Supplemental Material 1: Model and Derivations

Bayesian Hierarchical Model for Immune Responses to Leishmania - a tick borne Co-Infection Study

Felix M. Pabon-Rodriguez, Grant D. Brown, Breanna M. Scorza, Christine A. Petersen

1 Qualitative Categories for Disease Status

To assess CanL disease progression of subjects over time, we took into account LeishVet stage, pathogen load, and level of anti-*Leishmania* antibodies as explained in the paper. Each dog was classified based on the scoring proposed by Solano-Gallego et al. [1]. We further aggregate this scoring into the qualitative categories described below. For ith dog at time $t+1$, we define disease state $D_{i,t+1}$ as follows,

$$D_{i,t+1} = \begin{cases} 1\,(\text{Healthy}), & \text{if } LeishVet = 0 \text{ or } 1 \\ 2\,(\text{Asymptomatic}), & \text{if } LeishVet = 2 \\ 3\,(\text{Symptomatic}), & \text{if } LeishVet = 3 \text{ or } 4 \\ 4\,(\text{Removed}), & \text{if removed due to severe case of Leishmaniasis.} \end{cases} \tag{1}$$

2 Bayesian Hierarchical Model

For $N = 50$ dogs, and $T = 7$ time points, the proposed model is structured as follows

$$\begin{aligned} f_D(D_{i,t}, P_{i,t}, A_{i,t}, X_i) &= \left[\pi_{i,t}^{(1)}, \pi_{i,t}^{(2)}, \pi_{i,t}^{(3)}, \pi_{i,t}^{(4)}\right] \\ \pi_{i,t}^{(k)} &= \frac{exp\Big[M_{i,t}^{D(k)}\boldsymbol{\beta}_D^{(k)} + X_i\boldsymbol{\alpha}_D\Big]}{1 + \sum_{g=1,2,3} exp\Big[M_{i,t}^{D(g)}\boldsymbol{\beta}_D^{(g)} + X_i\boldsymbol{\alpha}_D\Big]} \\ \pi_{i,t}^{(4)} &= 1 - \pi_{i,t}^{(1)} - \pi_{i,t}^{(2)} - \pi_{i,t}^{(3)} \\ f_A(D_{i,t}, P_{i,t}, A_{i,t}, X_i) &= M_{i,t}^{A}\boldsymbol{\beta}_A + X_i\boldsymbol{\alpha}_A \\ f_P(D_{i,t}, P_{i,t}, A_{i,t}, X_i) &= M_{i,t}^{P}\boldsymbol{\beta}_P + X_i\boldsymbol{\alpha}_P \end{aligned} \tag{2}$$

$$\begin{aligned} D_{i,t+1} &\sim \text{Multinomial}\big(1; f_D(D_{i,t}, P_{i,t}, A_{i,t}, X_i)\big) \\ A_{i,t+1} &\sim \mathcal{N}\big(f_A(D_{i,t}, P_{i,t}, A_{i,t}, X_i),\, \sigma_A^2\big) \\ P_{i,t+1} &\sim \mathcal{N}\big(f_P(D_{i,t}, P_{i,t}, A_{i,t}, X_i),\, \sigma_P^2\big) \end{aligned} \tag{3}$$

for $k = 1, 2, 3, 4$ in the probability terms, and $i = 1, \ldots, N$ and $t = 1, \ldots, T$.

3 Derivations

3.1 Complete Data Likelihood

Using the model definition as described by Equations 2 and 3, the joint likelihood can be defined by the product of the each probability density (or probability mass) function corresponding to each model components. In this case, we have two continuous outcomes (pathogen load and antibody levels) and one categorical (disease status). To facilitate the definition of the likelihood presented in Equation 4, let us define $\boldsymbol{\theta}_P = (\boldsymbol{\beta}_P, \boldsymbol{\alpha}_P, \sigma_P^2)'$ as the set of vectors associated with pathogen load. Similarly, we can define the set of parameters associated with the antibody levels and disease status as $\boldsymbol{\theta}_A = (\boldsymbol{\beta}_A, \boldsymbol{\alpha}_A, \sigma_A^2)'$ and $\boldsymbol{\theta}_D = (\boldsymbol{\beta}_D^{(1)}, \boldsymbol{\beta}_D^{(2)}, \boldsymbol{\beta}_D^{(3)} \boldsymbol{\alpha}_D)'$, respectively.

$$
\begin{aligned}
\mathcal{L}(\theta_P, \theta_A, \theta_D | P, A, D) = \prod_{i=1}^{N} \Big[& f(P_{i,1}) \cdot f(A_{i,1}) \cdot f(D_{i,1}) \\
& \cdot \prod_{t=2}^{T} \Big\{ f(P_{i,t} | P_{i,t-1}, \theta_P) \cdot f(A_{i,t} | A_{i,t-1}, \theta_A) \cdot f(D_{i,t} | D_{i,t-1}, \theta_D) \Big\} \Big] \\
\propto \prod_{i=1}^{N} \Bigg[& f(P_{i,1}) \cdot f(A_{i,1}) \cdot f(D_{i,1}) \prod_{t=2}^{T} \Bigg\{ \Big(\frac{1}{\sigma_P} exp\Big\{ - \frac{1}{2\sigma_P^2} (P_{i,t} - \eta_{i,t-1}^P)^2 \Big\} \Big) \\
& \cdot \Big(\frac{1}{\sigma_A} exp\Big\{ - \frac{1}{2\sigma_A^2} (A_{i,t} - \eta_{i,t-1}^A)^2 \Big\} \Big) \\
& \cdot \Big({\pi_{i,t-1}^{(1)}}^{D_{i,t-1}^{(1)}} \cdot {\pi_{i,t-1}^{(2)}}^{D_{i,t-1}^{(2)}} \cdot {\pi_{i,t-1}^{(3)}}^{D_{i,t-1}^{(3)}} \cdot {\pi_{i,t-1}^{(4)}}^{1 - \sum_{g=1}^{3} D_{i,t-1}^{(g)}} \Big) \Bigg\} \Bigg]
\end{aligned}
\tag{4}
$$

Here we have that the mean expressions for each normal density distribution are defined as $\eta_{i,t-1}^P = M_{i,t}^P \boldsymbol{\beta}_P + X_i \boldsymbol{\alpha}_P$, and $\eta_{i,t-1}^A = M_{i,t}^A \boldsymbol{\beta}_A + X_i \boldsymbol{\alpha}_A$, respectively. The expressions for the $\pi_{i,t-1}^{(k)}$ are presented in Equation 2, and where $D_{i,t-1}^{(k)} = 1$ for $k = 1, 2, 3$ if $D_{i,t-1} = k$ and $D_{i,t-1}^{(k)} = 0$ otherwise, which are defined based in Equation 1. Note that a proportionality notation is used being in the second line of the likelihood, which helps on factoring out all of the constants or fixed quantities. In addition, notice that we defined our baseline category in the multinomial distribution as $D_{i,t-1}^{(4)} = 1 - \sum_{g=1}^{3} D_{i,t-1}^{(g)}$.

3.2 Full Conditionals (Model Components)

Since pathogen load, antibodies level, and disease status were not observed for some of the time points, which is known as latent variables, then we estimated the corresponding components of the model for those time points. Therefore, the full conditional for the three main components of the model are given below.

3.2.1 Disease Status

$$
\begin{aligned}
f_c(D | P, A, \theta_P, \theta_A, \theta_D) \propto \prod_{i=1}^{N} \Bigg[& f(P_{i,1}) f(A_{i,1}) f(D_{i,1}) \prod_{t=2}^{T} \Bigg\{ exp\Big\{ - \frac{1}{2\sigma_P^2} (P_{i,t} - \eta_{i,t-1}^P)^2 \Big\} \\
& \cdot \Big({\pi_{i,t-1}^{(1)}}^{D_{i,t-1}^{(1)}} \cdot {\pi_{i,t-1}^{(2)}}^{D_{i,t-1}^{(2)}} \cdot {\pi_{i,t-1}^{(3)}}^{D_{i,t-1}^{(3)}} \cdot {\pi_{i,t-1}^{(4)}}^{1 - D_{i,t-1}^{(1)} - D_{i,t-1}^{(2)} - D_{i,t-1}^{(3)}} \Big) \Bigg\} \Bigg]
\end{aligned}
\tag{5}
$$

3.2.2 Pathogen Load

$$
f_c(P | A, D, \theta_P, \theta_A, \theta_D) \propto \prod_{i=1}^{N} \Bigg[f(P_{i,1}) f(A_{i,1}) f(D_{i,1}) \prod_{t=2}^{T} \Bigg\{ exp\Big\{ - \frac{1}{2\sigma_P^2} (P_{i,t} - \eta_{i,t-1}^P)^2 \Big\} \Bigg] \tag{6}
$$

3.2.3 Antibodies Level

$$f_c(A|P,D,\theta_P,\theta_A,\theta_D) \propto \prod_{i=1}^{N}\left[f(P_{i,1})f(A_{i,1})f(D_{i,1})\prod_{t=2}^{T}\left\{exp\Big\{-\frac{1}{2\sigma_A^2}(A_{i,t}-\eta_{i,t-1}^{A})^2\Big\}\right\}\right] \tag{7}$$

3.3 Full Conditionals (Parameters)

If ω_1 is equal to one of the parameters in $\{\beta_P, \alpha_P, \sigma_P^2\}$, then the full conditional is given by

$$f_c(\omega_1|.) \propto \prod_{i=1}^{N}\left[f(P_{i,1})f(A_{i,1})f(D_{i,1})\prod_{t=2}^{T}\left\{exp\Big\{-\frac{1}{2\sigma_P^2}(P_{i,t}-\eta_{i,t-1}^{P})^2\Big\}\right\}\right]\cdot\pi(\omega_1) \tag{8}$$

If ω_2 is equals to one of the parameters in $\{\beta_A, \alpha_A, \sigma_A^2\}$, then the full conditional is given by

$$f_c(\omega_2|.) \propto \prod_{i=1}^{N}\left[f(P_{i,1})f(A_{i,1})f(D_{i,1})\prod_{t=2}^{T}\left\{exp\Big\{-\frac{1}{2\sigma_A^2}(A_{i,t}-\eta_{i,t-1}^{A})^2\Big\}\right\}\right]\cdot\pi(\omega_2) \tag{9}$$

Finally, if ω_3 is equals to one of the parameters in $\{\beta_D^{(k)}, \alpha_D\}$ for $k = 1, 2, 3$, then the full conditional is given by

$$\begin{aligned} f_c(\omega_3|.) \propto \prod_{i=1}^{N}\Bigg[f(P_{i,1})f(A_{i,1})f(D_{i,1})\Bigg\{\prod_{t=2}^{T}\Big(&{\pi_{i,t-1}^{(1)}}^{D_{i,t-1}^{(1)}}\cdot{\pi_{i,t-1}^{(2)}}^{D_{i,t-1}^{(2)}}\cdot{\pi_{i,t-1}^{(3)}}^{D_{i,t-1}^{(3)}} \\ &\cdot{\pi_{i,t-1}^{(4)}}^{1-D_{i,t-1}^{(1)}-D_{i,t-1}^{(2)}-D_{i,t-1}^{(3)}}\Big)\Bigg\}\Bigg]\cdot\pi(\omega_3) \end{aligned} \tag{10}$$

References

[1] Solano-Gallego L, Miró G, Koutinas A, et al. LeishVet guidelines for the practical management of canine leishmaniosis. Parasites & Vectors. 2011;4(86). doi:https://doi.org/10.1186/1756-3305-4-86.