Effects of Working, Menstrual Cycle Phase, and Family History of Hypertension on the Systolic Blood Pressure of Nurses

BIOSTAT 234: Data Analysis Project #1

Chad Pickering

March 7, 2019

Introduction.

The general physical health of nurses as a unique occupational group is of interest - they undergo a copious amount of stress which in turn could affect their long-term health. 203 female nurses had systolic blood pressure (SBP, mmHG), a critical measure of cardiovascular health, recorded several times over a varying length of time. Three additional characteristics of each nurse were taken to predict SBP, family history of hypertension, whether or not it is a working day, and current menstrual cycle phase. We will fit a Bayesian regression model to determine if family history of hypertension, working day, or menstrual cycle phase are associated with high SBP, and whether or not there is an interaction between SBP measures on working days and family history of hypertension.

Choice of model and priors.

Out of the 203 nurses, 113 (55.7%) were measured on a workday, 102 (50.2%) were in the luteal phase of their menstrual cycle, and 91 (44.8%) reported that one or both parents had or has hypertension (Table 1).

The density for SBP is approximately bell-shaped with mildly thick tails and a slight positive skew toward those who are hypertensive. As expected, quantile-quantile plots (Figure 1) show a better fit for a t-distribution on the SBP responses than does a normal. I fit a selection of normal models and t-distributed models with the informative priors described below and, based on the fit of the response, DIC values, and observed convergence behavior, decided on the model where the response is t-distributed and all fixed effects are Gaussian. Nurse-level random effects remained Gaussian with flat priors - no prior information about them is available. The model is:

$$
\begin{aligned}
Y &= \alpha + X^T\beta + \epsilon \quad \text{where} \\
\alpha &= [\alpha_1 \quad ... \quad \alpha_{203}], \\
\beta^T &= [\beta_{work} \quad \beta_{phase} \quad \beta_{famhist1} \quad \beta_{famhist2} \quad \beta_{work\times famhist1} \quad \beta_{work\times famhist2}], \\
\epsilon &\sim t(0, \tau, d)
\end{aligned}
$$

The priors selected via past literature or by convention are as follows:

$$
\begin{aligned}
\alpha_i &\sim No(0, 0.001) \quad \forall i \in \{1, ..., 203\} \\
\beta_0 &\sim No(114.7, 0.0035) \\
\beta_{work} &\sim No(7.9, 0.1036) \\
\beta_{phase} &\sim No(2.44, 0.238) \\
\beta_{famhist1} &\sim No(2.45, 0.0156) \\
\beta_{famhist2} &\sim No(4.9, 0.0156) \\
\beta_{work\times famhist1} &\sim No(0, 0.001) \\
\beta_{work\times famhist2} &\sim No(0, 0.001) \\
\tau &\sim \Gamma(a = 50, b = 14280.5) \\
d &\sim Unif(2, \infty)
\end{aligned}
$$

In Lin J-D, et al., 64198 women between the ages of 21 and 65 were enrolled in an observational study where SBPs were measured. The female group had a mean SBP of 114.7 with a SD of 16.9 (precision 0.0035), which will be used as the normal prior parameters for the intercept β_0. This precision is weak enough to allow for the sample intercept to overpower it if the average SBP for those women with all zero-value predictors is significantly different. The precision τ is Gamma distributed, specified such that the first parameter is the chosen prior sample size $a = 50$ (approximately 25% of the data's sample size) and $b = a \cdot \hat{\sigma^2}_{prior} = 14280.5$ where $\hat{\sigma} = 16.9$. Drastic changes to the shape and scale while keeping the mean the same did not change estimates significantly.

Theorell, et al. took a sample of 56 women between 20 and 59 years of age who worked in acute emergency care, child psychiatry, or a pediatric outpatient clinic and assessed SBP - the mean difference between work and rest/leisure groups was 7.9 with a combined variance of 9.648 (precision 0.1036). Mean SBPs measured in the luteal and follicular phases of women's menstrual cycles in Tsai, et al. were 119.26 and 116.82 respectively, a small difference of 2.44, with precision determined to be approximately 0.238. Although SBP increases with age in all individuals, before menopause many studies have shown that women appear to be protected against hypertension; only after menopause do SBP elevations in women become as prevalent as they are in men.

It appears that the interaction of family history and gender on hypertension is not well-defined in the literature. For example, Tozawa et al. asserts that for every additional family member with hypertension

mean SBP increases by 4.9 regardless of gender, while Goldstein et al. finds no significant difference between mean SBP among women with 0, 1, or 2 parents with hypertension. I set the prior at the average of the two findings, 2.45 and 4.9 for the two family history parameters, respectively, with the same precision, 0.0156, teased out of standard errors found in both papers. The interaction between working day and family history is a rather niche research question, so I set a flat Gaussian prior on both terms to let the data speak for itself.

Results and discussion.

The model was run in JAGS with 10000 iterations, 5000 of which were burn-in, were generated over 5 chains with a thinning parameter of 5 to eliminate some of the autocorrelation at adjacent lags. As is common in t-distributed models, mixing for the fixed and random effects were quite poor even after adjustments to some of the JAGS parameters to promote convergence (Figures 2-4). As seen in the figures, autocorrelation is high for many of the fixed and random effects out to double digit lags except for the degrees of freedom and sigma. Centering and scaling predictors was not possible as all fixed effects are factor variables, but that would have been the next attempted remedy. Some posterior densities like the interaction terms look bimodal, further suggesting poor mixing, as these priors are very weak (Figure 5).

With the priors specified, we have sufficient evidence that working day is predictive of the SBP measure of nurses - measures taken on working days are about 6.8 points higher on average as compared to non-working days (95% CI: 1.5-12.6; Table 2). No other fixed effect posterior estimates are significant at the 0.05 level - all other 95% credible intervals include 0 (Table 2; Figure 5). When a nurse is in the luteal phase, her SBP is on average 1.9 points higher than when she is in the follicular phase, but this is insignificant (95% CI: -1.4-4.9; Table 2). Having two hypertensive parents appears to raise a nurse's average SBP slightly more than if they have only one hypertensive parent, but both coefficients of family history appear to be insignificant, so evidence for an association is unsubstantiated. Finally, there is no significant interaction between working day and family history. The posterior estimate of the degrees of freedom on y at around 5 suggests that a t-distributed error is more appropriate than a Gaussian one; higher DIC values of all Gaussian models fit agrees with this claim.

A severe limitation to these results is the absence of an age covariate. We know that SBP tends to increase as a person gets older, so to not correct for this natural change over time is a noteworthy oversight. Additionally, looking at the QQ plots (Figure 1), it could be wise to consider a lognormal model; unfortunately, transforming the prior parameters via a Taylor series approximation requires the standard deviations to be small compared to the means for the logarithmic approximation to be accurate. A mixed normal-t model

may also be considered because of the skewed nature of only the right tail.

Conclusion.

After applying informative priors to all fixed effects, working day is the only significant covariate in the t-distributed model fit. Menstrual phase and family history of hypertension do not significantly predict SBP in female nurses, and there appears to be no interaction between working day and family history. Future models should factor in a log transformation or mixture, and attempt to further mend convergence issues by running JAGS or Stan models for 12+ hours.

References.

Goldstein IB, Shapiro D, Guthrie D: Ambulatory blood pressure and family history of hypertension in healthy men and women. Am J Hypertens 2006; 19: 486-491.

Lin J-D, Chen Y-L, Wu C-Z, Hsieh C-H, Pei D, Liang Y-J, et al. Identification of normal blood pressure in different age group. Med 2016; 95:e3188.

Theorell, Tores, et al. Influence of Job Strain and Emotion on Blood Pressure in Female Hospital Personnel during Workhours. Scandinavian Journal of Work, Environment & Health, Vol. 19, No. 5, 1993, pp. 313-318.

Tozawa M, Oshiro S, Iseki C, et al. Family history of hypertension and blood pressure in a screened cohort. Hypertens Res 2001; 24: 93-98.

Tsai PS, Yucha CB, Sheffield D, Yang, M. Effects of daily activities on ambulatory blood pressure during menstrual cycle in normotensive women. Appl Psychophysiol Biofeedback 2003;28:25-36.

Van Egeren LF. The relationship between job strain and blood pressure at work, at home, and during sleep. Psychosomatic Medicine 1992, 54, 337-343.

Appendix.

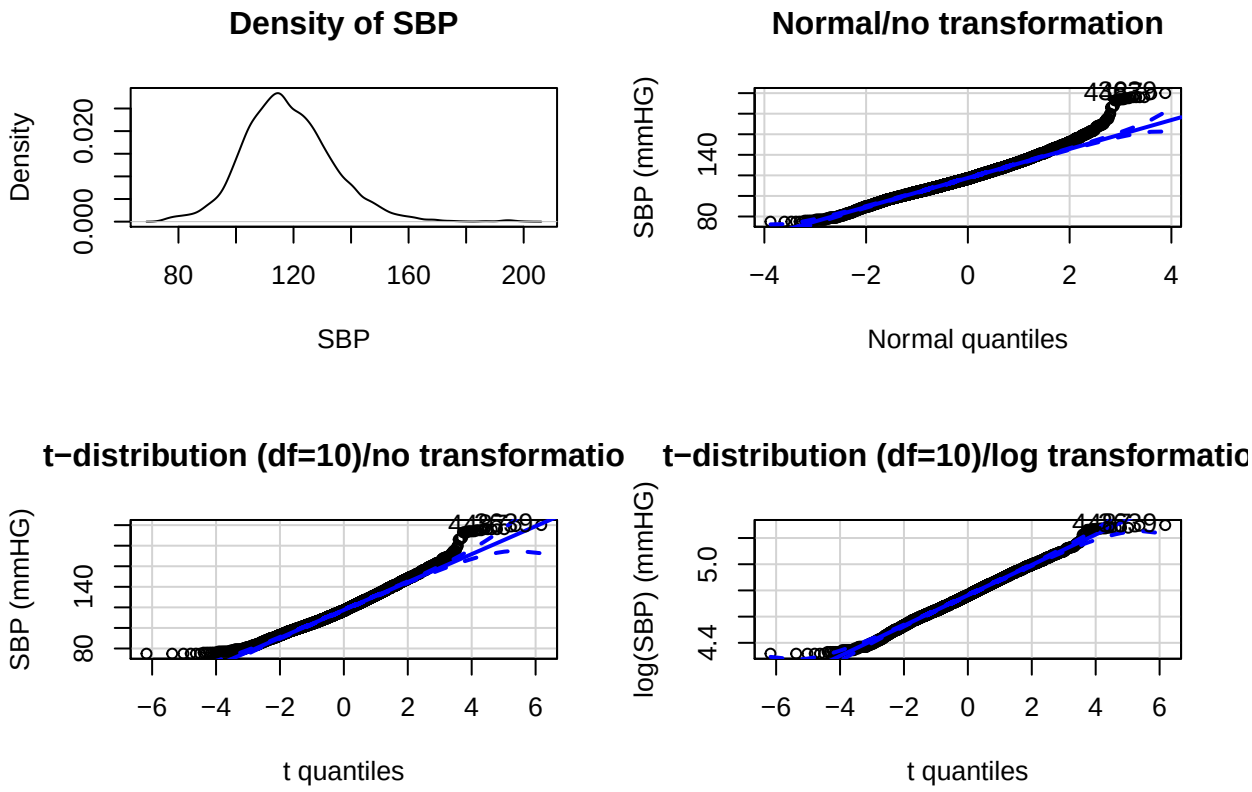
Figure 1. Characteristics of sample.

Variable	Level	Mean(SD)/Freq(%)
SBP (mmHg)		118.2 (15.5)
Working day	Working	113 (55.7)
	Not working	90 (44.3)
Menstrual cycle phase	Luteal	102 (50.2)
	Follicular	101 (49.8)
Family history of hypertension	Neither parent	112 (55.1)
	One parent	77 (37.9)
	Both parents	14 (6.9)

Table 2: Posterior estimates from Bayesian regression with t-distributed errors.

	Post. Est	SD Est	2.5%	97.5%
work	6.8	2.9	1.5	12.6
phase	1.9	1.6	-1.4	4.9
famhist1	2.9	4.8	-5.8	14.0
famhist2	6.7	7.6	-8.2	22.0
work x famhist1	-7.0	7.5	-24.0	5.3
work x famhist2	-2.1	14.0	-30.6	23.9
intercept	113.5	2.6	107.8	118.3
y df	4.9	0.2	4.4	5.4
sigma	10.2	0.1	10.0	10.5

Figure 1. Density and QQ plots of SBP.



Figures 2 and 3. ACF and time-series plots for fixed effects.

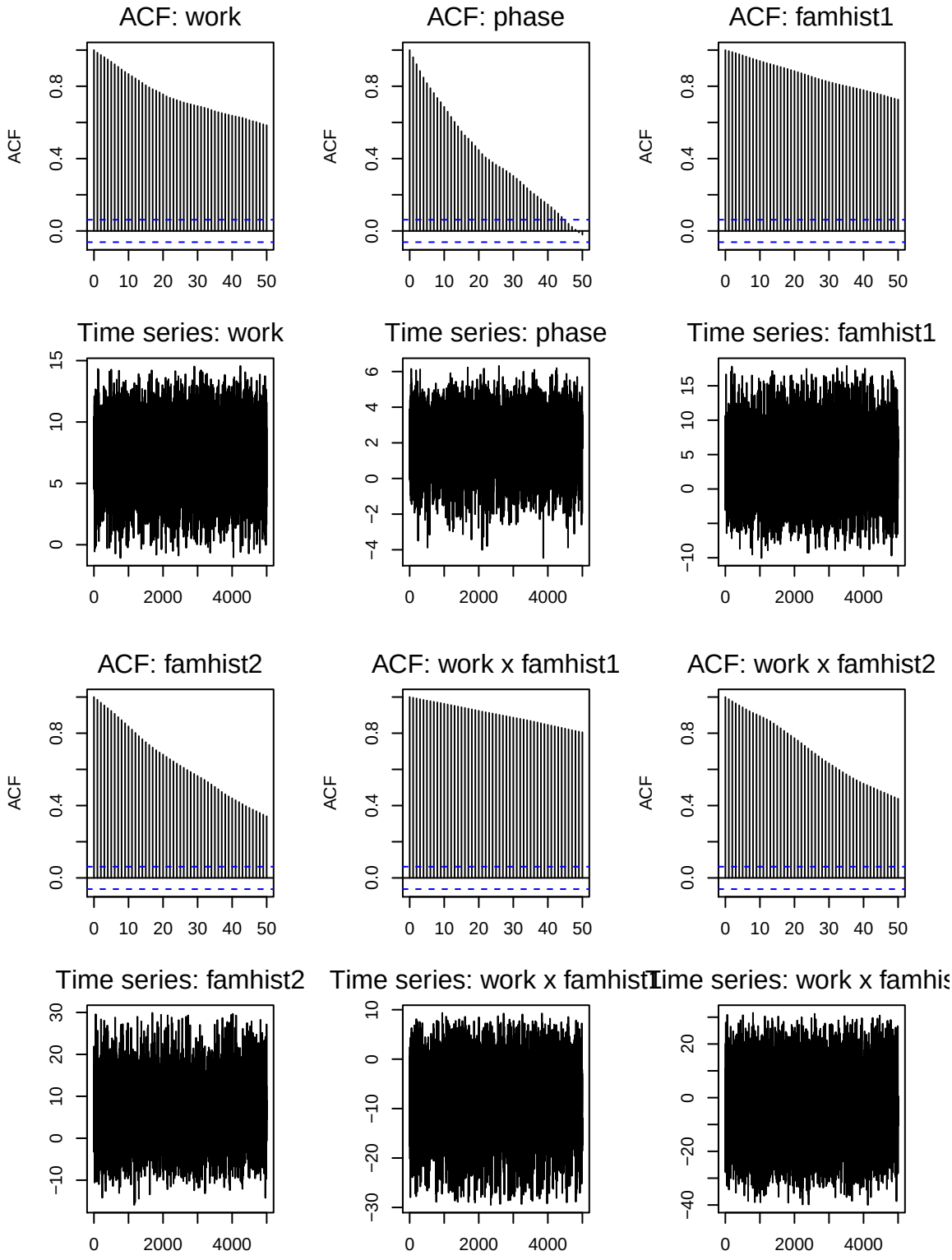


Figure 4. ACF and time-series plots for assorted random effects and other parameters.

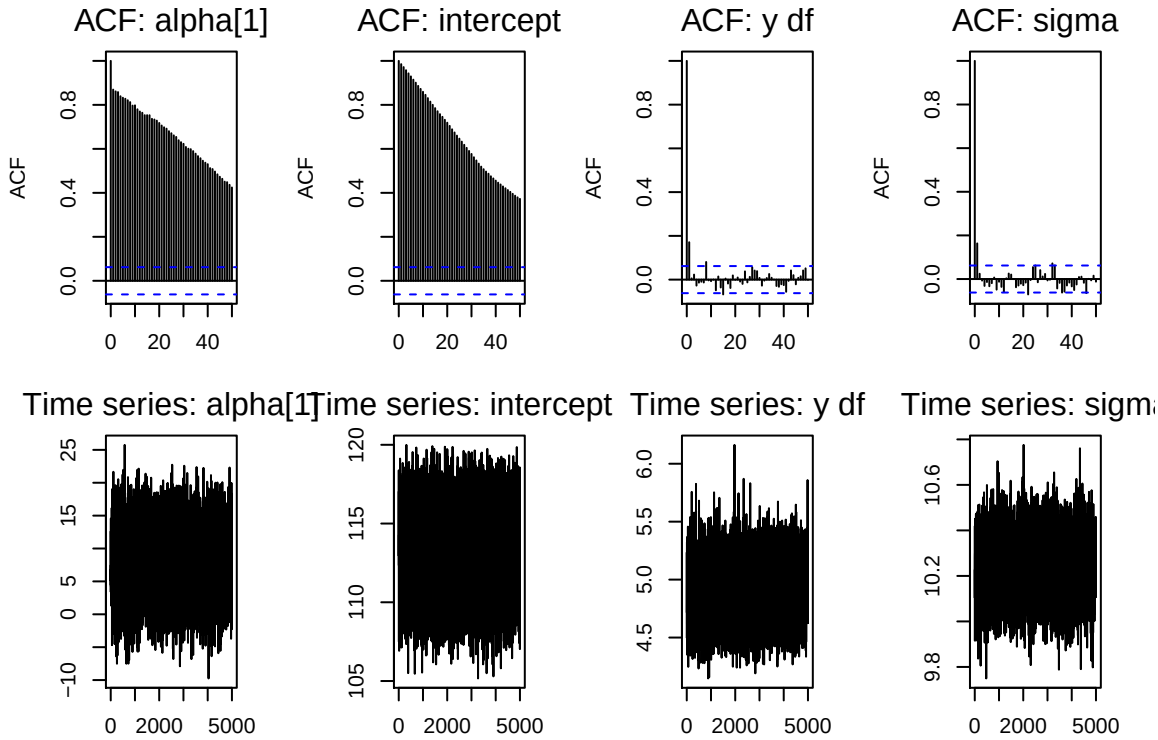
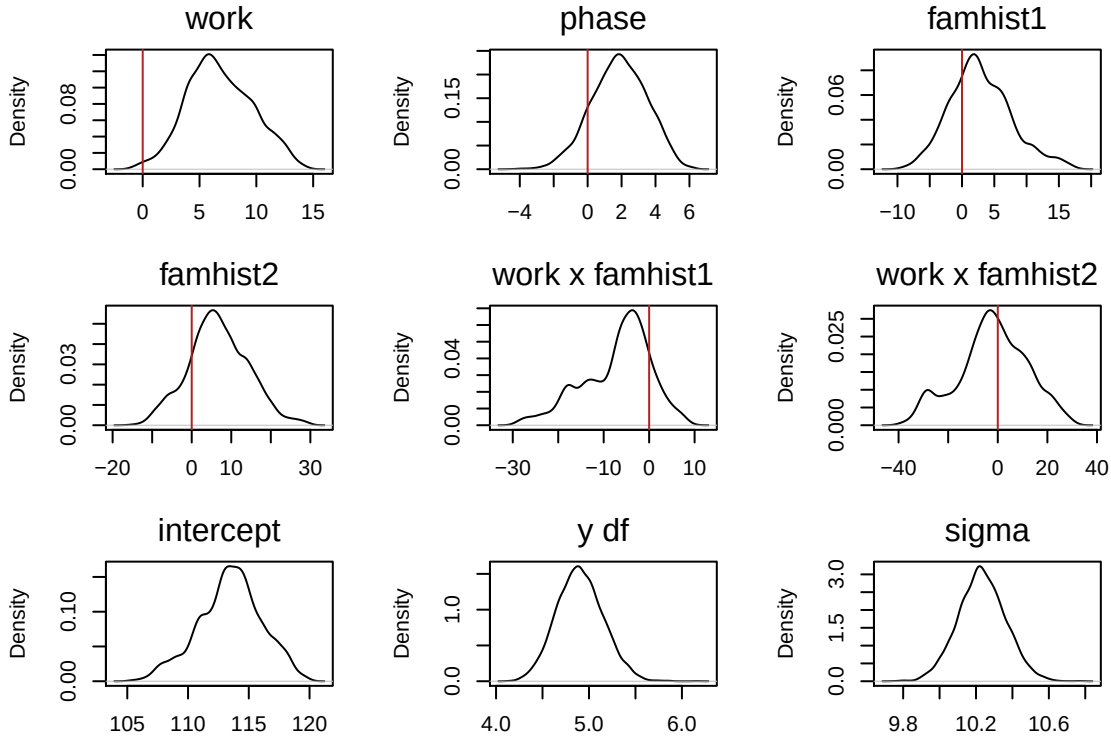


Figure 5. Posterior densities for fixed effects, intercept, y df, and sigma.



JAGS program.

```
cat("
model
{
   for(i in 1:N_obs){
      y[i] ~ dt(mu[i], tau, dfy)
        mu[i] <- beta0 + alpha[idnum[i]] + inprod(x[i,], beta[])
     }

   beta0 ~ dnorm(beta0_mean, beta0_prec)

   for(j in 1:N_betas){
         beta[j] ~ dnorm(beta_mean[j], beta_prec[j])
     }

   for(k in 1:N_ids){
       alpha[k] ~ dnorm(alpha_mean[k], alpha_prec[k])
   }

   tau ~ dgamma(tau.a, tau.b)
     sigma <- 1/sqrt(tau)

   dfy <- 1/invdfy
   invdfy ~ dunif(0, 0.5)
}
  ", file = "dap1_model2.txt", fill = TRUE, append = FALSE)

# betas: work, phase, famhist1, famhist2, work*famhist1, work*famhist2
prior_data_t <- list(N_obs=9573, N_betas=6, N_ids=203,
                     beta0_mean=114.7, beta0_prec=0.0035,
                     beta_mean=c(7.9, 2.44, 2.45, 4.9, 0, 0),
                     beta_prec=c(0.1036, 0.238, 0.0156, 0.0156, 0.001, 0.001),
                     alpha_mean=rep(0, 203),
                     alpha_prec=rep(0.001, 203),
                     tau.a=50, tau.b=14280.5, x=x[,-1], y=y, idnum=x[,1])

inits_t <- rep(list(list(alpha=rep(0, 203),
                         beta=rep(0, 6), beta0=0, tau=1)), 5)

# Parameters to monitor
parameters_t <- c("beta0", "alpha", "beta", "sigma",
                  "dfy")

# JAGS object
inform_t.sim <- jags(prior_data_t, inits_t, parameters_t, "dap1_model2.txt",
                     n.chains=5, n.iter=10000, n.burnin=5000,
                     n.thin=5, DIC=TRUE)
```