Appendices

A Log pooling Gaussian densities

We can exactly compute $p_{\text{pool}}$ when logarithmically pooling Gaussian densities. Noting that, in the one dimensional case, $N(\phi; \mu, \sigma^2)^{\lambda_m} = N(\phi; \mu, \frac{\sigma^2}{\lambda_m})$, we use the results of Bromiley (2003) and write

$$
\Omega_1 = \left( \begin{array}{cc} \lambda_1^{-1} & 0 \\ 0 & \lambda_3^{-1} \end{array} \right) \left[ \begin{array}{cc} \sigma_1^2 & 0 \\ 0 & \sigma_3^2 \end{array} \right]^{-1}, \quad \Omega_2 = \left( \begin{array}{cc} \lambda_2^{-1} & 0 \\ 0 & \lambda_3^{-1} \end{array} \right) \left[ \begin{array}{cc} \sigma_2^2 & \rho \sigma_2 \sigma_3 \\ \rho \sigma_2 \sigma_3 & \sigma_3^2 \end{array} \right]^{-1}
$$

(51)

$$
\Sigma_{\text{log}} = (\Omega_1 + \Omega_2)^{-1}, \quad \mu_{\text{log}} = \Sigma_{\text{log}} \left( \begin{array}{c} \mu_1 \\ \mu_3 \end{array} \right) + \Omega_2 \left( \begin{array}{c} \mu_2,1 \\ \mu_2,1 \end{array} \right)
$$

(52)

hence $p_{\text{pool}}(\phi_{1-2}, \phi_{2-3}) = N((\phi_{1-2}, \phi_{2-3}); \mu_{\text{log}}, \Sigma_{\text{log}})$. The choice of $\lambda_2$ is critical; by controlling the contribution of $p_2$ to $p_{\text{pool}}, \lambda_2$ controls the degree of correlation present in the latter. The left hand column of Figure 3 illustrates this phenomena. When $\lambda_1 = \lambda_3 = 0 \implies \lambda_2 = 1$, all correlation in $p_2$ is present in $p_{\text{pool}}$. The correlation decreases for increasing values of $\lambda_1$ until $\lambda_1 = \lambda_3 = 0.5 \implies \lambda_2 = 0$, where no correlation persists.

B Sequential sampler

![Figure 11: A DAG of the submodels and their common quantities, with the sequential sampling strategy overlaid. The stage one target ($s_1$) is encapsulated in the light blue dashed line; stages two and three ($s_2$, $s_3$) are in dark blue and red respectively.](image)

Figure 11 depicts graphically the strategy employed by the sequential sampler. The sequential sampler assumes that the pooled prior decomposes such that

$$
p_{\text{pool}}(\phi) = p_{\text{pool},1}(\phi_{1-2})p_{\text{pool},2}(\phi_{1-2}, \phi_{2-3})p_{\text{pool},3}(\phi_{1-2}, \phi_{2-3}).
$$

(53)

This is necessary to avoid sampling all components of $\phi$ in the first stage. All pooled priors trivially satisfy (53), as we can assume all but $p_{\text{pool},3}(\phi_{1-2}, \phi_{2-3})$ are improper, flat distributions. However, including some portion of the pooled prior in each stage of the sampler can improve performance, and eliminate computational instabilities when submodel likelihoods contain little information.

Stage one

Stage one of the sequential sampler targets

$$
P_{\text{meld},1}(\phi_{1-2}, \psi_1 | Y_1) \propto p_{\text{pool},1}(\phi_{1-2})p_{1}(\phi_{1-2}, \psi_1, Y_1) / p_{1}(\phi_{1-2}),
$$

(54)
using a generic proposal kernel for both $\phi_{1:2}$ and $\psi_1$. The corresponding acceptance probability for a proposed update from $(\phi_{1:2}, \psi_1)$ to $(\phi^*_{1:2}, \psi^*_1)$ is

$$
\alpha((\phi^*_{1:2}, \psi^*_1), (\phi_{1:2}, \psi_1)) = \frac{p_{pool,1}(\phi^*_{1:2}) p_{1}(\phi^*_{1:2}, \psi^*_1, Y_1)p_{1}(\phi_{1:2}) q(\phi_{1:2}, \psi_1 | \phi^*_{1:2}, \psi^*_1)}{p_{pool,1}(\phi_{1:2}) p_{1}(\phi_{1:2}, \psi_1, Y_1)p_{1}(\phi^*_{1:2}) q(\phi^*_{1:2}, \psi^*_1 | \phi_{1:2}, \psi_1)}. \quad (55)
$$

**Stage two**

The stage two target augments the stage one target by including the second submodel, corresponding prior marginal distribution, and an additional pooled prior term

$$
p_{meld,2}(\phi_{1:2}, \phi_{2:3}, \psi_1, \psi_2 | Y_1, Y_2) \propto p_{pool,1}(\phi_{1:2}) p_{pool,2}(\phi_{1:2}, \phi_{2:3}) \frac{p_{1}(\phi_{1:2}, \psi_1, Y_1)p_{2}(\phi_{1:2}, \phi_{2:3}, \psi_2, Y_2)}{p_1(\phi_{1:2}) p_2(\phi_{1:2}, \phi_{2:3})}. \quad (56)
$$

A Metropolis-within-Gibbs strategy is employed, where the stage one samples are used as a proposal for $\phi_{1:2}$, whilst a generic proposal kernel is used for $\psi_2$ and $\phi_{2:3}$. Thus the proposal distributions for $\phi^*_{1:2}$ and $\phi^*_{2:3}$ are

$$
\phi^*_{1:2}, \psi^*_1 | \phi_{2:3}, \psi_2 \sim p_{meld,1}(\phi^*_{1:2}, \psi^*_1 | Y_1) \quad (57)
$$

$$
\phi^*_{2:3}, \psi^*_2 | \phi_{1:2}, \psi_1 \sim q(\phi^*_{2:3}, \psi^*_2 | \phi_{2:3}, \psi_2). \quad (58)
$$

The acceptance probability for this proposal strategy is

$$
\alpha((\phi^*_{1:2}, \psi^*_1), (\phi_{1:2}, \psi_1)) = \frac{p_{pool,2}(\phi^*_{1:2}, \phi_{2:3}) p_{2}(\phi^*_{1:2}, \phi_{2:3}, \psi_2, Y_2)p_{2}(\phi_{1:2}, \phi_{2:3})}{p_{pool,2}(\phi_{1:2}, \phi_{2:3}) p_{2}(\phi_{1:2}, \phi_{2:3}, \psi_2, Y_2)p_{2}(\phi^*_{1:2}, \phi_{2:3})} \quad (59)
$$

$$
\alpha((\phi^*_{2:3}, \psi^*_2), (\phi_{2:3}, \psi_2)) = \frac{p_{pool,2}(\phi_{1:2}, \phi_{2:3}) p_{2}(\phi_{1:2}, \phi_{2:3}, \psi_2, Y_2)p_{2}(\phi_{1:2}, \phi_{2:3}) q(\phi_{2:3}, \psi_2 | \phi^*_{2:3}, \psi^*_2)}{p_{pool,2}(\phi_{1:2}, \phi_{2:3}) p_{2}(\phi_{1:2}, \phi_{2:3}, \psi_2, Y_2)p_{2}(\phi_{1:2}, \phi_{2:3}) q(\phi^*_{2:3}, \psi^*_2 | \phi_{2:3}, \psi_2)}. \quad (60)
$$

Our judicious choice of proposal distribution has resulted in a cancellation in Equation (59) which removes all terms related to $p_1$. Similarly, all terms related to $p_2$ are constant – hence cancel – in Equation (60). This eliminates any need to re-evaluate the first submodel.

**Stage three**

In stage three we target the full melded posterior

$$
p_{meld,3}(\phi_{1:2}, \phi_{2:3}, \psi_1, \psi_2, \psi_3 | Y_1, Y_2, Y_3) \propto p_{meld}(\phi_{1:2}, \phi_{2:3}, \psi_1, \psi_2, \psi_3 | Y_1, Y_2, Y_3). \quad (61)
$$
The target has now been broadened to include terms from the third submodel and the entirety of the pooled prior. Again, we employ a Metropolis-within-Gibbs sampler, with proposals drawn such that

\[
\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi^\tau_1, \psi^\tau_2 \mid \psi_3 \sim P_{meld, 2}(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi^\tau_1, \psi^\tau_2 \mid Y_1, Y_2) \tag{62}
\]

\[
\psi^\tau_3 \mid \phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi^\tau_1, \psi^\tau_2 \sim q(\psi^\tau_3 \mid \psi_3), \tag{63}
\]

which leads to acceptance probabilities of

\[
\alpha((\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi^\tau_1, \psi^\tau_2), (\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi, \psi_2)) = \frac{P_{pool, 3}(\phi_{1\tau, 2}, \phi_{2\tau, 3}) p_3(\phi_{2\tau, 3}, \psi_3, Y_3) p_3(\phi_{2\tau, 3})}{P_{pool, 3}(\phi_{1\tau, 2}, \phi_{2\tau, 3}) p_3(\phi_{2\tau, 3}, \psi, Y_3) p_3(\phi_{2\tau, 3})} \tag{64}
\]

\[
\alpha(\psi^\tau_3, \psi_3) = \frac{p_3(\phi_{2\tau, 3}, \psi^\tau_3, Y_3) q(\psi^\tau_3 \mid \psi_3)}{p_3(\phi_{2\tau, 3}, \psi_3, Y_3) q(\psi^\tau_3 \mid \psi_3)}. \tag{65}
\]

The informed choice of proposal distribution for \((\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi, \psi_2)\) has allowed us to target the full melded posterior without needing to evaluate all submodels simultaneously.

## C Normal approximation calculations

Substituting in the approximations of Section 3.2 to Equation (36) yields the approximate melded posterior

\[
\hat{p}_{meld}(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_1, Y_2, Y_3) \propto \frac{\hat{p}_1(\phi_{1\tau, 2} \mid \tilde{\mu}_1, \tilde{\Sigma}_1) \hat{p}_3(\phi_{2\tau, 3} \mid \tilde{\mu}_3, \tilde{\Sigma}_3) p_2(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_2)}{\hat{p}_1(\phi_{1\tau, 2} \mid \tilde{\mu}_1, 0, \tilde{\Sigma}_1) \hat{p}_3(\phi_{2\tau, 3} \mid \tilde{\mu}_3, 0, \tilde{\Sigma}_3) p_2(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_2)}. \tag{66}
\]

Noting that the product of independent normal densities is an unnormalised multivariate normal density with independent components, we rewrite Equation (66) as

\[
\hat{p}_{meld}(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_1, Y_2, Y_3) \propto \frac{\hat{p}_{nu}(\phi_{1\tau, 2}, \phi_{2\tau, 3} \mid \hat{\mu}_{nu}, \hat{\Sigma}_{nu}) p_2(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_2)}{\hat{p}_{de}(\phi_{1\tau, 2}, \phi_{2\tau, 3} \mid \hat{\mu}_{de}, \hat{\Sigma}_{de}) p_2(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_2)},
\]

\[
\hat{\mu}_{nu} = \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_3 \end{bmatrix}, \quad \hat{\Sigma}_{nu} = \begin{bmatrix} \hat{\Sigma}_{1,1} & \hat{\Sigma}_{1,3} \\ \hat{\Sigma}_{3,1} & \hat{\Sigma}_{3,3} \end{bmatrix}, \quad \hat{\mu}_{de} = \begin{bmatrix} \hat{\mu}_{1,0} \\ \hat{\mu}_{3,0} \end{bmatrix}, \quad \hat{\Sigma}_{de} = \begin{bmatrix} \hat{\Sigma}_{1,0} & 0 \\ 0 & \hat{\Sigma}_{3,0} \end{bmatrix}. \tag{67}
\]

The ratio of normal densities is also an unnormalised normal density, and hence Equation (67) simplifies to

\[
\hat{p}_{meld}(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_1, Y_2, Y_3) \propto \hat{p}(\phi_{1\tau, 2}, \phi_{2\tau, 3} \mid \hat{\mu}, \hat{\Sigma}) p_2(\phi_{1\tau, 2}, \phi_{2\tau, 3}, \psi_2 \mid Y_2),
\]

\[
\hat{\Sigma} = (\hat{\Sigma}_{nu}^{-1} - \hat{\Sigma}_{de}^{-1})^{-1}, \quad \hat{\mu} = \hat{\Sigma} \left( \hat{\Sigma}_{nu}^{-1} \hat{\mu}_{nu} - \hat{\Sigma}_{de}^{-1} \hat{\mu}_{de} \right). \tag{68}
\]
**D Calculating the cumulative fluid balance from the raw fluid data**

In the raw fluid data each patient has \( \tilde{I} = 1, \ldots, \tilde{L}_i \) observations. Each observation \( \tilde{x}_{i,l} \) is typically a small fluid administration (e.g. an injection of some medicine in saline solution), or a fluid discharge (almost always urine excretion). The observations have corresponding observation times \( \tilde{u}_{i,l} \), with \( \tilde{u}_i = \{ \tilde{u}_{i,1}, \ldots, \tilde{u}_{i,\tilde{L}_i} \} \) and \( \tilde{x}_i = \{ \tilde{x}_{i,1}, \ldots, \tilde{x}_{i,\tilde{L}_i} \} \). We code the fluid administrations/inputs as positive values, and the excretions/outputs as negative values. Each patient has an enormous number of raw fluid observations \( (\tilde{L}_i \ll \tilde{L}_i) \) and it is computationally infeasible to consider them all at once. We aggregate the raw fluid observations into 8-hourly changes in fluid balance. From these 8-hourly changes we calculate the cumulative fluid balance.

Mathematically, we define an ordered vector of boundary values

\[
v_i = ([\min\{\tilde{u}_i\}], [\min\{\tilde{u}_i\}] + \frac{1}{3}, \ldots, [\max\{\tilde{u}_i\}]),
\]

noting that \( \dim(v_i) = \tilde{L}_i + 1 \). Because the observation times encoded as *days since ICU admission* and we are interested in the 8-hourly changes, our floor and ceiling functions round down or up to the appropriate third respectively. The raw fluid observations are then divided up into \( \tilde{L}_i \) subsets of \( \{ \tilde{x}_i, \tilde{u}_i \} \) based on which boundary values the observation falls in between:

\[
V_{i,l} = \{ \{ \tilde{x}_i, \tilde{u}_i \} \mid v_{i,l} \leq \tilde{u}_i < v_{i,l+1} \},
\]

for \( l = 1, \ldots, \tilde{L}_i \). Denote \( N^V_{i,l} = |V_{i,l}| / 2 \) (dividing by two as \( V_{i,l} \) contains both the observation and the observation time). The \( l \)th 8-hourly fluid change \( \Delta_{i,l} \) and corresponding observation time \( u_{i,l} \) can then be computed as

\[
\Delta_{i,l} = \sum_{s=1}^{N^V_{i,l}} \tilde{x}_{i,s}, \text{ s.t. } \tilde{x}_{i,s} \in V_{i,l}, \quad u_{i,l} = \frac{1}{N^V_{i,l}} \sum_{s=1}^{N^V_{i,l}} \tilde{u}_{i,s}, \text{ s.t. } \tilde{u}_{i,s} \in V_{i,l}.
\]

Finally, the 8-hourly cumulative fluid balance data are computed by \( x_{i,l} = \sum_{s=1}^{l} \Delta_{i,s} \), and we assume they too are observed at \( u_{i,l} \).

**E Priors and justification for the cumulative fluid submodel**

The parameters for the gamma prior for \( \eta_{1,i}^b \) and \( \eta_{1,i}^a \) are obtained by assuming that the 2.5-, 50-, and 97.5- percentiles are at 0.5, 5, and 20 (Belgorodski et al., 2017). A slope of 0.5 (i.e. the change in cumulative fluid balance per day) is unlikely but possible due to missing data; a slope of 20 is also unlikely but possible as extremely unwell patients can have very high respiratory rates and thus require large fluid inputs.

The prior for the breakpoint \( \kappa_i \) is derived as follows. Define \( u_{i,(1)} = \min(u_i) \) and \( u_{i,(n)} = \max(u_i) \), with \( r_i = u_{i,(n)} - u_{i,(1)} \). We reparameterise the breakpoint by noting that \( \kappa_i = \kappa_i^{\text{raw}} r_i + u_{i,(1)} \), where \( \kappa_i^{\text{raw}} \in [0, 1] \). We then set \( \kappa_i^{\text{raw}} \sim \text{Beta}(5, 5) \) to regularise the breakpoint towards the middle of each individual’s stay in ICU. This is crucial to
ensure the submodel is identifiable when there is little evidence of a breakpoint in the data. Note that this results in the following analytic expression for \( p_{2,3} \) of \( (\phi_2^2 \cap 3) \)

\[
\begin{align*}
p_3(\phi_{2,3}) &= \prod_{i=1}^{N} p(\eta_{1,i}^b)p(\eta_{1,i}^a)p(\kappa_i), \quad \text{with} \quad p(\kappa_i) = p_{\kappa_i}^{raw} \left( \frac{\kappa_i - \mu_{i,(1)}}{r_i} \right) \frac{1}{r_i} \quad (72)
\end{align*}
\]

by the change of variables formula.

Specifying a prior for \( \eta_{0,i} \), the cumulative fluid balance at \( \kappa_i \), is difficult because it too depends on the length of stay. Instead, we reparameterise so that \( \eta_{0,i} \) is a function of the y-intercept \( \eta_{0,i}^{raw} \).

\[
\begin{align*}
\eta_{0,i} = (\eta_{0,i}^{raw} + \eta_{1,i}^b \kappa_i)1_{\{t < \kappa_i\}} + (\eta_{0,i}^{raw} + \eta_{1,i}^a \kappa_i)1_{\{t \geq \kappa_i\}}
\end{align*}
\]

We place a LogNormal(1.61, 0.47) prior on \( \eta_{0,i}^{raw} \). These values are obtained assuming that, a priori, the 2.5%, 50%, and 99% percentiles of \( \eta_{0,i}^{raw} \) are 0.5, 5, and 15 respectively (Belgorodski et al., 2017). This is a broad prior that reflects the numerous possible admission routes into the ICU. We expect those admitted from the wards to have little pre-admission fluid data. Those admitted from the operating theatre occasionally have their in-theatre fluid input recorded after admission into the ICU, with no easy way to distinguish these records in the data.

\section*{F Analytic form for the survival function}

The hazard at arbitrary time \( t \) is

\[
\begin{align*}
h_i(t) &= \gamma t^{\gamma-1} \exp \left\{ \begin{bmatrix} w_i^\top \theta + \alpha \frac{\partial}{\partial t} m_i(t) \end{bmatrix} \right\} \\
m_i(t) &= \eta_{0,i} + \eta_{1,i}^b (t - \kappa_i)1_{\{t < \kappa_i\}} + \eta_{1,i}^a (t - \kappa_i)1_{\{t \geq \kappa_i\}} \\
\frac{\partial}{\partial t} m_i(t) &= \eta_{1,i}^b 1_{\{t < \kappa_i\}} + \eta_{1,i}^a 1_{\{t \geq \kappa_i\}}.
\end{align*}
\]

Then, for \( t > \kappa_i \), the cumulative hazard is

\[
\begin{align*}
\int_0^t h_i(u)du &= \int_0^t \gamma u^{\gamma-1} \exp \left\{ \begin{bmatrix} w_i^\top \theta + \alpha \eta_{1,i}^b 1_{\{u < \kappa_i\}} + \alpha \eta_{1,i}^a 1_{\{u \geq \kappa_i\}} \end{bmatrix} \right\} du \\
&= \gamma \exp \{ w_i^\top \theta \} \int_0^\kappa u^{\gamma-1} \exp \left\{ \alpha \eta_{1,i}^b 1_{\{u < \kappa_i\}} + \alpha \eta_{1,i}^a 1_{\{u \geq \kappa_i\}} \right\} du \\
&= \gamma \exp \{ w_i^\top \theta \} \left[ \int_0^\kappa u^{\gamma-1} \exp \{ \alpha \eta_{1,i}^b \} du + \int_{\kappa_i}^t u^{\gamma-1} \exp \{ \alpha \eta_{1,i}^a \} du \right] \\
&= \exp \{ w_i^\top \theta \} \left[ \exp \{ \alpha \eta_{1,i}^b \} \kappa_i^{\gamma} + \exp \{ \alpha \eta_{1,i}^a \} (t^{\gamma} - \kappa_i^{\gamma}) \right]
\end{align*}
\]

and for \( t < \kappa_i \),
\[
\int_0^t h_i(u)du = \gamma \exp\{w_i^\top \theta\} \left[ \int_0^t u^{\gamma - 1} \exp\{\alpha \eta_{i,i}^b\} du \right]
= \exp\{w_i^\top \theta\} \left[ \exp\{\alpha \eta_{i,i}^b\} t_i^\gamma \right]
= t_i^\gamma \exp\{w_i^\top \theta + \alpha \eta_{i,i}^b\}.
\]

The survival functions then have corresponding definitions for \( t > \kappa_i \) and \( t < \kappa_i \) as \( S_i(t) = \exp\{-\int_0^t h_i(u)du\} \).

### G Survival submodel prior justification

Our prior for \((\gamma, \alpha, \theta)\) must result in a plausible distribution for \( p_{2,i}(T_i \mid d_i = 1) \), and a reasonable balance between \( d_i = 1 \) and \( d_i = 0 \) events. The primary concern is unintentionally specifying a prior for which the bulk of \( p_{2,i}(T_i \mid d_i = 1) \) is very close to zero. In addition, certain extreme configurations of \((\gamma, \alpha, \theta)\) cause issues for the methodology of Crowther and Lambert (2013), particularly the numerical root finding and numerical integration steps. We would like to rule out such extreme configurations \textit{a priori}. Ideally we would encode this information a joint prior for \((\gamma, \alpha, \theta)\), but specifying the appropriate correlation structure for these parameters is prohibitively challenging. Instead we focus on specifying appropriate marginals for each of \( \gamma, \alpha \), and \( \theta \), and create visual prior predictive checks (Gabry et al., 2019; Gelman et al., 2020) to ensure the induced prior for \((T_i, d_i)\) is acceptable.

Before justifying our chosen marginal prior, we note that the \( \exp\{x_i^\top \theta + \alpha \frac{\partial}{\partial T_i} m_i(T_i)\} \) term implies that the priors for \( \theta \) and \( \alpha \) are on the log-scale. Hence the magnitude of these parameters must be small, otherwise all event times would be very near zero or at infinity. The asymmetric effect of the transformation from the log scale also implies that symmetric priors are not obviously sensible. From these observations we deduce that \( \theta \) and \( \alpha \) must not be too large in magnitude, however if they are negative then they can be slightly larger. Hence, we specify the skew-normal priors detailed in Section 5.3, noting that the skewness parameter for \( \alpha \) is smaller, because \( \frac{\partial}{\partial T_i} m_i(T_i) \) is strictly positive and typically between 0.5 and 20, whilst \( w_i \) is standardised to be approximately standard normal. Lastly, if \( \gamma \) is too far away from 1 (in either direction), then the event times are very small either because the hazard increases rapidly \( (\gamma \gg 1) \), or because almost all of the cumulative hazard is in the neighbourhood of 0 \( (\gamma \ll 1) \). We specify a gamma distribution for \( \gamma \) with the 1\(^{\text{st}}\), 50\(^{\text{th}}\), and 99\(^{\text{th}}\)-percentiles of \( p_2(\gamma) \) at 0.2, 1, and 2, allowing for a wide range of hazard shapes, but removing many of the extremes.

### H Estimating submodel prior marginal distributions

For \( p_1 \), we note that \( p_1(\phi_{1-2}) = \prod_{i=1}^N p_{1,i}(T_i, d_i) \), and that \( p_{1,i}(T_i, d_i) \) conditions on each individual’s length of stay (in specifying the location of the knots), as well as the range, mean, and standard deviation of the P/F data (by standardising \( \tilde{z}_{i,j} \)). Simple Monte Carlo samples are drawn from \( p_1(\phi_{1-2}) \) and used to estimate \( \tilde{p}_1(\phi_{1-2}) \). Under
the second submodel we obtain samples from \( p_2(\phi_{1:2}, \phi_{2:3}) \) using the methodology of Crowther and Lambert (2013) as implemented in \texttt{simsurv} (Brilleman, 2021). These samples are used to estimate \( \tilde{p}_2(\phi_{1:2}, \phi_{2:3}) \).

**P/F submodel**

We approximate \( p_1(\phi_{1:2}) \) using a mixture of discrete and continuous distributions, with a discrete spike at \( C_i \) for the censored events and a beta distribution for the (rescaled) event times. Monte Carlo samples of \( T_i \) and \( d_i \) are obtained from \( p_{1,i}(T_i, d_i) \) by drawing \( \beta_{0,i} \) and \( \zeta_i \) from their respective prior distributions and then solving (46). Denoting the estimated mixture weight \( \hat{\pi}_i \in [0, 1] \), the density estimate is

\[
\hat{p}_{1,i}(T_i, d_i) = \hat{\pi}_i \text{Beta} \left( \frac{T_i}{C_i}; \hat{a}_i, \hat{b}_i \right) \frac{1}{C_i} \mathbb{1}_{\{d_i=1\}} + \left(1 - \hat{\pi}_i\right) \mathbb{1}_{\{d_i=0, T_i=C_i\}}
\]

(74)

where \( \hat{\pi}_i, \hat{a}_i \) and \( \hat{b}_i \) are maximum likelihood estimates obtained using the prior samples. Examples of \( \hat{p}_{1,i}(T_i, d_i) \) for a subset of individuals are displayed in Figure 12.

**Survival submodel**

Our estimate of \( p_2(\phi_{1:2}, \phi_{2:3}) \) relies on the fact that

\[
p_2(\phi_{1:2}, \phi_{2:3}) = \prod_{i=1}^{N} p_{2,i}(T_i, d_i, \kappa_i, \eta_{1,i}^b, \eta_{1,i}^a).
\]

(75)
As such we estimate \( p_{2,i}(T_i, d_i, \kappa_i, \eta_{1,i}^b, \eta_{1,i}^a) \) for each individual and take the product of these estimates. Drawing samples from \( p_{2,i}(T_i, d_i, \kappa_i, \eta_{1,i}^b, \eta_{1,i}^a) \) is challenging: we use the approach proposed by Crowther and Lambert (2013) as implemented in Brilleman (2021). Inspecting the samples reveals correlation between \( (T_i, d_i) \) and \( (\kappa_i, \eta_{1,i}^b, \eta_{1,i}^a) \) that we would like to capture in our estimate. To do so, we fit a mixture of multivariate normal distributions to transformations of the continuous parameters with support on \( \mathbb{R} \),

\[
\tilde{T}_i = \text{Logit} \left( \frac{T_i}{C_i} \right), \quad \tilde{\kappa}_i = \text{Logit} \left( \frac{\kappa_i - u_{i,(1)}}{u_{i,(n)} - u_{i,(1)}} \right),
\]

\[
\tilde{\eta}_{1,i}^b = \log(\eta_{1,i}^b), \quad \tilde{\eta}_{1,i}^a = \log(\eta_{1,i}^a).
\]

The resulting density estimate, with estimated mixture weight \( \tilde{\theta}_i \in [0, 1] \), is

\[
\tilde{p}_2(T_i, d_i, \kappa_i, \eta_{1,i}^b, \eta_{1,i}^a) = \left[ (\tilde{\theta}_i)N \left( \left[ \tilde{T}_i, \tilde{\kappa}_i, \tilde{\eta}_{1,i}^b, \tilde{\eta}_{1,i}^a \right]^T; \tilde{\mu}_{1,i}, \tilde{\Sigma}_{1,i} \right) \right] 1_{\{d_i = 1\}} + (1 - \tilde{\theta}_i)N \left( \left[ \tilde{\kappa}_i, \tilde{\eta}_{1,i}^b, \tilde{\eta}_{1,i}^a \right]^T; \tilde{\mu}_{0,i}, \tilde{\Sigma}_{0,i} \right) 1_{\{d_i = 0, T_i = C_i\}} \right],
\]

where \( \tilde{\theta}_i, \tilde{\mu}_{1,i}, \tilde{\Sigma}_{1,i}, \tilde{\mu}_{0,i}, \) and \( \tilde{\Sigma}_{0,i} \) are maximum likelihood estimates, and the Jacobian correction \( J_i \) is

\[
J_i = \left[ \left( \frac{1}{C_i - T_i} \right)^{d_i} \left( \frac{1}{u_{i,(n)} - \kappa_i} + \frac{1}{\kappa_i - u_{i,(1)}} \right) \right] \left( \frac{1}{\eta_{1,i}^b} \right) \left( \frac{1}{\eta_{1,i}^a} \right) .
\]

We assess the fit of this estimate by drawing samples from \( \tilde{p}_{2,i}(\phi_{1:-2}, \phi_{2:-3}) \) and comparing them to the Monte Carlo samples drawn using \texttt{simsurv}. Our visual assessment is displayed in Figure 14 for individual \( i = 19 \). The normal approximation seems to fit the samples well, with the shape of \( p_{2,i}(T_i \mid d_i = 1) \) closely matching that of the Monte Carlo samples, and with a similar mix of \( d_i = 0 \) and \( d_i = 1 \) samples.

We also require an estimate of \( p_2(\phi_{1:-2}) \) for experiments discussed in Section 4.7. This is obtained using the samples generated under the survival submodel prior and the methodology of Section H. The raw samples and fit are displayed in Figure 13.

## 1 Cohort selection criteria

This appendix details the cohort selection criteria and our rationale for them. In the text we speak of the \( i \)th individual. This is because in our final data set (the data that results from applying the following criteria) we are dealing with unique individuals, however some individuals in MIMIC have multiple ICU stays. In this appendix \( i \) represents a single stay in ICU.

1. Each ICU stay must have at least 12 \( \text{PaO}_2/\text{FiO}_2 \) observations \( (J_i \geq 12) \), with the first 6 being greater than 350 \( (z_{i,j} > 350 \text{ for } j = 1, \ldots, 1) \).
This is to ensure we have enough data to fit a B-spline with 7 internal knots. The restriction on the first 6 observations is to avoid selecting those who have already started to experience respiratory failure prior to ICU admission.

2. The time between any 2 consecutive PaO\textsubscript{2}/FiO\textsubscript{2} observations cannot exceed 2 days.

   - This is because we believe this lack of PaO\textsubscript{2}/FiO\textsubscript{2} observations is likely to be an error in the data: e.g. what appears as a single ICU stay in the data is actually two or more separate stays.

3. The fluid observations must be after ICU admission (some observations are entered as ‘Pre-admission intake’) and cannot be associated with fluid administered in the operating room (OR). Note that this does not mean all OR fluid administrations are removed, as some are mis-labelled.

4. There must be sufficient temporal overlap between the fluid data and the PaO\textsubscript{2}/FiO\textsubscript{2} data. Specifically,

\[
\frac{\max \{0, \min [\max (t_i), \max (\bar{u}_i)] - \max [\min (t_i), \min (\bar{u}_i)]\}}{\max [\max (t_i), \max (\bar{u}_i)] - \min [\min (t_i), \min (\bar{u}_i)]} > 0.9
\]

   - The numerator of (78) is strictly positive, and the denominator ensures that the quantity is bounded between 0 and 1.
   - We cannot investigate the relationship between the rate of fluid intake and respiratory failure if the latter occurs without sufficient fluid data surrounding the event.
Figure 14: Monte Carlo (MC) samples from $p_{2,i}(\phi_{1,2}, \phi_{2,3})$ obtained using simsurv and samples from the fitted normal approximation (NA) for $i = 19$. The panels on the off diagonal elements contain a 2D kernel density estimate for $d_i = 1$ and the samples for $d_i = 0$. Diagonal and lower-triangular panels are on their original scales, whilst the upper-triangular panels are on the log scale.

### J Baseline covariate information

The baseline covariate vector $w_i$ contains the median of the measurements taken in the first 24 hours of the ICU stay, which are then standardised, of the following covariates:
Anion gap, Bicarbonate, Creatinine, Chloride, Glucose, Hematocrit, Hemoglobin, Platelet, Partial thromboplastin time, International normalized ratio, Prothrombin time, Sodium, blood Urea nitrogen, White blood cell count, Age at ICU admission, and Gender.

K Updating $\phi_{1:2}$ and $\phi_{2:3}$ in stage two individual-at-a-time

The parallel sampler described in Section 3.1 is a MH-within-Gibbs scheme, with each iteration sampling from the conditionals of the melded model

$$
\begin{align*}
\pi_{\text{meld}}(\phi_{1:2}, \psi_1 \mid Y, \psi_2, \phi_{2:3}, \psi_3),
\pi_{\text{meld}}(\phi_{2:3}, \psi_3 \mid Y, \phi_{1:2}, \psi_1, \psi_2),
\pi_{\text{meld}}(\psi_2 \mid Y, \phi_{1:2}, \psi_1, \phi_{2:3}, \psi_3).
\end{align*}
$$

We would like to update the first two conditionals ‘individual-at-a-time’. For simplicity we will assume the pooled prior is formed using product-of-experts pooling, and focus on the first conditional $\pi_{\text{meld}}(\phi_{1:2}, \psi_1 \mid Y, \psi_2, \phi_{2:3}, \psi_3)$. Similar arguments apply to the second conditional and other pooling types. Recall that

$$
\phi_{1:2} = (\phi_{1:2,1}, \ldots, \phi_{1:2,N}) = ((T_1, d_1), \ldots, (T_N, d_N)),
$$

and $\psi_1 = (\psi_{1,1}, \ldots, \psi_{1,N}) = (\omega_i)_{i=1}^N$. We would like to update $\phi_{1:2}$ and $\psi_1$ by updating $(T_i, d_i, \omega_i)$ for one individual at a time, for a total of $N$ ‘sub-steps’.

Sub-step 1

Suppose we are at step $t-1$ of the Markov chain, and we are proposing values for step $t$. Without any loss of generality we assume that we are updating individual $i = 1$ in sub-step 1 – in practice we update the individuals in a random order for each iteration of the MH-within-Gibbs scheme.

Our target is

$$
\pi_1(\chi_{1,1}, \psi_1, Y_1, (\chi_{1,i}, \psi_{1,i})_{i=2}^N)p_2(\phi_{1:2,1}(\chi_{1,1}) \mid Y_2, (\phi_{1:2,i}(\chi_{1,i}))_{i=2}^N, \psi_2, \phi_{2:3}).
$$

The model, detailed in Section 5.1, uses the conditional independence between individuals to factorise such that

$$
\pi_1(\chi_1, \psi_1 \mid Y_1) = \prod_{i=1}^N \pi_1(\chi_{1,i}, \psi_{1,i} \mid Y_1),
$$

which implies\(^{11}\) $(\chi_{1,i}, \psi_{1,i} \perp \perp \chi_{1,i'}, \psi_{1,i'}) \mid Y_1$ for $i \neq i'$, hence

$$
\pi_1(\chi_{1,1}, \psi_{1,1} \mid Y_1, (\chi_{1,i}, \psi_{1,i})_{i=2}^N) = \pi_1(\chi_{1,1}, \psi_{1,1} \mid Y_1).
$$

\(^{11}\)This property is also true of $p_3$. 

Chained Markov melding

It will be convenient to rewrite (81) as

\[
p_1(\chi_{1,1}, \psi_{1,1} \mid Y_1) \frac{p_2(\phi_{1:2,1}(\chi_{1,1}), (\phi_{1:2,i}(\chi_{1,i}))_{i=2}^N, \psi_2, \phi_{2:3} \mid Y_2)}{p_2((\phi_{1:2,i}(\chi_{1,i}))_{i=2}^N, \psi_2, \phi_{2:3} \mid Y_2)},
\]  

(84)

Suppose there are \( K_1 \) stage one samples from \( p_1(\chi_1, \psi_1 \mid Y_1) \), and each sample has a corresponding \( \phi_{1:2} \). We propose \( \psi^*_{1,1} \) and \( \chi^*_{1,1} \) (hence \( \phi^*_{1:2,1} \)) by sampling a random integer \( k^*_1 \) from \{1, \ldots, K_1\}, retrieving the corresponding values of \( \psi_1 \) and \( \chi_1 \), and ignoring \( (\chi_1, \ldots, \chi_{1,N}, \psi_1, \ldots, \psi_N) \). Given stage one samples from the correct stationary distribution, obtained from the post-warmup samples of a well mixed set of Markov chains, such a proposal mechanism is approximately equivalent to proposing from \( p_1(\chi_{1,1}, \psi_{1,1} \mid Y_1) \). The quality of the approximation depends on the quality of the stage one samples. Under this proposal, noting that the denominator term in (84) does not depend on \( p_1 \), we store the value of \( k^*_1 \) associated with the proposal (if it is accepted) in order to resample the stage one values of \( \psi_{1,1} \) to reflect the information in the other submodels.

**Sub-step \( n \)**

At sub-step \( n \), for \( 1 < n \leq N \) with \( (\chi_{1,i})_{i=N+1}^N = \emptyset \), we have updated \( ((\chi_{1,i}, \psi_{1,i})_{i=1}^{n-1}) \) to \( ((\chi_{1,i}, \psi_{1,i})_{i=1}^{n-1}) \). The target is

\[
p_1(\chi_{1,n}, \psi_{1,n} \mid Y_1, (\chi_{1,i}, \psi_{1,i})_{i=1}^{n-1}, (\chi_{1,i}, \psi_{1,i})_{i=n+1}^N) \times p_2(\phi_{1:2,n}(\chi_{1,n}) \mid Y_2, \psi_2, \phi_{2:3}, (\phi_{1:2,i}(\chi_{1,i}))_{i=1}^{n-1}, (\phi_{1:2,i}(\chi_{1,i}))_{i=n+1}^N) \]

\[
= p_1(\chi_{1,n}, \psi_{1,n} \mid Y_1) \frac{p_2(\phi_{1:2,n}(\chi_{1,n}), \phi_{1:2,i}(\chi_{1,i}))_{i=1}^{n-1}, (\phi_{1:2,i}(\chi_{1,i}))_{i=n+1}^N \mid Y_2)}{p_2(\psi_2, \phi_{2:3}, (\phi_{1:2,i}(\chi_{1,i}))_{i=1}^{n-1}, (\phi_{1:2,i}(\chi_{1,i}))_{i=n+1}^N \mid Y_2)} \]

(87)
We propose $\chi^*_1, \psi^*_1, n$ in the same manner as the previous section. The corresponding acceptance probability is

$$
\alpha((\chi^*_1, \psi^*_1), (\chi_1, \psi_1, n)) = \frac{p_1(\chi^*_1, \psi^*_1 | Y_1)}{p_1(\chi_1, \psi_1, n | Y_1)} \times \frac{p_2(\phi_{1-2, n}(\chi^*_1, \psi^*_1), \phi_{2-3, n}(\chi_1, \psi_1, n) | Y_1)}{p_2(\phi_{1-2, n}(\chi_1, \psi_1, n) | Y_1)} \times \frac{p_1(\chi_1, \psi_1, n | Y_1)}{p_1(\chi^*_1, \psi^*_1 | Y_1)} = \frac{p_2(\phi_{1-2, n}(\chi^*_1, \psi^*_1), \phi_{2-3, n}(\chi_1, \psi_1, n) | Y_1)}{p_2(\phi_{1-2, n}(\chi_1, \psi_1, n) | Y_1)} \tag{88}
$$

References

Abadi, F., Gimenez, O., Ullrich, B., Arlettaz, R., and Schaub, M. (2010). “Estimation of Immigration Rate Using Integrated Population Models.” *Journal of Applied Ecology*, 47(2): 393–400. 3, 15, 17

Abbas, A. E. (2009). “A Kullback-Leibler View of Linear and Log-Linear Pools.” *Decision Analysis*. 11

Ades, A. E. and Sutton, A. J. (2006). “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. 2

Belgorodski, N., Greiner, M., Tolksdorf, K., and Schueller, K. (2017). “rriskDistributions: Fitting Distributions to given Data or Known Quantiles.” R package version 2.1.2. 32, 33

Besbeas, P., Freeman, S. N., Morgan, B. J. T., and Catchpole, E. A. (2002). “Integrating Mark–Recapture–Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters.” *Biometrics*, 58(3): 540–547. 2

Brilleman, S. (2021). “Simsurv: Simulate Survival Data.” R package version 1.0.0. 35, 36

Brilleman, S. L., Elci, E. M., Novik, J. B., and Wolfe, R. (2020). “Bayesian Survival Analysis Using the Rstanarm R Package.” *arXiv:2002.09633 [stat]*. 22

Bromiley, P. (2003). “Products and Convolutions of Gaussian Probability Density Functions.” *Tina-Vision Memo*, 3(4): 1. 15, 29

Brooks, S. P., King, R., and Morgan, B. J. T. (2004). “A Bayesian Approach to Combining Animal Abundance and Demographic Data.” *Animal Biodiversity and Conservation*, 27(1). 2
Chained Markov melding

Burke, D. L., Ensor, J., and Riley, R. D. (2017). “Meta-Analysis Using Individual Participant Data: One-Stage and Two-Stage Approaches, and Why They May Differ.” *Statistics in Medicine*, 36(5): 855–875. 14

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). “Stan: A Probabilistic Programming Language.” *Journal of Statistical Software*, 76(1): 1–32. 18

Carvalho, L. M., Villela, D. A. M., Coelho, F. C., and Bastos, L. S. (2022). “Bayesian Inference for the Weights in Logarithmic Pooling.” *Bayesian Analysis*, 1–29. 11

Crowther, M. J. and Lambert, P. C. (2013). “Simulating Biologically Plausible Complex Survival Data.” *Statistics in Medicine*, 32(23): 4118–4134. 34, 35, 36

Dawid, A. P. and Lauritzen, S. L. (1993). “Hyper Markov Laws in the Statistical Analysis of Decomposable Graphical Models.” *The Annals of Statistics*, 21(3): 1272–1317. 5

de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Lang, D. T., and Bodik, R. (2017). “Programming with Models: Writing Statistical Algorithms for General Model Structures with NIMBLE.” *Journal of Computational and Graphical Statistics*, 26(2): 403–413. 18

Donnat, C., Miolane, N., Bunbury, F., and Kreindler, J. (2020). “A Bayesian Hierarchical Network for Combining Heterogeneous Data Sources in Medical Diagnoses.” In *Proceedings of the Machine Learning for Health NeurIPS Workshop*, volume 136 of *Proceedings of Machine Learning Research*, 53–84. PMLR. 2

Finke, A., King, R., Beskos, A., and Dellaportas, P. (2019). “Efficient Sequential Monte Carlo Algorithms for Integrated Population Models.” *Journal of Agricultural, Biological and Environmental Statistics*, 24(2): 204–224. 15, 17

Gabry, J., Mahr, T., Bürkner, P.-C., Modrák, M., Barrett, M., Weber, F., Sroka, E. C., and Vehtari, A. (2021). “Bayesplot: Plotting for Bayesian Models.” R package version 1.8.0. 18

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2019). “Visualization in Bayesian Workflow.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2): 389–402. 11, 34

Gelman, A., Vehtari, A., Simpson, D., Margossian, C. C., Carpenter, B., Yao, Y., Kennedy, L., Gabry, J., Bürkner, P.-C., and Modrák, M. (2020). “Bayesian Workflow.” *arXiv:2011.01808 [stat]*. 11, 34

Genest, C., McConway, K. J., and Schervish, M. J. (1986). “Characterization of Externally Bayesian Pooling Operators.” *The Annals of Statistics*, 14(2): 487–501. 6

Giganti, M. J., Shaw, P. A., Chen, G., Bebawy, S. S., Turner, M. M., Sterling, T. R., and Shepherd, B. E. (2020). “Accounting for Dependent Errors in Predictors and Time-to-Event Outcomes Using Electronic Health Records, Validation Samples, and Multiple Imputation.” *Annals of Applied Statistics*, 14(2): 1045–1061. 27
“Joining and Splitting Models with Markov Melding.” *Bayesian Analysis*, 14(1): 81–109.

Hastie, T. and Tibshirani, R. (1999). *Generalized Additive Models*. Boca Raton, Fla: Chapman & Hall/CRC.

Hinton, G. E. (2002). “Training Products of Experts by Minimizing Contrastive Divergence.” *Neural Computation*, 14(8): 1771–1800.

Hooten, M. B., Johnson, D. S., and Brost, B. M. (2019). “Making Recursive Bayesian Inference Accessible.” *The American Statistician*, 1–10.

Jackson, D. and White, I. R. (2018). “When Should Meta-Analysis Avoid Making Hidden Normality Assumptions?” *Biometrical Journal*, 60(6): 1040–1058.

Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L.-w. H., Feng, M., Ghassemi, M., Moody, B., Szhoulovits, P., Anthony Celi, L., and Mark, R. G. (2016). “MIMIC-III, a Freely Accessible Critical Care Database.” *Scientific Data*, 3(1): 160035.

Kay, M. (2020). “Tidybayes: Tidy Data and Geoms for Bayesian Models.” R package version 2.0.2.

Kharratzadeh, M. (2017). “Splines in Stan.” *Stan Case Studies*, 4.

Lindsten, F., Johansen, A. M., Naesseth, C. A., Kirkpatrick, B., Schön, T. B., Aston, J. A. D., and Bouchard-Côté, A. (2017). “Divide-and-Conquer with Sequential Monte Carlo.” *Journal of Computational and Graphical Statistics*, 26(2): 445–458.
Chained Markov melding

Lu, C. J. and Meeker, W. Q. (1993). “Using Degradation Measures to Estimate a Time-to-Failure Distribution.” Technometrics, 35(2): 161–174. 4

Lunn, D., Barrett, J., Sweeting, M., and Thompson, S. (2013). “Fully Bayesian Hierarchical Modelling in Two Stages, with Application to Meta-Analysis.” Journal of the Royal Statistical Society Series C, 62(4): 551–572. 3, 11

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). “The BUGS Project: Evolution, Critique and Future Directions.” Statistics in Medicine, 28(25): 3049–3067. 18

Manderson, A. A. and Goudie, R. J. B. (2022). “A Numerically Stable Algorithm for Integrating Bayesian Models Using Markov Melding.” Statistics and Computing, 32(2): 24. 27

Massa, M. S. and Lauritzen, S. L. (2010). “Combining Statistical Models.” In Algebraic Methods in Statistics and Probability II, volume 516 of Contemp. Math., 239–259. Amer. Math. Soc., Providence, RI. 5

Mauff, K., Steyerberg, E., Kardys, I., Boersma, E., and Rizopoulos, D. (2020). “Joint Models with Multiple Longitudinal Outcomes and a Time-to-Event Outcome: A Corrected Two-Stage Approach.” Statistics and Computing, 30(4): 999–1014. 3, 4, 11, 28

Maunder, M. N. and Punt, A. E. (2013). “A Review of Integrated Analysis in Fisheries Stock Assessment.” Fisheries Research, 142: 61–74. 2

Meng, X.-L. (2014). “A Trio of Inference Problems That Could Win You a Nobel Prize in Statistics (If You Help Fund It).” In Lin, X., Genest, C., Banks, D. L., Molenberghs, G., Scott, D. W., and Wang, J.-L. (eds.), Past, Present, and Future of Statistical Science, 561–586. Chapman and Hall/CRC. 3

Nelsen, R. B. (2006). An Introduction to Copulas. Springer New York, second edition. 27

Nicholson, G., Blangiardo, M., Briers, M., Diggle, P. J., Fjelde, T. E., Ge, H., Goudie, R. J. B., Jersakova, R., King, R. E., Lehmann, B. C. L., Mallon, A.-M., Padellini, T., Teh, Y. W., Holmes, C., and Richardson, S. (2021). “Interoperability of Statistical Models in Pandemic Preparedness: Principles and Reality.” Statistical Science (forthcoming). 3

NIMBLE Development Team (2019). “NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling.” R package manual version 0.9.0. 18

Oh, E. J., Shepherd, B. E., Lumley, T., and Shaw, P. A. (2018). “Considerations for Analysis of Time-to-Event Outcomes Measured with Error: Bias and Correction with SIMEX.” Statistics in medicine, 37(8): 1276–1289. 27

Oh — (2021). “Raking and Regression Calibration: Methods to Address Bias from Correlated Covariate and Time-to-Event Error.” Statistics in Medicine, 40(3): 631–649. 27
O’Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, P., Jenkinson, D., Oakley, J., and Rakow, T. (2006). Uncertain Judgements: Eliciting Experts’ Probabilities. Statistics in Practice. Wiley. 6

Parsons, J., Niu, X., and Bao, L. (2021). “A Bayesian Hierarchical Modeling Approach to Combining Multiple Data Sources: A Case Study in Size Estimation.” arXiv:2012.05346 [stat]. 2

Plummer, M. (2019). “Rjags: Bayesian Graphical Models Using MCMC.” R package version 4-10. 18

Presanis, A. M., Pebody, R. G., Birrell, P. J., Tom, B. D. M., Green, H. K., Durnall, H., Fleming, D., and De Angelis, D. (2014). “Synthesising Evidence to Estimate Pandemic (2009) A/H1N1 Influenza Severity in 2009–2011.” Annals of Applied Statistics, 8(4): 2378–2403. 2

Rizopoulos, D. (2012). Joint Models for Longitudinal and Time-to-Event Data: With Applications in R. CRC Press. 4

Rosenberg, P. S. (1995). “Hazard Function Estimation Using B-splines.” Biometrics, 51(3): 874–887. 28

Royston, P. and Parmar, M. K. B. (2002). “Flexible Parametric Proportional-Hazards and Proportional-Odds Models for Censored Survival Data, with Application to Prognostic Modelling and Estimation of Treatment Effects.” Statistics in Medicine, 21(15): 2175–2197. 28

Rufo, M. J., Martín, J., and Pérez, C. J. (2012a). “Log-Linear Pool to Combine Prior Distributions: A Suggestion for a Calibration-Based Approach.” Bayesian Analysis, 7(2): 411–438. 11

Rufo, M. J., Pérez, C. J., and Martín, J. (2012b). “A Bayesian Approach to Aggregate Experts’ Initial Information.” Electronic Journal of Statistics, 6(none): 2362–2382. 11

Rutherford, M. J., Crowther, M. J., and Lambert, P. C. (2015). “The Use of Restricted Cubic Splines to Approximate Complex Hazard Functions in the Analysis of Time-to-Event Data: A Simulation Study.” Journal of Statistical Computation and Simulation, 85(4): 777–793. 28

Schaub, M. and Abadi, F. (2011). “Integrated Population Models: A Novel Analysis Framework for Deeper Insights into Population Dynamics.” Journal of Ornithology, 152(1): 227–237. 2

Schaub, M., Ullrich, B., Knötzsch, G., Albrecht, P., and Meisser, C. (2006). “Local Population Dynamics and the Impact of Scale and Isolation: A Study on Different Little Owl Populations.” Oikos, 115(3): 389–400. 3, 15

Seethala, R. R., Hou, P. C., Aisiku, I. P., Frendl, G., Park, P. K., Mikkelsen, M. E., Chang, S. Y., Gajic, O., and Sevransky, J. (2017). “Early Risk Factors and the Role of Fluid Administration in Developing Acute Respiratory Distress Syndrome in Septic Patients.” Annals of Intensive Care, 7(1): 11. 21
Chained Markov melding

Soetaert, K., Hindmarsh, A. C., Eisenstat, S. C., Moler, C., Dongarra, J., and Saad, Y. (2020). “Rootsolve: Nonlinear Root Finding, Equilibrium and Steady-State Analysis of Ordinary Differential Equations.” R package version 1.8.2.1. 20

Stan Development Team (2021). “RStan: The R Interface to Stan.” R package version 2.26. 18

The ARDS Definition Task Force (2012). “Acute Respiratory Distress Syndrome: The Berlin Definition.” JAMA, 307(23): 2526–2533. 4

Tom, J. A., Sinsheimer, J. S., and Suchard, M. A. (2010). “Reuse, Recycle, Reweigh: Combating Influenza through Efficient Sequential Bayesian Computation for Massive Data.” The Annals of Applied Statistics, 4(4): 1722–1748. 11

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2020). “Rank-Normalization, Folding, and Localization: An Improved $\hat{R}$ for Assessing Convergence of MCMC.” Bayesian Analysis. 18, 23

Wang, W., Aseltine, R., Chen, K., and Yan, J. (2020). “Integrative Survival Analysis with Uncertain Event Times in Application to a Suicide Risk Study.” Annals of Applied Statistics, 14(1): 51–73. 27

Wang, W. and Yan, J. (2021). “Shape-Restricted Regression Splines with R Package Splines2.” Journal of Data Science, 19(3): 498–517. 20

Zipkin, E. F. and Saunders, S. P. (2018). “Synthesizing Multiple Data Types for Biological Conservation Using Integrated Population Models.” Biological Conservation, 217: 240–250. 2, 3