Structural differences in mixing behavior informing the role of asymptomatic infection and testing symptom heritability

DOI: 10.1016/j.mbs.2016.12.004
Handle: http://hdl.handle.net/1942/24130

Structural differences in mixing behaviour informing the role of asymptomatic infection and testing symptom heritability

Eva Santermans[a], Kim Van Kerckhove[a], Amin Azmon[b], W. John Edmunds[c], Philippe Beutels[d], Christel Faes[a], Niel Hens[a,d]

[a] *Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Belgium*
[b] *Novartis Pharma AG, Oncology Business Unit / General Medical Affairs, Novartis Campus, Basel, Switzerland*
[c] *Centre for the Mathematical Modelling of Infectious Diseases, Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom*
[d] *Centre for Health Economics Research and Modelling Infectious Diseases, Vaccine & Infectious Disease Institute, University of Antwerp, Belgium*

Abstract

Most infectious disease data is obtained from disease surveillance which is based on observations of symptomatic cases only. However, many infectious diseases are transmitted before the onset of symptoms or without developing symptoms at all throughout the entire disease course, referred to as asymptomatic transmission. Fraser and colleagues [1] showed that this type of transmission plays a key role in assessing the feasibility of intervention measures in controlling an epidemic outbreak. To account for asymptomatic transmission in epidemic models, methods often rely on assumptions that cannot be verified given the data at hand.

The present study aims at assessing the contribution of social contact data from asymptomatic and symptomatic individuals in quantifying the contribution of (a)symptomatic infections. We use a mathematical model based on ordinary differential equations (ODE) and a likelihood-based approach followed by Markov Chain Monte Carlo (MCMC) to estimate the model parameters and

*Corresponding author
Email address: eva.santermans@uhasselt.be (Eva Santermans)

their uncertainty.

Incidence data on influenza-like illness in the initial phase of the 2009 A/H1N1pdm epidemic is used to illustrate that it is possible to estimate either the proportion of asymptomatic infections or the relative infectiousness of symptomatic versus asymptomatic infectives. Further, we introduce a model in which the chance of developing symptoms depends on the disease state of the person that transmitted the infection.

In conclusion, incorporating social contact data from both asymptomatic and symptomatic individuals allows inferring on parameters associated with asymptomatic infection based on disease data from symptomatic cases only.

Keywords: mathematical model, influenza, asymptomatic transmission, social contact data, symptom heritability

---

1. Introduction

In the absence of effective vaccines or treatment, controlling the spread of an infectious disease during the early stages of an outbreak, relies on (i) isolation of symptomatic cases and (ii) tracing and quarantining the contacts of these cases. Hence, the timing of onset of symptoms relative to the start of infectiousness is a crucial factor in the success of these public health interventions. It has been shown that the proportion of asymptomatic infections (i.e. transmission that occurs before symptom onset or without showing symptoms at all) is a key parameter to predict whether or not isolation and contact tracing will lead to containment [1]. It is therefore important to use an epidemic model that explicitly takes into account asymptomatic transmission. However, in many cases the available data is based on observations of symptomatic individuals only. To overcome this limitation, models often rely on untestable assumptions, e.g. assuming a fixed proportion of asymptomatic individuals [2] or ignoring pre-symptomatic transmission [3].

Data on social contacts of individuals in a population have already proven to be a valuable additional source of information when estimating the Who

Acquires Infection From Whom (WAIFW) matrix and the basic reproduction number R_0 (see e.g. [4] [5]). More recently, social contact data have also been used to gain insight in the impact of illness on social contact patterns [6]. It was found that individuals symptomatic with influenza-like illness (ILI) have less social contacts than asymptomatic individuals. Furthermore, the age distribution of contacts differs between symptomatic and asymptomatic cases. These differences in mixing behavior affect the expected distribution of infection during the early stages of an outbreak, which allowed Van Kerckhove and colleagues [7] to estimate the proportion of ILI infections caused by asymptomatic cases (34%; CI: 0% - 77%) from ILI incidence data.

Influenza viruses are highly infectious and cases can show a variety of symptoms such as fever, runny nose and sore throat. A substantial number of cases also show little to no apparent symptoms. Several challenge studies have looked at the dynamics of viral shedding and symptoms following influenza virus infections; for a review see [8]. Symptomatic cases are considered to be more infectious than asymptomatic cases, since it was found that clinical cases have a higher quantity of virus in nasal wash fluids compared to individuals who did not develop symptoms. In addition, a positive correlation was found between severity of illness and the mean quantity of virus. The link between administered dose and development or degree of symptoms is less clear. Carrat and colleagues [8] reported a negative correlation between inoculated dose and fever, whereas Huang et al. [9] did not find a dependency between inoculated dose and disease outcome. Their findings point to host factors leading to asymptomatic infections. Hence, it is clear that more research is needed to find the precise link between the amount and duration of viral shedding, the development and the degree of symptoms and the transmission of the virus.

In the current study we will extend the work of Van Kerckhove et al. [7] by incorporating social contact data from asymptomatic and symptomatic individuals to inform mixing patterns in a compartmental model described by a system of ordinary differential equations. We will illustrate inference on parameters related to asymptomatic infection using incidence data on influenza-like illness.

Furthermore, we will also investigate the possibility that the chance of developing ILI symptoms depends on whether infection came from a symptomatic or an asymptomatic case. The paper is organized as follows. In Section 2, we introduce the model structure, data and estimation procedure. In Section 3, the ILI data are analyzed, and, lastly, Section 4 summarizes our main results, conclusions, and avenues for further research.

2. Material and Methods

2.1. Data

2.1.1. ILI data

Weekly incidence data were obtained from general practitioners' weekly consultation data on influenza-like-illness (ILI) from England and Wales during the early part of the A/H1N1pdm influenza epidemic in 2009 (weeks 23-29) [10]. Pre-existing immunity to the pandemic strain was obtained from a serological study in England the year before the pandemic [11].

2.1.2. Social contact data

We use data from a social contact survey that was carried out during the A/H1N1pdm influenza epidemic in England. This survey is described in detail in [6]. Briefly, participants were recruited into the study through packs with antiviral medication distributed at thirty-one antiviral distribution centers throughout England during the epidemic. The packs contained a social contact diary to be filled in on one day during the time they were symptomatic with ILI. Two weeks later (by which time participants were expected to have recovered), participants were sent a similar, follow-up questionnaire. Thus, the study aimed to obtain two contact diaries from each participant: one completed when the participant was showing symptoms and one completed after he or she had recovered. In these contact diaries participants were asked to record details about each person they met during the course of a day: gender and (estimated) age of the contact, social setting and duration of the encounter, frequency with

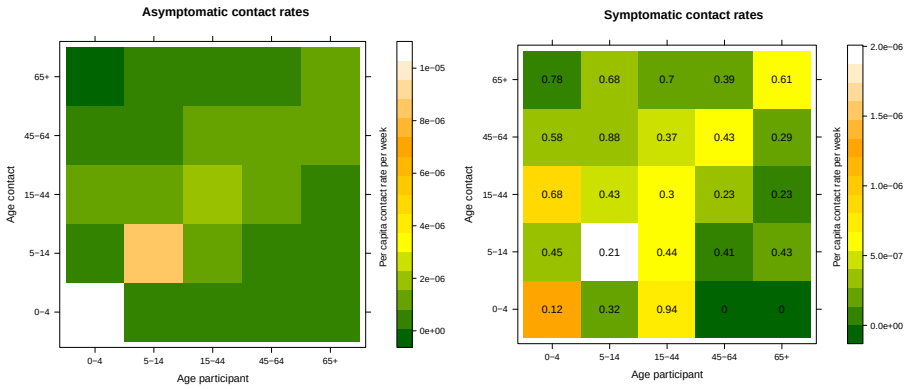

Figure 1: Age-specific contact rates for asymptomatic individuals (left) and symptomatic individuals (right) based on the age classes of the incidence data. The right plots displays the percentage reduction in contact rates between symptomatic versus asymptomatic individuals.

which that person was met, and whether the encounter involved any skin-to-skin contact (e.g., hand-shake, kiss, or contact sport). A total of 140 participants returned two completed contact diaries. Based on this information social contact matrices C^a and C^s for both recovered (assumed to be the same as asymptomatic) and symptomatic individuals were calculated, respectively [7]. These matrices are presented in Figure 1.

2.2. Transmission models

2.2.1. Non-preferential model

We use a compartmental model which describes the disease dynamics for influenza and infections with similar disease progress. In this model, individuals either develop symptoms or not after a pre-symptomatic stage. We will refer to this model as the non-preferential transmission model, since the development of symptoms is independent of the status of the infector. It is depicted as a flow diagram in Figure 2. Note that superscripts indicate clinical status of the infected individual: symptomatic 's' or asymptomatic 'a'.

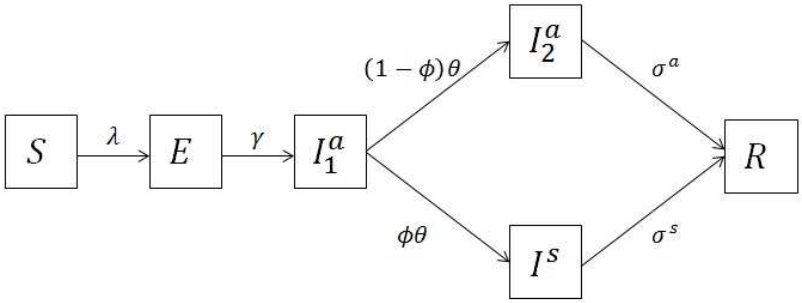

Figure 2: Schematic diagram of the non-preferential transmission model. Superscripts indicate presence (s) or absence (a) of symptoms.

Hence, we assume that susceptible individuals are infected at rate $\lambda(t)$. Following infection, individuals enter the exposed compartment (E) in which they are infected but not yet infectious. After a mean latent period $1/\gamma$ individuals become asymptomatic infectious, entering the compartment I_1^a. We define ϕ; $0 \leq \phi \leq 1$ to be the proportion of cases that will develop symptoms, and $1-\phi$ the proportion of cases that will remain asymptomatic. Infectious individuals move from the asymptomatic compartment I_1^a to the symptomatic I^s or asymptomatic I_2^a compartments at rates $\phi \times \theta$ and $(1-\phi) \times \theta$, respectively. Finally, individuals recover and move to the recovered compartment (R) at rates σ^a and σ^s, respectively. The corresponding system of ordinary differential equations (ODEs) is available in Appendix A.

2.2.2. Preferential model

Further, we extend this model by keeping track of whether a susceptible individual is infected by an asymptomatic or a symptomatic case; infection caused by an asymptomatic case occurs at rate $\lambda_a(t)$ and by a symptomatic case at rate, $\lambda_s(t)$, respectively. If the infector is asymptomatic, the infected individual will move from S to E_a; if the infector is symptomatic, the infected individual will move to E_s. Next, cases become asymptomatic infectious at rate γ and

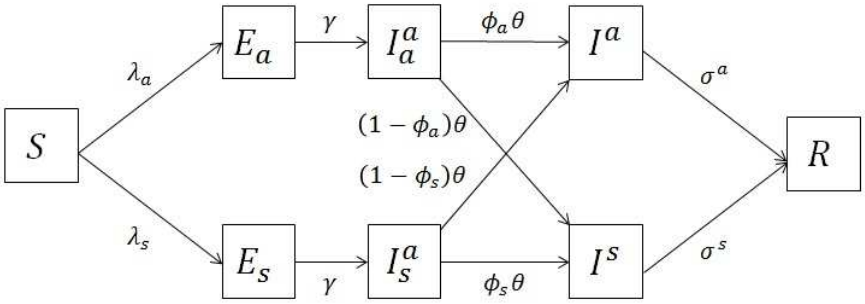

Figure 3: Schematic diagram of the preferential transmission model. Superscripts indicate clinical status of the infected individual: symptomatic (s) or asymptomatic (a). Subscripts indicate whether the infector was symptomatic (s) or asymptomatic (a).

move to I_a^a or I_s^a. Infected individuals then either develop symptoms or remain asymptomatic. We define ϕ_a as the probability that an individual infected by an asymptomatic case remains asymptomatic and ϕ_s as the probability that an individual infected by a symptomatic case develops symptoms. Figure 3 shows a schematic diagram of this model which we will call the preferential transmission model. The corresponding system of ODEs is available in Appendix B.

Note that we assume the length of the incubation period to be independent of the infector-type, since there is no information in the literature on possible differences in incubation and latent period between individuals infected by symptomatic and asymptomatic cases. Under this assumption, the preferential model simplifies to the non-preferential model if $\phi_s = 1 - \phi_a$.

2.3. Age structure and social contacts

Consider a population that is divided in K age categories. The rate at which a susceptible person in age group k acquires infection at time t is defined as $\lambda(k,t)$, the age-specific force of infection. Further, let $\beta(k,k')$ denote the time-independent average per capita rate at which an infectious individual in age group k' makes effective contact with a susceptible person in age group k,

per unit time. The force of infection is then given by (see e.g. [12])

$$\lambda(k,t) = \sum_{k'=0}^{K} \beta(k,k') I(k',t),$$

where $I(k',t)$ denotes the total number of infectious individuals in age group k' at time t. We follow the approach by Wallinga and colleagues [4] who express $\beta(k,k')$ as

$$\beta(k,k') = q \cdot c(k,k'), \tag{1}$$

where $c(k,k')$ is the per capita rate at which an individual in age group k makes contact with a person in age group k', per unit of time, and q a constant proportionality factor that may capture, among other effects, susceptibility and infectivity. The elements $c(k,k')$ form a $K \times K$ matrix C, which is called a social contact matrix. Equation 1 is referred to as the 'social contact hypothesis'. The social contact matrix describes how individuals of different age groups mix in a population.

In this paper, we distinguish between the asymptomatic social contact matrix C^a and the symptomatic social contact matrix C^s. Hence, $c^a(k,k')$ is the per capita rate at which an asymptomatic individual in age group k' makes contact with a person in age group k. We allow different proportionality factors for asymptomatic and symptomatic individuals and denote them by q^a and q^s, respectively. Hence,

$$\beta^a(k,k') = q^a \cdot c^a(k,k'),$$

$$\beta^s(k,k') = q^s \cdot c^s(k,k'),$$

with β^a and β^s the transmission rates of asymptomatic and symptomatic cases, respectively. Lastly, define $q^r = \frac{q^s}{q^a}$ as the relative infectiousness of symptomatic cases versus asymptomatic cases.

Then, the force of infection for the non-preferential transmission model is defined as

$$\lambda_{K\times 1}(t) = \beta^a_{K\times K} \times (I^a_{1,K\times 1}(t) + I^a_{2,K\times 1}(t)) + \beta^s_{K\times K} \times I^s_{K\times 1}(t),$$

where $\times$ denotes matrix multiplication. For the preferential transmission model the rate at which a susceptible individual acquires infection from an asymptomatic or symptomatic individual at time t, respectively, are given by

$$\lambda_{a,K\times 1}(t) = \beta^a_{K\times K} \times (I^a_{a,K\times 1}(t) + I^a_{s,K\times 1}(t) + I^a_{K\times 1}(t)),$$

$$\lambda_{s,K\times 1}(t) = \beta^s_{K\times K} \times I^s_{K\times 1}(t).$$

The total force of infection is then $\lambda_{K\times 1}(t) = \lambda_{a,K\times 1}(t) + \lambda_{s,K\times 1}(t)$.

The reproduction numbers for these models can be derived using the next-generation approach [13]. For the non-preferential model the expression for R is given by:

$$R = \max\left(\text{eigenvalues}\left(\frac{\beta^a \Delta \boldsymbol{S}^\top}{\theta} + \frac{(1-\phi)\beta^a \Delta \boldsymbol{S}^\top}{\sigma^a} + \frac{\phi\beta^s \Delta \boldsymbol{S}^\top}{\sigma^s}\right)\right),$$

where $\boldsymbol{A}_{c\times d}\Delta\boldsymbol{B}_{c\times 1}$ operates by multiplying the ith row of $\boldsymbol{A}$ with the ith element of $\boldsymbol{B}$. The expression for R in the preferential model is less straightforward and shown in Appendix B.

2.4. Estimation procedure

We divide the population in five age categories based on the age classes of the incidence data at hand: $0-4$, $5-14$, $15-44$, $45-65$ and 65+. The data consist of reported number of new symptomatic cases per age group per week. We take into account that not all ILI cases are reported via general practitioners and that these under-reporting rates can differ by age.

We use a likelihood-based approach by assuming

$$y_{k,j} \sim \text{Po}(\rho_k \cdot (I^s_{new,k}(j) - I^s_{new,k}(j-1))),$$

where $y_{k,j}$ is the observed number of new cases in age group k in week j. $I^s_{new,k}(t)$ is the expected cumulative number of new symptomatic cases in age group k at time t obtained by solving $\mathrm{d}\boldsymbol{I^s_{new}}(t)/\mathrm{d}t = \phi \cdot \theta \cdot \boldsymbol{I^a_1}(t)$ for the non-preferential model and $\mathrm{d}\boldsymbol{I^s_{new}}(t)/\mathrm{d}t = (1-\phi_a)\cdot\theta\cdot\boldsymbol{I^a_a}(t) + \phi_s \cdot \theta \cdot \boldsymbol{I^a_s}(t)$ in the preferential model. The age-specific reporting rate of ILI cases is denoted by $\rho_k (k = 1, ..., 5)$.

The system of differential equations is initiated by taking into account the pre-existing immunity to the pandemic strain $\boldsymbol{S}(0)$. Furthermore, since we observed a large impact of the initial number of symptomatic cases $\boldsymbol{I^s}(0)$, these five parameters (one for each age category) are included in the estimation procedure. The number of asymptomatic cases at time 0 is assumed to be 0. The initial number of recovered individuals is then $\boldsymbol{R}(0) = \boldsymbol{N} - \boldsymbol{S}(0) - \boldsymbol{I^s}(0)$, with $\boldsymbol{N}$ the population size.

Our aim is to estimate $\phi, \phi_a, \phi_s, q^a, q^r$, ρ_k, and $I_k^s(0)(k = 1, ..., 5)$. Other parameters are assumed known and were obtained from a literature review on influenza transmission models by Dorjee et al. [14]. In this review, parameter values were extracted from studies that estimate (or use) mean, minimum and/or maximum values. These were summarized into three categories: (1) estimated values, where an article attempted to estimate parameters from empirical data; (2) referenced values, where values were adopted from other papers; (3) assumed values, where values were based on expert opinion or unpublished data sources. Parameters were summarized as median and range of means, minimum and maximum values from the reviewed articles. An overview of these parameters is given in Table 1. Note that the subclinical and clinical infectious period refer to the period from infectiousness to recovery for asymptomatic $(1/\theta+1/\sigma^a)$ and symptomatic individuals $(1/\theta + 1/\sigma^s)$, respectively. We use the estimated values when available (median of means), otherwise the referenced values are assumed to be known.

Parameters are estimated via a Markov Chain Monte Carlo (MCMC) approach. This procedure was performed using the LaplacesDemon package [15] in R3.1.1 and R3.2.2. A two-phase approach was used, where the first phase consists of the Adaptive-Mixture Metropolis (AMM) algorithm to achieve stationary samples that seem to have converged to the target distribution. In the second phase Random-Walk Metropolis (RWM), a non-adaptive algorithm, is used to obtain final samples. In this phase 10,000,000 iterations were conducted retaining every 1,000th iteration. Burn-in period is based on the convergence diagnostic by Boone, Merrick and Krachey (BMK) [16]. Uninformative prior distributions

Table 1: An overview of parameters of pandemic influenza A/H1N1 2009 in humans obtained from a literature review [14]. These values were either estimated from empirical data of experimental or observational studies (Est.); or referenced for modeling (Ref.).

Parameter		**Median of means (range)**	**Median of min. values (range)**	**Median of max. values (range)**
Incubation period	Est.	2.0	1.0 (1.0 - 2.0)	-
$1/\gamma + 1/\theta$	Ref.	2.0 (1.5 - 3.0)	1	5
Latent period	Est.	-	-	-
$1/\gamma$	Ref.	1.5 (1 - 3.5)	0.9 (0.7 - 1.0)	4.0 (2.0 - 5.0)
Subclinical	Est.	-	-	-
infectious period	Ref.	1.0 (0.5 - 2.5)	-	2.0
Clinical	Est.	5.6	1.0	10.0 (8.0 - 12.0)
infectious period	Ref.	3.8 (2.5 - 7.0)	3.8 (1.9 - 4.0)	5.5 (2.9 - 10)

are specified for all parameters except for the initial number of cases $\boldsymbol{I^s}(0)$ for which informative priors are used based on the ILI incidence in week 22. A table of prior distributions can be found in Appendix C. To ensure that the estimates lay within their proper parameter space, logit transformations are applied for ϕ, ϕ_a, ϕ_s and $\rho_k (k = 1, ..., 5)$ and log transformations for q^a and q^r. Furthermore, since symptomatic cases are considered to be more infectious than asymptomatic cases, the infectiousness ratio q_r is restricted to be larger than 1.

2.5. Impact of home isolation

One of the possible interventions targeting symptomatic individuals is recommending time off from work. Van Kerckhove et al. [7] showed that contacts made at home are not a proxy for contacts made when symptomatic. Therefore, we assess the impact of individuals staying at home after symptom onset by assuming that a proportion p of symptomatic individuals stays home immediately after symptom onset. The contact matrix for symptomatic individuals C^s is replaced by $pC^s_h + (1-p)C^s$ in which C^s_h is the contact matrix obtained from contacts made at home by symptomatic individuals in the social contact survey.

Hence, we assume that these contact rates do not increase when individuals stay at home. The obtained posterior parameter samples from the (non-)preferential model are used to solve the system of ODEs associated with this isolation model for fixed values of p. This way we can assess the impact of p on the difference in the number of (a)symptomatic cases.

3. Results

Using the social contact matrices and the ILI incidence data, described in Section 2.1, we will look into the estimation of the proportionality factor, q^a, the infectiousness ratio q^r, the reporting rates, $\rho_k (k = 1, ..., 5)$, the proportion of symptomatic infections, ϕ, for the non-preferential model and the proportions ϕ_a and ϕ_s for the preferential model.

After exploratory analyses (see Appendix D), we found that one age-independent reporting rate ρ is not estimable from the data and its value does not affect other parameter estimates. Hence, it is only possible to estimate the relative differences in reporting rates between age categories. We set the reporting rate of a randomly chosen age category fixed as reference category: $\rho_4 = 0.2$. This value of 20% is based on a literature search for reporting rates on ILI and influenza. Since no information on reporting rates was found specifically for H1N1 in England, this search was conducted worldwide including seasonal influenza e.g. [17]. However, since there is so little information on under-reporting, we will only be interpreting the estimated relative differences.

3.1. Non-preferential model

Posterior medians and 95% posterior credible intervals for the estimated parameters and R are shown in Table 2. Trace plots and posterior distribution plots are shown in Appendix A.

The posterior credible intervals for ϕ and q^r are quite wide, indicating that it is difficult to estimate these parameters from the data. This is confirmed by the posterior density plots (Figure A.8). A scatter plot of ϕ versus q^r (Figure A.9)

Table 2: Posterior median, 95% posterior credible intervals and DIC value for the non-preferential and preferential model.

Parameter	Non-preferential	Preferential
$\boldsymbol{\phi}$	$0.15(0.04, 0.39)$	
$\boldsymbol{\phi_a}$		$0.98(0.89, 1.00)$
$\boldsymbol{\phi_s}$		$0.22(0.062, 0.58)$
$\boldsymbol{q^a}$	$0.082(0.069, 0.093)$	$0.10(0.092, 0.12)$
$\boldsymbol{q^r}$	$2.76(1.04, 9.21)$	$2.62(1.04, 9.20)$
$\boldsymbol{\rho_1}$	$0.21(0.18, 0.25)$	$0.20(0.17, 0.24)$
$\boldsymbol{\rho_2}$	$0.20(0.17, 0.22)$	$0.20(0.18, 0.23)$
$\boldsymbol{\rho_3}$	$0.23(0.20, 0.25)$	$0.21(0.19, 0.23)$
$\boldsymbol{\rho_4}$	0.2	0.2
$\boldsymbol{\rho_5}$	$0.15(0.12, 0.19)$	$0.15(0.12, 0.18)$
$\boldsymbol{R}$	$1.36(1.33, 1.40)$	$1.41(1.23, 1.63)$
DIC	298.75	288.49

shows a strong link between both parameters, indicating that we can either estimate the proportion of symptomatic cases or the relative infectiousness from the data at hand. When only 15% of cases develop symptoms, symptomatic cases are estimated to be about 2.76 times more infectious than asymptomatic cases. The reproduction number is estimated to be 1.36. Lastly, the reporting rate is estimated to be about 1.15 times higher in the $15-44$ years age group and 0.75 times lower in the 65+ years age group compared to the reporting rate in the $45-65$ years age group. The estimated incidence is shown in Figure 4 and the number of symptomatic and asymptomatic cases over time are plotted in Figure A.10.

3.2. Preferential model

Posterior medians and 95% posterior credible intervals for the estimated parameters and R are shown in Table 2. Trace plots and posterior distribution plots can be found in Appendix B.

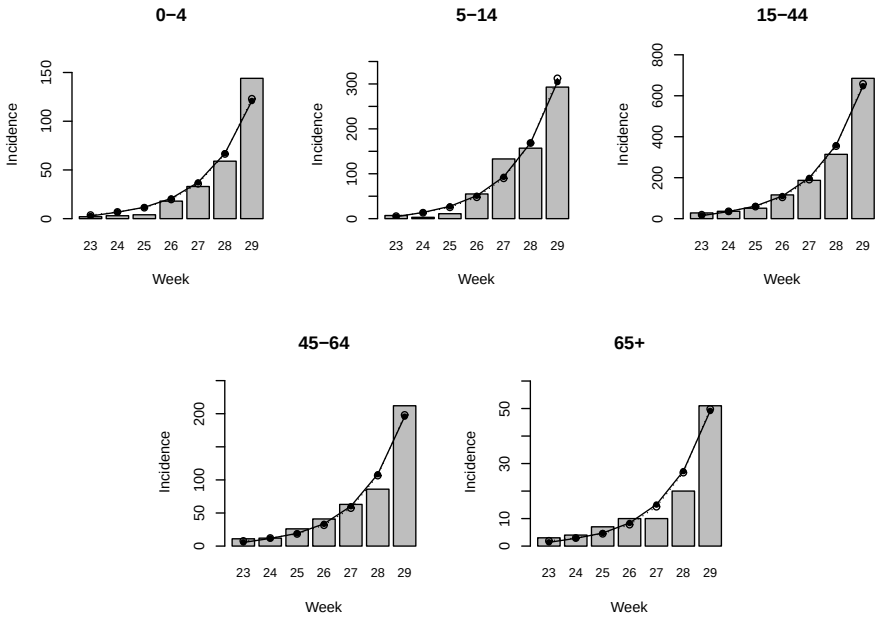

Figure 4: Observed (grey bars) and estimated (connected dots) reported weekly incidence for the five age categories. Full line and filled dots is the estimated incidence for the non-preferential model, dotted line and open dots are the estimates for the preferential model.

In this model, we see similar results as for the non-preferential model. Confidence intervals, posterior distribution plots and a correlation plot show that ϕ_s and q^r are strongly connected and we cannot estimate both from the data (Figures B.12 and B.13). When 22% of cases infected by a symptomatic case develop symptoms, symptomatic cases are estimated to be about 2.62 times more infectious than asymptomatic cases. Furthermore, this model confirms that the reporting rate in the 65+ age class is lower than in the other age categories. Reproduction number is estimated at 1.41. Lastly, this model has a smaller DIC value and, hence, a better fit than the non-preferential model. Plots of incidence and number of cases can be found in Figures 4 and B.14.

To check whether the preferential model simplifies to the non-preferential model ($\phi_s = (1 - \phi_a)$), the difference between ϕ_s and $1 - \phi_a$ is calculated for each

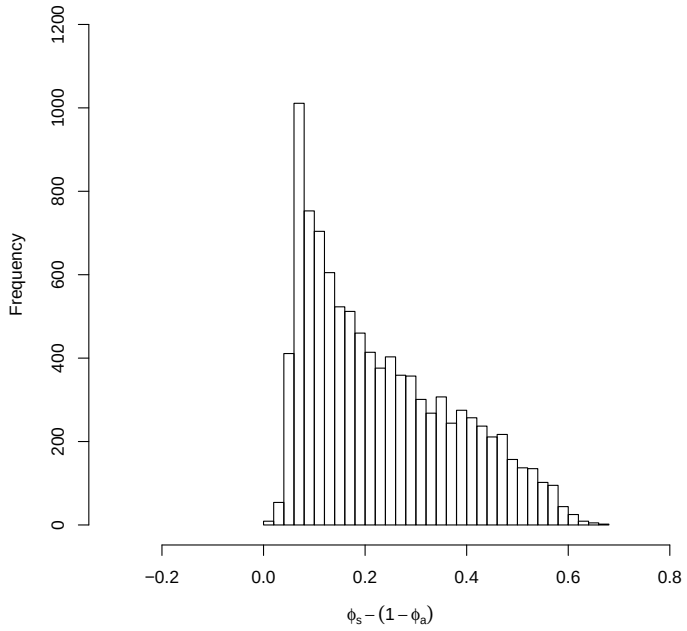

Figure 5: Histogram of MCMC samples for $\phi_s - (1 - \phi_a)$, with ϕ_s the proportion of individuals infected by a symptomatic case that develop symptoms and ϕ_a the proportion of individuals infected by an asymptomatic case that remain asymptomatic in the preferential model.

posterior sample. The histogram of this difference is shown in Figure 5. The 95% credible interval for the difference is $[0.05, 0.55]$. This shows that the preferential model does not simplify to the non-preferential model.

3.3. Impact of home isolation

Figure 6 shows the reduction of cases when a proportion p of symptomatic individuals stays home after symptom onset. As p increases, the reduction in cases also increases. For the non-preferential model, there is no visible difference between symptomatic and asymptomatic cases. Using the preferential model, we do see a larger reduction in symptomatic cases compared to asymptomatic cases. Note that the reduction of cases is larger according to the non-preferential

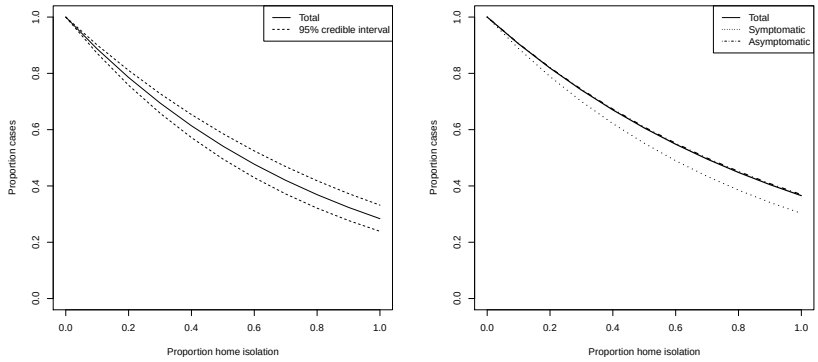

Figure 6: Proportion of cases plotted against the proportion of symptomatic individuals staying home immediately after symptom onset. Left panel: reduction in total number of cases for the non-preferential model with 95% confidence intervals. Right panel: reduction in the number of total, symptomatic and asymptomatic cases for the preferential model.

model in comparison with the preferential model.

4. Discussion

In this paper, we inferred parameters for an epidemic model accounting for asymptomatic transmission and age-dependent under-reporting based on weekly incidence data and social contact data from symptomatic and asymptomatic individuals. The differences in mixing behavior between these individuals affect the expected age-distribution of infection during the early stages of an outbreak [7]. This makes it possible to estimate parameters related to asymptomatic infection using data on symptomatic cases only. Furthermore, we compared a simple SEIR model with asymptomatic infection to a model in which the development of symptoms depends on the status of the infector.

Using a Bayesian approach on ILI data from England and Wales during the early stages of the 2009 epidemic [10], we showed that it is possible to either estimate the proportion of symptomatic infections or the relative infectiousness of symptomatic cases compared to asymptomatic cases in the non-preferential

model. Hence, when one has prior information on one of these parameters, it is possible to estimate the other one from incidence data. Furthermore, although the visual difference in fit between both models is minimal, we found that the data supports the preferential transmission hypothesis i.e. the development of ILI symptoms depends on whether one was infected by a symptomatic or asymptomatic case. Both models show a significantly larger under-reporting rate for people older than 65 years in comparison with $45-65$ year olds. This means that the discrepancy between consultation rates and symptomatic illness rates is larger for the elderly in comparison with the non-elderly adults, although consultation rates in this last age category were found to be lower. Also note that the reporting rates we estimate can possibly account for factors other than the propensity to visit a GP, e.g. the ability to better fit the data because of working with a hidden layer [18]. This can be due to discrepancies between the true contact rates underlying infection and the social contact proxies, age-specific differences in susceptibility and infectiousness, etc. Age-dependent proportionality parameters could also have been used to account for these age-specific differences. Lastly, we assessed the effect of symptomatic individuals staying at home. Following the preferential transmission hypothesis, we found a reduction in total number of cases of 39% $(0.30, 0.45)$ when 50% of individuals would stay home immediately after symptom onset. If all symptomatic individuals would stay home, a reduction of 63% $(0.53, 0.70)$ is observed. To assess more subtle scenarios of home isolation, we will use individual-based models in future research.

Recently, Lin and colleagues [19] explored the trade-off between contact rates and infectiousness (i.e. decreasing contact rates and increasing infectiousness with increasing symptom severity) using a model similar to our non-preferential model. They found that R_0 varies non-monotonically with symptom severity, implying that certain interventions such as antivirals for influenza, can increase R_0. Their research highlights the importance of using empirical data describing the relation between contact rates and symptom severity in epidemiological models.

The preferential model resembles the infector-dependent severity (IDS) model described by Ball and colleagues [20]. However, they assume a homogeneously mixing population and do not estimate model parameters. They derived a threshold limit theorem for their model and looked at the effect of vaccination. They showed that in certain scenarios the proportion of mildly (asymptomatic in our setting) infected individuals can increase with increasing vaccination coverage. This emphasizes the practical importance of our model for a wide range of pathogens with different levels of symptom severity.

There are some limitations to our approach that need to be discussed. One of them is that the reporting rates are not estimable from the data. Hence, one can only infer on the relative differences in reporting rates between age categories. To estimate the true number of cases, information on the reporting rate in at least one age class is needed. Further, we assumed constant reporting rates over time. Though we believe this is of limited impact, since we consider a relatively short period of time (first 7 weeks of the epidemic). Also, the obtained estimates rely on the values of the fixed parameters as found in the literature. Changing these parameters will affect the estimated target parameters. Lastly, we assumed that the contact behavior of asymptomatic cases is similar to the contacts of recovered individuals. We believe this a reasonable assumption, which is partly supported by the findings of Van Kerckhove et al. [7]. They found that the contact patterns from recovered individuals were similar to the contact patterns observed in the POLYMOD study. However it is possible that asymptomatic cases do not feel entirely in best shape and therefore might have a different contact behavior than healthy individuals. Also, we use social contact data and incidence data from A/H1N1pdm in 2009, thus it is uncertain how our conclusions would apply to other influenza strains.

Future research is needed to clarify the exact role of acquired viral dose in the development of influenza symptoms. Up until now, challenge studies have not given clear results able to confirm or reject our preferential transmission hypothesis [8, 9]. Lastly, to extend this model for other diseases more empirical data on how contact rates change with symptom severity are needed.

Acknowledgements

We gratefully acknowledge Ken Eames for fruitful discussions at the initial phase of the project. ES acknowledges support from a Methusalem research grant from the Flemish government. NH acknowledges support from the Antwerp University scientific chair in Evidence-Based Vaccinology, financed in 2009-2014 by an unrestricted gift from Pfizer. Support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy) is gratefully acknowledged. The computational resources and services used in this work were provided by by the VSC (Flemish Supercomputer Center), funded by the Hercules Foundation and the Flemish Government - department EWI.

References

[1] C. Fraser, S. Riley, R. M. Anderson, N. M. Ferguson, Factors that make an infectious disease outbreak controllable., Proceedings of the National Academy of Sciences of the United States of America 101 (16) (2004) 6146–6151. doi:10.1073/pnas.0307506101.

[2] H. Inaba, H. Nishiura, The state-reproduction number for a multistate class age structured epidemic system and its application to the asymptomatic transmission model., Mathematical biosciences 216 (1) (2008) 77–89. doi:10.1016/j.mbs.2008.08.005.

[3] K. Ejima, K. Aihara, H. Nishiura, The impact of model building on the transmission dynamics under vaccination: observable (symptom-based) versus unobservable (contagiousness-dependent) approaches., PloS one 8 (4) (2013) e62062. doi:10.1371/journal.pone.0062062.

[4] J. Wallinga, P. Teunis, M. Kretzschmar, Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents, American Journal of Epidemiology 164 (10) (2006) 936–944. doi:10.1093/aje/kwj317.

[5] N. Goeyvaerts, N. Hens, B. Ogunjimi, M. Aerts, Z. Shkedy, P. V. Damme, P. Beutels, Estimating infectious disease parameters from data on social contacts and serological status, Journal of the Royal Statistical Society: Series C (Applied Statistics) 59 (2) (2010) 255–277. doi:10.1111/j.1467-9876.2009.00693.x.
URL http://doi.wiley.com/10.1111/j.1467-9876.2009.00693.x

[6] K. T. D. Eames, N. L. Tilston, P. J. White, E. Adams, W. J. Edmunds, The impact of illness and the impact of school closure on social contact patterns, Health Technology Assessment 14 (34) (2010) 267–312. doi:10.3310/hta14340-04.

[7] K. Van Kerckhove, N. Hens, W. J. Edmunds, K. T. D. Eames, The impact of illness on social networks: implications for transmission and control of influenza., American journal of epidemiology 178 (11) (2013) 1655–1662. doi:10.1093/aje/kwt196.

[8] F. Carrat, E. Vergu, N. M. Ferguson, M. Lemaitre, S. Cauchemez, S. Leach, A. J. Valleron, Time lines of infection and disease in human influenza: A review of volunteer challenge studies, American Journal of Epidemiology 167 (7) (2008) 775–785. doi:10.1093/aje/kwm375.

[9] Y. Huang, A. K. Zaas, A. Rao, N. Dobigeon, P. J. Woolf, T. Veldman, N. C. Øien, M. T. McClain, J. B. Varkey, B. Nicholson, L. Carin, S. Kingsmore, C. W. Woods, G. S. Ginsburg, A. O. Hero, Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influen PLoS genetics 7 (8) (2011) e1002234. doi:10.1371/journal.pgen.1002234.
URL http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1002234{#}

[10] Public Health England, Weekly epidemiological updates archive, http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/PandemicInfluenza/H1N1P accessed: December 20, 2010 (2010).

[11] E. Miller, K. Hoschler, P. Hardelid, E. Stanford, N. Andrews, M. Zambon, Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study., Lancet 375 (9720) (2010) 1100–1108. doi:10.1016/S0140-6736(09)62126-7.

[12] N. Hens, Z. Shkedy, M. Aerts, C. Faes, P. Van Damme, P. Beutels, Modeling Infectious Disease Parameters Based on Serological and Social Contact Data: a Modern Statistical Perspective, Springer-Verlag New York Inc., 2012.

[13] O. Diekmann, J. Heesterbeek, , J. Metz, On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations., Journal of Mathematical Biology 28 (1990) 365–382.

[14] S. Dorjee, Z. Poljak, C. W. Revie, J. Bridgland, B. Mcnab, E. Leger, J. Sanchez, A review of simulation modelling approaches used for the spread of zoonotic influenza viruses in animal and human populations, Zoonoses and Public Health 60 (6) (2013) 383–411. doi:10.1111/zph.12010.

[15] B. Hall, LaplacesDemon: An R Package for Bayesian Inference, R package version.
URL ftp://ftp.yz.yamagata-u.ac.jp/mirror7/cran/web/packages/LaplacesDemon/vignettes/

[16] E. L. Boone, J. R. Merrick, M. J. Krachey, A Hellinger distance approach to MCMC diagnostics, Journal of Statistical Computation and Simulation 84 (4) (2012) 833–849. doi:10.1080/00949655.2012.729588.
URL http://www.tandfonline.com/doi/abs/10.1080/00949655.2012.729588?journalCode=gscs

[17] A. Lunelli, C. Rizzo, S. Puzelli, A. Bella, E. Montomoli, M. C. Rota, I. Donatelli, A. Pugliese, Understanding the dynamics of seasonal influenza in Italy: incidence, transmissibility and population susceptibility in a 9-year period, Influenza and Other Respiratory Viruses 7 (3) (2013) 286–295. doi:10.1111/j.1750-2659.2012.00388.x.

[18] A. Azmon, C. Faes, N. Hens, On the estimation of the reproduction number based on misreported epidemic data, Statistics in Medicine 33 (7) (2014) 1176–1192. doi:10.1002/sim.6015.

[19] C.-J. Lin, K. A. Deger, J. H. Tien, Modeling the trade-off between transmissibility and contact in infectious disease dynamics., Mathematical biosciences 277 (2016) 15–24. doi:10.1016/j.mbs.2016.03.010.
URL http://www.sciencedirect.com/science/article/pii/S0025556416300062

[20] F. Ball, T. Britton, An epidemic model with infector-dependent severity, Advances in Applied Probability 39 (4) (2007) 949–972.
URL http://projecteuclid.org/euclid.aap/1198177234

Appendix A. Non-preferential model

The system of ordinary differential equations (ODEs) for the non-preferential model is given by

$$\begin{cases} \frac{d\boldsymbol{S}(t)}{dt} &= -\lambda \boldsymbol{S}(t) \\ \frac{d\boldsymbol{E}(t)}{dt} &= \lambda \boldsymbol{S}(t) - \gamma \boldsymbol{E}(t) \\ \frac{d\boldsymbol{I_1^a}(t)}{dt} &= \gamma \boldsymbol{E}(t) - \theta \boldsymbol{I_1^a}(t) \\ \frac{d\boldsymbol{I_2^a}(t)}{dt} &= (1-\phi)\theta \boldsymbol{I_1^a}(t) - \sigma^a \boldsymbol{I_2^a}(t) \\ \frac{d\boldsymbol{I^s}(t)}{dt} &= \phi\theta \boldsymbol{I_1^a}(t) - \sigma^s \boldsymbol{I^s}(t) \\ \frac{d\boldsymbol{R}(t)}{dt} &= \sigma^a \boldsymbol{I_2^a}(t) + \sigma^s \boldsymbol{I^s}(t) \end{cases}$$

The next generation matrix for this model corresponding with the infected states (E, I_1^a, I_2^a, I^s) is given by

$$\boldsymbol{G}_{NP} = \begin{pmatrix} \frac{\beta^a \Delta \boldsymbol{S}^\top}{\theta} + \frac{(1-\phi)\beta^a \Delta \boldsymbol{S}^\top}{\sigma^a} + \frac{\phi\beta^s \Delta \boldsymbol{S}^\top}{\sigma^s} & \frac{\beta^a \Delta \boldsymbol{S}^\top}{\theta} + \frac{(1-\phi)\beta^a \Delta \boldsymbol{S}^\top}{\sigma^a} + \frac{\phi\beta^s \Delta \boldsymbol{S}^\top}{\sigma^s} & \frac{\beta^a \Delta \boldsymbol{S}^\top}{\sigma^a} & \frac{\beta^s \Delta \boldsymbol{S}^\top}{\sigma^s} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

Therefore the reproduction number for the non-preferential model is given by

$$R = \max\left(\text{eigenvalue}\left(\boldsymbol{G}_{NP}\right)\right) = \max\left(\text{eigenvalue}\left(\frac{\beta^a \Delta \boldsymbol{S}^\top}{\theta} + \frac{(1-\phi)\beta^a \Delta \boldsymbol{S}^\top}{\sigma^a} + \frac{\phi\beta^s \Delta \boldsymbol{S}^\top}{\sigma^s}\right)\right),$$

where $\boldsymbol{A}_{c\times d}\Delta\boldsymbol{B}_{c\times 1}$ operates by multiplying the ith row of $\boldsymbol{A}$ with the ith element of $\boldsymbol{B}$.

Trace plots, posterior and prior distributions for the parameters are plotted in Figures A.7 and A.8. A scatter plot of posterior samples for q^r versus ϕ is presented in Figures A.9. Furthermore, the estimated number of symptomatic and asymptomatic cases in each age category are shown in Figure A.10.

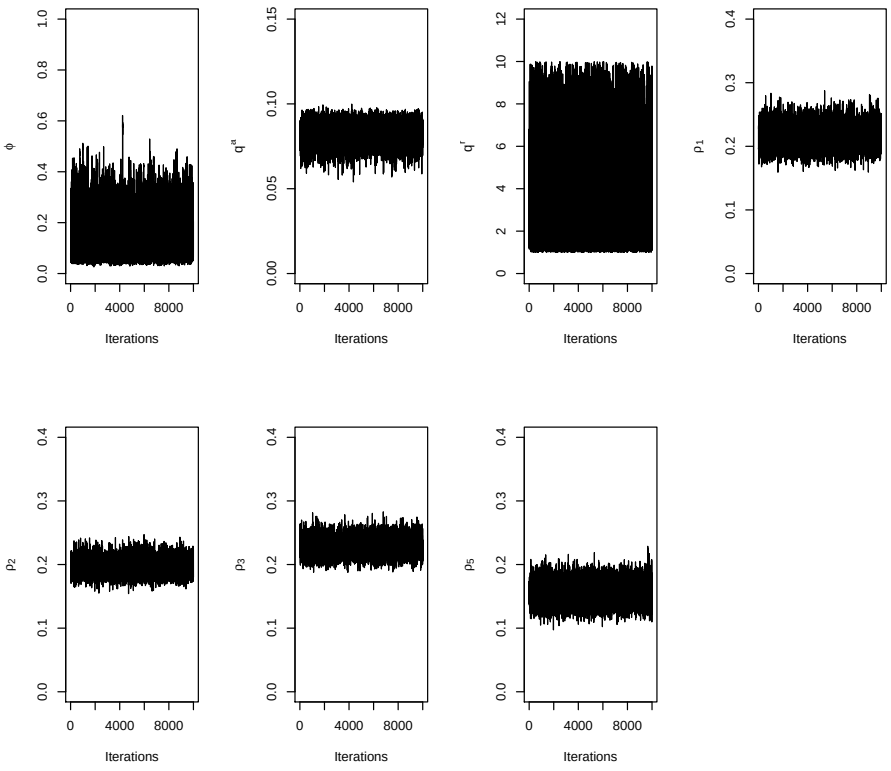

Figure A.7: Trace plot of the MCMC samples for the proportion of cases that develop symptoms (ϕ), the proportionality factor for asymptomatic individuals (q^a), the relative infectiousness of symptomatic cases versus asymptomatic cases (q^r) and the reporting rates ($\rho_i, i = 1, 2, 3, 5$).

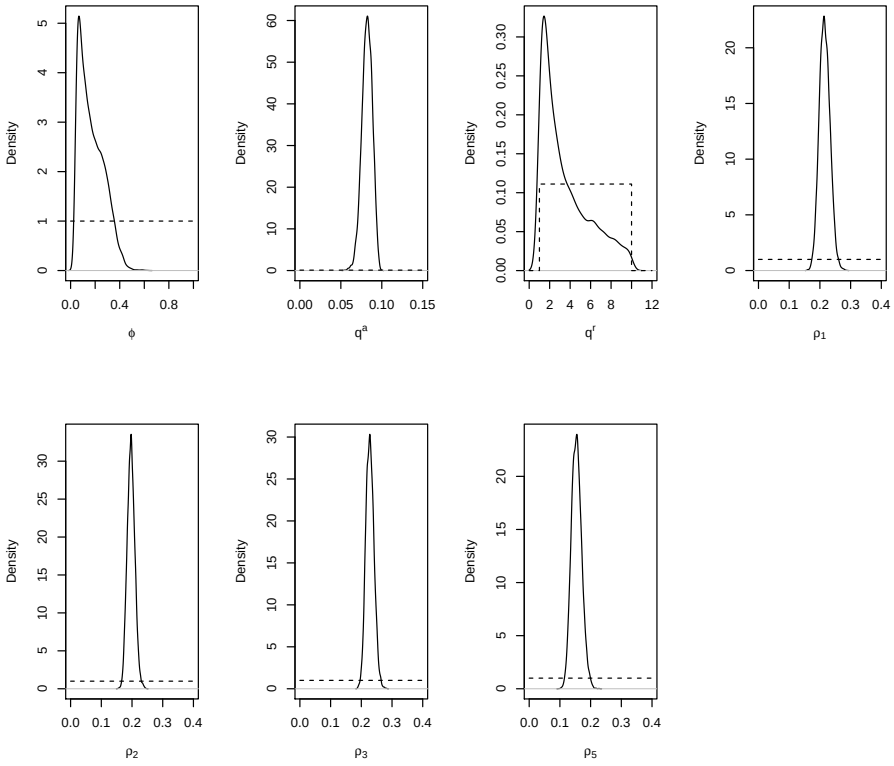

Figure A.8: Prior (dotted line) and posterior (full line) distributions for the proportion of cases that develop symptoms (ϕ), the proportionality factor for asymptomatic individuals (q^a), the relative infectiousness of symptomatic cases versus asymptomatic cases (q^r) and the reporting rates ($\rho_i, i = 1, 2, 3, 5$).

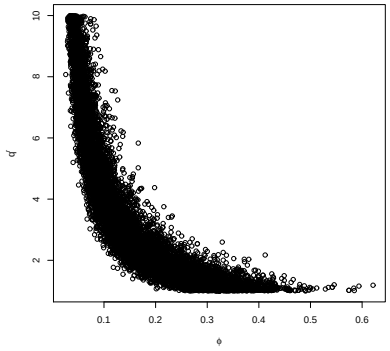

Figure A.9: Scatterplot of the infectiousness ratio q_r versus the proportion of symptomatic cases ϕ.

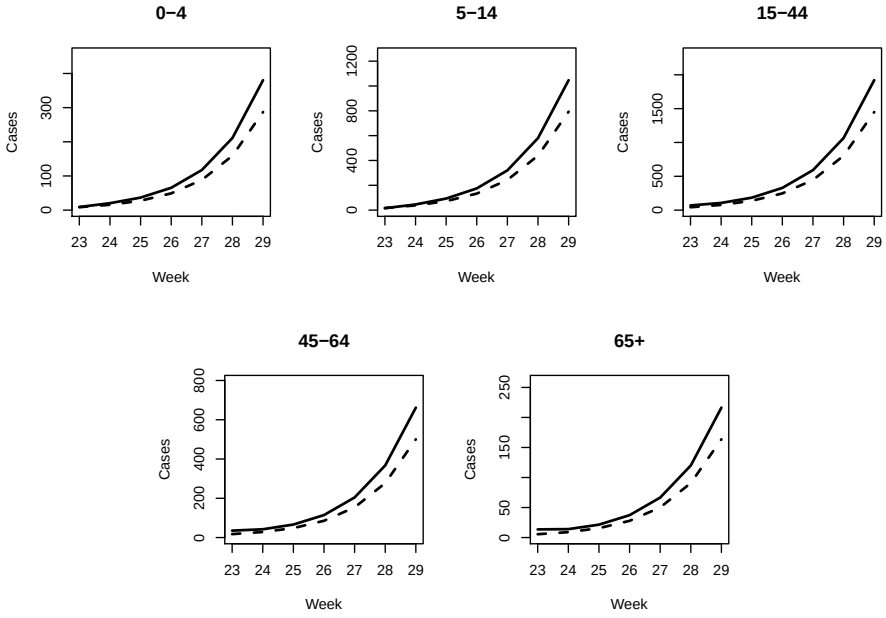

Figure A.10: Number of symptomatic (full line) and asymptomatic (dotted line) cases over time for the five age categories assuming a 20% reporting rate in the $45-65$ age class.

Appendix B. Preferential model

The system of ordinary differential equations (ODEs) for the preferential model is given by

$$\begin{cases} \frac{d\boldsymbol{S}(t)}{dt} &= -(\lambda_a+\lambda_s)\boldsymbol{S}(t) \\ \frac{d\boldsymbol{E_a}(t)}{dt} &= \lambda_a\boldsymbol{S}(t)-\gamma_a\boldsymbol{E_a}(t) \\ \frac{d\boldsymbol{E_s}(t)}{dt} &= \lambda_s\boldsymbol{S}(t)-\gamma_s\boldsymbol{E_s}(t) \\ \frac{d\boldsymbol{I_a^a}(t)}{dt} &= \gamma_a\boldsymbol{E_a}(t)-\theta_a\boldsymbol{I_a^a}(t) \\ \frac{d\boldsymbol{I_s^a}(t)}{dt} &= \gamma_s\boldsymbol{E_s}(t)-\theta_s\boldsymbol{I_s^a}(t) \\ \frac{d\boldsymbol{I^a}(t)}{dt} &= (\phi_a)\theta_a\boldsymbol{I_a^a}(t)+(1-\phi_s)\theta_s\boldsymbol{I_s^a}(t)-\sigma^a\boldsymbol{I^a}(t) \\ \frac{d\boldsymbol{I^s}(t)}{dt} &= (1-\phi_a)\theta_a\boldsymbol{I_a^a}(t)+\phi_s\theta_s\boldsymbol{I_s^a}(t)-\sigma^s\boldsymbol{I^s}(t) \\ \frac{d\boldsymbol{R}(t)}{dt} &= \sigma^a\boldsymbol{I^a}(t)+\sigma^s\boldsymbol{I^s}(t) \end{cases}$$

The next generation matrix for this model corresponding with the infected states $(E_a, E_s, I_a^a, I_s^a, I^a, I^s)$ is given by

$$\boldsymbol{G}_P = \begin{pmatrix} \frac{\beta^a\Delta\boldsymbol{S}^\top}{\theta}+\frac{\phi_a\beta^a\Delta\boldsymbol{S}^\top}{\sigma^a} & \frac{\beta^a\Delta\boldsymbol{S}^\top}{\theta}+\frac{(1-\phi_s)\beta^a\Delta\boldsymbol{S}^\top}{\sigma^a} & \frac{\beta^a\Delta\boldsymbol{S}^\top}{\theta}+\frac{\phi_a\beta^a\Delta\boldsymbol{S}^\top}{\sigma^a} & \frac{\beta^a\Delta\boldsymbol{S}^\top}{\theta}+\frac{(1-\phi_s)\beta^a\Delta\boldsymbol{S}^\top}{\sigma^a} & \frac{\beta^a\Delta\boldsymbol{S}^\top}{\sigma^a} & \mathbf{0} \\ \frac{(1-\phi_a)\beta^s\Delta\boldsymbol{S}^\top}{\sigma^s} & \frac{\phi_s\beta^s\Delta\boldsymbol{S}^\top}{\sigma^s} & \frac{(1-\phi_a)\beta^s\Delta\boldsymbol{S}^\top}{\sigma^s} & frac\phi_s\beta^s\Delta\boldsymbol{S}^\top\sigma^s & \mathbf{0} & \frac{\beta^s\Delta\boldsymbol{S}^\top}{\sigma^s} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Therefore the reproduction number for the preferential model is given by

$$R = \max\left(\text{eigenvalues}\left(\boldsymbol{G}_P\right)\right).$$

Trace plots, posterior and prior distributions for the parameters are plotted in Figures B.11 and B.12. A scatter plot of posterior samples for q^r versus ϕ_s is presented in Figures B.13. Furthermore, the estimated number of symptomatic and asymptomatic cases in each age category are shown in Figure B.14.

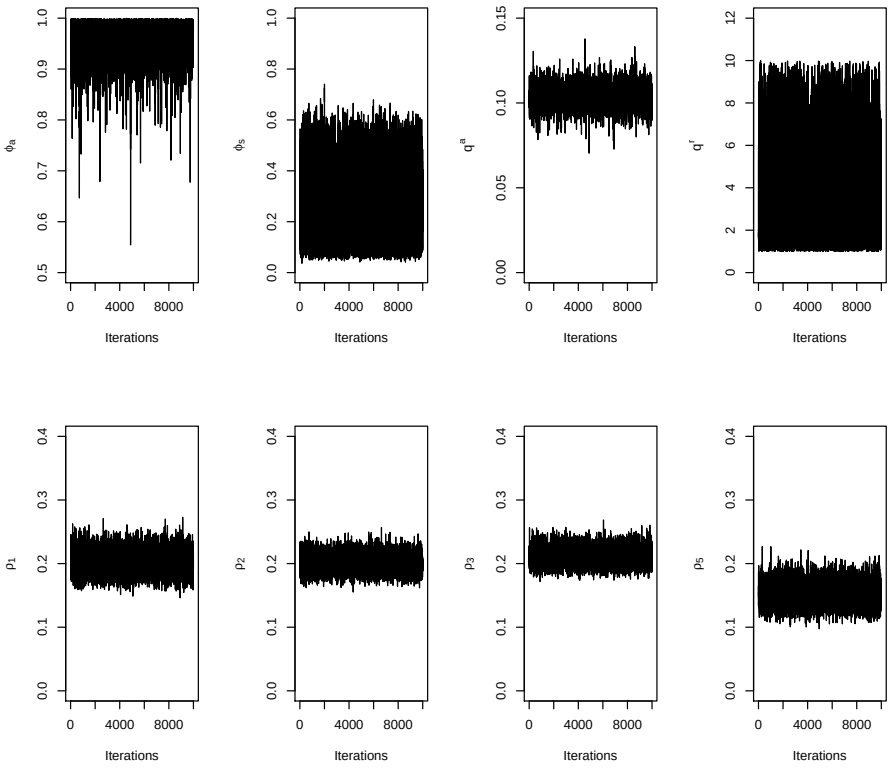

Figure B.11: Trace plot of the MCMC samples for the proportion of individuals infected by a symptomatic case that develop symptoms (ϕ_s), the proportion of individuals infected by an asymptomatic case that remain asymptomatic (ϕ_a), the proportionality factor for asymptomatic individuals (q^a), the relative infectiousness of symptomatic cases versus asymptomatic cases (q^r) and the reporting rates ($\rho_i, i = 1, 2, 3, 5$).

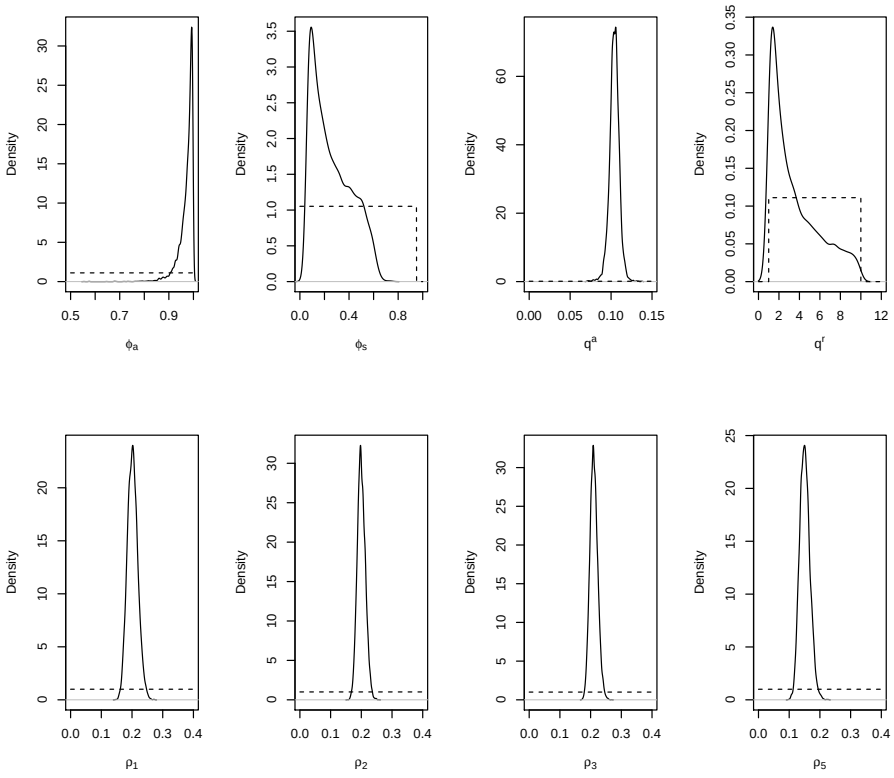

Figure B.12: Prior (dotted line) and posterior (full line) distributions for the proportion of individuals infected by a symptomatic case that develop symptoms (ϕ_s), the proportion of individuals infected by an asymptomatic case that remain asymptomatic (ϕ_a), the proportionality factor for asymptomatic individuals (q^a), the relative infectiousness of symptomatic cases versus asymptomatic cases (q^r) and the reporting rates ($\rho_i, i = 1, 2, 3, 5$).

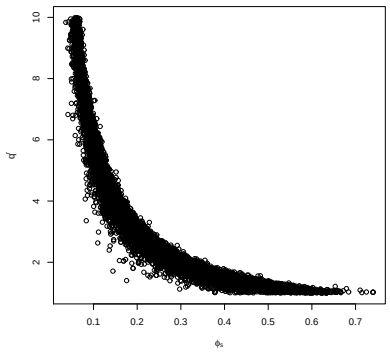

Figure B.13: Scatterplot of the infectiousness ratio q_r versus the proportion of symptomatic individuals infected by a symptomatic case ϕ_s.

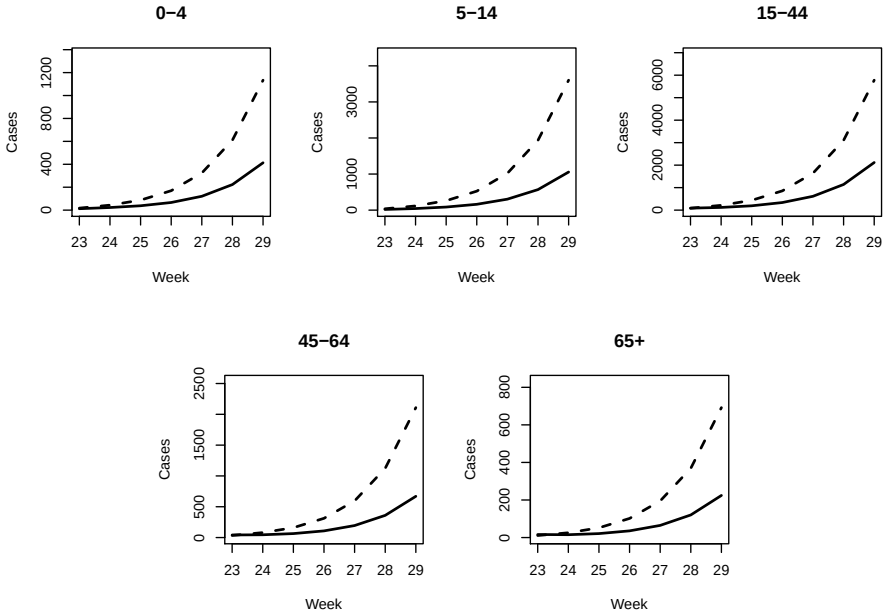

Figure B.14: Number of symptomatic (full line) and asymptomatic (dotted line) cases over time for the five age categories assuming a 20% reporting rate in the $45-65$ age class.

Table C.3: Prior distributions.

Parameter	Prior distribution
$I_k^s(0)$	$N(\mu_k, \delta_k)(k = 1, ..., 5)$; truncated$(0.1, 1000)$
ϕ	$U(0, 1)$
ϕ_a	$U(0.1, 1)$
ϕ_s	$U(0, 0.95)$
q^a	$U(0, 10)$
q^r	$U(1, 10)$
ρ_k	$U(0, 1)$

Appendix C. Prior distributions

The univariate prior distributions for all parameters are given in Table C.3. All of these are uninformative, except for the initial number of symptomatic cases. The number of symptomatic cases at time 0 in age class i is assumed to follow a truncated normal distribution, with mean μ_k based on the observed ILI incidence for age class k in week 22 and standard deviation δ_k.

Appendix D. Exploratory analyses

We use an age-independent reporting rate ρ and fix it at the values $0.3, 0.5, 1.0$ and repeat the estimation procedure for each value for both the non-preferential and preferential model.

For both models, the value of an age-independent reporting rate ρ does not affect model fit or other parameter estimates. For this reason, we use a reference reporting rate in one age category in our models with age-dependent reporting.

Table D.4: Posterior median, 95% posterior intervals and DIC value for the non-preferential model for different values of the reporting rate ρ.

Parameter	$\rho = 0.3$	$\rho = 0.5$	$\rho = 1.0$
$\boldsymbol{\phi}$	$0.34(0.10, 0.90)$	$0.34(0.083, 0.86)$	$0.35(0.083, 0.90)$
$\boldsymbol{q^a}$	$0.060(0.035, 0.079)$	$0.062(0.039, 0.081)$	$0.062(0.039, 0.080)$
$\boldsymbol{q^r}$	$2.79(1.05, 8.90)$	$2.68(1.04, 9.52)$	$2.67(1.05, 9.18)$
$\boldsymbol{\rho_1}$	0.3	0.5	1.0
$\boldsymbol{\rho_2}$	0.3	0.5	1.0
$\boldsymbol{\rho_3}$	0.3	0.5	1.0
$\boldsymbol{\rho_4}$	0.3	0.5	1.0
$\boldsymbol{\rho_5}$	0.3	0.5	1.0
$\boldsymbol{R}$	$1.47(1.39, 1.56)$	$1.46(1.39, 1.55)$	$1.46(1.39, 1.55)$
DIC	306.61	307.05	307.56

Table D.5: Posterior median, 95% posterior intervals and DIC value for the preferential model for different values of the reporting rate ρ.

Parameter	$\rho = 0.3$	$\rho = 0.5$	$\rho = 1.0$
$\boldsymbol{\phi_a}$	$0.98(0.91, 1.00)$	$0.99(0.92, 1.00)$	$0.99(0.92, 1.00)$
$\boldsymbol{\phi_s}$	$0.28(0.07, 0.64)$	$0.29(0.07, 0.65)$	$0.25(0.07, 0.63)$
$\boldsymbol{q^a}$	$0.11(0.093, 0.12)$	$0.11(0.092, 0.12)$	$0.11(0.092, 0.12)$
$\boldsymbol{q^r}$	$2.32(1.03, 8.97)$	$2.24(1.04, 9.05)$	$2.53(1.03, 9.35)$
$\boldsymbol{\rho_1}$	0.3	0.5	1.0
$\boldsymbol{\rho_2}$	0.3	0.5	1.0
$\boldsymbol{\rho_3}$	0.3	0.5	1.0
$\boldsymbol{\rho_4}$	0.3	0.5	1.0
$\boldsymbol{\rho_5}$	0.3	0.5	1.0
$\boldsymbol{R}$	$1.45(1.27, 1.62)$	$1.45(1.26, 1.62)$	$1.44\ (1.26,\ 1.61)$
DIC	290.97	291.78	291.43