

1 **Appendix A: Gibbs sampling procedure based on Pólya-Gamma** 2 **expansion**

3 **A.1 Random effects model**

4 Bayesian analysis relies on the posterior distribution of regression coefficients, which
5 is proportional to the likelihood of the data and the prior distribution:

$$p(\gamma_j^q, \gamma^q, \Sigma^q | \mathbf{y}) \propto p(\mathbf{y} | \gamma_j^q) p(\gamma_j^q | \gamma^q, \Sigma^q) p(\gamma^q) p(\Sigma^q). \quad (1)$$

6 The multinomial logistic likelihood (Equation 2) can be expanded with a Pólya-
7 Gamma auxiliary variable to suit a Gibbs sampling procedure. This expansion relies
8 on the following equality [1]:

$$\begin{aligned} p((\mathbf{y}_j = \mathbf{h}^q) | \gamma_j^q, \gamma_j^{-q}, \omega_j^q) &= \frac{\exp(\mathbf{x}_{ji} \gamma_j^q)}{\sum_{r=1}^{Q-1} \exp(\mathbf{x}_{ji} \gamma_j^r) + 1}, \\ &\propto \exp \left[-\frac{1}{2} (\boldsymbol{\kappa}_j^q - \boldsymbol{\eta}_j^q)^T \boldsymbol{\Omega}_j^q (\boldsymbol{\kappa}_j^q - \boldsymbol{\eta}_j^q) \right], \end{aligned} \quad (2)$$

9 where $\mathbf{X}_j$ is a matrix filled with n_j rows of covariate vectors $\mathbf{x}_{ji}$ and $\boldsymbol{\eta}_j^q = \mathbf{X}_j \gamma_j^q -$
10 $\ln \left[\sum_{m \neq q} \exp(\mathbf{X}_j \gamma_j^m) \right]$, $\boldsymbol{\kappa}_j^q = \frac{I(\mathbf{y}_j = \mathbf{h}^q) - \frac{1}{2}}{\omega_j^q}$.

11 Equation 2 can be recognized as the kernel of a multivariate Gaussian likelihood
12 of working variable $\boldsymbol{\kappa}_j^q$ [1]:

$$\boldsymbol{\kappa}_j^q \sim N(\boldsymbol{\eta}_j^q, \{\boldsymbol{\Omega}_j^q\}^{-1}) \quad (3)$$

13 Here, $\boldsymbol{\Omega}_j^q$ reflects the diagonal matrix of Pólya-Gamma distributed variables $\omega_j^q =$
14 $(\omega_{j1}^q, \dots, \omega_{jn_j}^q)$. A Gibbs sampler can be constructed when the likelihood in Equation
15 3 is combined with multivariate normal prior distributions on random regression
16 coefficients $\gamma_j^q | \gamma^q, \Sigma^q$ and mean random regression coefficients γ^q , and an inverse-
17 Wishart prior distribution on covariance matrix Σ^q :

$$\begin{aligned} \gamma_j^q &\sim N(\gamma^q, \Sigma^q) \\ \gamma^q &\sim N(\mathbf{g}^q, \mathbf{G}^q) \\ \Sigma^q &\sim \mathcal{W}^{-1}(j^0, \mathbf{S}^q) \end{aligned} \quad (4)$$

18 The resulting Gibbs sampler consists of the following steps:

1 Sample mean regression coefficients:

$$\gamma^{q(l)} \sim N \left(\mathbf{V}_\gamma^q (\{\Sigma^{q(l-1)}\}^{-1} \sum_{j=1}^J \gamma_j^{q(l-1)} + \mathbf{G}^q \mathbf{g}^q), \mathbf{V}_\gamma^q \right)$$

19 with prior mean vector $\mathbf{g}^q$, prior precision matrix $\mathbf{G}^q$ and posterior variance
20 matrix $\mathbf{V}_\gamma = (J \{\Sigma^{q(l-1)}\}^{-1} + \mathbf{G}^q)^{-1}$.

2 Sample covariance matrices of regression coefficients:

$$\Sigma^{q(l)} \sim \mathcal{W}^{-1} \left(j^0 + J, \mathbf{S}^q + \sum_{j=1}^J \left(\gamma_j^{q(l)} - \gamma^{q(l-1)} \right) \left(\gamma_j^{q(l)} - \gamma^{q(l-1)} \right)^T \right)$$

- 1 with prior hyperparameters $j^0 \geq P$ and $\mathbf{S}^q$.
 2 3 For each j , sample random regression coefficients:

$$\gamma_j^{q(l)} \sim N \left(\mathbf{V}_{\gamma_j^q}^q (\mathbf{X}_j \Omega_j^{q(l-1)} (\boldsymbol{\kappa}_j^{q(l-1)} + \ln[\sum_{m \neq q} \exp(\mathbf{X}_j \gamma_j^m)]) + \{\Sigma^{q(l)}\}^{-1} \gamma^{q(l)}), \mathbf{V}_{\gamma_j^q}^q \right)$$

- 3 with prior mean vector $\gamma^{q(l)}$, prior precision matrix $\Sigma^{q(l)}$, posterior variance
 4 matrix $\mathbf{V}_{\gamma_j^q}^q = (\mathbf{X}_j^T \Omega_j^{q(l-1)} \mathbf{X}_j + \{\Sigma^{q(l)}\}^{-1})^{-1}$, and diagonal matrix of Pólya-
 5 Gamma variables $\Omega_j^{q(l-1)} = \text{diag}(\omega_{j1}^{q(l-1)}, \dots, \omega_{jn_j}^{q(l-1)})$.
 4 For each j and i , sample Pólya-Gamma variables:

$$\omega_{ji}^{q(l)} \sim PG(1, \eta_{ji}^{q(l)})$$

- 6 The remainder of this section shows the derivations of the full conditional distribu-
 7 tions.

8 A.1.1 Deriving the likelihood function

- 9 The following equality forms the basis to rewrite the multinomial likelihood in
 10 Equation 2 as a Gaussian likelihood [1]:

$$\begin{aligned} p((\mathbf{y}_j = \mathbf{h}^q) | \gamma_j, \omega_j^q, \mathbf{x}_j) &= \frac{\exp(\mathbf{x}_{ji} \gamma_j^q)}{\sum_{r=1}^{Q-1} \exp(\mathbf{x}_{ji} \gamma_j^r) + 1}, \\ &= \prod_{i=1}^{n_j} 2 \exp \left[\kappa_{ji}^q \omega_{ji}^q \eta_{ji}^q \right] \int_0^\infty \exp \left[\frac{-\omega_{ji}^q (\eta_{ji}^q)^2}{2} \right] p(\omega_{ji}^q) d\omega_{ji}^q \end{aligned} \quad (5)$$

- 11 where $\omega_{ji}^q \sim PG(1, \eta_{ji}^q)$ is a Pólya-Gamma distributed variable,
 12 where $\eta_{ji}^q = \mathbf{x}_{ji} \gamma_j^q - \ln \left[\sum_{m \neq q} \exp(\mathbf{x}_{ji} \gamma_j^m) \right]$,
 13 and where working variable $\kappa_j^q = \frac{I(\mathbf{y}_j = \mathbf{h}^q) - \frac{1}{2}}{\omega_j^q}$.

1 Further algebraic transformation results in the kernel of a Gaussian likelihood:

$$\begin{aligned}
 p((\mathbf{y}_j = \mathbf{h}^q)|\cdot) &= \prod_{i=1}^{n_j} 2 \exp \left[\kappa_{ji}^q \omega_{ji}^q \eta_{ji}^q \right] \int_0^\infty \exp \left[\frac{-\omega_{ji}^q (\eta_{ji}^q)^2}{2} \right] p(\omega_{ji}^q) d\omega_{ji}^q \quad (6) \\
 &\propto \exp \left[\frac{1}{2} (\kappa_j^q \omega_j^q \eta_j^q - \omega_j^q (\eta_j^q)^2) \right] \\
 &\propto \exp \left[-\frac{1}{2} (\kappa_j^q - \eta_j^q)^T \Omega_j^q (\kappa_j^q - \eta_j^q) \right],
 \end{aligned}$$

2 Hence, working variable κ_j^q is multivariate normally distributed:

$$\kappa_j^q \sim N(\eta_j^q, \{\Omega_j^q\}^{-1}). \quad (7)$$

3 *A.1.2 Deriving conditional posterior distributions*

4 *Random regression coefficients γ_j^q* Using the likelihood in Equation 7 and prior
 5 distribution $\gamma_j^q \sim N(\gamma^q, \{\Sigma^q\})$, the conditional posterior distribution of random
 6 regression coefficients γ_j^q is also a multivariate normal distribution:

$$\begin{aligned}
 p(\gamma_j^q|\cdot) &\propto p(\mathbf{y}_j|\gamma_j^q, \gamma_j^{-q}, \omega_j^q, \mathbf{x}) p(\gamma_j^q) \quad (8) \\
 &\propto \exp \left[-\frac{1}{2} (\kappa_j^q - (\eta_j^q))^T \Omega_j^q (\kappa_j^q - (\eta_j^q)) \right] \times \\
 &\quad \exp \left[-\frac{1}{2} (\gamma_j^q - \gamma^q)^T \{\Sigma^q\}^{-1} (\gamma_j^q - \gamma^q) \right] \\
 &\propto \exp \left[-\frac{1}{2} \left(\{\gamma_j^q\}^T (\{\mathbf{X}_j\}^T \Omega_j^q \mathbf{X}_j + \{\Sigma^q\}^{-1}) \gamma_j^q - 2\{\gamma_j^q\}^T \right. \right. \\
 &\quad \left. \left. (\{\mathbf{X}_j\}^T \Omega_j^q \kappa_j^q + \ln[\sum_{m \neq q} \exp(\mathbf{X}_j \gamma_j^m)]) + \{\Sigma^q\}^{-1} \gamma^q \right) \right] \\
 &\propto \exp \left[-\frac{1}{2} \right. \\
 &\quad \left(\gamma_j^q - \mathbf{V}_{\gamma_j}^q (\{\mathbf{X}_j\}^T \Omega_j^q \kappa_j^q + \ln[\sum_{m \neq q} \exp(\mathbf{X}_j \gamma_j^m)]) + \{\Sigma^q\}^{-1} \gamma^q \right)^T \\
 &\quad \left. \{\mathbf{V}_{\gamma_j}^q\}^{-1} \right. \\
 &\quad \left. \left(\gamma_j^q - \mathbf{V}_{\gamma_j}^q (\{\mathbf{X}_j\}^T \Omega_j^q \kappa_j^q + \ln[\sum_{m \neq q} \exp(\mathbf{X}_j \gamma_j^m)]) + \{\Sigma^q\}^{-1} \gamma^q \right) \right] \\
 &\sim N \left(\mathbf{V}_{\gamma_j}^q (\mathbf{X}_j \Omega_j^q \kappa_j^q + \ln[\sum_{m \neq q} \exp(\mathbf{X}_j \gamma_j^m)]) + \{\Sigma^q\}^{-1} \gamma^q, \mathbf{V}_{\gamma_j}^q \right)
 \end{aligned}$$

7 with prior mean vector γ^q , prior variance matrix Σ^q and posterior variance matrix

8 $\mathbf{V}_{\gamma_j}^q = (\mathbf{X}_j^T \Omega_j^q \mathbf{X}_j + \{\Sigma^q\}^{-1})^{-1}$.

1 *Random mean γ^q* When the posterior distribution of γ_j^q (Equation 8) is included
 2 as a likelihood and combined with a $N(\mathbf{g}^q, \{\mathbf{G}^q\}^{-1})$ prior distribution, the con-
 3 ditional posterior distribution of random mean γ^q is another multivariate normal
 4 distribution:

$$\begin{aligned}
 p(\gamma^q | \cdot) &\propto \prod_{j=1}^J p(\gamma_j^q | \gamma^q, \Sigma^q) p(\gamma^q) \\
 &\propto \prod_{j=1}^J \exp \left[-\frac{1}{2} (\gamma_j^q - \gamma^q)^T \{\Sigma^q\}^{-1} (\gamma_j^q - \gamma^q) \right] \times \\
 &\quad \exp \left[-\frac{1}{2} (\gamma^q - \mathbf{g}^q)^T \mathbf{G}^q (\gamma^q - \mathbf{g}^q) \right] \\
 &\propto \exp \left[-\frac{1}{2} (\{\gamma^q\}^T (J\{\Sigma^q\}^{-1}) \gamma^q) - 2\{\gamma^q\}^T \left(\{\Sigma^q\}^{-1} \sum_{j=1}^J \gamma_j^q \right) \right] \times \\
 &\quad \exp \left[-\frac{1}{2} \{\gamma^q\}^T \mathbf{G}^q \gamma^q - 2\{\gamma^q\}^T \mathbf{G}^q \mathbf{g}^q \right] \\
 &\propto \exp \left[-\frac{1}{2} \{\gamma^q\}^T (J\{\Sigma^q\}^{-1} + \mathbf{G}^q) \gamma^q - \right. \\
 &\quad \left. 2\{\gamma^q\}^T \left(\{\Sigma^q\}^{-1} \sum_{j=1}^J \gamma_j^q + \mathbf{G}^q \mathbf{g}^q \right) \right] \\
 &\propto \exp \left[-\frac{1}{2} \left(\gamma^q - \mathbf{V}_\gamma^q \left(\{\Sigma^q\}^{-1} \sum_{j=1}^J \gamma_j^q + \mathbf{G}^q \mathbf{g}^q \right) \right)^T \{\mathbf{V}_\gamma^q\}^{-1} \right. \\
 &\quad \left. \left(\gamma^q - \mathbf{V}_\gamma^q \left(\{\Sigma^q\}^{-1} \sum_{j=1}^J \gamma_j^q + \mathbf{G}^q \mathbf{g}^q \right) \right) \right] \\
 &\sim N \left(\mathbf{V}_\gamma^q \left(\{\Sigma^q\}^{-1} \sum_{j=1}^J \gamma_j^q + \mathbf{G}^q \mathbf{g}^q \right), \mathbf{V}_\gamma^q \right),
 \end{aligned} \tag{9}$$

5 with prior mean vector $\mathbf{g}^q$, prior precision matrix $\mathbf{G}^q$, and posterior variance matrix
 6 $\mathbf{V}_\gamma^q = (J\{\Sigma^q\}^{-1} + \mathbf{G}^q)^{-1}$.

7 *Random variance Σ^q* When the posterior distribution of γ_j^q (Equation 8) is
 8 included as a likelihood and combined with an inverse Wishart $\mathcal{W}^{-1}(j^0, \mathbf{S}^q)$ prior,
 9 the conditional posterior distribution of random variance Σ^q is proportional to an

1 inverse Wishart distribution:

$$\begin{aligned}
 p(\Sigma^q | \cdot) &\propto p(\gamma_j^q | \gamma^q, \Sigma^q) p\{\Sigma^q\} \\
 &\propto \prod_{j=1}^J |\Sigma^q|^{\frac{1}{2}} \exp \left[-\frac{1}{2} (\gamma_j^q - \gamma^q)^T \{\Sigma^q\}^{-1} (\gamma_j^q - \gamma^q) \right] \times \\
 &\quad |\Sigma^q|^{\frac{1}{2}(j^0 + p^R + 1)} \exp \left[-\frac{1}{2} \text{tr}(\mathbf{S}^q \{\Sigma^q\}^{-1}) \right] \\
 &\propto |\Sigma^q|^{-\frac{1}{2}(j^0 + J + P^R + 1)} \times \\
 &\quad \exp \left[-\frac{1}{2} \text{tr} \left(\left(\mathbf{S}^q + \sum_{j=1}^J (\gamma_j^q - \gamma^q)(\gamma_j^q - \gamma^q)^T \right) \{\Sigma^q\}^{-1} \right) \right] \\
 &\sim \mathcal{W}^{-1} \left(j^0 + J, \mathbf{S}^q + \sum_{j=1}^J (\gamma_j^q - \gamma^q)(\gamma_j^q - \gamma^q)^T \right).
 \end{aligned} \tag{10}$$

2 A.2 Mixed effects model

3 A mixed effect model is defined as follows:

$$\phi_{ji}^q = f(\mathbf{x}_{ji}^F \boldsymbol{\beta}^q + \mathbf{x}_{ji}^R \boldsymbol{\gamma}_j^q) \tag{11}$$

4 where $\mathbf{x}_{ji}^F$ and $\mathbf{x}_{ji}^R$ are vectors of fixed and random covariates respectively. Vectors $\boldsymbol{\beta}^q$
 5 and $\boldsymbol{\gamma}_j^q$ reflect the accompanying fixed and random regression coefficients. Function
 6 f refers to the multinomial logistic likelihood function.

7 The multivariate normal distribution of working variable $\boldsymbol{\kappa}_j^q$ then has the following
 8 form:

$$\boldsymbol{\kappa}_j^q \sim N(\boldsymbol{\eta}_j^q, \{\boldsymbol{\Omega}_j^q\}^{-1}). \tag{12}$$

9 Here, $\boldsymbol{\eta}_j^q = \mathbf{X}_j^F \boldsymbol{\beta}^q + \mathbf{X}_j^R \boldsymbol{\gamma}_j^q - \ln \left[\sum_{m \neq q} \exp(\mathbf{X}_j^F \boldsymbol{\beta}^m + \mathbf{X}_j^R \boldsymbol{\gamma}_j^m) \right]$. The likelihood
 10 in Equation 12 can be combined with the prior distributions in Equation 4,
 11 complemented with a multivariate normally distributed prior on $\boldsymbol{\beta}^q$:

$$\boldsymbol{\beta}^q \sim N(\mathbf{b}^q, \mathbf{B}^q) \tag{13}$$

12 The Gibbs sampling algorithm in list A.1 is extended with a distinct step for the
 13 fixed regression coefficients:

14 1 Sample fixed regression coefficients:

$$\begin{aligned}
 \boldsymbol{\beta}^{q(l)} &\sim N \left(\mathbf{V}_{\boldsymbol{\beta}}^q \left(\sum_{j=1}^J \mathbf{X}_j^{FT} \boldsymbol{\Omega}_j^{q(l-1)} (\boldsymbol{\kappa}_j^{q(l-1)} - \mathbf{X}_j^R \boldsymbol{\gamma}_j^{q(l)} + \right. \right. \\
 &\quad \left. \left. \ln \left[\sum_{m \neq q} \exp(\mathbf{X}_j^F \boldsymbol{\beta}^{m(l)} + \mathbf{X}_j^R \boldsymbol{\gamma}_j^{m(l-1)}) \right] \right) + \mathbf{B}^q \mathbf{b}^q \right), \mathbf{V}_{\boldsymbol{\beta}}^q \right)
 \end{aligned}$$

1 with prior mean vector $\mathbf{b}^q$, prior precision matrix $\mathbf{B}^q$ and posterior variance
 2 matrix $\mathbf{V}_\beta^q = (\sum_{j=1}^J \mathbf{X}_j^{FT} \boldsymbol{\Omega}_j^{q(l-1)} \mathbf{X}_j^F + \mathbf{B}^q)^{-1}$.

2 Sample mean random regression coefficients:

$$\gamma^{q(l)} \sim N \left(\mathbf{V}_\gamma^q (\{\boldsymbol{\Sigma}^{q(l-1)}\}^{-1} \sum_{j=1}^J \gamma_j^{q(l-1)} + \mathbf{G}^q \mathbf{g}^q), \mathbf{V}_\gamma^q \right)$$

3 with prior mean vector $\mathbf{g}^q$, prior precision matrix $\mathbf{G}^q$ and posterior variance
 4 matrix $\mathbf{V}_\gamma = (J\{\boldsymbol{\Sigma}^{q(l-1)}\}^{-1} + \mathbf{G}^q)^{-1}$.

3 Sample covariance matrices of random regression coefficients:

$$\boldsymbol{\Sigma}^{q(l)} \sim \mathcal{W}^{-1} \left(j^0 + J, \boldsymbol{\Sigma}^0 + \sum_{j=1}^J (\gamma_j^{q(l-1)} - \gamma_j^{q(l)}) (\gamma_j^{q(l-1)} - \gamma_j^{q(l)})^T \right)$$

5 with prior hyperparameters $j^0 \geq P^R$ and $\boldsymbol{\Sigma}^0$.

6 4 For each j , sample random regression coefficients:

$$\gamma_j^{q(l)} \sim N \left(\mathbf{V}_{\gamma_j}^q (\mathbf{X}_j^R \boldsymbol{\Omega}_j^{q(l-1)} (\boldsymbol{\kappa}_j^{q(l-1)} - \mathbf{X}_j^F \boldsymbol{\beta}^{q(l)} + \right. \\ \left. \ln[\sum_{m \neq q} \exp(\mathbf{X}_j^F \boldsymbol{\beta}^{m(l)} + \mathbf{X}_j^R \gamma_j^{m(l)})]) + \{\boldsymbol{\Sigma}^q\}^{-1} \gamma^q), \mathbf{V}_{\gamma_j}^q \right)$$

7 with prior mean vector $\gamma^{q(l)}$, prior precision matrix $\boldsymbol{\Sigma}^{q(l)}$ and posterior
 8 variance matrix $\mathbf{V}_{\gamma_j}^q = (\mathbf{X}_j^{RT} \boldsymbol{\Omega}_j^{q(l-1)} \mathbf{X}_j^R + \{\boldsymbol{\Sigma}^{q(l)}\}^{-1})^{-1}$.

5 For each j and i , sample Pólya-Gamma variables:

$$\omega_{ji}^{q(l)} \sim PG(1, \eta_{ji}^{q(l)})$$

9 A.3 A note on prior specification

10 A.3.1 Regression parameters

11 In the Gibbs sampling framework, regression coefficients are normally distributed
 12 with a mean and covariance matrix. We shortly discuss the role of these parameters
 13 below. The covariance matrix defines the spread of the distribution and therefore has
 14 a substantial influence on informativity: Small variance parameters increase prior
 15 information. When non-informativity is preferable, large variance parameters are
 16 not the simple answer, as they may destabilize computations in Bayesian logistic
 17 regression analysis [2]. Jeffreys's prior could be an option, but sufficiently stable
 18 computation is not guaranteed [2, 3]. The challenge is therefore to specify prior
 19 variance parameters that are both sufficiently small to support stable analysis and
 20 to give a realistic support of the parameter and at the same time sufficiently large
 21 to be considered vague.

22 The mean hyperparameters defines the center of the distribution and becomes
 23 increasingly influential on the posterior distribution when the variance of the
 24 distribution is small. The relevance of adequate mean hyperparameters therefore
 25 increases with the informativity of the analysis. It should be noted that prior

information of mean regression coefficients is not always available in the required parametrization. Researchers may be more likely to have information available in terms of (success) probabilities rather than logistic regression parameters. Kavelaars et al. propose an approach to compute mean hyperparameters for the context of treatment comparison in the presence of a single patient characteristics, based on expected joint response probabilities [4].

A.3.2 Covariance matrices

The covariance matrix follows an inverse-Wishart distribution with parameters. Specifying a non-informative prior on covariance matrices and variance parameters in general is not straightforward [5, 6]. The informativity of the inverse-Wishart distribution is sensitive to the size of variance parameters: small variances make inverse-Wishart distributions more informative. Naively specifying standard prior hyperparameters without consideration of prior information or data at hand may result in an undesirably large prior influence. Weakly informative (data-based) prior specification may be superior, if not essential for computational stability [5].

Appendix B: Procedure for transformation to the probability scale and decision-making

Algorithm 1 Procedure for statistical decision-making with posterior regression coefficients

```

1: Step 1. Transform regression coefficients to treatment differences
2: Let  $\gamma_j^Q = (0, \dots, 0)$  and  $\mathbf{x} = (1, T, w, \dots)$ 
3: for draw  $(l) \leftarrow 1 : L$  do
4:   for cluster  $j \leftarrow 1 : J$  do
5:     Compute joint response probabilities
6:     for treatment  $T \leftarrow 0 : 1$  do
7:       for joint response category  $q \leftarrow 1 : Q$  do
8:         if Population of interest defined by a range of values of  $w$  then
9:
10:          Compute  $\phi_{Tj}^{q(l)} = \int_w \frac{\exp[\mathbf{x}_j' \gamma_j^{q(l)}]}{\sum_{r=1}^{Q-1} \exp[\mathbf{x}_j' \gamma_j^{r(l)}] + 1} dw$ 
11:        end if
12:        if Population of interest defined by a fixed value of  $w$  then
13:
14:          Compute  $\phi_{Tj}^{q(l)} = \frac{\exp[\mathbf{x}_j' \gamma_j^{q(l)}]}{\sum_{r=1}^{Q-1} \exp[\mathbf{x}_j' \gamma_j^{r(l)}] + 1}$ 
15:        end if
16:      end for
17:      Compute multivariate success probabilities
18:      for outcome  $k \leftarrow 1 : K$  do
19:        Compute  $\theta_{Tj}^{q(l)} = \sum_{q=1}^Q \phi_{Tj}^{q(l)} I(\mathbf{h}^q \in U_k)$ 
20:      Compute multivariate treatment difference

```

```

1  21:           Compute  $\delta_j^{k(l)} = \theta_{1j}^{k(l)} - \theta_{0j}^{k(l)}$ 
2  22:           end for
3  23:       end for
4  24:   end for
5  25:   for outcome  $k \leftarrow 1 : K$  do
6  26:       Pool  $\delta^{k(l)} = \sum_{j=1}^J \frac{n_j}{\sum_{j=1}^J n_j} \delta_j^{k(l)}$ 
7  27:   end for
8  28: end for
9  29: Step 2. Make superiority decision
10 30: Define superiority region  $\mathcal{S}_R$ 
11 31: Draw conclusion
12 32: if  $\frac{1}{L} \sum_{(l)=1}^L I(\delta^{(l)} \in \mathcal{S}_R) > p_{cut}$  then Conclude superiority
13 33: else Conclude non-superiority
14 34: end if

```

Author details

References

1. Polson NG, Scott JG, Windle J. Bayesian inference for logistic models using Pólya–Gamma latent variables. *Journal of the American statistical Association*. 2013;108(504):1339–1349.
2. Gelman A, Jakulin A, Pittau MG, Su YS. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*. 2008 dec;2(4):1360–1383.
3. Poirier D. Jeffreys’ prior for logit models. *Journal of Econometrics*. 1994 aug;63(2):327–339.
4. Kavelaars XM, Mulder J, Kaptein MC. Bayesian multivariate logistic regression for superiority and inferiority decision-making under treatment heterogeneity. Submitted for publication. 2022;.
5. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*. 2006;1(3):515 – 534. Available from: <https://doi.org/10.1214/06-BA117A>.
6. Schuurman NK, Grasman RPPP, Hamaker EL. A Comparison of Inverse-Wishart Prior Specifications for Covariance Matrices in Multilevel Autoregressive Models. *Multivariate Behavioral Research*. 2016 mar;51(2-3):185–206.