Biostatistics (2015), **0**, 0, *pp.* 1–12
doi:10.1093/biostatistics/nplcm˙20151028-supplimentary

Supplementary Materials to Nested Partially-Latent Class Models for Dependent Binary Data; Estimating Disease Etiology

ZHENKE WU*[,1], MARIA DELORIA-KNOLL[2], SCOTT L. ZEGER[1]

[1] *Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21205*

[2] *Department of International Health, Johns Hopkins University, Baltimore, MD 21205*

zhwu@jhu.edu

Appendix A. Mean and Covariance Structure

In this section, we present and discuss formulas for the model-based marginal observation rates and pairwise log odds ratios among cases and controls. They can be readily modified to accommodate "other" causes as discussed in Section 2.3 of the main text.

Appendix A.1 Marginal Observation Rate

The marginal observation rates are given by

$$\mathbb{P}(M_{i'j} = 1 \mid Y_{i'} = 1) = \pi_j \sum_{k=1}^{K} \theta_k^{(j)} \eta_k + (1 - \pi_j) \left\{ \sum_{k=1}^{K} \psi_k^{(j)} \eta_k \right\}, \tag{A1}$$

$$\mathbb{P}(M_{ij} = 1 \mid Y_i = 0) = \sum_{k=1}^{K} \psi_k^{(j)} \nu_k. \tag{A2}$$

In the context of childhood pneumonia problem, Equation (A1) indicates that the observed rate of pathogen j among cases comprises of two parts: cases whose disease is caused by pathogen

*To whom correspondence should be addressed.

j for which the observation is a true positive event, and those whose disease is caused by another pathogen for which the observation is a false positive.

The case and control mean observation rates for pathogen j are equal (i.e., non-interference submodel; Section 2.3 of main text) when either of Condition (I) or (II) below holds.

$$\text{(I) } \psi_1^{(j)} = \cdots = \psi_K^{(j)} = \psi^{(j)} \text{ and } \sum_{k=1}^{K} \theta_k^{(j)} \eta_k = \psi^{(j)};$$

$$\text{(II) } \boldsymbol{\eta} = \boldsymbol{\nu}, \text{ and } \sum_{k=1}^{K} \left[\theta_k^{(j)} \eta_k - \psi_k^{(j)} \nu_k \right] = 0.$$

The first part of Condition (I) says that the binary response on dimension j is constant across subclasses among controls, which implies independence of j-th dimension's measurement to other dimensions. The second part says, within the jth disease class, the marginal observation rate of dimension j equals the control rate.

Condition (II) means case and control subclass weights are equal and the observation rate in the jth case class equals that in the controls.

Appendix A.2 Marginal Pairwise Log Odds Ratios

The marginal pairwise log odds ratio $\omega_{j\ell}$ for pathogen pair (j, ℓ) among cases is given by:

$$\begin{aligned}
\omega_{j\ell} &= \log \left\{ \frac{\mathbb{P}(M_{ij}=1, M_{i\ell}=1)\mathbb{P}(M_{ij}=0, M_{i\ell}=0)}{\mathbb{P}(M_{ij}=1, M_{i\ell}=0)\mathbb{P}(M_{ij}=0, M_{i\ell}=1)} \right\} \\
&= \log \left(\sum_{c=1}^{J} \pi_c \left[\sum_{k=1}^{K} \left\{ \theta_k^{(j)} \right\}^{\mathbf{1}\{c=j\}} \left\{ \psi_k^{(j)} \right\}^{\mathbf{1}\{c\neq j\}} \left\{ \theta_k^{(\ell)} \right\}^{\mathbf{1}\{c=\ell\}} \left\{ \psi_k^{(\ell)} \right\}^{\mathbf{1}\{c\neq \ell\}} \eta_k \right] \right) \\
&\quad - \log \left(\sum_{c=1}^{J} \pi_c \left[\sum_{k=1}^{K} \left\{ 1-\theta_k^{(j)} \right\}^{\mathbf{1}\{c=j\}} \left\{ 1-\psi_k^{(j)} \right\}^{\mathbf{1}\{c\neq j\}} \left\{ \theta_k^{(\ell)} \right\}^{\mathbf{1}\{c=\ell\}} \left\{ \psi_k^{(\ell)} \right\}^{\mathbf{1}\{c\neq \ell\}} \eta_k \right] \right) \\
&\quad + \log \left(\sum_{c=1}^{J} \pi_c \left[\sum_{k=1}^{K} \left\{ 1-\theta_k^{(j)} \right\}^{\mathbf{1}\{c=j\}} \left\{ 1-\psi_k^{(j)} \right\}^{\mathbf{1}\{c\neq j\}} \left\{ 1-\theta_k^{(\ell)} \right\}^{\mathbf{1}\{c=\ell\}} \left\{ 1-\psi_k^{(\ell)} \right\}^{\mathbf{1}\{c\neq \ell\}} \eta_k \right] \right) \\
&\quad - \log \left(\sum_{c=1}^{J} \pi_c \left[\sum_{k=1}^{K} \left\{ \theta_k^{(j)} \right\}^{\mathbf{1}\{c=j\}} \left\{ \psi_k^{(j)} \right\}^{\mathbf{1}\{c\neq j\}} \left\{ 1-\theta_k^{(\ell)} \right\}^{\mathbf{1}\{c=\ell\}} \left\{ 1-\psi_k^{(\ell)} \right\}^{\mathbf{1}\{c\neq \ell\}} \eta_k \right] \right). \quad \text{(A3)}
\end{aligned}$$

Setting $K = 1$ in the formula gives log odds ratios for a locally independent model (Wu *and*

others, 2015). When $K > 1$, suppose nearly all of pneumonia is caused by pathogen j: $\pi_j \approx 1$, we calculate $\omega_{j\ell}$ under two scenarios:

a) If the true positive rates for pathogen j across subclasses, i.e. $\theta_k^{(j)}, k = 1, ..., K$, are equal, then $\omega_{j\ell} \approx 0$, that is, we have approximate marginal independence between measurements on the jth pathogen and the rest among the cases;

b) If the number of subclasses $K = 2$ and true positive rates $\theta_k^{(j)}, k = 1, 2$ are very different, say, 1 versus 0 as an extreme example, we can show that $\omega_{j\ell} = \mathsf{logit}(\psi_1^{(\ell)}) - \mathsf{logit}(\psi_2^{(\ell)})$, which means the pairwise log odds ratio between pathogen j and ℓ among cases is determined by the variation of control subclass FPRs for the ℓth pathogen.

Appendix B. Stick-Breaking Prior

This section briefly discusses the stick-breaking priors used in the Bayesian inference for the nested partially-latent class models. A stick-breaking mixture model in theory has countably infinite number of subclasses. However, because the ν_k and η_k decrease exponentially quickly in k, *a priori*, we expect that only a small number of subclasses will be used to model the data. The expected number of subclasses from a stick-breaking prior is logarithmic in the number of observations (Hjort *and others*, 2010). This is different than a finite mixture model, which uses a fixed number of clusters to model the data. In the stick-breaking mixture model, the actual number of clusters used to model data is not fixed, and can be automatically inferred from data using the usual Bayesian posterior inference framework (Neal, 2000).

Equations (2.12)-(2.14) in the text place exchangeable prior weight on the subclasses. Following Ishwaran and James (2002), in our computations, we truncate the infinite sum to the first K^* terms with K^* sufficiently large to balance computing speed and approximating performance of the model. In our simulations and data application $K^* = 10$ is usually deemed adequate. Most subclass measurement profiles are not assigned with meaningful weights either in the simula-

tions or in data application, so that a small number of effective subclasses are usually sufficient for approximation. Also, by placing hyperpriors on stick-breaking parameters α_0 and α_1 as in Equation (2.14) in the text, we can let the data inform us about the desired sparsity level for approximating the probability contingency tables for the control and each disease class. A small value of the estimate $\widehat{\alpha}_0$ ($\widehat{\alpha}_1$) suggests that only a small number of subclasses are necessary for the controls (cases). We have chosen hyperparameters in the Gamma hyperpriors for α_1 and α_1 to be $(0.25, 0.25)$ which gives good parameter estimation performance in simulations.

Appendix C. Gibbs Sampler Algorithm

We propose the following MCMC sampling steps, assuming the truncation level is $K^* = K$:

1. Update the class indicator $I_{i'}$ for cases $i' = 1, ..., n_1$, from a multinomial distribution with probabilities

$$\begin{aligned}\mathbb{P}(I_{i'} = j \mid \cdots) = p_{i'}^{(j)} &\propto [\mathbf{M}_{i'} \mid Z_{i'}, \boldsymbol{\Theta}, \boldsymbol{\Psi}, I_{i'} = j][Z_{i'} \mid \boldsymbol{\eta}, I_{i'} = j][I_{i'} = j \mid \boldsymbol{\pi}] \\ &\propto \left\{\theta_{Z_{i'}}^{(j)}\right\}^{M_{i'j}} \left\{1 - \theta_{Z_{i'}}^{(j)}\right\}^{1-M_{i'j}} \prod_{l \neq j} \left\{\psi_{Z_{i'}}^{(l)}\right\}^{M_{i'l}} \left\{1 - \psi_{Z_{i'}}^{(l)}\right\}^{1-M_{i'l}} \cdot \eta_{Z_{i'}} \cdot \pi_j,\end{aligned}$$

   for $j = 1, ..., J$.

2. Update subclass indicators $Z_{i'}$ for case $i' = 1, ..., n_1$, from a multinomial distribution with probabilities

$$\begin{aligned}\mathbb{P}(Z_{i'} = k \mid \cdots) = q_{i'k} &\propto [\mathbf{M}_{i'} \mid Z_{i'}, I_{i'}, \boldsymbol{\Theta}, \boldsymbol{\Psi}][Z_{i'} \mid I_{i'}, \boldsymbol{\eta}] \\ &\propto \eta_k \cdot \left\{\theta_k^{(I_{i'})}\right\}^{M_{i'I_{i'}}} \left\{1 - \theta_k^{(I_{i'})}\right\}^{1-M_{i'I_{i'}}} \prod_{l \neq I_{i'}} \left\{\psi_k^{(l)}\right\}^{M_{i'l}} \left\{1 - \psi_k^{(l)}\right\}^{1-M_{i'l}}.\end{aligned}$$

   Update subclass indicators Z_i for control $i = n_1 + 1, ..., n_1 + n_0$, from a multinomial distribution with probabilities

$$\mathbb{P}(Z_i = k \mid \cdots) = q_{ik} \propto [\mathbf{M}_i \mid Z_i = k, \boldsymbol{\Psi}][Z_i = k \mid \boldsymbol{\nu}]$$

$$\propto \nu_k \cdot \prod_{j=1}^{J} \left\{\psi_k^{(j)}\right\}^{M_{ij}} \left\{1-\psi_k^{(j)}\right\}^{1-M_{ij}}, k = 1, ..., K.$$

3. Update the case subclass weights $\boldsymbol{\eta}$ for $j = 1, ..., J$ from

$$pr(\boldsymbol{\eta} \mid \cdots) \propto \prod_{i':I_{i'}=j} [Z_{i'} \mid \boldsymbol{\eta}, I_{i'}][\boldsymbol{\eta} \mid \alpha_1]$$

which can be accomplished by first setting $u_K^* = 1$ and sampling

$$u_k^* \sim \mathsf{Beta}\left(1 + z_k', \alpha_1 + \sum_{l=k+1}^{K} z_l'\right), k = 1, ..., K-1,$$

where z_k' is the number of cases assigned to subclass k in class j. We write

$$z_k' = \#\left\{i' : Y_{i'} = 1, Z_{i'} = k, I_{i'} = j\right\},$$

for $k = 1, ..., K-1$, where "$\#A$" counts the number of elements in set A. We then construct $\eta_1 = u_k^*$, $\eta_k = u_k^* \prod_{l=1}^{k-1} \{1 - u_l^*\}$, $k = 2, ..., K$.

4. Update the control subclass weights $\boldsymbol{\nu} = (\nu_1, ..., \nu_K)^T$ from

$$pr(\boldsymbol{\nu} \mid \cdots) \propto \prod_{i:Y_i=0} [Z_i \mid \boldsymbol{\nu}] \cdot [\boldsymbol{\nu} \mid \alpha_0],$$

which can be accomplished by first setting $v_K^* = 1$ and sampling

$$v_k^* \sim \mathsf{Beta}\left(1 + z_k, \alpha_0 + \sum_{l=k+1}^{K} z_k\right), k = 1, ..., K-1,$$

where z_k is the number of controls assigned to subclass k, and then constructing $\nu_1 = v_k^*$, $\nu_k = v_k^* \prod_{l=1}^{k-1}(1 - v_l^*)$, $k = 2, ..., K$.

5. Update concentration parameter α_0 and α_1 for stick-breaking prior from

$$pr(\alpha_0 \mid \cdots) \propto [\boldsymbol{\nu} \mid \alpha_0][\alpha_0] \propto \alpha_0^{K-1} \exp(-\alpha_0 \cdot r) \cdot pr(\alpha_0),$$

where $r = -\left\{\sum_{k=1}^{K-1} \log(1 - \nu_k^*)\right\}$. If conditionally conjugate prior for α_0 is used, i.e. $\alpha_0 \sim \mathsf{Gamma}(a_{\alpha_0}, b_{\alpha_0})$ with mean $a_{\alpha_0}/b_{\alpha_0}$ and variance $a_{\alpha_0}/b_{\alpha_0}^2$, then the full conditional distribution reduces to $\mathsf{Gamma}\left(a_{\alpha_0} + K - 1, b_{\alpha_0} + r\right)$. Similarly for α_1 with $\boldsymbol{\nu}$ replaced by $\boldsymbol{\eta}$ and $(a_{\alpha_0}, b_{\alpha_0})$ replaced by $(a_{\alpha_1}, b_{\alpha_1})$.

6. Update the vector of subclass TPR for $j = 1, ..., J$ from

$$pr(\boldsymbol{\theta}^{(j)} \mid \cdots) \propto \prod_{\{i':I_{i'}=j\}} [\mathbf{M}_{i'} \mid \boldsymbol{\theta}^{(j)}, Z_{i'}, I_{i'}][\boldsymbol{\theta}^{(j)}]$$
$$\propto \prod_{k=1}^{K} \left\{\theta_k^{(j)}\right\}^{m_{k1}^{(j)}} \left\{1-\theta_k^{(j)}\right\}^{m_{k0}^{(j)}} \cdot [\boldsymbol{\theta}^{(j)}],$$

where $m_{kc}^{(j)} = \#\{i' : Y_{i'} = 1, Z_{i'} = k, I_{i'} = j, M_{i'j} = c\}$, $c = 0, 1$. If prior for TPRs are independent Beta distributions, then this is a product of Beta distributions.

7. Update subclass-specific FPRs $\psi_k^{(j)}$ for $j = 1, ..., J$, $k = 1, ..., K$ from

$$pr(\psi_k^{(j)} \mid \cdots) \propto \prod_{i':Y_{i'}=1, I_{i'}\neq j, Z_{i'}=k} [M_{i'j} \mid \boldsymbol{\psi}^{(j)}, Z_{i'}, I_{i'}] \prod_{i:Y_i=0} [M_{ij} \mid \boldsymbol{\psi}^{(j)}, Z_i] \cdot [\psi_k^{(j)}]$$
$$\propto \left\{\psi_k^{(j)}\right\}^{s_{k1}^{(-j)}} \left\{1-\psi_k^{(j)}\right\}^{s_{k0}^{(-j)}} \cdot pr(\psi_k^{(j)}),$$

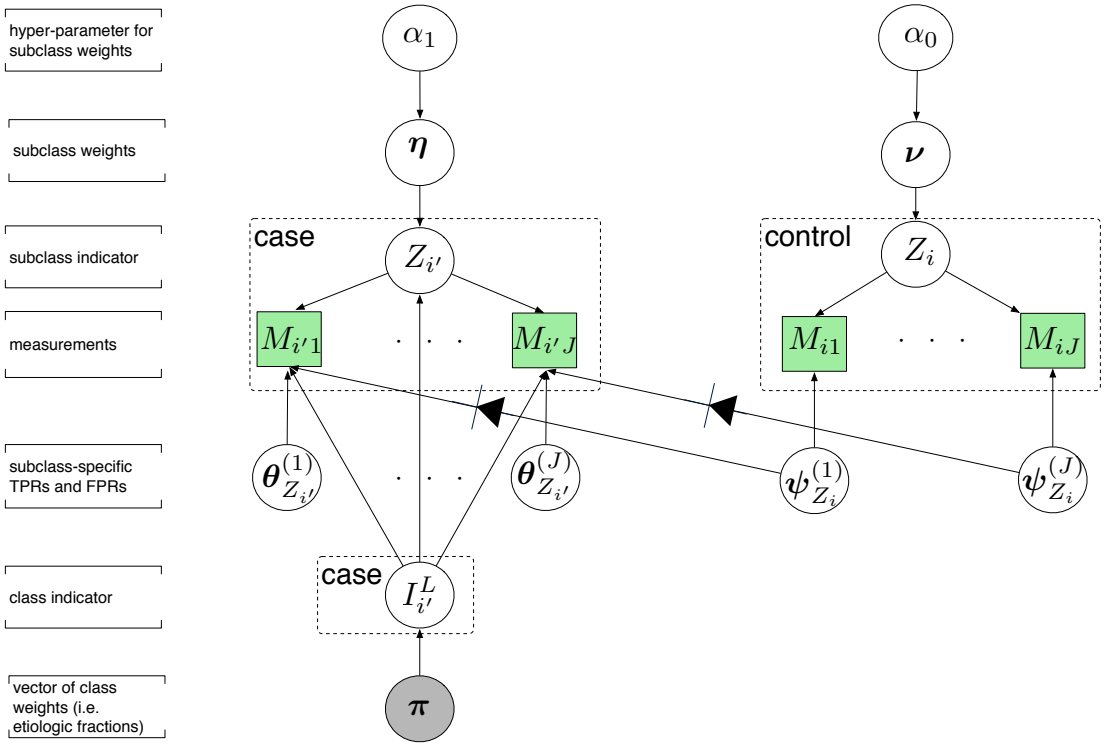
where $s_{kc}^{(-j)} = \#\{i' : Y_{i'} = 1, Z_{i'} = k, I_{i'} \neq j, M_{i'j} = c\} + \#\{i : Y_i = 0, Z_i = k, M_{ij} = c\}$, for $c = 0, 1$. If the prior on FPRs are $\mathsf{Beta}(a_1, b_1)$, then the above conditional distribution is $\mathsf{Beta}(a_1 + s_{k1}^{(j)}, b_1 + s_{k0}^{(j)})$.

8. Update $\boldsymbol{\pi}$ from $\mathsf{Dirichlet}\left(d_1 + t^{(j)}, ..., d_J + t^{(j)}\right)$, where $t^{(j)}$ is the number of cases assigned to class j, i.e. $t^{(j)} = \#\{i' : Y_{i'} = 1, I_{i'} = j\}$, $j = 1, .., J$.

Appendix D. Directed Acyclic Graph for Nested Partially-Latent Class Models

This section illustrates the model structure of nested partially-latent class models using a directed acyclic graph (DAG) and provides some details on posterior inference.

Because the false positive rate (FPR) parameters $\boldsymbol{\Psi}$ are in both the control and case likelihood (2.3) and (2.4) in the text, their posterior depend on both the control and case models. This is referred to as "feedback" because the case model will indirectly inform $\boldsymbol{\Psi}$. If we only want the control data to inform the case model but not vice versa, we can "cut" this source of feedback through approximate conditional updating in the Gibbs sampler (Lunn *and others*, 2009). That



Appendix Figure 1. Directed acyclic graph (DAG) for the npLCM. Quantities in circles are unknown parameters or auxiliary variables; quantities in solid squares are observables. The etiologic fraction $\boldsymbol{\pi}$ is of primary scientific interest. The solid arrows represent probabilistic relationship between the connected variables. The "cut" valve "A ⊣▶ B" means that when updating node A in the Gibbs sampler, we drop the likelihood terms that involve node B.

is, we update $\psi_k^{(j)}$ by $pr(\psi_k^{(j)} \mid M_{ij}; i : Y_i = 0)$ instead of Step 7 of the Gibbs sampler (see Appendix D). It will cut the information flow from the case model to the FPR parameters $\boldsymbol{\Psi}$ and is indicated by the check-bit valves in Appendix Figure 1. It is desirable when certain parts of the joint model are considered not reliable to inform a subset of parameters, and can be implemented by the `cut` function in `WinBUGS 1.4`. Such "cut-the-feedback" approximate Bayesian computation has both gains in computational speed and inferential robustness, and is also suggested in other contexts (Liu *and others*, 2009; Warren *and others*, 2012; Zigler and Dominici, 2014).

Appendix E. Parameter Settings and Coverage Rates in Simulation Studies

We present the true parameter values and the empirical coverage rates in simulation studies (Section 4).

Scenario I

$$\boldsymbol{\pi} = (0.5, 0.2, 0.15, 0.1, 0.05)'$$
$$\boldsymbol{\Theta}^T = \begin{bmatrix} 0.95 & 0.9 & 0.9 & 0.9 & 0.9 \\ 0.95 & 0.9 & 0.9 & 0.9 & 0.9 \end{bmatrix}$$
$$\boldsymbol{\Psi}^T = \begin{bmatrix} 0.25 & 0.25 & 0.2 & 0.15 & 0.15 \\ 0.2 & 0.2 & 0.25 & 0.1 & 0.1 \end{bmatrix}$$
$$\boldsymbol{\nu} = (0.5, 0.5)'$$
$$\boldsymbol{\eta} = (\eta_o, 1-\eta_o)',\ 0 \leqslant \eta_o \leqslant 1$$

Scenario II

$$\boldsymbol{\pi} = (0.5, 0.2, 0.15, 0.1, 0.05)'$$
$$\boldsymbol{\Theta}^T = \begin{bmatrix} 0.95 & 0.95 & 0.55 & 0.95 & 0.95 \\ 0.95 & 0.55 & 0.95 & 0.55 & 0.55 \end{bmatrix}$$
$$\boldsymbol{\Psi}^T = \begin{bmatrix} 0.4 & 0.4 & 0.05 & 0.2 & 0.2 \\ 0.05 & 0.05 & 0.4 & 0.05 & 0.05 \end{bmatrix}$$
$$\boldsymbol{\nu} = (0.5, 0.5)'$$
$$\boldsymbol{\eta} = (\eta_o, 1-\eta_o)',\ 0 \leqslant \eta_o \leqslant 1$$

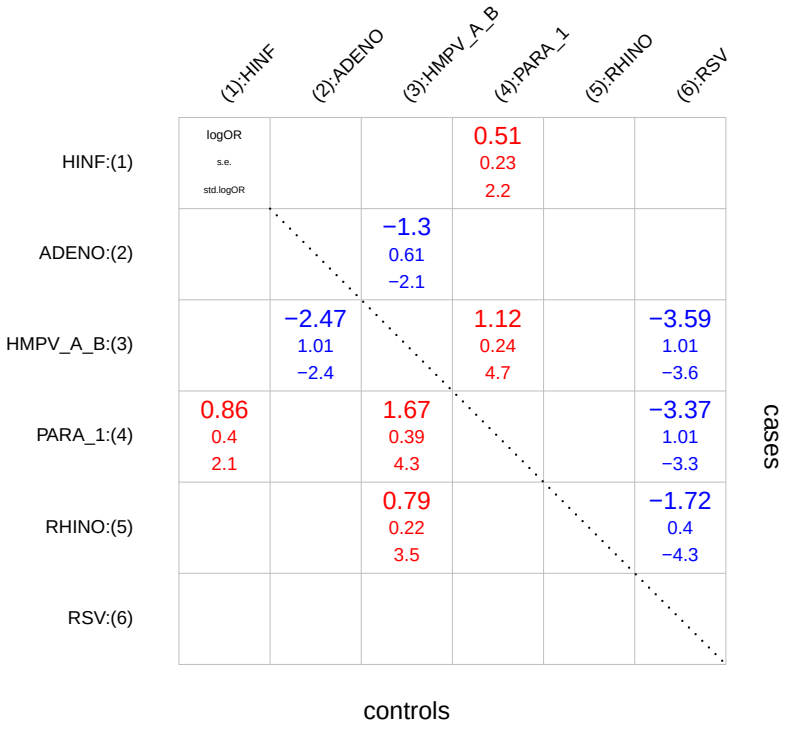
Appendix Table 1. Comparison of the actual coverage rates of 95% credible intervals for each disease class estimated by results fitted to 1,000 replication data sets.

			Truth: Cases' First Subclass Weight (η_o)				
		Model	0	0.25	0.5	0.75	1
	Class		100×Coverage (Standard Error)				
I	A	np	98.5(0.4)	99.3(0.3)	98.5(0.4)	98.1(0.4)	97.8(0.5)
		p	97.8(0.5)	97.6(0.5)	98.3(0.4)	97.7(0.5)	96.5(0.6)
	B	np	98.8(0.3)	97.9(0.5)	97.8(0.5)	97.3(0.5)	98.4(0.4)
		p	98.5(0.4)	98.2(0.4)	97.4(0.5)	97.7(0.5)	96.8(0.6)
	C	np	96.6(0.6)	98.5(0.4)	97.7(0.5)	97.7(0.5)	94.3(0.7)
		p	93.0(0.8)	96.6(0.6)	98.6(0.4)	97.5(0.5)	95.1(0.7)
	D	np	99.0(0.3)	99.1(0.3)	98.1(0.4)	98.1(0.4)	97.6(0.5)
		p	98.3(0.4)	98.6(0.4)	98.3(0.4)	96.9(0.5)	95.8(0.6)
	E	np	98.1(0.4)	98.5(0.4)	98.2(0.4)	98.0(0.4)	97.4(0.5)
		p	98.6(0.4)	97.1(0.5)	96.6(0.6)	96.3(0.6)	95.2(0.7)
II	A	np	95.4(0.7)	88.4(1.1)	88.2(1.1)	94.0(0.8)	98.8(0.3)
		p	99.6(0.2)	100.0(0.0)	96.7(0.6)	*85.6*(1.1)	**72.3**(1.4)
	B	np	*80.4*(1.3)	*84.8*(1.1)	*86.2*(1.1)	98.3(0.4)	98.1(0.4)
		p	**9.9**(0.9)	**62.5**(1.5)	92.1(0.9)	98.9(0.3)	*82.9*(1.2)
	C	np	89.2(1.0)	89.8(1.0)	97.2(0.5)	98.0(0.4)	*84.4*(1.1)
		p	**0.0**(0.0)	**6.1**(0.8)	91.0(0.9)	**75.3**(1.4)	**0.0**(0.0)
	D	np	93.5(0.8)	90.7(0.9)	95.4(0.7)	98.0(0.4)	95.1(0.7)
		p	**53.3**(1.6)	88.7(1.0)	97.2(0.5)	98.0(0.4)	93.9(0.8)
	E	np	95.4(0.7)	94.7(0.7)	96.1(0.6)	98.5(0.4)	96.5(0.6)
		p	**56.1**(1.6)	92.1(0.9)	97.8(0.5)	98.2(0.4)	92.0(0.9)

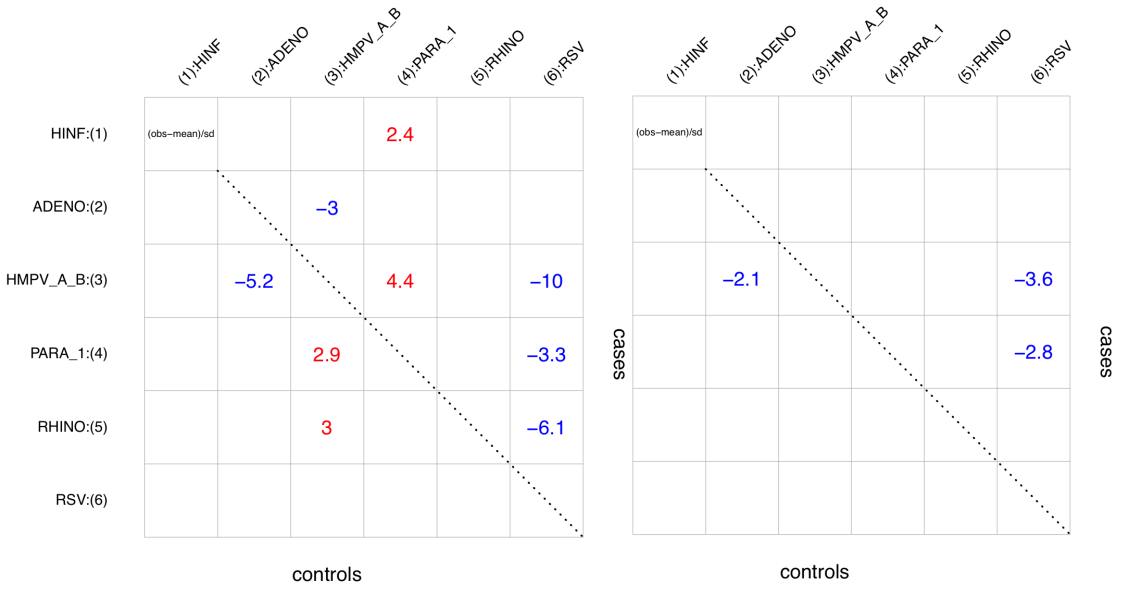
Appendix F. For Section 5: Analysis of PERCH Data

Full Pathogen Names and Abbreviations:

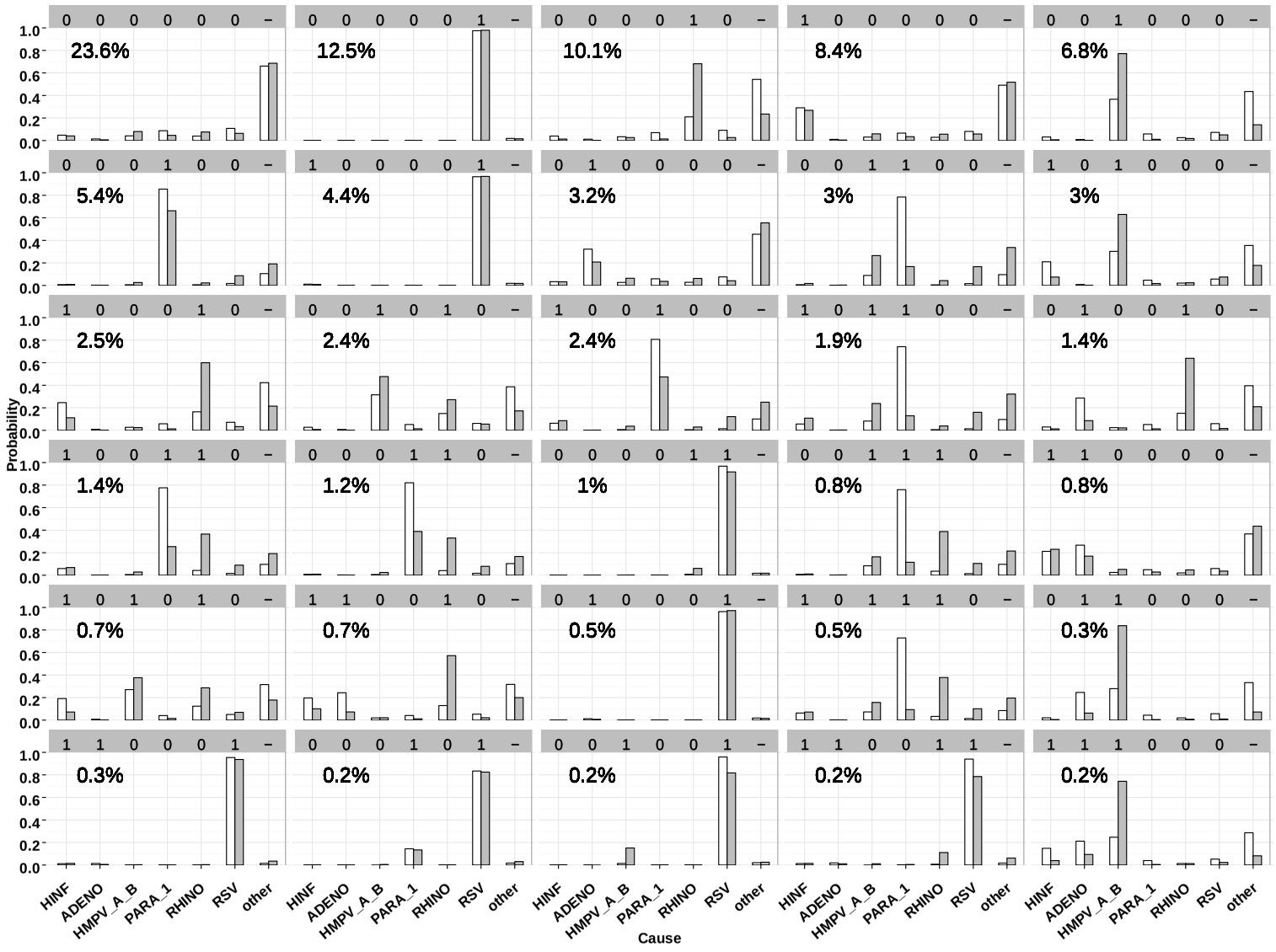
(1).HINF- *Haemophilus Influenzae*; (2). ADENO -Adenovirus; (3). HMPV-A/B - Human Metapneumovirus Type A or B; (4). PARA-1 - Parainfluenza Type 1 Virus; (5). RHINO - Rhinovirus; (6). RSV - Respiratory Syncytial Virus Type A or B.



Appendix Figure 2. Matrix of significant pairwise log odds ratios (LOR) for cases (upper) and controls (lower). LOR is at the top of the cell. Below it, its standard error is in smaller type, using the same color as the LOR. Then the estimate is divided by its standard error. We put the actual value when the Z-statistics has an absolute value greater than 2; a plus (red) or minus (blue) if between 1 and 2; blank otherwise.



Appendix Figure 3. Posterior predictive checking for pairwise odds ratios separately for cases (upper triangle) and controls (lower triangle) with expert priors on true positive rates. *Left*: pLCM; *Right*: npLCM. Each entry is a standardized log odds ratio difference (SLORD): the observed log odds ratio for a pair of measurements minus the mean LOR for the posterior predictive distribution divided by the standard deviation of the posterior predictive distribution. The first significant digit of absolute SLORs are shown in red for positive and blue for negative values, and only those greater than 2 are shown. On average, for a well fitting model, we expect $0.05 \times \binom{6}{2} \times 2 \approx 1.5(\pm 2.4)$ non-blank cells in cases and controls, respectively.



Appendix Figure 4. Each panel shows the individual etiology distribution estimated by the empirical distribution of MCMC samples of the disease class indicator. The height of a bar represents the probability of a case caused by each of the 7 causes labelled on the x-axis. For each cause, paired bars compare the estimates from the pLCM (left) and the npLCM (right); the binary codes at the top represent the NPPCR data with its observed frequency marked beneath (no measurements on "other" causes hence left as "-"). 30 most common patterns are shown here ordered by their observed frequencies.

References

Hjort, Nils Lid, Holmes, CC, Müller, Peter and Walker, Stephen G. (2010). Bayesian nonparametrics. *AMC* **10**, 12.

Ishwaran, Hemant and James, Lancelot F. (2002). Approximate Dirichlet Process computing in finite normal mixtures. *Journal of Computational and Graphical Statistics* **11**(3), 508–532.

Liu, Fei, Bayarri, MJ, Berger, JO *and others*. (2009). Modularization in bayesian analysis, with emphasis on analysis of computer models. *Bayesian Analysis* **4**(1), 119–150.

Lunn, David, Best, Nicky, Spiegelhalter, David, Graham, Gordon and Neuenschwander, Beat. (2009). Combining mcmc with 'sequential' PKPD modelling. *Journal of Pharmacokinetics and Pharmacodynamics* **36**(1), 19–38.

Neal, Radford M. (2000). Markov chain sampling methods for dirichlet process mixture models. *Journal of Computational and Graphical Statistics* **9**(2), 249–265.

Warren, Joshua, Fuentes, Montserrat, Herring, Amy and Langlois, Peter. (2012). Spatial-temporal modeling of the association between air pollution exposure and preterm birth: Identifying critical windows of exposure. *Biometrics* **68**(4), 1157–1167.

Wu, Zhenke, Deloria-Knoll, Maria, Hammitt, Laura L and Zeger, Scott L. (2015). Partially latent class models for case–control studies of childhood pneumonia aetiology. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, DOI: 10.1111/rssc.12101.

Zigler, Corwin Matthew and Dominici, Francesca. (2014). Uncertainty in propensity score estimation: Bayesian methods for variable selection and model-averaged causal effects. *Journal of the American Statistical Association* **109**(505), 95–107.