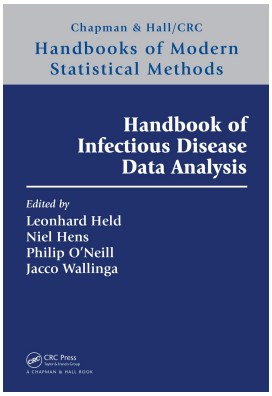

Handbook of Infectious Disease Data Analysis

Leonhard Held, Niel Hens, Philip O'Neill, Jacco Wallinga

Analysing Multiple Epidemic Data Sources

Publication details
https://www.routledgehandbooks.com/doi/10.1201/9781315222912-24
Daniela De Angelis, Anne M. Presanis
Published online on: 04 Nov 2019

24

Analysing Multiple Epidemic Data Sources

Daniela De Angelis and Anne M. Presanis

CONTENTS

24.1 Introduction

> "A catalogue of the number of deaths induced by the major epidemics of historical time is staggering, and dwarfs the total deaths on all battlefields." [1]

This quote sets the problem of infectious diseases in perspective. Although major historical threats have been defeated [2], new emerging ones continue to challenge humans. It is not surprising that increasing effort has been made by policy makers to assess and anticipate the consequence of epidemics. Evidence-based knowledge of disease burden, including

prevalence, incidence, severity, and transmission, in different population strata, in different locations and, if feasible, in real time, is becoming progressively key to the planning and evaluation of public health policies [2]. Direct observation of a disease process hardly ever is possible. However, retrospective and prospective estimation of the key aspects of burden just listed is feasible through the use of indirect information collected in administrative registries. The previous chapters in this handbook (and the rich literature that exists [2, 3] and references therein) provide plenty of examples of how surveillance information, together with statistical models, can be used to reconstruct the disease process underlying the pattern of the observed data, infer the unobserved (latent) characteristics of the epidemic, and forecast its evolution.

Here we focus, in particular, on statistical inference that makes simultaneous use of multiple data sources, including different streams of surveillance data, ad hoc studies, and expert opinion.

This "evidence synthesis" approach is not new in medical statistics. Meta-analysis and network meta-analysis are well established approaches to combine data from studies of similar design, typically clinical trials [4]. The idea has been generalized in the areas of medical decision-making [5], technology assessment [6, 7], and epidemiology (e.g., [8]) to assimilate data from sources of different types and studies of different designs and is becoming popular in other scientific fields as modern technologies enable the collection and storage of ever increasing amounts of information (e.g., [9, 10]). For infectious disease, in the last 10 years there has been a proliferation of papers employing multiple sources of information to reconstruct characteristics of epidemics of blood-borne and respiratory diseases, including estimation of prevalence (e.g., HIV [11], HCV [12, 13], and campylobacteriosis [14]), severity (e.g., [15, 16]), incidence (e.g., toxoplasmosis [17], influenza [18], and pertussis [19]) and, transmission (e.g., influenza [20, 21]).

The use of multiple data sources poses a number of statistical and computational challenges: the combination of various sources, most likely affected by selection and informative observation biases and heterogeneous in type, relevance, and granularity, leads to probabilistic models with complex structures that are difficult to build and fit and challenging to criticize [22].

In this chapter, we will use motivating case studies of influenza to introduce the evidence synthesis setting in infectious diseases, illustrate the building and fitting of relevant models, and highlight the opportunities offered and the challenges posed by the multiplicity of sources. The models we will concentrate on are typically Bayesian as this framework offers a natural setup for the synthesis of information.

The chapter is organized as follows. In Section 24.2, we describe our motivating examples; in Section 24.3, the generic framework for evidence synthesis for infectious diseases is introduced; the models developed for the chosen examples are presented in Sections 24.4 and 24.5; Section 24.6 is devoted to the challenges encountered in the building, fitting, and criticism of models that incorporate multiple sources; and we conclude with a final discussion in Section 24.7.

24.2 Motivating Example: Influenza

Public health responses aimed at mitigating the impact of an outbreak need reliable (and prompt) assessment of the likely severity and spread of the infection. This understanding is particularly key when a new pathogen emerges, potentially causing a pandemic, for example a new influenza strain as in 2009 [23] or more recently, the zika [24] and ebola [25] outbreaks.

We will use examples of influenza severity and transmission estimation, in particular referring to the 2009 A/H1N1 pandemic in the United Kingdom, to illustrate how, in the absence of ideal information, estimation of severity and transmission can be carried out by using data from a multiplicity of sources.

24.2.1 Severity

The severity of an infectious disease, such as influenza, can be thought as a "pyramid" (Figure 24.1), where with increasing severity, there are fewer and fewer infections. A proportion of infected individuals progress to symptoms, then to hospitalization, and the most severe end-points of either intensive care unit (ICU) admission or death. Severity is usually expressed as "case-severity risks" (CSRs); i.e. probabilities that an infection leads to a severe event such as being hospitalized, admitted to ICU, or dying. Quantification of such risks is necessary both prospectively, during an outbreak, to understand the severity and likely burden on health-care of an ongoing epidemic; and retrospectively, to assess the severe burden of the particular strain responsible for the epidemic, the adequacy of any public health response during the outbreak, and to inform responses in future outbreaks. However, such CSRs are challenging to directly observe, requiring therefore estimation.

Prospectively, estimation would require a cohort of cases, i.e. individuals with laboratory-confirmed influenza to be followed up over time. However, a representative sample of those who are infected is almost impossible to recruit, particularly as infections that are asymptomatic are less likely to be observed in health-care data than symptomatic infections. Even if it were possible, prospective estimation would have to account appropriately for censoring, as the end points of interest might take time to occur.

For retrospective estimation, censoring may not be an issue; however, differential probabilities of observing cases at different levels of severity (ascertainment/detection probabilities) may lead to biases [23]. "Multiplier" methods [26–28] therefore have been proposed when individual-level survival-type data are not available, combining aggregate case numbers at different levels of severity, for example. surveillance data on sero-prevalence (i.e., the proportion of blood samples testing positive for influenza), general practice [GP] consultations for influenza-like-illness [ILI]; hospital/ICU admissions, and mortality to obtain esti-

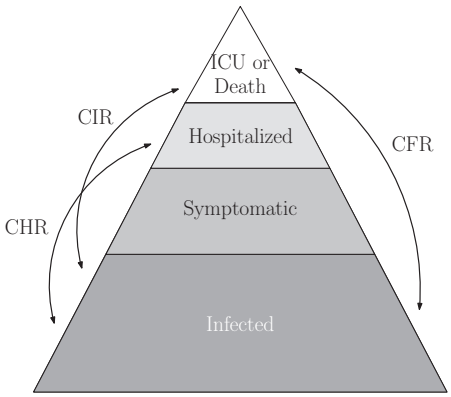

FIGURE 24.1
Severity of influenza as a "pyramid": Infected individuals progress from asymptomatic infection ("I") to symptomatic infection ("S"), hospitalization ("H"), intensive care unit (ICU)-admission ("ICU") and/or death ("D"). Case-severity risks, that is, the case-hospitalization (CHR), case-ICU-admission (CIR), and case-fatality (CFR) risks, are defined as probabilities of a severe event given infection.

mates of the CSRs. These methods account for the ascertainment/detection biases suffered by aggregate surveillance data through multiplication by inverse proportions detected with informal uncertainty quantification using Monte Carlo forward simulation (e.g., [26, 27]).

Hybrid methods that combine hierarchical models with multiplier methods in different stages have appeared recently [28]. However, multiplier methods to estimate severity were first formalized using Bayesian evidence synthesis to simultaneously account for uncertainty, prior information on some ascertainment/detection biases, and censoring [15, 29, 30]. The uncertainty inherent in each data source, together with prior uncertainty, is propagated formally through to posterior estimates of the CSRs. The estimated CSRs are derived as products of probabilities of being at one severity level conditional on being at lower severity level (Figure 24.1).

Such an evidence synthesis is presented in [15], for the A/H1N1 pandemic in the UK during each of three waves experienced in summer 2009, the 2009/2010, and 2010/2011 seasons. The available data sources include (1) cross-sectional sero-prevalence data from laboratory-tested residual sera from patients after diagnostic testing for various (non-respiratory) conditions—these data, over time, inform changes in the proportion of the population exposed to influenza strains (i.e., the population level of immunity, and hence indirectly inform incidence), (2) estimates of numbers symptomatic based on potentially underascertained GP consultations for "influenza-like-illness" (ILI) and corresponding data on the proportion of nasopharyngeal swabs from individuals with ILI that test virologically positive for the A/H1N1pdm strain, (3) underascertained retrospective and prospective daily hospital admissions from 129 hospital trusts in the first two waves, and a sentinel set of 23 trusts in the third wave (Figure 24.2a), and (4) underascertained numbers of deaths occurring in individuals with confirmed A/H1N1 infection (Figure 24.2b). Each source poses a number of challenges in addition to the above-mentioned ascertainment/detection biases. The sero-prevalence data, although available for all three waves, in the second and third waves does not allow separation of individuals with antibodies in response to vaccination from infected individuals. Point estimates of the number symptomatic from the Health Protection Agency (HPA) are only available in the first two waves, with an informal "uncertainty range" from sensitivity analyses. For the third wave, such estimates are instead obtained from a joint regression model of the GP ILI and virological positivity data based on a much smaller sentinel set of general practices than the more comprehensive sentinel system used in the first two waves. Both the GP and positivity datasets are required to disentangle ILI consultations due to "background" consultations for other respiratory illness from actual influenza consultations. The switch from a comprehensive to a sentinel hospital system between the second and third waves results in sparser data and changes in the age groups recorded (Figure 24.3), particularly affecting the number of severe outcomes (ICU admissions and deaths) reported in the hospitalization data: no deaths are observed in the third wave. This sparsity requires the use of an additional ICU data source, which poses its own challenge. The system measures prevalent cases present in ICU rather than incident or cumulative incident ICU admissions, and hence requires a model of the process of admissions and discharges to obtain estimates of cumulative admissions.

None of the data sources on their own can provide an estimate of all CSRs of interest. However, by combining them all in a Bayesian evidence synthesis, the challenges just described can be resolved to derive the necessary severity estimates, as presented in Section 24.4.

24.2.2 Transmission

Understanding the dynamics of an infectious disease amounts to estimation of the rate at which it spreads and the factors that are contributing to its spread. To acquire such

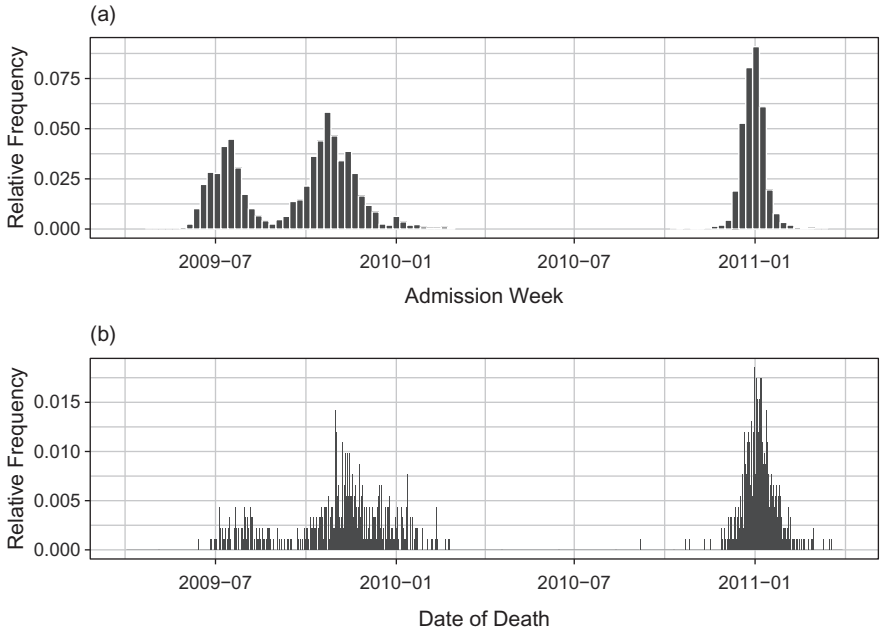

FIGURE 24.2
Frequency of observed severe events relative to the total number of events over the period April 2009-March 2011 of the three waves of infection among individuals with confirmed A/H1N1 pandemic influenza: (a) weekly hospitalizations by admission week; and (b) deaths by date of death.

knowledge, mechanistic transmission models are used [1], expressed by differential equations describing the disease dynamics resulting from the interaction between individuals at different disease stages. For example, in the susceptible-infectious-recovered (SIR) model new infections are generated through the contact between susceptible and infectious individuals (i.e., in state S and I, respectively). For influenza, and other respiratory infections, the relevant contact is between different age groups, since school-age children and their interactions with other children and with adults are known to be key drivers of transmission [31]. Historically, studies of influenza transmission have been carried out either by simulation [2] or by estimating the parameters of transmission models using direct information from a single time series of disease endpoints, such as confirmed cases (e.g., [31]).

However, in recent years the need and potential of combining data from multiple sources to infer latent characteristics of epidemics has been increasingly recognized. For influenza, in particular, since the 2009 A/H1N1 influenza pandemic, this recognition resulted in the development of a number of transmission models using data from either multiple surveillance time series (e.g., [20, 21, 32–35]) or a combination of surveillance and phylogenetic data (e.g., [36–38]). The integration of different sources of evidence can ensure identification of interpretable parameters in transmission models and a more comprehensive description of the evolution of an outbreak [22].

An example is given in [20] where, in the absence of a complete time series of confirmed influenza cases, various data sources are used to estimate retrospectively transmission during the first two waves of pandemic A/H1N1 influenza infection. Figure 24.4 shows the data for the London region: (a) GP consultations for ILI from May to December 2009, (b) a series of cross-sectional samples from sero-prevalence surveys (see Section 24.2.1),

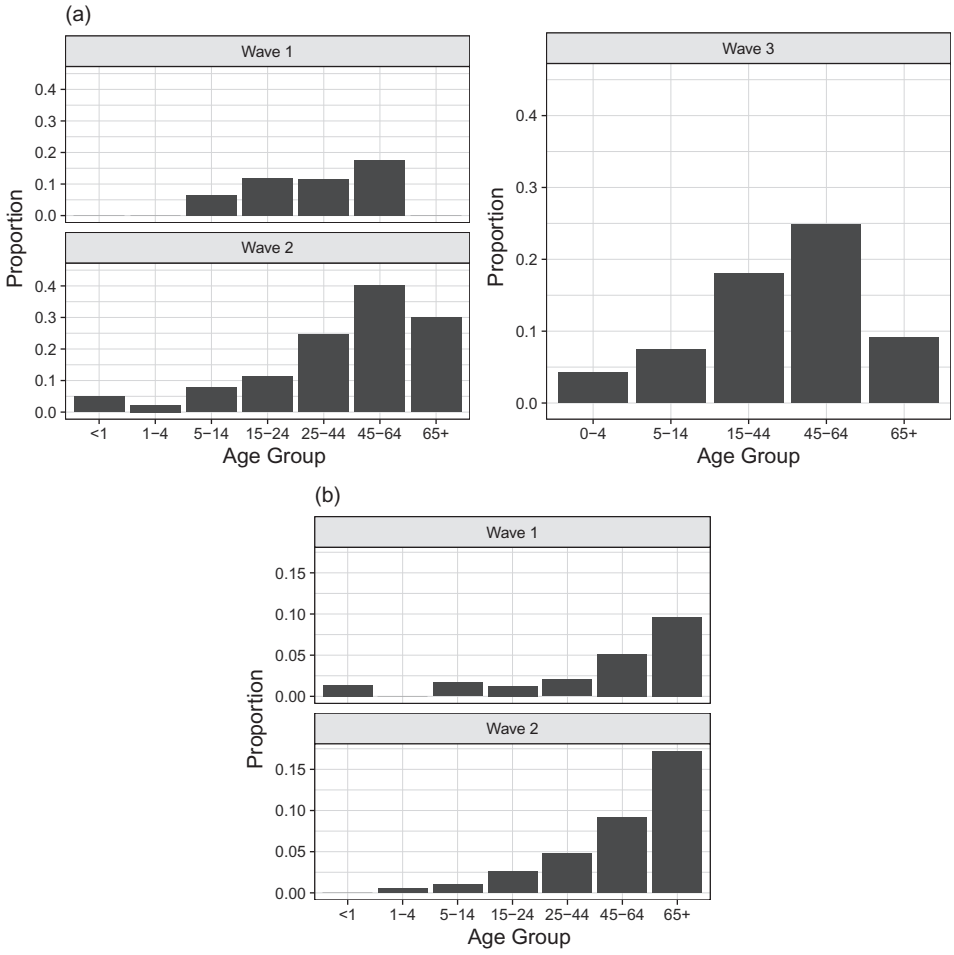

FIGURE 24.3
Proportion of observed individuals hospitalized with confirmed A/H1N1 pandemic influenza who experienced severe events during the three waves of infection: (a) ICU admissions per confirmed A/H1N1 hospitalization by age and wave and (b) deaths per confirmed A/H1N1 hospitalization by age and wave.

(c) virological data on nasopharyngeal positivity for A/H1N1 (again as in Section 24.2.1), and (d) a limited time series of confirmed cases in the first few weeks of the outbreak, up till June 2009, when contact tracing ceased. As in Section 24.2.1, GP consultation data are contaminated by individuals experiencing non-A/H1N1-related ILI, whose health-care seeking behavior is highly influenced by governmental advice and media reporting. To reconstruct the underlying pattern of A/H1N1 infections, GP data had to be combined with information on A/H1N1 virological positivity, on population immunity from the serological surveys, knowledge on the natural history of A/H1N1, including the probability of developing symptoms and data on the propensity of patients with symptomatic infections to consult a GP [39].

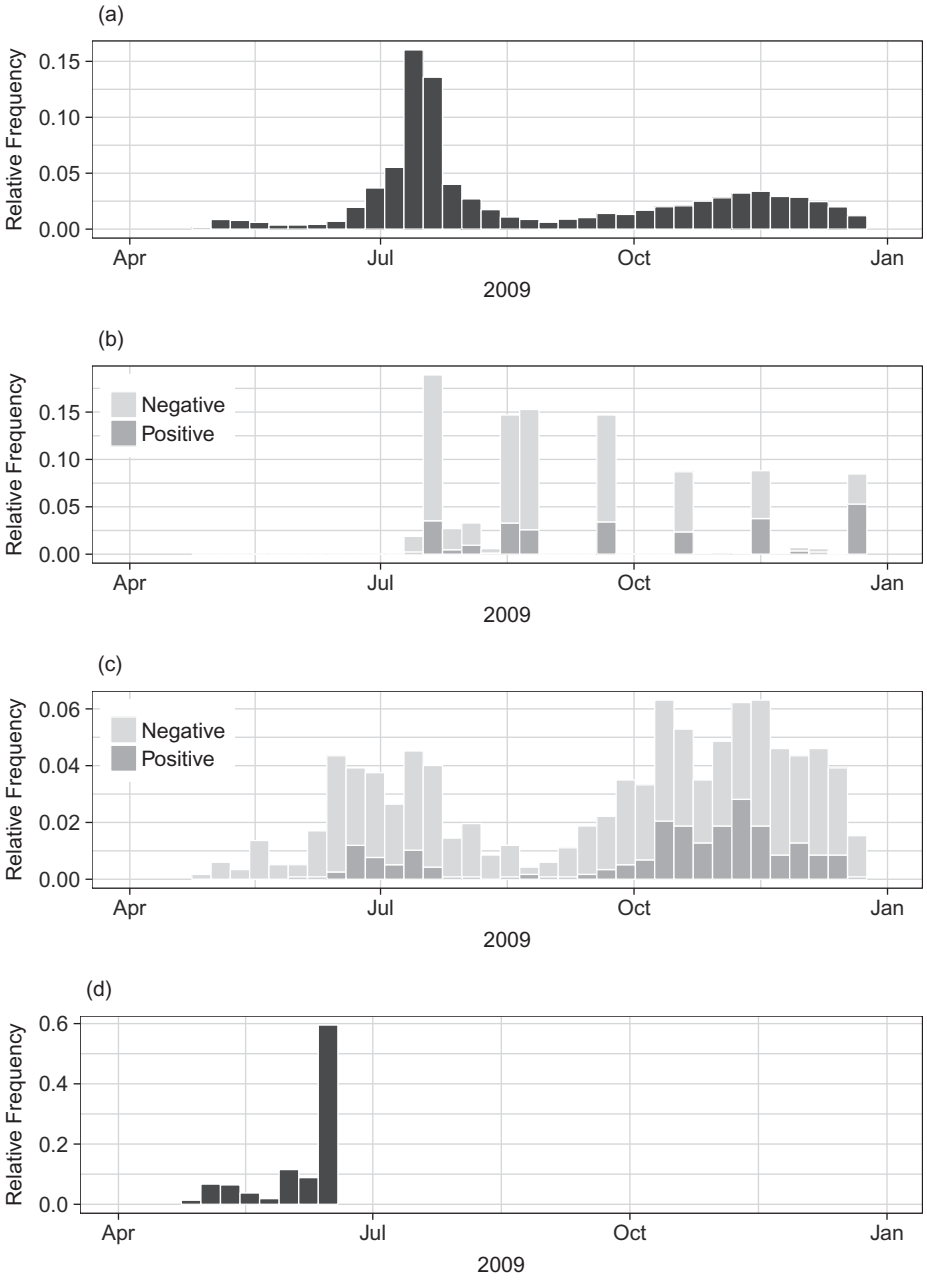

FIGURE 24.4
Four data streams used in the transmission model of [20]: (a) GP ILI consultations, (b) serological positivity, (c) virological positivity, and (d) confirmed cases. Each plot shows the frequency of events, (consultations per 100,000 population, positive and negative sera samples, positive and negative swabs, confirmed cases, respectively) relative to the total number of respective events over the period April 2009-December 2009 of the first two waves of infection.

24.3 Bayesian Evidence Synthesis

The notion of evidence synthesis is intrinsic to the Bayesian philosophy of assimilating information, Bayes' theorem being the basis for the combination of prior and new evidence. Generalizing the concepts of meta-analysis and network meta-analysis, the evidence synthesis described and used here combines information from different study designs through complex hierarchical models [8, 11, 22].

24.3.1 A useful graphical representation

Bayesian hierarchical models have a long history of being expressed as directed acyclic graphs (DAGs), encoding the dependency structure between variables in the model [40].

A generic evidence synthesis model can be represented graphically as in Figure 24.5. Square nodes represent observable quantities such as $\boldsymbol{y}_i$, whereas circles are latent quantities, such as ψ_i. Double circles such as θ_j are *founder* nodes (i.e., parameters to which a prior distribution is assigned). Dashed rectangles, or "plates," represent repetition over indices, such as $i \in 1, \ldots, n$. Dependencies between variables are indicated by direct arrows with solid and dashed arrows representing distributional (stochastic) and functional (deterministic) dependencies, respectively. The joint distribution of all quantities in the DAG is the product of the conditional distributions of each node given its direct parents. The aim of an evidence synthesis model such as Figure 24.5 is to estimate a set of k *basic* parameters $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_k\}$, based on a set of n independent datasets $\boldsymbol{y} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$, where n is not necessarily equal to k. Each dataset $\boldsymbol{y}_i, i \in 1, \ldots, n$ is assumed to inform a quantity (a *functional* parameter) $\psi_i = \psi_i(\boldsymbol{\theta})$ that can be expressed as a deterministic function of the basic parameters. If $\psi_i \equiv \theta_j$ for some $j \in 1, \ldots, k$, $\boldsymbol{y}_i$ is said to *directly* inform θ_j. Otherwise, for $\psi_i = \psi_i(\boldsymbol{\theta})$, $\boldsymbol{y}_i$ *indirectly* informs multiple parameters in the basic parameter set, in conjunction with all the other datasets. Further functional quantities $\psi_i = \psi_i(\boldsymbol{\theta}), i > n$ may be of interest to derive from the basic parameters, even if no data directly inform such functions. Assuming the independence of each dataset $\boldsymbol{y}_i, i \in 1, \ldots, n$ conditional on their common parents, the posterior distribution of the basic parameters $\boldsymbol{\theta}$ given the data $\boldsymbol{y}$ is

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto p(\boldsymbol{\theta}) \prod_{i=1}^{n} p(\boldsymbol{y}_i \mid \psi_i(\boldsymbol{\theta})) = p(\boldsymbol{\theta}) \prod_{i=1}^{n} p(\boldsymbol{y}_i \mid \boldsymbol{\theta}).$$

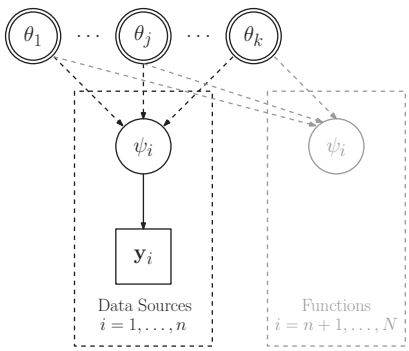

FIGURE 24.5
DAG of a generic evidence synthesis model.

Such an evidence synthesis model and DAG can clearly be extended both horizontally (more datasets) and vertically (hierarchical modelling).

Many evidence synthesis models used in the epidemic literature can be usefully represented graphically by DAGs [35, 41, 42]. Figure 24.6 is a generic DAG representation of a population-level multi-state disease transmission model where the vector-valued $\boldsymbol{X}(t) = (X_1(t), \ldots, X_K(t))$ corresponds to either the number or proportion of the population in each disease state $1, \ldots, K$ at time t. Movement between states is governed by transition rates $\boldsymbol{\lambda}(t)$, which are parameterized in terms of a collection of unknown basic parameters $\boldsymbol{\theta}(t)$, and the current state of the system $\boldsymbol{X}(t)$. If transmission is not explicitly modeled, the dependence of $\boldsymbol{\lambda}(t)$ on the states $\boldsymbol{X}(t)$, represented by the dashed arrow from $\boldsymbol{X}(t)$ to $\boldsymbol{\lambda}(t)$, is removed, and the model simplifies to a standard linear multi-state model. Typically, the $\boldsymbol{X}(t)$ and $\boldsymbol{\lambda}(t)$ are not directly observed. Instead, as in the simpler case of Figure 24.5, observations $\boldsymbol{y}_{t,i}, i = 1, \ldots, n_t$ are available at time t, with each $\boldsymbol{y}_{t,i}$ informing a functional parameter $\boldsymbol{\psi}_{t,i} = \boldsymbol{\psi}_{t,i}(\boldsymbol{X}(t), \boldsymbol{\lambda}(t)$. These relationships may be stochastic dependencies, or more usually, deterministic functions. Again assuming that the $\boldsymbol{y}_{t,i}$ for $i = 1, \ldots, n_t$ are independent conditional on their common parents, the likelihood of the data $\boldsymbol{y_t} = (\boldsymbol{y}_{t,i}, i = 1, \ldots, n_t)$ at time t is expressed as the product

$$L(\boldsymbol{y_t} \mid \boldsymbol{X}(t), \boldsymbol{\lambda}(t)) = \prod_{i=1}^{n_t} L(\boldsymbol{y}_{t,i} \mid \boldsymbol{\psi}_{t,i}(\boldsymbol{X}(t), \boldsymbol{\lambda}(t))).$$

When the data are also conditionally independent over time, the likelihood of all the data $\boldsymbol{y} = (\boldsymbol{y}_{0,1}, \ldots, \boldsymbol{y}_{t,n_t})$ given the basic parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_0, \ldots, \boldsymbol{\theta}_t, \ldots)$ is

$$L(\boldsymbol{y} \mid \boldsymbol{\theta}) = \prod_t \prod_{i=1}^{n_t} L(\boldsymbol{y}_{t,i} \mid \boldsymbol{\psi}_{t,i}(\boldsymbol{X}(t), \boldsymbol{\lambda}(t))) = \prod_t \prod_{i=1}^{n_t} L(\boldsymbol{y}_{t,i} \mid \boldsymbol{\theta}_t).$$

The posterior distribution of $\boldsymbol{\theta}$ given the data is then $p(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto L(\boldsymbol{y} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})$. Note that given the posterior distribution of the basic parameters and/or the states and transition rates, any function $\psi(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{X})$, even if not directly observed, can be derived.

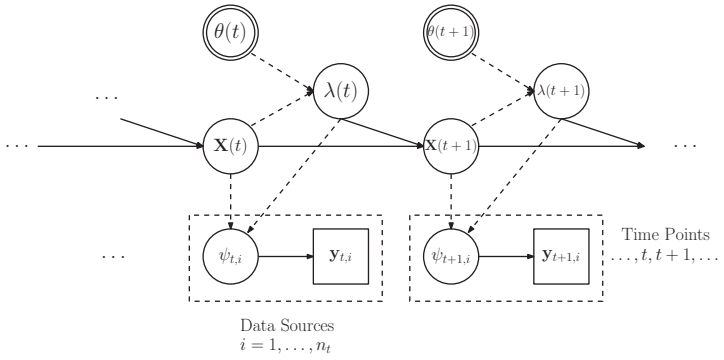

FIGURE 24.6
DAG of a multi-state model representing a transmission model embedded in an evidence synthesis.

24.4 Cross-Sectional Estimation: Severity

Static, cross-sectional models to estimate prevalence or cumulative incidence of (perhaps severe) disease are particular cases of the model in Figure 24.6. The influenza severity estimation of [15] can be seen as three cross-sectional analyses, one for each wave of infection, where severity is expressed in terms of ratios of cumulative incidence of infection at different levels of severity (Figure 24.1). The model is also stratified by age group $a \in \{<1, 1-4, 5-14, 15-24, 25-44, 45-64, 65+\}$. The three timepoints $t \in \{0, 1, 2\}$ are not completely independent, as they share some parameters.

24.4.1 Model specification

24.4.1.1 First and second waves

Figure 24.7 displays the DAG corresponding to the first wave of infection, in summer 2009 ($t = 0$ in the notation of Figure 24.6). The disease states $\boldsymbol{X}(0)$ correspond to the numbers N_l of infections at each severity level $l = \{\text{I,S,H,ICU,D}\}$ (Figure 24.1). Note that age and wave/time indices have been omitted in the DAG and what follows, for brevity.

Functional parameters Ideally, we would assume a nested binomial structure for the states $N_l \sim \text{Binomial}(N_m, p_{l|m})$ for a severity level m lower than l and for conditional probability $p_{l|m}$ of being a case at level l given infection at level m. However, for computational reasons, we instead assume a mean parameterization such that generically, the states N_l are deterministic functions

$$N_l = \lfloor p_{l|m} N_m \rfloor$$

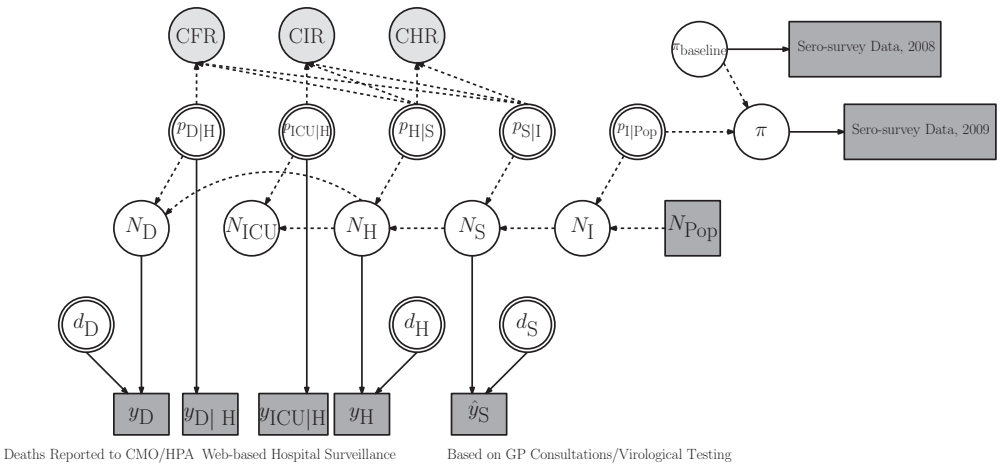

FIGURE 24.7
DAG representing the first wave of A/H1N1 pandemic influenza infection. (Adapted from Presanis et al. [15])

of the number N_m of infections at a lower level m of severity and the conditional probability $p_{l|m}$ of being a case at level l given infection at level m. The total population N_{Pop} is considered fixed. The CSRs of interest (i.e., the CHR, CIR, and CFR risks (Figure 24.1)), are products of component conditional probabilities:

$$\begin{aligned} \text{CHR} &= p_{\text{H|I}} = p_{\text{H|S}} \times p_{\text{S|I}} \\ \text{CIR} &= p_{\text{ICU|I}} = p_{\text{ICU|H}} \times \text{CHR} \\ \text{CFR} &= p_{\text{D|I}} = p_{\text{D|H}} \times \text{CHR}. \end{aligned}$$

The functional parameters $\boldsymbol{\psi}$ of Figure 24.6 are defined as the set $\boldsymbol{\psi} = \{N_l, l \in \{\text{I,S,H,ICU,D}\}\} \cup \{\text{CHR, CIR, CFR}\}$ of states and the corresponding CSRs.

Observational model The observations, $y_i, i = 1, \ldots, n$, where $n = 7$ per age group, are either reported numbers of infections at different levels of severity ($\hat{y}_{\text{S}}, y_{\text{H}}, y_{\text{D}}$ in Figure 24.7), reported numbers of hospitalizations that lead to severe events ($y_{\text{ICU|H}}, y_{\text{D|H}}$), or the cross-sectional sero-prevalence data before and after the first wave. Each observation is a realization of a binomial distribution with probability parameters representing either a detection probability d_l with size parameter the counts of the numbers of infections N_l at levels $l \in \{\text{S,H,D}\}$; a conditional probability $p_{\text{ICU|H}}$ or $p_{\text{D|H}}$ with size parameter the subset of observed hospitalizations y_{H} where the final outcome $y_{\text{ICU|H}}$ or $y_{\text{D|H}}$ or discharge was observed; or a sero-prevalence, either before the first wave (π_{baseline}) or after the first wave (π), with size parameter the number of sera samples tested. The latter sero-prevalence data inform, indirectly, the infection attack rate $p_{\text{I|Pop}}$ via their difference.

Basic parameters The basic parameters $\boldsymbol{\theta}$ of Figure 24.6 are the set $\boldsymbol{\theta} = \{p_{l|m}, l, m \in \{\text{I,S,H,ICU,D}\}\} \cup \{d_l, l \in \{\text{S,H,D}\}\}$ of conditional and detection probabilities in Figure 24.7. Each probability is given an independent flat prior, apart from the symptomatic CHR risk $p_{\text{H|S}}$ which is assigned an informative prior based on external data.

The second wave model is similar to that of the first wave excluding only the sero-prevalence data due to the challenges of disentangling vaccinated cases from true infections in the sero-samples in the absence of good data on vaccination in the dataset.

24.4.1.2 Third wave

The third wave model differs more substantially (Figure 24.8). Estimates of the number of symptomatic infections were derived from a joint Bayesian model of the GP consultation and virological positivity data from the smaller sentinel system, regressed on age group and time. This sub-model was fitted at a first stage, accounting for the probability of consulting a GP-given symptoms, and giving posterior mean (sd) estimates $\hat{y}_{\text{S}}$ ($\hat{\sigma}_{\text{S}}$) on a log scale. In contrast to the first two waves, these estimates are not considered underascertained. They are therefore incorporated in the third wave model, at a second stage, by a likelihood term, assuming $\hat{y}_{\text{S}} \sim \text{normal}(\log(N_{\text{S}}), \hat{\sigma}_{\text{S}})$. Finally, since the hospitalization data in the third wave (Figures 24.2 and 24.3) are sparse, they lead to uncertain and underascertained estimates of the numbers hospitalized and the proportion of hospitalizations leading to ICU admission. Extra prevalence-type data on the number $N_{t,\text{ILIC}}$ of suspected ILI cases present in all ICUs in the UK are therefore incorporated, through a sub-model for these data that assumes entrances at rate λ_t and exits at rate μ to/from ICU form an immigration-death stochastic process [15] (Figure 24.9). Note that this ICU sub-model is another example of a multi-state model as in Figure 24.6, where the state is $N_{t,\text{ILIC}}$. Since the observations are of ILI

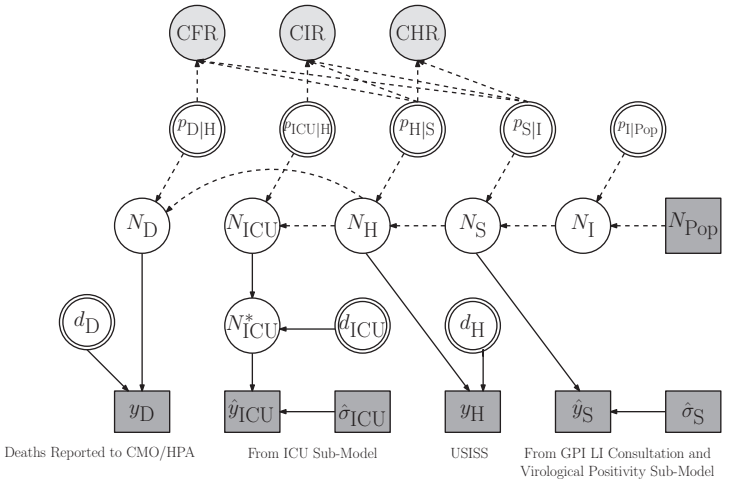

FIGURE 24.8
DAG representing the third wave of A/H1N1 pandemic influenza infection. (Adapted from Presanis et al. [15])

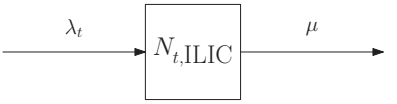

FIGURE 24.9
Immigration-death process model for ILI cases in ICU. λ_t is the daily rate of admissions to ICU, $N_{t,\text{ILIC}}$ the number of ILI cases present in ICU on day t, and μ the rate of exit (discharges or deaths), so that the expected length of stay in ICU is $1/\mu$.

rather than confirmed influenza, estimates of the ILI admission rate λ_t are combined with virological positivity data from secondary care to obtain estimates of the cumulative number y_{ICU} of new ICU admissions for A/H1N1 influenza. The posterior mean (sd) estimates $\hat{y}_{\text{ICU}}$ ($\hat{\sigma}_{\text{ICU}}$) from this sub-model, on a log scale, then are incorporated in the third wave model by a contribution to the likelihood:

$$\hat{y}_{\text{ICU}} \sim \text{Normal}(\log(N^*_{\text{ICU}}), \hat{\sigma}_{\text{ICU}}) \tag{24.1}$$

where N^*_{ICU} is considered a lower bound for the total number of A/H1N1 ICU admissions, N_{ICU}, since the prevalent ICU case data cover only a portion of the time period of the third wave. This lower bound is implemented through a binomial assumption

$$N^*_{\text{ICU}} \sim \text{Binomial}(N_{\text{ICU}}, d_{\text{ICU}})$$

with probability parameter representing a detection probability d_{ICU}.

In contrast to the first two waves, priors for the conditional probabilities $p_{t=3,l|m}$ were expressed hierarchically by centering these probabilities, on a logit scale, on their respective second wave versions, $p_{t=2,l|m}$.

24.4.2 Results

Figure 24.10 and the left-hand side of Figure 24.11 give posterior summaries of the symptomatic CFR (sCFR = $p_{D|S}$), CFR and number of symptomatic infections N_S, by age and wave. Note that for the first two waves of infection, the model giving these results assumes that the HPA-provided estimates of N_S are underestimates, with detection probability d_S.

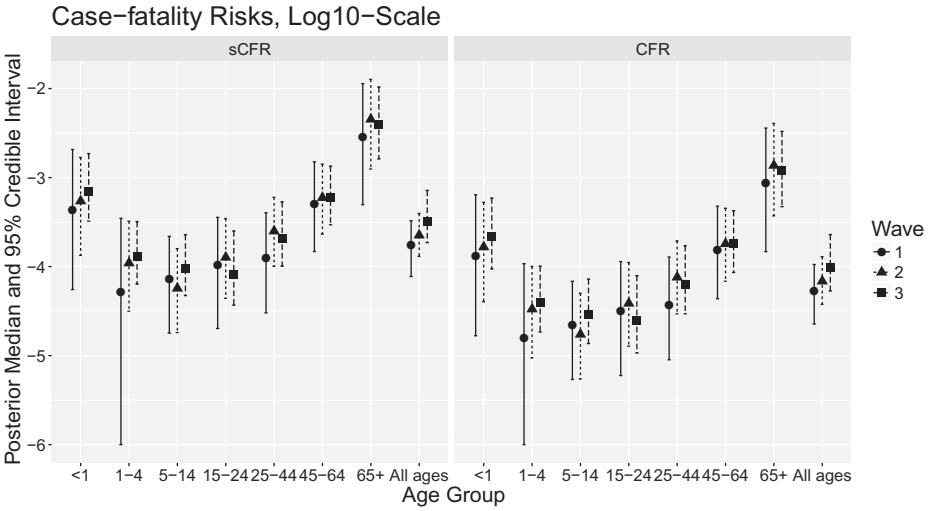

FIGURE 24.10
Posterior median and 95 percent credible intervals of symptomatic CFR and CFR by wave and age group, log-scale.

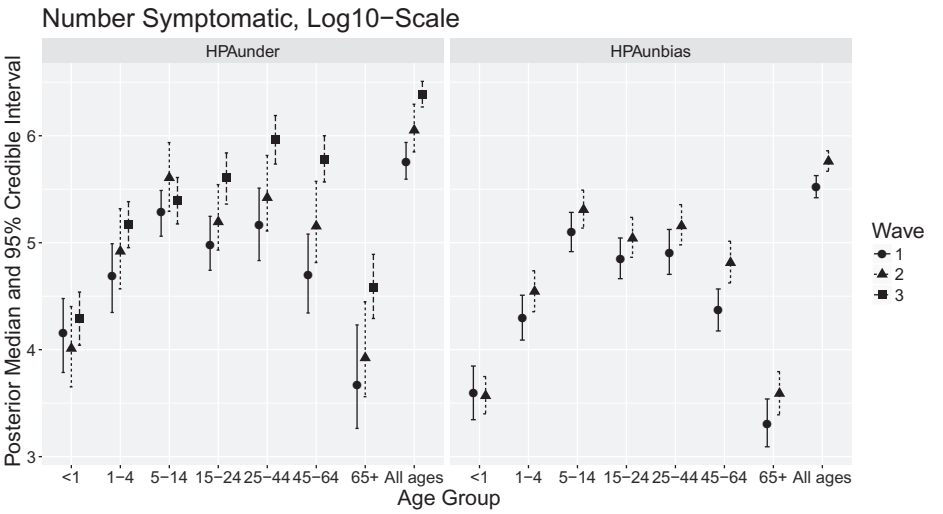

FIGURE 24.11
Left: Posterior median and 95 percent credible intervals of number of symptomatic infections by wave and age group, log-scale. Right: Same posterior summaries but for a sensitivity analysis assuming $d_S = 1$.

In [15], discussion of the assumptions about the potential under-estimation in HPA estimates of the number symptomatic led to a number of sensitivity analyses. An initial unpublished sensitivity analysis assumed the HPA estimates were unbiased, i.e., $d_S = 1$, giving the posterior estimates of N_S on the right-hand side of Figure 24.11 for waves 1 and 2 only.

Notable patterns include the "u"-shaped age distribution of the case-fatality risks and corresponding "n"-shaped age distribution of symptomatic cases; and an age shift towards older ages over the three waves of infection. The sensitivity analysis assuming $d_S = 1$ gives lower estimates of N_S (and hence higher CSRs, not shown), but with a similar age pattern.

Further sensitivity analyses, as reported in [15], suggested the key age patterns were more robust to prior choices than to the choice of bias/detection model for the HPA estimates.

24.5 Temporal Estimation: Transmission

Compartmental mechanistic models, typically used to describe the process of disease transmission (see Section 24.2 and Chapters 7 and 23 of this handbook), also can be represented as in Figure 24.6. The disease states $\boldsymbol{X}(t)$ are of the classic susceptible (S), infected (I), recovered (R) type and the transition rates $\boldsymbol{\lambda}(t)$ are functional parameters defined in terms of the current disease states. In particular, for the subset of $\boldsymbol{\lambda}(t)$ representing incidence rates, we define $\boldsymbol{\lambda}_{\text{Inc}}(t) = f(\boldsymbol{X}_{\text{I}}(t))$ where $\boldsymbol{X}_{\text{I}}(t)$ is the current size of the infected and/or infectious states.

In [20], the transmission of a novel pandemic A/H1N1 influenza strain among a fixed population stratified into A age groups is estimated through the combination of a deterministic age-structured transmission model with disease and reporting models, describing disease transmission, progression, and health-care seeking behavior of infected individuals, respectively.

24.5.1 Model specification

24.5.1.1 Transmission model

The transmission dynamics are governed by a system of differential equations of the type:

$$\begin{aligned}
\frac{dS(t,a)}{dt} &= -\lambda(t,a)S(t,a) \\
\frac{dE(t,a)}{dt} &= \lambda(t,a)S(t,a) - \frac{1}{d_L}E(t,a) \\
\frac{dI(t,a)}{dt} &= \frac{1}{d_L}E(t,a) - \frac{1}{d_I}I(t,a)
\end{aligned} \tag{24.2}$$

where $S(t,a)$, $E(t,a)$, $I(t,a)$ represent the number (or proportion) of the population of age group $a, (a = 1, \ldots, A)$ in the S (susceptible), E (exposed) and I (infectious) states at time t and d_L and d_I are the mean latent and infectious periods. Transmission is driven by the time- and age-varying rate $\lambda(t,a)$ at which susceptible individuals become infected. The system in (24.2) is evaluated using an Euler approximation at times $t_k = k\delta t, k = 0, \ldots, K$, where the choice of $\delta t = 0.5$ days is sufficiently small that the probability of more than one change of state per period is negligible. Under this discretization, at time t_k the vector

$(S_{t_k,a}, E_{t_k,a}, I_{t_k,a})$ gives the number of individuals in each state with the number of new infections in $[t_{k-1}, t_k)$ being $\Delta_{t_k,a} = \lambda_{t_{k-1},a} S_{t_{k-1},a}$, where

$$\lambda_{t_k,a} = 1 - \prod_{b=1}^{A} \left\{ \left(1 - M_{t_k}^{(a,b)} R_0(\phi)/d_I\right)^{I_{t_k,b}} \right\} \delta t. \tag{24.3}$$

Here R_0 is the basic reproduction number, the expected number of secondary infections caused by a single primary infection in a fully susceptible population, often parameterized in terms of the epidemic growth rate ϕ [43]; and the time-varying mixing matrices $\boldsymbol{M}_{t_k}$, express the pattern of transmission between age groups, with the generic entry $M_{t_k}^{(a,b)}$ being the relative rate of effective contacts between individuals of each pair of age groups (a, b) at time t_k. The quantity $1 - M_{t_k}^{(a,b)} R_0(\phi)/d_I$ gives the probability of an individual in age group a not being infected by an infectious individual in age group b in the interval $k + 1$. When raised to the power of all the infectious individuals in group b, the probability of not being infected by any individual in group b is obtained. Taking the product over all age groups gives the probability of not being infected at all. This expression for $\lambda_{t_k,a}$ is known as the Reed-Frost formulation [44]. The initial conditions of the system are determined by parameter I_0, the total number of infectious individuals across all age groups at time t_0; an assumed equilibrium distribution of infections over the age groups; and an assumption of initial exponential growth that determines the relationship between the numbers in the four disease states. The mean latent period d_L is taken as known, whereas the mean infectious period d_I is a parameter to be estimated. Therefore, the dynamics of the transmission model (24.2) depend on the *basic* parameter vector $\boldsymbol{\theta}_T = (\phi, I_0, d_I, \boldsymbol{m})$, where $\boldsymbol{m}$ parameterize the mixing matrices $\boldsymbol{M}_{t_k}$. The transmission model is represented schematically in DAG format in Figure 24.12. The dependency on age has been omitted in the DAG and in what follows for brevity.

24.5.1.2 Disease progression and health-care seeking

The newly infected individuals Δ_{t_k}, following an incubation time, develop ILI symptoms with probability p_{Sym} (Figure 24.12). With probability p_{C}, the symptomatic cases are virologically confirmed through contact tracing or hospitalization in the early phase of the epidemic; and with time-varying probability $p_{t,\text{G}}$, symptomatic patients choose to contact a primary care practitioner (GP). These processes result in the (latent) number of symptomatic cases $N_{t_s,\text{Sym}}$, confirmed cases $N_{t_u,\text{C}}$, and GP consultations $N_{t_v,\text{G}}$, which can each be expressed as a convolution of the new infections Δ_{t_k} with the distribution of the time delay between infection and the relevant health-care event. For instance, the number $N_{t_v,\text{GP}}$ of GP consultations in the interval $[t_{v-1}, t_v)$ is

$$N_{t_v,\text{GP}} = p_{Sym} p_{t_v,\text{G}} \sum_{k=0}^{v} \Delta_{t_k} f(v - k) \tag{24.4}$$

where the (discretized) delay probability mass function $f(\cdot)$ accounts for the time from infection to symptoms and the time from symptoms to GP consultation. The disease progression component of the model is specified by *basic* parameters $\boldsymbol{\theta}_D = (p_{\text{Sym}}, p_{\text{C}}, p_{t_v,\text{G}})$ (Figure 24.12) and, from equation (24.4), the quantities $N_{t_u,\text{C}}$ and $N_{t_v,\text{G}}$ are complex functions of both $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_D$.

24.5.1.3 Observational model

The goal here is to estimate the rate of infections $\lambda_{t_{k-1}}$ over time and predict the resulting burden on health-care facilities, through the estimation of the basic parameters $\boldsymbol{\theta} = \boldsymbol{\theta}_T \cup \boldsymbol{\theta}_D$

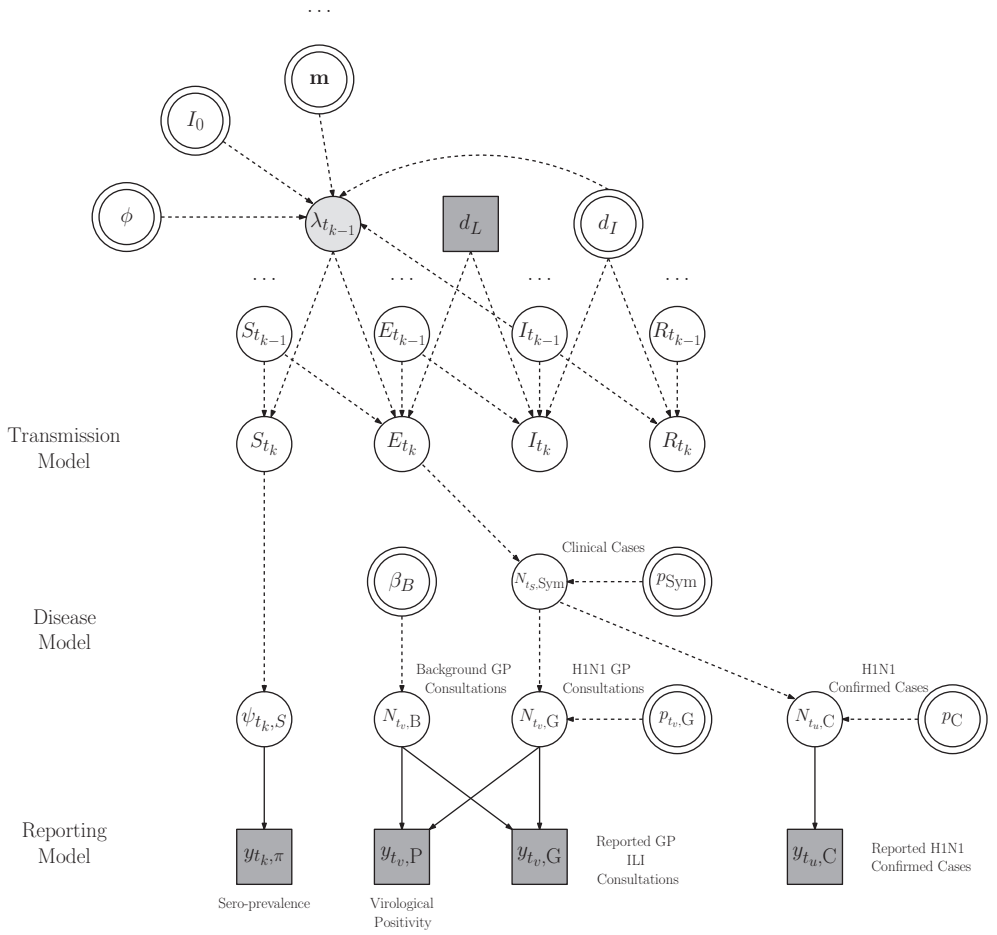

FIGURE 24.12
DAG representing the transmission model (From McDonald, S.A. et al., *BMC Infect. Dis.*, 14, 564, 2014.)

from observed data. As anticipated in Section 24.2.2, direct data on the number of new infections are not available. Therefore, a combination of a number of indirect evidence sources informing different aspects of the infection and disease processes needs to be used to estimate $\boldsymbol{\theta}$.

The observed data $\boldsymbol{y}_t = (y_{t_u,\mathrm{C}}, y_{t_v,\mathrm{G}}, y_{t_v,\mathrm{P}}, y_{t_k,\mathrm{S}})$ include $y_{t_u,\mathrm{C}}$, the counts of confirmed cases during the initial weeks of the outbreak (Figure 24.4d); $y_{t_v,\mathrm{G}}$, the number of primary care consultations for ILI, including the individuals attending for non-pandemic ILI (Figure 24.4a); $y_{t_v,\mathrm{P}}$, the complementary virological data on nasopharyngeal positivity for A/H1N1 (Figure 24.4c and Section 24.2.1); and $y_{t_k,\mathrm{S}}$, the cross-sectional sero-prevalence data (Figure 24.4b). Typically, there is some reporting delay between the disease diagnoses and their appearance in health-care surveillance, but for simplicity, here no such delay is assumed, so that $y_{t_i,i}$ is observed at the same time t_i as the disease endpoint $N_{t_i,i}$ for each i. Each item of data informs $\boldsymbol{\theta}$ through a probabilistic link.

More specifically, $y_{t_k,\mathrm{S}}$ is a realization of a binomial distribution,

$$y_{t_k,\mathrm{S}} \sim \text{Binomial}\left(n_{t_k,\mathrm{S}}, \psi_{t_k,\mathrm{S}}\right),$$

with sample size $n_{t_k,\mathrm{S}}$ and probability $\psi_{t_k,\mathrm{S}} = 1 - \frac{S_{t_k}}{N}$. The sero-prevalence data $y_{t_k,\mathrm{S}}$ therefore directly inform the number of susceptibles S_{t_k} and parameters $\boldsymbol{\theta}_T$ as $S_{t_k} = S_{t_k}(\boldsymbol{\theta}_T)$.

The counts of confirmed cases $y_{t_u,\mathrm{C}}$ and ILI consultations $y_{t_v,\mathrm{G}}$ are taken as realizations of negative binomial distributions, with means given by functional parameters $\psi_{t_u,\mathrm{C}}$ and $\psi_{t_v,\mathrm{G}}$, respectively, and time-varying over-dispersion parameter η_t,

$$y_{t_v,G} \sim \text{Negative Binomial}\left(\psi_{t_v,G}, \eta_{t_v}\right)$$

where

$$\psi_{t_u,\mathrm{C}} = N_{t_u,\mathrm{C}}$$

$$\psi_{t_v,\mathrm{G}} = N_{t_v,\mathrm{B}} + N_{t_v,\mathrm{G}}$$

and are, therefore, functions defined by convolution equations of the type in (24.4).

To disentangle the GP ILI consultations due to the pandemic strain, $N_{t_v,\mathrm{G}}$, from all other ILI consultations, $N_{t_v,\mathrm{B}}$, information is needed on the proportion of all GP ILI consultations that result from the pandemic strain. This information is provided by virological positivity data, where observed positive samples $y_{t_v,\mathrm{P}}$ are considered realizations of a binomial distribution,

$$y_{t_v,\mathrm{P}} \sim \text{Binomial}\left(n_{t_v,\mathrm{P}}, \psi_{t_v,\mathrm{P}}\right),$$

with sample size (number of tests) $n_{t_v,\mathrm{P}}$ and probability parameter $\psi_{t_v,\mathrm{P}} = 1 - \frac{N_{t_v,\mathrm{B}}}{N_{t_v,\mathrm{B}}+N_{t_v,\mathrm{G}}}$. The proportion positive, $\psi_{t_v,\mathrm{P}}$, is expressed as a function of the disentangled counts $N_{t_v,\mathrm{B}}$ and $N_{t_v,\mathrm{G}}$. As in Section 24.4.1.2, the proportion positive $\psi_{t_v,\mathrm{P}}$ and number of ILI consultations $\psi_{t_v,\mathrm{G}}$ are jointly regressed on age and time, on logit and log scales, respectively. The background counts $N_{t_v,\mathrm{B}}$ are therefore a function of the regression parameters $\boldsymbol{\beta}_B$ (Figure 24.12).

In each generic calendar time interval $[t_{j-1}, t_j]$, as indicated in Section 24.3.1, the likelihood of the data is the product

$$L(\boldsymbol{y}_{t_j} \mid \boldsymbol{\theta}) = \prod_{i \in \{\mathrm{C,G,P,S}\}} L(y_{t_j,i} \mid \boldsymbol{\theta}) \tag{24.5}$$

of the contributions $L(y_{t_j,i} \mid \boldsymbol{\theta})$ of the four data streams, as these are considered to be independent conditional on their common parents. The posterior distribution then is obtained by combining the likelihood with the prior distribution $p(\boldsymbol{\theta})$.

24.5.2 Results

Figure 24.13 shows the posterior distribution for the number of new A/H1N1 infections by age group in London, revealing the two waves of infection in summer and autumn/winter 2009. The first wave of infection has a higher peak, whereas the second wave, particularly by age group, is spread over a longer period of time, resulting in a higher attack rate in the second wave. R_0 is estimated to be 1.65 (95% credible interval 1.56–1.75).

Figure 24.14 shows predictions forward in time based on the data up to days 83 and 192 of the epidemic, respectively. Note how the uncertainty in the predictions is progressively reduced as data accumulate.

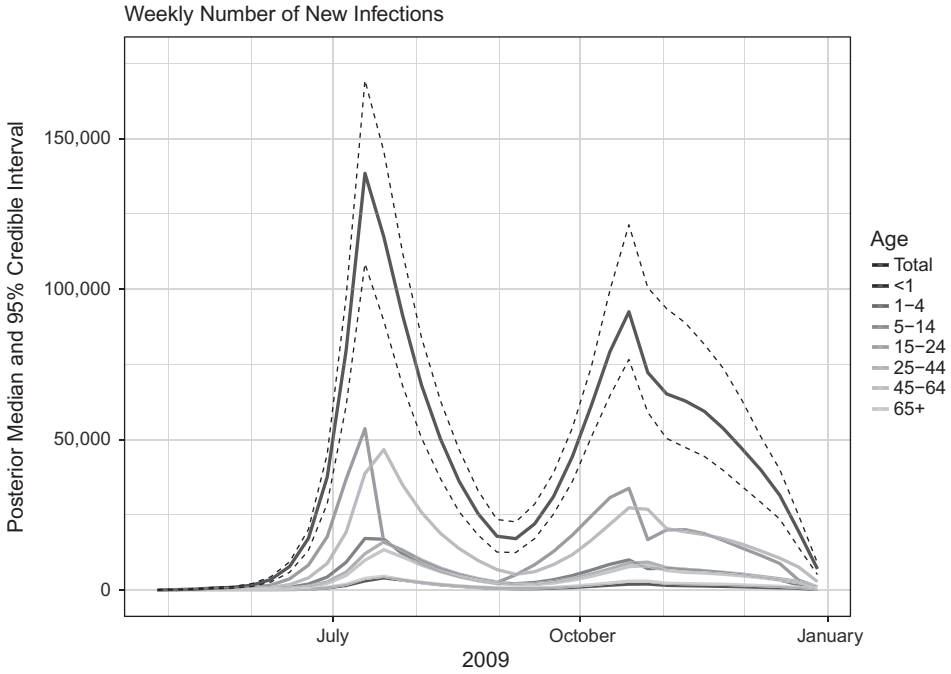

FIGURE 24.13
Posterior median (solid lines) and 95 percent credible interval (dashed) lines number of new infections per week in London overall and by age group.

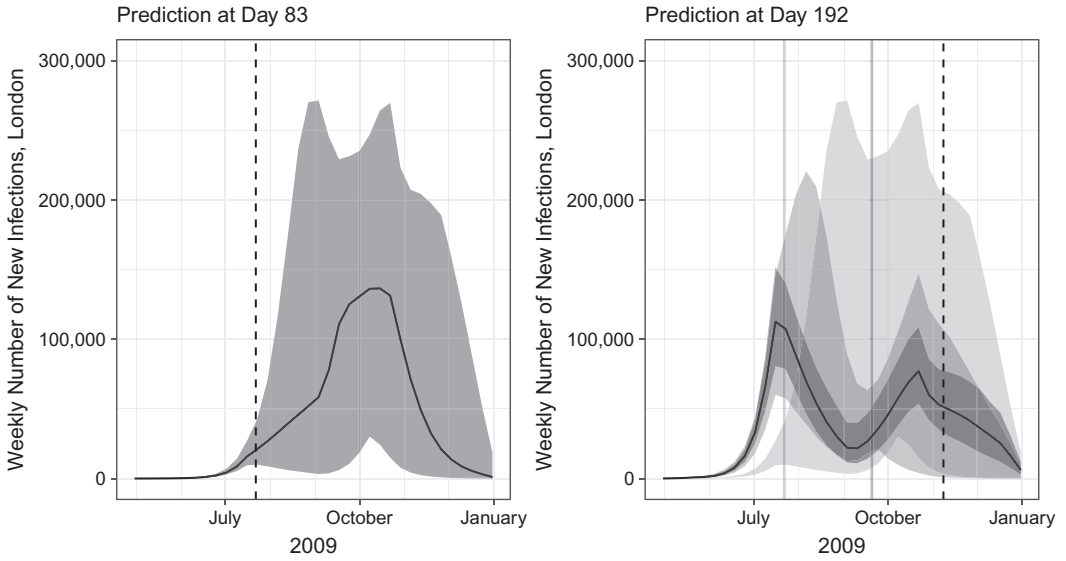

FIGURE 24.14
Predictions of the weekly number of new infections in London at days 83 and 192 of the epidemic (vertical dashed lines): posterior median (black line) and 95 percent predictive intervals (shaded areas). For the day 192 prediction, paler shaded areas represent the predictive intervals at two previous timepoints (days 83 and 143, respectively, shown by the grey vertical lines).

24.6 Model Building, Inference and Criticism

The two influenza case studies just discussed, although relatively simple, have demonstrated how the combination of multiple sources can easily lead to complex probability models. This complexity can challenge standard inferential tools, motivating the development of novel approaches. In what follows, we continue to use the two influenza examples to illustrate such new approaches to model building, efficient inference, and model criticism.

24.6.1 Strategies for model building

There are two strategies to build a complex evidence synthesis model: (1) all the data are combined simultaneously, as in joint models (e.g., [45]); or (2) the model is assembled in stages, using subsets of the evidence initially, before combining the results in a second stage, as in standard meta-analysis [4]. When a model is complex, the latter strategy is sensible, to understand what might be inferred from each dataset in isolation. A staged approach is used in both the influenza examples of Sections 24.4 and 24.5. For example, in the severity model, a sub-model of a stochastic process describing entries to/exits from ICU is first fitted to the ICU prevalent case data in the third wave, before combining the results in the full evidence synthesis model. In the transmission model, a joint regression model of GP consultation and virological positivity data was fitted initially, before the results were incorporated in the transmission model to disentangle "background" non-influenza noise from the signal of influenza consultations.

To illustrate a two-stage process, we use the severity example. Figure 24.15 shows a simplified schematic DAG of the stages. In both the stage 1 and stage 2 models, the cumulative number of ICU admissions measured over the time period of the ICU prevalence data, N_{I}^*, has a prior model. The stage 1 prior is in terms of the parameters of the ICU entry/exit process, λ_t and μ; whereas the stage 2 prior is in terms of the parameters of the period-prevalence-type severity model (i.e., the conditional and detection probabilities described in Section 24.4).

The existence of two prior models poses the question of how to combine the two sources of information. In [15], the problem was solved using an approximate method, transferring

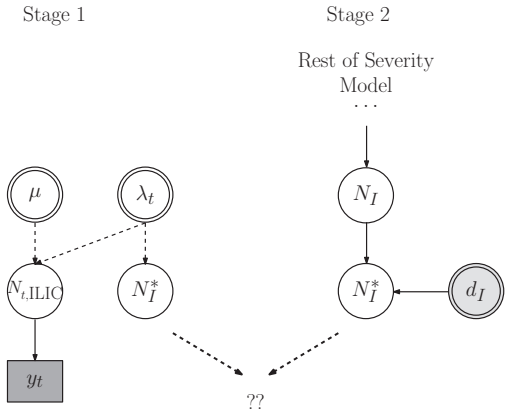

FIGURE 24.15
Two-stage modeling strategy for joining ICU sub-model with rest of severity model.

the posterior mean (sd) estimate of N_I^* from the stage 1 ICU sub-model to the stage 2 severity model by a likelihood term (equation (24.1) and Figure 24.8), such that the posterior of $\log(N_I^*)$ from the ICU sub-model is approximated by a normal distribution. This approximate approach is acceptable when the approximation is good (i.e., when the sample size in the stage 1 model is large enough to guarantee a Gaussian posterior distribution). If not, then in [46], an alternative, more general, exact method for joining (and splitting) models, "Markov melding," is proposed.

In general terms, suppose we have M probability submodels $p_m(\phi, \psi_m, Y_m)$, $m = 1, \ldots, M$, with parameters ϕ and ψ_m and observable random variables Y_m. Suppose further that ϕ is common to all modules, acting as a "link" between the submodels, and that the aim is to combine all modules into a single model $p_{\text{comb}}(\phi, \psi_1, \ldots, \psi_M, Y_1, \ldots, Y_M)$, so that the posterior distribution for the link parameter (ϕ) and the submodel-specific parameters ψ_m reflects all information and uncertainty. In some cases, such model joining is readily achievable using standard hierarchical modeling constructs. However, there are some contexts where this joining is not straightforward, in particular where (1) some sub-models may not be expressible conditional on the link parameter ϕ, particularly if ϕ is a non-invertible deterministic function of other parameters in a sub-model and (2) the prior marginal distributions $p_m(\phi)$ for ϕ differ in different submodels. Both of these situations arise in the severity estimation of Section 24.4 and Figure 24.15, where the link parameter ϕ is the cumulative number of ICU admissions N_I^*.

Markov melding [46] addresses model joining in these contexts by building on Markov combination [47–49] and Bayesian melding [50] ideas. Markov combination allows joining of sub-models under the restrictive constraint that the prior marginals $p_m(\phi)$ are identical for each m (i.e., the submodels $p_m(\phi, \psi_m, Y_m)$, $m = 1, \ldots, M$ are *consistent* in the link parameter ϕ: $p_m(\phi) = p(\phi)$ for all m). In practice, however, the marginal distributions $p_m(\phi)$ are usually *not* exactly identical, as in Figure 24.15. Markov melding therefore exploits the Bayesian melding approach [50] to replace each marginal with a pooled marginal distribution

$$p_{pool}(\phi) = g(p_1(\phi), \ldots, p_M(\phi)),$$

where g is a pooling function chosen such that $\int g(\phi)\, d\phi = 1$ and $p_{pool}(\phi)$ is an appropriate summary of the individual marginals. Since each replaced model is now consistent in ϕ, the Markov melded model can be obtained by a Markov combination of the replaced models:

$$p_{\text{MM}}(\phi, \psi_1, \ldots, \psi_M, Y_1, \ldots, Y_M) = p_{pool}(\phi) \prod_{m=1}^{M} \frac{p_m(\phi, \psi_m, Y_m)}{p_m(\phi)}.$$

Different possible pooling functions g for the melded marginal $p_{pool}(\phi)$ are discussed in [46]. After the new model p_{MM} has been formed, posterior inference given all the data $y_1, \ldots, y_M$ can be performed. Markov melding incorporates more data than any single submodel, and so will provide more precise inferences if the various components of evidence (priors and data) in each submodel do not substantially conflict. Otherwise, Markov melding may be misleading, so, before proceeding, the underlying reasons for the conflict should be investigated and resolved (see Section 24.6.3).

Note that the Markov melding method can, of course, be generalized to the case of multivariate link and submodel-specific parameters, $\boldsymbol{\phi}$ and $\boldsymbol{\psi}_m$ [46].

24.6.1.1 Markov melding in the influenza severity example

For the severity example, the use of Markov melding on the ICU and severity sub-model implied priors (Figure 24.16) results in greater posterior precision under different pooling

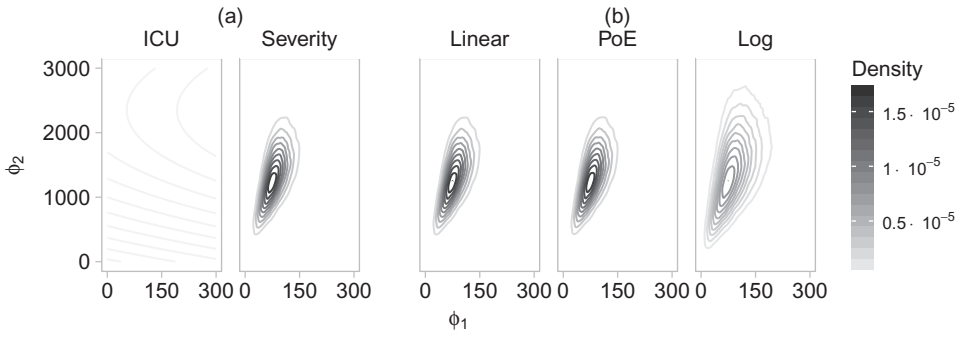

FIGURE 24.16
(a,b) Pooled prior marginal distributions under different pooling functions for the ICU and severity sub-models.

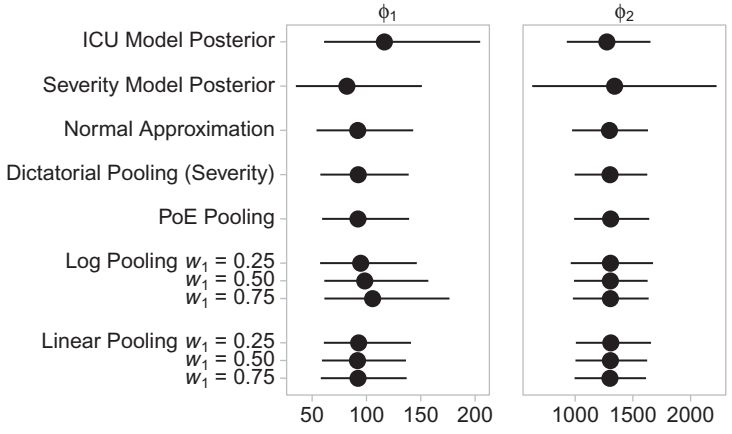

FIGURE 24.17
Posterior distributions (medians and 95% credible intervals) for the ICU and severity model parameters under the Markov melded model with different pooling functions in comparison to the separate sub-model posterior distributions.

functions, in comparison to both the sub-model posteriors alone and the normal approximation employed in [15] (Figure 24.17).

24.6.2 Computationally efficient inference

Markov chain Monte Carlo (MCMC) methods [51] have become standard tools in Bayesian inference to sample from posterior distributions. This sampling, in cross-sectional estimation problems of the type in Section 24.4 (e.g., [13]), can feasibly be carried out using available software implementing MCMC (e.g., [52]). On the other hand, inference for more complex models, such as transmission models similar to that of Section 24.5 (e.g., [21]), requires bespoke code and a tailored MCMC algorithm. However, classical MCMC often is not a computationally viable option in stochastic transmission models when the data structure is complex (e.g., [34]) and can become computationally inefficient even in inference for deterministic models, when inferences are required within a restricted time frame. Alternative

approaches to Bayesian inference are becoming popular, often combined with MCMC, to tackle such computational challenges. Examples include Approximate Bayesian Computation (ABC) (e.g., [37]), used to estimate the likelihood; Sequential Monte Carlo (SMC) (e.g., [53]), where inference is sequentially updated using only the most recent data; and emulation [54] and history matching [55], where a complex transmission model is replaced by a simpler approximate model.

As a concrete illustration, we re-visit the deterministic transmission model of Section 24.5, where the goal is to reconstruct retrospectively the evolution of the influenza A/H1N1 epidemic. In [20], an adaptive Metropolis Hasting algorithm is used to derive the posterior distribution of $\boldsymbol{\theta}$ using 245 days of GP consultation data, combined with the various additional sources of information. To reconstruct the epidemic in London, each run of the MCMC algorithm requires around 7×10^5 iterations to reach convergence, taking more than 4 hours for a single MCMC chain The bottleneck is the evaluation of the likelihood in equation (24.5), which involves calculation of the computationally expensive convolutions in equation (24.4). This time is an acceptable computational burden in the case of retrospective inference. However, during an epidemic, prospective estimation and prediction will be needed as new data arrive, for example, daily. Then the MCMC algorithm would have to be rerun each day to re-analyze the complete dataset, which would not be optimal, particularly if alternative models need to be explored. More efficiently, the posterior distribution at each new time point could be derived from the one at the previous time point. In [35], a hybrid SMC algorithm is developed to enable this more efficient use of the information and carry out inference and predictions in real time. The idea underlying SMC is to derive sequential posterior distributions $\pi_k(\boldsymbol{\theta}) = p(\boldsymbol{\theta} \mid \boldsymbol{y}_{t_1:t_k}) \propto p(\boldsymbol{\theta}) L(\boldsymbol{y}_{t_1:t_k} \mid \boldsymbol{\theta})$ for $k = 1, \ldots, K$ as data accumulate. At each time point t_k the distribution $\pi_k(\boldsymbol{\theta})$ is approximated by n_k particles $\{\boldsymbol{\theta}_k^{(1)}, \ldots, \boldsymbol{\theta}_k^{(n_k)}\}$ with corresponding weights $\{\omega_k^{(1)}, \ldots, \omega_k^{(n_k)}\}$. As data $\boldsymbol{y}_{t_{k+1}}$ arrive at t_{k+1}, $\pi_k(\boldsymbol{\theta})$ serves as an importance distribution and the updated $\pi_{k+1}(\boldsymbol{\theta})$ is obtained by re-weighting the sample $\{\boldsymbol{\theta}_k^{(1)}, \ldots, \boldsymbol{\theta}_k^{(n_k)}\}$ by the importance ratios $\frac{\pi_{k+1}\left(\boldsymbol{\theta}_k^{(j)}\right)}{\pi_k\left(\boldsymbol{\theta}_k^{(j)}\right)}$ for each $j \in 1, \ldots, n_k$.

In the model of [35] this ratio reduces to the likelihood of the new data, so that the j^{th} particle has weight

$$\omega_{k+1}^j \propto \omega_k^j \frac{\pi_{k+1}\left(\boldsymbol{\theta}_k^{(j)}\right)}{\pi_k\left(\boldsymbol{\theta}_k^{(j)}\right)} = \omega_k^j L(\boldsymbol{y}_{k+1} \mid \boldsymbol{\theta}_k^{(j)}).$$

This simple SMC scheme works well when the data follow a stable pattern, as demonstrated in settings where only one data stream is available (e.g., [56]). However, in the specific application of [35], the challenge is not particularly posed by the multiplicity of data, but rather by the sudden change in the pattern of health-seeking behavior produced by a public health intervention (see Figure 24.4a). Such a change introduces a shock to the system and complicates dramatically the tracking of the sequential distributions $\pi_k(\cdot)$ over time. On arrival of a particularly informative new batch of data $\boldsymbol{y}_{t_{k+1}}$, the sample $\{\boldsymbol{\theta}_k^{(1)}, \ldots, \boldsymbol{\theta}_k^{(n_k)}\}$ degenerates to the few particles consistent with the new information, which, carrying large weights, give a misleading estimate of $\pi_{k+1}(\cdot)$. The naive SMC algorithm then is adapted to handle these highly informative observations by introducing resampling and MCMC jittering steps [57] to rejuvenate the sample and by sequentially including only fractions of the new data to minimize the divergence between posterior distributions at consecutive times [58, 59]. The result is a hybrid semi-automatic SMC algorithm that is more computationally efficient than the original MCMC, is highly parallelizable, and can deal with sudden shocks in the observational patterns.

24.6.3 Model criticism: Conflict and influence

Model criticism is crucial to any analysis. However, specific to the context of multiple source evidence synthesis are the potential for conflicting evidence, with such conflicts needing to be detected, quantified, and resolved, and the critical assessment of what the role and influence of different sources is.

In the influenza severity example of Section 24.4, the initial sensitivity analysis shown on the right-hand side of Figure 24.11 did not include a detection probability d_S for the HPA estimates $\hat{y}_S$ of the number symptomatic (i.e., with $d_S = 1$). However, this "naive" model led to high posterior mean deviances, as shown in Table 24.1, for "data" $\hat{y}_S$ on a log scale for the first wave. By comparison, the model assuming the HPA estimates are underestimates has much lower posterior mean deviances and DIC contributions (Table 24.2). The lack of fit in the "naive" model motivated a closer look at the consistency of the different sources of evidence about the denominators, or infections at lower levels of severity (asymptomatic and symptomatic), resulting in both the sensitivity analyses of [15] and more formal conflict assessment in [60].

24.6.3.1 Conflict assessment methods

Bayesian predictive diagnostics (e.g., [61–64]) have been used for a long time in model assessment, comparing observations to predictions from the model. Posterior predictive

TABLE 24.1
Deviance Summaries for "data" on number symptomatic in first wave, log-scale, by age group, for the "naive" model: HPA estimate ($\hat{y}_S$) and corresponding standard deviation ($\hat{\sigma}_S$); posterior mean estimate (N_S) and corresponding 95 percent credible interval (CrI); posterior mean deviance contributions ($\overline{D}$); plug-in deviance at posterior mean of parameters ($D(\overline{\theta})$); effective number of parameters (p_D); deviance information criterion (DIC)

Age	$\hat{y}_S$	$\hat{\sigma}_S$	N_S	95% CrI		$\overline{D}$	$D(\overline{\theta})$	p_D	DIC
< 1	8.11	0.30	8.27	7.70	8.85	1.32	0.32	1.01	2.33
1-4	7.81	0.26	9.89	9.42	10.38	66.72	65.80	0.92	67.65
5-14	9.78	0.28	11.75	11.32	12.16	49.16	48.58	0.58	49.74
15-24	10.23	0.26	11.17	10.74	11.62	14.05	13.30	0.76	14.81
25-44	11.46	0.29	11.30	10.83	11.80	1.03	0.30	0.73	1.75
45-64	12.03	0.26	10.06	9.61	10.52	59.17	58.36	0.81	59.98
65+	11.25	0.27	7.62	7.12	8.15	175.41	174.49	0.92	176.33

TABLE 24.2
Deviance summaries for "data" on number symptomatic in first wave, log-scale, by age group, for the model assuming HPA estimates are under-estimates

Age	$\hat{y}_S$	$\hat{\sigma}_S$	N_S	95% CrI		$\overline{D}$	$D(\overline{\theta})$	p_D	DIC
< 1	8.11	0.30	8.11	7.56	8.66	0.89	0.00	0.89	1.78
1-4	7.81	0.26	9.71	9.20	10.23	0.92	0.06	0.85	1.77
5-14	9.78	0.28	11.45	10.93	11.95	0.82	0.00	0.82	1.64
15-24	10.23	0.26	11.11	10.64	11.58	1.00	0.23	0.77	1.76
25-44	11.46	0.29	11.22	10.71	11.72	0.90	0.00	0.90	1.80
45-64	12.03	0.26	9.91	9.41	10.41	0.86	0.00	0.86	1.73
65+	11.25	0.27	7.56	7.03	8.10	0.92	0.00	0.92	1.84

tests [65] are known to be conservative, due to using the data both to fit the model and to compare to model predictions, so a variety of (computationally intensive) post-processing, approximate, or cross-validatory methods have been proposed instead (e.g., [63, 64, 66–69]). Typically, each of these methods has been used to assess models of a single dataset, rather than for an evidence synthesis.

"Conflict p-values" [60, 69–71] have been proposed as a generalization of Bayesian cross-validatory predictive p-values that compare not only subsets of data to predictions resulting from the rest of the data, but also whole sub-models, comprising data, model structure, and prior information, with predictions from the rest of the model. The key idea, known as "node-splitting," is to split a DAG $\mathcal{G}(\phi, \boldsymbol{\theta}_{\setminus\phi}, \boldsymbol{y})$, comprising data $\boldsymbol{y}$ and latent quantites $\boldsymbol{\theta} = (\phi, \boldsymbol{\theta}_{\setminus\phi})$, into two independent partitions at any "separator" node ϕ, $\mathcal{G}(\phi_a, \boldsymbol{\theta}_{a\setminus\phi}, \boldsymbol{y}_a)$ and $\mathcal{G}(\phi_b, \boldsymbol{\theta}_{b\setminus\phi}, \boldsymbol{y}_b)$. Two copies of the separator ϕ are created, ϕ_a and ϕ_b, that are each identifiable in partitions $\mathcal{G}(\phi_a, \boldsymbol{\theta}_{a\setminus\phi}, \boldsymbol{y}_a)$ and $\mathcal{G}(\phi_b, \boldsymbol{\theta}_{b\setminus\phi}, \boldsymbol{y}_b)$, respectively. The aim is to compare the posterior distributions from each partition, $p(\phi_a \mid \boldsymbol{y}_a)$ and $p(\phi_b \mid \boldsymbol{y}_b)$.

This comparison is achieved by defining a difference function $\delta = h(\phi_a) - h(\phi_b)$, on an appropriate scale $h(\cdot)$, and considering where 0 lies in the posterior distribution of the difference, $p_\delta(\delta \mid \boldsymbol{y}_a, \boldsymbol{y}_b)$. A two-sided conflict p-value corresponding to the hypothesis test $H_0 : \delta = 0$ is defined as $c = 2 \times \min[\Pr\{p_\delta(\delta \mid \boldsymbol{y}_a, \boldsymbol{y}_b) < p_\delta(0 \mid \boldsymbol{y}_a, \boldsymbol{y}_b)\}, 1 - \Pr\{p_\delta(\delta \mid \boldsymbol{y}_a, \boldsymbol{y}_b) < p_\delta(0 \mid \boldsymbol{y}_a, \boldsymbol{y}_b)\}]$, with different methods for evaluating c and one-sided variations given in [60, 70]. The conflict p-value has been demonstrated to be uniform under the null model in a range of models by [71]. This setup can be generalized to multiple partitions and multiple node-splits [72].

24.6.3.2 Conflict in the influenza severity example

To assess whether the lack of fit to the HPA estimates of the number symptomatic in the "naive" model (Table 24.1) could be due to conflicting evidence, the model for the first wave only is split into two partitions at N_{S}, as in Figure 24.18. The "parent" partition comprises the data and priors informing the parent nodes of N_{S} (i.e., the sero-prevalence data and an informative prior for the proportion of infections that are symptomatic, $p_{\mathrm{S|Inf}}$). The rest

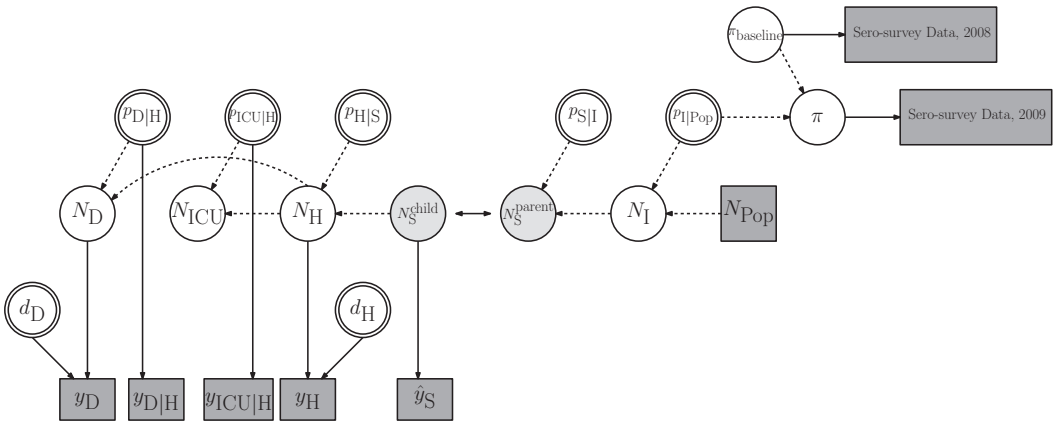

FIGURE 24.18
DAG showing node split at N_{S}. On the right is the "parent" model and on the left the "child" model. The double-headed arrow represents the comparison between the two.

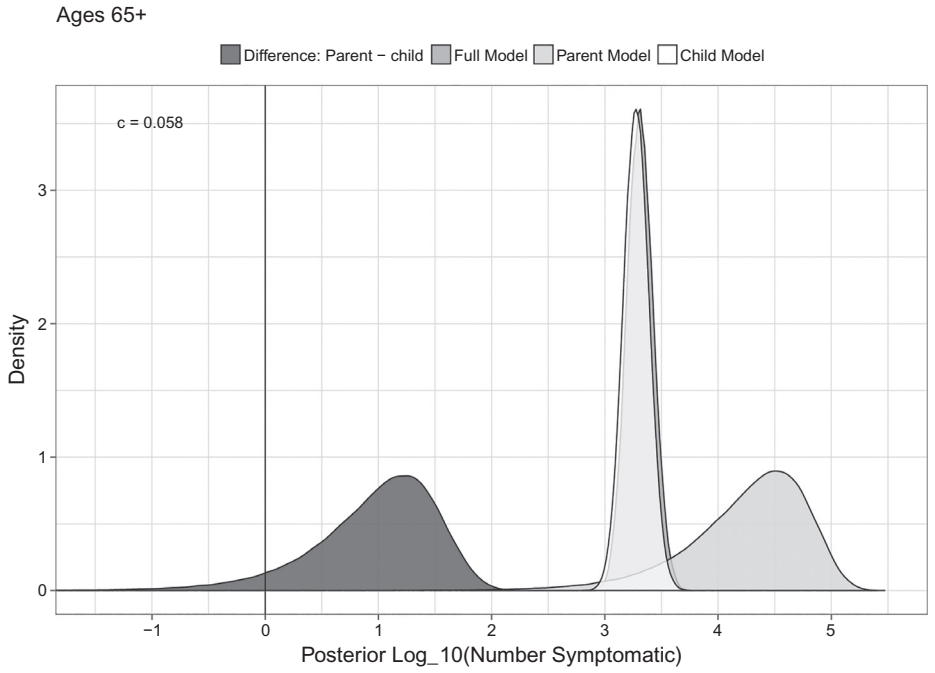

FIGURE 24.19
Conflict (posterior difference) at N_S between the parent and child models in age group 65+ in the first wave: difference function $\delta = \log_{10}(N_S^{\text{parent}}) - \log_{10}(N_S^{\text{child}})$ (dark grey); $\log_{10}(N_S)$ from full model (medium grey); $\log_{10}(N_S^{\text{parent}})$ from parent model (light grey); $\log_{10}(N_S^{\text{child}})$ from child model (ivory).

of the evidence (the severe case data and priors) comprises the "child" model. The test for conflict between the parent and child models demonstrated low posterior probabilities of no conflict, particularly for the youngest and oldest age groups [60]. The posterior difference function $\delta = \log_{10}(N_S^{\text{parent}}) - \log_{10}(N_S^{\text{child}})$ is plotted for the age group 65+ in Figure 24.19, together with the conflict p-value $c = 0.058$ and the corresponding posterior distributions for the two partitions and the full model. Note that the sero-prevalence data in the parent model imply a much higher, although also much more uncertain, number symptomatic than the severe case data in the child model. The lack of certainty in the sero-prevalence data means the severe case data and priors have more influence in the full model, so the full model posterior is much closer to the child model posterior than the parent one.

A similar investigation of conflict at a different node in the "naive" model, the number of hospitalizations N_H, leads to splitting the DAG into three partitions (not shown), based on the sero-prevalence data, the hospital data, and the mortality data, respectively. The influence of the evidence in the three partitions on the full model is shown in Figure 24.20. As in the previous example, the sero-prevalence (denominator) data in partition 1 are the most uncertain, with the hospital data in partition 2 having less uncertainty, and the mortality data in partition 3 even less uncertainty. The posterior distributions for all three partitions overlap substantially, however, so that conflict is not detectable. Instead, when the three partitions are combined in the full model, we obtain a much more precise estimate of the number hospitalized, which is a compromise between the three partitions, as would be expected.

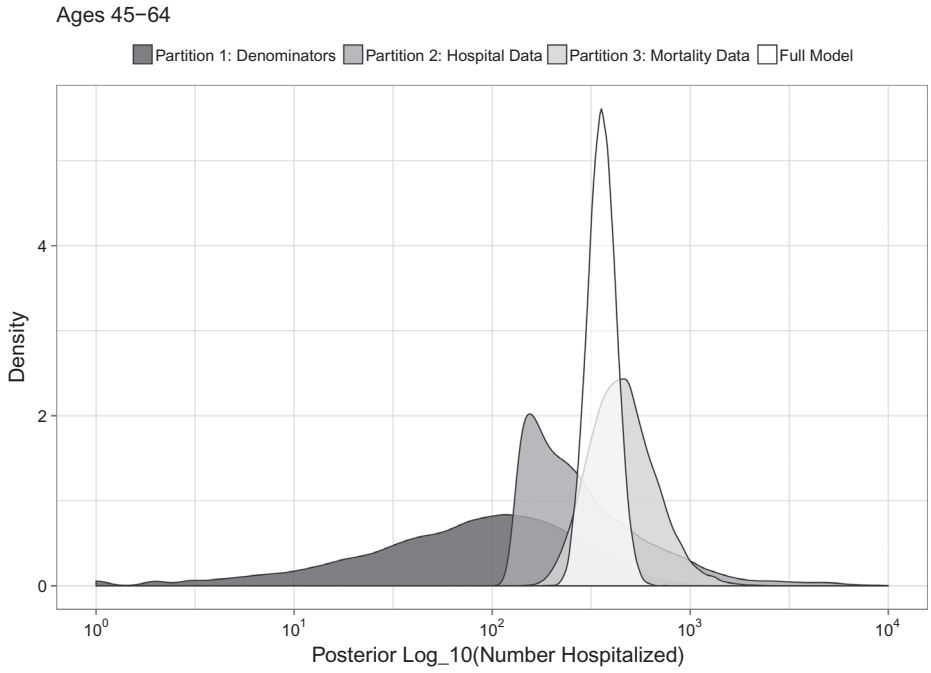

FIGURE 24.20
Influence and conflict at N_{H} between the three partitions in age group 45–64 in the first wave.

24.7 Discussion

This chapter has illustrated both the advantages and complexities of synthesizing multiple data sources to estimate various hidden characteristics of infectious disease, through the two examples of severity and transmission estimation for influenza. Section 24.6 introduced three sets of tools to approach the challenges of complex model building, computationally efficient inference and model criticism, respectively. However, these tools are a first step in resolving these challenges, with a number of questions remaining open.

Markov melding generalizes existing ideas that facilitate realistic evidence synthesis through a modular approach to model building. Some outstanding challenges include the choice of pooling function and the degree of heterogeneity between prior models that is acceptable for Markov melding to be appropriate – when does heterogeneity become conflict? Nevertheless, Markov melding is an important step in this era of big data, generalizing divide-and-conquer approaches [73].

Efficient inference for transmission models has been introduced using a sequential approach. Currently, the sequential approach is highly tailored to accommodate shocks to the system, so a question of interest is how to adapt a transmission model structure in real time to be able to generalize the SMC approach to any plausible epidemic scenario. The influenza model illustrated is deterministic in its dynamics, and the generic multi-state model description in Section 24.3.1 and Figure 24.6, while in theory accommodating stochastic dynamic transmission, is more focused on deterministic dynamics. An important area of research is to consider efficient, real-time, inference from multiple sources for stochastic epidemics,

particularly in the early stages of an outbreak [3]. The influenza transmission model of [34] includes several levels of stochasticity and data, but is highly computationally intensive, and therefore not feasible in real time. How much stochasticity is therefore necessary to realistically model an emerging outbreak?

The illustration of conflict assessment in Section 24.6.3 is targeted, in that conflict was assessed at particular nodes in the DAG of the influenza severity model, following suspected biases in particular data sources. However, in some contexts, it may not be so clear where to look for potential conflict, in which case a systematic search throughout a DAG for conflict may be warranted. However, such systematic assessment entails multiple tests, either through a multivariate difference function (as in the example of Figure 24.20) or through fitting multiple node-split models. A framework for systematic assessment, accounting for the multiple tests and their correlation, has therefore been proposed [72]. Further open questions in this area include how to improve power to detect conflict and how to make such methods more accessible by improving the computational feasibility of systematic conflict assessment. As with any cross-validatory framework, multiple node-splitting can be computationally burdensome, so for hierarchical models, [74] have proposed an integrated nested Laplace approximation (INLA) approach to fast conflict diagnostics. A final area of open research related to understanding the influence of different, potentially conflicting, evidence sources on inference is the adaptation of value of information methods to evidence synthesis [75].

Acknowledgments

The authors would like to thank Robert Goudie and Paul Birrell for help with plots. Anne Presanis and Daniela De Angelis are supported by the Medical Research Council (Unit program number MC_UU_00002/11).

References

[1] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford, UK, 1991.

[2] H. Heesterbeek, R. M. Anderson, V. Andreasen, S. Bansal, D. De Angelis, C. Dye, K. T. D. Eames et al. Modeling infectious disease dynamics in the complex landscape of global health. *Science*, 347(6227):aaa4339, 2015.

[3] P. J. Birrell, D. De Angelis, and A. M. Presanis. Evidence synthesis for stochastic epidemic models. *Statistical Science*, 33(1):34–43, 2018.

[4] M. Borenstein, L. V. Hedges, J. Higgins, and H. R. Rothstein. *Introduction to Meta-Analysis*. Wiley Online Library, Chichester, UK, 2009.

[5] D. M. Eddy, V. Hasselblad, and R. Shachter. *Meta-Analysis by the Confidence Profile Method*. Academic Press, London, UK, 1992.

[6] D. J. Spiegelhalter, K. R. Abrams, and J. P. Myles. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Statistics in Practice. Wiley, Chichester, UK, 2004.

[7] N. J. Welton, A. J. Sutton, N. J. Cooper, K. R. Abrams, and A. E. Ades. Evidence synthesis in a decision modelling framework. In *Evidence Synthesis for Decision Making in Healthcare*, pages 138–150. John Wiley & Sons, Hoboken, NJ, 2012.

[8] A. E. Ades and A. J. Sutton. Multiparameter evidence synthesis in epidemiology and medical decision-making: Current approaches. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(1):5–35, 2006.

[9] M. C. Wheldon, A. E. Raftery, S. J. Clark, and P. Gerland. Bayesian reconstruction of two-sex populations by age: Estimating sex ratios at birth and sex ratios of mortality. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(4):977–1007, 2015.

[10] J. S. Clark, D. Nemergut, B. Seyednasrollah, P. J. Turner, and S. Zhang. Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious data. *Ecological Monographs*, 87(1):34–56, 2017.

[11] D. De Angelis, A. M. Presanis, S. Conti, and A. E. Ades. Estimation of HIV burden through Bayesian evidence synthesis. *Statistical Science*, 29(1):9–17, 2014.

[12] R. J. Harris, M. Ramsay, V. D. Hope, L. Brant, M. Hickman, G. R. Foster, and D. De Angelis. Hepatitis C prevalence in England remains low and varies by ethnicity: An updated evidence synthesis. *European Journal of Public Health*, 22(2):187–192, 2011.

[13] S. A. McDonald, R. Mohamed, M. Dahlui, H. Naning, and A. Kamarulzaman. Bridging the data gaps in the epidemiology of hepatitis C virus infection in Malaysia using multiparameter evidence synthesis. *BMC Infectious Diseases*, 14:564, 2014.

[14] I. Albert, E. Espié, H. De Valk, and J. B. Denis. A Bayesian evidence synthesis for estimating campylobacteriosis prevalence. *Risk Analysis*, 31(7):1141–1155, 2011.

[15] A. M. Presanis, R. G. Pebody, P. J. Birrell, B. D. M. Tom, H. K. Green, H. Durnall, D. Fleming, and D. De Angelis. Synthesising evidence to estimate pandemic (2009) A/H1N1 influenza severity in 2009–2011. *Annals of Applied Statistics*, 8(4):2378–2403, 2014.

[16] M. Shubin, M. Virtanen, S. Toikkanen, O. Lyytikäinen, and K. Auranen. Estimating the burden of A(H1N1) pdm09 influenza in Finland during two seasons. *Epidemiology and Infection*, 142(5):964–974, 2014.

[17] N. J. Welton and A. E. Ades. A model of toxoplasmosis incidence in the UK: Evidence synthesis and consistency of evidence. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 54(2):385–404, 2005.

[18] S. A. McDonald, A. M. Presanis, D. De Angelis, W. Van der Hoek, M. Hooiveld, G. Donker, and M. E. Kretzschmar. An evidence synthesis approach to estimating the incidence of seasonal influenza in the Netherlands. *Influenza and other Respiratory Viruses*, 8(1):33–41, 2014.

[19] S. A. Mcdonald, P. Teunis, N. van der Maas, S. de Greeff, H. de Melker, and M. E. Kretzschmar. An evidence synthesis approach to estimating the incidence of symptomatic pertussis infection in the Netherlands, 2005-2011. *BMC Infectious Diseases*, 15:588, 2015.

[20] P. J. Birrell, G. Ketsetzis, N. J. Gay, B. S. Cooper, A. M. Presanis, R. J. Harris, A. Charlett et al. Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. *Proceedings of the National Academy of Sciences*, 108(45):18238–18243, 2011.

[21] I. Dorigatti, S. Cauchemez, and N. M. Ferguson. Increased transmissibility explains the third wave of infection by the 2009 H1N1 pandemic virus in England. *Proceedings of the National Academy of Sciences*, 110(33):13422–13427, 2013.

[22] D. De Angelis, A. M. Presanis, P. J. Birrell, G. S. Tomba, and T. House. Four key challenges in infectious disease modelling using data from multiple sources. *Epidemics*, 10:83–87, 2015.

[23] M. Lipsitch, S. Riley, S. Cauchemez, A. C. Ghani, and N. M. Ferguson. Managing and reducing uncertainty in an emerging influenza pandemic. *New England Journal of Medicine*, 361(2):112–115, 2009.

[24] A. J. Kucharski, S. Funk, R. M. Eggo, H.-P. Mallet, W. J. Edmunds, and E. J. Nilles. Transmission dynamics of Zika virus in island populations: A modelling analysis of the 2013–14 French Polynesia outbreak. *PLoS Neglected Tropical Diseases*, 10(5):e0004726, 2016.

[25] A. Camacho, R. M. Eggo, S. Funk, C. H. Watson, A. J. Kucharski, and W. J. Edmunds. Estimating the probability of demonstrating vaccine efficacy in the declining Ebola epidemic: A Bayesian modelling approach. *BMJ Open*, 5(12):e009346, 2015.

[26] C. Reed, F. J. Angulo, D. L. Swerdlow, M. Lipsitch, M. I. Meltzer, D. Jernigan, and L. Finelli. Estimates of the prevalence of pandemic (H1N1) 2009, United States, April-July 2009. *Emerging Infectious Diseases*, 15(12):2004–2007, 2009.

[27] S. S. Shrestha, D. L. Swerdlow, R. H. Borse, V. S. Prabhu, L. Finelli, C. Y. Atkins, K. Owusu-Edusei et al. Estimating the burden of 2009 pandemic influenza a (H1N1) in the United States (April 2009-April 2010). *Clinical Infectious Diseases*, 52(S1):75–82, 2011.

[28] C. Reed, S. S. Chaves, P. D. Kirley, R. Emerson, D. Aragon, E. B. Hancock, L. Butler et al. Estimating influenza disease burden from population-based surveillance data in the United States. *PLoS One*, 10(3):e0118369, 2015.

[29] A. M. Presanis, D. De Angelis, The New York City Swine Flu Investigation Team, A. Hagy, C. Reed, S. Riley, B. S. Cooper, L. Finelli, P. Biedrzycki, and M. Lipsitch. The severity of pandemic H1N1 influenza in the United States, from April to July 2009: A Bayesian analysis. *PLoS Medicine*, 6(12):e1000207+, 2009.

[30] A. M. Presanis, R. G. Pebody, B. J. Paterson, B. D. M. Tom, P. J. Birrell, A. Charlett, M. Lipsitch, and D. De Angelis. Changes in severity of 2009 pandemic A/H1N1 influenza in England: A Bayesian evidence synthesis. *BMJ*, 343(7824):d5408+, 2011.

[31] S. Cauchemez, A.-J. Valleron, P.-Y. Boelle, A. Flahault, and N. M. Ferguson. Estimating the impact of school closure on influenza transmission from sentinel data. *Nature*, 452(7188):750, 2008.

[32] M. Baguelin, S. Flasche, A. Camacho, N. Demiris, E. Miller, and W. J. Edmunds. Assessing optimal target populations for influenza vaccination programmes: An evidence synthesis and modelling study. *PLoS Medicine*, 10(10):e1001527+, 2013.

[33] D. E. Te Beest, P. J. Birrell, J. Wallinga, D. De Angelis, and M. van Boven. Joint modelling of serological and hospitalization data reveals that high levels of pre-existing immunity and school holidays shaped the influenza A pandemic of 2009 in the Netherlands. *Journal of the Royal Society, Interface / the Royal Society*, 12(103):20141244+, 2014.

[34] M. Shubin, A. Lebedev, O. Lyytikäinen, and K. Auranen. Revealing the true incidence of pandemic A(H1N1) pdm09 influenza in Finland during the first two seasons An analysis based on a dynamic transmission model. *PLoS Computational Biology*, 12(3):e1004803, 2016.

[35] P. J. Birrell, L. Wernisch, B. D. M. Tom, L. Held, G. O. Roberts, R. G. Pebody, and D. De Angelis. Efficient real-time monitoring of an emerging influenza epidemic: How feasible? To appear in *The Annals of Applied Statistics*, 2019.

[36] R. J. F. Ypma, A. M. A. Bataille, A. Stegeman, G. Koch, J. Wallinga, and W. M. van Ballegooijen. Unravelling transmission trees of infectious diseases by combining genetic and epidemiological data. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728):444–450, 2012.

[37] O. Ratmann, G. Donker, A. Meijer, C. Fraser, and K. Koelle. Phylodynamic inference and model assessment with approximate Bayesian computation: Influenza as a case study. *PLoS Computational Biology*, 8(12):e1002835, 2012.

[38] T. Jombart, A. Cori, X. Didelot, S. Cauchemez, C. Fraser, and N. Ferguson. Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. *PLoS Computational Biology*, 10(1):e1003457, 2014.

[39] E. Brooks-Pollock, N. Tilston, W. J. Edmunds, and K. T. D. Eames. Using an online survey of healthcare-seeking behaviour to estimate the magnitude and severity of the 2009 H1N1v influenza epidemic in England. *BMC Infectious Diseases*, 11:1–8, 2011.

[40] S. L. Lauritzen. *Graphical Models*. Clarendon Press, Oxford, UK, 1996.

[41] P. K. Andersen and N. Keiding. Multi-state models for event history analysis. *Statistical Methods in Medical Research*, 11(2):91–115, 2002.

[42] C. H. Jackson, M. Jit, L. D. Sharples, and D. De Angelis. Calibration of complex models through Bayesian evidence synthesis: A demonstration and tutorial. *Medical Decision Making*, 35(2):148–161, 2015.

[43] H. J. Wearing, P. Rohani, and M. J. Keeling. Appropriate models for the management of infectious diseases. *PLoS Medicine*, 2(7):e174, 2005.

[44] F. Ball. A threshold theorem for the Reed-Frost chain-binomial epidemic. *Journal of Applied Probability*, 20(1):153–157, 1983.

[45] D. Rizopoulos. *Joint Models for Longitudinal and Time-to-Event Data: with Applications in R*. CRC Press, Boca Raton, FL, 2012.

[46] R. J. B. Goudie, A. M. Presanis, D. Lunn, D. De Angelis, and L. Wernisch. Joining and splitting models with Markov melding. *Bayesian Analysis*, 14(1):81–109, 2019. doi:10.1214/18/BA1104.

[47] A. P. Dawid and S. L. Lauritzen. Hyper Markov laws in the statistical analysis of decomposable graphical models. *Annals of Statistics*, 21(3):1272–1317, 1993.

[48] M. S. Massa and S. L. Lauritzen. Combining statistical models. In M. A. G. Viana and H. P. Wynn, eds., *Contemporary Mathematics: Algebraic Methods in Statistics and Probability II*, pages 239–260. American Mathematical Society, Providence, RI, 2010. https://bookstore.ams.org/conm-516.

[49] M. S. Massa and E. Riccomagno. Algebraic representations of Gaussian Markov combinations. *Bernoulli*, 23(1):626–644, 2017.

[50] D. Poole and A. E. Raftery. Inference for deterministic simulation models: The Bayesian melding approach. *Journal of the American Statistical Association*, 95(452):1244–1255, 2000.

[51] D. Gamerman and H. F. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. CRC Press, Boca Raton, FL, 2006.

[52] D. Lunn, D. J. Spiegelhalter, A. Thomas, and N. Best. The BUGS project: Evolution, critique and future directions. *Statistics in Medicine*, 28(25):3049–3067, 2009.

[53] D. M. Sheinson, J. Niemi, and W. Meiring. Comparison of the performance of particle filter algorithms applied to tracking of a disease epidemic. *Mathematical Biosciences*, 255:21–32, 2014.

[54] M. Farah, P. J. Birrell, S. Conti, and D. De Angelis. Bayesian emulation and calibration of a dynamic epidemic model for A/H1N1 influenza. *Journal of the American Statistical Association*, 109(508):1398–1411, 2014.

[55] I. Andrianakis, N. McCreesh, I. Vernon, T. J. McKinley, J. E. Oakley, R. N. Nsubuga, M. Goldstein, and R. G. White. Efficient history matching of a high dimensional individual-based HIV transmission model. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):694–719, 2017.

[56] J. B. S. Ong, I. Mark, C. Chen, A. R. Cook, H. Chyi. Lee, V. J. Lee, R. T. P. Lin, P. A. Tambyah, and L. G. Goh. Real-time epidemic monitoring and forecasting of H1N1-2009 using influenza-like illness from general practice and family doctor clinics in Singapore. *PloS One*, 5(4):e10036, 2010.

[57] W. R. Gilks and C. Berzuini. Following a moving target: Monte Carlo inference for dynamic Bayesian models. *Journal of the Royal Statistical Society: Series B* (*Statistical Methodology*), 63(1):127–146, 2001.

[58] P. Del Moral, A. Doucet, and A. Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436, 2006.

[59] R. M. Neal. Sampling from multimodal distributions using tempered transitions. *Statistics and Computing*, 6(4):353–366, 1996.

[60] A. M. Presanis, D. Ohlssen, D. J. Spiegelhalter, and D. De Angelis. Conflict diagnostics in directed acyclic graphs, with applications in Bayesian evidence synthesis. *Statistical Science*, 28(3):376–397, 2013.

[61] G. E. P. Box. Sampling and Bayes' inference in scientific modelling and robustness. *Journal of the Royal Statistical Society. Series A* (*General*), 143(4):383–430, 1980.

[62] D. B. Rubin. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, 12(4):1151–1172, 1984.

[63] A. Gelman, X.-L. Meng, and H. Stern. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 6:733–807, 1996.

[64] M. J. Bayarri and M. E. Castellanos. Bayesian checking of the second levels of hierarchical models. *Statistical Science*, 22(3):322–343, 2007.

[65] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Texts in Statistical Science. Chapman and Hall/CRC Press London, UK, 2nd ed., 2003.

[66] M. J. Bayarri and J. O. Berger. P-values for composite null models. *Journal of the American Statistical Association*, 95(452):1127–1142, 2000.

[67] N. L. Hjort, F. A. Dahl, and G. H. Steinbakk. Post-processing posterior predictive p-values. *Journal of the American Statistical Association*, 101(475):1157–1174, 2006.

[68] G. H. Steinbakk and G. O. Storvik. Posterior predictive p-values in Bayesian hierarchical models. *Scandinavian Journal of Statistics*, 36(2):320–336, 2009.

[69] E. C. Marshall and D. J. Spiegelhalter. Identifying outliers in Bayesian hierarchical models: A simulation-based approach. *Bayesian Analysis*, 2:409–444, 2007.

[70] J. Gåsemyr and B. Natvig. Extensions of a conflict measure of inconsistencies in Bayesian hierarchical models. *Scandinavian Journal of Statistics*, 36(4):822–838, 2009.

[71] J. Gåsemyr. Uniformity of node level conflict measures in Bayesian hierarchical models based on directed acyclic graphs. *Scandinavian Journal of Statistics*, 43(1):20–34, 2016.

[72] A. M. Presanis, D. Ohlssen, K. Cui, M.a Rosinska, and D. De Angelis. Conflict diagnostics for evidence synthesis in a multiple testing framework. *arXiv preprint arXiv:1702.07304*, 2017.

[73] R. Bardenet, A. Doucet, and C. Holmes. On Markov chain Monte Carlo methods for tall data. *The Journal of Machine Learning Research*, 18(1):1515–1557, 2017.

[74] E. Ferkingstad, L. Held, and H. Rue. Fast and accurate Bayesian model criticism and conflict diagnostics using R-INLA. *Stat*, 6(1):331–344, 2017.

[75] C. H. Jackson, A. M. Presanis, S. Conti, and D. De Angelis. Value of information: Sensitivity analysis and research design in Bayesian evidence synthesis. *Journal of the American Statistical Association*, 2019. doi:10.1080/01621459.2018.1562932.