Bayesian Hierarchical Analysis of Blood Pressure Reduction

knitr::opts_chunk$set(echo = TRUE, message = FALSE, warning = FALSE)

```
library(rstan)
library(tidyverse)
library(bayesplot)
library(loo)
library(gridExtra)
library(bayesrules)
library(tidybayes)
library(coda)

options(mc.cores = parallel::detectCores())
rstan_options(auto_write = TRUE)

Expected object (not included in the repo):
presionarterial with columns: reduccion, hospital, edad, severidad
presionarterial <- read.csv("data/presionarterial.csv")

datos <- presionarterial %>%
mutate(hospital_id = as.integer(as.factor(hospital)))
n_hospitals <- length(unique(datos$hospital_id))
```

Image 1: Histogram

```
hist(
datos$reduccion,
main = "Distribution of Blood Pressure Reduction",
xlab = "Reduction"
)
```

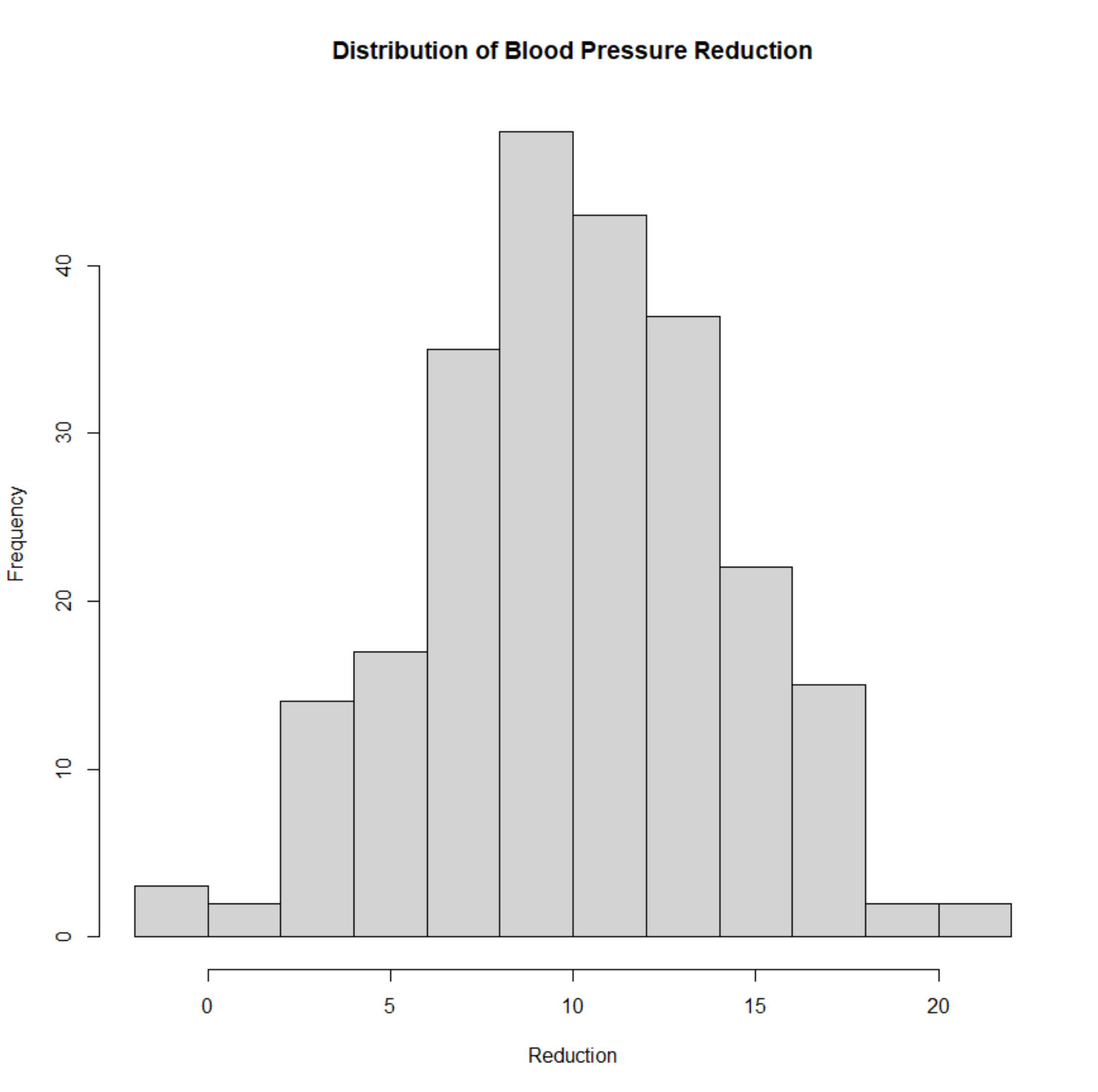


Image 2: Boxplot by hospital

```
ggplot(datos, aes(x = factor(hospital_id), y = reduccion)) +
geom_boxplot(alpha = 0.5) +
theme_minimal() +
labs(
title = "Blood Pressure Reduction by Hospital",
x = "Hospital",
y = "Reduction"
)
```

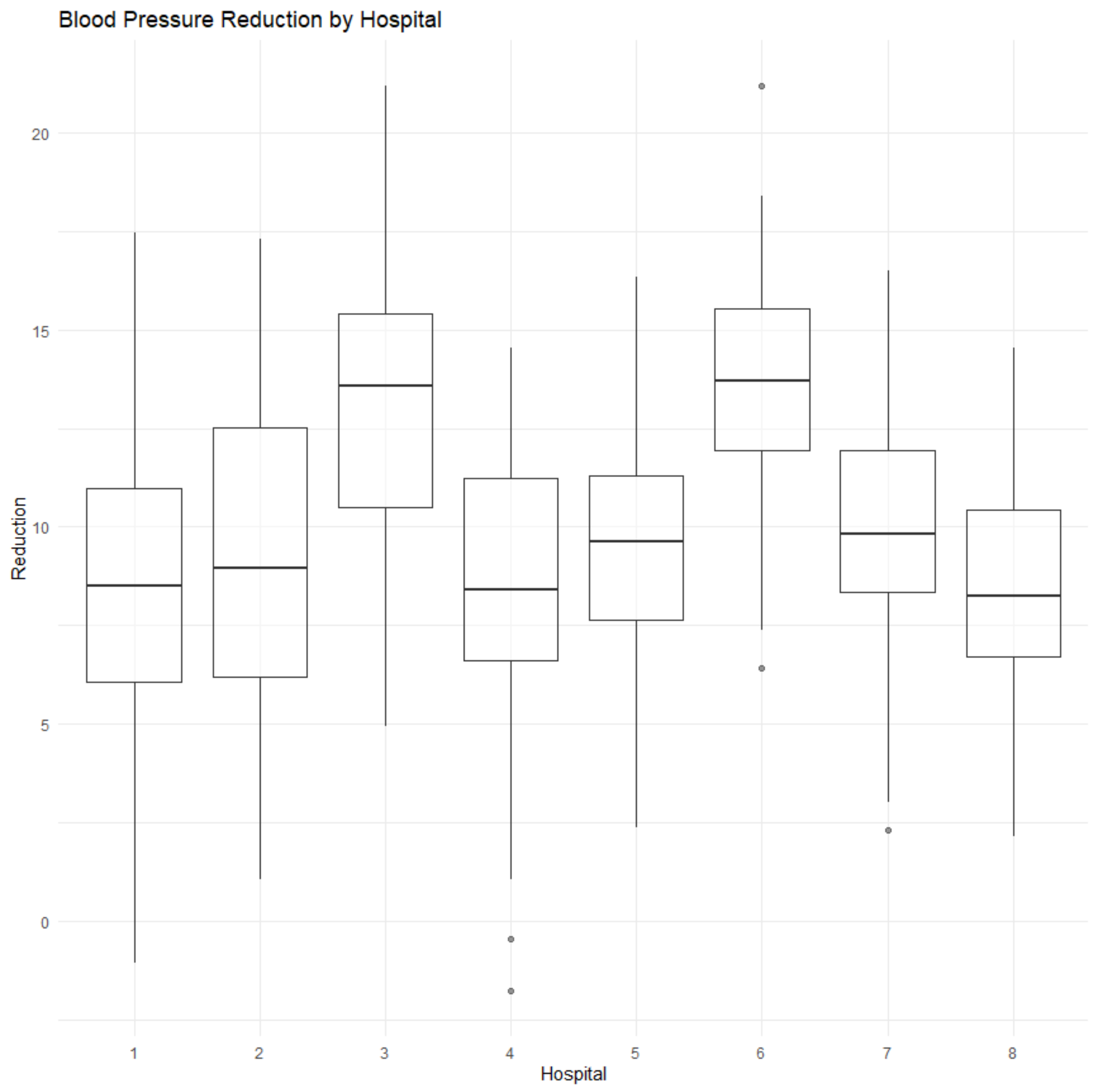


Image 3: Age vs reduction (colored by severity)

```
ggplot(datos, aes(x = edad, y = reduccion, color = factor(severidad))) +
geom_point(alpha = 0.5) +
geom_smooth(method = "lm", se = TRUE) +
theme_minimal() +
labs(
title = "Reduction vs Age by Baseline Severity",
x = "Age",
y = "Reduction",
color = "Severity"
)
```

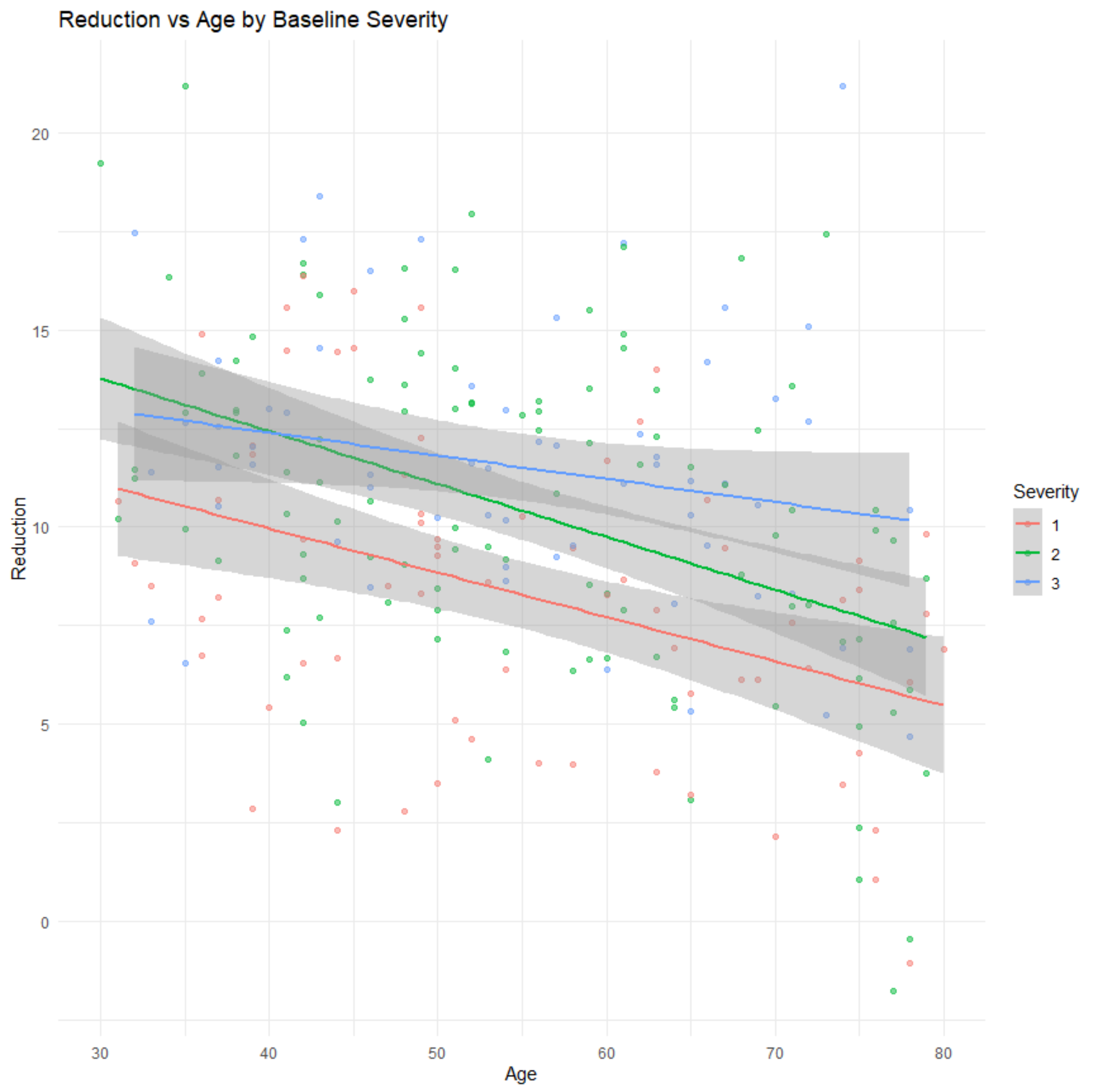
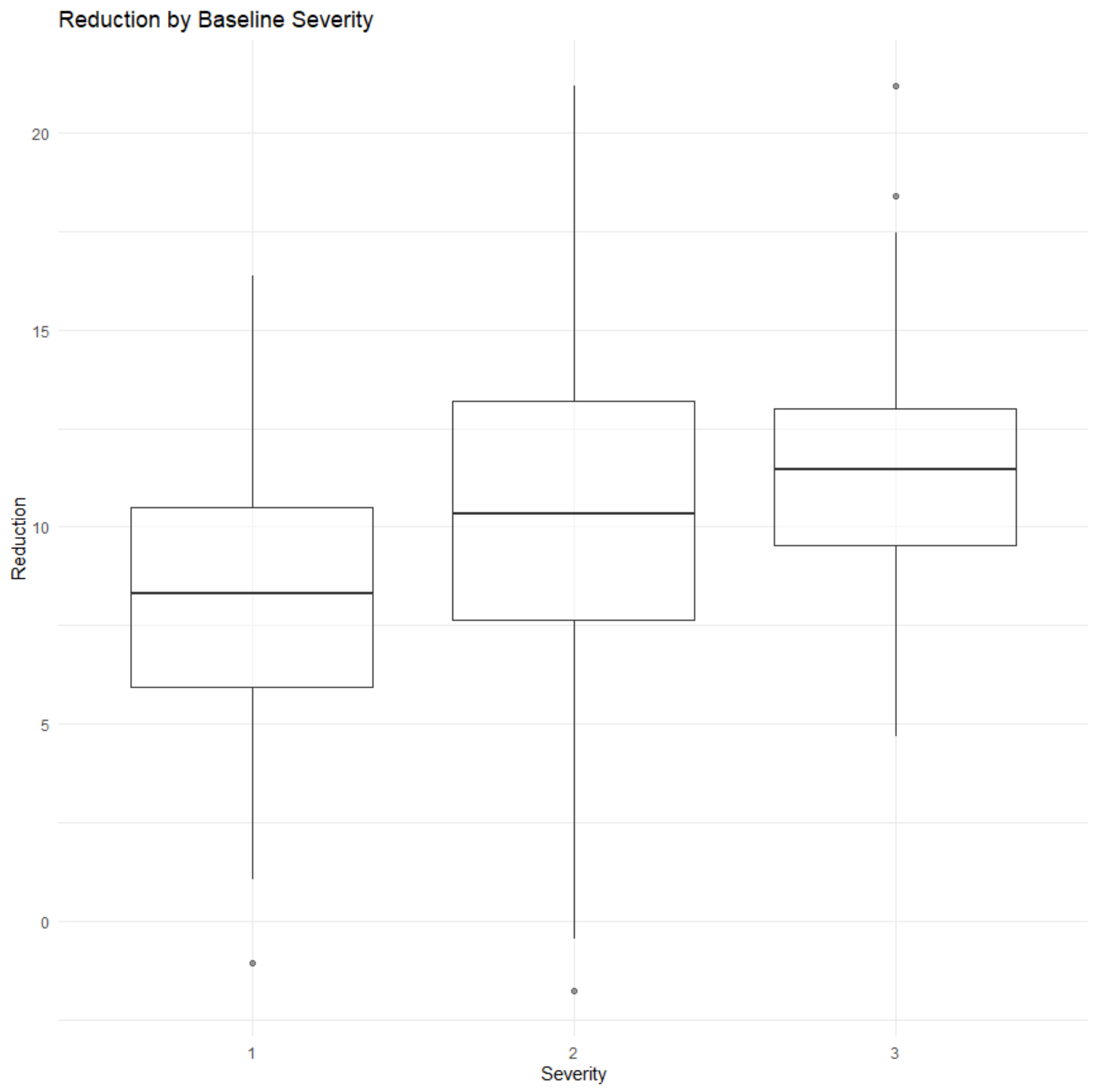


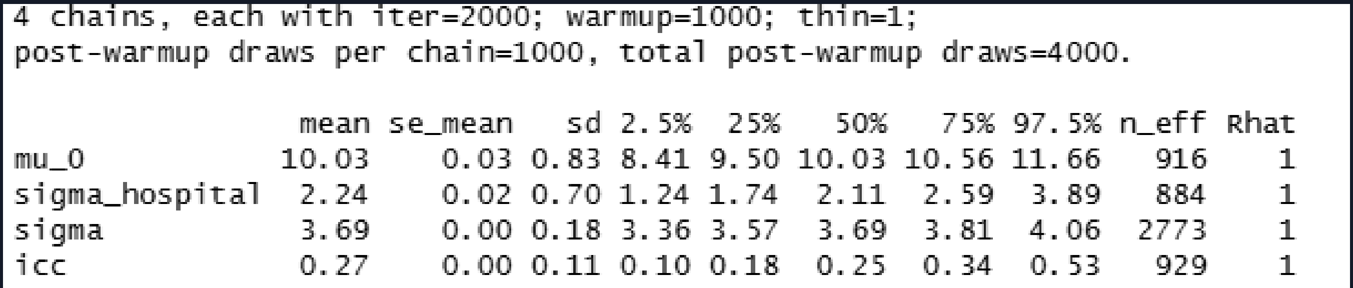
Image 4: Boxplot by severity

```
ggplot(datos, aes(x = factor(severidad), y = reduccion)) +
geom_boxplot(alpha = 0.5) +
theme_minimal() +
labs(
title = "Reduction by Baseline Severity",
x = "Severity",
y = "Reduction"
)
```



Summary by hospital

```
summary_by_hospital <- datos %>%
group_by(hospital_id) %>%
summarise(
n = n(),
mean_reduction = mean(reduccion),
sd_reduction = sd(reduccion),
min_reduction = min(reduccion),
max_reduction = max(reduccion),
.groups = "drop"
)

summary_by_hospital
```



```
4 chains, each with iter=2000; warmup=1000; thin=1;
post-warmup draws per chain=1000, total post-warmup draws=4000.

                mean se_mean   sd 2.5%  25%   50%   75% 97.5% n_eff Rhat
mu_0           10.03    0.03 0.83 8.41 9.50 10.03 10.56 11.66   916    1
sigma_hospital  2.24    0.02 0.70 1.24 1.74  2.11  2.59  3.89   884    1
sigma           3.69    0.00 0.18 3.36 3.57  3.69  3.81  4.06  2773    1
icc             0.27    0.00 0.11 0.10 0.18  0.25  0.34  0.53   929    1
```

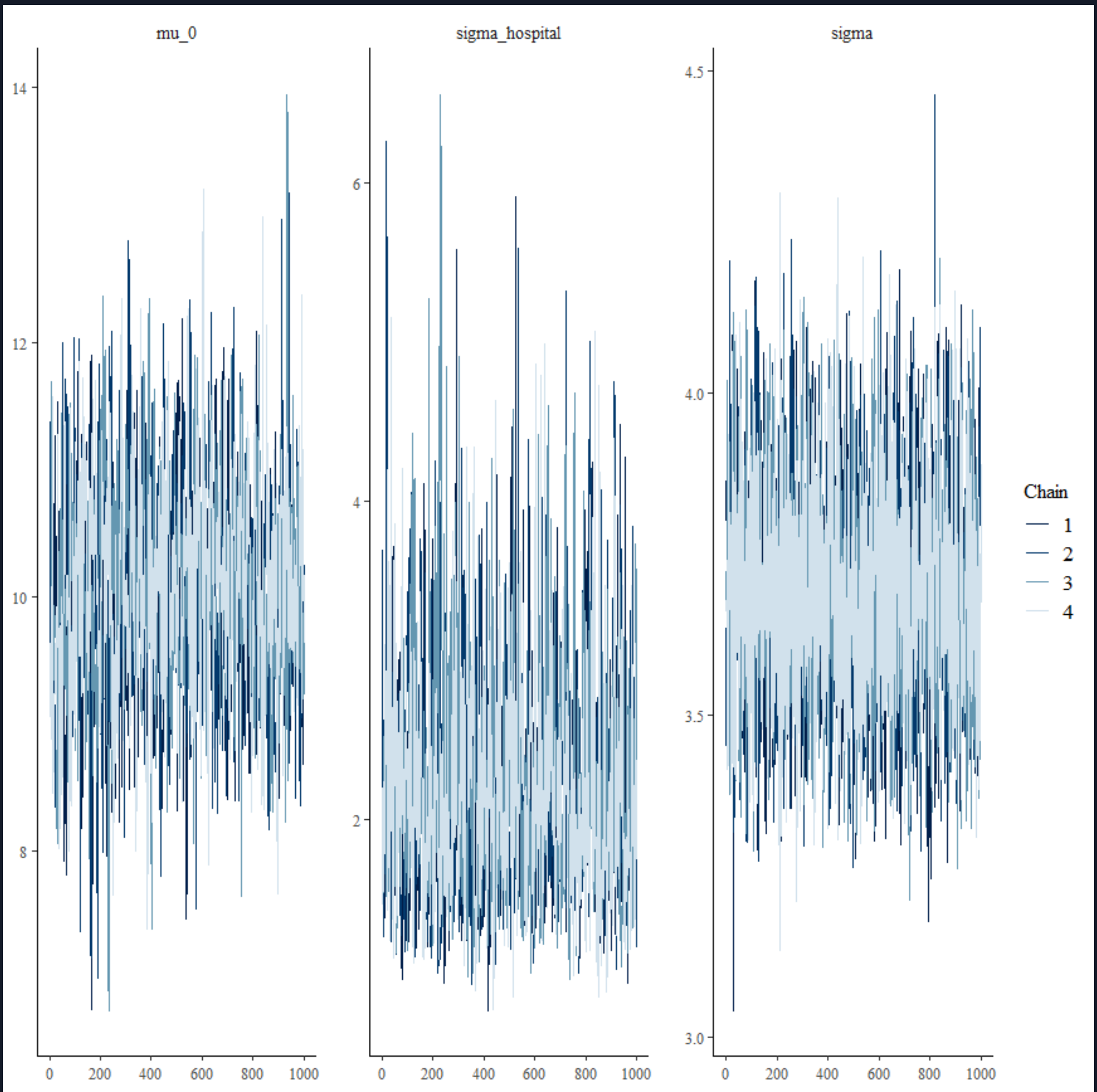
Stan model + fit

```
stan_code <- "
```

```
data {
int<lower=0> N;
int<lower=0> J;
int<lower=1,upper=J> hospital[N];
vector[N] y;
}

parameters {
real mu_0;
vector[J] eta;
real<lower=0> sigma_hospital;
real<lower=0> sigma;
}

transformed parameters {
vector[J] alpha;
alpha = sigma_hospital * eta;
}

model {
mu_0 ~ normal(10, 5);
eta ~ normal(0, 1);
sigma_hospital ~ cauchy(0, 2.5);
sigma ~ cauchy(0, 2.5);

y ~ normal(mu_0 + alpha[hospital], sigma);
}

generated quantities {
vector[N] y_pred;
real icc;

for (i in 1)
y_pred[i] = normal_rng(mu_0 + alpha[hospital[i]], sigma);

icc = square(sigma_hospital) /
(square(sigma_hospital) + square(sigma));
}
"

stan_data <- list(
N = nrow(datos),
J = n_hospitals,
hospital = datos$hospital_id,
y = datos$reduccion
)

compiled_model <- rstan::stan_model(model_code = stan_code)

fit <- rstan::sampling(
compiled_model,
data = stan_data,
chains = 4,
iter = 2000,
warmup = 1000,
seed = 123
)
```

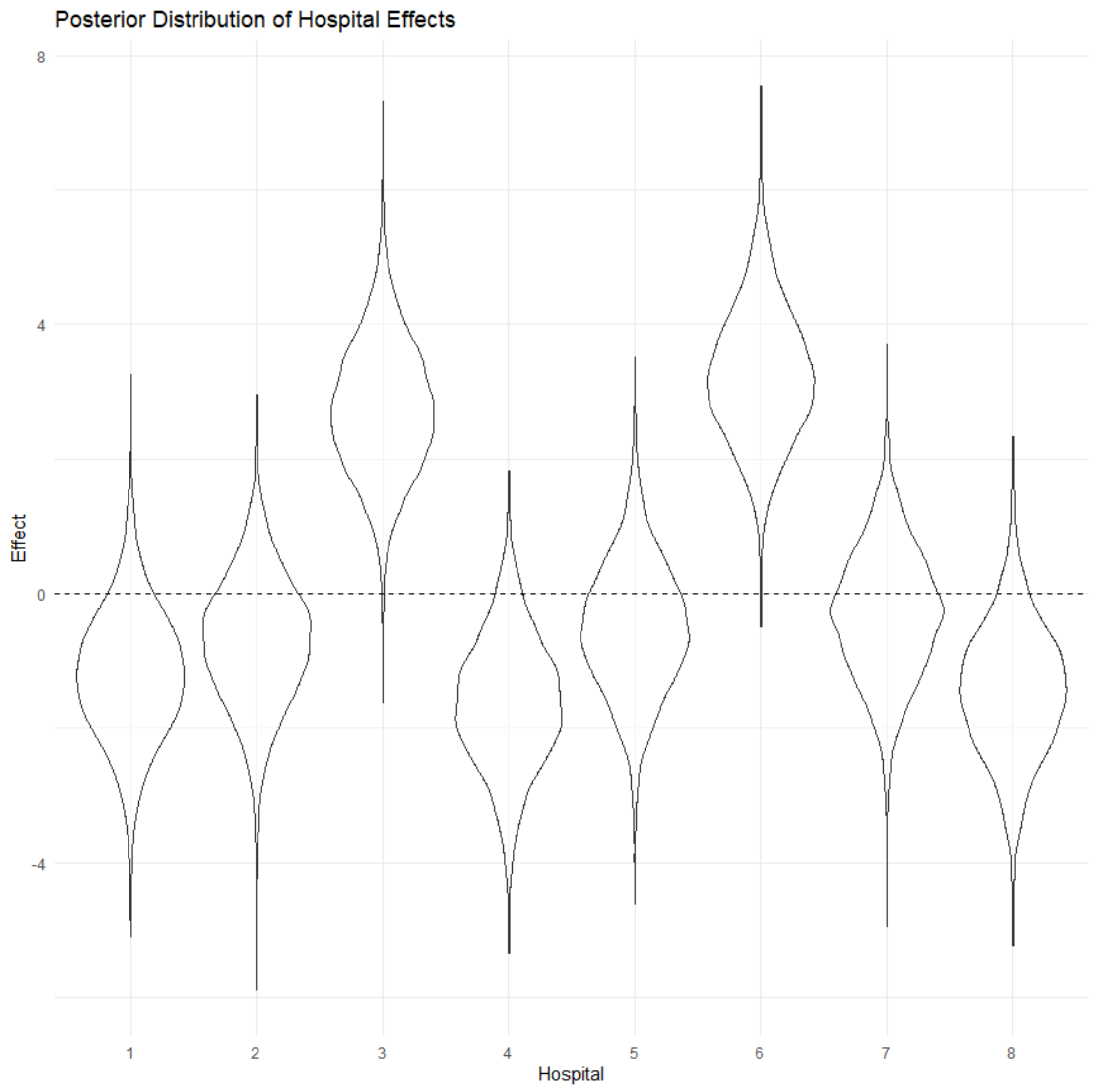
Trace plot

```
bayesplot::mcmc_trace(
as.array(fit),
pars = c("mu_0", "sigma_hospital", "sigma")
)

posterior <- rstan::extract(fit)

effects_df <- tibble(
hospital = rep(1, each = nrow(posterior$alpha)),
effect = as.vector(posterior$alpha)
)
```



Posterior hospital effects

```
ggplot(effects_df, aes(x = factor(hospital), y = effect)) +
geom_violin(alpha = 0.5) +
geom_hline(yintercept = 0, linetype = "dashed") +
theme_minimal() +
labs(
title = "Posterior Distribution of Hospital Effects",
x = "Hospital",
y = "Effect"
)
```


Posterior Distribution of Hospital Effects
Effect
8
4
0
-4
1
2
3
4
5
6
7
8
Hospital