1)$$

Note that the diagonal elements for student, father and mother were displayed only for completeness and not used in the analysis (households in our dataset did not contain more than one students/fathers/mothers). Parameter estimates in Model 12 are shown in Table S2. In this contact pattern matrix, as all adults are assumed to share the same contact intensity. Meanwhile, the estimates of c_{FC} and c_{OC} are relatively similar. We explored variant models that further stratify c_{AA} while

45

c_{FC} and c_{OC} are equated to keep the number of parameters unchanged (=5).

We considered the following submodels: Model 12a (intense contact within couples), Model 12b (mother acting as a hub) and Model 12c (generation-assortative).

$$
c_{kl} \text{ (Model 12a)} = \begin{bmatrix} 1 & 1 & 2 & 3 & 2 \\ 1 & 1 & 2 & 3 & 2 \\ 2 & 2 & 4 & 4 & 5 \\ 3 & 3 & 4 & 4 & 5 \\ 2 & 2 & 5 & 5 & 5 \end{bmatrix},
$$

$$
c_{kl} \text{(Model 12b)} = \begin{bmatrix} 1 & 1 & 2 & 3 & 2 \\ 1 & 1 & 2 & 3 & 2 \\ 2 & 2 & 5 & 4 & 5 \\ 3 & 3 & 4 & 4 & 4 \\ 2 & 2 & 5 & 4 & 5 \end{bmatrix}, \qquad\qquad (S2)
$$

$$
c_{kl} \text{(Model 12c)} = \begin{bmatrix} 1 & 1 & 5 & 3 & 5 \\ 1 & 1 & 5 & 3 & 5 \\ 5 & 5 & 2 & 2 & 5 \\ 3 & 3 & 2 & 2 & 5 \\ 5 & 5 & 5 & 5 & 4 \end{bmatrix}
$$

Estimated contact pattern matrices are shown in Tables S3-S5. Models 12a and 12c had much better WBIC than Model 12 (ΔWBIC = -14.4 and ΔWBIC = -17.3, respectively), while that of Model 12b was slightly worse than Model 12. Of the two models exhibiting improved WBICs, Model 12c was selected with a significant WBIC difference of 2.9. Parameter estimates other than c_{kl} did not vary between compared models to the first significant figure.

Selection of the scaling factor

In our baseline model, the total amount of contacts $C_k = \sum_l c_{kl}$ was used to scale the effective household contact (i.e., $\eta_{kl} \propto C_k^{-\gamma}$) to reflect heterogeneous contact patterns. On the other hand, previous modelling studies often used household size N in place of C_k (2–5). Although C_k and N are correlated (C_k and N-1 coincide in homogeneous settings) and may work as a good proxy with each other, we considered comparison between these two approaches to be of interest. We tested a variant of Model 12c where C_k is replaced with N (i.e., $\eta_{kl} \propto N^{-\gamma}$), but the model performance was significantly worsened (ΔWBIC = 8.8). The estimated value of gamma did not change (γ = 0.52; CrI: 0.34-0.75). The use of the total amount of contacts is preferred to household size as a scaling factor for the within-household transmission, and even when household size is used as variable, the semi-density-dependent model may still be applicable.

Table S1. WBIC of models with different sets of assumptions.

Model ID	c_{kl}	ε_k	γ	Single parent	WBIC	ΔWBIC
1	Hom	Unif	DD	N	33269.16	2134.96
2	Het	Unif	DD	N	33054.70	1920.50
3	Hom	Unif	FD	N	33259.36	2125.16
4	Het	Unif	FD	N	32731.10	1596.90
5	Hom	Unif	IM	N	33243.32	2109.12
6	Het	Unif	IM	N	32277.92	1143.72
7	Hom	Strat	DD	N	31215.16	80.96
8	Het	Strat	DD	N	31151.08	16.88
9	Hom	Strat	FD	N	31205.44	71.24
10	Het	Strat	FD	N	31150.78	16.58
11	Hom	Strat	IM	N	31186.64	52.44
12	Het	Strat	IM	N	31134.20	0
13	Hom	Unif	DD	Y	33267.72	2133.52
14	Het	Unif	DD	Y	33061.14	1926.94
15	Hom	Unif	FD	Y	33256.72	2122.52
16	Het	Unif	FD	Y	32752.26	1618.06
17	Hom	Unif	IM	Y	33241.46	2107.26
18	Het	Unif	IM	Y	32182.18	1047.98
19	Hom	Strat	DD	Y	31223.68	89.48
20	Het	Strat	DD	Y	31167.16	32.96
21	Hom	Strat	FD	Y	31212.52	78.32
22	Het	Strat	FD	Y	31168.08	33.88
23	Hom	Strat	IM	Y	31194.82	60.62
24	Het	Strat	IM	Y	31151.98	17.78

Hom: homogeneous mixing, Het: Heterogeneous mixing

Unif: uniform risk of external infection, Strat: stratified risk of external infection

DD: density-dependent, FD: frequency-dependent, IM: intermediate

Single-Parent: whether the "single parent" category has a unique parameter. Y=Yes, N=No.

WBIC: Widely-applicable Bayesian information criterion, ΔWBIC: WBIC difference from the best model

Table S2. Estimated contact pattern matrix (c_{kl}) in Model 12.

	Student	Sibling	Father	Mother	Other
Student	1.28		0.54	1.40	0.45
Sibling					
Father	0.54		1		
Mother	1.40				
Other	0.45				

WBIC = 31134.20; ΔWBIC = 0 (baseline)

Table S3. Estimated contact pattern matrix (c_{kl}) in Model 12a.

	Student	Sibling	Father	Mother	Other
Student	0.97		0.39	1.09	0.39
Sibling					
Father	0.39		1		0.39
Mother	1.09				
Other	0.39		0.39		1

WBIC = 31119.78; ΔWBIC = -14.42

Table S4. Estimated contact pattern matrix (c_{kl}) in Model 12b.

	Student	Sibling	Father	Mother	Other
Student	1.25		0.49	1.37	0.49
Sibling					
Father	0.49		1.01		1.01
Mother	1.37			1	
Other	0.49		1.01		1.01

WBIC = 31134.72; ΔWBIC = 0.54

Table S5. Estimated contact pattern matrix (c_{kl}) in Model 12c.

	Student	Sibling	Father	Mother	Other
Student	1.04		0.43	1.16	0.43
Sibling					
Father	0.43		1		
Mother	1.16				
Other	0.43				1.97

WBIC = 31116.88; ΔWBIC = -17.32 (best model)

2. Source-stratified risk of infection and risk attributable to the introduction of influenza into a household

We quantified the risk of infection attributable to external and within-household infection from the parameter estimates. Three family compositions were selected as model cases: (a) "nuclear family": father, mother and two children, (b) "many-siblings family": father, mother and four children, and (c) "three-generation family": father, mother, two children and two grandparents. We assumed that one of the children in each model case households was "student", and the others were "siblings". The overall risk of infection for type k individual is given by

$$r_k = \sum_{\boldsymbol{n}} \frac{n_k}{N_k} \pi(\boldsymbol{n}; \boldsymbol{N}, \boldsymbol{\varepsilon}, H), \tag{S3}$$

(the sum is taken for all possible $\boldsymbol{n}$), and $r_k - \varepsilon_k$ corresponds to the additional infection risk due to within-household transmission.

We also compared the risk of infection after the introduction of influenza into households with the initial overall risk. We defined post-introduction risk as the conditional probability that an individual experience infection by the end of the season, given that one index case is already observed in the same family. Here, for simplicity, we limited the analysis to introductions by primary school students (i.e., individual type "student") only.

Suppose that $k=1$ corresponds to the type "student". Post-introduction risk obtained by modifying the formula for r_k as

$$r_k^{\text{pos}} = \sum_{\{\boldsymbol{n}|n_1=0\}} \frac{n_k}{N_k} \cdot \frac{\pi(\boldsymbol{n}; \boldsymbol{N}, \boldsymbol{\varepsilon} + H_1, H)}{S_1(\boldsymbol{n}, \boldsymbol{\varepsilon})}. \tag{S4}$$

H_1 is the (additional) force of infection arising from the infected student, i.e., $(H_1)_k = \frac{c_{k1}}{c_k{}^{\gamma}}$. The sum is taken for all possible $\boldsymbol{n}$ whose first component $n_1 = 0$ (because the force of infection from the student is incorporated in H_1). Note that by dividing $h(\boldsymbol{n}; \boldsymbol{N}, \boldsymbol{\varepsilon} + H_1, H)$ by $S_1(\boldsymbol{n}, \boldsymbol{\varepsilon})$, we can yield the conditional probability that $\boldsymbol{n}$ individuals (other than the "student") are infected given the presence of the force of infection H_1.

3. Sensitivity analysis

Procedures for the sensitivity analysis

(i) Ascertainment bias

In order to account for potential ascertainment bias, we incorporated reporting probabilities into the model. We assumed that infections are reported with a certain probability p_k. Epidemiological properties, such as infectiousness, were assumed to be identical between reported and unreported cases. The likelihood of observing a household final size outcome $(\boldsymbol{n}; \boldsymbol{N})$ given the reporting probability vector $\boldsymbol{p}$ is obtained by using the binomial distribution:

$$L(\boldsymbol{\varepsilon}, H, \boldsymbol{p}; (\boldsymbol{n}; \boldsymbol{N})) = \sum_{\boldsymbol{n'} \geq \boldsymbol{n}} h(\boldsymbol{n'}; \boldsymbol{N}, \boldsymbol{\varepsilon}, H) \prod_k \mathrm{Bin}(n_k; n'_k, p_k). \tag{S5}$$

The sum $\sum_{\boldsymbol{n'} \geq \boldsymbol{n}}$ is taken for all vector $\boldsymbol{n'}$ satisfying $n_k \leq n_k' \leq N_k$ ($\forall k$).

In this sensitivity analysis, we assumed that the reporting probability p for children ("student" and "sibling") is 0.8. The reporting probability for adults was varied from 0.5 to 0.8.

(ii) Different susceptibility in children

Susceptibility to influenza infection was differentiated between children and adults. Let σ be the susceptibility of children relative to that of adults. The effect of σ was employed in the model by differentiating the transmissibility β as

$$\beta_k = \begin{cases} \beta\sigma & (k = \text{"Student", "Sibling"}) \\ \beta & \text{(otherwise)} \end{cases}. \tag{S6}$$

Five different values of σ: 0.75, 1.25, 1.5, 1.75, 2.0 were tested. $\sigma = 0.75$ corresponds to the assumption that children may have less risk of infection per exposure (e.g., due to potentially high vaccination coverage). The value of σ greater than 1 reflects the assumption that children are more vulnerable than adults.

(iii) Multiple counting of households

We identified all the possible combinations of respondents who might be from the same household by the following process. (1) Respondents were classified by their school and family composition. We assumed that siblings usually go to the same primary school. (2) Data were matched up, and consistency was checked between the sex and grade of the respondent and the reported composition of siblings. For instance, a second-grade boy who has no older brother should not be from the same household as a fourth-grade girl who has an older brother (the girl's older brother should also be the boy's older brother). Here, we assumed that siblings should be in different grades, and neglected the possibility of twins or siblings in the same grade. (3) Individuals potentially from the same household were grouped together. Combination of grouping was chosen so that as many individuals as possible are grouped together in total. Individuals in each matched group were assumed to be from the same household, and their data were integrated to represent one household data. Respondents were classified as "students", and siblings who were not found in the dataset was classified as "siblings". Because sex, school and grade were not used in the parameter estimation, individual-level details of the grouping arrangement (who was grouped with whom) did not affect the subsequent analysis. Through this whole process, 1,294 individuals identified as candidates potentially from multiple-counted households were processed, reducing the number of households from 10,486 to 9,763 (-6.9%). Note that this is an extreme case where as many consistent siblings as possible are grouped together, and that the reality may lie between the two extremes (no-grouping and maximum-grouping).

50

(iv) Case censoring

Participants reported in the survey the total number of siblings, siblings in four categories (older brother, older sister, younger brother or younger sister) and whether siblings in each category had influenza. Let $\boldsymbol{M} = (M_1, M_2, M_3, M_4)$ and $\boldsymbol{m} = (m_1, m_2, m_3, m_4)$ be the true composition of siblings and the number of siblings with an influenza episode in each category (1: older brother; 2: older sister, 3: younger brother; 4: younger sister), respectively. Due to the questions in the survey, the dataset did not include either $\boldsymbol{M}$ or $\boldsymbol{m}$. Instead, we have $M = \sum_i M_i$ and censored sibling data $\boldsymbol{M}'$ ($M_i' = \min(M_i, 1)$) and $\boldsymbol{m}'$ ($m_i' = \min(m_i, 1)$), as the questions on sibling categories were yes-no questions.

We constructed a modified likelihood function to address this censoring issue. The basic idea was to generate all possible patterns of $\boldsymbol{M}$ and $\boldsymbol{m}$ that are consistent with the observation and aggregate the corresponding probabilities to obtain the likelihood for the censored data. First, we defined a conditional probability $\pi(\boldsymbol{M}; M)$, the probability that the true sibling composition is $\boldsymbol{M}$ given M. Assuming that the probability of being the n-th child in given M siblings is equally $\frac{1}{M+1}$ and that the sex of a child is evenly distributed, we get

$$\pi(\boldsymbol{M}; M) = \frac{1}{(M+1) \cdot 2^M} \binom{M_1 + M_2}{M_1} \binom{M_3 + M_4}{M_3} \sigma(\boldsymbol{M}, M), \tag{S7}$$

where $\sigma(\boldsymbol{M}, M)$ is an indicator function that takes 1 if $\boldsymbol{M}$ is consistent with M (i.e., $\sum_i M_i = M$), and 0 otherwise.

Let $\pi(\boldsymbol{m}; \boldsymbol{M}, \varphi)$ be the probability of observing a sibling outcome pattern $\boldsymbol{m}$ given $\boldsymbol{M}$. This is also conditional to the existence of other family members and their outcomes, and those conditions are represented by φ.

Using $\pi(\boldsymbol{M}; M)$ and $\pi(\boldsymbol{m}; \boldsymbol{M}, \varphi)$, the likelihood of observing $\{\boldsymbol{M}', \boldsymbol{m}'\}$ given M and φ is

$$l(\boldsymbol{M}', \boldsymbol{m}'; M, \varphi) = \sum_{\boldsymbol{M}} \pi(\boldsymbol{M}; M)\, \sigma(\boldsymbol{M}, \boldsymbol{M}') \sum_{\boldsymbol{m}} \pi(\boldsymbol{m}; \boldsymbol{M}, \varphi)\, \sigma(\boldsymbol{m}, \boldsymbol{m}'), \tag{S8}$$

where $\sigma(\boldsymbol{M}, \boldsymbol{M}')$ and $\sigma(\boldsymbol{m}, \boldsymbol{m}')$ are indicator functions checking if $\boldsymbol{m}$ and $\boldsymbol{M}$ are consistent with the observation.

Since we assume all siblings exhibit identical epidemiological behaviour, considering the effect of loss of distinguishability, $\pi(\boldsymbol{m}; \boldsymbol{M}, \varphi)$ is substituted with

$$\pi(\boldsymbol{m}; \boldsymbol{M}, \varphi) = \frac{\prod_i \binom{M_i}{m_i}}{\binom{M}{m}} \pi(m; M, \varphi), \tag{S9}$$

where $m = \sum_i m_i$. $\pi(m; M, \varphi)$ is equivalent to π in equation (1) in the main text, and thereby we get the likelihood accounting for possible case censoring in siblings.

Results of the sensitivity analysis

The estimates from the sensitivity analysis were compared in Figure S1. Equally lowering the reporting probability for children and adults slightly increased some of the parameters while the overall relative magnitude was almost conserved. When the reporting probability for adults was set lower than children, parameters which involve adults increased and those involving children decreased (Figures S1A and S1B). Increasing the relative susceptibility in children resulted in lower child-involved contact intensities (Figures S1C and S1D). Multiple counting of data reported by students from the same household did not seem to have affected the result, but some changes were caused by addressing censored cases in siblings, which may be resulted from the possibility of unobserved sibling cases (Figures S1E and S1F). Except that the contact intensity between children was substantially lowered by either underreporting of adults or high susceptibility in children, the relative trend remained almost similar throughout our sensitivity analysis. Especially, the risk of external infection in children and the contact intensity between children and adults remained at a sufficient level, such that the secondary transmission from children is still of paramount importance. The exponent parameter γ was stable throughout the sensitivity analysis (median within 0.50 ± 0.02), except that it was slightly higher (0.59; CrI: 0.39-0.79) when the case censoring (iv) was considered.
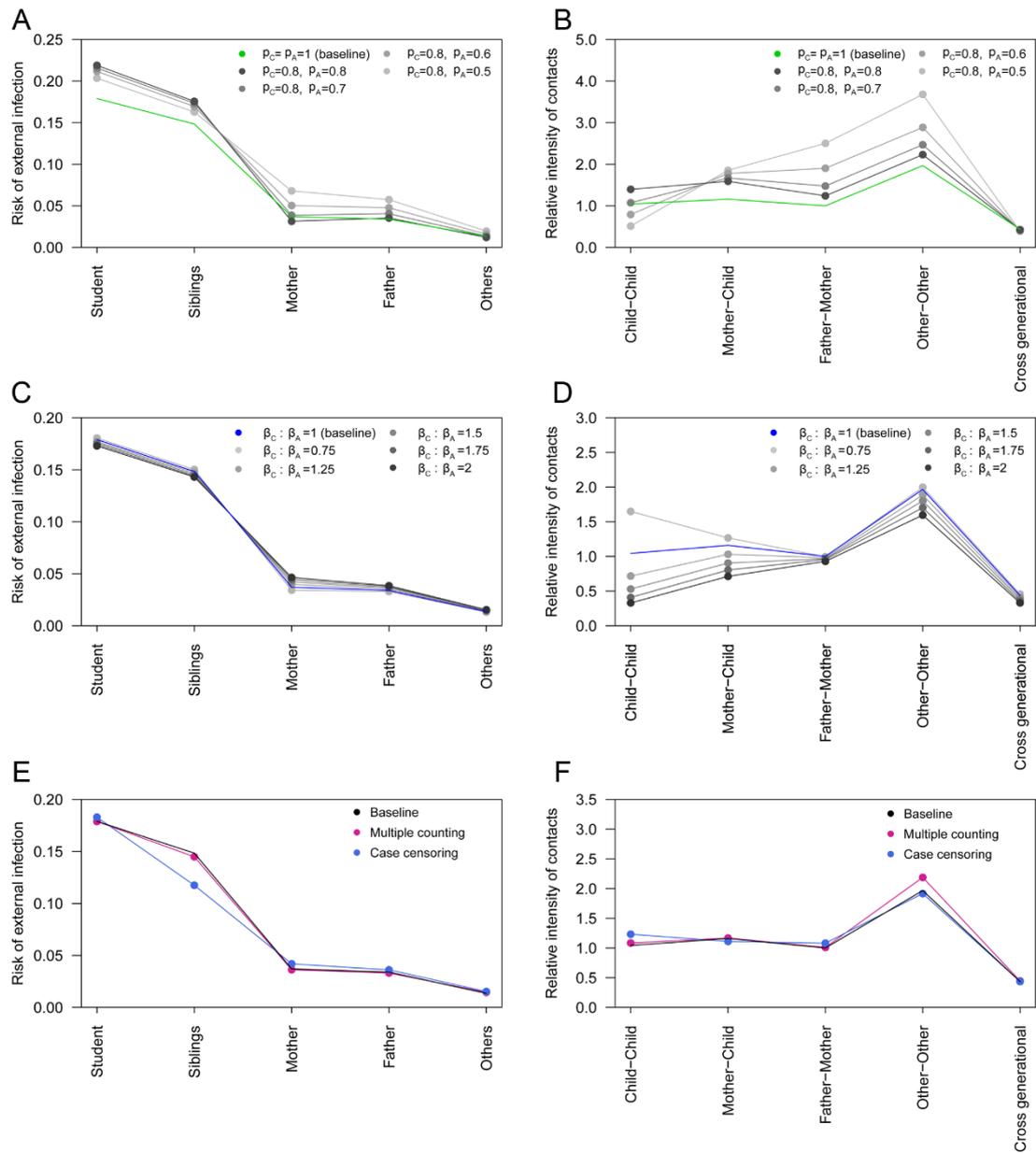
Figure S1. Parameter estimates from the sensitivity analysis. The estimated risk of external infection and relative intensity of household contacts are compared with the baseline estimates. The relative intensity of contacts in the figures is multiplied by the relative change in the estimated transmissibility parameter β for comparability.

(A), (B) Various reporting probabilities in children (p_C) and adults (p_A).

(C), (D) Various ratios between susceptibility in children (β_C) and adults (β_A).

(E), (F) Estimates from the modified dataset addressing multiple counting of households and censoring of cases in siblings.

4. Model fit

To evaluate the goodness-of-fit of our model, the model prediction was compared with the observed data. Let $\hat{\theta}$ be the set of median parameter estimates. $\pi(\boldsymbol{n}; \boldsymbol{N}, \hat{\theta})$, the probability of observing outcome pattern $\boldsymbol{n}$ given household composition $\boldsymbol{N}$, is obtained from Equation (1) in the main text. Assuming that the distribution of $\boldsymbol{N}$ in dataset D is given as observed ($\pi_D(\boldsymbol{N})$), the predictive distribution of the outcome patterns $(\boldsymbol{N}_i, \boldsymbol{n}_i)$ (approximated by the point estimate $\hat{\theta}$) is

$$\pi(\boldsymbol{N}_i, \boldsymbol{n}_i; \hat{\theta}) = \pi(\boldsymbol{n}_i; \boldsymbol{N}_i, \hat{\theta})\pi_D(\boldsymbol{N}_i), \tag{S10}$$

Figure S2 compares the predictive distribution with the actual frequency in the dataset. The 95% intervals are approximated by the 95% quantiles of a binomial distribution

$$F_D(\boldsymbol{N}, \boldsymbol{n}) \sim \text{Binom}\left(F_D(\boldsymbol{N}), \pi(\boldsymbol{n}; \boldsymbol{N}, \hat{\theta})\right), \tag{S11}$$

where F_D is the frequency in data D of size S_D. The predicted and observed frequency show good accordance despite the relatively modest parameter space dimension (=11). The similarity between the two distributions are also supported by the empirical Kullback-Leibler divergence of 0.05, where

$$\widehat{\text{KL}} = \sum_d \frac{F_D(\boldsymbol{N}, \boldsymbol{n})}{S_D} \cdot \log\left(\frac{F_D(\boldsymbol{N}, \boldsymbol{n})}{\pi(\boldsymbol{n}; \boldsymbol{N}, \hat{\theta})F_D(\boldsymbol{N})}\right). \tag{S12}$$
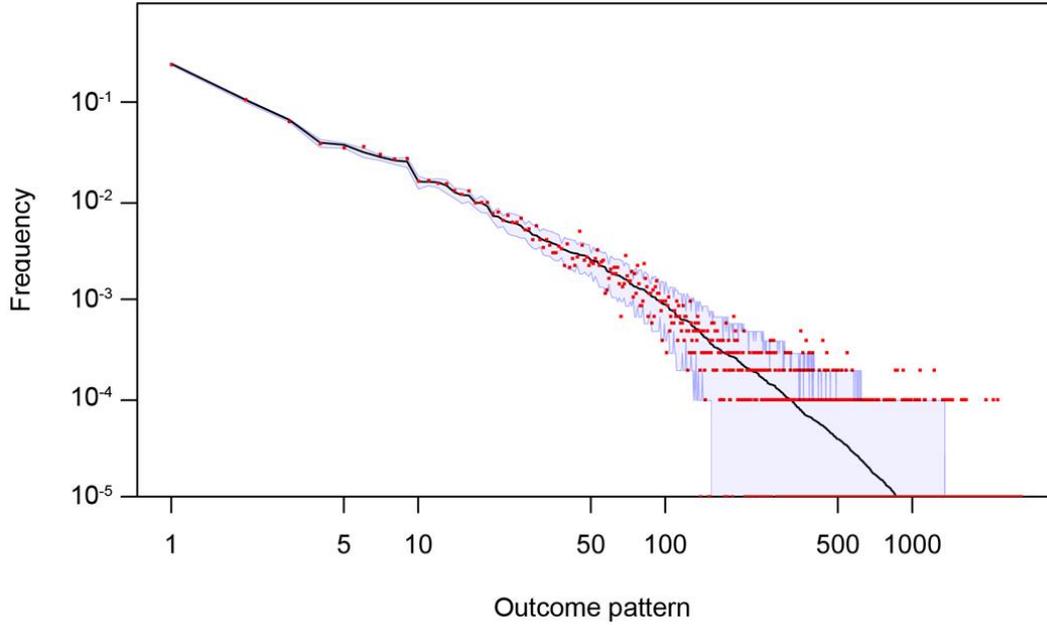

Figure S2. Comparison between the predicted and observed household final outcomes.
Red dots correspond to the observed relative frequency of data (household compositions and final outcomes of the household members), where the x-axis denotes the numbering of outcome patterns $(\boldsymbol{N}, \boldsymbol{n})$. With the sample size of ~10,000, 10^{-4} on the y-axis denotes frequency 1; dots for frequency 0 are shown on the x-axis. The black line indicates the probability of observation predicted by the model, and the shaded area shows 95% intervals. Both x- and y-axes are on a logarithmic scale.

We also compared the predicted and observed distributions of the final attack size (the total number of household cases during the season) for specific compositions in Figure S3. The observed distribution was right-skewed from the "binomial scenario", where within-household transmission is not present and individuals are assumed to be exposed to the external risk of infection only.
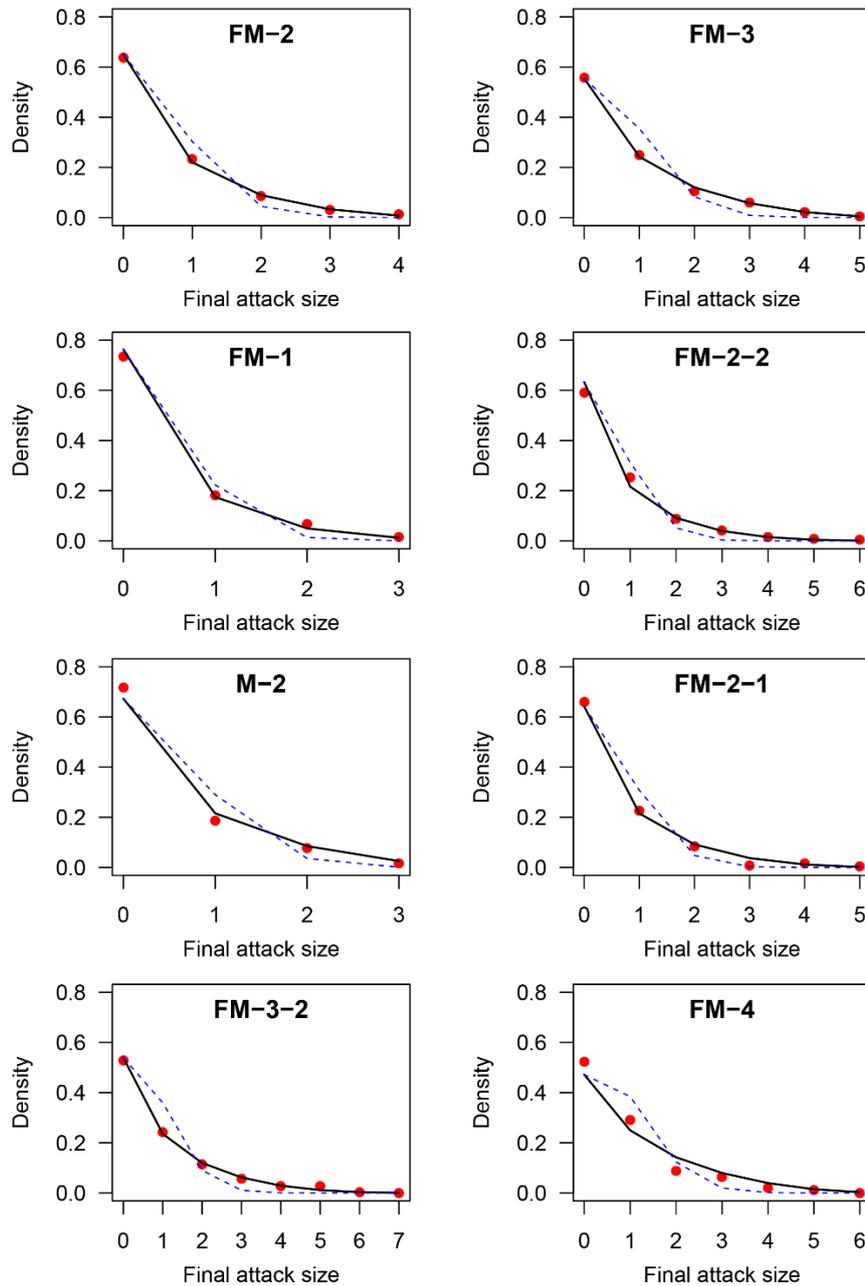

Figure S3. The observed and predicted final attack size distributions.

Red dots and black lines denote the observed and predicted relative frequencies. Blue lines represent "binomial scenario", where within-household transmission is not present. Eight major household compositions (accounting for 84% of the total households in the dataset) are shown.

4.3   Additional notes: techniques used for fast computation of likelihood function

One of the challenges in the estimation process in this study was to minimise computation time for evaluating the likelihood function. The likelihood function for the heterogeneous Longini-Koopman model involves recursive calculation for each household, and this had to be applied to over 10,000 households in the dataset to produce a single total likelihood value. This likelihood function needed to be repeatedly evaluated in MCMC with up to millions of iterations, which was then repeated for model selection. To achieve this within realistic time, the following techniques were used in this study.

4.3.1   Use of compiled programming language

The main code for the analysis was written in R language. However, being an interpreter programming language, the computation speed of R is inferior to compiled language. For faster computation, the internal algorithm for the likelihood function was written in C++ via {Rcpp} and {RcppArmadillo} package. RcppArmadillo is an R package that provides an access to Armadillo, a powerful C++ library for linear algebra in conjunction with Rcpp. Replacing the main likelihood function with C++ improved the computation speed by a factor of around 100-1000.

4.3.2   Memoisation

Although introducing Rcpp was very effective for faster computation, the likelihood evaluation still incurred 1-2 seconds of computation time, which was impractical for multiple MCMC runs required for exhaustive model selection. To further speed up the process, memoisation technique was used. Due to the recursive nature of the likelihood and relatively localised data space (i.e. the likelihood is uniquely determined by the household composition and infection pattern, and the variation of these in the dataset was relatively limited), the original likelihood function involved unnecessary repetition of computation where a value was recalculated even if it was already known in the previous iterations. Memoisation is a useful approach to such settings. Every time a household-wise likelihood value is to be calculated, the code checks whether the same value has already been calculated and stored in a "dictionary". If the stored value exists, the value is fetched instead of repeating the same calculation; if the value does not exist in the dictionary (i.e. the likelihood was first to be computed), the likelihood function was executed and the value was newly stored in the dictionary. Memoisation achieved a speedup by a factor of around 20.

4.3.3   Parallelisation

Finally, MCMC runs for multiple model comparison were implemented in parallel. The MCMC

package {LaplacesDemon} used in this study has a function *LaplacesDemon.hpc*, which handles parallelisation of MCMC chains. Each chain of MCMC requires a certain burn-in period, which serves as a substantial overhead of parallelism, and thus the benefit of unnecessarily extensive parallelisation will be marginal. In Paper 1, 8 parallel chains were run by *LaplacesDemon.hpc* for each model implementation.

5   Paper 2: Bias correction methods for test-negative designs in the presence of misclassification

CrossMark

Bias correction methods for test-negative designs in the presence of misclassification

A. Endo[1,2,3] (iD) , S. Funk[1,2] and A. J. Kucharski[1,2]

[1]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel St., London WC1E 7HT, UK; [2]Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, Keppel St., London WC1E 7HT, UK and [3]The Alan Turing Institue, Euston Rd., London NW1 2DB, UK

Abstract

The test-negative design (TND) has become a standard approach for vaccine effectiveness (VE) studies. However, previous studies suggested that it may be more vulnerable than other designs to misclassification of disease outcome caused by imperfect diagnostic tests. This could be a particular limitation in VE studies where simple tests (e.g. rapid influenza diagnostic tests) are used for logistical convenience. To address this issue, we derived a mathematical representation of the TND with imperfect tests, then developed a bias correction framework for possible misclassification. TND studies usually include multiple covariates other than vaccine history to adjust for potential confounders; our methods can also address multivariate analyses and be easily coupled with existing estimation tools. We validated the performance of these methods using simulations of common scenarios for vaccine efficacy and were able to obtain unbiased estimates in a variety of parameter settings.

CAMBRIDGE
UNIVERSITY PRESS

Introduction

Vaccine effectiveness (VE) is typically estimated as the vaccine-induced risk reduction of the target disease (TD) and has been traditionally studied using cohort or case–control designs. However, the test-negative design (TND) is becoming a popular alternative design for VE studies [1, 2]. This is a modified version of the case–control study with an alternative definition of the control group; traditional case–control studies usually define controls as non-disease individuals in the study population, while TND studies use individuals with similar symptoms to the TD but presenting negative test results (i.e. patients of non-target diseases; ND). The TND can therefore minimise ascertainment bias by including only medically-attended patients in both case and control groups. Many TND studies have focused on influenza vaccination, but recent studies have also considered other diseases including pneumococcal disease [3, 4] and rotavirus disease [5–7].

Despite its increasing popularity, a TND can be more vulnerable than other study designs to misclassification of disease outcome. Multiple studies have shown that VE is underestimated when the diagnostic tests used in the study are imperfect (i.e. have a sensitivity and/or a specificity less than 100%) [8–10]. This can be a particular issue when simple tests (e.g. rapid diagnostic tests) are used for logistical convenience, as simple tests tend to have lower diagnostic performance than more advanced tests (e.g. polymerase chain reaction; PCR). Previous studies evaluated the expected degree of bias and concluded that specificity had a more important effect on bias than sensitivity [8–11]. These findings appear to support the use of rapid tests, despite limited sensitivity, because the specificity of these tests is typically high [2]. However, theoretical studies to date have been based on a limited range of assumptions about efficacy and pathogen epidemiology; it is therefore unclear whether such conclusions hold for all plausible combinations of scenarios.

If a study is expected to generate a non-negligible bias in estimation, such bias needs to be assessed and – if possible – corrected before the estimate is reported. Greenland [12] proposed a bias correction method for cohort studies where the sensitivity and specificity of the test are known (or at least assumed). However, it has been pointed out that bias correction in case–control studies is in general difficult because of differential recruitment, whereby the probability of recruiting (test-positive) cases and (negative) controls may be different [12, 13]. Although TND studies are often considered to be special cases of case–control studies, they are free from the issue of differential recruitment because the recruitment and classification are mutually-independent [14]. This means that, while Greenland's method does not apply to TND as-is, another type of bias correction may still be possible. For example, De Smedt *et al.* have characterised the misclassification bias in VE in the TND in a simulation study [10]. One limitation of their formulation was it relies on the unobserved âœtrueâ disease risk being known, where in reality this is not usually measurable in field studies. As a result,

bias correction methods for TND studies that are directly applicable to field data have not yet been proposed. Moreover, previous analysis of misclassification bias has not considered the impact of multivariate analysis, where potential confounders (e.g. age and sex) are also included in the model used to estimate VE.

To address these issues, we develop a bias correction method for the test-negative VE studies that uses only data commonly available in field studies. We also apply these methods to multivariate analyses. As our approach uses the so-called multiple over-imputation (MO) framework (generalisation of multiple imputation) [15], it can easily be coupled with a wide range of estimation tools without modifying their inside algorithms. Finally, we evaluate the performance of our methods by simulations of plausible epidemiological scenarios.

Methods and results

Characterising bias in TND studies

First, we consider the case where only vaccination history is included as a risk factor of acquiring the TD (i.e. the univariate setting). Following the approach of Haber *et al.* [16], we consider four steps in the case reporting process: vaccination, onset of symptoms, seeking of medical care and diagnosis. For simplicity, let us assume that occurrence of TD and ND are mutually independent, where their prevalences in the unvaccinated population are represented as r_1 and r_0, respectively[1]. Let n_V and n_U be the vaccinated and unvaccinated population size. The target variable in VE studies is γ, the relative risk of TD in the vaccinated population relative to the unvaccinated (i.e. $VE = 1 - \gamma$). Vaccinated and unvaccinated population can have different likelihoods of seeking medical treatment given disease. We denote by m_V and m_U the probability of medical attendance given ND in vaccinated and unvaccinated population, respectively.

As our focus in the present study is the bias in VE estimation caused by imperfect tests, we made two key assumptions following Haber *et al.* [16]. One assumption is that vaccines have no effect on the risk of ND. This enables ND patients to be eligible for a control group and is a key assumption in TND studies. The other assumption is that the probability of medical attendance in vaccinated and unvaccinated population given infection is constant regardless of the disease (TD or ND). The probability of medical attendance given TD may be different from ND (m_V and m_U), potentially due to difference in severity; we assume that these probabilities are obtained by multiplying a constant factor μ (i.e. μm_V and μm_U)[2]. These assumptions may not always hold and TND can be biased in such cases. However, we assume that they do in the following analysis to keep our focus on misclassification bias; namely, the study was assumed to be able to provide an unbiased VE estimate if tests are perfect.

Following the above notations, we can classify the vaccinated and unvaccinated population into multiple categories shown in Table 1. We can characterise different VE study designs (cohort, case–control and test-negative) by the categories in Table 1 from which each design tries to sample: the cohort design samples from populations n_V and n_U and follows them up to see what proportions fall into x_V and x_V; the case–control design samples from

medically-attended cases $(x_V + x_U)$ and non-diseased controls $(n_V - x_V + n_U - x_U)$ and calculate the odds ratio to approximate the relative risk (however, the actual studies can mismeasure these variables when misclassificaiton is present).

In TND studies, medically-attended patients $(x_V + x_U + y_V + y_U)$ are sampled and classified into four categories based on the test result and vaccine history. Let q be the proportion sampled relative to the population. Denoting the observed case counts with misclassification by X and Y, the process of data collection in TND can be represented by the following matrix expression:

$$\begin{bmatrix} X_V & X_U \\ Y_V & Y_U \end{bmatrix} = q \begin{bmatrix} \alpha & 1 - \beta \\ 1 - \alpha & \beta \end{bmatrix} \begin{bmatrix} x_V & x_U \\ y_V & y_U \end{bmatrix}, \qquad (1)$$

where α and β are the sensitivity and specificity of the test, respectively. Matrix

$$C = \begin{bmatrix} \alpha & 1 - \beta \\ 1 - \alpha & \beta \end{bmatrix}$$

describes the conversion from the true disease state to the observed result. We hereafter refer to C as the classification matrix. The determinant $c = |C| = \alpha + \beta - 1$ is the Youden index of the test and satisfies $0 < c \leq 1$ (if $c < 0$, the test is not predictive and the definitions of positive/negative should be swapped). Youden index indicates the level of information retained in the potentially misclassified test results. Youden index of 0 indicates that the information is completely lost and the test is no better than random guesses.

We define bias in the VE estimate to be the absolute difference between the (raw) estimate, derived from the misclassified observation, and the true value. Let $\delta = ((r_1 \mu)/r_0)$ be the odds of the (medically-attended) TD in the unvaccinated population. Then the expected bias B is given as a function of four independent parameters, α, β, γ and δ:

$$\begin{aligned} B(\alpha, \beta, \gamma, \delta) &= VE_{raw} - VE_{true} \\ &= (1 - \gamma_{raw}) - (1 - \gamma_{true}) \\ &= \gamma - \frac{[\alpha\gamma\delta + (1-\beta)][(1-\alpha)\delta + \beta]}{[(1-\alpha)\gamma\delta + \beta][\alpha\delta + (1-\beta)]}. \end{aligned} \qquad (2)$$

This suggests that the influence of sensitivity/specificity on the degree of bias varies depending on the case ratio $\delta/(1 + \delta)$, i.e. the ratio between the incidence of medical attendance for TD and ND in the unvaccinated study population[3] (Fig. 1). The degree of bias also depends on γ but is independent of m_V and m_U. The degree of bias is largely determined by the test specificity when the case ratio is small, but the influence of sensitivity and specificity is almost equivalent to a case ratio of 0.6. It is notable that high specificity does not always assure that the bias is negligible. This may be true if specificity is strictly 100% and the case ratio is low to moderate, but a slight decline to 97% can cause a bias up to 10–15 percentage points. The effect of sensitivity is also non-negligible when the case ratio is high.

When the expected bias is plotted against the case ratio with various combinations of test performance, we find that VE estimates can be substantially biased for certain case ratios (especially

---

[1]It has been suggested that a possible violation of this assumption occur as a result of virus interference [17], but conclusive evidence for this is currently lacking [18, 19] and the effect on VE estimates may be limited in any case [20].

[2]This may not be true, for example, if vaccination reduces the severity of TD and hence reduces the likelihood of medical attendance.

[3]For example, a case ratio of 0.5 indicates TD:ND = 1:1 in the unvaccinated. The value is smaller than 0.5 when TD < ND and greater than 0.5 when TD > ND.

Table 1. Population classified into different categories of interest in VE studies

	Vaccinated		Unvaccinated	
	Notation	Mean	Notation	Mean
All	n_V	n_V	n_U	n_U
Infected by TD	−	$\gamma r_1 n_V$	−	$r_1 n_U$
Medically-attended (true) TD patients	x_V	$\mu m_V \gamma r_1 n_V$	x_U	$\mu m_U r_1 n_U$
Test-positive TD patients	x_{+V}	$\alpha \mu m_V \gamma r_1 n_V$	x_{+U}	$\alpha \mu m_U r_1 n_U$
Test-negative TD patients	x_{-V}	$(1-\alpha)\mu m_V \gamma r_1 n_V$	x_{-U}	$(1-\alpha)\mu m_U r_1 n_U$
Infected by ND	−	$r_0 n_V$	−	$r_0 n_U$
Medically-attended (true) ND patients	y_V	$m_V r_0 n_V$	y_U	$m_U r_0 n_U$
Test-positive ND patients	y_{+V}	$(1-\beta)m_V r_0 n_V$	y_{+U}	$(1-\beta)m_U r_0 n_U$
Test-negative ND patients	y_{-V}	$\beta m_V r_0 n_V$	y_{-U}	$\beta m_U r_0 n_U$
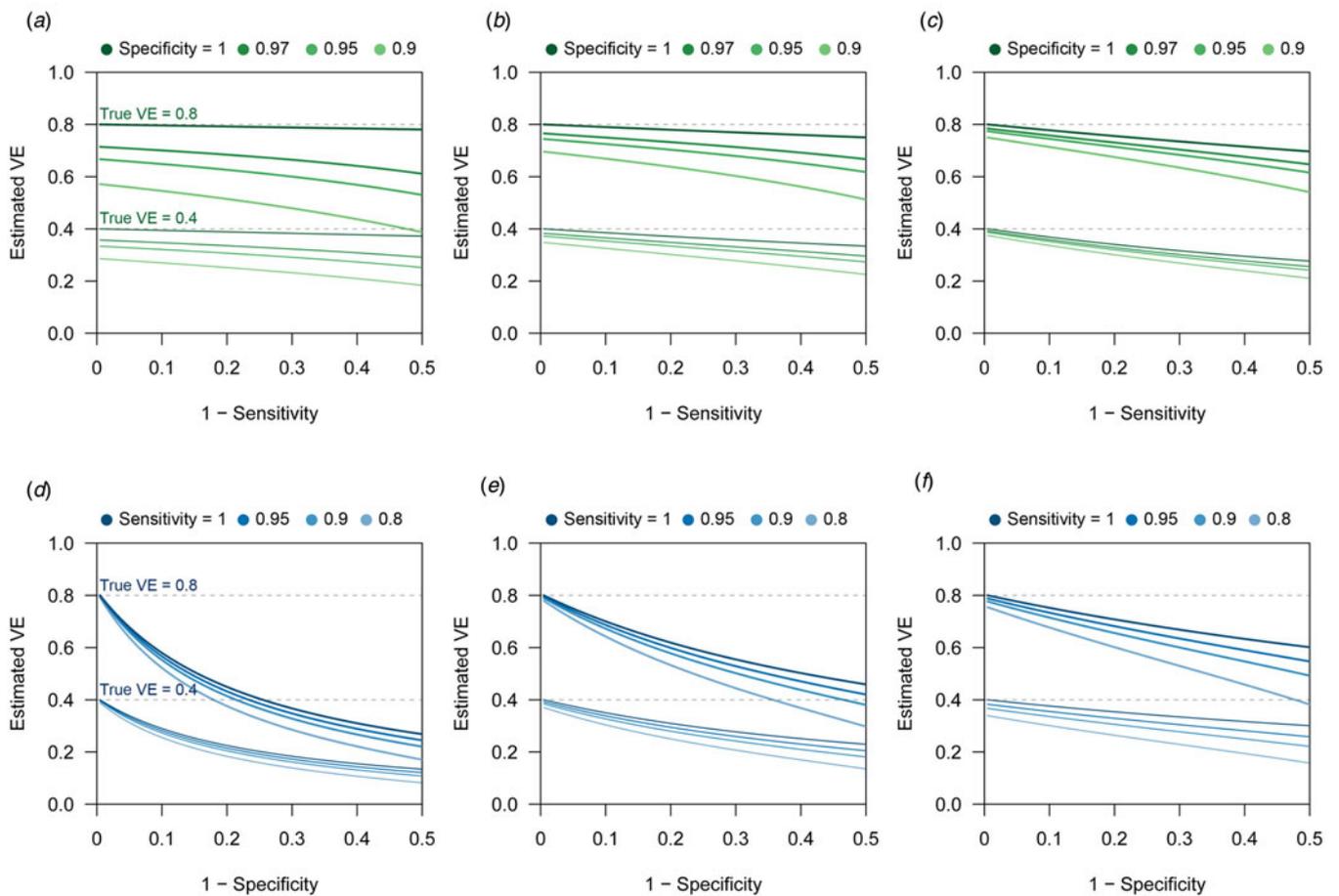
Fig. 1. Bias in VE estimates caused by misclassification for different combinations of parameter values. (a)–(c) Estimated VE plotted against sensitivity. (a) True case ratio (the true ratio between TD and ND cases included in the study) = 0.2 (b) 0.4 (c) 0.6. Each two sets of lines respectively correspond to different true VEs (80% and 40%, denoted by the dotted lines). (d)–(f) Estimated VE plotted against specificity. (d) True case ratio = 0.2 (e) 0.4 (f) 0.6.

when the ratio is far from 1:1), even with reasonably high sensitivity and specificity (Fig. 2b). In TND studies, researchers have no control over the case ratio because the study design requires that all tested individuals be included in the study. We found that the proportion of TD-positive patients in previous TND studies (retrieved from three systematic reviews [21–23]) varied

considerably, ranging from 10% to 70% (Fig. 2a)[4]. Because of this large variation in the case ratio, it would be difficult to predict the degree of bias before data collection. Post-hoc assessment and

[4]Strictly speaking, proportion positive is a different quantity from case ratio, but it should serve as a reasonable proxy of the case ratio in most settings.
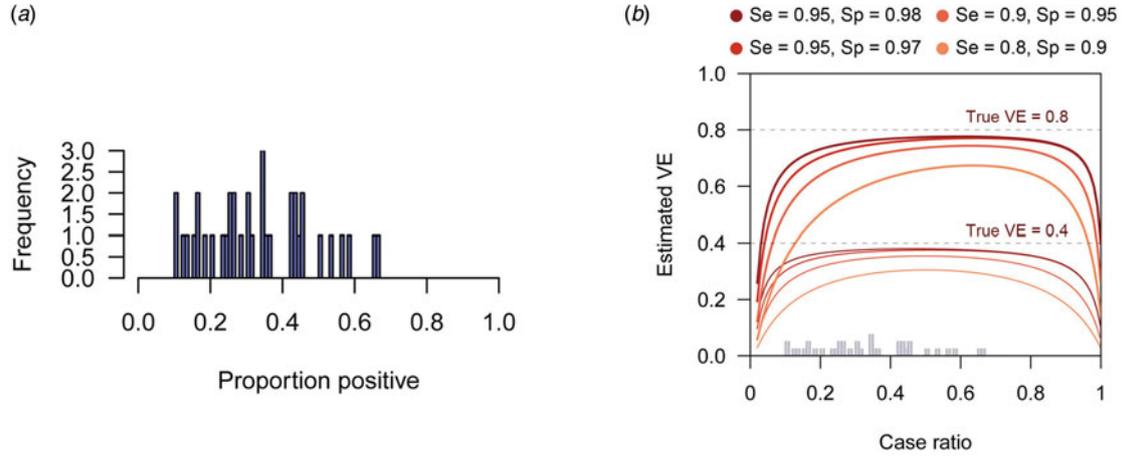
Fig. 2. Biased VE estimates with varying case ratio and the observed proportion of positive patients. (a) The proportion of test-positive patients in TND studies from systematic reviews. The proportions were retrieved from three systematic reviews [21–23]. (b) Estimated VE plotted against case ratio. Two sets of lines respectively correspond to different true VEs (80% and 40%, denoted by the dotted lines). The histogram in Panel (a) is overlaid on the *x*-axis.

correction therefore need to be considered. See the Supplementary Document for further analysis of the degree of bias.

Bias correction in univariate analysis

Corrected odds ratio

Although TND has sometimes been referred to as a special case of the case–control design, there is a distinct feature in the sampling procedure of TND. It has been pointed out that the adjustment methods for misclassification bias developed for cohort studies do not apply to case–control studies because the sampling ratio in case–control studies (differential recruitment) varies between case and control groups [12, 13]. However, as we have shown in Section 'Characterising bias in TND studies', both cases (TD patients) and controls (ND patients) are sampled at the same ratio (q) in TND studies. This suggests that the existing bias correction formulas developed for a hypothetical setting [13] where the whole population is evenly sampled (which is unrealistic in traditional studies) may be applicable to TND studies.

By left-multiplying Equation (1) with the inverted matrix C^{-1}, we can obtain the corrected odds ratio γ^* as

$$\gamma^* = \frac{X_V - ((1-\beta)/\beta)Y_V}{Y_V - ((1-\alpha)/\alpha)X_V} \cdot \frac{Y_U - ((1-\alpha)/\alpha)X_U}{X_U - ((1-\beta)/\beta)Y_U}, \quad (3)$$

which adjusts for misclassification to give an asymptotically-unbiased estimate of γ. This result can also be derived by maximising the likelihood accounting for misclassification in the observed TND data (see the Supplementary Document).

All four components of (3) (two numerators and two denominators) are usually expected to be non-negative with moderate VE (less than 100%) because these components are considered to be proportional to reconstructed true case counts. However, in some (relatively rare) cases, one or more components may become negative due to random fluctuations in observation. Theoretically, negative values are not permitted as true case counts, and thus such negative quantities would need to be truncated to 0. As a result, the corrected odds ratio can be either 0 or infinity. It is unrealistic in clinical settings that vaccines have absolute 100% or −100% effectiveness. Uncertainty around such MLEs should be carefully considered; increasing sample size or

redesigning the study might be recommended where possible. Alternatively, the Bayesian framework may be used to yield an interval estimate with the likelihood shown in section 'Direct likelihood method for the logistic regression model' adapted for a univariate model.

The confidence interval for VE can be obtained by assuming log-normality of the odds ratio γ, i.e.

$$\gamma = \gamma^* \exp{(\pm 1.96\sigma^*)},$$

where σ is the shape parameter of the log-normal distribution and is empirically given as

$$\sigma^* = SD(\log{(\gamma^*)})$$

$$= c \sqrt{\frac{\dfrac{X_V Y_V (X_V + Y_V)}{[\alpha Y_V - (1-\alpha)X_V]^2 [\beta X_V - (1-\beta)Y_V]^2} +}{\dfrac{X_U Y_U (X_U + Y_U)}{[\alpha Y_U - (1-\alpha)X_U]^2 [\beta X_U - (1-\beta)Y_U]^2}}}. \quad (4)$$

See the Supplementary Document for details of the MLE and confidence intervals.

Simulation

To assess the performance of the corrected odds ratio given in Equation (3) and uncertainty around it, we used simulation studies. TND study datasets were drawn from Poisson distributions (see the Supplementary Document for model settings and the likelihood function) as it is a reasonable assumption when medically-attended cases are recruited over the study period. We parameterised the mean incidence in the dataset by the 'baseline medical attendance' $\lambda_V = qm_V(r_1\mu + r_0)n_V$ and $\lambda_U = qm_U(r_1\mu + r_0)n_U$, so that λ_V and λ_U correspond to the mean number of vaccinated/unvaccinated patients when vaccine has no effect (i.e. $\gamma = 1$ and VE = 0). The mean total sample size (given as $((1 + \gamma\delta)/(1 + \delta))\lambda_V + \lambda_U$) was set to be 3000. Parameter values were chosen according to a range of scenarios shown in Table 2, and the true VE=1 −γ was compared with the estimates obtained from the simulated data. For each scenario, simulation was repeated 500 times to yield the distribution of estimates. Reproducible codes (including those

Table 2. Simulation settings

ID	Scenario	True VE (γ)	λ_V/λ_U	Case ratio ($\gamma/(1+\gamma)$)	Sensitivity (α)	Specificity (β)
1	Baseline: low VE	0.4	0.5	0.5	0.8	0.95
2	Baseline: high VE	0.8	0.5	0.5	0.8	0.95
3	High quality test: low VE	0.4	0.5	0.5	**0.95**	**0.97**
4	High quality test: high VE	0.8	0.5	0.5	**0.95**	**0.97**
5	Low quality test: low VE	0.4	0.5	0.5	**0.6**	0.9
6	Low quality test: high VE	0.8	0.5	0.5	**0.6**	0.9
7	High TD incidence: low VE	0.4	0.5	**0.7**	0.8	0.95
8	High TD incidence: high VE	0.8	0.5	**0.7**	0.8	0.95
9	Low TD incidence: low VE	0.4	0.5	**0.3**	0.8	0.95
10	Low TD incidence: high VE	0.8	0.5	**0.3**	0.8	0.95
11	High vaccine coverage: low VE	0.4	**0.7**	0.5	0.8	0.95
12	High vaccine coverage: high VE	0.8	**0.7**	0.5	0.8	0.95
13	Low vaccine coverage: low VE	0.4	**0.3**	0.5	0.8	0.95
14	Low vaccine coverage: high VE	0.8	**0.3**	0.5	0.8	0.95
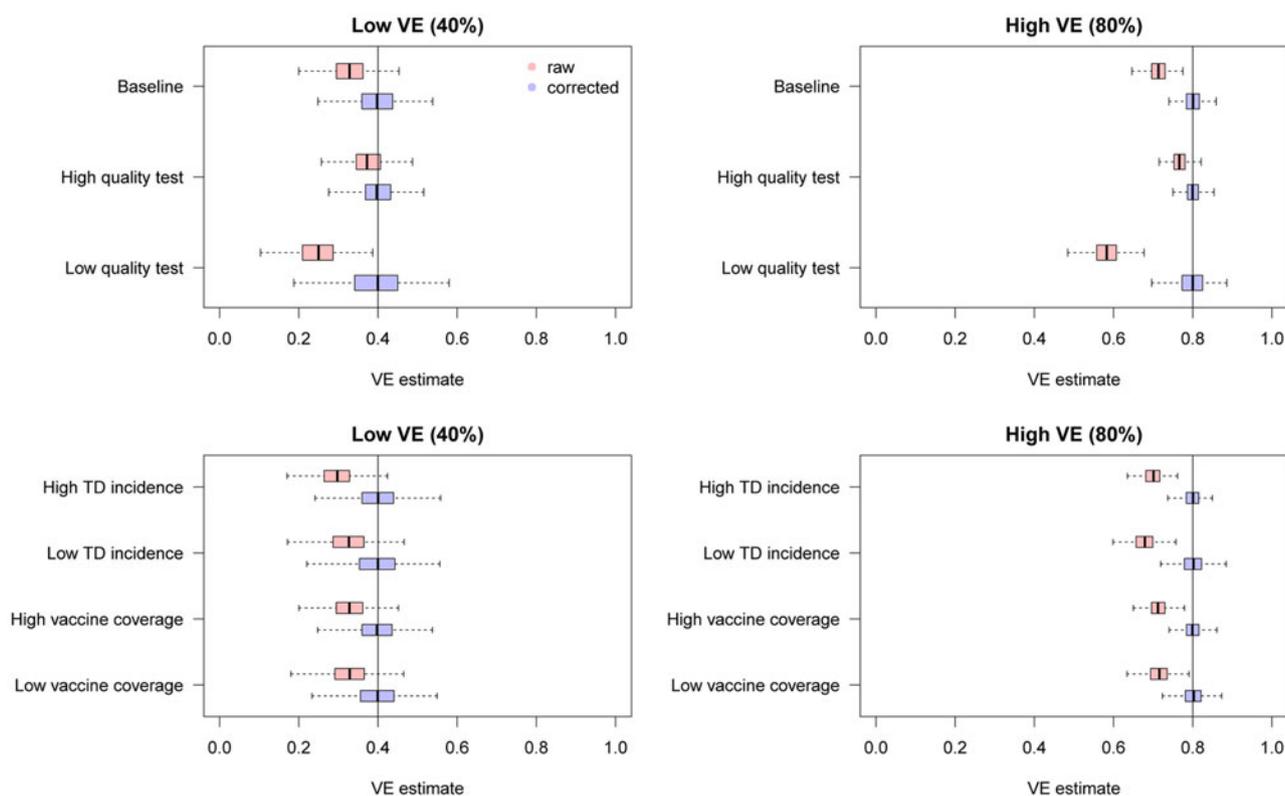
Fig. 3. Bias correction for simulated data in the univariate setting. The distributions of bias-corrected VE estimates (boxplots in blue) are compared with those of raw VE estimates without correction (red). Five hundred independent datasets were randomly generated for each set of parameter values, and the corrected and uncorrected VE estimates are compared with the true value (black solid line). See Table 2 for parameter settings in each scenario.

for simulations in later sections) are reposited on GitHub (https://github.com/akira-endo/TND-biascorrection/).

We found that the uncorrected estimates, directly obtained from the raw case counts that were potentially misclassified, exhibited substantial underestimation of VE for most parameter values (Fig. 3). On the other hand, our bias correction method

was able to yield unbiased estimates in every setting, whose median almost correspond to the true VE. Although the corrected and uncorrected distributions were similar (with a difference in median ~5%) when VE is relatively low (40%) and the test has sufficiently high sensitivity and specificity (95% and 97%, respectively), they became distinguishable with a higher VE (80%). With

lower test performances, the bias in the VE estimates can be up to 10–20%, which may be beyond the level of acceptance in VE studies.

Bias correction of VEs reported in previous studies

We have seen that the degree of bias for uncorrected VE estimates depends on parameter values. To explore the possible degree of bias in existing VE studies, we extracted the reported crude VEs (i.e. VEs unadjusted for potential confounders) from two systematic reviews [21, 23][5] and applied our bias correction method assuming different levels of test sensitivity and specificity. The case counts for each study summarised in the reviews were considered eligible for the analysis if the total sample size exceeded 200. Varying the assumed sensitivity and specificity, we investigated the possible discrepancy between the reported VE (or crude VE derived from the case counts if unreported in the reviews) and bias-corrected VE. We did not consider correcting adjusted VEs because it requires access to the original datasets.

Figure 4 displays the discrepancy between the reported VE and bias-corrected VE corresponding to a range of assumptions on the test performance. Many of the extracted studies employed PCR for the diagnostic test, which is expected to have a high performance. However, the true performance of PCR cannot be definitively measured as there is currently no other gold-standard test available. Figure 4 suggests that even a slight decline in the test performance can introduce a non-negligible bias in some parameter settings. Our bias correction methods may therefore also be useful in TND studies using PCR, which would enable a sensitivity analysis accounting for potential misdiagnosis by PCR tests. In this light, it is useful that the corrected odds ratio

$$\gamma^* = \frac{X_V - ((1-\beta)/\beta)Y_V}{Y_V - ((1-\alpha)/\alpha)X_V} \cdot \frac{Y_U - ((1-\alpha)/\alpha)X_U}{X_U - ((1-\beta)/\beta)Y_U}$$

is a monotonic function of both α and β (given that all the four components are positive). The possible range of VE in a sensitivity analysis is obtained by supplying γ^* with the assumed upper/lower limits of sensitivity and specificity.

Bias correction in multivariate analysis

Theoretical framework

TND studies often employ a multivariate regression framework to address potential confounding variables such as age. The most widespread approach is to use generalised linear models (e.g. logistic regression) and include vaccination history as well as other confounding variables as covariates. The estimated linear coefficient for vaccination history can then be converted VE (in the logistic regression model, the linear coefficient for vaccination history corresponds to $\log(1-VE)$). In this situation, the likelihood function now reflects a regression model and thus the bias-corrected estimate in the univariate analysis (Equation (3)) is no longer applicable. We therefore need to develop a separate multivariate TND study framework to correct for bias in multivariate analysis.

Suppose that covariates $\xi = (\xi^1, \xi^2, \ldots, \xi^n)$ are included in the model, and that ξ^1 corresponds to vaccination history (1: vaccinated, 0: unvaccinated). These covariates ξ_i, as well as outcome variable Z_i (i.e. test results) are available for each individual i included in the study. In TND studies, it is often convenient to
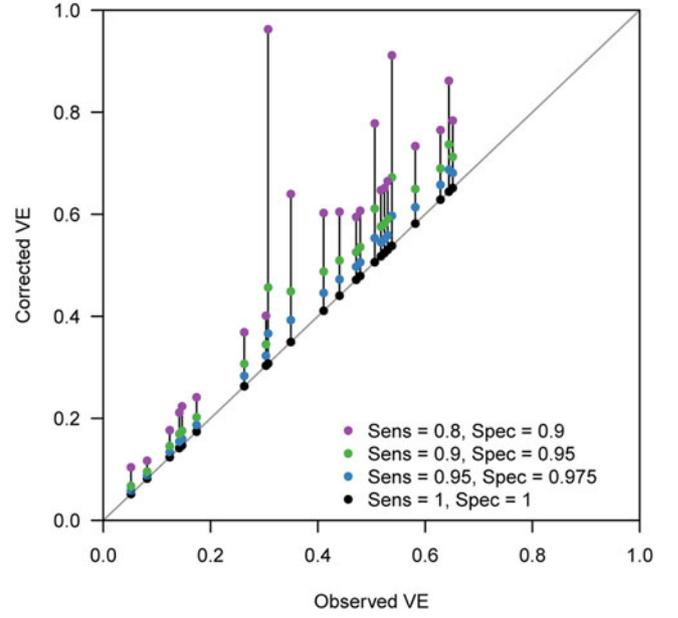
Fig. 4. Bias correction method applied to published VE estimates assuming various test sensitivity and specificity. Case count data were extracted from two systematic reviews [21, 23]. Each connected set of dots show how (crude) VE estimates reported in the review varies when imperfect sensitivity and specificity are assumed. Black dots on the grey diagonal line denote the original VEs reported in the reviews. This should correspond to the true value if sensitivity = specificity = 1. Coloured dots show the bias-corrected VE considering potential misclassification.

model the binomial probability for the true outcome $p_1(\xi_i)$, i.e. the conditional probability that the true outcome is TD as opposed to ND given an individual has covariates ξ_i. Let us use parameter set θ to model the binomial probabilities p_1 (and $p_0 = 1 - p_1$). Using the binomial probability π_{Z_i} for observed (potentially misclassified) outcome Z_i, we can obtain the MLE for θ by maximising

$$\mathcal{L}(\theta; D) = \prod_{S}^{i=1} \pi_{Z_i}(\xi_i; \theta)$$

$$= \prod_{i \in \{+\}} \left[\alpha p_1(\xi_i; \theta) + (1-\beta) p_0(\xi_i; \theta) \right] \quad (5)$$

$$\times \prod_{i \in \{-\}} \left[(1-\alpha) p_1(\xi_i; \theta) + \beta p_0(\xi_i; \theta) \right].$$

With the estimate θ^*, the VE estimate for an individual with covariates $\xi^{2:n} = (\xi^2, \xi^3, \ldots, \xi^n)$ is given as (1−odds ratio):

$$VE(\xi^{2:n}) = 1 - \frac{p_1(\xi^1 = 1, \xi^{2:n}; \theta^*)}{p_0(\xi^1 = 1, \xi^{2:n}; \theta^*)} \Big/ \frac{p_1(\xi^1 = 0, \xi^{2:n}; \theta^*)}{p_0(\xi^1 = 0, \xi^{2:n}; \theta^*)}. \quad (6)$$

See the Supplementary Document for further details.

Direct likelihood method for the logistic regression model

The logistic regression model is well-suited for modelling binomial probabilities p_1 and p_0. The log-odds $(\log(\frac{p_1}{p_0}))$ is characterised by a linear predictor as:

$$\log\left(\frac{p_1(\xi; \theta)}{p_0(\xi; \theta)}\right) = \theta_0 + \theta_1 \xi^1 + \cdots + \theta_n \xi^n. \quad (7)$$

---

[5]Young *et al.* [22] was not included because they did not report case counts.

In the logistic regression model where covariate ξ_1 indicates vaccination history, the corresponding coefficient θ_1 gives the VE estimate: $\text{VE} = 1 - \exp(\theta_1)$. Due to the assumed linearity, the estimated VE value is common across individuals regardless of covariates $\xi^{2:n}$.

We can employ the direct likelihood method by combining Equations (5) and (7). The usual logistic regression optimises θ by assuming that the test results follow Bernoulli distributions $Z_i \sim \text{Bernoulli}(p_1(\xi_i; \theta))$ ($Z_i = 1$ for positive test results and 0 for negative). To correct the misclassification bias, we instead need to use the modified probabilities to construct the likelihood accounting for diagnostic error, i.e.

$$Z_i \sim \text{Bernoulli}(\pi_+(\xi_i; \theta))$$
$$= \text{Bernoulli}(\alpha p_1(\xi_i; \theta) + (1 - \beta) p_0(\xi_i; \theta)). \qquad (8)$$

Parameter θ is estimated by directly maximising the probability of observing $\{Z_i\}$ based on Equation (8)

Note that as long as the binomial probability is the modelling target, other type of models (e.g. machine learning classifiers) could also be employed under a similar framework.

Multiple overimputation combined with existing tools

The direct likelihood method presented in the previous section is the most rigorous MLE approach and would therefore be preferable whenever possible. However, it is often technically-demanding to implement such approaches as it involves re-defining the likelihood; if we wanted to use existing tools for logistic regression (or other models), for example, we would need to modify the internal algorithm of such tools. This is in particular complicated in tools for generalised linear models including logistic regression, whose standard algorithm is the iteratively reweighted least squares method [24], which does not involve the explicit likelihood. To ensure that our correction methods can be employed without losing access to substantial existing software resources, we also propose another method, which employs a MO framework [15] to account for misclassification. Whereas multiple imputation only considers missing values, MO is proposed as a more general concept which includes overwriting mismeasured values in the dataset by imputation. In our multivariate bias correction method, test results in the dataset (which are potentially misclassified) are randomly overimputed.

Let M be an existing estimation software tool whose likelihood specification cannot be reprogrammed. Given data $d = \{z_i, \xi_i\}_{i=1,2,\ldots,S}$, where z_i denotes the true disease state ($z = 1$ for TD and $z = 0$ for ND), M would be expected to return at least the following two elements: the point estimate of VE (ε_d) and the predicted binomial probability $\hat{p}_1(\xi_i)$ for each individual i. From the original observed dataset D, J copies of imputed datasets $\{\tilde{D}^j\} = \{\tilde{D}^1, \tilde{D}^2, \ldots, \tilde{D}^J\}$ are generated by the following procedure.

(1) For $i = 1, 2, \ldots, S$, impute disease state $\tilde{z}_i^j$ based on the test result Z_i. Each $\tilde{z}_i^j$ is sampled from a Bernoulli distribution conditional to Z_i:

$$\tilde{z}_i^j \sim \begin{cases} \text{Bernoulli}(1 - \tilde{\varphi}_{i+}) & (Z_i = 1) \\ \text{Bernoulli}(\tilde{\varphi}_{i-}) & (Z_i = 0) \end{cases}. \qquad (9)$$

(2) $\tilde{\varphi}_{i+}$ and $\tilde{\varphi}_{i-}$ are estimated probabilities that the test result for individual i is incorrect (i.e. $z_i \neq Z_i$) given Z_i. The sampling procedure (9) is therefore interpreted as the test result Z_i

being 'flipped' at a probability $\tilde{\varphi}_{i+}$ or $\tilde{\varphi}_{i-}$. Later we will discuss possible procedures to obtain these probabilities.

(3) Apply M to $\tilde{D}^j = \{\tilde{z}_i^j, \xi_i\}$ to yield a point estimate of VE (ε^j).
(4) Repeat (1) and (2) for $j = 1, 2, \ldots, J$ to yield MO estimates $\{\varepsilon^j\}_{j=1,\ldots,J}$.

Once MO estimates $\{\varepsilon^j\}$ are obtained, the pooled estimate and confidence intervals of VE are obtained by appropriate summary statistics, e.g. Rubin's rules [25]. As long as the estimated 'flipping' probabilities $\tilde{\varphi}_{i\pm} = (\tilde{\varphi}_{i+}, \tilde{\varphi}_{i-})$ are well chosen, this MO procedure should provide an unbiased estimate of VE with a sufficiently large number of iterations J.

As a method to estimate the flipping probability $\tilde{\varphi}_{i\pm}$, here we propose the parametric bootstrapping described as follows. Given the observed test result Z_i, $\tilde{\varphi}_{i\pm}$ is given as the Bayesian probability

$$\mathcal{P}(z_i = Z_i | Z_i)$$
$$= \begin{cases} \dfrac{\alpha \mathcal{P}(z_i = 1)}{\alpha \mathcal{P}(z_i = 1) + (1 - \beta)\mathcal{P}(z_i = 0)} = \dfrac{\alpha p_1(\xi_i)}{\alpha p_1(\xi_i) + (1 - \beta)p_0(\xi_i)} & (Z_i = 1) \\[3mm] \dfrac{\beta \mathcal{P}(z_i = 0)}{(1 - a)\mathcal{P}(z_i = 1) + \beta \mathcal{P}(z_i = 0)} = \dfrac{\beta p_0(\xi_i)}{(1 - \alpha)p_1(\xi_i) + \beta p_0(\xi_i)} & (Z_i = 0) \end{cases}$$
$$(10)$$

Although the true binomial probabilities $p_0(\xi_i)$, $p_1(\xi_i)$ are unknown, their estimators are derived with the inverted classification matrix in the same manner as Equation (3). By substituting

$$\begin{bmatrix} p_1(\xi_i) \\ p_0(\xi_i) \end{bmatrix}$$

with

$$C^{-1} \begin{bmatrix} \pi_+(\xi_i) \\ \pi_-(\xi_i) \end{bmatrix},$$

we get

$$\tilde{\varphi}_{i+} = 1 - \mathcal{P}(z_i = 1 | Z_i = 1) = \frac{1 - \beta}{\alpha + \beta - 1} \left[\alpha \cdot \frac{\pi_-(\xi_i)}{\pi_+(\xi_i)} - (1 - \alpha) \right]$$
$$\tilde{\varphi}_{i-} = 1 - \mathcal{P}(z_i = 0 | Z_i = 0) = \frac{1 - \alpha}{\alpha + \beta - 1} \left[\beta \cdot \frac{\pi_+(\xi_i)}{\pi_-(\xi_i)} - (1 - \beta) \right]$$
$$(11)$$

These probabilities can be computed provided the odds of the test results $\pi_+(\xi_i)/\pi_-(\xi_i)$. We approximate this odds by applying estimation tool M to the original data D; i.e. the predicted binomial probability $\hat{p}_1(\xi_i)$ obtained from D is used as a proxy of $\pi_+(\xi_i)$. Generally it is not assured that true and observed probabilities $p_1(\xi_i)$ and $\pi_+(\xi_i)$ have the same mechanistic structure captured by M; however, when our concern is limited to the use of model-predicted probabilities to smooth the data D, we may expect for M to provide a sufficiently good approximation. The above framework can be regarded as a variant of parametric bootstrapping methods as MO datasets are generated from data D assuming a parametric model M. The whole bias correction procedure is presented in pseudocode (Fig. 5); sample R code is also available on GitHub (https://github.com/akira-endo/TND-biascorrection/).

EM algorithm

Another possible approach to addressing misclassification is the use of the EM algorithm, which has been proposed for case–

Algorithm 1 Multiple imputation with parametric bootstrapping

Input: $D = \{Z_i, \xi_i\}_{i=1,2,\ldots S}$

Fit model M to D to obtain a predictive model $\pi_+ = \hat{p}_1(\xi)$

for $j = 1, 2, \ldots, J$ **do**
    for $i = 1, \ldots, S$ **do**
        $\pi_+ \leftarrow \hat{p}_1(\xi_i)$ ▷ Predict the binomial probability π_+ by model M
        $\tilde{z}_i^j \leftarrow Z_i$ ▷ Copy Z_i, and then flip at a probability φ to impute $\tilde{z}_i^j$
        if $Z_i = 1$ **then**
            $\varphi \leftarrow \frac{1-\beta}{\alpha+\beta-1}\left[\alpha \cdot \frac{1-\pi_+}{\pi_+} - (1-\alpha)\right]$
            $u \leftarrow \text{Unif}(0,1)$
            if $u < \varphi$ **then**
                $\tilde{z}_i^j \leftarrow 0$
            end if
        else
            $\varphi \leftarrow \frac{1-\alpha}{\alpha+\beta-1}\left[\beta \cdot \frac{\pi_+}{1-\pi_+} - (1-\beta)\right]$
            $u \leftarrow \text{Unif}(0,1)$
            if $u < \varphi$ **then**
                $\tilde{z}_i^j \leftarrow 1$
            end if
        end if
    end for
    Fit model M to $\tilde{D}^j = \{\tilde{z}_i^j, \xi_i\}_{i=1,2,\ldots S}$ to estimate parameter ε^j
end for
Output: MO estimates $\{\varepsilon^j\}_{j=1,2,\ldots,J}$

Fig. 5. Multiple imputation with parametric bootstrapping.

control studies in a previous study (where differential recruitment was not considered) [26]. Because of its methodological similarity, the algorithm can also be applied to the TND. The original EM algorithm presented in [26] would produce, if properly implemented, the result equivalent to the direct likelihood approach. However, the original EM algorithm requires that the model can handle non-integer sample weights (which may not always be assured). Moreover, computing confidence intervals in EM algorithm can be complicated. We therefore recommend parametric bootstrapping as the first choice of bias correction method when the direct likelihood approach is inconvenient.

*Simulation of bias correction with parametric bootstrapping*
To assess the performance of this method, we used the same simulation framework as in the univariate analysis (Table 2). In addition to vaccination history (denoted by ξ^1), we consider one categorical and one continuous covariate. Let us assume that ξ_2 represents the age group (categorical; 1: child, 0: adult) and ξ_3 the pre-infection antibody titre against TD (continuous). Suppose that the population ratio between children and adults is 1:2, and that ξ_3 is scaled so that it is standard normally distributed in the population. For simplicity, we assumed that all the covariates are mutually independent with regard to the distribution and effects (i.e. no association between covariates and no interaction effects). The relative risk of children was set to be 2 and 1.5 for TD and ND, respectively, and a unit increase in the antibody titre was assumed to halve the risk of TD (and not to affect the risk of ND). The mean total sample size λ was set to be 3000, and 500 sets of simulation data were generated for each scenario. VE estimates were corrected by the parametric

bootstrapping approach (the number of iterations $J = 100$) and were compared with the raw (uncorrected) VE estimates.

Figure 6 shows the distributions of estimates with and without bias correction in the multivariate setting. Our bias correction (parametric bootstrapping) provided unbiased estimates for all the scenarios considered. Overall, biases in the uncorrected estimates were larger than those in the univariate setting. In some scenarios, the standard error of the bias-corrected estimates was extremely wide. This was primarily because of the uncertainty already introduced before misclassification rather than the failure of bias correction (as can be seen in the Supplementary Fig. S3). Larger sample size is required to yield accurate estimates in those settings, as the information loss due to misclassification will be added on top of the inherent uncertainty in the true data.

The number of confounding variables
We investigated how the bias in uncorrected VE estimates can be affected by the number of confounding variables. In addition to the vaccine history ξ^1, we added a set of categorical/continuous confounding variables to the model and assessed the degree of bias caused by misclassification. The characteristics of the variables were inherited from those in section 'Simulation of bias correction with parametric bootstrapping': categorical variable 'age' and continuous variable 'pre-infection antibody titre'. That is, individuals were assigned multiple covariates (e.g. 'categorical variable A', 'categorical variable B', ..., 'continuous variable A', 'continuous variable B', ...) whose distribution and effect were identical to 'age' (for categorical variables) and 'antibody titre' (for continuous variables) in section 'Simulation of bias correction with parametric bootstrapping'. No interaction between
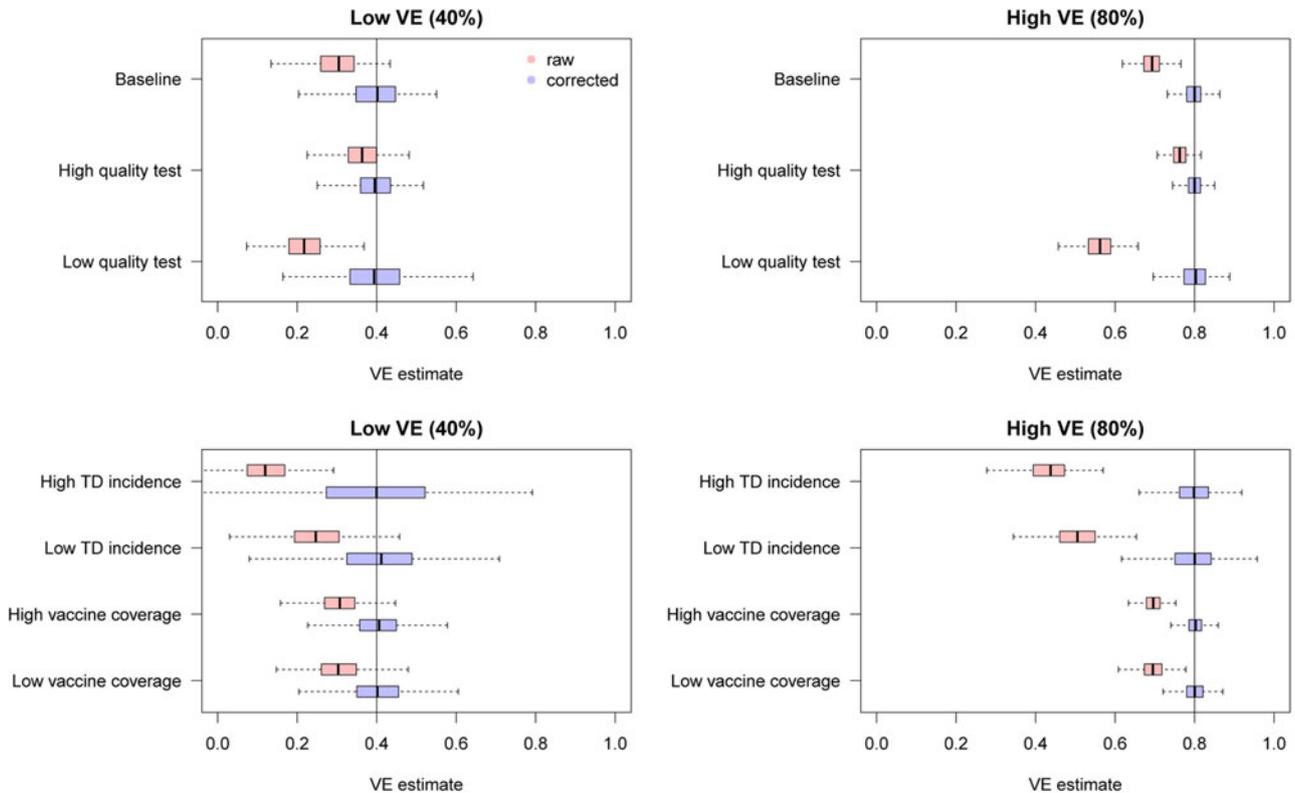
Fig. 6. Bias correction for simulated data in the multivariate setting. The distributions of bias-corrected (blue) and uncorrected (red) VE estimates from 500 simulations are compared. Dotted lines denote median and black solid lines denote the true VE. The parametric bootstrapping bias correction method was used for bias correction.

covariates was assumed. The covariate set in section 'Simulation of bias correction with parametric bootstrapping' being baseline (the number of covariates: (vaccine, categorical, continuous) = (1, 1, 1)), we employed two more scenarios with a larger number of covariates: (1, 3, 3) and (1, 5, 5).

The simulation results are presented in Figure 7. Overall, additional confounding variables led to more severe bias in the uncorrected VE estimates towards underestimation. These results further highlight the importance of bias correction when heterogeneous disease risks are expected; VE estimates adjusted for many confounding variables can exhibit substantial misclassification bias.

Discussion

Misclassification caused by imperfect diagnostic tests can potentially lead to substantial biases in TND studies. By considering the processes involved in VE estimation, we have characterised the degree of bias potentially caused by misclassification in different parameter settings, finding that VE can be noticeably underestimated, particularly when the ratio between TD and ND cases in the study data is unbalanced. To address this potential bias, we developed multiple bias correction methods that provide unbiased VE estimates in both univariate and multivariate settings. When the test sensitivity and specificity are known or assumed, those values can be used to restore the true VE estimate by a relatively simple statistical procedure. Using simulations, we showed that our methods could successfully eliminate the bias in VE estimates obtained from misclassified data, although some uncertainty was introduced as a result of the information loss.

We believe that our methods could therefore enable researchers to report unbiased VE estimates even when imperfect tests had to be used. Such methods could also help in the scaling up of TND studies, as tests with limited performance are usually inexpensive and logistically convenient.

Although TND is a relatively new study design, first appearing in a publication in 2005 [27], it has gained broad popularity and is becoming a standard approach in VE studies. One of the largest factors that have contributed to its widespread use is the fact that data collection can be completed within clinical setups [1]. Whereas cohort or case–control studies usually require additional efforts including follow-up or recruitment of non-patients, TND studies only involve patients visiting healthcare facilities with suspects of certain diseases and thus routinely collected clinical data can be easily adapted for analysis. VE studies of influenza, for which TND is most frequently used, often use PCR as a diagnostic tool for better data quality [23]. However, such studies usually involve intensive effort and cost, and thus may only be feasible by large-scale research bodies. Our bias correction methods may open a possibility of wider use of clinical data especially in settings where rapid tests are routinely used for diagnosis. For example, rapid influenza diagnostic tests are routinely used for outpatient clinics and hospitals in Japan, and such clinical data have facilitated a number of TND studies [28–33]. Such studies based on rapid tests could benefit from our methods, as it would provide strong support for the validity of their estimates. Our methods may also be useful in resource-limited settings or for diseases without high-performance diagnostic tools.

Even in resourceful settings where high-performance tests are available, the slight possibility of misclassification might not
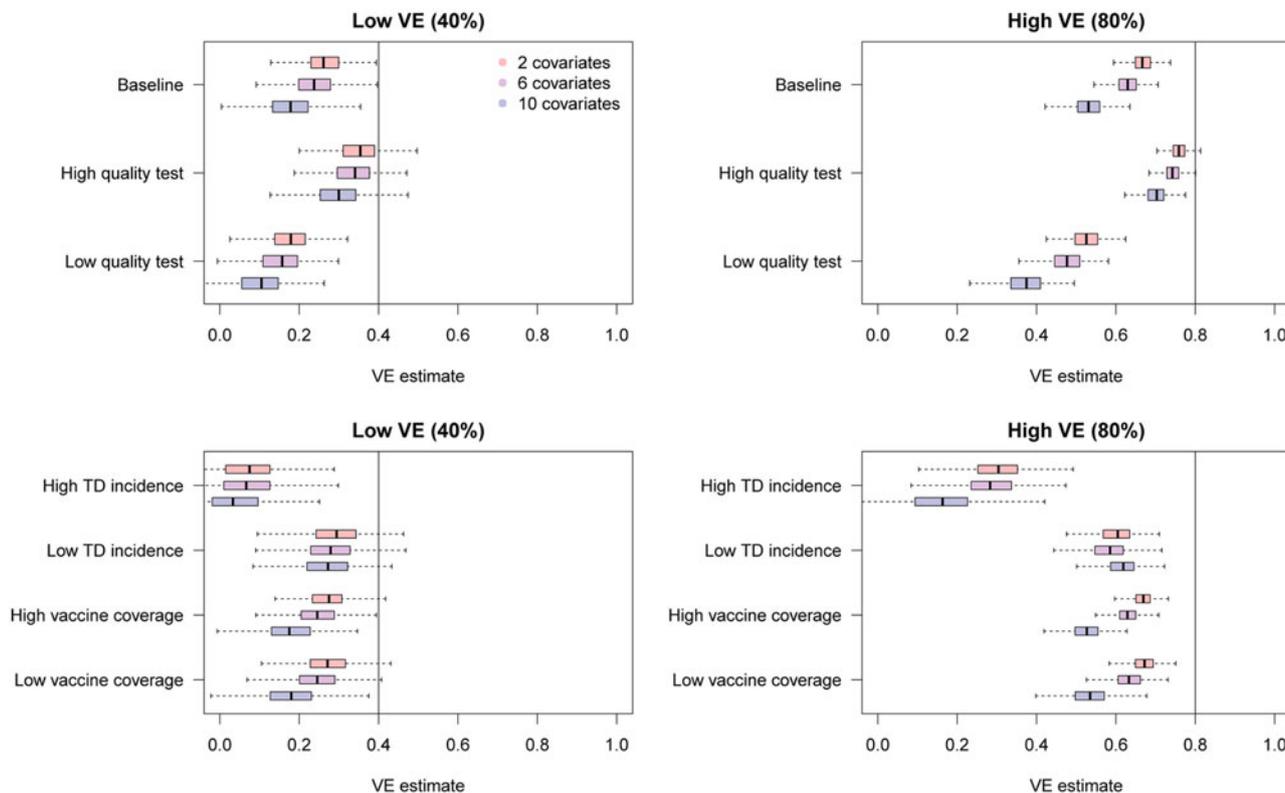
Fig. 7. Bias in raw VE estimates from simulated data in the presence of different numbers of confounding variables. The distributions in red, purple and blue correspond to uncorrected VE estimates in the presence of 2, 6 and 10 confounding variables in addition to the vaccination history.

always be neglected. Although PCR tests are currently used as a gold-standard for influenza diagnosis, their sensitivity and specificity may not be exact 100%; especially, the sensitivity of the test depends not only on microbiological technique but also on the quality of swab samples. In addition, it is suggested that the sensitivity of PCR tests may change during the time course of infection [34] and be sufficiently high only during a limited time window. Our simulation study also indicated that a high heterogeneity in individual characteristics in the study population might increase the bias. Our methods could enable researchers to implement sensitivity analysis by assuming the possible test sensitivity and specificity in such cases.

Our bias correction methods are also intended to be reasonably straightforward for researchers to introduce. Existing estimation tools including software libraries and packages are often used in epidemiological analyses. Incorporating the MO approach, our parametric bootstrapping bias correction method only involves data manipulation and does not require modification of the estimation algorithm. Once multiple sets of data are randomly generated, any type of analysis can be performed as long as the results can be summarised over the MO datasets. Of particular note is that our methods for multivariate analysis allow stratification of sensitivity and specificity among individuals. Therefore, the users can employ more complex misclassification mechanisms including time-varying test performance or test performance affected by individual characteristics. Datasets with a mixture of different diagnostic tools [3, 35] can also be handled by applying different values for each test.

There are some limitations to our study. We only focused on misclassification of diagnosis (i.e. misclassified outcomes) and

did not consider misclassification of covariates (e.g. vaccine history and other confounding variables), which is another important type of misclassification in TND studies [10]. Further, it is generally not easy to plausibly estimate the sensitivity and specificity for measurement of covariates (e.g. recall bias), which must be known or assumed to implement bias correction. However, if reliable estimates are available, an extension of our approach may yield bias-corrected VE estimates in the presence of covariate misclassification. Moreover, to keep our focus only on diagnostic misclassification, our methods rested on the assumption that other sources of bias in TND studies are non-existent or properly addressed. Potential sources of bias in TND studies have been discussed elsewhere [16, 36, 37], and the researchers conducting TND studies need to carefully consider the possibility of such biases in addition to the diagnostic misclassification. Lastly, it must be noted that our methods depend on the assumed test sensitivity and specificity, and that misspecifying those values can result in an improper correction. The sensitivity and specificity of tests are usually reported by manufacturers in a comparison of the test results with gold-standard tests; however, when such gold-standard tests themselves are not fully reliable or when no available test has satisfactory performance to be regarded as gold-standard, specifying sensitivity and specificity of a test is in principle impossible. Further, test performances reported by manufacturers might lack sufficient sample size or might not be identical to those in the actual study settings. Use of composite reference standards [38, 39] or external/internal validation approaches [40] may help overcome these problems.

Although the presence of imperfect diagnosis limits the quality of clinical data, such data can still hold useful information, and

this information can be transformed into useful insights by appropriate statistical processing. Our bias correction methods were developed primarily for TND studies, but a similar approach could be applied to broader classes of estimation problems with misclassification. Potential areas for future analysis include extension to test data involving continuous quantitative measurements, and coupling with dynamic transmission models. The value of routinely collected data in healthcare settings has become widely recognised with the advancement of data infrastructure, and we believe our methods could help support the effective use of such data.

Conclusion

Bias correction methods for the TND studies were developed to address potential misclassification bias due to imperfect tests.

References

1. **De Serres G *et al.*** (2013) The test-negative design: validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. *Eurosurveillance.*
2. **Fukushima W and Hirota Y** (2017) Basic principles of test-negative design in evaluating influenza vaccine effectiveness. *Vaccine.*
3. **Suzuki M *et al.*** (2017) Serotype-specific effectiveness of 23-valent pneumococcal polysaccharide vaccine against pneumococcal pneumonia in adults aged 65 years or older: a multicentre, prospective, test-negative design study. *The Lancet Infectious Diseases* **17**, 313–321.
4. **Prato R *et al.*** (2018) Effectiveness of the 13-valent pneumococcal conjugate vaccine against adult pneumonia in Italy: a case–control study in a 2-year prospective cohort. *BMJ Open* **8**.
5. **Araki K *et al.*** (2018) Effectiveness of monovalent and pentavalent rotavirus vaccines in Japanese children. *Vaccine* **36**, 5187–5193.
6. **Lopez AL *et al.*** (2018) Effectiveness of monovalent rotavirus vaccine in the Philippines. *Scientific Reports* **8**, 14291.
7. **Muhsen K *et al.*** (2018) Effectiveness of rotavirus pentavalent vaccine under a universal immunization programme in Israel, 2011–2015: a case control study. *Clinical Microbiology and Infection* **24**, 53–59.
8. **Orenstein EW *et al.*** (2007) Methodologic issues regarding the use of three observational study designs to assess influenza vaccine effectiveness. *International Journal of Epidemiology.*
9. **Jackson ML and Rothman KJ** (2015) Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. *Vaccine.*
10. **De Smedt T *et al.*** (2018) Bias due to differential and non-differential disease – and exposure misclassification in studies of vaccine effectiveness. *PLoS ONE.*
11. **Haber M *et al.*** (2018) A comparison of the test-negative and traditional case-control study designs with respect to the bias of estimates of rotavirus vaccine effectiveness. *Vaccine* **36**, 5071–5076.
12 **Greenland S** (1996) Basic methods for sensitivity analysis of biases.
13. **Kleinbaum DG, Kupper LL and Morgenstern H** (1982) *Epidemiologic Research: Principles and Quantitative Methods.* John Wiley & Sons.
14. **Jackson ML and Nelson JC** (2013) The test-negative design for estimating influenza vaccine effectiveness. *Vaccine* **31**, 2165–2168.
15. **Blackwell M, Honaker J and King G** (2017) A unified approach to measurement error and missing data: overview and applications. *Sociological Methods and Research.*
16. **Haber M *et al.*** (2015) A probability model for evaluating the bias and precision of influenza vaccine effectiveness estimates from case-control studies. *Epidemiology and Infection* **143**, 1417–1426.
17. **Cowling BJ and Nishiura H** (2012) Virus interference and estimates of influenza vaccine effectiveness from test-negative studies. *Epidemiology (Cambridge, Mass.).*
18. **Cowling BJ *et al.*** (2012) Increased risk of noninfluenza respiratory virus infections associated with receipt of inactivated influenza vaccine. *Clinical Infectious Diseases.*
19. **Sundaram ME *et al.*** (2013) Influenza vaccination is not associated with detection of noninfluenza respiratory viruses in seasonal studies of influenza vaccine effectiveness. *Clinical Infectious Diseases.*
20. **Suzuki M, Camacho A and Ariyoshi K** (2014) Potential effect of virus interference on influenza vaccine effectiveness estimates in test-negative designs. *Epidemiology and Infection.*
21. **Leung VK *et al.*** (2016) Concordance of interim and final estimates of influenza vaccine effectiveness: a systematic review. *Eurosurveillance.*
22. **Young B *et al.*** (2018) Duration of influenza vaccine effectiveness: a systematic review, meta-analysis, and meta-regression of test-negative design case-control studies. *Journal of Infectious Diseases.*
23. **Darvishian M *et al.*** (2014) Effectiveness of seasonal influenza vaccine in community-dwelling elderly people: a meta-analysis of test-negative design case-control studies. *The Lancet Infectious Diseases.*
24. **Sidney Burrus C** (2012) Iterative reweighted least squares.
25. **Rubin DB** (2004) *Multiple Imputation for Nonresponse in Surveys.* Wiley Classics Library. Wiley.
26. **Magder LS and Hughes JP** (1997) Logistic regression when the outcome is measured with uncertainty. *American Journal of Epidemiology* **146**, 195–203.
27. **Skowronski DM *et al.*** (2005) Effectiveness of vaccine against medical consultation due to laboratory-confirmed influenza: results from a sentinel physician pilot project in British Columbia, 2004–2005. *Canada Communicable Disease Report.*
28. **Seki Y *et al.*** (2016) [Effectiveness of influenza vaccine in adults using A test-negative, case-control design −2013/2014 and 2014/2015 seasons-] (Japanese). *Kansenshogaku zasshi. The Journal of the Japanese Association for Infectious Diseases.*
29. **Seki Y, Onose A and Sugaya N** (2017) Influenza vaccine effectiveness in adults based on the rapid influenza diagnostic test results, during the 2015/16 season. *Journal of Infection and Chemotherapy.*
30. **Saito N *et al.*** (2017) Negative impact of prior influenza vaccination on current influenza vaccination among people infected and not infected in prior season: a test-negative case-control study in Japan. *Vaccine.*
31. **Shinjoh M *et al.*** (2018) Inactivated influenza vaccine effectiveness and an analysis of repeated vaccination for children during the 2016/17 season. *Vaccine.*
32. **Ando S** (2018) Effectiveness of quadrivalent influenza vaccine based on the test-negative control study in children during the 2016–2017 season. *Journal of Infection and Chemotherapy.*
33. **Sugaya N *et al.*** (2018) Three-season effectiveness of inactivated influenza vaccine in preventing influenza illness and hospitalization in children in Japan, 2013–2016. *Vaccine.*
34. **Sullivan SG, Feng S and Cowling BJ** (2014) Potential of the test-negative design for measuring influenza vaccine effectiveness: a systematic review. *Expert Review of Vaccines.*

35. **Franke MF** *et al.* (2017) Comparison of two control groups for estimation of oral cholera vaccine effectiveness using a case-control study design. *Vaccine*.

36. **Sullivan SG, Tchetgen EJ and Cowling BJ** (2016) Theoretical basis of the test-negative study design for assessment of influenza vaccine effectiveness. *American Journal of Epidemiology* **184**, 345–353.

37. **Lewnard JA** *et al.* (2018) Measurement of vaccine direct effects under the test-negative design. *American Journal of Epidemiology*.

38. **Limmathurotsakul D** *et al.* (2012) Fool's gold: why imperfect reference tests are undermining the evaluation of novel diagnostics: a reevaluation of 5 diagnostic tests for leptospirosis. *Clinical Infectious Diseases*.

39. **Naaktgeboren CA** *et al.* (2013) Value of composite reference standards in diagnostic research. *BMJ (Clinical research ed.)*.

40. **Kahn SR** *et al.* (2018) Comparing external and internal validation methods in correcting outcome misclassification bias in logistic regression: a simulation study and application to the case of postsurgical venous thromboembolism following total hip and knee arthroplasty. *Pharmacoepidemiology and Drug Safety*.

Appendix

Mathematical notations in section 'Methods and results'

Characterising bias in TND studies

r_1, r_0:	prevalence of TD/ND in the unvaccinated population
n_V, n_U:	vaccinated and unvaccinated population size
γ:	the relative risk of TD in the vaccinated population relative to the unvaccinated ($VE = 1 - \gamma$)
m_V, m_U:	the probability of medical attendance given ND in vaccinated/unvaccinated population
μ:	factor for the probability of medical attendance given TD in relative to that given ND

$x_V, x_U, x_{+V}, x_{+U}, x_{-V}, x_{-U},$ $y_V, y_U, y_{+V}, y_{+U}, y_{-V}, y_{-U}$:	the expected number of cases in the population; see Table 1 for definitions
X_V, X_U, Y_V, Y_U:	the observed case counts subject to misclassification
q:	the proportion of study samples relative to the total population
α, β:	sensitivity and specificity of the test
C:	the classification matrix
c:	the Youden index of the test; the determinant of C δ: odds of medically-attended TD in the unvaccinated population

Bias correction in univariate analysis

$\gamma^\star$:	corrected odds ratio
$\sigma^\star$:	shape parameter of the log-normal distribution that gives the confidence interval of $\gamma^\star$ (see Equation (4))
λ_V, λ_U:	'baseline medical attendance', the mean number of vaccinated/unvaccinated patients when vaccine has no effect ($VE = 0$)

Bias correction in multivariate analysis

$\xi = (\xi^1, \xi^2, \ldots, \xi^n)$:	covariates included in the model, where xi^1 denotes vaccination history
Z_i, z_i:	observed test result and true disease state for individual i
$p_1(\xi), p_0(\xi)$:	probability that the true test result is positive/negative for an individual with covariates ξ
$\theta, \theta^\star$:	model parameter for $p_1(\xi)$ and its estimate
$\pi_{Z_i}(\xi_i)$:	probability for the observed test result Z_i
$\tilde{z}_i^j$:	imputed disease state of individual i in the j-th imputed dataset $\tilde{D}^j$
$\tilde{\phi}_{i+}, \tilde{\phi}_{i-}$:	estimated probability that the test result for individual i is incorrect
ε^j:	point estimate for VE from the j-th imputed dataset $\tilde{D}^j$

5.2 Supplementary materials

SUPPLEMENTARY DOCUMENT: BIAS CORRECTION METHODS FOR TEST-NEGATIVE DESIGN IN THE PRESENCE OF MISCLASSIFICATION

Akira Endo,   **Sebastian Funk,**   **Adam J. Kucharski**

1 Expected degree of bias and parameter settings

Based on Equation (3) in the main text, we extensively explored the relationship between the bias in vaccine effectiveness (VE) estimates and the parameter settings. Figure S1 shows the expected degree of bias for different combinations of parameter values (sensitivity: α, specificity: β, true VE: $1 - \gamma$ and case ratio: $\frac{\delta}{1+\delta}$). A nonlinear relationship can be seen in the contour plots. In general, the influence of sensitivity is larger in the high case ratio region, and that of specificity is larger in the low case ratio region. Both sensitivity and specificity affect the bias when the case ratio is intermediate. Interestingly, the true VE have different effects depending on the case ratio: high VE leads to larger bias when the case ratio is low, while moderate VE gives the largest bias when the case ratio is intermediate to high.

As the specificity of diagnostic tests tend to be higher than the sensitivity, we further explored the cases of higher specificity (> 95%) in Figure S2. Even if the specificity is sufficiently high, the level of bias can be larger than one might expect within the realistic range of case ratio (0.1-0.7).

2 Univariate model and parameter estimation

The sample size of a TND study is usually unconstrained as the study design requires every patient with TD-like symptoms to be included, where the study population is typically limited to patients visiting specific medical institutions and is sufficiently smaller than the total population. Therefore, it is natural that we assume that the reported incidence of TD and ND both follow Poisson-distributions. The mean number of unvaccinated patients in the dataset is given as $\lambda_U = q m_U (r_1 \mu + r_0) n_U$. Let $\lambda_V = \frac{m_V n_V}{m_U n_U} \lambda_U$ so that λ_V corresponds to the mean number of vaccinated patients when $\gamma = 1$, i.e. VE = 0. This definition is to ensure that parameters γ and λ_V are mutually independent. Let $\delta = \frac{r_1 \mu}{r_0}$ be the odds of the (medically-attended) target disease in the unvaccinated population. Using these four parameters $\gamma, \delta, \lambda_V, \lambda_U$, we get the following table for (potentially mis-classified) mean case counts:

	Vaccinated	Unvaccinated
Test positive	$\frac{\alpha\gamma\delta+(1-\beta)}{1+\delta}\lambda_V$	$\frac{\alpha\delta+(1-\beta)}{1+\delta}\lambda_U$
Test negative	$\frac{(1-\alpha)\gamma\delta+\beta}{1+\delta}\lambda_V$	$\frac{(1-\alpha)\delta+\beta}{1+\delta}\lambda_U$
Subtotal	$\frac{1+\gamma\delta}{1+\delta}\lambda_V$	λ_U

When data $D = (X_V, Y_V, X_U, Y_U)$ is obtained following this misclassified pattern, we can construct the likelihood of obtaining such data, given underlying parameters, as

$$\mathcal{L}(\gamma, \delta, \lambda_V, \lambda_U; D) = \mathrm{Pois}\left(X_V; \frac{\alpha\gamma\delta + (1-\beta)}{1+\delta}\lambda_V\right) \mathrm{Pois}\left(Y_V; \frac{(1-\alpha)\gamma\delta + \beta}{1+\delta}\lambda_V\right)$$

$$\mathrm{Pois}\left(X_U; \frac{\alpha\delta + (1-\beta)}{1+\delta}\lambda_U\right) \mathrm{Pois}\left(Y_U; \frac{(1-\alpha)\delta + \beta}{1+\delta}\lambda_U\right). \tag{S1}$$

By expanding this equation we get

$$
\mathcal{L}(\gamma, \delta, \lambda_V, \lambda_U; D)
$$
$$
= \frac{[\alpha\gamma\delta + (1-\beta)]^{X_V}[(1-\alpha)\gamma\delta + \beta]^{Y_V}[\alpha\delta + (1-\beta)]^{X_U}[(1-\alpha)\delta + \beta]^{Y_U}\lambda_V^{S_V}\lambda_U^{S_U}}{(1+\delta)^{S_V}(1+\delta)^{S_U}X_V!Y_V!X_U!Y_U!\exp\left(\frac{1+\gamma\delta}{1+\delta}\lambda_V\right)\exp(\lambda_U)}, \tag{S2}
$$

where $S_V = X_V + Y_V$ and $S_U = X_U + Y_U$.

For mathematical convenience, we change the variable λ_V to $\lambda_V' = \frac{1+\gamma\delta}{1+\delta}\lambda_V$. Let $l = \log\mathcal{L}(\gamma, \delta, \lambda_V', \lambda_U; X)$. Partial derivatives of l are

$$
\begin{aligned}
\frac{\partial l}{\partial \gamma} &= \frac{\alpha\delta X_V}{\alpha\gamma\delta + (1-\beta)} + \frac{(1-\alpha)\delta Y_V}{(1-\alpha)\gamma\delta + \beta} - \frac{\delta S_V}{1+\gamma\delta}\\
\frac{\partial l}{\partial \delta} &= \frac{\alpha\gamma X_V}{\alpha\gamma\delta + (1-\beta)} + \frac{(1-\alpha)\gamma Y_V}{(1-\alpha)\gamma\delta + \beta} + \frac{\alpha X_U}{\alpha\delta + (1-\beta)} + \frac{(1-\alpha)Y_U}{(1-\alpha)\delta + \beta} - \frac{\gamma S_V}{1+\gamma\delta} - \frac{S_U}{1+\delta}\\
\frac{\partial l}{\partial \lambda_V'} &= \frac{S_V}{\lambda_V'} - 1\\
\frac{\partial l}{\partial \lambda_U} &= \frac{S_U}{\lambda_U} - 1
\end{aligned} \tag{S3}
$$

Equation (S3) gives the maximum likelihood estimates:

$$
\begin{aligned}
\gamma^* &= \frac{X_V - \frac{1-\beta}{\beta}Y_V}{Y_V - \frac{1-\alpha}{\alpha}X_V} \cdot \frac{Y_U - \frac{1-\alpha}{\alpha}X_U}{X_U - \frac{1-\beta}{\beta}Y_U}\\
\delta^* &= \frac{\beta}{\alpha} \cdot \frac{X_U - \frac{1-\beta}{\beta}Y_U}{Y_U - \frac{1-\alpha}{\alpha}X_U}\\
\lambda_V'^* &= S_V\\
\lambda_U^* &= S_U
\end{aligned} \tag{S4}
$$

The confidence intervals for parameters can be constructed using the Fisher's information matrix from Equation (S3). λ_V' and λ_U are independent from other parameters and

$$
\mathrm{Var}(\lambda_V') = -\frac{\partial^2 l}{\partial \lambda_V'^2} = \frac{S_V}{\lambda_V'^2} - \frac{\partial^2 l}{\partial \lambda_U^2} = \frac{S_U}{\lambda_U^2} \tag{S5}
$$

We log-transform γ and δ for mathematical convenience. Noting that $\frac{\partial v}{\partial(\log u)} = u\frac{\partial v}{\partial u}$, we get

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial(\log(\gamma))^2} &= \frac{\gamma\delta}{(1+\gamma\delta)^2}S_V - \frac{\alpha(1-\beta)\gamma\delta}{[\alpha\gamma\delta + (1-\beta)]^2}X_V - \frac{(1-\alpha)\beta\gamma\delta}{[(1-\alpha)\gamma\delta + \beta]^2}Y_V\\
-\frac{\partial^2 l}{\partial\log\gamma\partial\log\delta} &= \frac{\gamma\delta}{(1+\gamma\delta)^2}S_V - \frac{\alpha(1-\beta)\gamma\delta}{[\alpha\gamma\delta + (1-\beta)]^2}X_V - \frac{(1-\alpha)\beta\gamma\delta}{[(1-\alpha)\gamma\delta + \beta]^2}Y_V\\
-\frac{\partial^2 l}{\partial(\log\delta)^2} &= \frac{\gamma\delta}{(1+\gamma\delta)^2}S_V - \frac{\alpha(1-\beta)\gamma\delta}{[\alpha\gamma\delta + (1-\beta)]^2}X_V - \frac{(1-\alpha)\beta\gamma\delta}{[(1-\alpha)\gamma\delta + \beta]^2}Y_V\\
&+ \frac{\delta}{(1+\delta)^2}S_U - \frac{\alpha(1-\beta)\delta}{[\alpha\delta + (1-\beta)]^2}X_U - \frac{(1-\alpha)\beta\delta}{[(1-\alpha)\delta + \beta]^2}Y_U
\end{aligned} \tag{S6}
$$

With the parameter values estimated in Eq. (S4), we get the following information matrix

$$
\begin{bmatrix}
\frac{\hat{x}_V\hat{y}_V}{S_V}\left[1 - S_V\left(\frac{\alpha(1-\beta)}{X_V} + \frac{(1-\alpha)\beta}{Y_V}\right)\right] & \frac{\hat{x}_V\hat{y}_V}{S_V}\left[1 - S_V\left(\frac{\alpha(1-\beta)}{X_V} + \frac{(1-\alpha)\beta}{Y_V}\right)\right]\\
\frac{\hat{x}_V\hat{y}_V}{S_V}\left[1 - S_V\left(\frac{\alpha(1-\beta)}{X_V} + \frac{(1-\alpha)\beta}{Y_V}\right)\right] & \frac{\hat{x}_V\hat{y}_V}{S_V}\left[1 - S_V\left(\frac{\alpha(1-\beta)}{X_V} + \frac{(1-\alpha)\beta}{Y_V}\right)\right] + \frac{\hat{x}_U\hat{y}_U}{S_U}\left[1 - S_U\left(\frac{\alpha(1-\beta)}{X_U} + \frac{(1-\alpha)\beta}{Y_U}\right)\right]
\end{bmatrix}
$$

72

where $\hat{x}_\xi = \frac{1}{c}[\beta X_\xi - (1-\beta)Y_\xi]$ and $\hat{y}_\xi = \frac{1}{c}[\alpha Y_\xi - (1-\alpha)X_\xi]$ are the estimated true case counts for $\xi = V, U$. Let $p_V = x_V/(x_V + y_V)$ and $p_U = x_U/(x_U + y_U)$ be the corresponding true binomial probabilities.

The inverse of the information matrix provides variance of estimates: in particular, for $\log\gamma$ we get

$$
\begin{aligned}
\mathrm{Var}(\log\gamma^*) &= \frac{S_V}{x_V y_V} \cdot \frac{1}{\left[1 - \left(\frac{\alpha(1-\beta)}{\pi_V} + \frac{(1-\alpha)\beta}{1-\pi_V}\right)\right]} + \frac{S_U}{x_U y_U} \cdot \frac{1}{\left[1 - \left(\frac{\alpha(1-\beta)}{\pi_U} + \frac{(1-\alpha)\beta}{1-\pi_U}\right)\right]} \\
&= \frac{S_V}{x_V y_V} \cdot \frac{\pi_V(1-\pi_V)}{(1 - \pi_V - (1-\alpha))(\pi_V - (1-\beta))} + \frac{S_U}{x_U y_U} \cdot \frac{\pi_U(1-\pi_U)}{(1 - \pi_U - (1-\alpha))(\pi_U - (1-\beta))} \\
&= \frac{c^2}{S_V} \frac{\pi_V(1-\pi_V)}{(1-\pi_V-(1-\alpha))^2(\pi_V-(1-\beta))^2} + \frac{c^2}{S_U} \frac{\pi_U(1-\pi_U)}{(1-\pi_U-(1-\alpha))^2(\pi_U-(1-\beta))^2},
\end{aligned}
$$ (S7)

equivalent to the Eq. (7) in the main text. We can relate this to the true standard error that would be obtained with perfect tests,

$$
\mathrm{SD}(\log\gamma_{\text{true}}) = \sqrt{\frac{1}{S_V p_V(1-p_V)} + \frac{1}{S_U p_U(1-p_U)}} = \sqrt{\frac{\sigma_V{}^2}{S_V} + \frac{\sigma_U{}^2}{S_U}},
$$ (S8)

or to the observed standard error (without correction),

$$
\mathrm{SD}(\log\gamma_{\text{raw}}) = \sqrt{\frac{1}{S_V \pi_V(1-\pi_V)} + \frac{1}{S_U \pi_U(1-\pi_U)}} = \sqrt{\frac{\Sigma_V{}^2}{S_V} + \frac{\Sigma_U{}^2}{S_U}},
$$ (S9)

where $\sigma_V = [p_V(1-p_V)]^{-1/2}$ and $\sigma_U = [p_U(1-p_U)]^{-1/2}$ are the components of the true standard error and $\Sigma_V = [\pi_V(1-\pi_V)]^{-1/2}$ and $\Sigma_U = [\pi_U(1-\pi_U)]^{-1/2}$ are those of uncorrected standard error. We get

$$
\begin{aligned}
\sigma^* = \mathrm{SD}(\log(\gamma^*)) &= \sqrt{\frac{\sigma_V^2}{S_V} \cdot \frac{1}{\left(1 - \frac{1-\alpha}{1-\pi_V}\right)\left(1 - \frac{1-\beta}{\pi_V}\right)} + \frac{\sigma_U^2}{S_U} \cdot \frac{1}{\left(1 - \frac{1-\alpha}{1-\pi_U}\right)\left(1 - \frac{1-\beta}{\pi_U}\right)}} \\
&= \frac{1}{c}\sqrt{\frac{\Sigma_V^2}{S_V} \cdot \left(\frac{\pi_V(1-\pi_V)}{p_V(1-p_V)}\right)^2 + \frac{\Sigma_U^2}{S_U} \cdot \left(\frac{\pi_U(1-\pi_U)}{p_U(1-p_U)}\right)^2}.
\end{aligned}
$$ (S10)

This equation indicates that the confidence intervals diverge when the true outcome is bipolarised ($p_V, p_U \simeq 0$ or 1).

3 Multivariate model and likelihood

Suppose that covariates $\xi = (\xi^1, \xi^2, ..., \xi^n)$ are included in the model, and that ξ^1 corresponds to vaccination history (1: vaccinated, 0: unvaccinated). ξ is expected to have a certain distribution over the total population N, and let us denote the frequency density of covariates ξ by $N(\xi)$, where $\int N(\xi)d\xi = N$. Let $\rho_1(\xi)$ and $\rho_0(\xi)$ be the conditional probabilities that an individual is included in the study with TD and ND, respectively, given covariates ξ. Incorporating misclassification, the probability of an individual i with covariates ξ_i being included and tested positive/negative will be

$$
\begin{aligned}
\rho_+(\xi_i) &= \alpha\rho_1(\xi_i) + (1-\beta)\rho_0(\xi_i) \\
\rho_-(\xi_i) &= (1-\alpha)\rho_1(\xi_i) + \beta\rho_0(\xi_i)
\end{aligned}
$$ (S11)

Assuming that disease incidences follow Poisson distributions, as in the univariate case, we can obtain the probability density of observing data $D = \{Z_i, \xi_i\}_{i=1,2,...S}$ (Z_i denotes the test result of individual i) as

$$
\mathcal{P}(D) = \mathrm{Pois}(S_+; \lambda_+)\,\mathrm{Pois}(S_-; \lambda_-) \prod_{i\in\{+\}} \frac{\rho_+(\xi_i)N(\xi_i)}{\lambda_+} \prod_{i\in\{-\}} \frac{\rho_-(\xi_i)N(\xi_i)}{\lambda_-}.
$$ (S12)

where λ_+ and λ_- are the mean incidence of being included in the study and tested positive/negative: $\lambda_\pm = \int \rho_\pm(\xi)N(\xi)d\xi$. The first two Poisson distributions on the right-hand side of Eq. (S12) give the probability that the study yields S_+ positive and S_- negative subjects. The products that follow represent the probability density for covariates ξ_i observed in the positive/negative group.

Suppose that we model this system using a parameter set θ. We could directly model $\rho_1(\xi_i; \theta)$ and $\rho_0(\xi_i; \theta)$; however, it is often more convenient to model the binomial probability for the true outcome $p_1(\xi_i) = \frac{\rho_1(\xi_i)}{\rho(\xi_i)}$ and $p_0(\xi_i) = \frac{\rho_0(\xi_i)}{\rho(\xi_i)}$, where $\rho(\xi_i) = \rho_1(\xi_i) + \rho_0(\xi_i) = \rho_+(\xi_i) + \rho_-(\xi_i)$ is the probability density of being included in the study given covariates ξ, because the absolute scale of incidence is rarely of a primary concern. The binomial probabilities for the respective observed outcomes (with errors) are then given by:

$$\begin{aligned} \pi_+(\xi_i; \theta) &= \alpha p_1(\xi_i; \theta) + (1 - \beta)p_0(\xi_i; \theta) \\ \pi_-(\xi_i; \theta) &= (1 - \alpha)p_1(\xi_i; \theta) + \beta p_0(\xi_i; \theta) \end{aligned} \tag{S13}$$

Let us use parameter set θ to model the binomial probabilities π_+ (and π_-) and assume that another set of parameters η (nuisance parameters) characterise $\rho(\xi_i)$. Then our objective is reduced to the estimation of θ and η.

Rearranging Equation (S12), we get the joint likelihood for θ and η:

$$\mathcal{L}(\theta, \eta; D) = \binom{S}{S_+} \text{Pois}(S; \lambda(\eta)) \prod_{i=1}^{S} \frac{\rho(\xi_i; \eta)N(\xi_i)}{\lambda(\eta)} \prod_{i=1}^{S} \pi_{Z_i}(\xi_i; \theta), \tag{S14}$$

where $\lambda(\eta)$ is the overall mean incidence: $\lambda(\eta) = \int \rho(\xi; \eta)N(\xi)d\xi$. The factor outside the products on the right-hand side of Eq. (S14) is the probability that the study yields S subjects of which S_+ are positives and S_- are negatives. The first product is the probability density for covariates ξ_i observed in data D, and the second product is the binomial probabilities for the test results Z_i. When only θ is of our concern, we can obtain the MLE for θ by maximising

$$\mathcal{L}(\theta; D) = \prod_{i=1}^{S} \pi_{Z_i}(\xi_i; \theta) = \prod_{i \in \{+\}} [\alpha p_1(\xi_i; \theta) + (1 - \beta)p_0(\xi_i; \theta)] \prod_{i \in \{-\}} [(1 - \alpha)p_1(\xi_i; \theta) + \beta p_0(\xi_i; \theta)], \tag{S15}$$

as θ and η are separate in the likelihood (S14).

4 Increased uncertainty introduced by misclassification

Although our bias correction methods provide unbiased VE estimates from potentially misclassified test results, the resulting uncertainty is larger than that which would be obtained from estimates using the true disease status. In Figure S3, we compared bias-corrected estimates obtained from misclassified data (by the direct likelihood method in the multivariate setting) with those obtained from the true data (i.e., 100% sensitivity and specificity). Although both estimates are unbiased around the true value, the results from the misclassified data exhibit higher variability (by a factor of 1.1-3.0) due to the loss of information caused by misdiagnosis. Increased uncertainty due to misclassification should be carefully considered when one calculates the power of test-negative design studies. Overestimated test performance may not only underestimate the true VE but also lead to overconfidence.
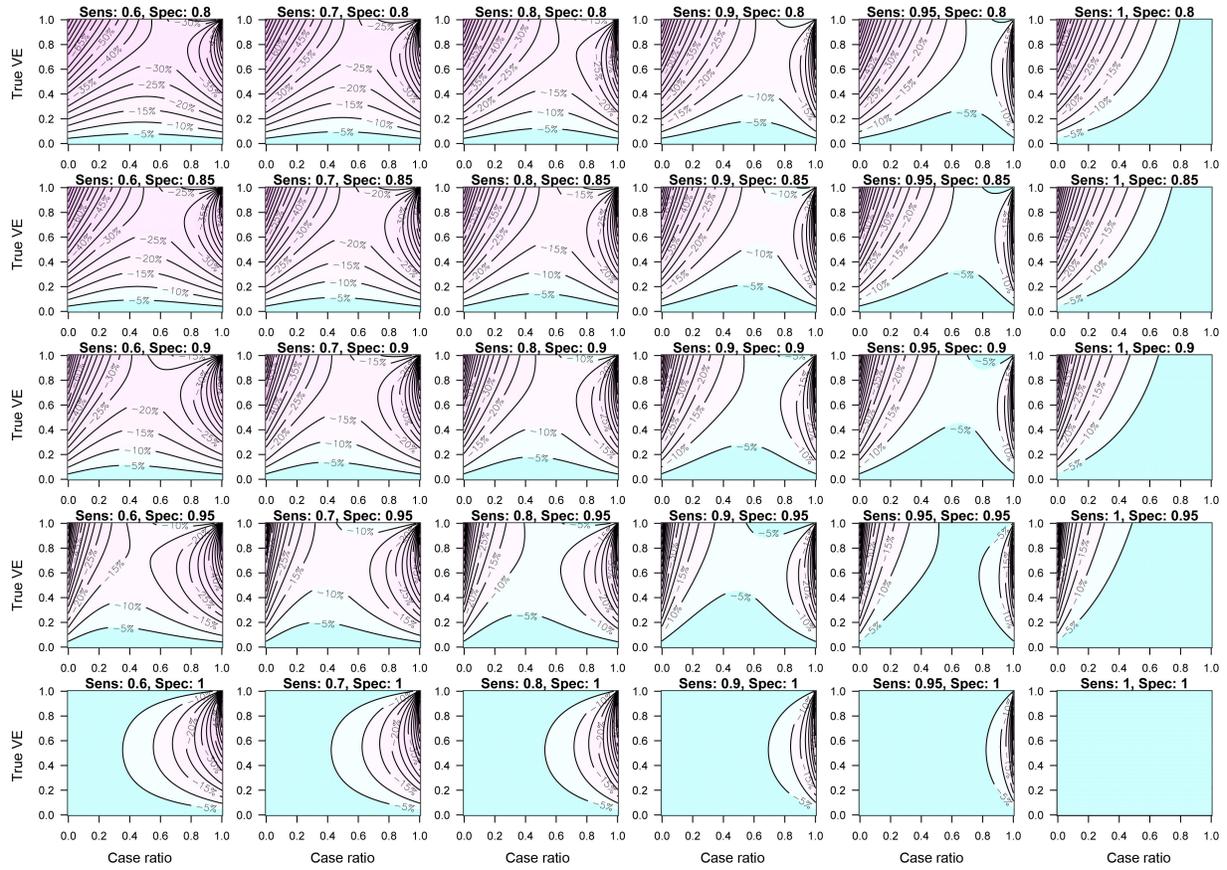
Figure S1: Bias in VE estimates caused by misclassification. Absolute bias (difference between the estimated VE and the true VE) for a set of parameter values is displayed in contour plots in percentage points. Negative figures indicate that the estimated VE is lower than the true VE.
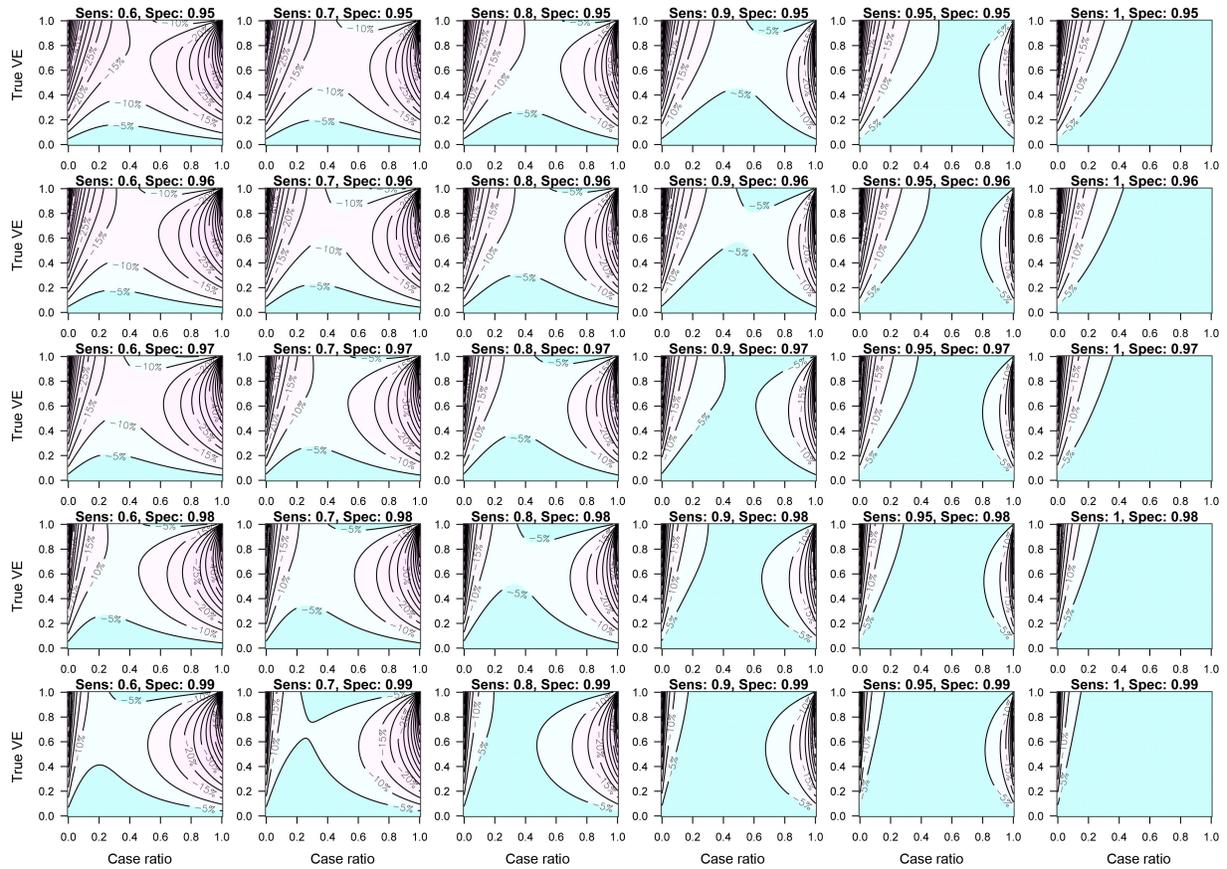
Figure S2: Bias in VE estimates caused by misclassification (high specificity). Absolute bias (difference between the estimated VE and the true VE) for a set of parameter values is displayed in contour plots in percentage points. Negative figures indicate that the estimated VE is lower than the true VE.
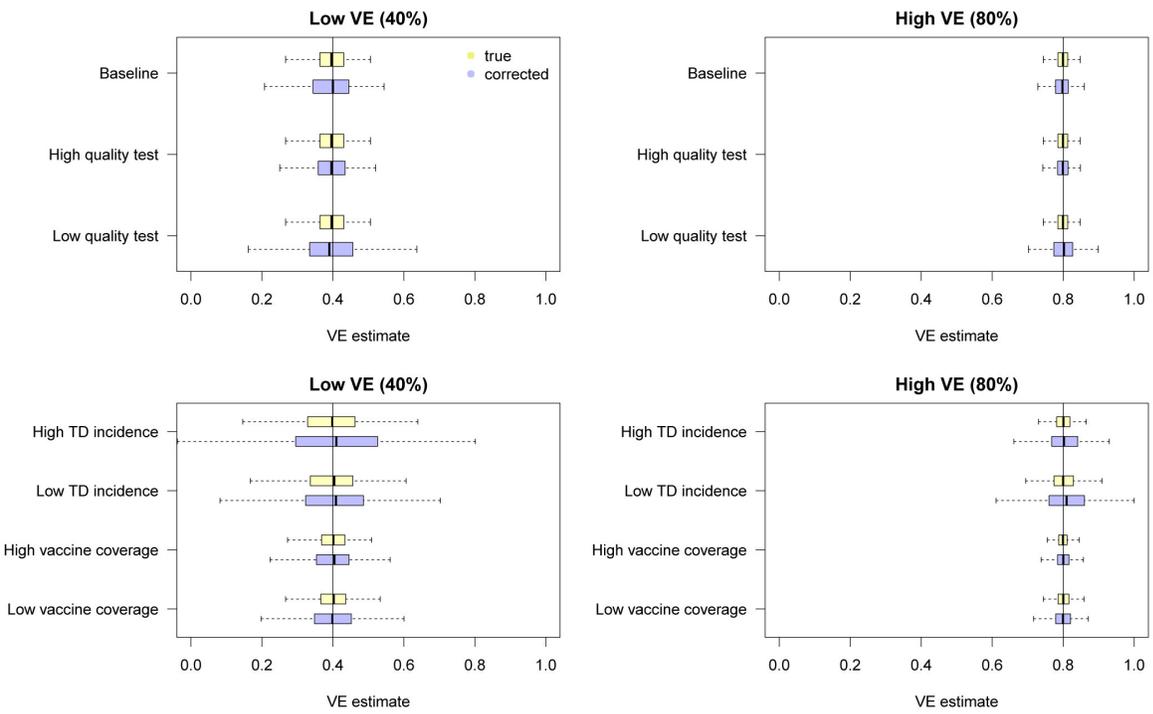
Figure S3: Uncertainty in VE estimates obtained from the true/misclassified datasets in the multivariate setting. The distributions of VE estimates from the simulated true (yellow) and misclassified (light blue) data are shown. The direct likelihood method was employed to correct biases in the misclassified data.

6 Paper 3: Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1700902	Title	Dr
First Name(s)	Akira		
Surname/Family Name	Endo		
Thesis Title	Roles of heterogeneity in infectious disease epidemiology: implications on dynamics, inference and control of influenza and COVID-19		
Primary Supervisor	Sebastian Funk		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Wellcome Open Research		
When was the work published?	9 April 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate conceived the study, designed the model, performed the analysis and wrote the original draft of the manuscript.

SECTION E

Student Signature	Akira Endo
Date	10 December 2020

Supervisor Signature	Sebastian Funk
Date	10 December 2020

Check for updates

RESEARCH ARTICLE

REVISED **Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China** [version 3; peer review: 2 approved]

Akira Endo [iD] [1,2],
Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group,
Sam Abbott [iD] [1,3], Adam J. Kucharski[1,3], Sebastian Funk [iD] [1,3]

[1]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK
[2]The Alan Turing Institute, London, NW1 2DB, UK
[3]Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK

Abstract
Background: A novel coronavirus disease (COVID-19) outbreak has now spread to a number of countries worldwide. While sustained transmission chains of human-to-human transmission suggest high basic reproduction number R_0, variation in the number of secondary transmissions (often characterised by so-called superspreading events) may be large as some countries have observed fewer local transmissions than others.

Methods: We quantified individual-level variation in COVID-19 transmission by applying a mathematical model to observed outbreak sizes in affected countries. We extracted the number of imported and local cases in the affected countries from the World Health Organization situation report and applied a branching process model where the number of secondary transmissions was assumed to follow a negative-binomial distribution.

Results: Our model suggested a high degree of individual-level variation in the transmission of COVID-19. Within the current consensus range of R_0 (2-3), the overdispersion parameter k of a negative-binomial distribution was estimated to be around 0.1 (median estimate 0.1; 95% CrI: 0.05-0.2 for R0 = 2.5), suggesting that 80% of secondary transmissions may have been caused by a small fraction of infectious individuals (~10%). A joint estimation yielded likely ranges for R_0 and k (95% CrIs: R_0 1.4-12; k 0.04-0.2); however, the upper bound of R_0 was not well informed by the model and data, which did not notably differ from that of the prior distribution.

Open Peer Review

Reviewer Status ✔ ✔

	Invited Reviewers	
	1	**2**
version 3 (revision) 10 Jul 2020	✔ report	
version 2 (revision) 03 Jul 2020	✔ report	✔ report
version 1 09 Apr 2020	✔ report	? report

1. **Lin Wang** [iD], Institut Pasteur, Paris, France

2. **Kaiyuan Sun**, National Institutes of Health, Bethesda, USA

Any reports and responses or comments on the article can be found at the end of the article.

Conclusions: Our finding of a highly-overdispersed offspring distribution highlights a potential benefit to focusing intervention efforts on superspreading. As most infected individuals do not contribute to the expansion of an epidemic, the effective reproduction number could be drastically reduced by preventing relatively rare superspreading events.

Keywords
COVID-19, SARS-CoV-2, novel coronavirus, overdispersion, superspreading, branching process

This article is included in the Coronavirus (COVID-19) collection.

Introduction

A novel coronavirus disease (COVID-19) outbreak, which is considered to be associated with a market in Wuhan, China, is now affecting a number of countries worldwide[1,2]. A substantial number of human-to-human transmission has occurred; the basic reproduction number R_0 (the average number of secondary transmissions caused by a single primary case in a fully susceptible population) has been estimated around 2–3[3–5]. More than 100 countries have observed confirmed cases of COVID-19. A few countries have already been shifting from the containment phase to the mitigation phase[6,7], with a substantial number of locally acquired cases (including those whose epidemiological link is untraceable). On the other hand, there are countries where a number of imported cases were ascertained but fewer secondary cases have been reported than might be expected with an estimated value of R_0 of 2–3.

This suggests that not all symptomatic cases cause a secondary transmission, which was also estimated to be the case for past coronavirus outbreaks (SARS/MERS)[8,9]. High individual-level variation (i.e. overdispersion) in the distribution of the number of secondary transmissions, which can lead to so-called superspreading events, is crucial information for epidemic control[9]. High variation in the distribution of secondary cases suggests that most cases do not contribute to the expansion of the epidemic, which means that containment efforts that can prevent superspreading events have a disproportionate effect on the reduction of transmission.

We estimated the level of overdispersion in COVID-19 transmission by using a mathematical model that is characterised by R_0 and the overdispersion parameter k of a negative binomial branching process. We fit this model to worldwide data on COVID-19 cases to estimate k given the reported range of R_0 and interpret this in the context of superspreading.

Methods

Data source

We extracted the number of imported/local cases in the affected countries (Table 1) from the WHO situation report 38[10] published on 27 February 2020, which was the latest report of the number of imported/local cases in each country (as of the situation report 39, WHO no longer reports the number of cases stratified by the site of infection). As in the WHO situation reports,

we defined imported cases as those whose likely site of infection is outside the reporting country and local cases as those whose likely site of infection is inside the reporting country. Those whose site of infection was under investigation were excluded from the analysis (Estonia had no case with a known site of infection and was excluded). In Egypt and Iran, no imported cases have been confirmed, which cause the likelihood value to be zero; data in these two countries were excluded. To distinguish between countries with and without an ongoing outbreak, we extracted daily case counts from an online resource[11] and determined the dates of the latest case confirmation for each country (as of 27 February).

Model

Assuming that the offspring distributions (distribution of the number of secondary transmissions) for COVID-19 cases are identically- and independently-distributed negative-binomial distributions, we constructed the likelihood of observing the reported number of imported/local cases (outbreak size) of COVID-19 for each country. The probability mass function for the final cluster size resulting from s initial cases is, according to Blumberg $et\ al.$[12], given by

$$c(x;s) = P(X = x;s) = \frac{ks}{kx + x - s}\binom{kx + x - s}{x - s}\frac{\left(\frac{R_0}{k}\right)^{x-s}}{\left(1 + \frac{R_0}{k}\right)^{kx+x-s}}.$$

If the observed case counts are part of an ongoing outbreak in a country, cluster sizes may grow in the future. To address this issue, we adjusted the likelihood for those countries with ongoing outbreak by only using the condition that the final cluster size of such a country has to be larger than the currently observed number of cases. The corresponding likelihood function is

$$c_o(x;s) = P(X \geq x;s) = 1 - \sum_{m=0}^{x-1} c(m;s),$$

with a convention $\sum_{m=0}^{-1} c(m;s) = 0$. We assumed that the growth of a cluster in a country had ceased if 7 days have passed since the latest reported case (denoted by set A). We applied the final size likelihood $c(x;\ s)$ to those countries and $c_o(x;\ s)$ to the rest of the countries (countries with an ongoing outbreak: B). The total likelihood is

$$L(R_0, k) = \prod_{i \in A} P(X = x_i; s_i) \prod_{i \in B} P(X \geq x_i; s_i).$$

Statistical analysis

Varying the assumed R_0 between 0–5 (fixed at an evenly-spaced grid of values), we estimated the overdispersion parameter k using the likelihood function described above. We used the Markov-chain Monte Carlo (MCMC) method to provide 95% credible intervals (CrIs). The reciprocal of k was sampled where the prior distribution for the reciprocal was weakly-informed half-normal (HalfNormal(σ = 10)). We employed the adaptive hit-and-run Metropolis algorithm[13]

Table 1. The number of confirmed COVID-19 cases reported (as of 27 February 2020).

Country	Total cases	Imported cases	Local cases	Site of infection unknown	Deaths	Latest date of case confirmation
South Korea	1766	17	605	1144	13	27/02/2020
Japan	186	39	129	18	3	27/02/2020
Singapore	93	24	69	0	0	27/02/2020
Australia	23	20	3	0	0	26/02/2020
Malaysia	22	20	2	0	0	27/02/2020
Vietnam*	16	8	8	0	0	13/02/2020
Philippines*	3	3	0	0	1	05/02/2020
Cambodia*	1	1	0	0	0	30/01/2020
Thailand	40	23	7	10	0	26/02/2020
India*	3	3	0	0	0	03/02/2020
Nepal*	1	1	0	0	0	24/01/2020
Sri Lanka	1	1	0	0	0	27/01/2020
USA	59	56	2	1	0	26/02/2020
Canada	11	9	1	1	0	27/02/2020
Brazil	1	1	0	0	0	26/02/2020
Italy	400	3	121	276	12	27/02/2020
Germany	21	3	14	4	0	27/02/2020
France	18	8	7	3	2	27/02/2020
UK	13	12	1	0	0	27/02/2020
Spain	12	10	1	1	0	27/02/2020
Croatia	3	2	1	0	0	26/02/2020
Austria	2	2	0	0	0	27/02/2020
Finland	2	2	0	0	0	26/02/2020
Israel	2	2	0	0	0	27/02/2020
Russia*	2	2	0	0	0	31/01/2020
Sweden	2	2	0	0	0	27/02/2020
Belgium*	1	1	0	0	0	04/02/2020
Denmark	1	1	0	0	0	27/02/2020
Estonia†	1	0	0	1	0	27/02/2020
Georgia	1	1	0	0	0	26/02/2020
Greece	1	1	0	0	0	27/02/2020
North Macedonia	1	1	0	0	0	26/02/2020
Norway	1	1	0	0	0	27/02/2020
Romania	1	1	0	0	0	26/02/2020
Switzerland	1	1	0	0	0	27/02/2020
Iran†	141	0	28	113	22	27/02/2020
Kuwait	43	43	0	0	0	27/02/2020
Bahrain	33	33	0	0	0	26/02/2020
UAE	13	8	5	0	0	27/02/2020
Iraq	6	6	0	0	0	27/02/2020
Oman	4	4	0	0	0	27/02/2020
Lebanon	1	1	0	0	0	27/02/2020
Pakistan	2	1	0	1	0	26/02/2020
Afghanistan	1	1	0	0	0	24/02/2020
Egypt*†	1	0	1	0	0	14/02/2020
Algeria	1	1	0	0	0	25/02/2020

* Countries considered to be without an ongoing outbreak

† Countries excluded from analysis

and obtained 500 thinned samples from 10,000 MCMC steps (where the first half of the chain was discarded as burn-in). We confirmed that the final 500 samples have an effective sample size of at least 300, indicating sufficiently low auto-correlation.

We also performed a joint-estimation of R_0 and k by the MCMC method (with a weakly-informed normal prior $N(\mu = 3, \sigma = 5)$ for R_0 and the weakly-informed half-normal prior (HalfNormal($\sigma = 10$)) for the reciprocal of k.

Statistical analysis was implemented in R-3.6.1 with a package {LaplacesDemon}-16.1.1. The reproducible code for this study is available on GitHub[14].

Proportion responsible for 80% of secondary transmissions

Using the estimated R_0 and k, we computed the estimated proportion of infected individuals responsible for 80% of the total secondary transmissions. Such proportion $p_{80\%}$ is given as

$$1 - p_{80\%} = \int_0^X \text{NB}\left(\lfloor x \rfloor; k, \frac{k}{R_0 + k}\right) dx,$$

where X satisfies

$$1 - 0.8 = \frac{1}{R_0} \int_0^X \lfloor x \rfloor \text{NB}\left(\lfloor x \rfloor; k, \frac{k}{R_0 + k}\right) dx.$$

Here, $\text{NB}\left(x; k, \frac{k}{R_0 + k}\right)$ represents the probability mass of a negative-binomial distribution with a mean R_0 and an overdispersion parameter k. This calculation is eased by the following rearrangement:

$$\frac{1}{R_0} \int_0^X \lfloor x \rfloor \text{NB}\left(\lfloor x \rfloor; k, \frac{k}{R_0 + k}\right) dx = \int_0^{X-1} \text{NB}\left(\lfloor x \rfloor; k+1, \frac{k}{R_0 + k}\right) dx.$$

We computed $p_{80\%}$ for each MCMC (Markov-chain Monte Carlo) sample to yield median and 95% CrIs.

Model comparison with a Poisson branching process model

To test if our assumption of overdispersed offspring distribution better describes the data, we compared our negative-binomial branching process model with a Poisson branching process model, which assumes that the offspring distribution follows a Poisson distribution instead of negative-binomial. Since a negative-binomial distribution converges to a Poisson distribution as $k \rightarrow \infty$, we approximately implemented a Poisson branching process model by fixing k of the negative-binomial model at 10^{10}. We compared the two models by the widely-applicable Bayesian information criterion (WBIC)[15].

### Simulation of the effect of underreporting

We used simulations to investigate potential bias caused by underreporting, one of the major limitations of the present study. Underreporting in some countries may be more frequent than others because of limited surveillance and/or testing capacity, causing heterogeneity in the number of cases that could have affected the estimated overdispersion. See *Extended data* (Supplementary materials)[16] for detailed methods.

### The effect of a differential reproduction number for imported cases

Due to interventions targeting travellers (e.g. screening and quarantine), the risk of transmission from imported cases may be lower than that from local cases. As part of the sensitivity analysis in *Extended data*, we estimated k assuming that the reproduction number of imported cases is smaller than that of local cases.

Results

Our estimation suggested substantial overdispersion ($k \ll 1$) in the offspring distribution of COVID-19 (Figure 1A and Figure 2). Within the current consensus range of R_0 (2–3), k was estimated to be around 0.1 (median estimate 0.1; 95% CrI: 0.05–0.2 for $R_0 = 2.5$). For the R_0 values of 2–3, the estimates suggested that 80% of secondary transmissions may

(A)

(B)
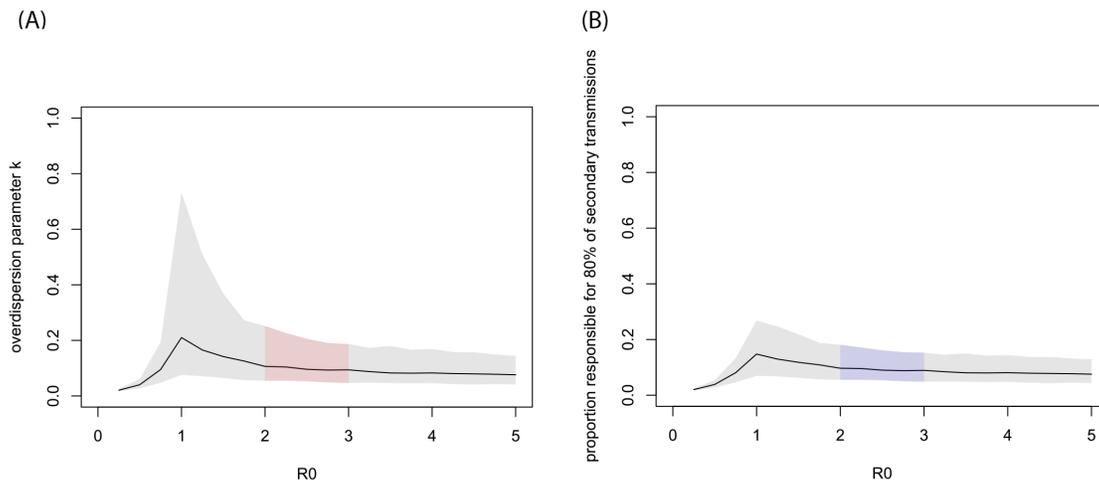
Figure 1. MCMC estimates given assumed R_0 values. (A) Estimated overdispersion parameter for various basic reproduction number R_0. **(B)** The proportion of infected individuals responsible for 80% of the total secondary transmissions ($p_{80\%}$). The black lines show the median estimates given fixed R_0 values and the grey shaded areas indicate 95% CrIs. The regions corresponding to the likely range of R_0 (2–3) are indicated by colour.

(A)

(B)
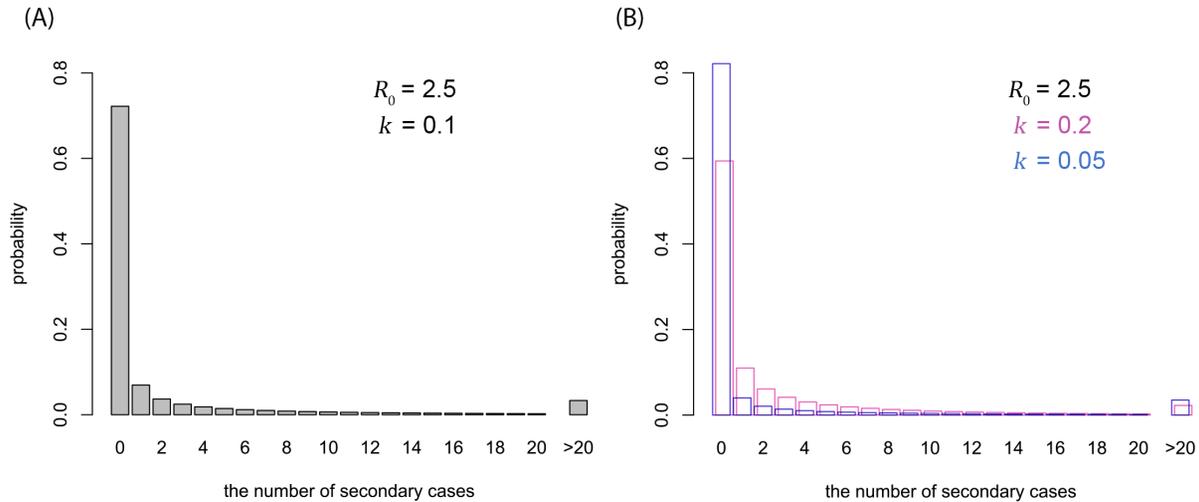
Figure 2. Possible offspring distributions of COVID-19. (A) Offspring distribution corresponding to $R_0 = 2.5$ and $k = 0.1$ (median estimate). **(B)** Offspring distribution corresponding to $R_0 = 2.5$ and $k = 0.05$ (95% CrI lower bound), 0.2 (upper bound). The probability mass functions of negative-binomial distributions are shown.

have been caused by a small fraction of infectious individuals (~10%; Figure 1B).

The result of the joint estimation suggested the likely bounds for R_0 and k (95% CrIs: R_0 1.4–12; k 0.04–0.2). The upper bound of R_0 did not notably differ from that of the prior distribution (=13.5), suggesting that our model and the data only informed the lower bound of R_0. This was presumably because the contribution of R_0 to the shape of a negative-binomial distribution is marginal when k is small (*Extended data*, Figure S1)[16]. A scatterplot (*Extended data*, Figure S2)[16] exhibited a moderate correlation between R_0 and k (correlation coefficient -0.4).

Model comparison between negative-binomial and Poisson branching process models suggested that a negative-binomial model better describes the observed data; WBIC strongly supported the negative-binomial model with a difference of 11.0 (Table 2). The simulation of the effect of underreporting suggested that possible underreporting is unlikely to cause underestimation of overdispersion parameter k (*Extended data*, Figure S3)[16]. A slight increase in the estimate of k was observed

Table 2. Model comparison between negative-binomial and Poisson branching process models.

Model	Parameter 95% CrIs		WBIC	ΔWBIC
	R_0	k		
Negative-binomial	1.4–12	0.04-0.2	45.6	0
Poisson	0.95–1.2	10^{10} (fixed)	56.6	11.0

when the reproduction number for imported cases was assumed to be lower due to interventions (*Extended data*, Table S1).

Discussion

Our results suggested that the offspring distribution of COVID-19 is highly overdispersed. For the likely range of R_0 of 2–3, the overdispersion parameter k was estimated to be around 0.1, suggesting that the majority of secondary transmission may be caused by a very small fraction of individuals (80% of transmissions caused by ~10% of the total cases). These results are consistent with a number of observed superspreading events observed in the current COVID-19 outbreak[17], and also in line with the estimates from the previous SARS/MERS outbreaks[8].

The overdispersion parameter for the current COVID-19 outbreak has also been estimated by stochastic simulation[18] and from contact tracing data in Shenzhen, China[19]. The former study did not yield an interpretable estimate of k due to the limited data input. In the latter study, the estimates of R_e (the effective reproduction number) and k were 0.4 (95% confidence interval: 0.3–0.5) and 0.58 (0.35–1.18), respectively, which did not agree with our findings. However, these estimates were obtained from pairs of cases with a clear epidemiological link and therefore may have been biased (downward for R_0 and upward for k) if superspreading events had been more likely to be missed during the contact tracing.

Although cluster size distributions based on a branching process model are useful in inference of the offspring distribution from limited data[12,20], they are not directly applicable to an ongoing outbreak because the final cluster size may not yet have been observed. In our analysis, we adopted an alternative approach which accounts for possible future

growth of clusters to minimise the risk of underestimation. As of 27 February 2020, the majority of the countries in the dataset had ongoing outbreaks (36 out of 43 countries analysed, accounting for 2,788 cases of the total 2,816). Even though we used the case counts in those countries only as the lower bounds of future final cluster sizes, which might have only partially informed of the underlying branching process, our model yielded estimates with moderate uncertainty levels (at least sufficient to suggest that k may be below 1). Together with the previous finding suggesting that the overdispersion parameter is unlikely to be biased downwards[21], we believe our analysis supports the possibility of highly-overdispersed transmission of COVID-19.

A number of limitations need to be noted in this study. We used the confirmed case counts reported to WHO and did not account for possible underreporting of cases. Heterogeneities between countries in surveillance and intervention capacities, which might also be contributing to the estimated overdispersion, were not considered (although we investigated such effects by simulations; see *Extended data*, Figure S3)[16]. Reported cases whose site of infection classified as unknown, which should in principle be counted as either imported or local cases, were excluded from analysis. Some cases with a known site of infection could also have been misclassified (e.g., cases with travel history may have been infected locally). The distinction between countries with and without ongoing outbreak (7 days without any new confirmation of cases) was arbitrary. However, we believe that our conclusion is robust because the distinction does not change with different thresholds (4–14 days), within which the serial interval of SARS-CoV-2 is likely to fall[22,23].

Our finding of a highly-overdispersed offspring distribution suggests that there is benefit to focusing intervention efforts on superspreading. As most infected individuals do not contribute to the expansion of transmission, the effective reproduction number could be drastically reduced by preventing relatively rare superspreading events. Identifying characteristics of settings that could lead to superspreading events will play a key role in designing effective control strategies.

Data availability
Source data
Zenodo: Extended data: Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. https://doi.org/10.5281/zenodo.3740348[16].

This project contains the following source data taken from references 10 and 11:
- bycountries_27Feb2020.csv. (Imported/local case counts by country from WHO situation report 38[10].)
- dailycases_international_27Feb2020.csv. (Daily case counts by country from COVID2019.app[11].)

Extended data
Zenodo: Extended data: Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. https://doi.org/10.5281/zenodo.3911576[16].

This project contains the following extended data
- supplementarymaterials.pdf. (Supplementary material: Estimating the amount of superspreading using outbreak sizes of COVID-19 outside China.)
- figS1.tif. (Figure S1. Offspring distributions for different R_0 values. The probability mass functions of negative-binomial distributions are shown. The overdispersion parameter k is fixed at 0.1.)
- figS2.tif. (Supplementary Figure 2. Scatter plot of MCMC samples from a joint estimation of R_0 and k. The dotted line represents the threshold $R_0 = 1$)
- figS3.tif. (Supplementary Figure 3. Estimates of overdispersion from simulations with underreporting. (A) Maximum-likelihood estimates (MLEs) of overdispersion parameter k with different distributions for country-specific reporting probability q_i (including constant $q_i = 1$). Both imported and local cases are assumed to be reported at probability q_i in country i. The blue dotted line indicates the true value $k = 0.1$. (B) MLEs where imported cases were assumed to be fully reported and local cases were reported at probability q_i. (C) Probability density functions for beta distributions used in the simulation.)

Code availability
The reproducible code is available at: https://github.com/akira-endo/COVID19_clustersize.

Archived code at time of publication: https://doi.org/10.5281/zenodo.3741743[14].

License: MIT.

References

1. Zhu N, Zhang D, Wang W, *et al.*: **A Novel Coronavirus from Patients with Pneumonia in China, 2019.** *N Engl J Med.* 2020; **382**(8): 727–733.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Lai CC, Shih TP, Ko WC, *et al.*: **Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges.** *Int J Antimicrob Agents.* 2020; **55**(3): 105924.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Zhao S, Lin Q, Ran J, *et al.*: **Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak.** *Int J Infect Dis.* 2020; **92**: 214–217.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. Zhang S, Diao M, Yu W, *et al.*: **Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis.** *Int J Infect Dis.* 2020; **93**: 201–204.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Abbott S, Hellewell J, Munday J, *et al.*: **The transmissibility of novel Coronavirus in the early stages of the 2019-20 outbreak in Wuhan: Exploring initial point-source exposure sizes and durations using scenario analysis [version 1; peer review: 1 approved].** *Wellcome Open Res.* 2020; **5**: 17.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Headquarters for Novel Coronavirus Disease Control; Ministry of Health Labour and Welfare: **Basic Policies for Novel Coronavirus Disease Control.** 2020.
   Reference Source

7. Department of Health and Social Care, Hancock M: **Press release: Government outlines further plans to support health and social care system in fight against COVID-19.** 2020; [cited 9 Mar 2020].
   Reference Source

8. Kucharski AJ, Althaus CL: **The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission.** *Euro Surveill.* 2015; **20**(25): 14–8.
   PubMed Abstract | **Publisher Full Text**

9. Lloyd-Smith JO, Schreiber SJ, Kopp PE, *et al.*: **Superspreading and the effect of individual variation on disease emergence.** *Nature.* 2005; **438**(7066): 355–359.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

10. World Health Organization: **Coronavirus disease 2019 (COVID-19) Situation Report – 38.** 2020.
    Reference Source

11. **COVID2019.app - LIVE stats and graphs.** 2020; [cited 4 Mar 2020].
    Reference Source

12. Blumberg S, Funk S, Pulliam JR: **Detecting differential transmissibilities that affect the size of self-limited outbreaks.** Wilke CO, editor. *PLoS Pathog.* 2014; **10**(10): e1004452.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. Chen MH, Schmeiser B: **Performance of the Gibbs, Hit-and-Run, and Metropolis Samplers.** *J Comput Graph Stat.* 1993; **2**(3): 251–272.
    Publisher Full Text

14. Endo A, Abbott S, Kucharski AJ, *et al.*: **Estimating the amount of superspreading using outbreak sizes of COVID-19 outside China (Version v1.0.0).** *Zenodo.* 2020.
    http://www.doi.org/10.5281/zenodo.3741743

15. Watanabe S: **A Widely Applicable Bayesian Information Criterion.** 2013; **14**: 867–897.
    Publisher Full Text

16. Endo A, Abbott S, Kucharski AJ, *et al.*: **Extended data: Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China.** *Zenodo.* 2020.
    http://www.doi.org/10.5281/zenodo.3911576

17. Liu Y, Eggo RM, Kucharski AJ: **Secondary attack rate and superspreading events for SARS-CoV-2.** *Lancet.* 2020; **395**(10227): e47.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Riou J, Althaus CL: **Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020.** *Euro Surveill.* 2020; **25**(4): 2000058.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Bi Q, Wu Y, Mei S, *et al.*: **Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts.** *medRxiv.* 2020; 2020.03.03.20028423.
    Publisher Full Text

20. Blumberg S, Lloyd-Smith JO: **Inference of R(0) and transmission heterogeneity from the size distribution of stuttering chains.** Ferguson N, editor. *PLoS Comput Biol.* 2013; **9**(5): e1002993.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Lloyd-Smith JO: **Maximum Likelihood Estimation of the Negative Binomial Dispersion Parameter for Highly Overdispersed Data, with Applications to Infectious Diseases.** Rees M, editor. *PLoS One.* 2007; **2**(2): e180.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Li Q, Guan X, Wu P, *et al.*: **Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia.** *N Engl J Med.* 2020; **382**(13): 1199–1207.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. Nishiura H, Linton NM, Akhmetzhanov AR: **Serial interval of novel coronavirus (COVID-19) infections.** *Int J Infect Dis.* 2020; **93**: 284–286.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Supplementary material: Estimating the amount of superspreading using outbreak sizes of COVID-19 outside China

Akira Endo, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Sam Abbott, Adam J Kucharski, Sebastian Funk

1. Negative-binomial offspring distributions for different R_0 values

We compared negative-binomial offspring distributions for different R_0 values where the overdispersion parameter k is fixed at 0.1 (Figure S1). When k is small, different R_0 values barely change the offspring distribution except for the mass for 0 and for large (> 20) secondary cases.
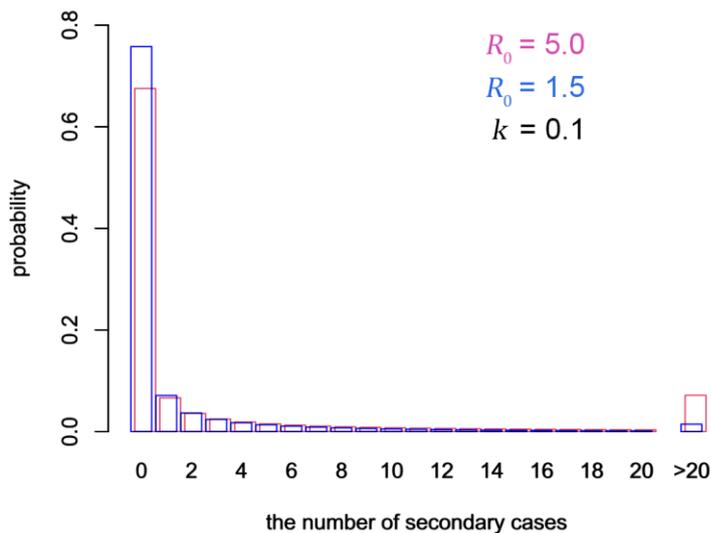

Figure S1. Offspring distributions for different R_0 values. The probability mass functions of negative-binomial distributions are shown. The overdispersion parameter k is fixed at 0.1.

2. Joint estimation of R_0 and k

We performed a joint-estimation of R_0 and k by the MCMC method (with a weakly-informed normal prior $N(\mu = 3, \sigma = 5)$ for R_0 to prevent a divergence; the prior for the reciprocal of k was a weakly-informed half-normal (HalfNormal($\sigma = 10$)). The estimated range of R_0 was wide (median 4.4; 95% CrI 1.4-12)

and the upper bound did not notably differ from that of the prior distribution (=13.5). The estimated range of k was low (median 0.08; 95% CrI 0.04-0.2), suggesting a highly heterogeneous offspring distribution. A scatterplot (Figure S2) exhibited a moderate correlation between R_0 and k (correlation coefficient 0.4).
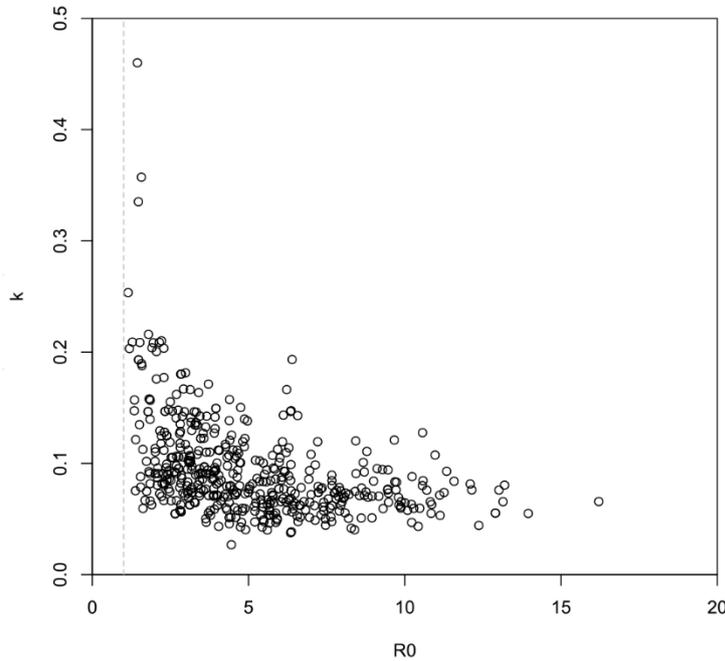

Figure S2. Scatter plot of MCMC samples from a joint estimation of R_0 and k. The dotted line represents the threshold $R_0 = 1$.

3. Simulation of the effect of underreporting

One of the major limitations of the present study is potential underreporting in the dataset. Underreporting in some countries may be more frequent than others because of limited surveillance and/or testing capacity, causing heterogeneity in the number of cases that could have affected the estimated overdispersion. An existing study suggested 38% as an optimistic global estimate of the detection probability for imported cases from Wuhan, China, with a substantial variation between countries [1]. We used simulations to investigate potential bias caused by underreporting. First, we assumed that the same probability of reporting applies to both imported and local cases in a country. We represented the data-generating process in the presence of underreporting as a binomial sampling. Let s_i and x_i^0 be the observed and true number of cases in country i, respectively.

$$s_i \sim \text{Binom}\big(x_i^0, q_i\big),$$

where q_i is the reporting probability for country i. When s_i is observed, by assuming that the prior probability for x_i^0 is (improper-) uniformly distributed, we get

$$x_i^0 - s^i \sim \text{NegBinom}(s_i + 1, q_i). \qquad (*)$$

We generated simulation datasets in the following steps.

1. Set s_i as the number of observed imported cases from the WHO situation report (Table 1 in the main text); sample reporting probability q_i for each country from a beta distribution (see Figure S3C) and then sample x_i^0 based on Equation $(*)$.

2. Sample two generations of cases where x_i^0 is the number of index cases using a negative-binomial-distrusted offspring distribution. Namely, for $t = 1,2$,

    $$x_i^t \sim \text{NegBinom}\left(kx_i^{t-1}, \frac{k}{R_0 + k}\right).$$

    We used $R_0 = 2.5$ and $k = 0.1$ for our simulations.

3. Sample the observed number of local cases by binomial sampling:

    $$X_i^t \sim \text{Binom}(x_i^t, q_i). \ (t = 1,2)$$

4. Apply the likelihood-based model in the main text to the observed imported/local cases: $\big(s_i, X_i^1 + X_i^2\big)$, where countries with non-zero X_i^2 are treated as "countries with an ongoing outbreak".

We used the maximum-likelihood approach here (as opposed to MCMC used in the main analysis) for simplicity. $R_0 = 2.5$ was assumed to be known and overdispersion parameter k was estimated. We ran 500 simulations for each assumed distribution of q_i and plotted the estimates (Figure S3A). Lower reporting probability introduced an upward bias in the estimates.

Next, we repeated the simulation with another scenario where the imported cases were assumed to be fully reported (100% reporting probability for imported cases) due to their awareness of the travel history. This can be implemented by skipping step 1 and using s_i as x_i^0. The degree of bias introduced in this simulation was relatively small (Figure S3B).
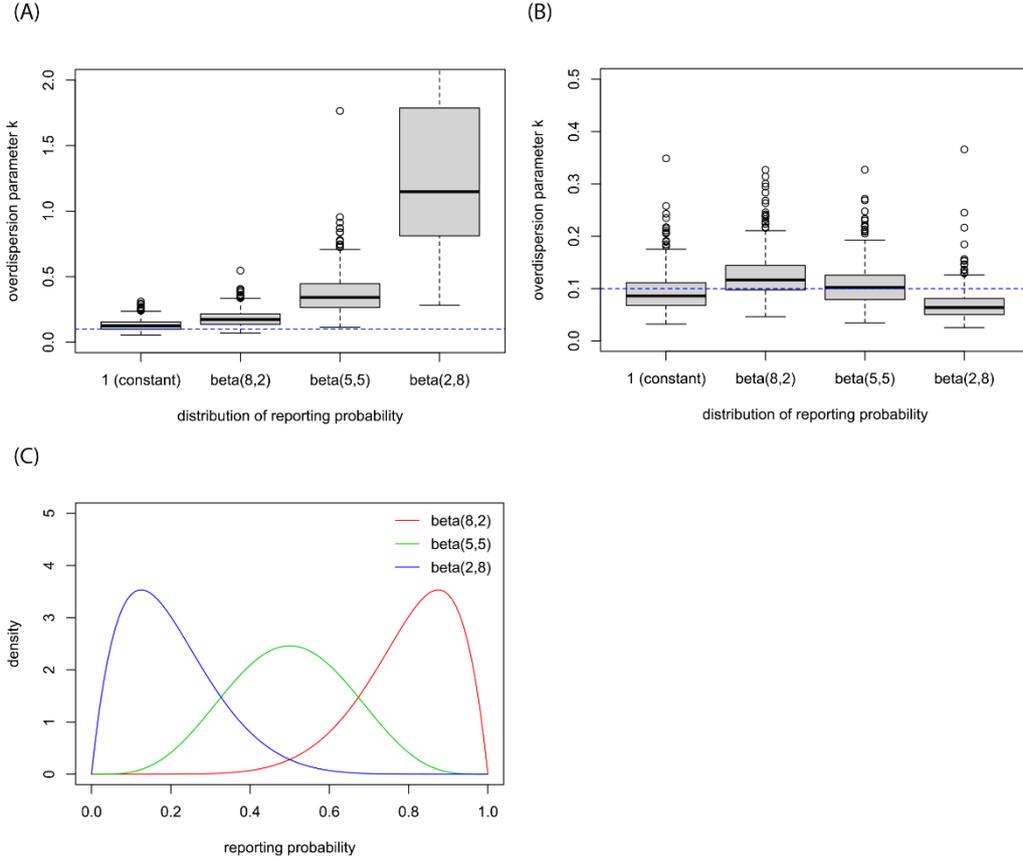
Figure S3. Estimates of overdispersion from simulations with underreporting. (A) Maximum-likelihood estimates (MLEs) of overdispersion parameter k with different distributions for country-specific reporting probability q_i (including constant $q_i = 1$). Both imported and local cases are assumed to be reported at probability q_i in country i. The blue dotted line indicates the true value $k = 0.1$. (B) MLEs where imported cases were assumed to be fully reported and local cases were reported at probability q_i. (C) Probability density functions for beta distributions used in the simulation.

4. The effect of a differential reproduction number for imported cases

Due to interventions targeting travellers (e.g. screening and quarantine), the risk of transmission from imported cases may be lower than that from local cases. To account for the effect of a differential reproduction number for imported cases, we modified the likelihood function $c(x; s)$ in the main text as

$$c_{\mathrm{I}}(x; s) = \sum_{j=0}^{x} \mathrm{NB}(j; ks, \mu = sR_{\mathrm{I}})c(x - s; j),$$

where $NB(j; ks, \mu = sR_I)$ is a negative binomial distribution with an overdispersion parameter ks and a mean sR_I., which corresponds to the distribution of the total number of secondary cases generated by s imported cases with an effective reproduction number R_I. Assuming that R_0 for local cases is 2.5, we estimated k for three R_I values: 0.5, 0.8 and 1.2. We found that the estimates of k were higher than our baseline estimates ($k = 0.1$) when R_I is below 1 ($R_I = 0.5, 0.8$), whereas the estimate for $R_I = 1.2$ was not very distinct from the baseline result (Table S1).

Table S1. The median estimates and 95% CrIs of the overdispersion parameter k with differential effective reproduction numbers for imported cases.

Assumed reproduction number		Estimated overdispersion parameter (k)
Imported cases (R_I)	Local cases (R_0)	
0.5	2.5	0.29 (0.10-1.24)
0.8	2.5	0.18 (0.08-0.54)
1.2	2.5	0.14 (0.06-0.32)

Reference

1.   Niehus R, De Salazar PM, Taylor AR, Lipsitch M. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. Lancet Infect Dis. 2020. doi:10.1016/S1473-3099(20)30229-2

6.3 Additional notes: definition of cluster and parameter identifiability

6.3.1 Clarifications on the definition of 'cluster'

The term 'cluster' is often loosely defined in infectious disease epidemiology. Usually a cluster represent a group of cases that are associated with each other, e.g. cases arising from a common superspreading event. It is often assumed that cases belonging to the same cluster are linked by transmission and therefore these cases constitute a connected subset ('branch') of a transmission chain. However, the presence or absence of actual transmission between cases are not necessarily easily identified in practice, and a case that links multiple cases in a transmission tree can sometimes be unobserved. Moreover, when a cluster is defined based on association with specific settings, cases involved in onward transmissions can be excluded from the cluster even if there is a transmission link. Overall, it is not straightforward or practical to give a strict definition to 'clusters' in epidemiology, and therefore one needs to note what a cluster in a specific context refers to. In this paper, a cluster is defined as a collection of cases that are linked to the transmission chain(s) originating from any one of the initial cases. All reported cases in a country are considered as a cluster originating from the imported cases, regardless of what actual transmission links between them were. The likelihood for the final cluster size used in the estimation is valid even if the cluster consists of multiple independent transmission chains linked to separate initial cases.

6.3.2 Identifiability between R_0 and k in joint estimation

In the main analysis, overdispersion parameter k was estimated given fix R_0 because the available data did not allow for reliable joint estimation of R_0 and k as shown in Supplementary materials section 2. Here, the posterior samples for R_0 ranged from 1 to 15, and given that the upper bound almost corresponds to what is expected from the prior distribution alone, it is suggested that the data barely informed the upper bound of R_0. This can be explained as follows. In this analysis, data from the countries with an 'ongoing outbreak' was assumed to reflect only the lower limit of the final outbreak size. As the majority of the countries included for analysis at the time of writing were countries with ongoing outbreaks, very large R_0 and small k could also explain the observation; that is, observing a certain proportion of naturally extinct outbreaks (in countries highlighted with * in Table 1) does not contradict a large R_0 if k is small, and as the data in countries with ongoing outbreaks does not provide information on the upper bound R_0 of because the outbreaks in these countries can continue to grow onwards. In fact, when a wider prior distribution was used for R_0, the upper bound of posterior samples became even unrealistically greater (~200). One possible approach to preventing such inflation of R_0 without using informative prior may be using not only the outbreak sizes at time of data collection but also the

temporal information of the dataset, i.e. the number of reported cases by dates. The initial growth of reported cases in each country, combined with the estimated serial interval from elsewhere, could have provided additional constraints on the possible range of R_0. However, this approach can result in other limitations, e.g. uncertainty and delays in reporting and uncertainty in the estimated serial interval, which may diminish the strengths of the current simplistic approach. Moreover, this approach is unlikely to provide more reliable estimates than our original approach referencing the estimated R_0 from the initial growth of outbreaks in the previous studies (mostly from China).

7 Paper 4: Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks

LONDON
SCHOOL of
HYGIENE
&TROPICAL
MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4846
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1700902	Title	Dr
First Name(s)	Akira		
Surname/Family Name	Endo		
Thesis Title	Roles of heterogeneity in infectious disease epidemiology: implications on dynamics, inference and control of influenza and COVID-19		
Primary Supervisor	Sebastian Funk		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Wellcome Open Research	
When was the work published?	13 October 2020	
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A	
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Improving health worldwide   www.lshtm.ac.uk

Stage of publication	Choose an item.

SECTION D – Multi-authored work

| For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate conceived the study, designed the model, performed the analysis and wrote the original draft of the manuscript.

SECTION E

Student Signature	Akira Endo
Date	10 December 2020

Supervisor Signature	Sebastian Funk
Date	10 December 2020

Improving health worldwide   Page 2 of 2   www.lshtm.ac.uk

Check for updates

RESEARCH ARTICLE

REVISED **Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks** [version 3; peer review: 2 approved]

Akira Endo [iD] [1-3],
Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group,
Quentin J. Leclerc [iD] [1,3], Gwenan M. Knight [iD] [1,3], Graham F. Medley [iD] [3,4],
Katherine E. Atkins [iD] [1,3,5], Sebastian Funk [iD] [1,3], Adam J. Kucharski [1,3]

[1]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK
[2]The Alan Turing Institute, London, NW1 2DB, UK
[3]Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK
[4]Department of Global Health and Development, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK
[5]Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, EH16 4UX, UK

Abstract

Introduction: Contact tracing has the potential to control outbreaks without the need for stringent physical distancing policies, e.g. civil lockdowns. Unlike forward contact tracing, backward contact tracing identifies the source of newly detected cases. This approach is particularly valuable when there is high individual-level variation in the number of secondary transmissions (overdispersion).

Methods: By using a simple branching process model, we explored the potential of combining backward contact tracing with more conventional forward contact tracing for control of COVID-19. We estimated the typical size of clusters that can be reached by backward tracing and simulated the incremental effectiveness of combining backward tracing with conventional forward tracing.

Results: Across ranges of parameter values consistent with dynamics of SARS-CoV-2, backward tracing is expected to identify a primary case generating 3-10 times more infections than a randomly chosen case, typically increasing the proportion of subsequent cases averted by a factor of 2-3. The estimated number of cases averted by backward tracing became greater with a higher degree of overdispersion.

Conclusion: Backward contact tracing can be an effective tool for

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
version 3 (revision) 31 Mar 2021		
version 2 (revision) 18 Jan 2021	✓ report	
version 1 13 Oct 2020	✓ report	✓ report

1. **Edward Hill** [iD], University of Warwick, Coventry, UK

   University of Warwick, Coventry, UK

   University of Warwick, Coventry, UK

outbreak control, especially in the presence of overdispersion as is observed with SARS-CoV-2.

Keywords
COVID-19, SARS-CoV-2, contact tracing, backward tracing, overdispersion

2. **Amy Winter** iD , Johns Hopkins University, Baltimore, USA

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Coronavirus (COVID-19) collection.

Introduction

Isolation of symptomatic cases and tracing and quarantine of their contacts is a staple public health control measure, and has the potential to prevent the need for stringent physical distancing policies that result in detrimental impacts on the society (e.g., civil lockdowns)[1,2]. Contact tracing is typically triggered by a confirmed index case identified via symptom-based surveillance. Contacts of this index case are identified via interviews by public health officials (manual contact tracing) or by tracking proximity records on digital devices (digital contact tracing), and asked to quarantine in order to prevent further transmissions.

Contact tracing often targets 'downstream' individuals, who may have been infected by the index case ('forward tracing'); i.e. those who have been in contact with the index case after the index case likely became infectious (often assumed as 2 days before illness onset for COVID-19[3,4]). However, 'backward tracing' can also be used to identify the upstream primary case who infected the index case (or a setting or event at which the index case was infected) by retracing history of contact to the likely

point of exposure up to the upper bound of the incubation period. For example, contact history of 14 days prior to symptom onset is collected in Japan, where backward tracing has been operated from the early phase of the COVID-19 outbreak[5,6]. If this primary case is identified, a larger fraction of the transmission chain can be detected by forward tracing each of the contacts of this primary case.

Unlike forward tracing, backward tracing is more effective when the number of onward transmissions is highly variable, because index cases are disproportionately more likely to have been generated by primary cases who also infected others (an example of the "friendship paradox"[7–9]). Because there is evidence that the number of secondary transmissions of SARS-CoV-2 per case exhibits substantial individual-level variation (i.e. overdispersion), often resulting in so-called superspreading events[10–12], a large proportion of infections may be linked to a small proportion of original clusters. As a result, finding and targeting originating clusters in combination with reducing onwards infection may substantially enhance the effectiveness of tracing methods[9,13,14].

In the present study, using a simple branching process model, we explore the incremental effectiveness of combining 'backward' tracing with conventional 'forward' tracing in the presence of overdispersion in SARS-CoV-2 transmission.

Methods

Overdispersion and the coverage of contact tracing

We used a branching process model to compare the performance of forward and backward contact tracing triggered by an index case found by symptom-based surveillance (Figure 1). We enumerate generations of transmission chains linked to the

(A) Forward contact tracing only

(B) Forward + backward contact tracing
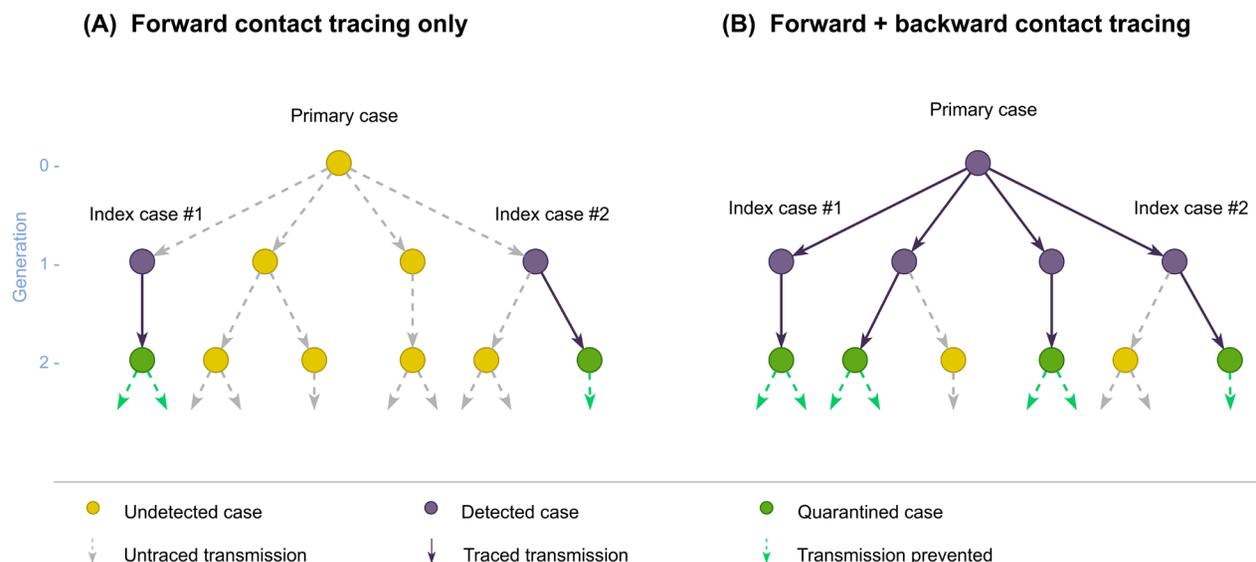
Figure 1. Schematic illustration of forward and backward contact tracing. Two cases (index cases #1 and #2) from a transmission tree originating from an (initially) undetected primary case are assumed to be detected by surveillance. Possible results of contact tracing are shown where (**A**) only forward tracing is performed or (**B**) both forward and backward tracing are performed. Some cases may remain undetected because contact tracing can miss cases.

index case so that the index case belongs to generation-1 (G1). Backward tracing first identifies the primary case (G0) that infected the index case and then applies forward tracing to those infected by the primary case (G1). We represent the transmission chains of COVID-19 by a branching process where $p(x)$ denotes the offspring distribution, i.e. the probability mass function of the number of secondary transmissions caused by a single case. If an individual is identified as a primary case, they are more likely to have generated more cases than any random case because the probability that a primary case is identified is proportional to the number of cases it generates. Therefore, the number of offspring of the identified primary case follows $p(x \mid G_0) = \dfrac{xp(x)}{\mathbb{E}(x)}$, where $\mathbb{E}(x) = \sum_{x=0}^{\infty} xp(x)$. The mean number of G1 cases able to be identified by backward tracing (including the index case) is $\quad \mathbb{E}(x \mid G_0) = \sum_{x=0}^{\infty} x \dfrac{xp(x)}{\mathbb{E}(x)} = \dfrac{\mathbb{E}(x^2)}{\mathbb{E}(x)} = R(1 + v^2),\quad$ where $\mathbb{E}(x) = R$ is the reproduction number and v is the coefficient of variation (the standard deviation of x divided by its mean). With a high overdispersion (large v), backward tracing of the index case can substantially increase the number of G1 cases to trace. Conversely, the mean number of cases that can be identified by forward tracing is R regardless of the degree of overdispersion.

When we assume $p(x)$ follows a negative-binomial distribution[11,15] with an overdispersion parameter k, backward tracing on average identifies $\quad \mathbb{E}(x \mid G_0) = R(1 + v^2) = 1 + R\left(1 + \dfrac{1}{k}\right)\quad$ G1 cases. Existing studies suggest k for SARS-CoV-2 transmission is small and likely to lie within the range of 0.1–0.5[11,16,17]. A small k indicates that the primary case identified through backward tracing typically generates more secondary cases than does a randomly selected case (i.e. $\mathbb{E}(x \mid G_0) > E(x) = R$).

The higher probability of identifying a large cluster by backward tracing can also be demonstrated by looking at the tail probability of the offspring distribution. Given a negative-binomial offspring distribution $p(x, R, k) = \dbinom{x + k - 1}{x}\left(\dfrac{R}{R + k}\right)^x \left(\dfrac{k}{R + k}\right)^k$, the probability that a cluster of secondary infections caused by a G0 case identified by backward tracing has a size of X or larger is

$$\sum_{x=X}^{\infty} p(x \mid G_0) = \frac{1}{R}\sum_{x=X}^{\infty} xp(x; R, k) = \sum_{x=X}^{\infty}\binom{x + k - 1}{x - 1}\left(\frac{R}{R + k}\right)^{x-1}\left(\frac{k}{R + k}\right)^{k+1}$$

$$= \sum_{x=X-1}^{\infty} p\left(x; \frac{k+1}{k}R, k+1\right).$$

For different combinations of the reproduction number R and overdispersion parameter k, we estimated the mean size of an identified cluster in backward tracing and the probability of observing a size of at least 5, 10 and 25.

Simulation of the effectiveness of forward and backward contact tracing

Using our simple branching process model with a negative-binomial offspring distribution, we assessed the potential effectiveness of forward and backward contact tracing. We assumed that contact tracing is triggered by the detection of an index case whose primary case is initially unknown so that our simulation would guide decision making at the operational level (i.e. whether it is worthwhile to implement contact tracing when a case is found). We compared two scenarios: forward tracing only and the combination of forward and backward tracing (Figure 1). In the forward only scenario, generation-2 (G2) cases resulting from an index case are potentially traced and quarantined; in the combined scenario, more G1 cases can be identified through backward tracing of the primary infection and thus a larger number of G2 cases can be traced and quarantined. As the infectious period of G1 cases is likely to have already passed when they are identified by contact tracing because tracing only starts after the index case is confirmed, we assumed that secondary transmissions caused by G1 cases would not be prevented and that only G2 cases successfully traced could be put in quarantine (which confers a relative reduction c in transmission). To account for potential limitations in the effectiveness of contact tracing, we assumed that the primary case is identified with probability b and that each offspring of identified cases are traced with probability q. G1 cases not traced may be independently found by symptom-based surveillance; we accounted for such independent case finding with a detection probability d (although we excluded backward tracing triggered by these cases from analysis), which is expected to be low due to frequent subclinical infections[18]. All parameters used for simulation are listed in Table 1.

We estimated the expected number of generation-3 (G3) cases averted and defined the effectiveness of contact tracing by the relative reduction in the total number of G3 cases. Assuming a negative-binomial branching process with a mean R and overdispersion parameter k, the mean total number of G3 cases given an index case found by surveillance is

$$C_3 = \sum_{x_0, x_1, x_2 = 0}^{\infty} x_0 x_1 x_2\, p(x_0 \mid G_0) p(x_1) p(x_2) = \mathbb{E}(x \mid G_0)\mathbb{E}(x)^2 = R^2\left(1 + R\left(1 + \frac{1}{k}\right)\right).$$

In the forward only scenario, the expected number of G1 cases excluding the initially found index case is $\mathbb{E}(x \mid G_0) - 1 = R\left(1 + \dfrac{1}{k}\right)$, of which proportion d is independently detected by symptom-based surveillance. Therefore, the total number of G1 cases targeted by forward tracing (including the index case) is $1 + Rd\left(1 + \dfrac{1}{k}\right)$.

Of the secondary cases generated by these G1 cases (R cases each on average), proportion q are successfully traced, i.e. $Rq(1+Rd(1+1/k))$ G2 cases are traced and asked to quarantine on average. The effective reproduction number of quarantined G2 cases is assumed to be $R(1-c)$; therefore, the estimated number of G3 cases averted is given as

$$\Delta_F = R^2 qc\left(1 + Rd\left(1 + \frac{1}{k}\right)\right).$$

Table 1. Parameter notations and values assumed in simulation.

Parameter	Notation	Assumed value in Figure 2 and *Extended data*, Figures S1 and S2[19]
Reproduction number	R	1.2, 2.5
Overdispersion parameter	k	0.2, 0.5
Relative reduction in transmission due to quarantine	c	0.2 – 1.0
Probability of identifying the primary (G0) case by backward tracing	b	0.5, 0.8
Probability of identifying each offspring of an already identified case	q	0.0– 1.0
Probability of a G1 case identified by surveillance independently of contact tracing	d	0.1, 0.2, 0.5

In the combined (forward + backward) scenario, G1 cases can also be detected by backward tracing. Of the mean $R\left(1+\dfrac{1}{k}\right)$ G1 cases potentially under the scope of backward tracing, a proportion $(1 - d)(1 - bq)$ will remain undetected either by backward tracing or independent detection. As a result, $(1-(1-d)(1-bq))$ $R(1+1/k)$ G1 cases are identified on average in addition to the index case, leading to tracing of $Rq(1+(1-(1-d)(1-bq))R(1+1/k))$ G2 cases. By asking these traced G2 cases to quarantine, G3 cases are expected to be averted by

$$\Delta_{F+B} = R^2 qc\left[1 + (1 - (1-d)(1-bq))R\left(1 + \frac{1}{k}\right)\right].$$

The effectiveness of contact tracing in the forward and combined scenarios are obtained as $\dfrac{\Delta_F}{C_3}$ and $\dfrac{\Delta_{F+B}}{C_3}$, respectively. The simulation was implemented in R-3.6.1. The replication code and *Extended data* are reposited on GitHub (https://github.com/akira-endo/COVID19_backwardtracing) and archived with Zenodo[19].

An earlier version of this article can be found on medRxiv (DOI: https://doi.org/10.1101/2020.08.01.20166595).

Results

Larger clusters are likely to be detected through backward tracing in the presence of overdispersion

The estimated mean and the tail probabilities of the secondary transmissions caused by a primary case identified via backward contact tracing suggests the potential strength of this tracing approach (Table 2). With a substantial individual-level variation in the number of secondary transmissions per case, characterised by a small overdispersion parameter k of a negative-binomial distribution ranging between 0.1–0.5, backward tracing typically leads to a primary case generating 3–10 times more infections than a randomly chosen case (whose mean defines the reproduction number R). The tail probabilities, ranging from 25%

to 88% for 5 or more offspring (Table 2), suggest that backward tracing is likely to find a relatively large cluster (≥5) under the plausible parameter settings. These values are striking because the probability of finding such clusters in forward tracing will be much lower. In a case of $R = 1.2$ and $k = 0.2$, only 6% of random cases results in 5 or more secondary infections, as opposed to 53% of primary cases identified by backward tracing.

Backward tracing typically results in multiple-fold increases in the overall effectiveness of contact tracing

Using a branching process model, we simulated the effectiveness of contact tracing. Across plausible ranges of parameter values, we found that introducing backward tracing in addition to forward tracing increased the effectiveness of contact tracing by a factor of 2–3 (Figure 2 and *Extended data*, S1 and S2[19]). Although the relative improvement in effectiveness by introducing backward tracing is similar between different values of k (0.2 and 0.5), the coverage of backward tracing scales up with overdispersion. We found that a higher degree of overdispersion (i.e. small k) resulted in a larger absolute number of cases averted by backward tracing (Figure 3 and *Extended data*, S3). In the presence of substantial overdispersion ($k = 0.1$), backward tracing is expected to avert 2–3 times more G3 cases than it does in a less-dispersed outbreak ($k = 0.5$).

Discussion

Using a simple branching process model, we showed that backward contact tracing has the potential to identify a large proportion of infections because of the observed overdispersion in COVID-19 transmission. For each index cases detected, forward tracing alone can, on average, identify at most the mean number of secondary infections (i.e. R). In contrast, backward tracing increases this maximum number of traceable individuals by a factor of 2–3, as index cases are more likely to come from clusters than a case is to generate a cluster. Furthermore, backward tracing contributes to epidemiological understanding of high-risk settings because transmission events with a common source are more likely to be identified. While standard tracing mostly

Table 2. Characteristics of transmissions from a primary case identified by backward contact tracing for different combinations of the reproduction number (R) and overdispersion parameter (k).

Reproduction number (R)	Overdispersion parameter (k)	Mean number of transmissions from primary case ($\mathbb{E}(x \mid G_0)$)	Probability ($x \geq 5 \mid G_0$)	Probability ($x \geq 10 \mid G_0$)	Probability ($x \geq 25 \mid G_0$)
	0.1	9.8	67%	39%	7%
	0.2	5.8	49%	18%	0.7%
0.8	0.3	4.5	38%	9%	0.1%
	0.4	3.8	30%	5%	0.02%
	0.5	3.4	25%	3%	0.003%
	0.1	14.2	77%	53%	17%
	0.2	8.2	62%	32%	4%
1.2	0.3	6.2	53%	20%	0.9%
	0.4	5.2	45%	13%	0.2%
	0.5	4.6	40%	9%	0.07%
	0.1	28.5	88%	74%	43%
	0.2	16.0	81%	59%	21%
2.5	0.3	11.8	75%	48%	11%
	0.4	9.8	71%	40%	6%
	0.5	8.5	67%	34%	3%

$\mathbb{E}(x \mid G_0)$: the mean number of offspring generated by a primary case identified by backward tracing (G0 case). Note that this is larger than the mean number of offspring of a random case.

Probability ($x \geq n \mid G_0$): the probability that the number of offspring generated by a G0 case is n or greater.

focuses on forward tracing[3,4], there has been increasing interest in a possible combination of forward and backward tracing to control COVID-19[14,20]. Our results provide further evidence for this approach by quantifying the possible benefit of backward tracing, especially when the offspring distribution is highly variable, as is the case with SARS-CoV-2.

There are a number of operational challenges to implementing such contact tracing approaches. Since the number of contacts that lead to transmission is likely to be only a fraction of total contacts experienced by detected cases, expanding the coverage of contact tracing may involve a substantial logistical burden[21,22]. Engagement with contact tracing systems and adherence to quarantine may not necessarily reach sufficient levels[23–25]. With a longer timeline of contact history to be interviewed, recall bias may affect the success rate of backward tracing. In practice, interviewed cases might be asked not only for specific individuals they know to have contacted but also for a history of locations or events visited, as happens during outbreak investigations so that those who were present can be notified and/or tested. Backward tracing can in effect be viewed as an outbreak investigation process in which new cases and their contacts can be routinely linked via their shared exposure events, supported by

cross-referencing over epidemiological, diagnostic and quarantine datasets, with additionally identified infections triggering further tracing. Due to the difficulty in determining the direction of transmission, backward tracing may find a cluster of cases rather than a single primary case. However, our results still apply as long as subsequent forward tracing is conducted for all of the identified cases.

Our model makes some simplifying assumptions. Delays in confirmation and tracing were such that only generation-2 (G2) cases were assumed to be traced and quarantined before becoming infectious. In reality, cases are identified at different points in time and the reduction in infectiousness may be partial if cases are quarantined after becoming infectious (which can be a concern for backward tracing with an additional generation to trace). To allow intuitive comparison, the effectiveness of tracing was measured by the proportion of G3 cases averted given an index case detected by surveillance, and long-term dynamics were not considered. We believe our focus on assessing the effectiveness of a single practice of contact tracing triggered by a detected case is more relevant to operational-level decision making given finite resources. We also did not consider in our model that independently detected multiple index cases may have the same
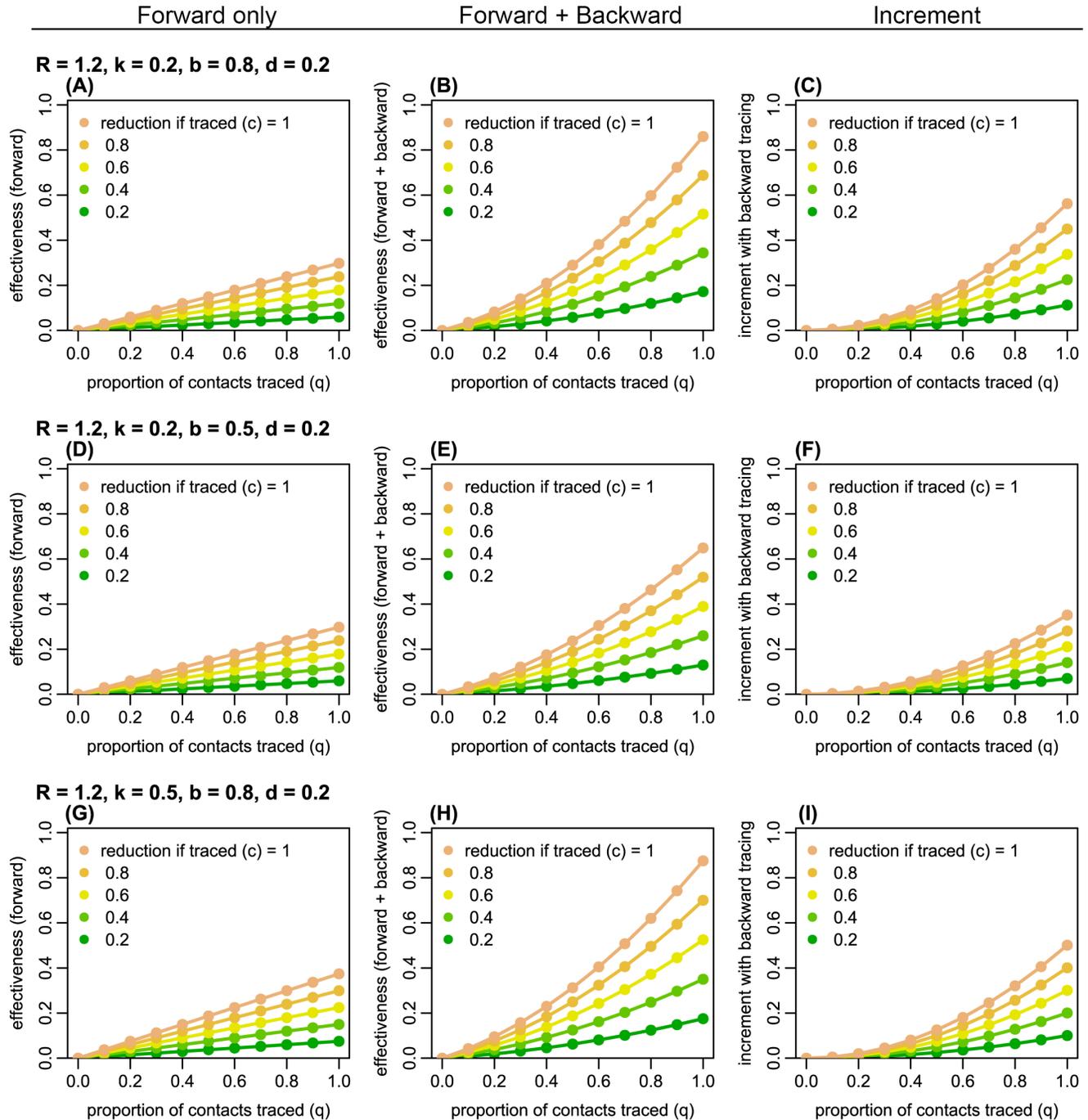
Figure 2. The estimated proportion of generation-3 (G3) cases averted by forward and backward contact tracing for different parameter values. Left panels (**A**, **D**, **G**): the effectiveness (the proportion of G3 cases averted) of forward tracing alone; middle panels (**B**, **E**, **H**): the effectiveness of a combination of forward and backward tracing; right panels (**C**, **F**, **I**): incremental effectiveness by combining backward tracing with forward tracing. Colours represent the relative reduction in transmission from G2 cases if traced and held in quarantine (*c*). *R*: the reproduction number; *k*: overdispersion parameter; *q*: proportion of secondary infections caused by a detected case successfully traced; *b*: probability of successful identification of the primary case; *d*: probability of detection of generation-1 (G1) cases independent of contact tracing.
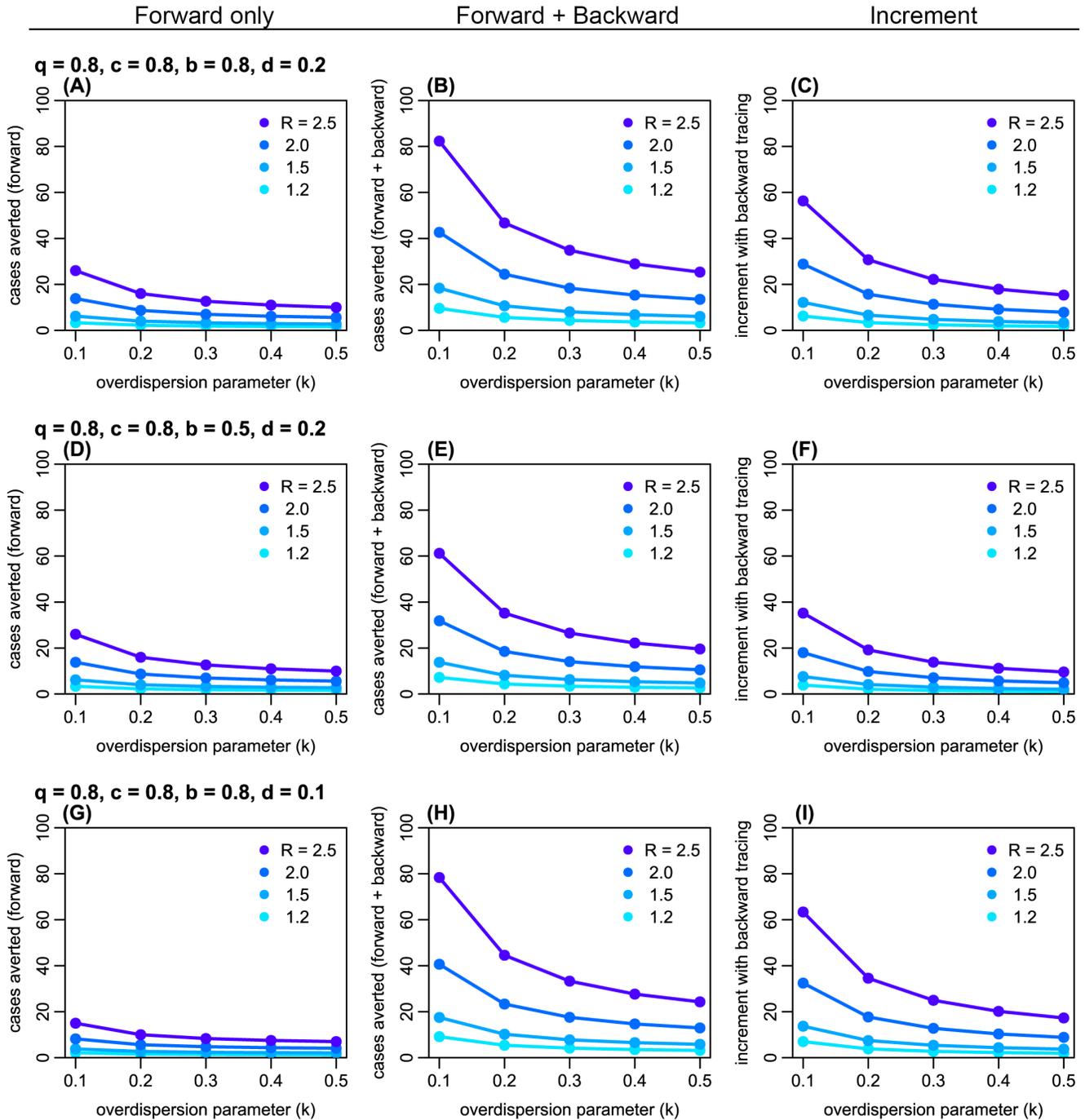
Figure 3. The estimated absolute number of generation-3 (G3) cases averted by forward and backward contact tracing. Left panels (**A**, **D**, **G**): the number of cases averted by forward tracing alone; middle panels (**B**, **E**, **H**): the number of cases averted by a combination of forward and backward tracing; right panels (**C**, **F**, **I**): additional cases averted by combining backward tracing with forward tracing. Colours represent the assumed reproduction number R. k: overdispersion parameter; q: proportion of secondary infections caused by a detected case successful traced; c: relative reduction in transmission from quarantined cases; b: probability of successful identification of the primary case; d: probability of detection of generation-1 (G1) cases independent of contact tracing.

primary case, which can cause duplicated effort of backward tracing. However, such duplication may be minimised if information of each index case is shared among health officials; moreover, overlapping backward tracing still has a benefit because it increases confidence in the identification of primary cases or infection settings.

With these limitations, our results suggest a significant potential benefit to backward tracing, which should be balanced against finite resources. Because backward tracing is operationally a set of forward tracing measures targeting multiple G1 cases in parallel, additional effectiveness requires a proportional amount of effort, in addition to the 'overhead' investigation effort to identify other G1 cases. Cost-effectiveness analysis combined with finer-scale dynamic modelling would help further identify the conditions under which backward tracing is most efficient and feasible.

Data availability
Underlying data
All data underlying the results are available as part of the article and no additional source data are required.

Extended data
Zenodo: akira-endo/COVID19_backwardtracing: Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks. https://doi.org/10.5281/zenodo.4062208[19].

Supplementary text 'COVID19_backwardtracing.html' contains the following supplementary figures (files stored in subfolder 'figs'), which are also available on GitHub:

- Figure S1: The estimated effectiveness with R = 2.5.

- Figure S2: The estimated effectiveness with R = 1.2 and d = 0.5.

- Figure S3: The number of generation 3 cases averted with 60% success rate of tracing and 60% relative reduction in transmission during quarantine.

Software availability
The reproducible code is available from: https://github.com/akira-endo/COVID19_backwardtracing.

Archived repository at time of publication: https://doi.org/10.5281/zenodo.4062208[19].

License: MIT

Members of the Centre for Mathematical Modelling of Infectious Diseases (CMMID) COVID-19 Working Group (random order)
Billy J Quilty, Matthew Quaife, Amy Gimma, Charlie Diamond, Rosalind M Eggo, Kiesha Prem, W John Edmunds, Fiona Yueqian Sun, Emily S Nightingale, James W Rudge, Simon R Procter, Rein M G J Houben, Sophie R Meakin, Christopher I Jarvis, James D Munday, Kevin van Zandvoort, Georgia R Gore-Langton, Stéphane Hué, Thibaut Jombart, Damien C Tully, Samuel Clifford, Nicholas G. Davies, Kathleen O'Reilly, Sam Abbott, C Julian Villabona-Arenas, Rachel Lowe, Megan Auzenbergs, David Simons, Nikos I Bosse, Jon C Emery, Yang Liu, Stefan Flasche, Mark Jit, Hamish P Gibbs, Joel Hellewell, Carl A B Pearson, Alicia Rosello, Timothy W Russell, Anna M Foss, Arminder K Deol, Oliver Brady, Petra Klepac

References

1. Ferretti L, Wymant C, Kendall M, *et al.*: **Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing.** *Science.* 2020; **368**(6491): eabb6936.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Kucharski AJ, Klepac P, Conlan AJK, *et al.*: **Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study.** *Lancet Infect Dis.* 2020; **20**(10): 1151–1160.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. World Health Organization: **Contact tracing in the context of COVID-19: Interim guidance.** 2020.
Reference Source

4. Centers for Disease Control and Prevention: **Health Departments: Interim Guidance on Developing a COVID-19 Case Investigation & Contact Tracing Plan.** 2020.
Reference Source

5. National Institute of Infectious Diseases: **[Implementation guideline for active epidemiological surveillance of novel coronavirus disease cases (tentative): addendum regarding implementation of rapid identification of cluster cases] (Japanese).** 2020.
Reference Source

6. Jindai K, Furuse Y, Oshitani H: **[Novel coronavirus disease cluster intervention] (Japanese).** *Infect Agents Surveill Rep.* 2020; **41**: 108–110.

7. Salathe M, Kazandjieva M, Lee JW, *et al.*: **A high-resolution human contact network for infectious disease transmission.** *Proc Natl Acad Sci.* 2010; **107**(51): 22020–22025.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Allard A, Moore C, Scarpino SV, *et al.*: **The role of directionality, heterogeneity and correlations in epidemic risk and spread.** 2020.
Reference Source

9. Kojaku S, Hébert-Dufresne L, Mones E, *et al.*: **The effectiveness of backward contact tracing in networks.** *arXiv.* 2020; 2005.02362.
Reference Source

10. Liu Y, Eggo RM, Kucharski AJ: **Secondary attack rate and superspreading events for SARS-CoV-2.** *Lancet.* 2020; **395**(10227): e47.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. Endo A, ; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Abbott S, *et al.*: **Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China [version 3; peer review: 2 approved].** *Wellcome Open Res.* 2020; **5**: 67.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Leclerc QJ, Fuller NM, Knight LE, *et al.*: **What settings have been linked to SARS-CoV-2 transmission clusters? [version 2; peer review: 2 approved].** *Wellcome Open Res.* 2020; **5**: 83.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. Klinkenberg D, Fraser C, Heesterbeek H: **The effectiveness of contact tracing in emerging epidemics.** Getz W, editor. *PLoS One.* 2006; **1**(1): e12.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

14. Bradshaw WJ, Alley EC, Huggins JH, *et al.*: **Bidirectional contact tracing dramatically improves COVID-19 control.** *medRxiv.* 2020.
Publisher Full Text

15. Lloyd-Smith JO, Schreiber SJ, Kopp PE, *et al.*: **Superspreading and the effect of individual variation on disease emergence.** *Nature.* 2005; **438**(7066): 355–359.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Lau MSY, Grenfell B, Nelson K, *et al.*: **Characterizing super-spreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA.** *medRxiv.* 2020.
Publisher Full Text

17. Adam D, Wu P, Wong J, *et al.*: **Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong.** Prepr (Version 1) available Res Sq. 2020.
Publisher Full Text

18. Oran DP, Topol EJ: **Prevalence of Asymptomatic SARS-CoV-2 Infection : A Narrative Review.** *Ann Intern Med.* 2020; **173**(5): 362–367.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Endo A: **akira-endo/COVID19_backwardtracing: Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks (Version 1.01).** *Zenodo.* 2020.
http://www.doi.org/10.5281/zenodo.4062208

20. Scientific Pandemic Influenza Group on Modelling Operational sub-group: **SPI-M-O: Consensus Statement on COVID-19, 3 June 2020.** 2020.
Reference Source

21. Keeling MJ, Hollingsworth TD, Read JM: **Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19).** *J Epidemiol Community Health.* 2020; **74**(10): 861–866.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Hellewell J, Abbott S, Gimma A, *et al.*: **Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts.** *Lancet Glob Heal.* 2020; **8**(4): e488–e496.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. Smith LE, Potts HWW, Amlot R, *et al.*: **Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK (the COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study).** *medRxiv.* 2020; 2020.09.15.20191957.
Publisher Full Text

24. Amann J, Sleigh J, Vayena E: **Digital contact-tracing during the Covid-19 pandemic: an analysis of newspaper coverage in Germany, Austria, and Switzerland.** *medRxiv.* 2020; 2020.10.22.20216788.
Publisher Full Text

25. Jonker M, de Bekker-Grob E, Veldwijk J, *et al.*: **COVID-19 Contact Tracing Apps: Predicted Uptake in the Netherlands Based on a Discrete Choice Experiment.** *JMIR Mhealth Uhealth.* 2020; **8**(10): e20741.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Supplementary materials: Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreak

Methods

We computed the effectiveness of contact tracing as the proportion of generation 3 cases averted. Assuming a negative-binomial branching process with a mean R and overdispersion paramter k, The mean total number of generation 3 cases given an index case found by surveillance is $C_3 = R^2(1 + R(1 + \frac{1}{k}))$.

Forward tracing

The expected number of generation 1 cases excluding the initially found index case is $R(1 + \frac{1}{k})$, of which proportion d is independently detected by symptom-based surveillance. Therefore, the total number of generation 1 cases targeted by forward tracing is $1 + Rd(1 + \frac{1}{k})$. Each of them would result in R^2 generation 3 cases on average if not traced, however, proportion qc are averted by contact tracing (success rate q) and quarantine (relative reduction c in infectiousness). The number of generation 3 cases averted is thus given as

$$\Delta_F = R^2 qc(1 + Rd(1 + \frac{1}{k})).$$

Backward tracing

We assume that backward contact tracing successfully identifies the primary case at probability b. Of the mean $R(1 + \frac{1}{k})$ generation 1 cases which is potentially under the scope of backward tracing, mean proportion $(1 - d)(1 - bq)$ will remain undetected either by backward tracing or independent detection. The total number of generation 1 cases detected is $1 + (1 - (1 - d)(1 - bq))R(1 + \frac{1}{k})$, which gives the number of generation 3 cases averted as

$$\Delta_{F+B} = R^2 qc(1 + (1 - (1 - d)(1 - bq))R(1 + \frac{1}{k})).$$

We computed the effectiveness of tracing as $\frac{\Delta_F}{C_3}$ and $\frac{\Delta_{F+B}}{C_3}$.

In [1]:

```
source("basic_funcs.R") # load functions
#options(jupyter.plot_mimetypes = "image/svg+xml")
```

105

The effectiveness of forward and backward contact tracing

Figures show the estimated effectiveness of forward and backward contact tracing. Left panels (A, D, G): the effectiveness (the proportion of generation 3 cases averted) of forward tracing alone; middle panels (B, E, H): the effectiveness of a combination of forward and backward tracing; right panels (C, F, I): additional effectiveness by combining backward tracing with forward tracing. Colours represent the relative reduction in infectiousness of generation 3 cases if traced and put in quarantine.

Figure 2 (main text). The estimated effectiveness with $R = 1.2$.

In [2]:

```
par(mfrow=c(3,3),mar=c(4,3,2.5,1),mgp=c(2,0.6,0),las=0,yaxs="i")

tertiary<-plotbypc(R=1.2,k=0.2,b=0.8,d=0.2, panel=c("(A)","(B)","(C)"))
tertiary<-plotbypc(R=1.2,k=0.2,b=0.5,d=0.2, panel=c("(D)","(E)","(F)"))
tertiary<-plotbypc(R=1.2,k=0.5,b=0.8,d=0.2, panel=c("(G)","(H)","(I)"))
```


Figure S1. The estimated effectiveness with $R = 2.5$.

106
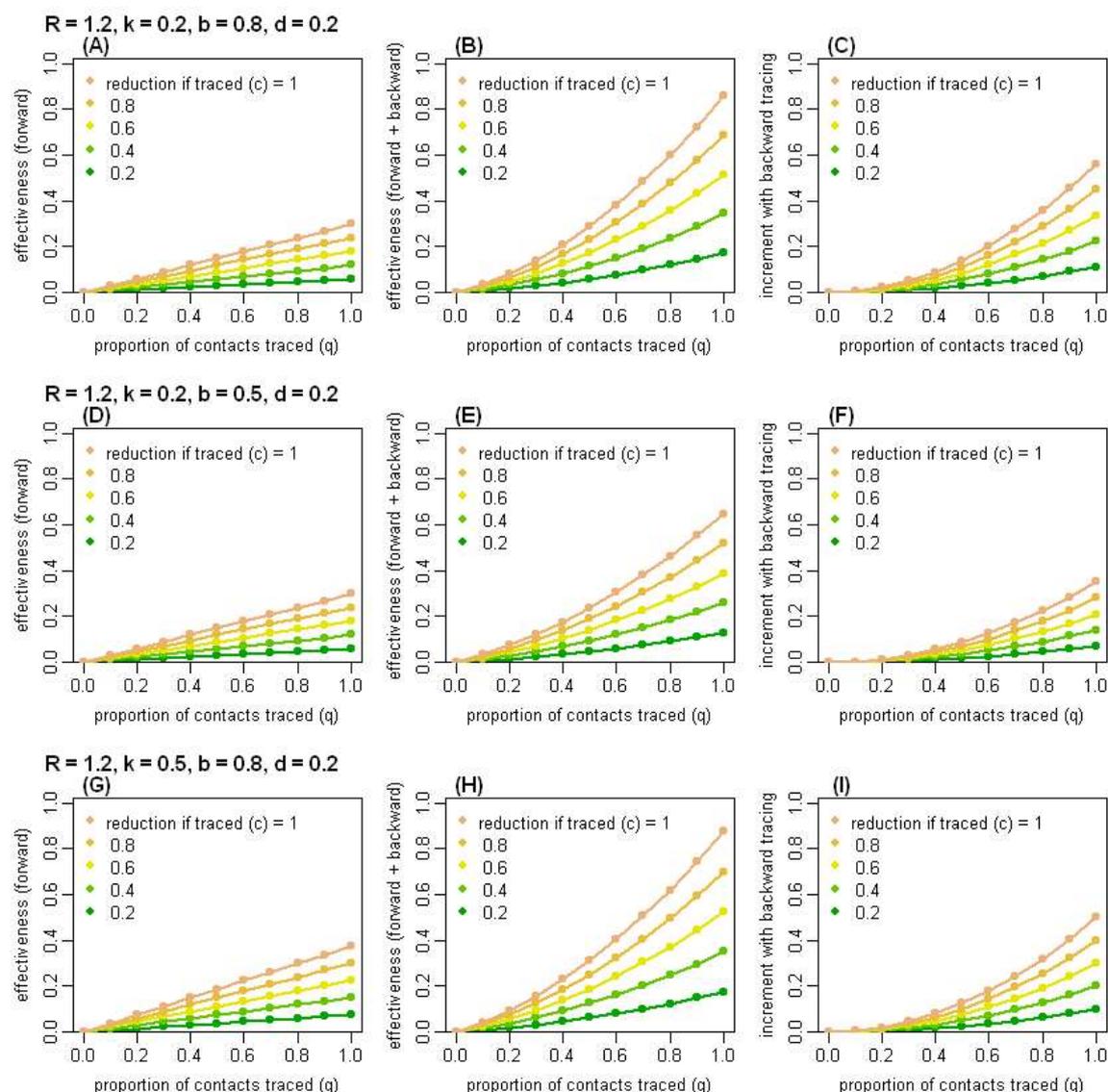
In [3]:

```
par(mfrow=c(3,3),mar=c(4,3,2.5,1),mgp=c(2,0.6,0),las=0,yaxs="i")

tertiary<-plotbypc(R=2.5,k=0.2,b=0.8,d=0.2, panel=c("(A)","(B)","(C)"))
tertiary<-plotbypc(R=2.5,k=0.2,b=0.5,d=0.2, panel=c("(D)","(E)","(F)"))
tertiary<-plotbypc(R=2.5,k=0.5,b=0.8,d=0.2, panel=c("(G)","(H)","(I)"))
```


Figure S2. The estimated effectiveness with $R = 1.2$ and $d = 0.5$.

107
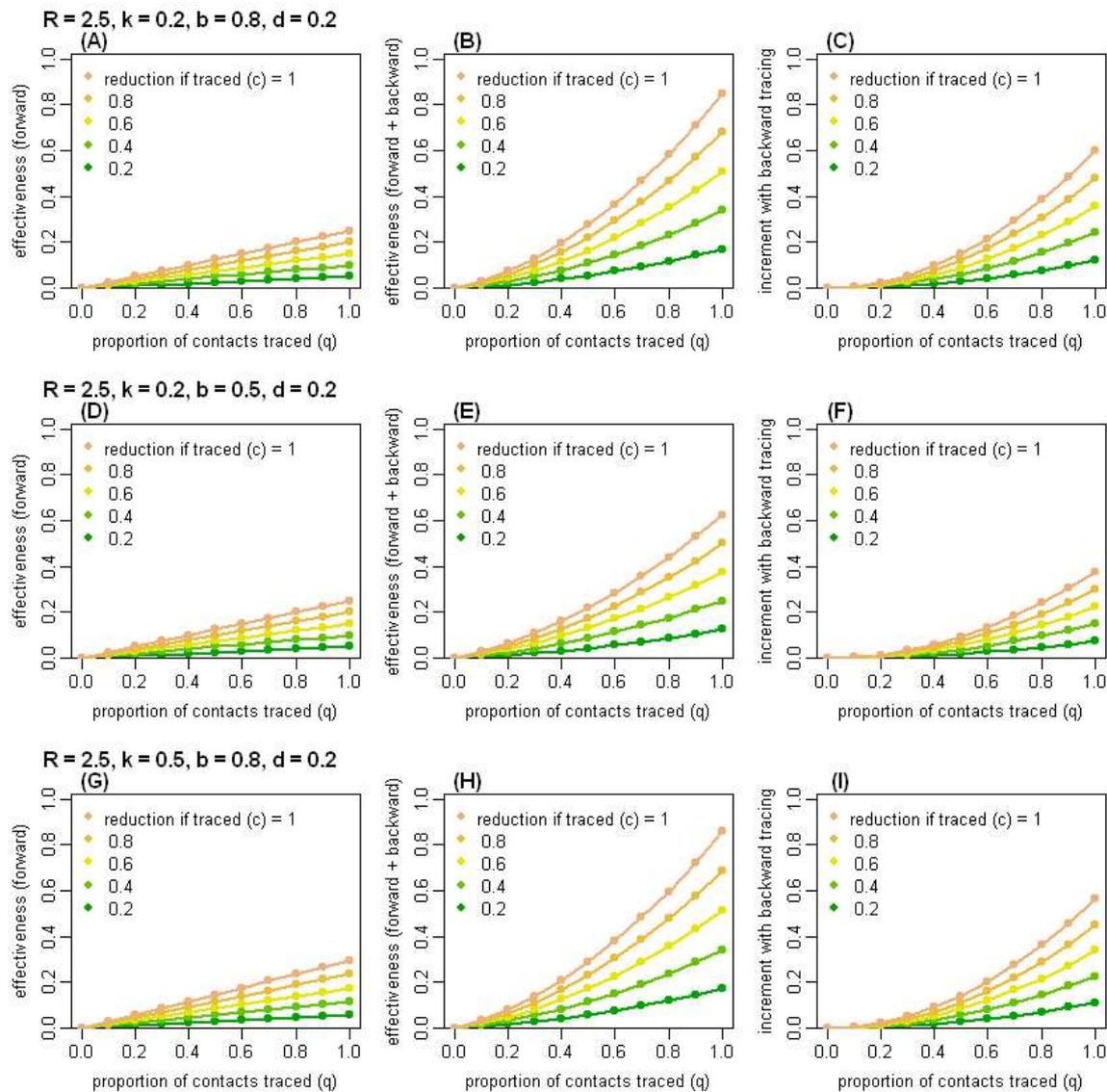
In [4]:

```
par(mfrow=c(3,3),mar=c(4,3,2.5,1),mgp=c(2,0.6,0),las=0,yaxs="i")

tertiary<-plotbypc(R=1.2,k=0.2,b=0.8,d=0.1, panel=c("(A)","(B)","(C)"))
tertiary<-plotbypc(R=1.2,k=0.2,b=0.5,d=0.1, panel=c("(D)","(E)","(F)"))
tertiary<-plotbypc(R=1.2,k=0.5,b=0.8,d=0.1, panel=c("(G)","(H)","(I)"))
```


108
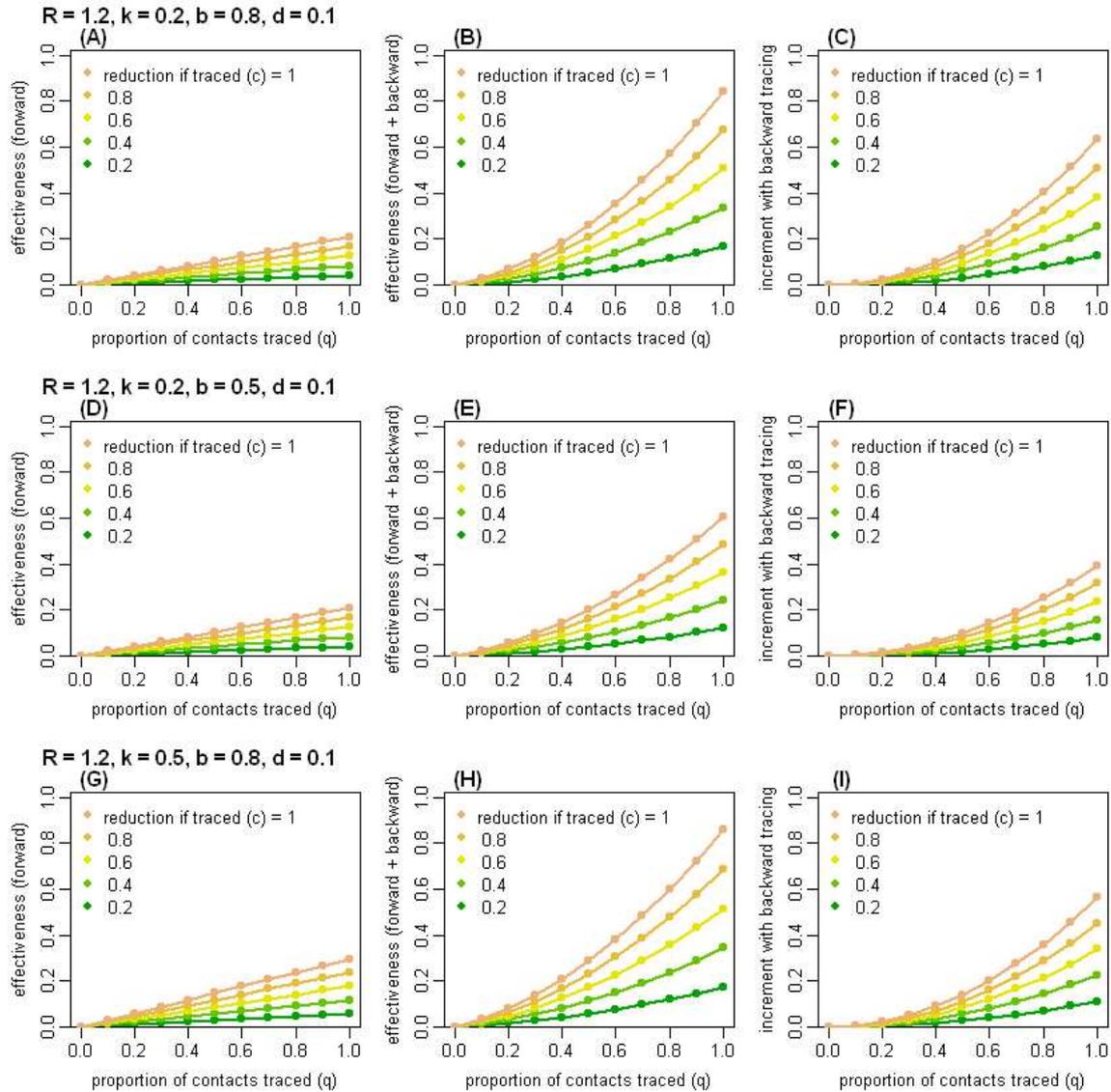
The absolute number of cases averted by backward tracing

Figures show the estimated number of generation-3 cases averted by forward and backward contact tracing. Left panels (A, D, G): the number of cases averted by forward tracing alone; middle panels (B, E, H): the number of cases averted by a combination of forward and backward tracing; right panels (C, F, I): additional cases averted by combining backward tracing with forward tracing. Colours represent the assumed reproduction number (R).

Figure 3 (main text). The number of generation 3 cases averted with 80% success rate of tracing and 80% relative reduction in transmission during quarantine.

In [5]:

```
par(mfrow=c(3,3),mar=c(4,3,2.5,1),mgp=c(2,0.6,0),las=0,yaxs="i")

tertiary<-plotbyRk(q=0.8,c=0.8,b=0.8,d=0.2, panel=c("(A)","(B)","(C)"), ylim=c(0,100))
tertiary<-plotbyRk(q=0.8,c=0.8,b=0.5,d=0.2, panel=c("(D)","(E)","(F)"), ylim=c(0,100))
tertiary<-plotbyRk(q=0.8,c=0.8,b=0.8,d=0.1, panel=c("(G)","(H)","(I)"), ylim=c(0,100))
```
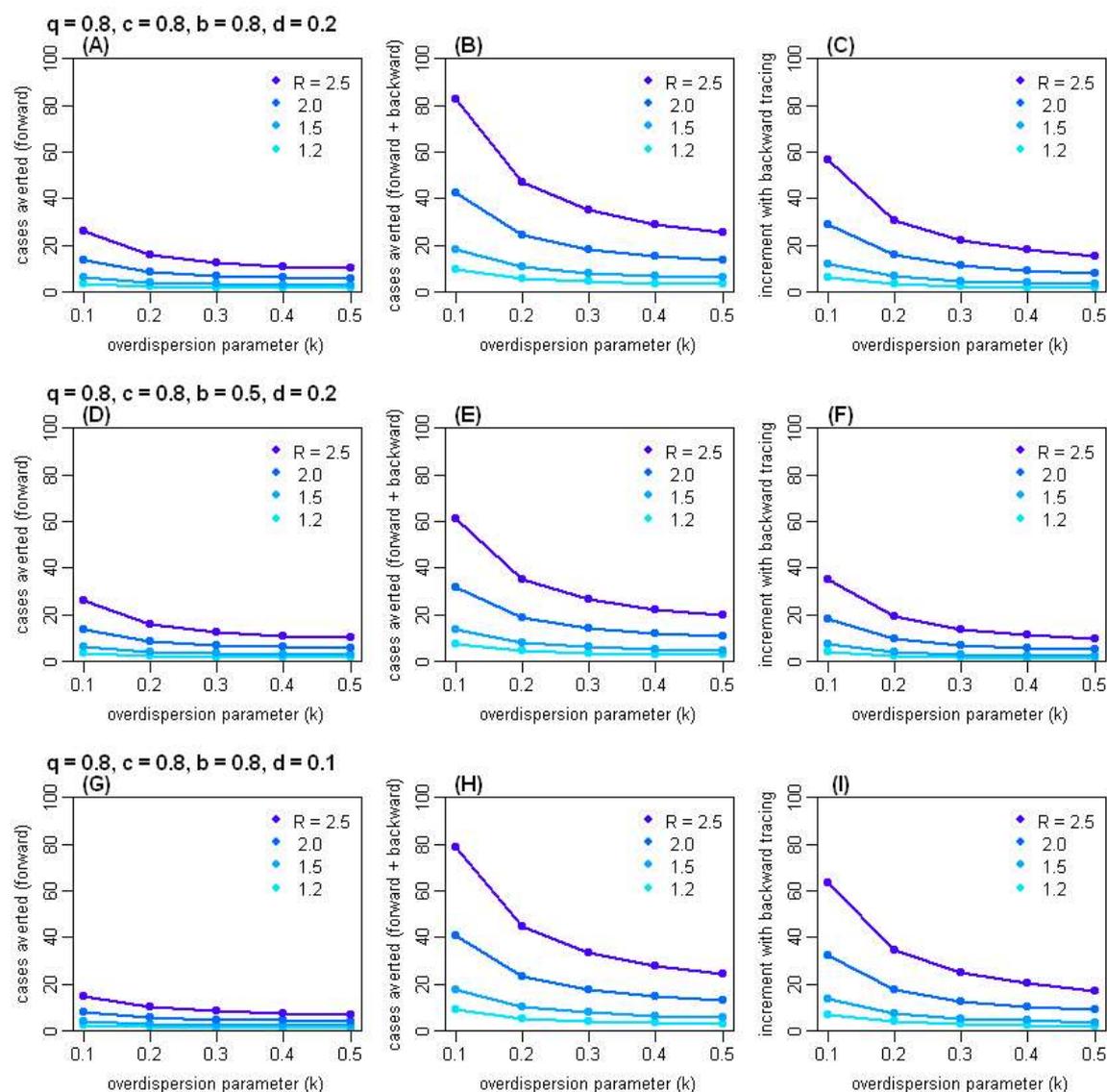
Figure S3. The number of generation 3 cases averted with 60% success rate of tracing and 60% relative reduction in transmission during quarantine.

In [6]:

```
par(mfrow=c(3,3),mar=c(4,3,2.5,1),mgp=c(2,0.6,0),las=0,yaxs="i")

tertiary<-plotbyRk(q=0.6,c=0.6,b=0.8,d=0.2, panel=c("(A)","(B)","(C)"), ylim=c(0,50))
tertiary<-plotbyRk(q=0.6,c=0.6,b=0.5,d=0.2, panel=c("(D)","(E)","(F)"), ylim=c(0,50))
tertiary<-plotbyRk(q=0.6,c=0.6,b=0.8,d=0.1, panel=c("(G)","(H)","(I)"), ylim=c(0,50))
```


110
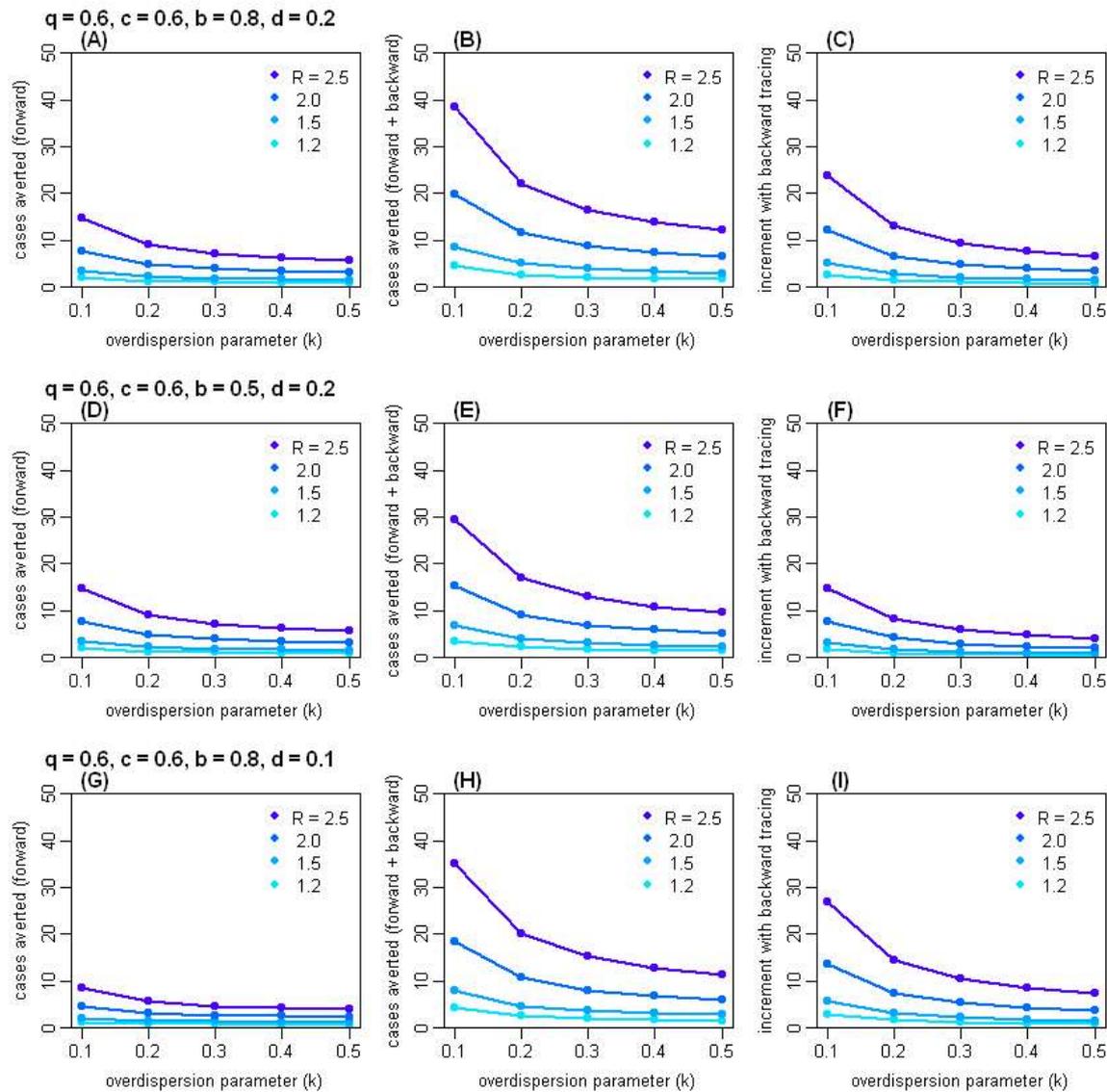
8 Paper 5: Within and between classroom transmission patterns of seasonal influenza and pandemic management strategies at schools

Within and between classroom transmission patterns of seasonal influenza and pandemic management strategies at schools

Akira Endo, CMMID COVID-19 Working Group, Mitsuo Uchida, Katherine E. Atkins, Adam J. Kucharski, Sebastian Funk

Abstract

Schools can play a central role in driving pandemics. While the contribution of schoolchildren to the overall transmission dynamics of COVID-19 remains unclear, the potential risks and interventions need to be assessed based on the best available data to inform management strategies in school settings. Heterogeneous social contact patterns associated with the social structures of schools (i.e. classes/grades) are likely to influence the within-school transmission dynamics; however, empirical evidence on the fine-scale transmission patterns between students has been limited. Using a mathematical model, we analysed a large-scale dataset of seasonal influenza outbreaks in Matsumoto city, Japan to estimate the transmission patterns as proxies for social interaction within and between classes/grades. We then applied these patterns to COVID-19 and pandemic influenza and simulated school outbreaks under multiple intervention scenarios. The overall within-school reproduction number, estimated to be around 0.8-0.9 for seasonal influenza, was minimally associated with class sizes and the number of classes per grade. Simulations suggested that with such transmission patterns, interventions changing class structures (e.g. reduced class sizes) may not be effective in preventing school outbreaks and that other precautionary measures (e.g. screening and isolation) need to be introduced. Class-level closures in response to detection of a case were suggested to be effective when regular screening tests for students are not available.

Background

With the emergence and rapid growth of the coronavirus disease 2019 (COVID-19) outbreak in early 2020, many countries decided to enforce school closures to prevent schools from becoming hotspots of transmission and thereby mitigate the further spread in the population [1,2]. Worldwide, there is much variability in how to balance these epidemic control measures and access to education. For example, many countries (including the UK and most European countries) had reopened schools by late 2020, employing a range of precautionary measures such as increased ventilation, enhanced hygiene, reduced class sizes and the introduction of 'social bubbles' (i.e. limiting contacts to small groups of students) [3]. Other countries took a different approach; as of 19 November 2020, in some countries (including the US and Canada) schools are only partially open, while in 23 countries schools remain closed [2]. Such diverse policies may in part reflect our still limited understanding of the potential role of schoolchildren in the transmission of severe acute respiratory syndrome virus 2 (SARS-CoV-2). While most infections are mild or asymptomatic [4–7], serology and outbreak investigations have shown that children can contribute to transmission [8–12]. Reports of COVID-19 outbreaks in school settings are relatively rare even after the full reopening of schools [13–15]; however, these data need to be interpreted with caution as multiple factors including asymptomatic infections, variability in transmission and enforcement of precaution measures could have been involved. There have been sporadic reports of large outbreaks associated with schools in various countries [16–18]. There is therefore uncertainty in the public health risk to students, teachers and the wider community when schools eventually reopen as normal.

A decision whether and how to open schools during a pandemic will need to weight the rights and welfare of children and their families against the public health implications. Although such policy should ideally be evidence-based, supporting data on outbreak risks and possible interventions in school settings are scarce. The few existing studies that assess the effect of school reopening plans on COVID-19 have three important limitations [19,20]. First, they assume that the contact rates of students are proportional to the number of students attending, which is not empirically validated. This assumption of density-dependent mixing necessarily entails that reducing student attendance (e.g. by introducing small class sizes or staggered attendance) would uniformly scale down the transmission

risk, which may overestimate the effect of interventions. Second, while school-based interventions aim to reduce social contact within and between classes and grades, previous studies have neglected these fine grain within-school contact patterns. Third, previous studies on within-school transmission dynamics were based on a limited number of schools and thus did not provide robust findings across schools or capture the full range of heterogeneity present [22,23].

To overcome these limitations, we first quantified social interaction within and between classes and grades by calibrating a model of seasonal influenza to an outbreak data from over 10,000 primary school students. This calibration allowed us to capture granular social contact patterns relevant to respiratory virus transmission in schools. We then embedded these estimates of the impact of class, grade and school sizes on social contact patterns into a dynamic model of SARS-CoV-2. Using this model, we assessed the risk and size of outbreaks under current COVID-19 interventions in use globally such as changes in class structure, screening and isolation, intermittent schooling and responsive class closures and we evaluated the optimal school-based pandemic management strategies. We also adapted the model to pandemic influenza to assess the robustness of these optimal strategies to a different pathogen.

Results

Transmission patterns of seasonal influenza in primary schools and estimated effects of school interventions on the reproduction number

We analysed the citywide survey data of 10,923 primary school students in Matsumoto city, Japan in 2014/15, which included 2,548 diagnosed influenza episodes of students (Figure 1A). The dataset was obtained from 29 schools with a range of class structures (sizes and the number of classes per grade), allowing for fine-scale analysis of within and between class transmission patterns (Figure 1B). Using a mathematical model that accounts for different levels of interaction within and between classrooms and grades, we estimated the within-school effective reproduction number R_S of seasonal

influenza in primary schools along with the breakdown of transmission risks associated with class/grade relationships (Figure 1C). The relationship between any pair of students in the same school was classified as either "classmates", "grademates" (in the same grade but not classmates) or "schoolmates" (not in the same grade). The estimated R_S was broken down as a sum of the contributions from these students, where the class size (n) and the number of classes per grade (m) was assumed to affect the risk of transmission. The reconstructed overall R_S in a 6-year primary school was estimated to be around 0.8-0.9 consistently and was minimally affected by n or m. Namely, an infected student was suggested to generate a similar number of secondary cases irrespective of the class structure. Transmission to classmates accounted for about two-thirds of R_S when each grade has only one class and was partially replaced by transmission to grademates as the number of classes per grade increases, while the sum of within-grade transmission (i.e. transmission to either classmates or grademates) remained stable.
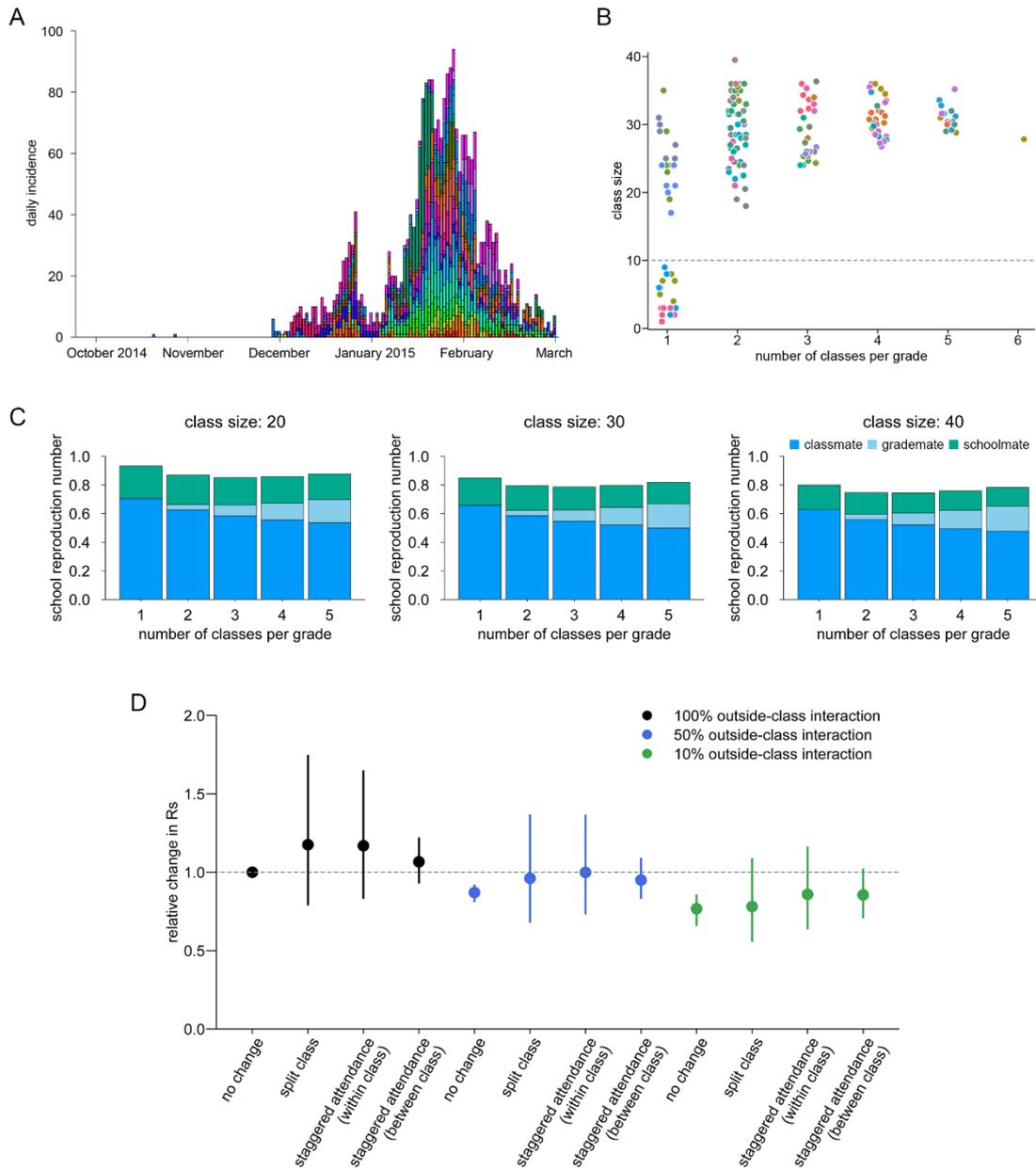
Figure 1. Transmission dynamics of seasonal influenza in primary schools in Matsumoto city and estimated effects of interventions for SARS-CoV-2. (A) Epidemic curve of seasonal influenza by illness onset in primary schools in Matsumoto city, 2014/15. Colours represent different schools. Month names denote the 1st day of the month. (B) Scatterplot of the class sizes and the number of classes per grade in the dataset. Each dot represents a class in the dataset and colours represent different schools. Dots are jittered along the x-axis. Classes of fewer than 10 students (denoted by dotted horizontal line) are excluded from the analysis. (C) School reproduction number (R_S) and its breakdown by the class/grade relationship (median estimates). (D) Relative change in the school reproduction number under school-based interventions. Dots represent medians and whiskers 95% credible intervals. Reduced outside-class transmissions (i.e. from grademates or schoolmates) were also considered (50% reduction: blue; 90% reduction: green).

To account for potential confounders that may affect the susceptibility or infectiousness of students, a log-linear regression was incorporated into the transmission model; the logarithm of the susceptibility and infectiousness of each student was given as a linear predictor of individual-level covariates (see Supplementary materials for detailed methods) and the effect of precautionary measures on susceptibility and infectiousness were simultaneously estimated with R_S. (Table 1). The results suggested that vaccines reduce susceptibility while masks reduce both susceptibility and infectiousness. Conversely, hand washing was suggested to increase susceptibility, in line with the earlier report of the original analysis [24] which attributed it to the congregation of students washing hands at school. Reduced chance of transmission during the winter break (27 December 2014–7 January 2015) was captured as a 76% decline in infectiousness of cases whose onset dates were during the break.

Table 1. Potential confounders and effects estimated in the log-linear regression

Individual-level covariate	Frequency in data	Relative susceptibility	Relative infectiousness
School grade (1 year increase)	—	1.04 (0.99-1.09)	0.93* (0.87-0.99)
Vaccine	47.7%	0.89* (0.81-0.98)	0.97 (0.82-1.17)
Mask	51.4%	0.76* (0.69-0.83)	0.66* (0.55-0.78)
Hand washing	80.1%	1.56* (1.34-1.78)	1.26 (0.94-1.78)
Onset in winter break	5.9% (of cases)	—	0.20* (0.11-0.33)

Values are median estimates and 95% credible intervals.

* 95% credible intervals not crossing 1.

Assuming that the relative contribution of class/grade relationship to the transmission risk is generally conserved in the dynamics of directly-transmitted diseases, we predicted the potential effects of interventions altering the school population structure (e.g. class sizes). The estimated relative effects of school-based interventions (summarised in Table 2) on R_S in a hypothetical setting of 6-year school with 2 classes per grade (40 students each) showed that splitting classes or staggered

attendance alone is unlikely to reduce R_S (Figure 1D), which is consistent with the aforementioned estimates of R_S minimally associated with class sizes and the number of classes. By reducing interactions between students from different classes by 90%, R_S could be reduced by up to around 20%. Combining split classes/staggered attendance and reduced interactions outside classes did not suggest incremental benefit.

Table 2. Summary of interventions that changes the size/number of classes

Interventions	Class size	The number of classes per grade
Baseline	40	2
Split class	20	4
Staggered attendance (within class)	20	2
Staggered attendance (between class)	40	1

Simulation of COVID-19 outbreaks in school

We reconstructed the time-dependent infection profile (i.e. the temporal distribution of secondary transmissions as a function of time after infection of the primary case) of SARS-CoV-2 from distributions reported in the literature [25,26] and assessed the possible reduction in transmission by screening either by symptoms or regular testing (Figure 2). If every student showing COVID-19-like symptoms is asked to isolate, post-symptomatic transmission within the school will be prevented. Post-symptomatic transmission is estimated to account for about half of the total secondary transmission of symptomatic individuals [25] and therefore expected to suppress the right tail of the infection profile. However, since symptom-based isolation will not apply to asymptomatic infections, the proportion of preventable transmission decreases with smaller assumed symptomatic proportions (Figure 2A). A recent study estimated 50% of seropositive children aged 2-15 years were symptomatic [27], and the performance of symptom screening could be even lower if some mild/atypical symptoms were missed in screening.

In addition to symptom screening, we also considered screening by regular testing. The daily rate of infectious students detected by a test (who will be asked to isolate from the next day) is given by the product of the frequency and sensitivity of the test ('effective testing rate'). Combined with symptom screening, regular testing could further reduce the risk of infection (Figure 2B). Of note, 10-

20% (roughly corresponding to performing an over 90% sensitive test once a week) effective testing rate is suggested to be sufficient to reduce the reproduction number by 40-70% and the effect saturates after the rate exceeds 30% (Figure 2C).
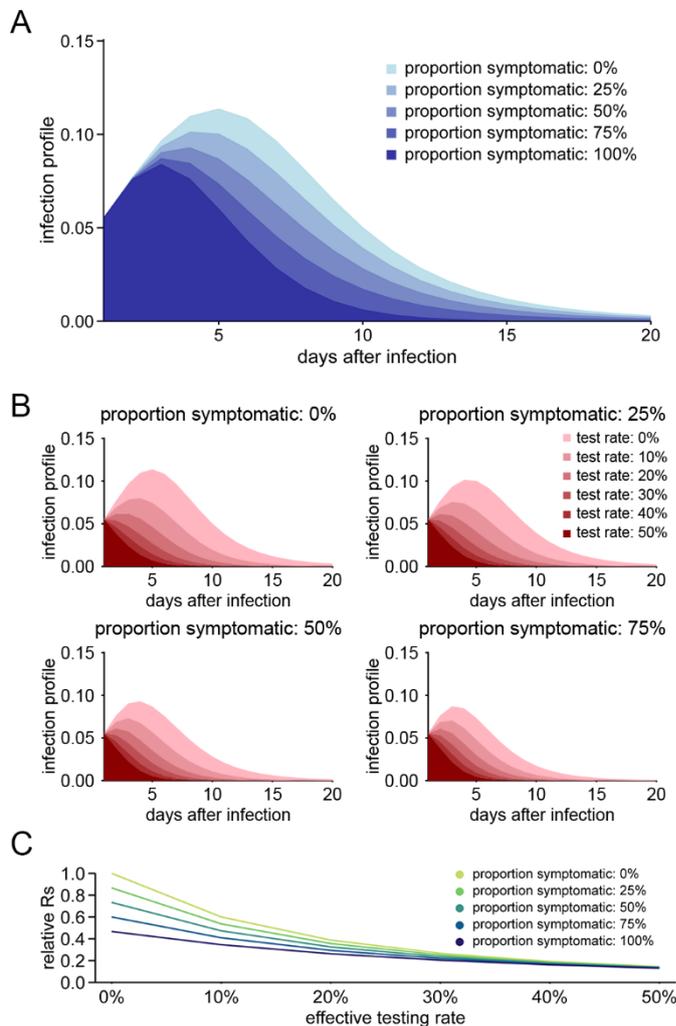
Figure 2. Time-dependent infection profile of SARS-CoV-2 and possible reduction by screening. (A) The effective infection profile for various symptomatic proportions where symptomatic students are isolated from the next day of symptom onset and do not contribute to further transmission (symptom screening). (B) The effective infection profile where students are screened by both symptoms and regular tests. Students are assumed to be isolated from the next day of presenting either symptoms or a positive test result. (C) The relative change in the reproduction number with combinations of symptom and regular test screening.

We used the estimated within-school transmission patterns of seasonal influenza as proxies for social interaction among students which can inform the dynamics of infectious diseases that share similar modes of transmission. By combining these patterns with the infection profile of SARS-CoV-

2, we simulated possible outbreaks of COVID-19 in a 6-year school triggered by a single case introduced from outside (Figure 3). For each of the three R_S values assumed (1.8, 1.2, 0.8), temporal patterns of disease spread across the school were simulated and compared between interventions. Without interventions, an outbreak can reach the whole school 50-60 days after the introduction at earliest in scenarios where R_S is above one. Given the estimated transmission patterns, the simulation also suggested that infection in a class can quickly spread to classes in different grades, indicating that an outbreak may be less likely to be contained within a grade. Interventions that change the size or the number of classes (split class and staggered attendance) were not predicted to contribute to the outbreak control (Figure 3A). Screening by symptoms and regular testing was suggested to be effective (Figure 3B). If 50% of infected students can be detected by symptoms at some point during their infectiousness period, symptom screening alone could render the scale of an outbreak with R_S = 1.2 comparable to one with R_S = 0.8. A combination of symptom screening and regular testing (effective test rate of 10%) could even bring an outbreak with R_S = 1.8 to a similar level. Intermittent schooling (setting regular "off" days on which students do not attend on-site classes) was also suggested to be effective. Alternating 'on' and 'off' every day was sufficient to suppress the outbreak. If combined with symptom screening, more days could be spent on-site (2 days 'on': 1 day 'off') while achieving the equivalent control of the outbreak.

To simulate outbreaks with 'class distancing', where between-class interactions are reduced (e.g. by enforcing 'social bubbles' within school), we employed two different assumptions on the change in within-class interactions. While within-class interaction may remain constant when interactions with students outside the class are restricted, it could also cause an increase in the within-class interaction to compensate for the reduction outside the class, just as was observed for transmission risks between grademates. Although the scale of the outbreaks became smaller than the baseline irrespective of the presence of compensation, the effect was smaller than other interventions (screening or intermittent schooling), especially in the presence of compensation effects (Figure 3C).

From the simulation results with a single initial case for each intervention scenario, we estimated the risk of outbreaks involving over 10 or 30 secondary transmissions given multiple

120

introductions of cases from outside the school (Figure 3D). With an increase in the number of

introductions, the risk of large outbreaks was suggested to rapidly increase. The results imply that

when multiple introductions are expected due to high levels of community transmission, it may be

safe to ensure that the school reproduction number R_S is around 0.5 or below; otherwise only up to 10

introductions would be sufficient to pose a non-negligible risk of a large school outbreak. The risk

decreases if we assume an excessive overdispersion as is observed with SARS-CoV-2 [28]; however,

the results exhibit qualitatively similar patterns and R_S of 0.5 should remain to be the primary target.
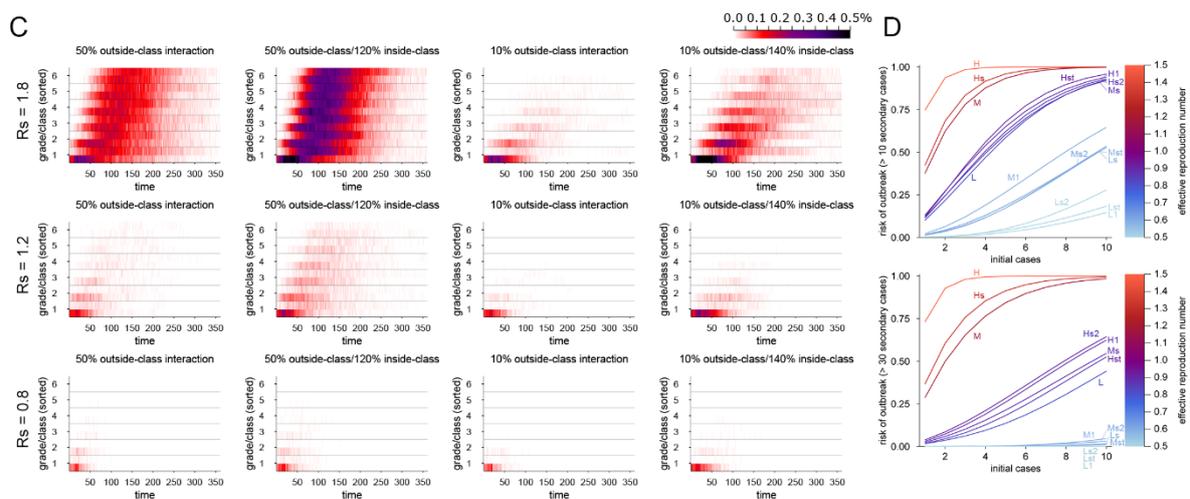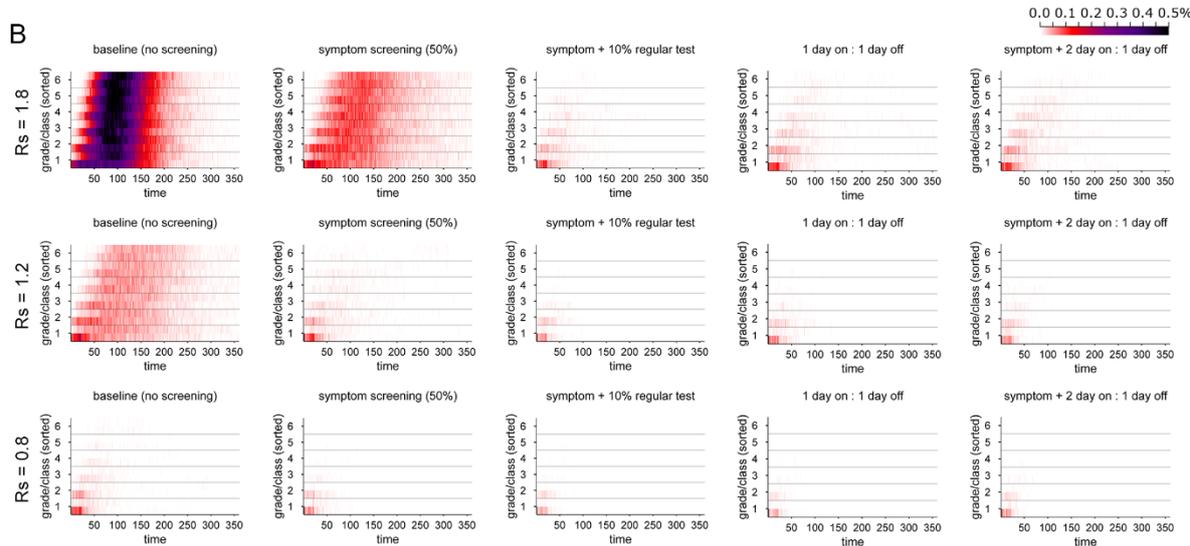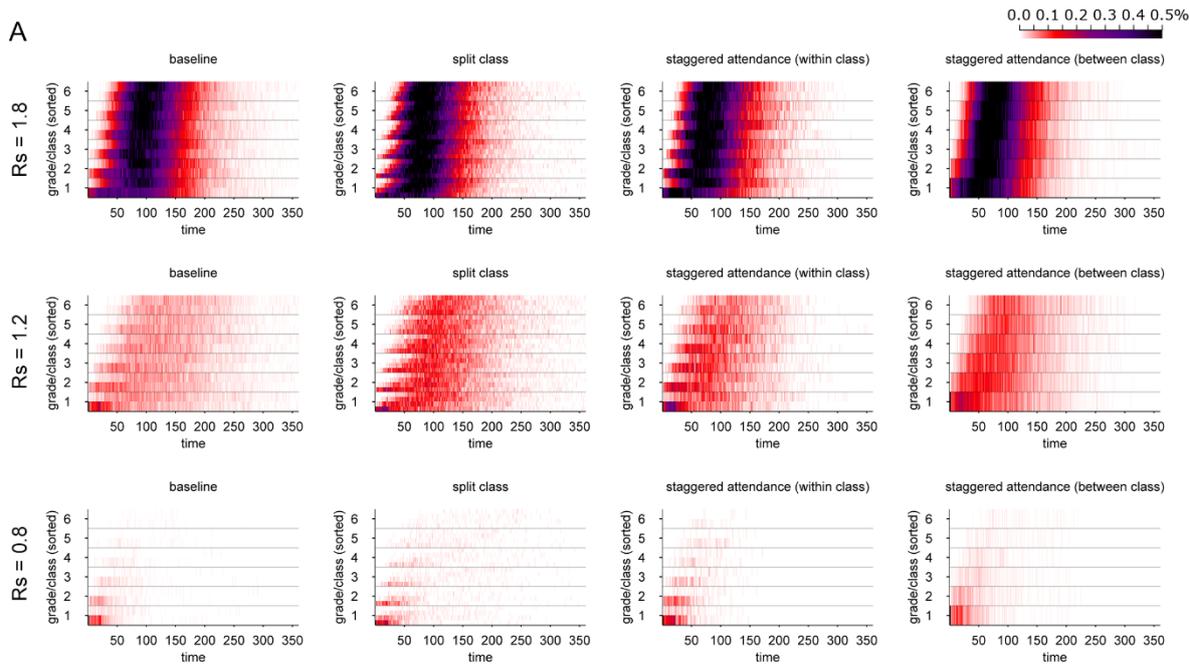
Figure 3. Outbreak simulations of SARS-CoV-2 in schools under interventions. (A) Simulated temporal patterns of outbreaks under interventions changing class structures. Colours represent the mean class incidence rate (the number of new infections on a single day in each class divided by the class size) over the 500 simulations. For each simulation, grades and classes are sorted by the date of the first case in the class so that the spread of infections in classes is time ordered from the bottom to the top. (B) Simulations with screening and intermittent schooling (C) Simulations with reduced outside-class interactions. Compensatory increases in the within-class interactions (20% and 40% increase in within-class interactions to compensate for 50% and 90% reductions in outside-class interactions, respectively) were also considered as part of the simulation. (D) The estimated risk of large outbreaks with multiple introductions. Curves show the probability that the eventual number of secondary transmissions within school exceeds 10 or 30 cases in the intervention scenarios, given multiple introductions of infected students from outside the school. Interventions are labelled by the following notations. H: the school reproduction number (R_S) = 1.8; M: R_S = 1.2; L: R_S = 0.8; s: screening by symptoms; t: screening by regular testing (effective rate 10%); 1: "1 day on: 1 day off" intermittent schooling; 2: "2 days on: 1 day off" intermittent schooling. Colours denote the effective reproduction number within the school for each intervention.

Managing school outbreak of COVID-19 by single-class closures

We explored the conditions that allow for effective control by class closures instead of a whole school closure by assessing the simulated spread of infections by the time the outbreak is first recognised either by symptoms or regular tests (Figures 4A-C). If the case finding depends only on symptoms, it is fairly likely (~50% or more) that more infections have spread unnoticed when the first case is recognised. Moreover, there is a chance of 25% or more that unnoticed infections also exist outside the class of the first detected case ('spillover'), which suggests that closure of that class alone may be insufficient for containment. If the proportion symptomatic is lower than 50% reported in [29], the outbreak could reach a substantial size (even over 10 or 20 infections) by the time the first case shows symptoms. Introducing regular testing, even at the effective rate of 10%, could markedly reduce the risk of undetected spread. The risk of outside-class spillover by the time of detection is limited to around 10%, which opens a possibility for control by closing only one class (or a few additional classes in the case of a rare event). If regular testing is not available and thus case finding needs to depend on the presence of symptoms, another possible option is to implement class distancing well before an outbreak is recognised to reduce the risk of spillover upon detection. When 50% of infections are symptomatic, reducing outside-class interaction by 50% is predicted to render

the spillover risk comparable to the 10% testing scenario. Similar results were obtained even in the

presence of a compensatory increase in the within-class interactions (Figure S6A).

Figure 4. Likely scales of COVID-19 outbreak at recognition and simulations of single-class closure strategies. (A) The

predicted distributions of the number of unnoticed infections by the time of the first identification of a case in school. (B)

The predicted distributions of unnoticed infections outside the class of the first identified case ('spillover'). (C) The

predicted distributions of unnoticed infections by the first identification of a case under the class distancing interventions

(blue: overall; red: spillover infections). (D) The final size of simulated outbreaks with and without single-class closure

strategies and the total days of class closures. Top panels: comparison of the cumulative number of infections with and

without class closures in each setting. Bottom panels: the distribution of the number of days of class closures aggregated across the school. Bars represent the upper 95% bound and middle lines show the mean over the simulations. Note that y-axes have different scales between panels. (E) Simulated temporal patterns of outbreaks and class closures with different closure strategies (symptomatic proportion: 50%). Colours represent the mean class incidence rate and the proportion of a class being closed over the 500 simulations.

We then simulated outbreaks of COVID-19 in schools where the single-class closure strategy is in operation, i.e. a class is closed for 14 days if any student in the class is found to be infected (either by showing symptoms or testing positive) while other classes with no detected infection keep operating (Figures 4D, 3E). Although the single-class closure strategies were suggested to be effective in outbreak containment across the settings considered, the 'naive' strategy with no regular testing or class distancing tended to result in a larger outbreak and more class closures, indicating the loss of education opportunities. This difference was particularly marked when the proportion of symptomatic infections is smaller (Figure S5). Incorporating regular testing or class distancing showed better performance both in terms of outbreak containment and education opportunities; regular testing resulted in smaller outbreak sizes while class distancing required less class closure, although the differences were minor. However, the results of regular testing combined with single-class closures warrant caution because the outcome was not substantially different from the isolation-only scenario with 10% regular tests; such marginal benefit may not be worth the loss of education opportunities. These results suggest that when regular tests are available, asking only test-positive students to isolate may be preferable to a class closure. Regular testing can identify infected students early in their infectiousness period; therefore, it becomes more likely that isolation alone is sufficient to prevent further transmissions.

Simulation of pandemic influenza outbreaks in schools

We applied our pandemic management approaches discussed as above in the context of COVID-19 to another potential threat—pandemic influenza. Compared with COVID-19, influenza tends to exhibit a shorter time course (i.e. shorter generation time and incubation period), which may affect the effectiveness of screening by symptoms/regular tests. Although empirical data is relatively

scarce on the symptomatic ratio of past pandemic influenza strains, that of seasonal A/H1N1 or A/H3N2 influenza strains in primary school-age children has been estimated to be around the range of 25-50% [30–32], in line with that of SARS-CoV-2 in children [29].

The infection profile constructed from the serial interval distribution used for the inference of the Matsumoto city data (mean: 2.2 days [33]) and the incubation period distribution of influenza A (median: 1.4 days [34]) reflected the possible scenario where screening by symptoms or regular tests may be less effective than SARS-CoV-2 because the majority of infections can occur before isolation due to shorter infection cycles (Figure 5B). In this setting, screening by symptoms and regular testing with 10-20% effective testing rates could reduce the reproduction number by only up to 30-40%: about half of what was estimated for SARS-CoV-2.

Outbreak simulations with various interventions overall showed similar patterns to COVID-19 except that screening by symptoms/regular tests was suggested to be less effective for pandemic influenza than for COVID-19 (Figure 5A). Notably, another difference was that combining "2 days on: 1 day off" intermittent schooling and symptom screening no longer had an equivalent effect to "1 day on: 1 day off" intermittent schooling. Single class closure strategies improved the outcome in most cases, although they resulted in larger outbreak sizes and more closures than in the COVID-19 simulation (Figures 5C, 4D). Combining class closures with regular testing, which was not suggested to be cost-effective for COVID-19, exhibited a plausible level of performance for pandemic influenza.

126
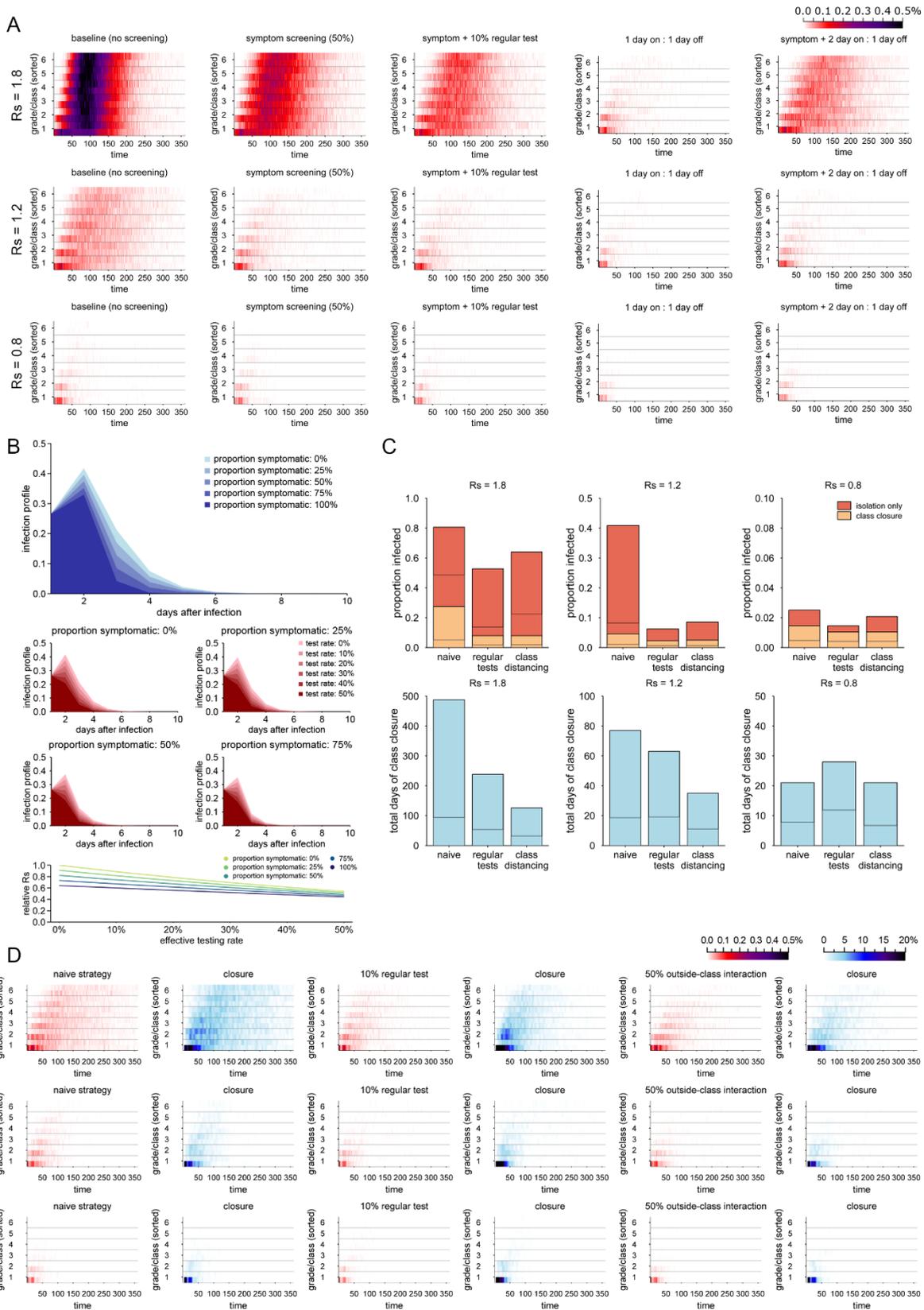
Figure 5. Simulated patterns of pandemic influenza outbreaks in schools. (A) Simulated temporal patterns of outbreaks with screening and intermittent schooling. Colours represent the mean class incidence rate (the number of new infections on a single day in each class divided by the class size) over the 500 simulations. (B) Assumed time-dependent infection profile of

pandemic influenza and possible reduction by screening. The effective infection profile is shown where infectious students identified either by symptoms or regular testing are isolated and thus do not contribute to the infection profile. (C) The final size of simulated outbreaks with and without single-class closure strategies and the total days of class closures. Top panels: comparison of the cumulative number of infections with and without class closures in each setting. Bottom panels: the distribution of the number of days of class closure aggregated across the school. Bars represent the upper 95% bound and middle lines show the mean over the simulations. (D) Simulated temporal patterns of outbreaks and class closures with different closure strategies (symptomatic proportion: 50%). Colours represent the mean class incidence rate and the proportion of a class being closed over the 500 simulations.

Discussion

We employed a mathematical model that stratifies transmission within and between classes/grades to understand and simulate the dynamics of directly-transmitted diseases at school. A citywide primary school seasonal influenza epidemic data was used to calibrate the model and a range of interventions were assessed in simulations of COVID-19 and pandemic influenza outbreaks. We used one of the largest datasets of school outbreaks with over 10,000 students and 2,500 cases, which we believe provides the best available evidence on transmission patterns within schools. Moreover, given that COVID-19 in schools has scarcely been found and documented, this dataset likely remains to be one of the most important resources to assess school transmission risks in the current COVID-19 pandemic as well as future pandemics.

The inferred transmission dynamics of seasonal influenza in Matsumoto city suggested that the within-school reproduction number R_S was, unlike as is often assumed, almost constant regardless of the size or the number of classes ('frequency-dependent mixing' [21]). The estimated R_S of 0.8-0.9, more than half of which was attributable to within-class transmissions, is consistent with a previous study in the United States [22] and also in line with the reported R_0 of 1.2-1.3 for seasonal influenza [35] provided that students in this dataset were previously estimated to have infected 0.3-0.4 household member on average [36]. The value of R_S below 1 suggests that the outbreak cannot sustain itself within school alone and that interactions through importing and exporting infections between

households and the general community is likely to play a crucial role in the overall transmission dynamics.

The estimated breakdown of R_S showed intriguing patterns. As the number of classes per grade increases, the contribution of within-class transmission risk declines and is replaced by within-grade transmission. Combined with the almost constant overall R_S, this might indicate that contact behaviour between students that contributes to transmission is rather inherent and minimally affected by the student population density. Namely, students may have a certain number of 'close friends' with whom they have more intimate interactions that could facilitate transmission. In a school with more classes per grade, some of such friendship comes from grademates instead of classmates, but the total number of close friends remains similar. This interpretation is in line with our understanding of influenza spreading predominantly in close proximity [37] and can have a non-negligible impact on the expected effect of interventions on not only influenza but also SARS-CoV-2, which shares similar routes and range of transmission [38,39]. Our results suggested that interventions such as reducing class sizes or the number of students present (staggered attendance) may be less effective than what is expected under the density-dependent mixing assumption. If interventions altering class structures are not accompanied by additional precaution measures and students try to resume their 'natural' behaviours (i.e. the same contact patterns as those in school with the resulting class structures) through so-called social contact 'rewiring' [40], the effect of such interventions can diminish or even reverse. For example, if other classes are absent due to staggered attendance, students may increase their interactions with classmates instead of their previous close friends in other classes. We propose that reducing the class sizes or the number of attending students should be considered only if they enable effective implementation of precaution measures such as physical distancing, environmental cleaning or forming social bubbles (although we believe this is often the case of school outbreak management).

Simulated school outbreaks of COVID-19 and pandemic influenza suggested two possible directions of management strategies. One of them is the 'passive' approach, which tries to reduce R_S before the emergence of an outbreak by interventions. If R_S is kept sufficiently small during everyday

operation by incorporating various intervention methods (e.g. screening based on symptoms or regular tests, reducing outside-class interactions and intermittent schooling), a school would be resistant to sustained transmission. We believe that the passive approach should aim R_S of at most around 0.5 such that the risk of large outbreaks is kept at an acceptable level even with multiple introductions from outside school. This approach is likely to require combining multiple intervention methods if the baseline R_S is high; however, if successfully implemented, it may also ensure that schools can operate nearly as normal even amid ongoing community transmission. Alternatively, schools could also decide to operate with less stringent measures and take a 'responsive' approach, where only students in classes with at least one confirmed case will isolate (single-class closure). This strategy requires less intensive baseline measures and thus could be more efficient in low community transmission settings. Moreover, it allows ramping up control efforts according to the actual intensity of outbreaks (i.e. the scale of closure follows that of an outbreak). For the responsive approach to work, the outbreak needs to be recognised before it spreads outside the initially-affected class. Reduced outside-class interactions will assist this and are expected to reduce both the scale of outbreaks and class closures. While regular testing combined with the responsive approach could also bring a similar effect, it did not prove to be cost-effective for COVID-19 because isolating only test-positive students (without class closure) was predicted to achieve similar outcomes with minimal loss in education opportunity (except for some of the pandemic influenza scenarios assuming shorter infection cycles). That is, if a school can afford regular testing of students, the passive approach will allow sufficient control and class closures may not be necessary. Alternatively, in such resource-rich settings, intensive testing of a whole class where a positive case is found may achieve the same effect as a class closure in the responsive approach, which may be preferable in some settings as it allows uninfected students to remain at school.

When designing an overall management plan, the strengths and weaknesses of intervention measures should be recognised and compared against each other. Regular testing is a powerful intervention that enables prompt detection and isolation of cases, which leaves responsive class closures almost unnecessary. In our simulations, the effective daily testing rate of 10% exhibited

sufficient performance in most cases. Using tests with a reasonable sensitivity, this means the frequency of tests can be no more than once a week. Although this would ease the required logistical burden, the option may not always be available to every school. The issue of false-positive also needs to be noted. Effective testing rate of 10% means that 99.9% specific tests may produce a false positive more than once a month on average, although this may still be worth the benefit if only positive students are isolated. Testing kits have different levels of specificity and should be selected considering the overall benefit given the risk of false positives. Meanwhile, screening by symptoms is unlikely to suffer from this issue since it will be reasonable to let symptomatic students stay home regardless of the actual cause, which renders this option probably the easiest to implement. However, symptom lists should be properly outlined such that mild symptoms (including those not typically considered as illness, e.g. loss of smell/taste for SARS-CoV infection [27]) will not be missed. Intermittent schooling is another powerful intervention better coupled with the passive approach, which reduces the number of days infectious students spend at school, irrespective of whether they are identified as cases. In our simulations, students were assumed to stay home once in two or three days. While this can result in missed opportunities for education, if considered in combination with weekends, it may be achievable only at the cost of one or two days of in-class teaching. Missed opportunities could even be mitigated by introducing online teaching on these "off" days. Compared with staggered attendance, intermittent schooling may be logistically more feasible as it does not require separate courses for split groups of students. Choice and combination of intervention measures should consider the risk assessment, current situation of community transmission and practical constraints.

Several limitations of this study should be noted. First, the transmission patterns within schools were estimated from a single dataset of seasonal influenza in primary schools (aged 5-12 years), and it is unclear to what extent the results can be extrapolated to other settings. Qualitative patterns may still be informative to predict possible transmission dynamics in different types of schools if they reflect social contact behaviours of schoolchildren, e.g. small effects of class sizes on R_S; however, the relative contribution of within-class/within-grade interactions may become smaller

for older students [41]. The data points used in the inference mostly consisted of classes of size 20-40 (those with a size smaller than 10 were excluded as they might be operated differently) and most schools had no more than 5 classes per grade. Our simulation was also limited to within this range for internal consistency and thus is not necessarily applicable to class structures outside this range (e.g. splitting a class of 20 students into two). Since the illness data of teachers were not available, they were not considered throughout the analysis, although their role in seasonal influenza transmission may have been minor. Second, there were potential sources of bias inherent to the nature of the dataset which was not necessarily fully addressed. Most importantly, the original study was an observational study and thus the differences in the transmission patterns between schools of different class structures might not be causal. We assumed that if class structures were altered by interventions, students would rewire their contacts according to the new class structure. However, students may respond differently in interventional settings, which could not be validated in observational studies. The effects of potential confounders were estimated primarily to minimise bias in the transmission patterns associated with the class structures, and the regression coefficients themselves (although overall in line with our understanding) should be interpreted with caution due to limited adjustment for biases. Although we believe our student incidence data had a better quality than most existing studies given encouraged medical attendance and confirmation by rapid diagnostic kits [36], a certain proportion of infections (e.g. asymptomatic or very mild) may have been missing from data. We believe that students feeling unwell due to influenza mostly attended medical institutions and received a test as it was encouraged by schools. Nonetheless, it should be noted that this could have been a source of bias in the estimated transmission patterns. Students with very mild symptoms (e.g. only slight sore throat) may visit a medical institution only if they know of other classmates also diagnosed with influenza. If such cases were common, the contribution of within-class transmissions in our results might have been an overestimate. Third, the epidemiological properties used in our simulations were subject to a number of assumptions. Within and between class/grade transmission patterns of COVID-19 and pandemic influenza were assumed to be proportional to those of seasonal influenza and scaled by the chosen R_S in the simulation. However, modelling studies often use similar assumptions of proportionality between transmission and social contacts [19,42] and we believe our

approach has strength over such studies as it could indirectly measure social contacts in the context of transmission. Infections acquired from household and general community were not explicitly modelled and simply treated as external introductions. Temporal profiles of infectiousness were based on limited data and also neglected individual-level variation. These may need to be updated in the future to reflect newer data; currently, the simulation results should be interpreted as a scenario analysis rather than conclusive predictions.

The present study offers novel insights into the transmission patterns in school settings reflecting class/grade structures. We believe these results would not only inform modelling studies that incorporate transmission dynamics in schools but also aid planning and assessment of outbreak management strategies at school for the current and future pandemics.

Materials and methods

Data

We analysed a citywide school-based influenza survey data from 2014/15 season. The survey was conducted in Matsumoto city, Japan, enrolling 13,217 students from all 29 public primary schools in the city. During the survey period (from October 2014 to February 2015), the participants were asked to fill out a questionnaire when they are back from the suspension of attendance due to diagnosed influenza (prospective survey). In March, the participants were asked to respond to another survey on their experience during the study period, regardless of whether they had contracted influenza (retrospective survey). A total of 2,548 diagnosed influenza episodes were reported in the prospective survey, which accounted for 96% of the cases officially recognised by the schools during the study period. Primary schools in Japan often requested students suspected of influenza to seek for diagnosis at a medical institution. All students reporting an influenza episode in the prospective survey answered that they had received a diagnosis and at least 95% of them were noticed of type A influenza (indicating that they were lab-confirmed). In the retrospective survey, 11,390 (86%) participants responded, among which 8,375 reported that they did not have influenza during the study period.

We combined those who responded to the prospective survey ("case group") and those who reported no influenza experience in the retrospective survey ("control group") and obtained a dataset of 10,923 students. Of those, 71 students from 3 schools with less than 10 students per grade were excluded because they may have different schooling patterns from other schools (e.g. some students in different grades shared classrooms). We used individual profiles (sex, school, grade, class, household composition), onset dates, influenza episodes of household members and precaution measures students engaged in (vaccine, mask, hand washing) in the subsequent analysis. Further details of the dataset can be found in the original studies [24,43].

The secondary data analysis conducted in the present study was approved by the ethics committee at the London School of Hygiene & Tropical Medicine (reference number: 14599).

Inference model

We modelled within-school transmission considering class structures as follows. We defined the "school proximity" d between a pair of students i and j attending the same school as

$$d_{ij} = \begin{cases} 1 & \text{(different grades, same school)} \\ 2 & \text{(different classes, same grade)} \\ 3 & \text{(different sex, same class)} \\ 4 & \text{(same sex, same class)} \end{cases} \tag{1}$$

To investigate the potential effect of reduced class sizes and the number of attending students, we modelled the transmission between students as a function of two variables: the class size n and the number of classes per grade m (i.e. the number of students per grade is nm). Namely, we assumed that in the absence of any individual covariate effects, the transmissibility between student i and j in proximity d is represented as

$$\beta_{ij} = \beta_d \big(n_{i,d} \big)^{\gamma_d} \big(m_{i,d} \big)^{\delta_d}, \tag{2}$$

where β_d, γ_d, δ_d are parameters to be estimated. When i and j are in the same grade (i.e. $d = 2, 3, 4$), the average class size and the number of classes in that grade were used as $n_{i,d}$ and $m_{i,d}$. When $d = 1$, the school average was used as $n_{i,d}$ and $m_{i,d}$. The exponent parameters within the same class were assumed to be equal: $\gamma_3 = \gamma_4$ and $\delta_3 = \delta_4$.

We modelled the daily hazard of incidence for student i as a renewal process. Let h_τ be the onset-based transmission hazard as a function of serial interval s (normalised such that $\sum_{s=1}^{\infty} h_s = 1$; $h_s = 0$ for $s \leq 0$). We used a gamma distribution of a mean of 1.7 and a standard deviation of 1.0 for influenza, which resulted in the mean serial interval of 2.2 days [33]. The daily hazard of disease onset attributed to school transmission is given as

$$\lambda_i^S(T) = v_i \sum_j w_j \beta_{ij} h_{T-T_j},$$

(3)

where v_i and w_i represent the relative susceptibility and infectiousness, respectively, which are specified for each individual by a log-linear model (see Table 1 for a list of covariates included).

In addition to the above within-school transmission, we also considered within-household transmission and general community transmission. The within-household transmission was incorporated as the Longini-Koopman model [44] with parameters from a previous study on the same cohort of students [36]. General community transmission was modelled as a logistic curve fitted to the total incidence in the dataset to reflect the overall trend of the epidemic. See Supplementary materials for further details of the model

We constructed the likelihood function and estimated the parameters by the Markov-chain Monte Carlo (adaptive mixture Metropolis) method. We obtained 1,000 thinned samples from 100,000 iterations after 100,000 iterations of burn-in, which yielded the effective sample size of at least 300 for each parameter. Using the posterior samples, we computed the proximity-specific reproduction number R_d in a hypothetical 6-year school with given n and m (assumed to be constant schoolwide) as

$$R_d = \begin{cases} 5nm \cdot \beta_1 n^{\gamma_1} m^{\delta_1} & (d = 1) \\ n(m-1) \cdot \beta_2 n^{\gamma_2} m^{\delta_2} & (d = 2) \\ n \cdot \dfrac{\beta_3 + \beta_4}{2} n^{\gamma_3} m^{\delta_3} & (d = 3, 4) \end{cases}$$

(4)

and defined the within-school reproduction number R_S as a sum of them.

We predicted the relative reduction in R_S under intervention measures changing the number of attending students and class structures by using posterior samples. Interventions were assumed to change n and m as shown in Table 2, and the predictive distribution of the relative change in R_S was computed for each intervention. The estimated R_S represents the value in a hypothetical condition where an infectious student spends the whole infectious period at school; the effect of absence due to symptoms or the staggered attendance was not included in this reduction.

Temporal infection profile of SARS-CoV-2 and influenza

We reconstructed the temporal infection profile of SARS-CoV-2 using distributions estimated originally in He et al. [25] and recalculated in Ashcroft et al. [26]. Since the estimated infection profile used the date of symptom onset as a reference point, we convolved the distributions of infection profile incubation period to reconstruct the infection profile as a function of time from infection. We approximated the infection profile reported in [26] as a normal distribution N($\mu = 0.53$, $\sigma = 2.65$) and convolved it with the lognormal incubation period distribution LN($\mu = 1.43$, $\sigma = 0.66$) [25] to obtain the infection profile. This resulted in a 6.8% chance of generation time being 0 day or less due to convolution; we thus used a truncated distribution at $t = 0$.

The modification of infection profile h_τ by screening was modelled as follows. Let U_τ represent the survival function against screening, i.e. the probability that an infected individual remains undetected by day τ post-infection. The infection profile under symptom screening is represented as

$$h'_\tau = h_\tau U_\tau = h_\tau (1 - \sigma)^\tau (1 - F_\tau), \tag{5}$$

where σ is the effective testing rate and F_τ is the cumulative distribution of the incubation period.

Similarly, we obtained the infection profile for influenza by using the gamma distribution described earlier (from [33]) and the estimated incubation period distribution of H1N1 pandemic influenza [34].

Simulation of COVID-19 and pandemic influenza outbreaks in school

We simulated school outbreaks using the estimated transmission patterns within and between classes/grades and infection profiles of SARS-CoV-2 and H1N1 pandemic influenza. For simplicity, we assumed that transmission risks between students are determined by class/grade structures and neglected the effect of other potential confounders such as sex, age and precaution measures (therefore, grades in the simulations were only for labelling purpose and did not necessarily correspond to actual school years). The inference model used for the Matsumoto city data and posterior samples were used for simulation, where external infection from outside the school (i.e. transmission from households and general community) was excluded except for the initial case. Starting from a single initial case on day 1, the simulation of transmission over 360 days (we did not consider weekends and school holidays for simplicity) was repeated 500 times, each with a different set of posterior samples of parameters.

For each of the assumed value of R_S (1.8, 1.2 and 0.8), we rescaled the posterior samples of the proximity-specific reproduction number R_d such that the relative magnitude between R_d is conserved and that $\sum_{d=1}^{3} R_d = R_S$. Different types of interventions (see Table 2) were incorporated into the simulation as follows. R_d values corresponding to different n and m were used to simulate the effect of changes in the size and the number of classes. Screening by symptoms and regular testing was implemented by using the modified infection profile in Equation (5). To represent intermittent schooling interventions, the values of infection profile on "off" days were manually set to zero. For reduced outside-class interactions scenarios, we reduced R_d values corresponding to outside-class interactions by either 50% or 90%. In addition to the "pure reduction" scenarios where outside-class interactions are reduced without counter-effects, we also accounted for a possible compensatory increase in the within-class interactions. We assumed that within-class interactions may increase by 20% to compensate for a 50% reduction outside-class and by 40% to compensate for a 90% reduction.

Using the distribution of final outbreak size with a single initial case $q_1(x)$ obtained in the simulation, we also estimated the risk of large outbreaks (i.e. > 10 and > 30 secondary transmissions)

137

given multiple introductions. The final outbreak size distribution given z introductions $q_z(x)$ is obtained as a z-fold convolution of $q_1(x)$:

$$q_z(x) = \sum_{x_1,x_2,\ldots,x_z=0}^{x} \delta_{(x)\left(\sum_{k=1}^{z} x_k\right)} \prod_{l=1}^{z} q_1(x - x_l), \tag{6}$$

where $\delta_{(x)(y)}$ is the Kronecker delta.

Assessing the risk of undetected spread of infection

We computed the distribution of the number of unnoticed infections by the detection of the first case in school by sampling the date of detection in each of the 500 simulation results. Let I_t be the number of new infections on day t. The cumulative distribution function (CDF) for the date of detection T_D is given as

$$\text{CDF}(T_D) = 1 - \prod_{t=1}^{T_D} \left((1 - \sigma)^{T_D - t} \left(1 - F_{T_D - t}\right) \right)^{I_t}. \tag{6}$$

We sampled T_D according to this CDF and obtained the number of undetected infections as $\sum_{t=1}^{T_D} I_t - 1$. The class which the first detected student belongs to was also sampled to provide the number of undetected infections outside that class, which was used to specify the spillover risk.

Simulation of single-class closure strategy

The single-class closure strategy was simulated using the same approach as previously described, except that classes have either an 'open' or 'closed' state each day. Students in closed classes were considered to be isolating at home and thus do not transmit to or receive infection from others on that day. For each infected student, the date of detection was sampled with the distribution in Equation (6) and the class closure started from the day after the first date of detection among the class. The class closure was assumed to last for 14 days (COVID-19) or 7 days (pandemic influenza). To assess the effectiveness of closure strategies, we compared the proportion of students experiencing infection by the end of the outbreak against the simulation results in the same settings but without closures.

All analysis was performed in Julia 1.2.0. Replication code is available on GitHub (https://github.com/akira-endo/schooldynamics_FluMatsumoto14-15).

Acknowledgement

Conflict of interest

AE received a research grant from Taisho Pharmaceutical Co., Ltd.

References

1.   European Centre for Disease Prevention and Control. COVID-19 in children and the role of
        school settings in COVID-19 transmission, 6 August 2020. Stockholm; 2020.

2.   United Nations Educational Scientific and Cultural Organization. COVID-19 Impact on
        Education: Global monitoring of school closures caused by COVID-19 (interactive map).
        [cited 20 Nov 2020]. Available: https://en.unesco.org/covid19/educationresponse

3.   Department for Education; UK Government. Guidance for full opening: schools (Updated 5
        November 2020). 2020 [cited 20 Nov 2020]. Available:
        https://www.gov.uk/government/publications/actions-for-schools-during-the-coronavirus-
        outbreak/guidance-for-full-opening-schools

4.   Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus
        Disease 2019 (COVID-19) Outbreak in China. JAMA. 2020;323: 1239.
        doi:10.1001/jama.2020.2648

5.   Centers for Disease Control and Prevention. Demographic Trends of COVID-19. 2020 [cited
        20 Nov 2020]. Available: https://www.cdc.gov/covid-data-tracker/index.html#demographics

6.   He J, Guo Y, Mao R, Zhang J. Proportion of asymptomatic coronavirus disease 2019: A
        systematic review and meta-analysis. J Med Virol. 2020; jmv.26326. doi:10.1002/jmv.26326

7.   Götzinger F, Santiago-García B, Noguera-Julián A, Lanaspa M, Lancella L, Calò Carducci FI,
        et al. COVID-19 in children and adolescents in Europe: a multinational, multicentre cohort
        study. Lancet Child Adolesc Heal. 2020;4: 653–661. doi:10.1016/S2352-4642(20)30177-2

8.   Wyllie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M, Vijayakumar P, et
        al. Saliva or Nasopharyngeal Swab Specimens for Detection of SARS-CoV-2. N Engl J Med.
        2020;383: 1283–1286. doi:10.1056/NEJMc2016359

9.   Szablewski CM, Chang KT, Brown MM, Chu VT, Yousaf AR, Anyalechi N, et al. SARS-

CoV-2 Transmission and Infection Among Attendees of an Overnight Camp — Georgia, June 2020. MMWR Morb Mortal Wkly Rep. 2020;69: 1023–1025. doi:10.15585/mmwr.mm6931e1

10. Stein-Zamir C, Abramson N, Shoob H, Libal E, Bitan M, Cardash T, et al. A large COVID-19 outbreak in a high school 10 days after schools' reopening, Israel, May 2020. Eurosurveillance. 2020;25. doi:10.2807/1560-7917.ES.2020.25.29.2001352

11. Jones TC, Mühlemann B, Veith T, Biele G, Zuchowski M, Hoffmann J, et al. An analysis of SARS-CoV-2 viral load by patient age. medRxiv. 2020; 2020.06.08.20125484. doi:10.1101/2020.06.08.20125484

12. Fontanet A, Tondeur L, Madec Y, Grant R, Besombes C, Jolly N, et al. Cluster of COVID-19 in northern France: A retrospective closed cohort study. medRxiv. 2020; 2020.04.18.20071134. doi:10.1101/2020.04.18.20071134

13. Buonsenso D, De Rose C, Moroni R, Valentini P. SARS-CoV-2 infections in Italian schools: preliminary findings after one month of school opening during the second wave of the pandemic. medRxiv. 2020; 2020.10.10.20210328. doi:10.1101/2020.10.10.20210328

14. Ismail SA, Saliba V, Lopez Bernal JA, Ramsay ME, Ladhani SN. SARS-CoV-2 infection and transmission in educational settings: cross-sectional analysis of clusters and outbreaks in England. medRxiv. 2020; 2020.08.21.20178574. doi:10.1101/2020.08.21.20178574

15. Russell F, Ryan KE, Snow K, Danchin M, Mulholland K, Goldfeld S. COVID-19 in Victorian Schools: An analysis of child-care and school outbreak data and evidence-based recommendations for opening schools and keeping them open. Melbourne; 2020. Available: https://www.dhhs.vic.gov.au/sites/default/files/documents/202009/Report-summary-COVID-19-in-victorian-schools-pdf.pdf

16. Goldstein E, Lipsitch M, Cevik M. On the effect of age on the transmission of SARS-CoV-2 in households, schools and the community. J Infect Dis. 2020. doi:10.1093/infdis/jiaa691

17. Kommuneoverlegen in Lillestrøm municipality, Sagdalen school in Lillestrøm municipality,

Nowegian Institute of Public Health. [Outbreak report 26.08.2020: Outbreak of COVID-19, Sagdalen school, 2020] (Norwegian). 2020. Available: https://www.lillestrom.kommune.no/contentassets/1d174acbc73f49b8be7266f82300c19f/2008 26_utbruddsrapport_sagdalen2020.pdf

18.   Ministry of Education Culture Sports Science and Techonogy Japan. [Hygiene management manual regarding novel coronavirus disease: new life standards at schools (2020.9.3 Ver.4)] (Japanese). 2020. doi:https://doi.org/10.1186/s12976-016-0045-2

19.   Panovska-Griffiths J, Kerr CC, Stuart RM, Mistry D, Klein DJ, Viner RM, et al. Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study. Lancet Child Adolesc Heal. 2020;4: 817–827. doi:10.1016/S2352-4642(20)30250-9

20.   Keeling MJ, Tildesley MJ, Atkins BD, Penman B, Southall E, Guyver-Fletcher G, et al. The impact of school reopening on the spread of COVID-19 in England. medRxiv. 2020; 2020.06.04.20121434. doi:10.1101/2020.06.04.20121434

21.   Begon M, Bennett M, Bowers RG, French NP, Hazel SM, Turner J. A clarification of transmission terms in host-microparasite models: Numbers, densities and areas. Epidemiol Infect. 2002. doi:10.1017/S0950268802007148

22.   Cauchemez S, Bhattarai A, Marchbanks TL, Fagan RP, Ostroff S, Ferguson NM, et al. Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. Proc Natl Acad Sci. 2011;108: 2825–2830. doi:10.1073/pnas.1008895108

23.   Clamer V, Dorigatti I, Fumanelli L, Rizzo C, Pugliese A. Estimating transmission probability in schools for the 2009 H1N1 influenza pandemic in Italy. Theor Biol Med Model. 2016. doi:10.1186/s12976-016-0045-2

24.   Uchida M, Kaneko M, Hidaka Y, Yamamoto H, Honda T, Takeuchi S, et al. Effectiveness of vaccination and wearing masks on seasonal influenza in Matsumoto City, Japan, in the

2014/2015 season: An observational study among all elementary schoolchildren. Prev Med Reports. 2017;5: 86–91. doi:10.1016/j.pmedr.2016.12.002

25. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med. 2020. doi:10.1038/s41591-020-0869-5

26. Ashcroft P, Huisman JS, Lehtinen S, Bouman JA, Althaus CL, Regoes RR, et al. COVID-19 infectivity profile correction. 2020. Available: http://arxiv.org/abs/2007.06602

27. Waterfield T, Watson C, Moore R, Ferris K, Tonry C, Watt A, et al. Seroprevalence of SARS-CoV-2 antibodies in children: a prospective multicentre cohort study. Arch Dis Child. 2020; archdischild-2020-320558. doi:10.1136/archdischild-2020-320558

28. Endo A, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Res. 2020;5: 67. doi:10.12688/wellcomeopenres.15842.3

29. Waterfield T, Watson C, Moore R, Ferris K, Tonry C, Watt AP, et al. Seroprevalence of SARS-CoV-2 antibodies in children - A prospective multicentre cohort study. medRxiv. 2020; 2020.08.31.20183095. doi:10.1101/2020.08.31.20183095

30. Hsieh YH, Tsai CA, Lin CY, Chen JH, King CC, Chao DY, et al. Asymptomatic ratio for seasonal H1N1 influenza infection among schoolchildren in Taiwan. BMC Infect Dis. 2014. doi:10.1186/1471-2334-14-80

31. Wang TE, Lin CY, King CC, Lee WC. Estimating pathogen-specific asymptomatic ratios. Epidemiology. 2010. doi:10.1097/EDE.0b013e3181e94274

32. Weinstein RA, Bridges CB, Kuehnert MJ, Hall CB. Transmission of Influenza: Implications for Control in Health Care Settings. Clin Infect Dis. 2003;37: 1094–1101. doi:10.1086/378292

33. Vink MA, Bootsma MCJ, Wallinga J. Serial Intervals of Respiratory Infectious Diseases: A Systematic Review and Analysis. Am J Epidemiol. 2014;180: 865–875. doi:10.1093/aje/kwu209

34. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DA. Incubation periods of acute respiratory viral infections: a systematic review. Lancet Infect Dis. 2009;9: 291–300. doi:10.1016/S1473-3099(09)70069-6

35. Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature. BMC Infect Dis. 2014. doi:10.1186/1471-2334-14-480

36. Endo A, Uchida M, Kucharski AJ, Funk S. Fine-scale family structure shapes influenza transmission risk in households: Insights from primary schools in Matsumoto city, 2014/15. PLoS Comput Biol. 2019. doi:10.1371/journal.pcbi.1007589

37. Killingley B, Nguyen-Van-Tam J. Routes of influenza transmission. Influenza Other Respi Viruses. 2013;7: 42–51. doi:10.1111/irv.12080

38. Banik RK, Ulrich A. Evidence of Short-Range Aerosol Transmission of SARS-CoV-2 and Call for Universal Airborne Precautions for Anesthesiologists During the COVID-19 Pandemic. Anesth Analg. 2020;131: e102–e104. doi:10.1213/ANE.0000000000004933

39. Klompas M, Baker MA, Rhee C. Airborne Transmission of SARS-CoV-2. JAMA. 2020;324: 441. doi:10.1001/jama.2020.12458

40. Leung KY, Ball F, Sirl D, Britton T. Individual preventive social distancing during an epidemic may have negative population-level outcomes. J R Soc Interface. 2018;15: 20180296. doi:10.1098/rsif.2018.0296

41. Guclu H, Read J, Vukotich CJ, Galloway DD, Gao H, Rainey JJ, et al. Social Contact Networks and Mixing among Students in K-12 Schools in Pittsburgh, PA. Barrat A, editor. PLoS One. 2016;11: e0151139. doi:10.1371/journal.pone.0151139

42. Kucharski AJ, Klepac P, Conlan AJK, Kissler SM, Tang ML, Fry H, et al. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. Lancet Infect Dis. 2020;20:

1151–1160. doi:10.1016/S1473-3099(20)30457-6

43. Uchida M, Kaneko M, Hidaka Y, Yamamoto H, Honda T, Takeuchi S, et al. Prospective epidemiological evaluation of seasonal influenza in all elementary schoolchildren in Matsumoto city, Japan, in 2014/2015. Jpn J Infect Dis. 2017. doi:10.7883/yoken.JJID.2016.037

44. Longini IM, Koopman JS. Household and community transmission parameters from final distributions of infections in households. Biometrics. 1982;38: 115–126.

45. Blackwell M, Honaker J, King G. A Unified Approach to Measurement Error and Missing Data: Overview and Applications. Sociol Methods Res. 2017. doi:10.1177/0049124115585360

46. Cowling BJ, Fang VJ, Riley S, Malik Peiris JS, Leung GM. Estimation of the serial interval of influenza. Epidemiology. 2009. doi:10.1097/EDE.0b013e31819d1092

47. Levy JW, Cowling BJ, Simmerman JM, Olsen SJ, Fang VJ, Suntarattiwong P, et al. The serial intervals of seasonal and pandemic influenza viruses in households in Bangkok, Thailand. Am J Epidemiol. 2013. doi:10.1093/aje/kws402

48. Leclerc QJ, Fuller NM, Knight LE, Funk S, Knight GM. What settings have been linked to SARS-CoV-2 transmission clusters? Wellcome Open Res. 2020;5: 83. doi:10.12688/wellcomeopenres.15889.2

8.2 Supplementary materials

Supplementary materials: Within and between classroom transmission patterns of seasonal influenza and pandemic management strategies at schools

Akira Endo, CMMID COVID-19 Working Group, Mitsuo Uchida, Katherine E. Atkins, Adam J. Kucharski, Sebastian Funk

Model specifications

Transmission model

Our dataset consisted of four components: final disease outcome D (1 for cases and 0 for controls), onset date of students T (NA for controls), household data H (i.e. household composition and how many members have had disease) and individual covariates X. The likelihood for student i with data components $\{D_i, T_i, H_i, X_i\}$ is given as

$$\begin{cases} L_i = p(H_i|\theta^H)\Gamma_i^S(T_{\max})\Gamma_i^H(T_{\max})\Gamma_i^C(T_{\max}) & (D_i = 0) \\ L_i = p(T_i, H_i|\theta^H, \theta)\Gamma_i^S(T_i - 1)\Gamma_i^H(T_i - 1)\Gamma_i^C(T_i - 1) & (D_i = 1) \end{cases} \quad \text{(S1)}$$

where $\Gamma_i^X(T)$ is the probability that student i survives the force of infection in settings X (S: school, H: household, C: general community) until time T. The first term of each product, $p(H_i|\theta^H)$ or $p(T_i, H_i|\theta^H, \theta)$, represents the probability of observing household data H_i and onset date T_i (if $D_i = 1$) given sets of parameters θ^H and θ. The parameter set θ^H consists of fixed parameters governing the within-household transmission, which we retrieved from the previous study on the same study cohort [36]. All other parameters θ are estimated. Since θ^H are assumed to be fixed, the likelihoods in Equation (S1) can be simplified as

$$\begin{cases} L_i \propto \Gamma_i^S(T_{\max})\Gamma_i^C(T_{\max}) & (D_i = 0) \\ L_i \propto p(T_i, H_i|\theta^H, \theta)\Gamma_i^S(T_i - 1)\Gamma_i^C(T_i - 1) & (D_i = 1) \end{cases} \quad \text{(S2)}$$

The survival probabilities for the school and community settings are modelled as

$$\Gamma_i^{S}(T) = \exp\left(-v_i \sum_j w_j \beta_{ij} \sum_{t=T_j+1}^{T} h_{t-T_j}\right),$$

$$\Gamma_i^{C}(T) = \exp(-v_i r_{C} \Lambda_{C}(T)),$$
$$\tag{S3}$$

where $\Lambda_{C}(T)$ is the cumulative density function of a logistic curve representing the time trend of community outbreak (see Section "Community transmission" for details)

The likelihood for student's onset and household episodes $p(T_i, H_i | \theta^{H}, \theta)$ is obtained as follows. First, the probability that student i has illness onset on T_i due to infection either from school or general community is

$$p_i^{S+C}(T_i) = 1 - \exp[\lambda_i^{S}(T_i) + \lambda^{C}(T_i)], \tag{S4}$$

where $\lambda_i^{S}(T_i)$ is as specified in Equation (3) in the main text and $\lambda^{C}(T_i)$ represents the hazard from community outbreak given as $\lambda^{C}(T) = r_{C}\frac{d}{dT}\Lambda_{C}(T)$. This probability $p_i^{S+C}(T_i)$ is then plugged into the household transmission model. In the prospective survey, household data H_i consisted of household cases simultaneously reported with the student's influenza episode. Since the illness onset dates were not reported for household cases, it was not possible to determine the direction of within-household transmissions. We assumed that the reported household cases represent those who could be linked to the student's onset (household cases infecting the student or vice versa) and thus their onset dates should be close enough to that of the student. It is also possible that they had been coprimary cases, i.e. unlinked infections separately obtained from outside the household, but we expect the onset dates to be within the same range even in such cases. Given the probability of a student acquiring disease from outside the household $p^{S+C}(T_i)$, the likelihood of household data H_i is given as

$$p_i(T_i, H_i) = p_i^{S+C}(T_i)\pi\left((0, H_i)\big|\theta^{H}\right) + \left(1 - p_i^{S+C}(T_i)\right)\pi\left((1, H_i)\big|\theta^{H}\right), \tag{S5}$$

where π is the likelihood of observing household final outcome (x, H_i) (i.e. which members of the household had influenza) used in the previous study [36]. The household likelihood model was parameterised using the median estimates reported in [36] except the external risk of infection (see "Community transmission"). The use median estimates to summarise the posterior samples could induce underestimation; we adopted this approximation for computational convenience, which was

148

unlikely to have substantially affected our conclusions as the contribution of the household likelihood to the qualitative results was relatively minor (see "Additional analysis").

Community transmission

We fitted a logistic curve $\hat{\lambda}_C(T) = \frac{a_3}{1+\exp(-a_1(T-a_2))}$ to the aggregated incidence data of students to represent the time trend of community outbreak. The Poisson likelihood was maximised to infer the parameters a_1, a_2 and a_3. The fitted logistic curve was normalised, i.e. $\lambda_C(T) = \frac{d}{dT}\Lambda_C(T) = \frac{\hat{\lambda}_C(T)}{\int_0^\infty \hat{\lambda}_C(T)dT}$, and used in Equation S2 and as a part of household likelihood π. Since the parameter estimates for the external risk of infection (probability of infection from outside the household) in [36] corresponded to the cumulative risk over the season, we rescaled them to account for the shorter time windows in the current analysis. We assumed that influenza episodes of household members accompanying the reported episode of a student occurred within the 3 days' range of the reported onset date of students; otherwise they might not have been reported as "coincided episodes". The external risk of infection ε_k for a type k household member ('sibling', 'father', 'mother', or 'other') as estimated in [36] was rescaled as

$$\varepsilon_k' = 3\varepsilon_k\lambda_C(T), \tag{S6}$$

where T is the onset date of the student. As a sensitivity analysis, we confirmed that different choice of the window range (7 days instead of 3) minimally affected the estimates of R_d (see "Additional analysis).
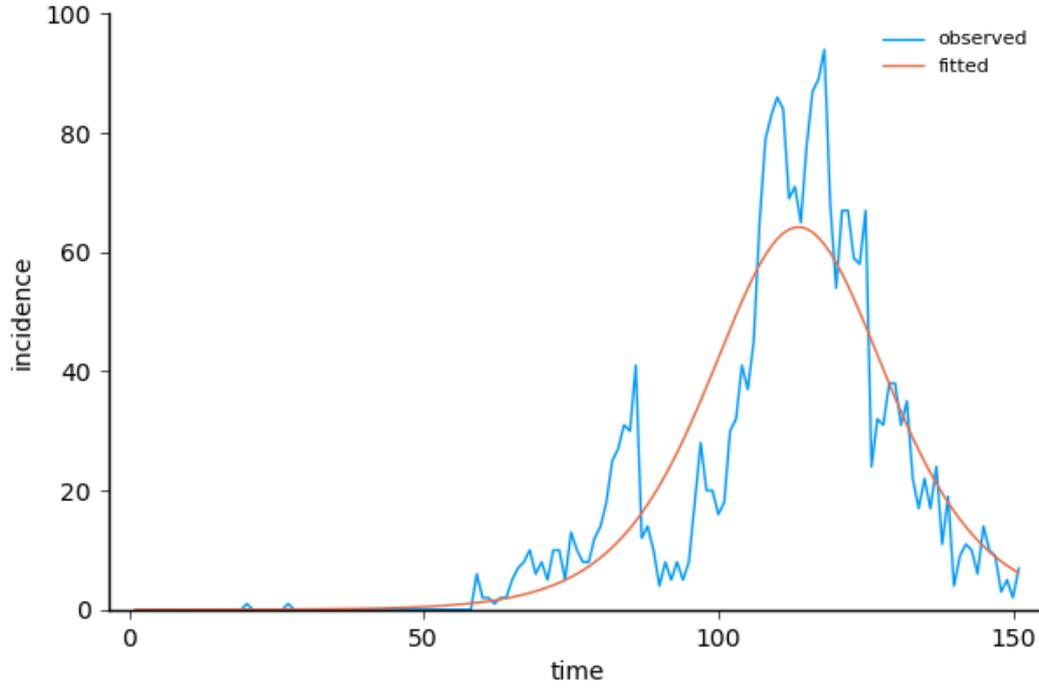
Figure S1. Logistic curve fitted to the observed incidence to model temporal trend in community transmission. The blue curve shows the aggregated incidence of students. The x-axis represents dates of illness onset, where day 1 corresponds to 1 October 2014.

Adjustment for potential confounders

Potential confounders $\boldsymbol{\xi}_i = \left(\xi_i^1, \xi_i^2, \dots\right)$ that may affect the susceptibility or infectiousness (shown in Table 1 in the main text) were addressed by loglinear regression:

$$v_i = \exp\left(\boldsymbol{\xi}_i^{\mathrm{T}} \boldsymbol{\alpha}_i^v\right),$$
$$w_i = \exp\left(\boldsymbol{\xi}_i^{\mathrm{T}} \boldsymbol{\alpha}_i^w\right),$$

(S7)

where $\boldsymbol{\alpha}_i^v$ and $\boldsymbol{\alpha}_i^w$ denote vectors of regression coefficients. We assumed that the variables v_i and w_i are involved in school and community transmission and not in household transmissions for methodological convenience, based on our assumption that protective effects including precaution measures may not be as effective inside households as outside. We believe this assumption is unlikely to have biased our inference on within and between class transmission patterns because exclusion of the household likelihood barely changed the estimates in our sensitivity analysis (see Supplementary materials "Additional analysis").

150

Since the potential confounders included self-reported precaution measures (vaccination, mask and hand washing), there were mismatches between the responses in the prospective and retrospective surveys that warranted our attention. Due to the different nature of recruitment (prospective survey accompanied leave-of-absence forms while retrospective survey was a mass questionnaire), the retrospective survey had a lower respondent rate (86%) than the prospective survey (96%). Limiting to those reporting influenza episodes ('cases') who were eligible to both the prospective and retrospective surveys, the retention rate (the proportion of prospective survey respondents who remained in the retrospective survey) was 84% overall. However, the retention rate by responses on the precaution measures showed inconsistent patterns (Table S1); more positive responses remained in the prospective surveys on "vaccination" and "mask" and vice versa on "hand washing". Notably, the retention rate of "hand washing: No" is over 100%, suggesting that there were students who answered "No" only in the retrospective survey (as the records were unlinkable between the surveys, it was not possible to confirm which students had mismatched responses). Although this might be explained by differential retention rates (i.e. those answering "Yes" on vaccination and masks and those answering "No" on hand washing were more likely to remain in the retrospective survey), more plausible would be that in the retrospective survey, students can give answers different from the prospective survey. This may be due to the recall bias because the retrospective survey was conducted in March, which could have been up to 4-5 months after the students' onset, and can cause biased estimates because the responses of the control group, only available for the retrospective survey, may also be inconsistent with their actual behaviour during the outbreak. We assumed that the responses of the case group in the prospective survey reflected the actual behaviours of students during the epidemic season and that the answers of both case and control groups in the retrospective survey are potentially misclassified. By treating the responses in the prospective survey as reference, the sensitivity and specificity of responses in the retrospective survey can be estimated. Assuming nondifferential misclassification with shared sensitivity and specificity across variables, we assessed and adjusted for recall bias in the dataset.

The sensitivity and specificity of the responses in the retrospective survey was jointly estimated with the retention rates in a form of a matrix M representing probabilities of retention and classification. Let $N^{\mathrm{p}}_{\xi=x}$ and $N^{\mathrm{r}}_{\xi=x}$ be the number of responses x ("Yes": $x = 1$; "No": 0) on a covariate ξ in the prospective and retrospective surveys, respectively. The expectancy of $N^{\mathrm{r}}_{\xi=x}$ can be given as

$$\begin{pmatrix} E\left(N^{\mathrm{r}}_{\xi=0}\right) \\ E\left(N^{\mathrm{r}}_{\xi=1}\right) \end{pmatrix} = M \begin{pmatrix} N^{\mathrm{p}}_{\xi=0} \\ N^{\mathrm{p}}_{\xi=1} \end{pmatrix}, \tag{S8}$$

and we maximised the corresponding Poisson likelihood to estimate M. The responses in the prospective survey reconstructed from those in the retrospective survey and the inverted matrix M^{-1} were overall consistent with the observed responses (Table S1).

Table S1. Comparison of observed and reconstructed confounders

Covariate	Survey responses			Model prediction
	Prospective	Retrospective	Retention rate	Prospective (reconstructed)
Vaccination	Yes: 1122	Yes: 978	87%	Yes: 1102
	No: 1426	No: 1171	82%	No: 1446
Mask	Yes: 1204	Yes: 1069	89%	Yes: 1226
	No: 1344	No: 1080	80%	No: 1322
Hand washing	Yes: 2200	Yes: 1778	81%	Yes: 2199
	No: 348	No: 371	107%	No: 349
Total	2548 cases	2149 cases	84%	—

We used the estimated parameter matrix M to adjust the likelihood function of the control group, whose responses are, by definition, missing in the prospective survey. We used the covariates of the case group as reported in the prospective survey and thus adjustment was not necessary for them. The component of the likelihood which can be affected by this adjustment is given as $\Gamma^{\mathrm{S}}_i(T_{\max}) \, \Gamma^{\mathrm{C}}_i(T_{\max})$ as in Equation (S3). Noting that only v_i is relevant to the adjustment, we get

$$\Gamma_i^{\text{S}}(T_{\max}) \, \Gamma_i^{\text{C}}(T_{\max}) = \exp(-v_i \Lambda_i^{\text{Total}}) = \exp\left(-\Lambda_i^{Total} \, exp(\xi_i^T \alpha_i^v)\right), \tag{S9}$$

where $\Lambda_i^{\text{Total}} = r_{\text{C}} \Lambda_{\text{C}}(T_{\max}) + \sum_j w_j \beta_{ij} \sum_{t=T_j+1}^{T_{\max}} h_{t-T_j}$ (hereafter, let us limit ξ to the three covariates shown in Table S1, where the value of 1/0 indicates Yes/No, respectively).

We then accounted for potential misclassification in the recorded covariates that determined v_i by incorporating an adapted version of the multiple overimputation method [45]. For each covariate, we estimated the Bayesian probabilities $p(\xi_i | \hat{\xi}_i)$, the conditional probability of the true values of binomial variables given the data. Although variables are repeatedly imputed in the original multiple overimputation method, we instead directly obtained the (approximated) adjusted likelihood for computational convenience as

$$\begin{aligned}
\Gamma_i^{\text{S}}(T_{\max}) \, \Gamma_i^{\text{C}}(T_{\max}) &= \sum_{\xi_i} p(\xi_i | \hat{\xi}_i) \exp\left(-\Lambda_i^{Total} \, exp(\xi_i^T \alpha_i^v)\right) \\
&\approx \exp\left(-\Lambda_i^{Total} \sum_{\xi_i} p(\xi_i | \hat{\xi}_i) \, exp(\xi_i^T \alpha_i^v)\right) \\
&= \exp\left(-\Lambda_i^{Total} \prod_k \left[p(\xi_i^k = 0 | \hat{\xi}_i^k) + p(\xi_i^k = 1 | \hat{\xi}_i^k) \, exp(\alpha_i^{v,k})\right]\right).
\end{aligned} \tag{S10}$$

We used the property of approximate linearity assuming $\left| \Lambda_i^{Total} \, exp(\xi_i^T \alpha_i^v) \right| \ll 1$ and also assumed an independence in misclassification.

As a sensitivity analysis, we also performed estimation without the adjustment of recall bias. The results were overall similar to the main results (Table S2), although the negative effect of hand washing on the susceptibility was slightly exaggerated.

Table S2. Loglinear regression results without the recall bias adjustment

Covariate	Frequency in data	Relative susceptibility	Relative infectiousness
Grade (1 year)	—	1.07* (1.02-1.13)	0.95 (0.89-1.01)
Vaccine	47.7%	0.88* (0.79-0.96)	1.00 (0.83-1.19)
Mask	51.4%	0.71* (0.65-0.77)	0.67* (0.55-0.80)
Hand washing	80.1%	1.70* (1.48-1.97)	1.25 (0.94-1.73)
Onset in winter break	5.9% (of cases)	—	0.24* (0.13-0.37)

Addressing sampling bias between case and control groups

Due to the lower respondent rate in the retrospective survey, the original likelihood directly constructed from the raw data underrepresented the control group. Although individual-level data (e.g. confounders and household episodes) was not available for students missing in the control group, it was still possible to estimate the number of such students as both the class sizes and the number of cases in each class were known. To avoid the overestimation of transmission risks this sampling bias could cause, we rescaled the likelihood of the control groups assuming that the individual-level data of included students are also representative of those missing. We did not consider students missing in the case group as the response rate was sufficiently high in the prospective survey (>95%).

The adjusted likelihood of students in a class A of size n_A, where x_A cases and y_A controls are observed is given as

$$\prod_{i \in A} L'_i = \left(\prod_{i \in A, D_i=0} L_i^{\frac{n_A-x_A}{y_A}} \right) \left(\prod_{i \in A, D_i=1} L_i \right). \tag{S11}$$

The first product represents the likelihood of the control group in the class and the second product represents that of the case group. Although this can lead to overconfidence in the loglinear regression results, the degree of such effect is minimal (e.g. 15% inflation of samples only cause 5% underestimation of standard error) and we prioritised reducing bias in transmission risk estimates.

Additional analysis

Sensitivity analysis of within and between class transmission patterns

We assessed the robustness of the estimated transmission patterns of seasonal influenza within and between classes to variations in the following assumptions: (i) serial interval distribution (ii) control of confounders (iii) use of household transmission model (iv) window period length for the household model. For each of these, we performed sensitivity analysis as described below and compared the results of estimated transmission patterns.

(i)   Serial interval distribution: We used the mean serial interval of 2.2 days as estimated in [33], in the main analysis, which is slightly shorter than other estimates [46,47]. We instead used a longer serial interval (3.5 days) as a sensitivity analysis.

(ii)   Control of confounders: The loglinear regression to adjust for potential confounders was excluded.

(iii)   Use of household transmission model: The likelihood compartment accounting for household transmission was excluded from analysis.

(iv)   Window period length for the household model: We assumed a specific window period during which influenza episodes of household members acquired from elsewhere were assumed to be reported along with the student's episode. We used a longer period (7 days) than in the main analysis (3 days).

Overall, our sensitivity analysis suggested that these assumptions had limited effects on the qualitative interpretation of our results on the within and between class transmission patterns of seasonal influenza.
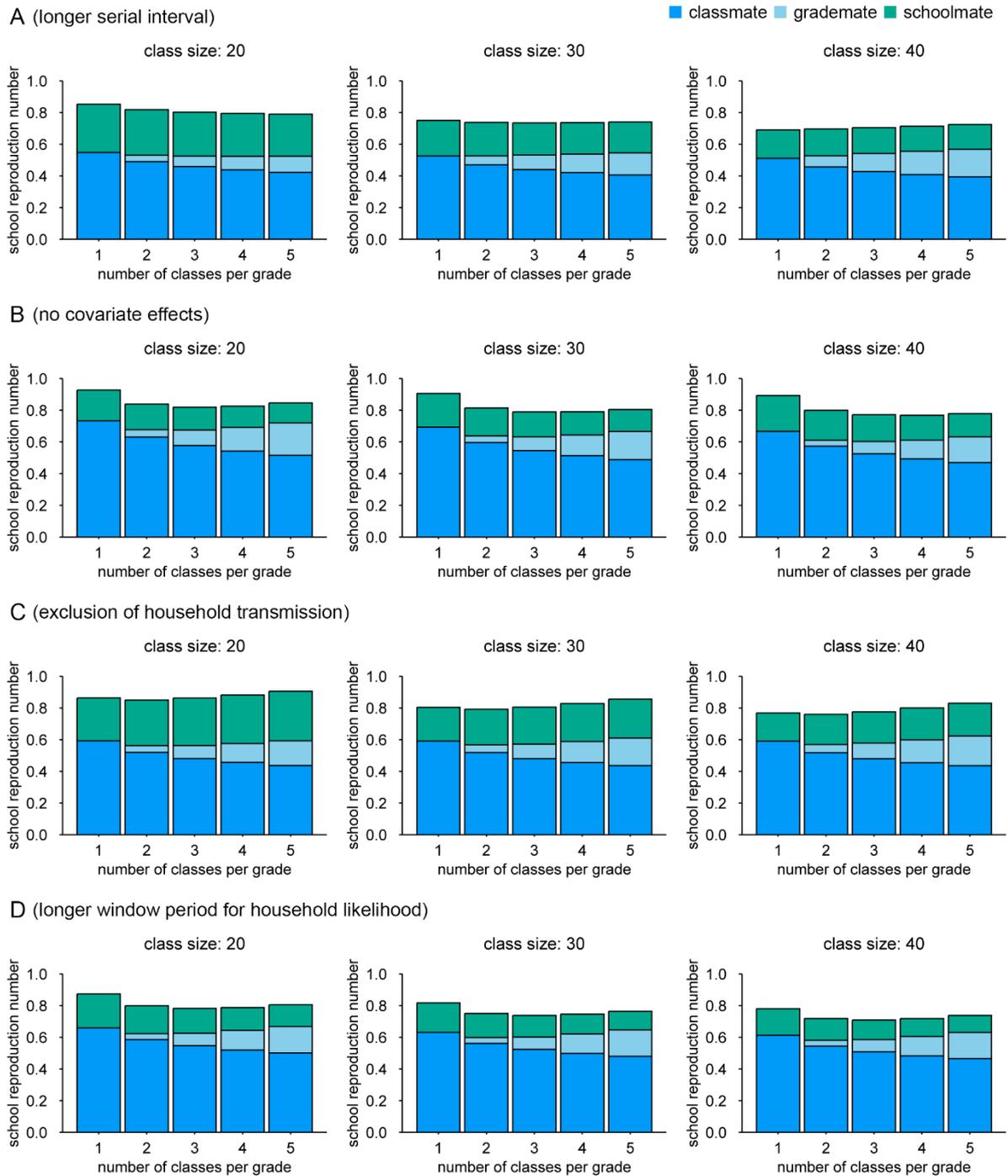
Figure S2. School reproduction number (R_S) and its breakdown by the class/grade relationship (median estimates) corresponding to various settings in the sensitivity analysis. (A) A longer serial interval (mean 3.5 days) was used instead of a mean 2.2 days used in the main analysis. (B) The loglinear regression used to adjust for potential confounders was excluded from analysis. (C) Household transmission model was excluded from analysis. (D) A longer time window for household influenza episodes to be reported along with students' episodes were assumed (7 days instead of 3 days in the main analysis).

156

Overdispersion and the risk of outbreaks

It has been suggested that the transmission of SARS-CoV-2 exhibit a high degree of individual-level variation (overdispersion) [28,48]. As a sensitivity analysis, we considered a negative-binomial overdispersion parameter $\kappa = 0.2$ in the simulation to account for potential superspreading. For each infectious student i, we rescaled the infectiousness variable w_i with a factor which follows a gamma distribution with a shape κ and a scale 1:

$$w_i' = \rho w_i, \tag{S5}$$

$$\rho \sim \text{Gamma}(\kappa, 1).$$

In the presence of overdispersion, the risk of large outbreaks generally becomes smaller given the same reproduction number. However, the differences diminished with multiple introductions, and it was suggested that 10 introductions would pose an almost similar level of outbreak risks to the 'no overdispersion' scenario.
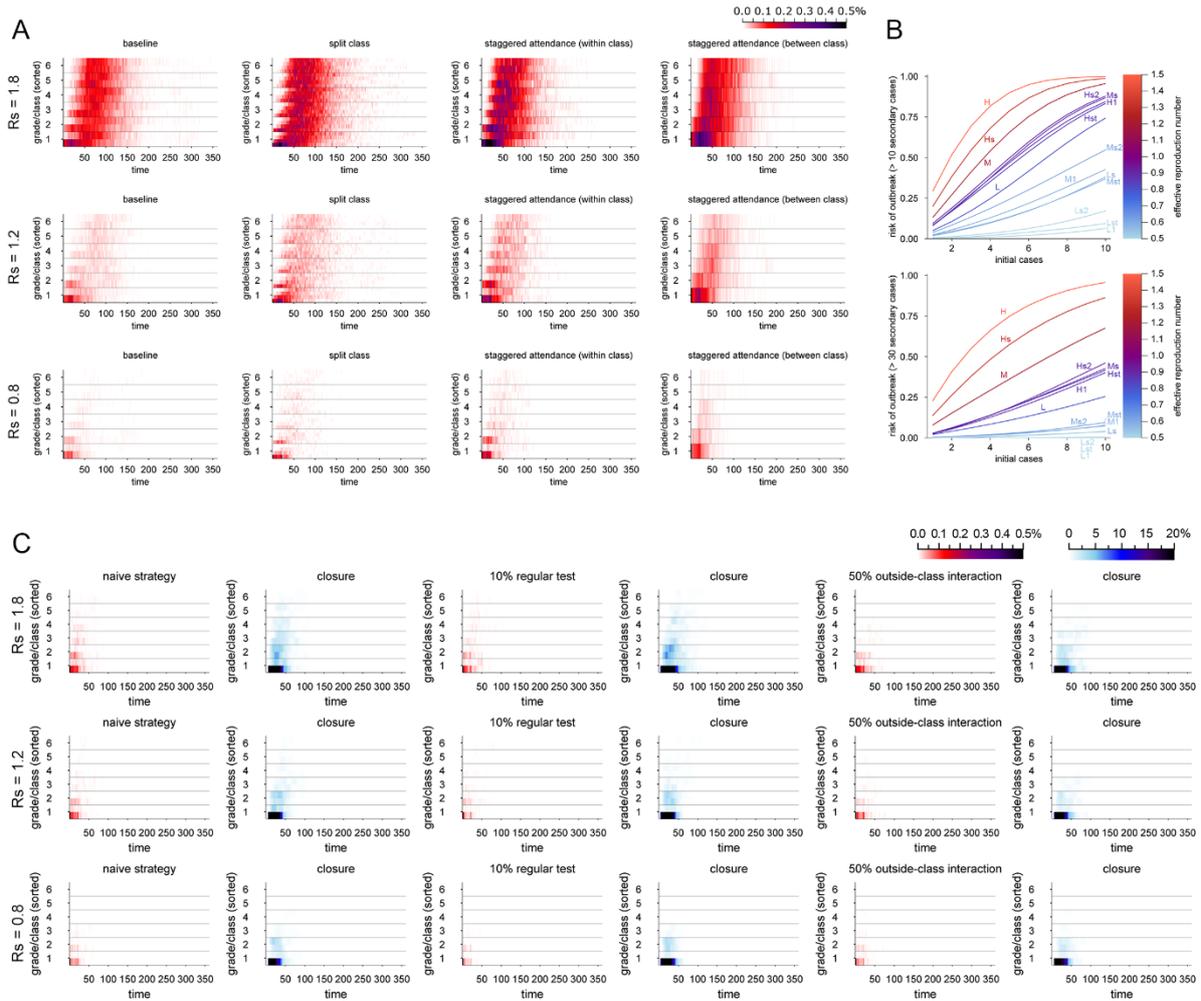
Figure S3. Outbreak simulations accounting for overdispersion in SARS-CoV-2 transmission. (A) Simulated patterns of outbreaks in schools under interventions changing class structures. Colours represent the mean class incidence rate (the number of new infections on a single day in each class divided by the class size) over the 500 simulations. (B) The estimated risk of large outbreaks with multiple introductions. Curves show the probability that the eventual number of secondary transmissions within school exceeds 10 or 30 cases in the intervention scenarios, given multiple introductions of infected student from outside the school. Interventions are labelled by the following notations. H: the school reproduction number (R_S) = 1.8; M: R_S = 1.2; L: R_S = 0.8; s: screening by symptoms; t: screening by regular testing (effective rate 10%); 1: "1 day on: 1 day off" intermittent schooling; 2: "2 days on: 1 day off" intermittent schooling. Colours denote the effective reproduction number within school for each intervention. (D) Simulated temporal patterns of outbreaks and class closures with different closure strategies (symptomatic proportion: 50%). Colours represent the mean class incidence rate and the proportion of a class being closed over the 500 simulations.

Simulation results for pandemic influenza

Here we show some of the simulation results for pandemic influenza (Figure S4). The main results were presented in the main text and the remaining results which showed almost equivalent patterns to the SARS-CoV-2 simulations are displayed for the sake of completeness.
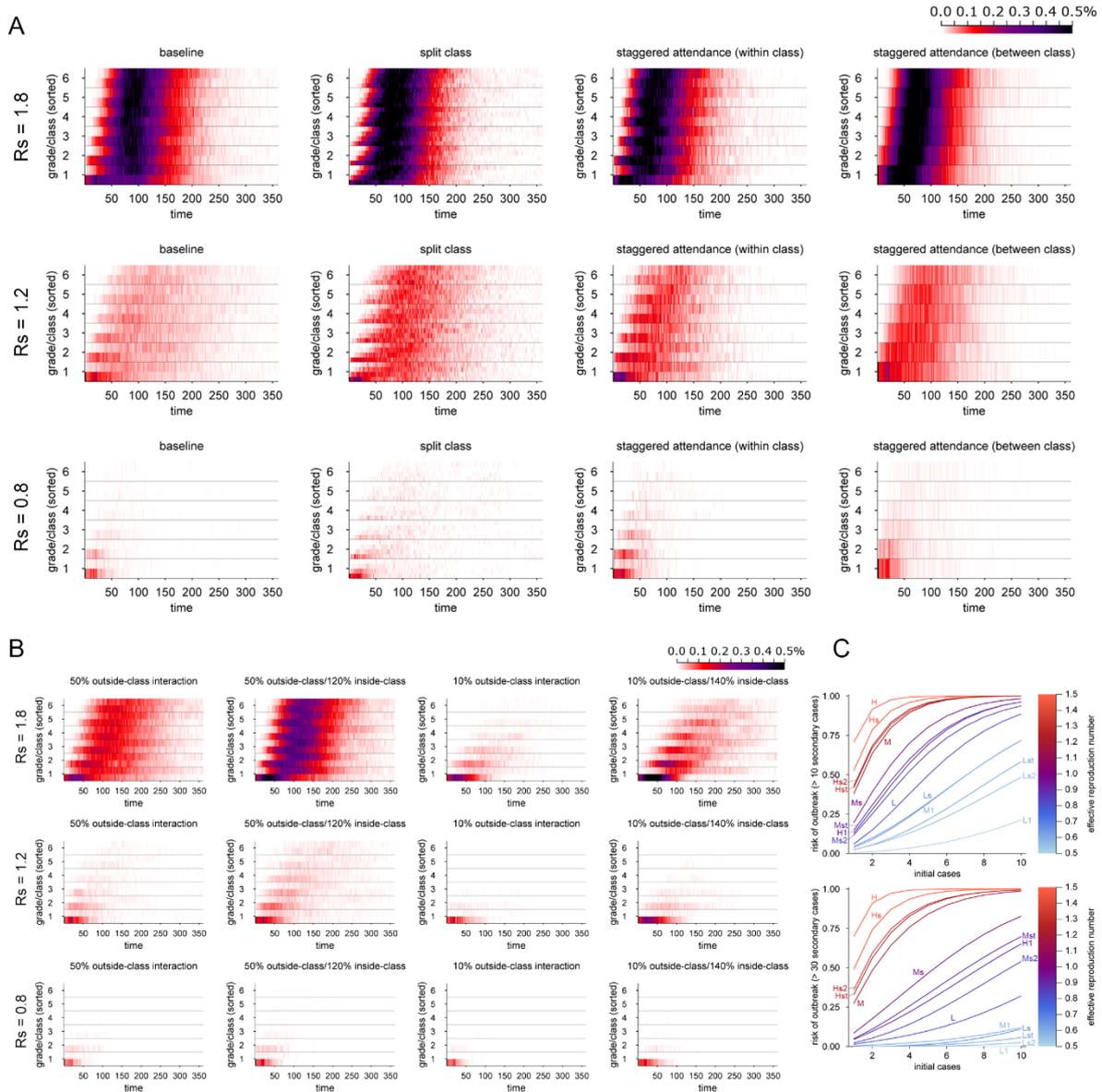
Figure S4. Additional simulation results of pandemic influenza outbreaks. (A) Simulated patterns of outbreaks in schools under interventions changing class structures. Colours represent the mean class incidence rate (the number of new infections on a single day in each class divided by the class size) over the 500 simulations. (B) Simulations with reduced outside-class interactions. Compensatory increases in the within-class interactions (20% and 40% increase in within-class interactions to compensate 50% and 90% reduction in outside-class interactions, respectively) were also considered as part of the simulation. (C) The estimated risk of large outbreaks with multiple introductions. Curves show the probability that the

eventual number of secondary transmissions within school exceeds 10 or 30 cases in the intervention scenarios, given multiple introductions of infected student from outside the school. Interventions are labelled by the following notations. H: the school reproduction number (R_S) = 1.8; M: R_S = 1.2; L: R_S = 0.8; s: screening by symptoms; t: screening by regular testing (effective rate 10%); 1: "1 day on: 1 day off" intermittent schooling; 2: "2 days on: 1 day off" intermittent schooling. Colours denote the effective reproduction number within school for each intervention.

Lower symptomatic proportion and closure strategies

We assumed a lower symptomatic proportion (25%) than used in the main analysis and assessed how it may affect the results of closure strategies. Both in the simulations of SARS-CoV-2 and pandemic influenza, lower symptomatic proportion resulted in larger outbreak sizes and broader class closure (Figure S3). Symptomatic screening becomes less likely to identify cases before they transmit and regular testing was suggested to be more effective.
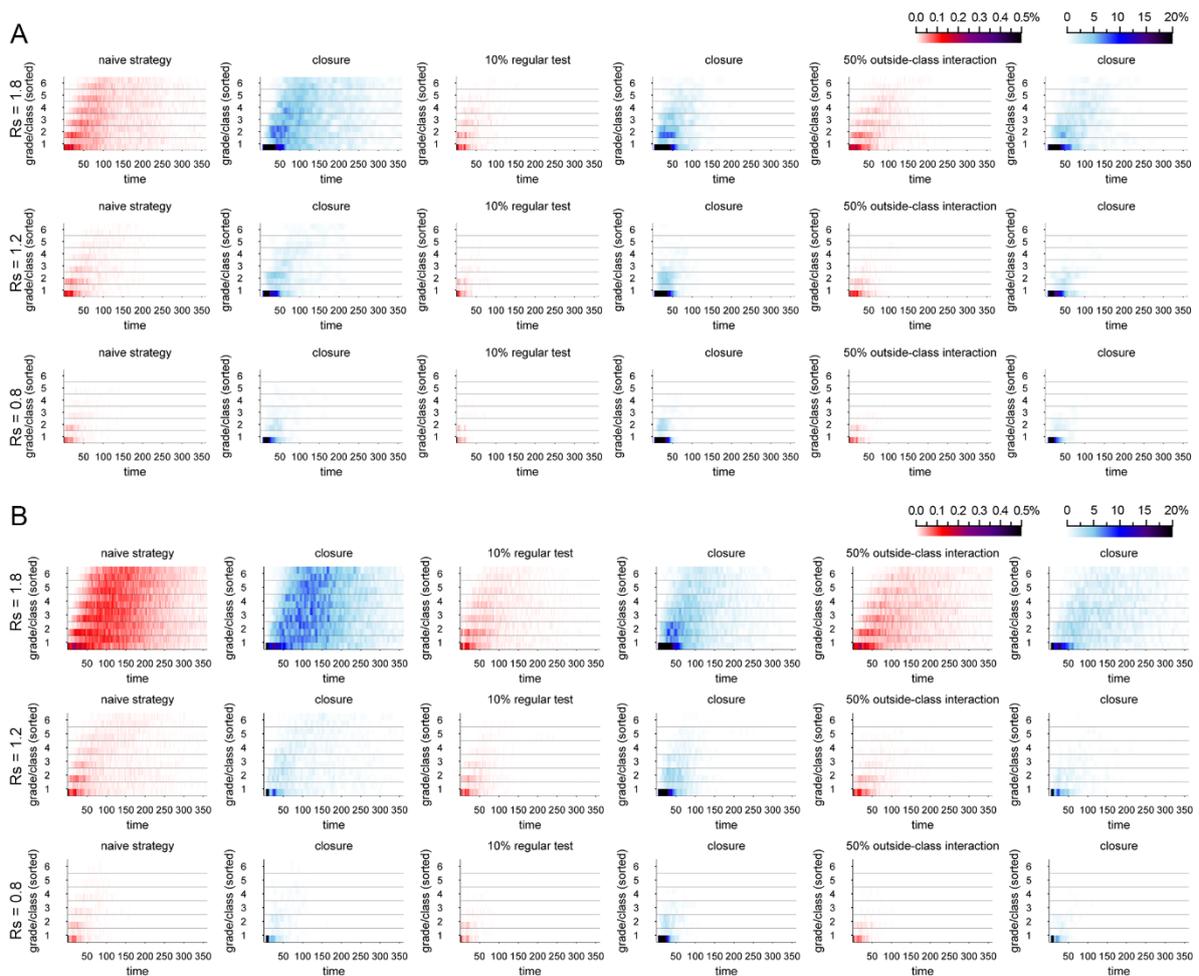
Figure S5. Simulation of single-class closure strategies with a lower symptomatic proportion of 25%. Colours represent the mean class incidence rate and the proportion of a class being closed over the 500 simulations. (A). Simulation of single-class closure strategies for SARS-CoV-2. (B). Simulation of single-class closure strategies for pandemic influenza.

Compensatory increase in within-class interactions and closure strategies

In the main analysis where we compared single-class closure strategy combined with class distancing, we excluded the potential compensatory increase in the within-class interaction. We reflected this effect on the class closure simulations where the inside-class interactions were increased by 20% and 40% respectively in scenarios with 50% and 90% reduction in outside-class interactions (Figure S6). With responsive class closures, increase in the within-class reproduction number hardly affected the outcome because the classes with at least one detected case was closed in the simulation regardless of how far infections had spread within the class, and further transmissions to outside the class were prevented.
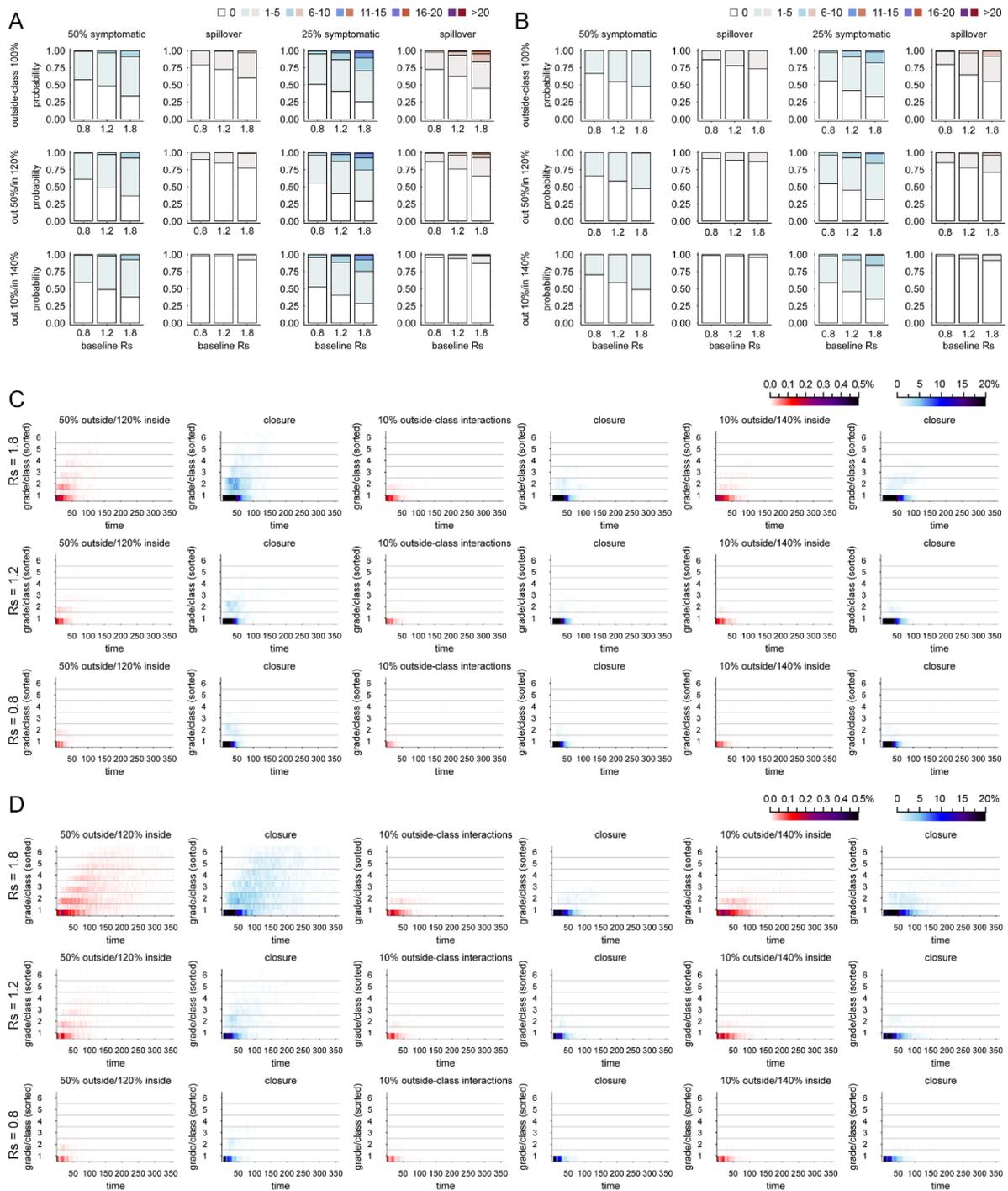
Figure S6. Predicted single-class closure outcomes with class distancing in the presence of compensatory increase in within-class interactions. (A)The predicted distributions of the number of unnoticed infections with SARS-CoV-2 by the first identification of a case (blue: overall; red: spillover infections). (B) The predicted distributions of the number of unnoticed infections with pandemic influenza by the first identification of a case. (C) Simulated temporal patterns of SARS-CoV-2 outbreaks and class closures with different closure strategies (symptomatic proportion: 50%). Colours represent the mean class incidence rate and the proportion of a class being closed over the 500 simulations. (D) Simulated temporal patterns of pandemic influenza outbreaks and class closures with different closure strategies (symptomatic proportion: 50%).

162

9   Discussion and conclusion

9.1   Overview of study findings and strengths

This thesis studied the role of heterogeneity in infectious disease epidemiology—specifically, how individual-level and network-level heterogeneities characterise the transmission patterns of respiratory infectious diseases, how accounting for heterogeneity could alter the handling and interpretation of the data and what benefits quantification of such heterogeneity could bring to infectious disease control. Particular focus was placed on two major acute respiratory infectious diseases both of which have been involved in previous and current (and potentially future) pandemics: influenza and COVID-19. By investigating epidemiological and public health-oriented research questions with mathematical models, implications of heterogeneity on the disease dynamics were presented and discussed from multiple aspects.

In Paper 1, the heterogeneous transmission dynamics of seasonal influenza within households of primary school students were modelled using an extended version of the Longini-Koopman model. The risk of acquiring influenza from outside the household and the risk of within-household transmission were separately estimated from the household-level final outcome data specifying who in each household experienced an influenza episode during the study period. The model selection suggested the importance of heterogeneity in both external and within-household transmission risks. The parameter estimation yielded markedly high risk of infection from outside the household for children (~20%) compared with adults (1-3%). Within-household transmission was suggested to be frequent within the same generation (between siblings, between parents and between grandparents) and between mother-child pairs. By combining the external risk of infection and within-household transmission risk, children were estimated to be responsible for most secondary transmission in households. These results highlight the importance of preventing influenza outbreaks in schools to protect not only students but also their family members, who may include vulnerable groups such as infants and the elderly. In this study, heterogeneity was incorporated into the model by a categorical variable for 'familial roles' (e.g. sibling, father, mother and grandparents). Although the model itself is relatively simple with only 11 parameters and the individuals in the dataset were characterised by one categorical variable (with five possible values), the model explained well the observed pattern of influenza episodes in 10,000 households. The following two factors may have in particular contributed to model performance. Firstly, the applied model was well-suited for the problem and had minimal but sufficient complexity to represent the observed patterns in the dataset: transmission patterns of influenza in households. In the model used in Paper 1, the within-household transmission risks were represented as a weighted network of transmission potential β_{ij}. As β_{ij} was determined not only

by the combination of the type (familial role) of *i* and *j* but also the 'effective household size' (which can be interpreted as the degree of a node in weighted networks), within-household transmission patterns modelled as a collection of small weighted networks successfully captured the substantial part of heterogeneity in the dataset. Secondly, the dataset had a strong signal, i.e. the selected variable (familial roles) was a dominant determinant of the transmission patterns and the dataset had a sufficient sample size to detect that determinant in a fine-scale. Due to a large number of households included, the dataset had a wide range of household compositions and disease outcomes of the household members. These factors may have assisted the most suitable and robust model to be identified.

Paper 2 proposed a statistical method to correct misclassification bias in the test-negative design (TND) vaccine effectiveness studies, a design that has been frequently used for influenza. Individual-level heterogeneity in the study participants is often accounted for by including potential confounding variables into the model. In TND studies, misclassification due to an imperfect test performance can lead to a bias in the estimated effectiveness of vaccines and is especially relevant when rapid diagnostic tests are in use. Paper 2 first quantified the expected degree of bias in the vaccine effectiveness (VE) estimates in TND studies in a wide range of settings and showed that the estimates can be significantly biased in the presence of misclassification. Despite the common perception that specificity is more important in causing misclassification bias than sensitivity, Paper 2 also showed that the contribution of sensitivity can be more substantial under certain parameter settings. Moreover, it was suggested that the degree of bias in the estimates tend to be larger when the study participants exhibit higher heterogeneity, i.e. when more covariates are associated with the outcome. To address this issue, bias correction methods that yields asymptotically unbiased estimates given known sensitivity and specificity were proposed. The method designed for the heterogeneous population adopted a multiple overimputation approach (Blackwell et al., 2017), which enabled the users to couple this method with existing estimation tools with minimal effort. The performance of the methods was assessed by simulation. The simulation results suggested that the proposed methods reliably estimated true VE from misclassified data, while the raw estimates without correction underestimated VE. This study presented important implications on the possible caveats of VE studies in heterogeneous populations and how they could be addressed statistically. Although the applications of these methods in school settings as originally planned were excluded from the scope of this thesis to incorporate more COVID-19 related studies, some of the key concepts of this study were partly incorporated in the following studies, e.g. multiple overimputation of confounding variables in Paper 5. Moreover, TND studies may also play an important role in VE studies for SARS-CoV-2 vaccines, although this was outside the scope of this thesis. After vaccines become available

population-wide, bottom-up VE estimates from field TND studies will provide post-hoc evaluation of vaccines. Such data may be subject to misclassification bias, and the proposed bias correction methods can prevent VE from being underestimated.

Paper 3 was initiated in response to the emergence of the 2020 COVID-19 pandemic. To quantify the degree of heterogeneity in the transmission of SARS-CoV-2, a final outbreak size model was applied to the international case count data. Assuming a basic reproduction number R_0 of 2-3, the results suggested a substantial degree of individual-level heterogeneity represented as a negative-binomial overdispersion parameter k of around 0.1. This value was in line with the early estimates for SARS-CoV and MERS-CoV, both exhibiting substantial overdispersion. When $R_0 = 2.5$, the overdispersion parameter $k = 0.1$ suggests that 10% of primary cases are responsible for 80% of secondary transmissions; therefore, identifying and controlling settings that promote superspreading are of paramount importance for containment of COVID-19. In contrast to the types of heterogeneities discussed in Papers 1 and 2, the heterogeneity quantified in this paper was purely empirical and was not explicitly related to specific characteristics of individuals. While multiple factors have been suggested as potential determinants of overdispersion (Asadi et al., 2019; Leclerc et al., 2020; Nishiura et al., 2020; Riediker & Tsai, 2020), their relative contributions to the observed heterogeneity remain to be unclear. Still, quantifying the degree of heterogeneity without specifying the determinants could be useful in understanding transmission dynamics. Paper 3 provided one of the earliest estimates of the degree of overdispersion in the transmission of SARS-CoV-2. While the possibility of substantial overdispersion had been suggested from the very beginning of the outbreak, formal quantitative analyses based on empirical data had been scarce. Prior to Paper 3, Nishiura et al. and Grantz et al. provided insights into the overdispersion. Nishiura et al. obtained the empirical offspring distribution based on the contact tracing of identified cases in Japan. They showed that 80% of cases did not transmit the virus to anyone and that those who generated a substantial number of cases were associated with the closed indoor environment (odds ratio: 18.7) (Nishiura et al., 2020). However, their data was likely to be underrepresented because the mean number of secondary cases was 0.4, much lower than both the regional and global estimates. Grantz et al. published an online report in which they estimated k using international case counts (Grantz et al., 2020). However, their approach was intended to be for demonstration purpose rather than a formal estimation, and thus its methodological validity was not fully established. Paper 3 was substantially inspired by this report and attempted to further develop the concept and methodology. Namely, Paper 3 shared the aim and underlying data with Grantz et al. but introduced (i) a formal likelihood-based estimation and (ii) adjustment for censoring of the observed case counts (i.e. the current case counts only serve as the lower bound of the final outbreak size if countries have an ongoing outbreak) to yield a

more reliable estimate of k. As a result, Paper 3 achieved real-time estimation of overdispersion in the initial phase of an ongoing outbreak by a statistically sound procedure. This study was followed by a number of studies that quantified the overdispersion in SARS-CoV-2 transmission using various methods and used as a reference estimate in many of those studies.

Based on the findings of Paper 3, Paper 4 discussed practical implications of the substantial overdispersion in the SARS-CoV-2 transmission on epidemic control. Specifically, the relationship between the degree of overdispersion and the effectiveness of contact tracing was mathematically quantified and a possible benefit of combining backward contact tracing was assessed by simulation. Using a simple branching process model, Paper 4 showed that, unlike the traditional forward contact tracing, backward contact tracing could reach large clusters of cases at a high probability in the presence of overdispersion. By combining backward contact tracing with traditional forward contact tracing, the number of cases averted by tracing and isolating the recently infected cases could typically be increased up to 2-3 fold. Backward contact tracing was originally used in only a few countries including South Korea and Japan as an early response to the COVID-19 local outbreak but the idea has been shared and discussed globally; some countries have introduced or been planning to introduce backward tracing as part of their outbreak response policy (Crozier et al., 2020; Oshitani, 2020; Queensland Health: Queensland Government, 2020; Scientific Pandemic Influenza Group on Modelling Operational sub-group, 2020). Paper 4, originally prepared as a working paper to inform the UK policy and later published as an academic paper, provided a useful context to the broad discussion on this contact tracing approach. Backward contact tracing is resource-intensive because it requires a strong involvement of public health officials. Therefore, it is expected to be most effective when the number of cases is relatively low and could be used as an exit strategy from stringent intervention measures such as lockdowns.

Lastly, Paper 5 focused on the school transmission dynamics of respiratory infectious diseases. The study first estimated the transmission patterns of seasonal influenza in primary schools over a heterogeneous network of classrooms. Transmission risks between a pair of students were characterised according to their relationship of classes/grades, modified by the class sizes and the number of classes per grade. The estimated within-school reproduction number and its breakdowns suggested that the estimated reproduction number was very similar (0.8-0.9) between classes and schools, irrespective of class sizes or the number of classes per grade. The relative contribution of within-class transmission was slightly smaller in grades with more classes and the contribution of inter-grade transmission increased to compensate for this reduction. These results indicated that, if this relationship between the transmission patterns between students and class

structures was conserved under school-based interventions which change the size and composition of classes (e.g. reduced class sizes and staggered attendance), the effectiveness of such interventions might be smaller than otherwise expected. Based on this scenario, possible school outbreaks of COVID-19 and pandemic influenza were simulated to assess multiple control strategies in school settings. Paper 5 proposed two directions for outbreak management strategies at schools: the 'passive approach', which continuously keep the within-school reproduction number within the subcritical level, and the 'responsive approach', which allows more relaxed prevention during the normal operation and enforces class closure upon detection of a case. For various ranges of parameter settings, the expected effectiveness of the management strategies was compared and discussed. In particular, screening by regular testing and class-level closure are proposed as potentially effective control measures. This paper was built upon a collection of findings and implications from the other studies in this thesis. The likelihood values corresponding to students' risk of infection from their households was included as part of the analysis, which was informed by the model and the estimation results of Paper 1. The idea of multiple overimputation and its implementation derived from Paper 2. The negative-binomial offspring distribution of SARS-CoV-2 as modelled in Papers 3 and 4 was incorporated into the simulation of COVID-19 school outbreaks as one of the additional scenario analyses.

9.2 Implications and limitations

9.2.1 Social roles and structures as drivers of heterogeneity

Heterogeneity in transmission patterns of infectious diseases analysed in this thesis was mostly related to social roles and structures. In households, familial roles such as sibling, father, mother and grandparent (and the household compositions as combinations of these roles) were suggested to be strong determinants of the risk of infection from outside the household and transmission potential within the household, which well explained the observed patterns of seasonal influenza occurrences in households. It should be noted that the familial roles are associated with multiple other factors including age, sex, and occupation and thus may not necessarily reflect the social role in the household alone. However, the performance of the household transmission model in Paper 1 suggested that such a variable as a summary of the overall epidemiological characteristics of individuals could be useful in fine-scale inference of infectious disease outbreaks. Appropriately selecting such variables that determine a major part of the transmission dynamics will likely require both field knowledge and thorough understanding of data generating processes. Statistical model selection processes should also guide the selection of variables to objectively assess the performance of the model.

Once the key variables are identified, the estimated transmission dynamics related to social roles and structures can bring useful implications on the epidemiological understanding and directions for control. Within school transmission dynamics estimated in Paper 5 was explained by class/grade structures, and the results led to a hypothesis that the transmission of seasonal influenza at schools might be related to close contacts between a limited number of students (i.e. between close friends) because the estimated reproduction number was almost constant regardless of the population density. Although this is just one possible interpretation of the results and needs to be empirically validated in the future study, it can be regarded as an example of a mathematical model accounting for social roles and structures that proposed an intriguing insight into underlying mechanisms of infectious disease dynamics.

9.2.2 Heterogeneity and outbreak extinction

An important effect of heterogeneity on the transmission dynamics of infectious diseases is that it often renders the outbreak prone to extinction. One of the basic mechanisms behind this is that the outbreak tends to be localised in the presence of heterogeneity. That is, if certain individuals are more susceptible to infection than others, those individuals are preferentially infected in the early phase of an outbreak and become immune, leaving only less susceptible individuals in the uninfected population (Britton et al., 2020). Heterogeneity in a form of network clustering (i.e. stronger within-group connections than between-group) as was observed in the within-school influenza outbreak also increases the chance of extinction by a similar mechanism. Because the virus is more likely to be transmitted within the group, the outbreak may become extinct before it spreads to other groups. Changing class structures and class distancing interventions analysed in Paper 5 are also motivated by this property.

Overdispersion in transmission, another form of heterogeneity modelled as a substantial variation in the offspring distribution, is also associated with the probability of extinction. Especially when the number of cases is small, the outbreak may well become extinct in the presence of overdispersion, even if the reproduction number is over one. Due to the high overdispersion in SARS-CoV-2 transmission, a number of countries did not see a large local outbreak in the earlier phase of the pandemic despite observing multiple introductions of cases from abroad. As a result, there were a wide spectrum of countries with different patterns in the number of imported and local cases, which allows for estimation of the overdispersion parameter k from the country-level case count data in Paper 3.

9.2.3 Heterogeneity and control measures

The optimal choice of control measures and their effectiveness may be affected by heterogeneity.

In Paper 4, the effectiveness of backward contact tracing was found to be strongly influenced by the degree of overdispersion. In the absence of overdispersion, the mean number of secondary transmissions is the same between the initially identified index case and the source case who infected the index case. Therefore, backward contact tracing is no more beneficial than randomly identifying one more case and the additional benefit may be marginal. On the other hand, backward contact tracing leads to large clusters at a disproportionate frequency if the transmission is overdispersed. Quantifying and modelling the heterogeneities in the transmission dynamics can thus result in a drastic shift in the direction of control strategies.

The importance of heterogeneity in the assessment of control measures was also highlighted in Paper 2. The degree of bias in the estimated VE in TND studies became larger when the study population was more diverse. As the VE is likely to be underestimated in TND studies when the test performance is imperfect; in such cases, vaccines might be erroneously considered as ineffective if the bias is not properly corrected.

The suggested responsive class closure approach in Paper 5 for school outbreak management utilises the heterogeneous transmission patterns between students according to class/grade structures. Since within-class transmission accounts for a large fraction of overall school reproduction number, an outbreak could be contained within a class if promptly detected and intervened. This approach may achieve management of school outbreaks without needing to enforce extensive school closure which can cause unnecessary loss of opportunity for education. Paper 5 suggests that the cost-benefit may be improved if class-level closure is combined with a reduction in interaction between students from different classes, which results in increased heterogeneity in transmission patterns within and between classes and grades.

9.2.4 Limitations

Aside from the specific limitations of each study detailed in the corresponding chapter as part of the paper, a number of limitations of this thesis should be acknowledged. Firstly, the datasets used in the studies in this thesis, i.e. Matsumoto influenza data and WHO situation report data, shared an important caveat of potential underreporting. Both datasets mainly consisted of symptomatic cases (while some asymptomatic infections may also have been reported) and the reporting may thus have been biased towards those with more severe symptoms. Cases were confirmed by lab tests (rapid tests or PCR tests), either of which has perfect sensitivity. Viral load of infected individuals varies over time and the sensitivity of tests may have depended on when cases are identified and tested. For these reasons, cases recorded in the datasets may have underrepresented the underlying infections. Although such potential underreporting was addressed by sensitivity

analyses where possible, these were based on specific assumptions on the mechanism of underreporting (e.g. constant probability of reporting). Secondly, although the modelled heterogeneous transmission dynamics in Paper 1 and 5 can be interpreted as weighted networks, the thesis did not put a focus on the properties used in the network theory (e.g. degree distribution, betweenness, etc.). This was partly because the dataset did not contain direct measures of social network such as friendship data and social roles; instead, social structures such as familial roles and class/grade memberships were used as a proxy of social networks. However, reanalysing the available datasets in the context of the network theory may provide additional insights in a future study. Thirdly, consideration of the associated cost of intervention measures considered in the thesis was mostly lightweight. The cost and effort required for backward contact tracing or most interventions at schools were not quantified, except for the cost of single-class closures in the 'responsive approach' in Paper 5 (quantified by the total number of days of class closure). When planning the actual implementation of these interventions, detrimental effects that interventions might cause should be assessed and compared with the expected benefit using established metrics (e.g. disability-adjusted life years).

9.3 Future work

In this section, I discuss research questions related to the studies in this thesis that remain to be answered, along with perspectives on the future work.

9.3.1 Household transmission patterns to inform public health (Paper 1)

In Paper 1, within-household transmission patterns of influenza were quantified. While I believe this study provided useful insights into the transmission dynamics of influenza in households, its primary focus was on inference and the study did not explore public health applications informed by the estimation results. A possible direction for public health applications is optimization of household-level vaccination strategies. Using the estimated within-household transmission patterns, optimal vaccine allocation strategies may be refined using mathematical models. The benefit to vaccinating household members includes protection not only for vaccinated members but also for unvaccinated individuals via the herd immunity effect. Therefore, when the number of vaccines available to each household is limited, the optimal allocation can be a nontrivial problem which requires to account for heterogeneity in transmission, external risk of infection expected severity and vaccine efficacy.

9.3.2 Bias correction for test-negative design with unknown test performance (Paper 2)

Bias correction methods proposed in Paper 2 assumed that the test performance of diagnostic tests

(sensitivity and specificity) is known. However, the true sensitivity and specificity of the tests in specific test-negative design studies are often uncertain, especially given that these values may vary between settings (Bruning et al., 2017). The proposed methods in Paper 2 cannot be applied as-is in such cases; however, by devising a study design that enables validation of test results, the bias correction methods in Paper 2 could be extended to allow for unknown test performance. Ideally, such extended methods should also be able to produce unbiased estimates in combination with existing estimation tools, which was one of the strengths of the methods presented in Paper 2.

9.3.3 Disentangling the source of overdispersion and tailored approach to assess the performance of contact tracing (Papers 3 and 4)

Overdispersion in transmission can arise from multiple factors (e.g. virological, biomechanical and sociobehavioural). In Papers 3 and 4, the specific sources of overdispersion were not discussed and the overall variation was only handled phenomenologically in a form of an offspring distribution. Although this simplification was inevitable due to limited availability of empirical data, the distributions of contacts, transmission and other factors should ideally be separately handled and discussed in the context of contact tracing. The offspring distribution estimated in Paper 3 is considered to reflect not only individual-level variation in the number of contacts but also variation in other factors such as viral load or duration of infectiousness. If additional data and analysis informs the relationship between these factors and, more importantly, the degree of variation in the number of contacts at an individual level, assessment of the effectiveness of contact tracing could be more realistic. In the practice of contact tracing, a tracer cannot distinguish contacts that result in transmission and that do not. Therefore, the required effort for contact tracing would be proportional to the number of contacts made by cases, not to the actual number of secondary transmissions. If only very small proportion of contacts actually lead to transmission, the contact tracing may suffer low cost effectiveness. With datasets informing the distribution of the number of contacts as well as secondary transmissions, more practical cost effectiveness assessment of forward and backward contact tracing would be possible, which will provide crucial information to guide the direction of public health policy.

9.3.4 Potential role of teachers/staff in school transmission dynamics (Paper 5)

As the Matsumoto influenza dataset did not contain information on teachers/staff, our analysis and simulation focused solely on students. Given the substantial number of students affected by influenza in the dataset, we believe that the relative contribution of teachers/staff to our estimates, if any, would have been minor and the estimated relationship between the within-school reproduction number and the size and number of classes were not substantially biased (unless

there were large outbreaks among teachers/staff, which were at least not reported in the original study). However, exclusion of teachers/staff from simulation may underestimate their potential role in mediating between-class transmission. Teachers/staff may visit multiple classes on a daily basis, and also can cause within-teachers/staff outbreak in their office. Especially given that many of the reported COVID-19 outbreaks at school involved teachers/staff (European Centre for Disease Prevention and Control, 2020), estimating their relative contribution to outbreaks and incorporating them in school outbreak simulation would be important future work. A minimum element required to allow for such an approach is estimates of the relative transmission risk between student-student, staff-student and staff-staff pairs. If these are estimated from school outbreak data of COVID-19, an additional group of "teachers/staff" could be added to the simulation model to provide more realistic assessment of outbreak risks.

9.4 Conclusion

Mathematical models are necessary simplifications of reality. As the real-world phenomena are essentially heterogeneous at infinitely fine scales in every aspect, mathematical modelling studies that try to capture the complexity of the reality better than the simplest models should almost inevitably account for heterogeneity. One of the most basic models of infectious disease dynamics, the Susceptible-Infectious-Recovered (SIR) model, is often introduced in a form of a homogeneous dynamical model; every individual in the model is assumed to have identical epidemiological properties and behaviour, although the SIR model was initially proposed as an age-dependent heterogeneous model. This may be partly because a homogeneous model is easy to analyse and explain, and can be fitted to data with minimal details (i.e. only a single time series of daily case counts). Moreover, homogeneous models are often sufficient to provide simple insights required in practice, especially when the overall population-level dynamics (such as the nationwide time-varying reproduction number) is of interest. Nonetheless, homogeneous models sometimes overlook important aspects of reality and can lead to inappropriate decisions, e.g. the role of superspreading cannot be analysed by an SIR model. By appropriately introducing essential heterogeneity into the model, a study may be able to find nontrivial insights that can improve our understanding and inform better control strategies. For example, the analysis of heterogeneous within-household transmission of seasonal influenza in this thesis emphasised the impact of schoolchildren on household outbreaks. Modelling individual-level heterogeneity in transmission suggested backward contact tracing can be an effective tool for control of SARS-CoV-2. Through these topics, this thesis has shown the importance of heterogeneity in understanding disease dynamics for seasonal and pandemic respiratory infections.

172

References

Addy, C. L., Longini, I. M., & Haber, M. (1991). A generalized stochastic model for the analysis of infectious disease final size data. *Biometrics*.

Asadi, S., Wexler, A. S., Cappa, C. D., Barreda, S., Bouvier, N. M., & Ristenpart, W. D. (2019). Aerosol emission and superemission during human speech increase with voice loudness. *Scientific Reports*, *9*(1), 2348. https://doi.org/10.1038/s41598-019-38808-z

Bailey, N. T. J. (1975). *The mathematical theory of infectious diseases and its applications. 2nd edition.* Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE.

Ball, F. (1986). A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. *Advances in Applied Probability*. https://doi.org/10.2307/1427301

Ball, F., Mollison, D., & Scalia-Tomba, G. (1997). Epidemics with two levels of mixing. In *The Annals of Applied Probability* (Vol. 7, Issue 1). https://doi.org/10.1515/9781400841356.436

Ball, F., & Neal, P. (2002). A general model for stochastic SIR epidemics with two levels of mixing. *Mathematical Biosciences*, *180*, 73–102. https://doi.org/10.1016/S0025-5564(02)00125-6

Banik, R. K., & Ulrich, A. (2020). Evidence of Short-Range Aerosol Transmission of SARS-CoV-2 and Call for Universal Airborne Precautions for Anesthesiologists During the COVID-19 Pandemic. *Anesthesia & Analgesia*, *131*(2), e102–e104. https://doi.org/10.1213/ANE.0000000000004933

Becker, N. G., & Britton, T. (1999). Statistical studies of infectious disease incidence. *Journal of the Royal Statistical Society B*, *61*(2), 287–307. https://doi.org/10.1111/1467-9868.00177

Begon, M., Bennett, M., Bowers, R. G., French, N. P., Hazel, S. M., & Turner, J. (2002). A clarification of transmission terms in host-microparasite models: Numbers, densities and areas. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268802007148

Bi, Q., Lessler, J., Eckerle, I., Lauer, S. A., Kaiser, L., Vuilleumier, N., Cummings, D. A. T., Flahault, A., Petrovic, D., Guessous, I., Stringhini, S., & Azman, A. S. (2021). Household Transmission of SARS-CoV-2: Insights from a Population-based Serological Survey. *MedRxiv*, 2020.11.04.20225573. https://doi.org/10.1101/2020.11.04.20225573

Bi, Q., Wu, Y., Mei, S., Ye, C., Zou, X., Zhang, Z., Liu, X., Wei, L., Truelove, S. A., Zhang, T., Gao, W., Cheng, C., Tang, X., Wu, X., Wu, Y., Sun, B., Huang, S., Sun, Y., Zhang, J., …

Feng, T. (2020). Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *The Lancet Infectious Diseases*, *20*(8), 911–919. https://doi.org/10.1016/S1473-3099(20)30287-5

Biggerstaff, M., Cauchemez, S., Reed, C., Gambhir, M., & Finelli, L. (2014). Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature. *BMC Infectious Diseases*. https://doi.org/10.1186/1471-2334-14-480

Blackwell, M., Honaker, J., & King, G. (2017). A Unified Approach to Measurement Error and Missing Data: Overview and Applications. *Sociological Methods and Research*. https://doi.org/10.1177/0049124115585360

Boni, M. F. (2008). Vaccination and antigenic drift in influenza. *Vaccine*, *26*, C8–C14. https://doi.org/10.1016/j.vaccine.2008.04.011

Brauner, J. M., Mindermann, S., Sharma, M., Johnston, D., Salvatier, J., Gavenčiak, T., Stephenson, A. B., Leech, G., Altman, G., Mikulik, V., Norman, A. J., Monrad, J. T., Besiroglu, T., Ge, H., Hartwick, M. A., Teh, Y. W., Chindelevitch, L., Gal, Y., & Kulveit, J. (2020). Inferring the effectiveness of government interventions against COVID-19. *Science*, eabd9338. https://doi.org/10.1126/science.abd9338

Britton, T., Ball, F., & Trapman, P. (2020). A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science*, *369*(6505), 846–849. https://doi.org/10.1126/science.abc6810

Bruning, A. H. L., Leeflang, M. M. G., Vos, J. M. B. W., Spijker, R., de Jong, M. D., Wolthers, K. C., & Pajkrt, D. (2017). Rapid Tests for Influenza, Respiratory Syncytial Virus, and Other Respiratory Viruses: A Systematic Review and Meta-analysis. *Clinical Infectious Diseases*. https://doi.org/10.1093/cid/cix461

Buchholz, U., Brockmann, S., Duwe, S., Schweiger, B., Heiden, M. an Der, Reinhardt, B., & Buda, S. (2010). Household transmissibility and other characteristics of seasonal oseltamivir-resistant influenza a(H1N1) viruses, Germany, 2007-8. *Eurosurveillance*.

Buonsenso, D., De Rose, C., Moroni, R., & Valentini, P. (2020). SARS-CoV-2 infections in Italian schools: preliminary findings after one month of school opening during the second wave of the pandemic. *MedRxiv*, 2020.10.10.20210328. https://doi.org/10.1101/2020.10.10.20210328

Carrat, F., Vergu, E., Ferguson, N. M., Lemaitre, M., Cauchemez, S., Leach, S., & Valleron, A. J. (2008). Time lines of infection and disease in human influenza: A review of volunteer challenge studies. In *American Journal of Epidemiology*.

https://doi.org/10.1093/aje/kwm375

Cauchemez, S., Bhattarai, A., Marchbanks, T. L., Fagan, R. P., Ostroff, S., Ferguson, N. M., Swerdlow, D., Sodha, S. V., Moll, M. E., Angulo, F. J., Palekar, R., Archer, W. R., & Finelli, L. (2011). Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proceedings of the National Academy of Sciences*, *108*(7), 2825–2830. https://doi.org/10.1073/pnas.1008895108

Cauchemez, S., Donnelly, C. A., Reed, C., Ghani, A. C., Fraser, C., Kent, C. K., Finelli, L., & Ferguson, N. M. (2009). Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. *The New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa0905498

Cauchemez, S., Ferguson, N. M., Fox, A., Mai, L. Q., Thanh, L. T., Thai, P. Q., Thoang, D. D., Duong, T. N., Minh Hoa, L. N., Tran Hien, N., & Horby, P. (2014). Determinants of Influenza Transmission in South East Asia: Insights from a Household Cohort Study in Vietnam. *PLoS Pathogens*, *10*(8), 2–9. https://doi.org/10.1371/journal.ppat.1004310

Cauchemez, S., Ferguson, N. M., Wachtel, C., Tegnell, A., Saour, G., Duncan, B., & Nicoll, A. (2009). Closure of schools during an influenza pandemic. In *The Lancet Infectious Diseases*. https://doi.org/10.1016/S1473-3099(09)70176-8

Centers for Disease Control and Prevention. (2020a). *CDC Seasonal Flu Vaccine Effectiveness Studies*. https://www.cdc.gov/flu/vaccines-work/effectiveness-studies.htm

Centers for Disease Control and Prevention. (2020b). *Demographic Trends of COVID-19*. https://www.cdc.gov/covid-data-tracker/index.html#demographics

Centers for Disease Control and Prevention. (2020c). *Operating schools during COVID-19: CDC's Considerations*. https://www.cdc.gov/coronavirus/2019-ncov/community/schools-childcare/schools.html

Chartrand, C., Leeflang, M. M. G., Minion, J., Brewer, T., & Pai, M. (2012). Accuracy of rapid influenza diagnostic tests: A meta-analysis. In *Annals of Internal Medicine*. https://doi.org/10.7326/0003-4819-156-7-201204030-00403

Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0012948

Clamer, V., Dorigatti, I., Fumanelli, L., Rizzo, C., & Pugliese, A. (2016). Estimating transmission probability in schools for the 2009 H1N1 influenza pandemic in Italy. *Theoretical Biology and Medical Modelling*. https://doi.org/10.1186/s12976-016-0045-2

Conlan, A. J. K., Eames, K. T. D., Gage, J. A., von Kirchbach, J. C., Ross, J. V., Saenz, R. A., & Gog, J. R. (2011). Measuring social networks in british primary schools through scientific engagement. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.2010.1807

Crozier, A., Mckee, M., & Rajan, S. (2020). Fixing England's COVID-19 response: learning from international experience. *Journal of the Royal Society of Medicine*, *113*(11), 422–427. https://doi.org/10.1177/0141076820965533

De Serres, G., Skowronski, D. M., Wu, X. W., & Ambrose, C. S. (2013). The test-negative design: Validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. *Eurosurveillance*. https://doi.org/10.2807/1560-7917.ES2013.18.37.20585

De Smedt, T., Merrall, E., Macina, D., Perez-Vilar, S., Andrews, N., & Bollaerts, K. (2018). Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0199180

Demiris, N., & O'Neill, P. D. (2005). Bayesian inference for stochastic multitype epidemics in structured populations via random graphs. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. https://doi.org/10.1111/j.1467-9868.2005.00524.x

Department for Education; UK Government. (2020). *Guidance for full opening: schools (Updated 5 November 2020)*. https://www.gov.uk/government/publications/actions-for-schools-during-the-coronavirus-outbreak/guidance-for-full-opening-schools

Eames, K. T. D. (2014). The influence of school holiday timing on epidemic impact. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268813002884

Eubank, S., Guclu, H., Kumar, V. S. A., Marathe, M. V., Srinivasan, A., Toroczkai, Z., & Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature*. https://doi.org/10.1038/nature02541

European Centre for Disease Prevention and Control. (2020). *COVID-19 in children and the role of school settings in COVID-19 transmission, 6 August 2020.*

Fateh-Moghadam, P., Battisti, L., Molinaro, S., Fontanari, S., Dallago, G., Binkin, N., & Zuccali, M. (2020). Contact tracing during Phase I of the COVID-19 pandemic in the Province of Trento, Italy: key findings and recommendations. *MedRxiv*, 2020.07.16.20127357. https://doi.org/10.1101/2020.07.16.20127357

Ferguson, N. M., Cummings, D. A. T., Fraser, C., Cajka, J. C., Cooley, P. C., & Burke, D. S.

(2006). Strategies for mitigating an influenza pandemic. *Nature*. https://doi.org/10.1038/nature04795

Fine, P. E. M. (1977). A COMMENTARY ON THE MECHANICAL ANALOGUE TO THE REED-FROST EPIDEMIC MODEL. *American Journal of Epidemiology*, *106*(2), 87–100. https://doi.org/10.1093/oxfordjournals.aje.a112449

Fournet, J., & Barrat, A. (2014). Contact patterns among high school students. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0107878

Frost, W. H. (1976). SOME CONCEPTIONS OF EPIDEMICS IN GENERAL. *American Journal of Epidemiology*, *103*(2), 141–151. https://doi.org/10.1093/oxfordjournals.aje.a112212

Fukushima, W., & Hirota, Y. (2017). Basic principles of test-negative design in evaluating influenza vaccine effectiveness. *Vaccine*, *35*(36), 4796–4800. https://doi.org/10.1016/j.vaccine.2017.07.003

Fumanelli, L., Ajelli, M., Manfredi, P., Vespignani, A., & Merler, S. (2012). Inferring the Structure of Social Contacts from Demographic Data in the Analysis of Infectious Diseases Spread. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1002673

Gemmetto, V., Barrat, A., & Cattuto, C. (2014). Mitigation of infectious disease at school: Targeted class closure vs school closure. *BMC Infectious Diseases*. https://doi.org/10.1186/s12879-014-0695-9

Glezen, W. P. (1996). Emerging infections: Pandemic influenza. In *Epidemiologic Reviews*. https://doi.org/10.1093/oxfordjournals.epirev.a017917

Goeyvaerts, N., Santermans, E., Potter, G., Torneri, A., Van Kerckhove, K., Willem, L., Aerts, M., Beutels, P., & Hens, N. (2017). Household Members Do Not Contact Each Other at Random: Implications for Infectious Disease Modelling. *BioRxiv*. http://biorxiv.org/content/early/2017/11/16/220202.abstract

Goldstein, E., Lipsitch, M., & Cevik, M. (2020). On the effect of age on the transmission of SARS-CoV-2 in households, schools and the community. *The Journal of Infectious Diseases*. https://doi.org/10.1093/infdis/jiaa691

Götzinger, F., Santiago-García, B., Noguera-Julián, A., Lanaspa, M., Lancella, L., Calò Carducci, F. I., Gabrovska, N., Velizarova, S., Prunk, P., Osterman, V., Krivec, U., Lo Vecchio, A., Shingadia, D., Soriano-Arandes, A., Melendo, S., Lanari, M., Pierantoni, L., Wagner, N., L'Huillier, A. G., … Riordan, A. (2020). COVID-19 in children and adolescents in Europe: a multinational, multicentre cohort study. *The Lancet Child & Adolescent Health*, *4*(9), 653–

661. https://doi.org/10.1016/S2352-4642(20)30177-2

Grantz, K., Metcalf, C. J. E., & Lessler, J. (2020). *Dispersion vs. Control*. https://hopkinsidd.github.io/nCoV-Sandbox/DispersionExploration.html

Greenland, S. (1996). Basic methods for sensitivity analysis of biases. In *International Journal of Epidemiology*. https://doi.org/10.1093/ije/25.6.1107

Guclu, H., Read, J., Vukotich, C. J., Galloway, D. D., Gao, H., Rainey, J. J., Uzicanin, A., Zimmer, S. M., & Cummings, D. A. T. (2016). Social contact networks and mixing among students in K-12 Schools in Pittsburgh, PA. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0151139

Haber, M., An, Q., Foppa, I. M., Shay, D. K., Ferdinands, J. M., & Orenstein, W. A. (2015). A probability model for evaluating the bias and precision of influenza vaccine effectiveness estimates from case-control studies. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268814002179

Hara, M, Takao, S., Fukuda, S., Shimazu, Y., & Miyazaki, K. (2004). [Comparison of three rapid diagnostic kits using immunochromatography for detection of influenza A virsuses]. *Kansenshogaku Zasshi*. https://doi.org/10.11150/kansenshogakuzasshi1970.78.935

Hara, Michimaru, Sadamasu, K., Takao, S., Shinkai, T., Kai, A., Fukuda, S., Shimazu, Y., Kuwayama, M., & Miyazaki, K. (2006). [Evaluation of immunochromatography test for rapid detection of influenza A and B viruses using real-time PCR]. *Kansenshogaku Zasshi. The Journal of the Japanese Association for Infectious Diseases*. https://doi.org/10.11150/kansenshogakuzasshi1970.80.522

He, X., Lau, E. H. Y., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X., Mo, X., Chen, Y., Liao, B., Chen, W., Hu, F., Zhang, Q., Zhong, M., Wu, Y., Zhao, L., … Leung, G. M. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*. https://doi.org/10.1038/s41591-020-0869-5

Heald-Sargent, T., Muller, W. J., Zheng, X., Rippe, J., Patel, A. B., & Kociolek, L. K. (2020). Age-Related Differences in Nasopharyngeal Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Levels in Patients With Mild to Moderate Coronavirus Disease 2019 (COVID-19). *JAMA Pediatrics*, *174*(9), 902. https://doi.org/10.1001/jamapediatrics.2020.3651

Hébert-Dufresne, L., Althouse, B. M., Scarpino, S. V., & Allard, A. (2020). Beyond R 0 : heterogeneity in secondary infections and probabilistic epidemic forecasting. *Journal of The Royal Society Interface*, *17*(172), 20200393. https://doi.org/10.1098/rsif.2020.0393

Hens, N., Goeyvaerts, N., Aerts, M., Shkedy, Z., Van Damme, P., & Beutels, P. (2009). Mining social mixing patterns for infectious disease models based on a two-day population survey in Belgium. *BMC Infectious Diseases*. https://doi.org/10.1186/1471-2334-9-5

Heymann, A. D., Hoch, I., Valinsky, L., Kokia, E., & Steinberg, D. M. (2009). School closure may be effective in reducing transmission of respiratory viruses in the community. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268809002556

House, T., Inglis, N., Ross, J. V., Wilson, F., Suleman, S., Edeghere, O., Smith, G., Olowokure, B., & Keeling, M. J. (2012). Estimation of outbreak severity and transmissibility: Influenza A(H1N1)pdm09 in households. *BMC Medicine*, *10*(1), 117. https://doi.org/10.1186/1741-7015-10-117

Ibuka, Y., Ohkusa, Y., Sugawara, T., Chapman, G. B., Yamin, D., Atkins, K. E., Taniguchi, K., Okabe, N., & Galvani, A. P. (2016). Social contacts, vaccination decisions and influenza in Japan. *Journal of Epidemiology and Community Health*. https://doi.org/10.1136/jech-2015-205777

Ip, D. K. M., Lau, L. L. H., Leung, N. H. L., Fang, V. J., Chan, K. H., Chu, D. K. W., Leung, G. M., Peiris, J. S. M., Uyeki, T. M., & Cowling, B. J. (2017). Viral Shedding and Transmission Potential of Asymptomatic and Paucisymptomatic Influenza Virus Infections in the Community. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*. https://doi.org/10.1093/cid/ciw841

Ismail, S. A., Saliba, V., Lopez Bernal, J. A., Ramsay, M. E., & Ladhani, S. N. (2020). SARS-CoV-2 infection and transmission in educational settings: cross-sectional analysis of clusters and outbreaks in England. *MedRxiv*, 2020.08.21.20178574. https://doi.org/10.1101/2020.08.21.20178574

Iuliano, A. D., Roguski, K. M., Chang, H. H., Muscatello, D. J., Palekar, R., Tempia, S., Cohen, C., Gran, J. M., Schanzer, D., Cowling, B. J., Wu, P., Kyncl, J., Ang, L. W., Park, M., Redlberger-Fritz, M., Yu, H., Espenhain, L., Krishnan, A., Emukule, G., … Mustaquim, D. (2018). Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *The Lancet*, *391*(10127), 1285–1300. https://doi.org/10.1016/S0140-6736(17)33293-2

Jackson, M. L., & Rothman, K. J. (2015). Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. *Vaccine*. https://doi.org/10.1016/j.vaccine.2015.01.069

Kass, R. E., Carlin, B. P., Gelman, A., Radford, &, Neal, M., Kass Panelists, R. E., & Neal, R. M.

(2012). *Markov Chain Monte Carlo in Practice: A Roundtable Discussion Statistical Practice Markov Chain Monte Carlo in Practice: A Roundtable Discussion.* https://doi.org/10.1080/00031305.1998.10480547

Keeling, M. J., & Eames, K. T. D. (2005). Networks and epidemic models. In *Journal of the Royal Society Interface*. https://doi.org/10.1098/rsif.2005.0051

Killingley, B., & Nguyen-Van-Tam, J. (2013). Routes of influenza transmission. *Influenza and Other Respiratory Viruses*, *7*, 42–51. https://doi.org/10.1111/irv.12080

Klick, B., Leung, G. M., & Cowling, B. J. (2012). Optimal design of studies of influenza transmission in households I: case-ascertained studies. *Epidemiology and Infection*, *140*(1), 106–114. https://doi.org/10.1017/S0950268811000392

Klompas, M., Baker, M. A., & Rhee, C. (2020). Airborne Transmission of SARS-CoV-2. *JAMA*, *324*(5), 441. https://doi.org/10.1001/jama.2020.12458

Kommuneoverlegen in Lillestrøm municipality, Sagdalen school in Lillestrøm municipality, & Nowegian Institute of Public Health. (2020). *[Outbreak report 26.08.2020: Outbreak of COVID-19, Sagdalen school, 2020] (Norwegian).* https://www.lillestrom.kommune.no/contentassets/1d174acbc73f49b8be7266f82300c19f/200826_utbruddsrapport_sagdalen2020.pdf

Kucharski, A. J., & Althaus, C. L. (2015). The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission. *Eurosurveillance*, *20*(25). https://doi.org/10.2807/1560-7917.ES2015.20.25.21167

Lau, L. L. H., Nishiura, H., Kelly, H., Ip, D. K. M., Leung, G. M., & Cowling, B. J. (2012). Household transmission of 2009 pandemic influenza A (H1N1): a systematic review and meta-analysis. *Epidemiology (Cambridge, Mass.)*. https://doi.org/10.1097/EDE.0b013e31825588b8

le Polain de Waroux, O., Flasche, S., Kucharski, A. J., Langendorf, C., Ndazima, D., Mwanga-Amumpaire, J., Grais, R. F., Cohuet, S., & Edmunds, W. J. (2018). Identifying human encounters that shape the transmission of Streptococcus pneumoniae and other acute respiratory infections. *Epidemics*. https://doi.org/10.1016/j.epidem.2018.05.008

Leclerc, Q. J., Fuller, N. M., Knight, L. E., Funk, S., & Knight, G. M. (2020). What settings have been linked to SARS-CoV-2 transmission clusters? *Wellcome Open Research*, *5*, 83. https://doi.org/10.12688/wellcomeopenres.15889.2

Lee, P.-I., Hu, Y.-L., Chen, P.-Y., Huang, Y.-C., & Hsueh, P.-R. (2020). Are children less

susceptible to COVID-19? *Journal of Microbiology, Immunology and Infection*, *53*(3), 371–372. https://doi.org/10.1016/j.jmii.2020.02.011

Leecaster, M., Toth, D. J. A., Pettey, W. B. P., Rainey, J. J., Gao, H., Uzicanin, A., & Samore, M. (2016). Estimates of Social Contact in a Middle School Based on Self-Report and Wireless Sensor Data. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0153690

Leung, N. H. L., Xu, C., Ip, D. K. M., & Cowling, B. J. (2015). Review Article: The Fraction of Influenza Virus Infections That Are Asymptomatic: A Systematic Review and Meta-analysis. *Epidemiology (Cambridge, Mass.)*. https://doi.org/10.1097/EDE.0000000000000340

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., & Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, *438*(7066), 355–359. https://doi.org/10.1038/nature04153

Longini, I. M., & Koopman, J. S. (1982). Household and community transmission parameters from final distributions of infections in households. *Biometrics*, *38*(1), 115–126.

Longini, I. M., Koopman, J. S., Haber, M., & Cotsonis, G. A. (1988). Statistical inference for infectious diseases: Risk-specific household and community transmission parameters. *American Journal of Epidemiology*. https://doi.org/10.1093/oxfordjournals.aje.a115038

Ludwig, D. (1975). Final size distribution for epidemics. *Mathematical Biosciences*, *23*(1–2), 33–46. https://doi.org/10.1016/0025-5564(75)90119-4

Madewell, Z. J., Yang, Y., Longini, I. M., Halloran, M. E., & Dean, N. E. (2020). Household Transmission of SARS-CoV-2. *JAMA Network Open*, *3*(12), e2031756. https://doi.org/10.1001/jamanetworkopen.2020.31756

Matsuzaki, Y., Katsushima, N., Nagai, Y., Shoji, M., Itagaki, T., Sakamoto, M., Kitaoka, S., Mizuta, K., & Nishimura, H. (2006). Clinical Features of Influenza C Virus Infection in Children. *The Journal of Infectious Diseases*, *193*(9), 1229–1235. https://doi.org/10.1086/502973

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., & More, S. J. (2020). Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ Open*, *10*(8), e039652. https://doi.org/10.1136/bmjopen-2020-039652

Melegaro, A., Fava, E. Del, Poletti, P., Merler, S., Nyamukapa, C., Williams, J., Gregson, S., & Manfredi, P. (2017). Social contact structures and time use patterns in the manicaland province of Zimbabwe. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0170459

Meyers, L. A., Newman, M. E. J., Martin, M., & Schrag, S. (2003). Applying network theory to epidemics: Control measures for Mycoplasma pneumoniae outbreaks. *Emerging Infectious Diseases*. https://doi.org/10.3201/eid0902.020188

Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M., & Brunham, R. C. (2005). Network theory and SARS: Predicting outbreak diversity. In *Journal of Theoretical Biology*. https://doi.org/10.1016/j.jtbi.2004.07.026

Ministry of Education Culture Sports Science and Techonogy Japan. (2020). *[Hygiene management manual regarding novel coronavirus disease: new life standards at schools (2020.9.3 Ver.4)] (Japanese)*. https://doi.org/https://doi.org/10.1186/s12976-016-0045-2

Minodier, L., Charrel, R. N., Ceccaldi, P.-E., van der Werf, S., Blanchon, T., Hanslik, T., & Falchi, A. (2015). Prevalence of gastrointestinal symptoms in patients with influenza, clinical significance, and pathophysiology of human influenza viruses in faecal samples: what do we know? *Virology Journal*, *12*(1), 215. https://doi.org/10.1186/s12985-015-0448-4

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., & Edmunds, W. J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.0050074

Newman, M. (2002). The spread of epidemic disease on networks. *Journal Reference: Phys. Rev. E*.

Nishiura, H., Ejima, K., Mizumoto, K., Nakaoka, S., Inaba, H., Imoto, S., Yamaguchi, R., & Saito, M. M. (2014). Cost-effective length and timing of school closure during an influenza pandemic depend on the severity. *Theoretical Biology and Medical Modelling*. https://doi.org/10.1186/1742-4682-11-5

Nishiura, H., Oshitani, H., Kobayashi, T., Saito, T., Sunagawa, T., Matsui, T., Wakita, T., & Suzuki, M. (2020). Closed environments facilitate secondary transmission of coronavirus disease 2019 (COVID-19). *MedRxiv*, 2020.02.28.20029272. https://doi.org/10.1101/2020.02.28.20029272

O'Neill, P. D., Balding, D. J., Becker, N. G., Eerola, M., & Mollison, D. (2000). Analyses of infectious disease data from household outbreaks by Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society Series C-Applied Statistics*, *49*(4), 517–542. https://doi.org/10.1111/1467-9876.00210

O'Neill, P. D., & Roberts, G. O. (1999). Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *162*(1),

Meyers, L. A., Newman, M. E. J., Martin, M., & Schrag, S. (2003). Applying network theory to epidemics: Control measures for Mycoplasma pneumoniae outbreaks. *Emerging Infectious Diseases*. https://doi.org/10.3201/eid0902.020188

Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M., & Brunham, R. C. (2005). Network theory and SARS: Predicting outbreak diversity. In *Journal of Theoretical Biology*. https://doi.org/10.1016/j.jtbi.2004.07.026

Ministry of Education Culture Sports Science and Techonogy Japan. (2020). *[Hygiene management manual regarding novel coronavirus disease: new life standards at schools (2020.9.3 Ver.4)] (Japanese)*. https://doi.org/https://doi.org/10.1186/s12976-016-0045-2

Minodier, L., Charrel, R. N., Ceccaldi, P.-E., van der Werf, S., Blanchon, T., Hanslik, T., & Falchi, A. (2015). Prevalence of gastrointestinal symptoms in patients with influenza, clinical significance, and pathophysiology of human influenza viruses in faecal samples: what do we know? *Virology Journal*, *12*(1), 215. https://doi.org/10.1186/s12985-015-0448-4

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., & Edmunds, W. J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.0050074

Newman, M. (2002). The spread of epidemic disease on networks. *Journal Reference: Phys. Rev. E*.

Nishiura, H., Ejima, K., Mizumoto, K., Nakaoka, S., Inaba, H., Imoto, S., Yamaguchi, R., & Saito, M. M. (2014). Cost-effective length and timing of school closure during an influenza pandemic depend on the severity. *Theoretical Biology and Medical Modelling*. https://doi.org/10.1186/1742-4682-11-5

Nishiura, H., Oshitani, H., Kobayashi, T., Saito, T., Sunagawa, T., Matsui, T., Wakita, T., & Suzuki, M. (2020). Closed environments facilitate secondary transmission of coronavirus disease 2019 (COVID-19). *MedRxiv*, 2020.02.28.20029272. https://doi.org/10.1101/2020.02.28.20029272

O'Neill, P. D., Balding, D. J., Becker, N. G., Eerola, M., & Mollison, D. (2000). Analyses of infectious disease data from household outbreaks by Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society Series C-Applied Statistics*, *49*(4), 517–542. https://doi.org/10.1111/1467-9876.00210

O'Neill, P. D., & Roberts, G. O. (1999). Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *162*(1),

121–129. https://doi.org/10.1111/1467-985X.00125

Orenstein, E. W., De Serres, G., Haber, M. J., Shay, D. K., Bridges, C. B., Gargiullo, P., & Orenstein, W. A. (2007). Methodologic issues regarding the use of three observational study designs to assess influenza vaccine effectiveness. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dym021

Oshitani, H. (2020). Cluster-Based Approach to Coronavirus Disease 2019 (COVID-19) Response in Japan, from February to April 2020. *Japanese Journal of Infectious Diseases*, *73*(6), 491–493. https://doi.org/10.7883/yoken.JJID.2020.363

Osterhaus, A. D. (2000). Influenza B Virus in Seals. *Science*, *288*(5468), 1051–1053. https://doi.org/10.1126/science.288.5468.1051

Park, Y. J., Choe, Y. J., Park, O., Park, S. Y., Kim, Y.-M., Kim, J., Kweon, S., Woo, Y., Gwack, J., Kim, S. S., Lee, J., Hyun, J., Ryu, B., Jang, Y. S., Kim, H., Shin, S. H., Yi, S., Lee, S., Kim, H. K., … Jeong, E. K. (2020). Contact Tracing during Coronavirus Disease Outbreak, South Korea, 2020. *Emerging Infectious Diseases*, *26*(10), 2465–2468. https://doi.org/10.3201/eid2610.201315

Paules, C., & Subbarao, K. (2017). Influenza. In *The Lancet*. https://doi.org/10.1016/S0140-6736(17)30129-0

Petersen, E., Koopmans, M., Go, U., Hamer, D. H., Petrosillo, N., Castelli, F., Storgaard, M., Al Khalili, S., & Simonsen, L. (2020). Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet Infectious Diseases*, *20*(9), e238–e244. https://doi.org/10.1016/S1473-3099(20)30484-9

Punpanich, W., & Chotpitayasunondh, T. (2012). A review on the clinical spectrum and natural history of human influenza. *International Journal of Infectious Diseases*, *16*(10), e714–e723. https://doi.org/10.1016/j.ijid.2012.05.1025

Queensland Health: Queensland Government. (2020). *Contact tracing for coronavirus (COVID-19): how it's done in Queensland*. https://www.health.qld.gov.au/news-events/news/contact-tracing-novel-coronavirus-covid-19-confirmed-case-notify-Queensland

Rahman, H. S., Aziz, M. S., Hussein, R. H., Othman, H. H., Salih Omer, S. H., Khalid, E. S., Abdulrahman, N. A., Amin, K., & Abdullah, R. (2020). The transmission modes and sources of COVID-19: A systematic review. *International Journal of Surgery Open*, *26*, 125–136. https://doi.org/10.1016/j.ijso.2020.08.017

Riediker, M., & Tsai, D.-H. (2020). Estimation of Viral Aerosol Emissions From Simulated

Individuals With Asymptomatic to Moderate Coronavirus Disease 2019. *JAMA Network Open*, *3*(7), e2013807. https://doi.org/10.1001/jamanetworkopen.2020.13807

Russell, F., Ryan, K. E., Snow, K., Danchin, M., Mulholland, K., & Goldfeld, S. (2020). *COVID-19 in Victorian Schools: An analysis of child-care and school outbreak data and evidence-based recommendations for opening schools and keeping them open*. https://www.dhhs.vic.gov.au/sites/default/files/documents/202009/Report-summary-COVID-19-in-victorian-schools-pdf.pdf

Salez, N., Mélade, J., Pascalis, H., Aherfi, S., Dellagi, K., Charrel, R. N., Carrat, F., & de Lamballerie, X. (2014). Influenza C virus high seroprevalence rates observed in 3 different population groups. *Journal of Infection*, *69*(2), 182–189. https://doi.org/10.1016/j.jinf.2014.03.016

Scientific Pandemic Influenza Group on Modelling Operational sub-group. (2020). *SPI-M-O: Consensus Statement on COVID-19, 3 June 2020*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/897526/S0471_SAGE_40_200603_SPI-M-O_Consensus_Statement.pdf

Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J. F., Quaggiotto, M., van den Broeck, W., Régis, C., Lina, B., & Vanhems, P. (2011). High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0023176

Stein-Zamir, C., Abramson, N., Shoob, H., Libal, E., Bitan, M., Cardash, T., Cayam, R., & Miskin, I. (2020). A large COVID-19 outbreak in a high school 10 days after schools' reopening, Israel, May 2020. *Eurosurveillance*, *25*(29). https://doi.org/10.2807/1560-7917.ES.2020.25.29.2001352

Su, S., Fu, X., Li, G., Kerlin, F., & Veit, M. (2017). Novel Influenza D virus: Epidemiology, pathology, evolution and biological characteristics. *Virulence*, *8*(8), 1580–1591. https://doi.org/10.1080/21505594.2017.1365216

Thai, P. Q., Mai, L. Q., Welkers, M. R. A., Hang, N. L. K., Thanh, L. T., Dung, V. T. V., Yen, N. T. T., Duong, T. N., Hoa, L. N. M., Thoang, D. D., Trang, H. T. H., de Jong, M. D., Wertheim, H., Hien, N. T., Horby, P., & Fox, A. (2014). Pandemic H1N1 virus transmission and shedding dynamics in index case households of a prospective Vietnamese cohort. *Journal of Infection*. https://doi.org/10.1016/j.jinf.2014.01.008

Torres, J. P., Piñera, C., De La Maza, V., Lagomarcino, A. J., Simian, D., Torres, B., Urquidi, C., Valenzuela, M. T., & O'Ryan, M. (2020). Severe Acute Respiratory Syndrome Coronavirus

2 Antibody Prevalence in Blood in a Large School Community Subject to a Coronavirus Disease 2019 Outbreak: A Cross-sectional Study. *Clinical Infectious Diseases*. https://doi.org/10.1093/cid/ciaa955

Tsang, T. K., Cauchemez, S., Perera, R. A. P. M., Freeman, G., Fang, V. J., Ip, D. K. M., Leung, G. M., Malik Peiris, J. S., & Cowling, B. J. (2014). Association between antibody titers and protection against influenza virus infection within households. *The Journal of Infectious Diseases*. https://doi.org/10.1093/infdis/jiu186

Tsang, T. K., Lau, L. L. H., Cauchemez, S., & Cowling, B. J. (2016). Household Transmission of Influenza Virus. *Trends in Microbiology*, *24*(2), 123–133. https://doi.org/10.1016/j.tim.2015.10.012

Uchida, M., Kaneko, M., Hidaka, Y., Yamamoto, H., & Honda, T. (2017). *Prospective Epidemiological Evaluation of Seasonal Influenza in All Elementary Schoolchildren in Matsumoto City , Japan , in 2014 / 2015*. 333–339. https://doi.org/10.7883/yoken.JJID.2016.037

Uchida, M., Kaneko, M., Hidaka, Y., Yamamoto, H., Honda, T., Takeuchi, S., Saito, M., & Kawa, S. (2017a). Effectiveness of vaccination and wearing masks on seasonal influenza in Matsumoto City, Japan, in the 2014/2015 season: An observational study among all elementary schoolchildren. *Preventive Medicine Reports*, *5*, 86–91. https://doi.org/10.1016/j.pmedr.2016.12.002

Uchida, M., Kaneko, M., Hidaka, Y., Yamamoto, H., Honda, T., Takeuchi, S., Saito, M., & Kawa, S. (2017b). Prospective epidemiological evaluation of seasonal influenza in all elementary schoolchildren in Matsumoto city, Japan, in 2014/2015. *Japanese Journal of Infectious Diseases*. https://doi.org/10.7883/yoken.JJID.2016.037

UK Office for National Statistics. (2020). *Deaths due to coronavirus (COVID-19) compared with deaths from influenza and pneumonia, England and Wales Deaths due to coronavirus (COVID-19) compared with deaths from influenza and pneumonia, England and Wales: deaths occurring between 1 January and 31*. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsduetocoronaviruscovid19comparedwithdeathsfrominfluenzaandpneumonia englandandwales/deathsoccurringbetween1januaryand31august2020

United Nations Educational Scientific and Cultural Organization. (2020). *COVID-19 Impact on Education: Global monitoring of school closures caused by COVID-19 (interactive map)*. https://en.unesco.org/covid19/educationresponse

Van Boven, M., Koopmans, M., Van Beest Holle, M. D. R., Meijer, A., Klinkenberg, D., Donnelly, C. A., & Heesterbeek, H. (2007). Detecting emerging transmissibility of avian influenza virus in human households. *PLoS Computational Biology*, *3*(7), 1394–1402. https://doi.org/10.1371/journal.pcbi.0030145

Van Kerkhove, M. D., Vandemaele, K. A. H., Shinde, V., Jaramillo-Gutierrez, G., Koukounari, A., Donnelly, C. A., Carlino, L. O., Owen, R., Paterson, B., Pelletier, L., Vachon, J., Gonzalez, C., Hongjie, Y., Zijian, F., Chuang, S. K., Au, A., Buda, S., Krause, G., Haas, W., … Mounts, A. W. (2011). Risk Factors for Severe Outcomes following 2009 Influenza A (H1N1) Infection: A Global Pooled Analysis. *PLoS Medicine*, *8*(7), e1001053. https://doi.org/10.1371/journal.pmed.1001053

Vetter, P., Vu, D. L., L'Huillier, A. G., Schibler, M., Kaiser, L., & Jacquerioz, F. (2020). Clinical features of covid-19. *BMJ*, m1470. https://doi.org/10.1136/bmj.m1470

Viner, R. M., Mytton, O. T., Bonell, C., Melendez-Torres, G. J., Ward, J., Hudson, L., Waddington, C., Thomas, J., Russell, S., van der Klis, F., Koirala, A., Ladhani, S., Panovska-Griffiths, J., Davies, N. G., Booy, R., & Eggo, R. M. (2020). Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults. *JAMA Pediatrics*. https://doi.org/10.1001/jamapediatrics.2020.4573

Volz, E., & Meyers, L. A. (2009). Epidemic thresholds in dynamic contact networks. *Journal of the Royal Society Interface*. https://doi.org/10.1098/rsif.2008.0218

Wang, L., Chu, C., Yang, G., Hao, R., Li, Z., Cao, Z., Qiu, S., Li, P., Wu, Z., Yuan, Z., Xu, Y., Zeng, D., Wang, Y., & Song, H. (2014). Transmission Characteristics of Different Students during a School Outbreak of (H1N1) pdm09 Influenza in China, 2009. *Scientific Reports*. https://doi.org/10.1038/srep05982

Wardell, R., Prem, K., Cowling, B. J., & Cook, A. R. (2017). The role of symptomatic presentation in influenza A transmission risk. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268816002740

Watanabe, S. (2013). *A Widely Applicable Bayesian Information Criterion*. *14*, 867–897. https://doi.org/10.1088/0953-8984/23/18/184115

Watanabe, T., Watanabe, S., Maher, E. A., Neumann, G., & Kawaoka, Y. (2014). Pandemic potential of avian influenza A (H7N9) viruses. *Trends in Microbiology*, *22*(11), 623–631. https://doi.org/10.1016/j.tim.2014.08.008

Waxman, D., & Nouvellet, P. (2019). Sub- or supercritical transmissibilities in a finite disease outbreak: Symmetry in outbreak properties of a disease conditioned on extinction. *Journal*

of Theoretical Biology, *467*, 80–86. https://doi.org/10.1016/j.jtbi.2019.01.033

Welge-Lüssen, A., & Wolfensberger, M. (2006). Olfactory Disorders following Upper Respiratory Tract Infections. In *Taste and Smell* (pp. 125–132). KARGER. https://doi.org/10.1159/000093758

World Health Organization. (2012). Vaccines against influenza WHO position paper — November 2012. In *Weekly Epidemiological Record = Relevé épidémiologique hebdomadaire* (Vol. 87, Issue 47).

World Health Organization. (2019). *Global influenza strategy 2019-2030*. World Health Organization.

World Health Organization. (2020a). *Coronavirus disease 2019 (COVID-19) Situation Report – 38*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200227-sitrep-38-covid-19.pdf

World Health Organization. (2020b). *Novel Coronavirus(2019-nCoV) Situation Report 11*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200131-sitrep-11-ncov.pdf

World Health Organization. (2020c). *Novel Coronavirus (2019-nCoV) Situation Report 1*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf

World Health Organization. (2020d). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Wu, Z., & McGoogan, J. M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China. *JAMA*, *323*(13), 1239. https://doi.org/10.1001/jama.2020.2648

Yamazaki, M., Mitamura, K., Ichikawa, M., Kimura, K., Komiyama, O., Shimizu, H., Kawakami, C., Watanabe, S., Imai, M., Cho, H., & Takeuchi, Y. (2004). [Evaluation of flow-through immunoassay for rapid detection of influenza A and B viruses]. *Kansenshogaku Zasshi. The Journal of the Japanese Association for Infectious Diseases*. https://doi.org/10.11150/kansenshogakuzasshi1970.78.865

Yonker, L. M., Neilan, A. M., Bartsch, Y., Patel, A. B., Regan, J., Arya, P., Gootkind, E., Park, G., Hardcastle, M., St. John, A., Appleman, L., Chiu, M. L., Fialkowski, A., De la Flor, D., Lima,

R., Bordt, E. A., Yockey, L. J., D'Avino, P., Fischinger, S., … Fasano, A. (2020). Pediatric Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): Clinical Presentation, Infectivity, and Immune Responses. *The Journal of Pediatrics*, *227*, 45-52.e5. https://doi.org/10.1016/j.jpeds.2020.08.037