Pre-commitment of analysis procedure

Pontus P. Sigray

20 February 2017

1) What's the main question being asked or hypothesis being tested in this study?

The study will examine whether patients with psychosis disorder have 1. increased, 2. decreased or 3. unchanged levels of Translocator Protein (TSPO) distribution volume (V_T) in the central nervous system compared to healthy controls.

2) Describe the key dependent variable(s) specifying how they will be measured.

TSPO V_T in three regions of interest (ROIs): hippocampus, temporal cortex (TC) and frontal cortex (FC).

Literature search

A systematic literature search will be performed on PubMed with the goal of identifying all published articles that:

1. Report TSPO V_T values for patients with psychosis disorder as compared to healthy controls
2. Uses a second generation radioligand,with arterial input function, to quantify TSPO V_T.
3. Report affinity-type (HAB, MAB and/or LAB) of participants.

Data retrieval and quality control

The corresponding authors of all studies fulfilling the inclusion criteria will be contacted and asked to share individual participant V_T values from hippocampus, TC and FC, as well as gender, age and TSPO genotype status. The data will then be quality controlled, by comparing the descriptive statistics of the shared data with those reported in the original articles.

3) How many and which conditions will participants be assigned to?

Two main conditions: Patients with psychosis disorder and healthy controls.

4) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Statistical model specification

A Bayesian linear mixed effects (LME) model will be specified:

- Individual participant ROI V_T values as dependent variable.
- patient-control group variable as fixed effect.
- Genotype (HAB or MAB) as lvl 2 random effect.
- Study as lvl 3 random effect.

The out of sample deviance (fits) of the above model will be evaluated for increasing complexity of the model specification (allowing slopes to vary in addition to intercepts). The model specification corresponding to the best fit will be selected.

Hypotheses testing

The best fitting model will then be used to evaluate the following hypotheses:

- H0: No difference in V_T between patients and controls
- H1: Patients have higher V_T as compared to controls
- H2: Patients have lower V_T as compared to controls

for each ROI.

The prior on patient-control fixed effect will be a truncated Gaussian distribution with mean = 0 and SD = 0.5, with a lower bound of 0 for H1 and and upper bound of zero for H2. A SD of 0.5 is chosen since it corresponds to a expected medium effect size of the fixed effect (Dienes 2014). The prior over the random effects will be weakly informative half-Cauchys (scale = 0.707) since the number of groups (both for genotype and for study) is small (Gelman 2006).

For each ROI, Bayes factor will be computed in favor of H1 (there is an increased V_T in patients) compared to the H0 (the fixed effect of patient-control variable is zero). Bayes factor will also be computed in favor of H2 (there is a decreased V_T in patients) compared to H0. Finally, if there is support for H1 or H2 over H0, Bayes factor in favor of H1 over H2 will be computed. This allows us to fully explore the three hypotheses of interest for each ROI.

A standard robustness check of the prior over the fixed effect will also be performed for both H1 and H2 (SD = 0.2 and SD = 0.8 corresponding an expected small and large effect size).

5) Any secondary analyses?

Demographics of sample

If possible, demographic data, such as age, gender, disease duration and clinical ratings will be extracted from all studies in order to assess patient-control and study heterogeneity.

Parameter estimation

For parameter estimation, model averaging of H1, H2 and H0 will *not* be performed. Rather, the effect size of difference between patients and healthy controls (and its credible interval) will be assessed for each ROI using a LME model with a weakly regularizing non-truncated prior distribution over the fixed effect. The

reason for this is to provide the field with a robust, accessible and easy to interpret effect size for future power estimation.

Drug free and medicated patients

If possible, the patient sample will be divided into: patients on medication and patients that are drug free. The same model as above can then be run with an additional predictor: medication status. The posterior of this predictor can be extracted and parameter estimation can be performed to conclude how plausible an effect of medication is.

6) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

The sample size will depend on how many articles from the systematic literature search are deemed to fulfill the inclusion criteria, and how many individual participant V_T values we can obtain from these articles.

7) Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)

All VT values (dependent variable) will be z-transformed (centered and divided by the SD) within each genotype group, within each study, to account for range differences between different TSPO radioligands used in the studies.

8) Have any data been collected for this study already?

Yes, all data have been collected.

9) Appendix

Specification of models to be evaluated:

$y_i \sim N(\mu_i, \sigma)$

$M1: \ \mu_i = \beta_0 + \beta_1 x_i + u_{0j} + w_{0k}$

$M2: \ \mu_i = \beta_0 + \beta_1 x_i + u_{0j} + u_{1j} x_i + w_{0k}$

$M3: \ \mu_i = \beta_0 + \beta_1 x_i + u_{0j} + w_{0k} + w_{1k} x_i$

$M4: \ \mu_i = \beta_0 + \beta_1 x_i + u_{0j} + u_{1j} x_i + w_{0k} + w_{1k} x_i$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim MV\ Normal\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} {\sigma_{u0}}^2 & \rho_u \sigma_{u0} \sigma_{u1} \\ \rho_u \sigma_{u0} \sigma_{u1} & {\sigma_{u1}}^2 \end{pmatrix} \right)$$

$$\begin{pmatrix} w_{0k} \\ w_{1k} \end{pmatrix} \sim MV\ Normal\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} {\sigma_{w0}}^2 & \rho_w \sigma_{w0} \sigma_{w1} \\ \rho_w \sigma_{w0} \sigma_{w1} & {\sigma_{w1}}^2 \end{pmatrix} \right)$$

$\sigma \sim Student's\ t(3, 0, 10) T(0, \infty)$

$\beta_0 \sim Student's\ t(3, 0, 10)$

$\beta_1 \sim Normal(0, 0.5) T(0, \infty)^{\mathrm{A}} \ \vee \ \beta_1 \sim N(0, 0.5) T(\infty, 0)^{\mathrm{B}}$

$\sigma_{u_0}, \sigma_{w_0} \sim Cauchy(0, \sqrt{2}/2) T(0, \infty)$

$\sigma_{u_1}, \sigma_{w_1} \sim Cauchy(0, \sqrt{2}/2) T(0, \infty)$

$\rho_u, \rho_w \sim Uniform(-1, 1)$

$i \in \{1, 2 \ldots 168\}$

$j \in \{HAB, MAB\}$

$k \in \{Kenk, Bloomfield, Coughlin, Collste, Hafizi\}$

where y_i is ROI V_T for subject i, $beta_1$ is the fixed effect for diagnostic group (ΔV_T), u_{0j} and u_{1j} are the intercept and slope for genotype j, w_{0k} and w_{1k} are the intercept and slope for Study k. A corresponds to $H1$ and B corresponds to $H2$. $M1$, $M2$, $M3$ and $M4$ are linear functions that corresponds to model 1, 2, 3 and 4.

Please note that priors on random effects are not final! These may be changed in order to facilitate model convergence.

Pubmed search terms

((("psychotic disorders"[MeSH Terms] OR ("psychotic"[All Fields] AND "disorders"[All Fields]) OR "psychotic disorders"[All Fields] OR "psychosis"[All Fields]) OR ("psychotic disorders"[MeSH Terms] OR ("psychotic"[All Fields] AND "disorders"[All Fields]) OR "psychotic disorders"[All Fields] OR ("psychotic"[All Fields] AND "disorder"[All Fields]) OR "psychotic disorder"[All Fields]) OR ("schizophrenia"[MeSH Terms] OR "schizophrenia"[All Fields])) AND ((translocator[All Fields] AND ("proteins"[MeSH Terms] OR "proteins"[All Fields] OR "protein"[All Fields])) OR (translocator[All Fields] AND ("antigens, cd59"[MeSH Terms] OR ("antigens"[All Fields] AND "cd59"[All Fields]) OR "cd59 antigens"[All Fields] OR "protein 18"[All Fields]) AND kDa[All Fields]) OR TSPO[All Fields] OR (peripheral[All Fields] AND ("receptors, gaba-a"[MeSH Terms] OR ("receptors"[All Fields] AND "gaba-a"[All Fields]) OR "gaba-a receptors"[All Fields] OR ("benzodiazepine"[All Fields] AND "receptor"[All Fields]) OR "benzodiazepine receptor"[All Fields])) OR PBR[All Fields])) AND (("positron-emission tomography"[MeSH Terms] OR ("positron-emission"[All Fields] AND "tomography"[All Fields]) OR "positron-emission tomography"[All Fields] OR ("positron"[All Fields] AND "emission"[All Fields] AND "tomography"[All Fields]) OR "positron emission tomography"[All Fields]) OR PET[All Fields]) AND 2004[EDAT] : 2017[EDAT]

References

Dienes, Zoltan. 2014. "Using Bayes to get the most out of non-significant results." *Frontiers in Psychology* 5. Frontiers: 781.

Gelman, Andrew. 2006. "Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)." *Bayesian Analysis* 1 (3). International Society for Bayesian Analysis: 515–34.