# A Role for Enhanced Functions of Sleep in Psychedelic Therapy?

# Supplementary Material: Full Details of the Statistical Analysis

**Tom Froese, Iwin Leenen, and Tomas Palenicek**

## *Data*

The data for this reanalysis, provided to us by Rambousek et al. (2014), are the following: The latencies (times in seconds) that each of 24 rats needed to reach the platform in the Morris water maze on 24 trials. Trials 1–16 were administered on Day 1; Trials 17–24 on Day 2. The 24 rats were randomly divided in 3 experimental conditions (8 rats in each condition), which differed with respect to the amount of psilocin administered after the last trial on Day 1: 0 mg/kg (control), 1 mg/kg, and 4 mg/kg. Hence, in total, we have $24 \times 24 = 576$ observations.

Let $y_{ij}$ denote the observed latency for rat $i$ (with $i = 1, \ldots, 24$) on trial $j$ ($j = 1, \ldots, 24$). Importantly, 33% of the $y_{ij}$ are censored as the trials were interrupted whenever the rat did not reach the platform within 60 seconds. We define $c_{ij} = 1$ when the observation was censored, and $c_{ij} = 0$ otherwise.

## *Model*

We specify a hierarchical linear model, where the $Y_{ij}$ (the uncensored random variables associated with the observations $y_{ij}$) are assumed to follow a normal distribution:

$$Y_{ij} \sim \mathrm{N}(\theta_{ij}, \tau^2), \tag{1}$$

with $\tau^2$ a free to-be-estimated variance parameter and $\theta_{ij}$ being the expected value for the time required by rat $i$ on trial $j$, which is modeled as:

$$\theta_{ij} = [\alpha_{1i} + \beta_{1i}(j - 8.5)]\, I(j \in \mathrm{Day}_1) + [\alpha_{2i} + \beta_{2i}(j - 20.5)]\, I(j \in \mathrm{Day}_2). \tag{2}$$

The indicator function $I(\textit{condition})$ in the latter equation returns 1 if *condition* is true and 0 otherwise; $\mathrm{Day}_1$ and $\mathrm{Day}_2$ denote the set of trials on the first day and second day, respectively (i.e., $\mathrm{Day}_1 = \{1, \ldots, 16\}$ and $\mathrm{Day}_2 = \{17, \ldots, 24\}$).

Equation (2) specifies, for each individual rat $i$ and separately for Day 1 and Day 2, a simple linear regression model for its latencies in function of the trial number, with intercepts $\alpha_{1i}$ and $\alpha_{2i}$, for Day 1 and Day 2, respectively, and slopes $\beta_{1i}$ and $\beta_{2i}$. Let $\boldsymbol{\vartheta}_i$ denote the vector of regression parameters associated with rat $i$, that is, $\boldsymbol{\vartheta}_i = (\alpha_{1i}, \beta_{1i}, \alpha_{2i}, \beta_{2i})'$. Then, the $\boldsymbol{\vartheta}_i$ are considered random effects, which are assumed to be independently drawn from a multivariate normal (MVN) distribution:

$$\boldsymbol{\vartheta}_i \sim \mathrm{MVN}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), \tag{3}$$

where the vector $\boldsymbol{\mu}_i$ is given by

$$\boldsymbol{\mu}_i = \begin{pmatrix} \tilde{\alpha}_1 \\ \tilde{\beta}_1 \\ \tilde{\alpha}_{20}\, I(\mathrm{Dose}_i = 0) + \tilde{\alpha}_{21}\, I(\mathrm{Dose}_i = 1) + \tilde{\alpha}_{24}\, I(\mathrm{Dose}_i = 4) \\ \tilde{\beta}_{20}\, I(\mathrm{Dose}_i = 0) + \tilde{\beta}_{21}\, I(\mathrm{Dose}_i = 1) + \tilde{\beta}_{24}\, I(\mathrm{Dose}_i = 4) \end{pmatrix} \tag{4}$$

and $\boldsymbol{\Sigma}$ is a free to-be-estimated $4 \times 4$ covariance matrix.

The parameters $\tilde{\alpha}_1$ and $\tilde{\beta}_1$ can be interpreted as, respectively, the mean intercept and slope across all rats on the first day. (As all rats received the same treatment on Day 1, we do not assume the intercept and slope to differ in function of the experimental condition.) On the other hand, $\tilde{\alpha}_{2k}$ and $\tilde{\beta}_{2k}$ (for $k = 0, 1, 4$) are the mean intercept and slope on the second day across the rats that were injected $k$ mg/kg of psilocin. Note further that the model specification in Equation (2) implies that the intercept parameters ($\tilde{\alpha}_1$ and $\tilde{\alpha}_{2k}$) can be interpreted as the mean latency across trials within the given day, while the slope parameters ($\tilde{\beta}_1$ and $\tilde{\beta}_{2k}$) model a possible learning effect across trials within days.

*Sampling distribution of the observed data* Equation (1) specifies that the latencies depend on the $\theta_{ij}$ and $\tau^2$. By assuming independence, it follows that the sampling distribution of the observed data is given by:

$$p(\mathbf{y} \mid \boldsymbol{\theta}, \tau^2) \;=\; \prod_{i=1}^{24} \prod_{j=1}^{24} \Big[ \big(\phi(y_{ij}; \theta_{ij}, \tau^2)\big)^{1-c_{ij}} \; \big(1 - \Phi(60; \theta_{ij}, \tau^2)\big)^{c_{ij}} \Big], \tag{5}$$

where $\mathbf{y}$ denotes the vector of observed data and $\boldsymbol{\theta}$ the corresponding vector of parameters $\theta_{ij}$. The $\phi(\cdot; \theta_{ij}, \tau^2)$ and $\Phi(\cdot; \theta_{ij}, \tau^2)$ denote the univariate normal probability density function and the normal cumulative distribution function, respectively, with mean $\theta_{ij}$ and variance $\tau^2$. Note that, by Eq. (5), censoring is taken into account by the model.

By Eq. (2), each $\theta_{ij}$ is fully determined by the regression parameters in the $\boldsymbol{\vartheta}_i$ (and the trial number). Hence, using $\boldsymbol{\vartheta}$ to denote the vector of the individual $\boldsymbol{\vartheta}_i$, we can interchangeably use $p(\mathbf{y} \mid \boldsymbol{\theta}, \tau^2)$ and $p(\mathbf{y} \mid \boldsymbol{\vartheta}, \tau^2)$.

## Estimation

The free parameters of the above described model are $\tilde{\alpha}_1$, $\tilde{\beta}_1$, $\tilde{\alpha}_{20}$, $\tilde{\beta}_{20}$, $\tilde{\alpha}_{21}$, $\tilde{\beta}_{21}$, $\tilde{\alpha}_{24}$, $\tilde{\beta}_{24}$, $\tau^2$, and $\boldsymbol{\Sigma}$. To obtain estimates for these parameters, we considered the model in a Bayesian framework and looked at the posterior distribution of the parameters. By Bayes theorem, the posterior distribution is proportional to the product of the likelihood function and the prior distribution. That is,

$$\begin{aligned} p(\tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_{20}, \tilde{\beta}_{20}, \tilde{\alpha}_{21}, \tilde{\beta}_{21}, \tilde{\alpha}_{24}, \tilde{\beta}_{24}, \tau^2, \boldsymbol{\Sigma} \mid \mathbf{y}) \;\propto\; & p(\mathbf{y} \mid \tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_{20}, \tilde{\beta}_{20}, \tilde{\alpha}_{21}, \tilde{\beta}_{21}, \tilde{\alpha}_{24}, \tilde{\beta}_{24}, \tau^2, \boldsymbol{\Sigma}) \\ & \times\; p(\tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_{20}, \tilde{\beta}_{20}, \tilde{\alpha}_{21}, \tilde{\beta}_{21}, \tilde{\alpha}_{24}, \tilde{\beta}_{24}, \tau^2, \boldsymbol{\Sigma}), \end{aligned}$$

where the symbol $\propto$ is read as "is proportional to".

The likelihood function can be decomposed as follows:

$$\begin{aligned} p(\mathbf{y} \mid \tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_{20}, \tilde{\beta}_{20}, \tilde{\alpha}_{21}, \tilde{\beta}_{21}, \tilde{\alpha}_{24}, \tilde{\beta}_{24}, \tau^2, \boldsymbol{\Sigma}) \;\propto\; & p(\mathbf{y} \mid \boldsymbol{\vartheta}, \tau^2) \\ & \times\; p(\boldsymbol{\vartheta} \mid \tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_{20}, \tilde{\beta}_{20}, \tilde{\alpha}_{21}, \tilde{\beta}_{21}, \tilde{\alpha}_{24}, \tilde{\beta}_{24}, \boldsymbol{\Sigma}), \end{aligned}$$

where $p(\mathbf{y} \mid \boldsymbol{\vartheta}, \tau^2)$ is given by Eq. (5). Assuming independence, it follows further from Eq. (3) that

$$p(\boldsymbol{\vartheta} \mid \tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_{20}, \tilde{\beta}_{20}, \tilde{\alpha}_{21}, \tilde{\beta}_{21}, \tilde{\alpha}_{24}, \tilde{\beta}_{24}, \boldsymbol{\Sigma}) \;=\; \prod_{i=1}^{24} \boldsymbol{\phi}(\boldsymbol{\vartheta}_i; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}).$$

In the latter equation $\boldsymbol{\phi}(\cdot; \boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ denotes the multivariate normal density function with means $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}$. Note that, by Eq. (4), the $\boldsymbol{\mu}_i$ are fixed given the parameters $\tilde{\alpha}_1$, $\tilde{\beta}_1$, $\tilde{\alpha}_{20}$, $\tilde{\beta}_{20}$, $\tilde{\alpha}_{21}$, $\tilde{\beta}_{21}$, $\tilde{\alpha}_{24}$, and $\tilde{\beta}_{24}$ (and knowledge about the experimental condition that rat $i$ belongs to).

With respect to the prior distribution, we assume that the parameters $\tilde{\alpha}_1$, $\tilde{\beta}_1$, $\tilde{\alpha}_{20}$, $\tilde{\beta}_{20}$, $\tilde{\alpha}_{21}$, $\tilde{\beta}_{21}$, $\tilde{\alpha}_{24}$, and $\tilde{\beta}_{24}$ are identically and independently distributed (iid) as follows:

$$\tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_{20}, \tilde{\beta}_{20}, \tilde{\alpha}_{21}, \tilde{\beta}_{21}, \tilde{\alpha}_{24}, \tilde{\beta}_{24} \overset{\text{iid}}{\sim} \text{N}(0, 1000^2).$$

By specifying a large variance, the prior distribution essentially does not provide information about these parameters.

For the variance parameters $\tau^2$ and $\boldsymbol{\Sigma}$, we assume independent vague priors as follows:

$$\tau^2 \;\sim \text{Gamma}(0.001, 0.001),$$

with the numbers between parentheses being the shape and inverse-scale parameters of the gamma distribution;

$$\boldsymbol{\Sigma} \;\sim \text{Inverse-Wishart}(4, \mathbf{I}),$$

that is, an Inverse-Wishart distribution with 4 degrees of freedom and the $4 \times 4$ identity matrix as the scale matrix.

We obtained 5,000 draws from the specified posterior distribution by running a Markov chain Monte Carlo procedure with 5,000,000 iterations (apart from 5,000 burn-in draws) of which each 1,000th iteration was saved. We implemented the model and conducted the analysis using the procedure PROC MCMC of SAS Software Version 9.4 (SAS Institute, 2015). Examination of the value-by-iteration plots showed that the chain for each of the parameters had rapidly reached its stationary distribution.

## Results: Posterior summary

Table 1 shows for each of the free parameters the posterior mean, standard deviation, and the associated 95%-high posterior density interval, based on the 5,000 draws from the simulated posterior distribution.

**Table 1.** Posterior summaries for each of the parameters, based on 5,000 draws obtained from a Markov chain Monte Carlo simulation procedure.

| Parameter | Mean | Standard deviation | 95%-high posterior density interval |
|---|---|---|---|
| $\tilde{\alpha}_1$ | 45.95 | 3.43 | [ 39.04; 52.52] |
| $\tilde{\beta}_1$ | −1.76 | 0.43 | [ −2.59; −0.90] |
| $\tilde{\alpha}_{20}$ | 43.82 | 4.26 | [ 35.18; 51.95] |
| $\tilde{\alpha}_{21}$ | 38.57 | 3.99 | [ 30.67; 46.33] |
| $\tilde{\alpha}_{24}$ | 27.14 | 4.00 | [ 19.14; 38.88] |
| $\tilde{\beta}_{20}$ | −5.08 | 1.57 | [ −8.11; −2.00] |
| $\tilde{\beta}_{21}$ | −5.22 | 1.47 | [ −7.97; −2.18] |
| $\tilde{\beta}_{24}$ | −2.52 | 1.45 | [ −5.46; 0.20] |
| $\tau^2$ | 23.32 | 0.97 | [ 21.47; 25.22] |
| $\mathbf{\Sigma}_{11}$ | 238.6 | 92.01 | [ 98.84; 423.7 ] |
| $\mathbf{\Sigma}_{12}$ | 6.05 | 8.29 | [ −9.44; 22.92] |
| $\mathbf{\Sigma}_{13}$ | 161.5 | 65.96 | [ 57.32; 291.3 ] |
| $\mathbf{\Sigma}_{14}$ | −18.72 | 17.57 | [−56.21; 11.07] |
| $\mathbf{\Sigma}_{22}$ | 2.40 | 1.49 | [ 0.20; 5.25] |
| $\mathbf{\Sigma}_{23}$ | 6.19 | 6.30 | [ −5.37; 19.58] |
| $\mathbf{\Sigma}_{24}$ | −1.48 | 1.67 | [ −4.83; 1.46] |
| $\mathbf{\Sigma}_{33}$ | 121.3 | 64.15 | [ 22.25; 247.6 ] |
| $\mathbf{\Sigma}_{34}$ | −15.02 | 13.75 | [−44.08; 7.71] |
| $\mathbf{\Sigma}_{44}$ | 4.00 | 3.80 | [ 0.10; 11.48] |