PAFs for inclusion health

Rob Aldridge & Darwin Del Castillo

Summary of analysis

Packages to be used

The packages used in this project were the following:
- googlesheets4: to import data from Google Sheets
- metafor: to calculate the SMR based on the meta-analysis data
- tidyverse: data wrangling and visualization
- skimr: to summarize the data after importation or manipulation
- mc2d: to run Monte Carlo simulations and setting prior distributions for Gibbs sampling
- patchwork: to arrange multiple plots
- rjags: to run the Bayesian Gibbs Sampling for estimating the overlapping proportions
- coda: to summarize the results of the Bayesian Gibbs Sampling
- ggpubfigs: to use color-blind-friendly palettes

Data import and basic cleaning

First, I imported the data from the Google Sheets verifying the data types and the structure of the data. Additionally, I assigned the number of iterations to the vector `sim_run` and deactivated the scientific notation for the analysis.

Then, I filtered the SMR data from the `input_smr_all` object into the new object `smr_filter_all_cause` to include only the non-excluded or duplicate data and calculated the standard error estimates from the confidence intervals.

Further cleaning and Monte Carlo simulations

I cleaned the total mortality data, the homeless population data, the prison population data, the substance use disorder (SUD) population data, and the total population data. I also ran Monte Carlo simulations for the homeless, prison, and SUD populations to estimate the total population affected by these conditions per country with 1000 iteration to account for uncertainty.

Estimating the total inclusion health groups for each country and calculating overlaps between populations

I set the Bayesian network assumptions for the Gibbs Sampling model. The full set of assumptions for the prior distributions of the overlap between inclusion health groups are:
- 0.2 overlap between homelessness and substance drug abusers
- 0.4 overlap between substance drug abuse and incarceration
- 0.2 overlap between incarceration and homelessness
- 0.05 overlap between the three conditions
The assumption for the overlap between homelessness and substance drug disorder come from previous literature (PMID: 36660275). The other overlaps are based on good faith (no better literature was found). I also assumed independence between populations.

The vector `alpha` contains the hyperparameters to specify the Dirichlet distribution for the prior probabilities of the Gibbs Sampling model. Given the prior probabilities assumed, the counts for the Dirichlet distribution sums 60. Then, the JAGS model string was defined, and the parameters to monitor were set. I ran the Gibbs Sampling model for each country using a loop and stored the results. Finally, I summarized the results of the posterior means for each country in a data frame.

After obtaining the proportions, I calculated the total population affected by any of these conditions per country and the proportion of the population affected by inclusion health.

Calculating SMRs and PAFs

I also calculated the SMR for each country using data from the systematic review and applying a random effects meta-analysis. The results were stored in the `res_smr_all_cause_both`,

I calculated the PAFs for each country based on the inclusion health proportions and the SMR data.
The functions to calculate the PAFs and the attributable deaths are stored in the `functions.R` file.
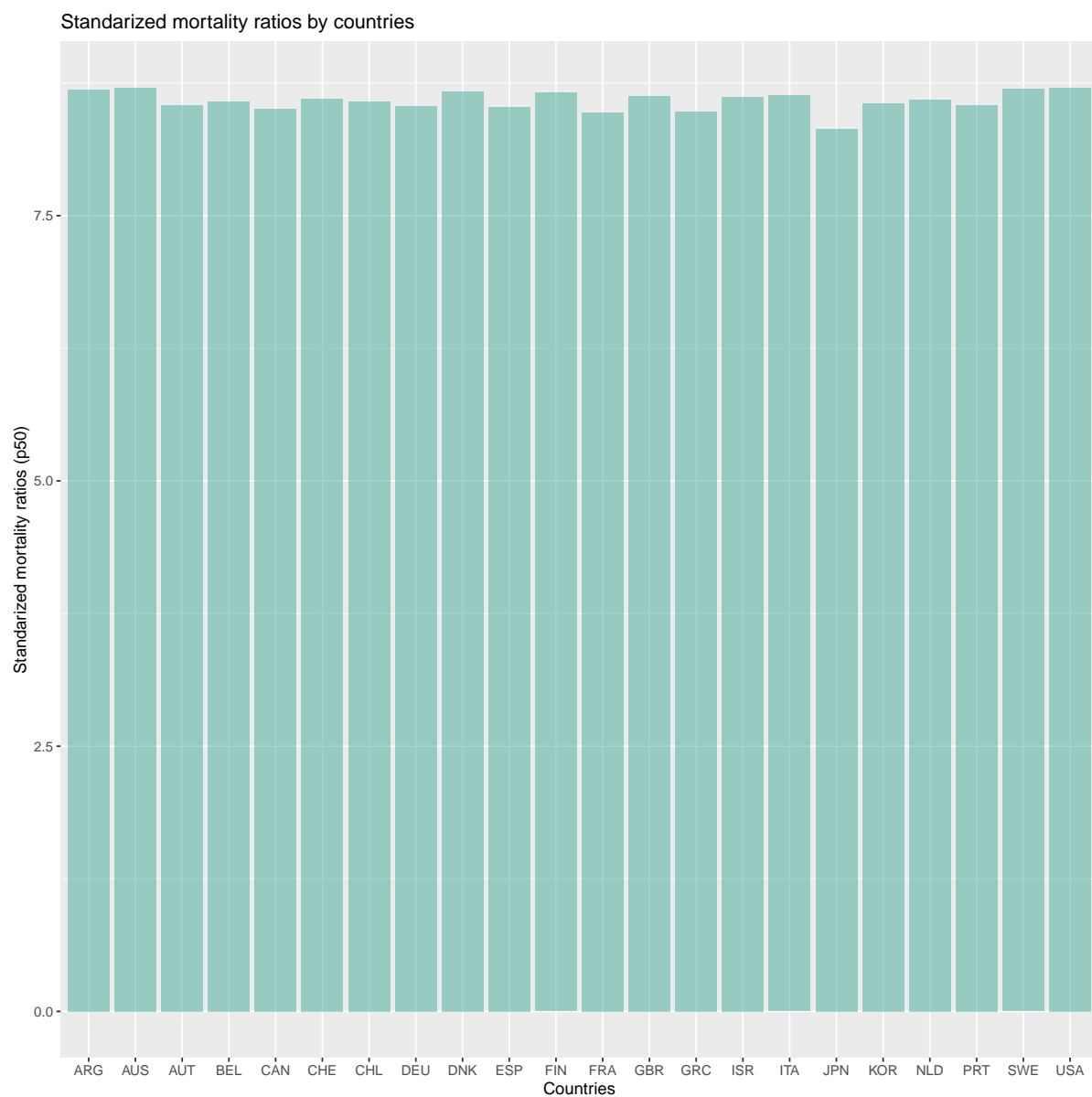
Graphs

Lastly, I created graphs that will help with the visualization of the results.
The first set of graphs shows the PAFs for inclusion health for each country stratified by sex.
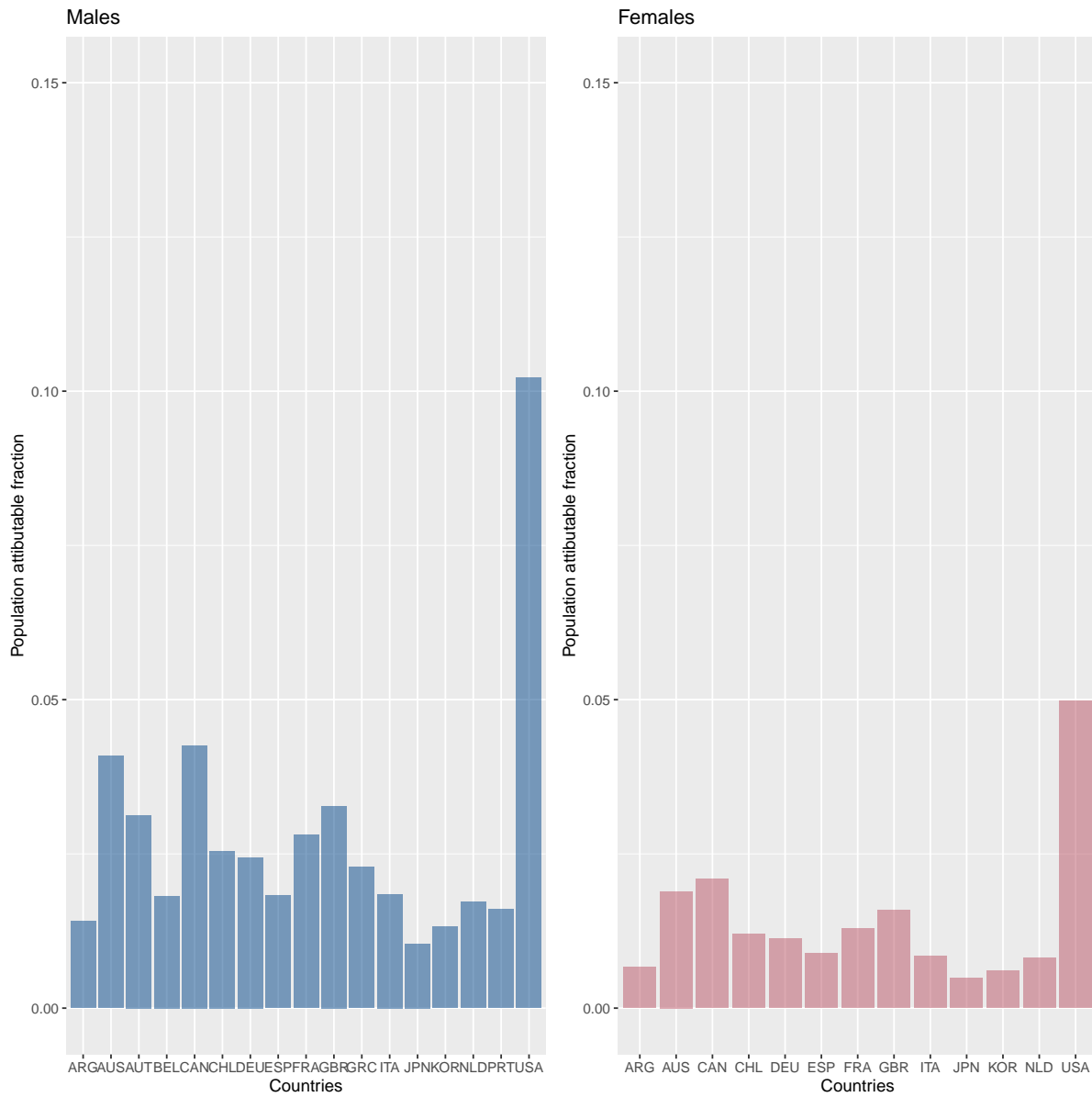
Second, I created a graph showing the proportion of inclusion health for each country.

Proportion of socially–excluded groups by countries
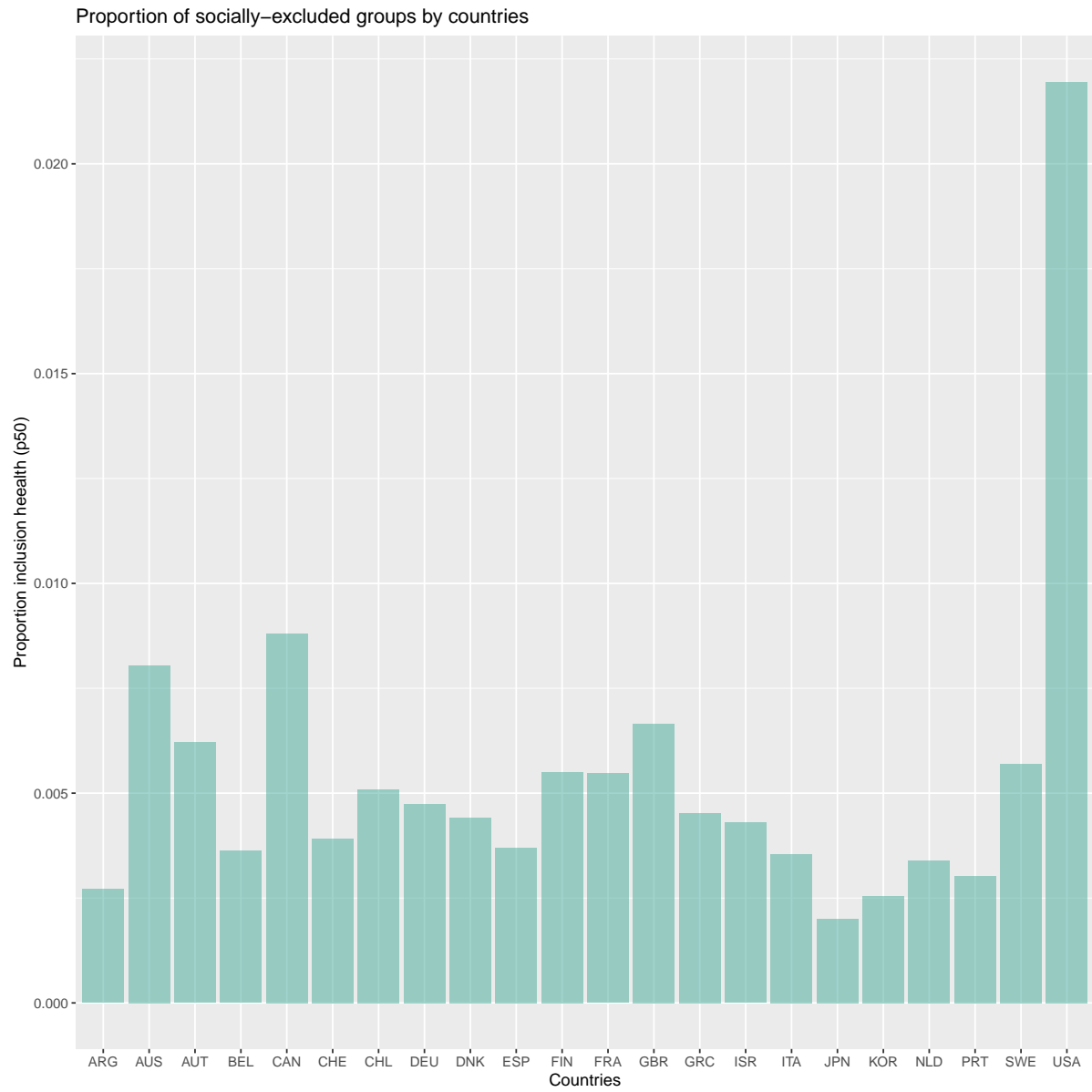
Third, I created a graph showing the SMR for each country.

Standarized mortality ratios by countries

Finally, I created a graph showing the total and attributable deaths for each country stratified by sex.