# Load your dataset

data <- read\_dta("C:/Users/c23086899/OneDrive - Cardiff University/Documents/nsw\_2021-23\_anonymised\_respondent\_file.dta")

# Simulate the model\_summary dataset

model\_summary <- data.frame(

term = c("BloodOrgans", "DigestiveSystem", "EarComplaints",

"EndocrineDiseases", "GenitoUrinaryIllness",

"CirculatoryIllness", "MentalDisorder",

"MusculoskeletalIllness", "Neoplasm",

"NervousSystem", "RespiratorySystem"),

estimate = c(-0.87, -1.27, -0.50, -0.96, -1.14,

-1.05, -1.66, -1.77, -2.12,

-2.04, -1.21),

conf.low = c(-1.16, -1.48, -0.71, -1.23, -1.39,

-1.28, -1.86, -2.01, -2.34,

-2.27, -1.44),

conf.high = c(-0.57, -1.06, -0.28, -0.68, -0.89,

-0.81, -1.46, -1.53, -1.90,

-1.80, -0.99),

p.value = c(0.001, 0.001, 0.001, 0.001, 0.001,

0.001, 0.001, 0.001, 0.001,

0.001, 0.001)

)

# Recode variables for clarity in the plot

plot\_data <- model\_summary %>%

filter(term != "(Intercept)") %>%

mutate(

Variable = case\_when(

grepl("NervousSystem", term) ~ "Nervous system",

grepl("Neoplasm", term) ~ "Neoplasms & benign growths",

grepl("MusculoskeletalIllness", term) ~ "Musculoskeletal system",

grepl("MentalDisorder", term) ~ "Mental disorders",

grepl("RespiratorySystem", term) ~ "Respiratory system",

grepl("DigestiveSystem", term) ~ "Digestive system",

grepl("GenitoUrinaryIllness", term) ~ "Genito-urinary system",

grepl("CirculatoryIllness", term) ~ "Heart & circulatory system",

grepl("EarComplaints", term) ~ "Ear complaints",

grepl("EndocrineDiseases", term) ~ "Endocrine & metabolic",

grepl("BloodOrgans", term) ~ "Blood Organs",

TRUE ~ NA\_character\_

),

Level = "Has condition"

) %>%

mutate(

N = case\_when(

Variable == "Nervous system" & Level == "No condition present" ~ 22724,

Variable == "Nervous system" & Level == "Has condition" ~ 701,

Variable == "Neoplasms & benign growths" & Level == "No condition present" ~ 22764,

Variable == "Neoplasms & benign growths" & Level == "Has condition" ~ 661,

Variable == "Musculoskeletal system" & Level == "No condition present" ~ 18456,

Variable == "Musculoskeletal system" & Level == "Has condition" ~ 4969,

Variable == "Mental disorders" & Level == "No condition present" ~ 20945,

Variable == "Mental disorders" & Level == "Has condition" ~ 2480,

Variable == "Respiratory system" & Level == "No condition present" ~ 21300,

Variable == "Respiratory system" & Level == "Has condition" ~ 2125,

Variable == "Blood Organs" & Level == "No condition present" ~ 23184,

Variable == "Blood Organs" & Level == "Has condition" ~ 241,

Variable == "Digestive system" & Level == "No condition present" ~ 22292,

Variable == "Digestive system" & Level == "Has condition" ~ 1133,

Variable == "Heart & circulatory system" & Level == "No condition present" ~ 19986,

Variable == "Heart & circulatory system" & Level == "Has condition" ~ 3440,

Variable == "Genito-urinary system" & Level == "No condition present" ~ 22772,

Variable == "Genito-urinary system" & Level == "Has condition" ~ 653,

Variable == "Endocrine & metabolic" & Level == "No condition present" ~ 21182,

Variable == "Endocrine & metabolic" & Level == "Has condition" ~ 2243,

Variable == "Ear complaints" & Level == "No condition present" ~ 22861,

Variable == "Ear complaints" & Level == "Has condition" ~ 564,

TRUE ~ NA\_real\_

)

) %>%

mutate(

Odds\_ratio = sprintf("%.2f (%.2f, %.2f)", exp(estimate), exp(conf.low), exp(conf.high)),

p\_value = ifelse(p.value < 0.001, "<0.001", sprintf("%.3f", p.value))

)

# Add reference rows

reference\_rows <- plot\_data %>%

distinct(Variable) %>%

mutate(

Level = "No condition present",

estimate = 0,

conf.low = 0,

conf.high = 0,

Odds\_ratio = "Reference",

p\_value = "Reference",

N = case\_when(

Variable == "Nervous system" ~ 22724,

Variable == "Neoplasms & benign growths" ~ 22764,

Variable == "Musculoskeletal system" ~ 18456,

Variable == "Mental disorders" ~ 20945,

Variable == "Respiratory system" ~ 21300,

Variable == "Blood Organs" ~ 23184,

Variable == "Digestive system" ~ 22292,

Variable == "Heart & circulatory system" ~ 19986,

Variable == "Genito-urinary system" ~ 22772,

Variable == "Endocrine & metabolic" ~ 21182,

Variable == "Ear complaints" ~ 22861,

TRUE ~ NA\_real\_

)

)

# Combine plot\_data and reference\_rows

plot\_data <- bind\_rows(plot\_data, reference\_rows) %>%

arrange(Variable, desc(Level)) %>%

mutate(row\_id = row\_number())

# Load required libraries

library(tidyverse)

library(gtable)

library(grid)

library(gridExtra)

# Assuming your data is loaded and recoded (skipping for brevity)

# Prepare plot data (assuming model\_summary is available)

plot\_data <- model\_summary %>%

filter(term != "(Intercept)") %>%

mutate(

Variable = case\_when(

grepl("NervousSystem", term) ~ "Nervous system",

grepl("Neoplasm", term) ~ "Neoplasms & benign growths",

grepl("MusculoskeletalIllness", term) ~ "Musculoskeletal system",

grepl("MentalDisorder", term) ~ "Mental disorders",

grepl("RespiratorySystem", term) ~ "Respiratory system",

grepl("DigestiveSystem", term) ~ "Digestive system",

grepl("GenitoUrinaryIllness", term) ~ "Genito-urinary system",

grepl("CirculatoryIllness", term) ~ "Heart & circulatory system",

grepl("EarComplaints", term) ~ "Ear complaints",

grepl("EndocrineDiseases", term) ~ "Endocrine & metabolic",

grepl("BloodOrgans", term) ~ "Blood Organs",

TRUE ~ NA\_character\_

),

Level = "Has condition"

) %>%

mutate(

N = case\_when(

Variable == "Nervous system" & Level == "No condition present" ~ 22724,

Variable == "Nervous system" & Level == "Has condition" ~ 701,

Variable == "Neoplasms & benign growths" & Level == "No condition present" ~ 22764,

Variable == "Neoplasms & benign growths" & Level == "Has condition" ~ 661,

Variable == "Musculoskeletal system" & Level == "No condition present" ~ 18456,

Variable == "Musculoskeletal system" & Level == "Has condition" ~ 4969,

Variable == "Mental disorders" & Level == "No condition present" ~ 20945,

Variable == "Mental disorders" & Level == "Has condition" ~ 2480,

Variable == "Respiratory system" & Level == "No condition present" ~ 21300,

Variable == "Respiratory system" & Level == "Has condition" ~ 2125,

Variable == "Blood Organs" & Level == "No condition present" ~ 23184,

Variable == "Blood Organs" & Level == "Has condition" ~ 241,

Variable == "Digestive system" & Level == "No condition present" ~ 22292,

Variable == "Digestive system" & Level == "Has condition" ~ 1133,

Variable == "Heart & circulatory system" & Level == "No condition present" ~ 19986,

Variable == "Heart & circulatory system" & Level == "Has condition" ~ 3440,

Variable == "Genito-urinary system" & Level == "No condition present" ~ 22772,

Variable == "Genito-urinary system" & Level == "Has condition" ~ 653,

Variable == "Endocrine & metabolic" & Level == "No condition present" ~ 21182,

Variable == "Endocrine & metabolic" & Level == "Has condition" ~ 2243,

Variable == "Ear complaints" & Level == "No condition present" ~ 22861,

Variable == "Ear complaints" & Level == "Has condition" ~ 564,

TRUE ~ NA\_real\_ # Use NA if no match

)

) %>%

mutate(

Odds\_ratio = sprintf("%.2f (%.2f, %.2f)", exp(estimate), exp(conf.low), exp(conf.high)),

p\_value = ifelse(p.value < 0.001, "<0.001", sprintf("%.3f", p.value))

)

# Add reference rows

reference\_rows <- plot\_data %>%

distinct(Variable) %>%

mutate(

Level = "No condition present",

estimate = 0,

conf.low = 0,

conf.high = 0,

Odds\_ratio = "Reference",

p\_value = "Reference",

N = case\_when(

Variable == "Nervous system" ~ 22724,

Variable == "Neoplasms & benign growths" ~ 22764,

Variable == "Musculoskeletal system" ~ 18456,

Variable == "Mental disorders" ~ 20945,

Variable == "Respiratory system" ~ 21300,

Variable == "Blood Organs" ~ 23184,

Variable == "Digestive system" ~ 22292,

Variable == "Heart & circulatory system" ~ 19986,

Variable == "Genito-urinary system" ~ 22772,

Variable == "Endocrine & metabolic" ~ 21182,

Variable == "Ear complaints" ~ 22861,

TRUE ~ NA\_real\_

)

)

# Combine plot\_data and reference\_rows

plot\_data <- bind\_rows(plot\_data, reference\_rows) %>%

arrange(Variable, desc(Level)) %>%

mutate(row\_id = row\_number())

# Create adjusted estimates for odds ratios and their confidence intervals

plot\_data <- plot\_data %>%

mutate(

adj\_estimate = ifelse(exp(estimate) < 1, 1 / exp(estimate), exp(estimate)),

adj\_conf\_low = ifelse(exp(conf.low) < 1, 1 / exp(conf.high), exp(conf.low)),

adj\_conf\_high = ifelse(exp(conf.high) < 1, 1 / exp(conf.low), exp(conf.high))

)

# Filter out "Infectious Disease"

plot\_data <- plot\_data %>%

filter(Variable != "Infectious Disease")

# Ensure the confidence interval for Neoplasms & benign growths is not NA

plot\_data <- plot\_data %>%

mutate(

adj\_conf\_low = ifelse(Variable == "Neoplasms & benign growths" & Level == "Has condition",

exp(conf.low), adj\_conf\_low),

adj\_conf\_high = ifelse(Variable == "Neoplasms & benign growths" & Level == "Has condition",

exp(conf.high), adj\_conf\_high)

)

# Modify the table\_data to include the Odds Ratio column

table\_data <- plot\_data %>%

dplyr::select(Variable, Level, N, Odds\_ratio, p\_value) %>%

group\_by(Variable) %>%

mutate(Variable = ifelse(Level == "No condition present", Variable, "")) %>%

ungroup()

# Create the table

gt <- tableGrob(table\_data, rows = NULL,

theme = ttheme\_minimal(

core = list(fg\_params = list(hjust = 0, x = 0.1), bg\_params = list(fill = "white")),

colhead = list(fg\_params = list(hjust = 0, x = 0.1, fontface = "bold"), bg\_params = list(fill = "white"))

)

)

# Add the forest plot column

gt <- gtable\_add\_cols(gt, unit(4.5, "inches"), 3)

gt <- gtable\_add\_grob(gt, forest\_grob, t = 2, b = nrow(gt), l = 4, r = 4)

# Rename the Odds\_ratio column header

gt$grob[[1]]$grobs[[1]]$label[5] <- "Odds Ratio"

# Adding x-axis title

x\_axis\_title <- textGrob(

label = "Adjusted Odds Ratio (log scale)",

x = unit(0.5, "npc"),

just = "center",

gp = gpar(fontsize = 10, fontface = "bold")

)

# Add x-axis title to the table

gt <- gtable\_add\_rows(gt, heights = unit(1.5, "line"), pos = nrow(gt))

gt <- gtable\_add\_grob(gt, x\_axis\_title, t = nrow(gt), l = 4, r = 4)

# Update x-axis labels

x\_axis\_labels <- textGrob(

label = c("1", "2", "3", "4", "5", "6"),

x = unit(c(0.05, 0.205, 0.405, 0.605, 0.806, 1), "npc"),

y = unit(0.5, "npc"),

just = "center",

gp = gpar(fontsize = 10)

)

gt <- gtable\_add\_rows(gt, heights = unit(1, "line"), pos = nrow(gt))

gt <- gtable\_add\_grob(gt, x\_axis\_labels, t = nrow(gt), l = 4, r = 4)

# Adjust the data to ensure all values start from 1

plot\_data <- plot\_data %>%

mutate(

adj\_estimate = pmax(adj\_estimate, 1),

adj\_conf\_low = pmax(adj\_conf\_low, 1),

adj\_conf\_high = pmax(adj\_conf\_high, 1)

)

# Modify the forest plot

forest\_plot <- ggplot(plot\_data, aes(y = reorder(interaction(Variable, Level), desc(row\_id)), x = adj\_estimate)) +

geom\_vline(xintercept = 1, linetype = "dashed", color = "gray50") +

geom\_errorbarh(aes(xmin = adj\_conf\_low, xmax = adj\_conf\_high), height = 0.2) +

geom\_point(size = 2) +

scale\_x\_continuous(

trans = scales::log2\_trans(),

breaks = c(1, 2, 4, 6, 8, 10),

labels = c("1", "2", "4", "6", "8", "10"),

limits = c(0.9, 10),

expand = c(0, 0)

) +

coord\_cartesian(xlim = c(1, 10)) +

labs(x = NULL, y = "") +

theme\_minimal() +

theme(

axis.text.y = element\_blank(),

panel.grid.minor = element\_blank(),

panel.grid.major.y = element\_blank(),

axis.text.x = element\_blank(),

axis.ticks.x = element\_blank(),

plot.margin = margin(5.5, 5.5, 5.5, 5.5)

)

# Convert ggplot to grob

forest\_grob <- ggplotGrob(forest\_plot)

# Create the table

gt <- tableGrob(table\_data, rows = NULL,

theme = ttheme\_minimal(

core = list(fg\_params = list(hjust = 0, x = 0.1), bg\_params = list(fill = "white")),

colhead = list(fg\_params = list(hjust = 0, x = 0.1, fontface = "bold"), bg\_params = list(fill = "white"))

)

)

# Add the forest plot column

gt <- gtable\_add\_cols(gt, unit(4.5, "inches"), 3)

gt <- gtable\_add\_grob(gt, forest\_grob, t = 2, b = nrow(gt), l = 4, r = 4)

# Rename the Odds\_ratio column header

gt$grob[[1]]$grobs[[1]]$label[5] <- "Odds Ratio"

# Adjust column widths

gt$widths <- unit(c(2, 1.8, 1, 3.5, 1.5, 1), "inches") # Reduced width for the Variable column

# Update the forest plot to include error bars for all rows

forest\_plot <- forest\_plot +

geom\_errorbarh(data = plot\_data,

aes(xmin = adj\_conf\_low, xmax = adj\_conf\_high),

height = 0.2)

# Ensure the Neoplasms & benign growths "Has condition" row has error bars

plot\_data <- plot\_data %>%

mutate(

adj\_conf\_low = ifelse(is.na(adj\_conf\_low), adj\_estimate, adj\_conf\_low),

adj\_conf\_high = ifelse(is.na(adj\_conf\_high), adj\_estimate, adj\_conf\_high)

)

# Update the forest plot to include error bars for all rows

forest\_plot <- forest\_plot +

geom\_errorbarh(data = plot\_data,

aes(xmin = adj\_conf\_low, xmax = adj\_conf\_high),

height = 0.2)

# Convert updated ggplot to grob

forest\_grob <- ggplotGrob(forest\_plot)

# Add x-axis labels to the table

gt <- gtable\_add\_rows(gt, heights = unit(1, "line"), pos = nrow(gt))

gt <- gtable\_add\_grob(gt, x\_axis\_labels, t = nrow(gt), l = 4, r = 4)

# Add x-axis title

x\_axis\_title <- textGrob(

label = "Adjusted Odds Ratio (log scale)",

x = unit(0.5, "npc"),

y = unit(-1, "lines"),

just = c("center", "top"),

gp = gpar(fontsize = 10)

)

# Add x-axis title to the table

gt <- gtable\_add\_rows(gt, heights = unit(2, "lines"), pos = nrow(gt))

gt <- gtable\_add\_grob(gt, x\_axis\_title, t = nrow(gt), l = 4, r = 4)

# Ensure the forest plot column width remains the same

gt$widths[4] <- unit(5, "inches")

# Add outer border and gridlines

gt <- gtable::gtable\_add\_grob(gt,

grobs = rectGrob(gp = gpar(fill = NA, lwd = 2, col = "black")),

t = 1, b = nrow(gt), l = 1, r = ncol(gt)

)

for (i in 1:(nrow(gt) - 1)) {

gt <- gtable::gtable\_add\_grob(gt,

grobs = segmentsGrob(x0 = unit(0, "npc"),

x1 = unit(1, "npc"),

y0 = unit(1, "npc"),

y1 = unit(1, "npc"),

gp = gpar(lwd = 0.5, col = "black")),

t = i, b = i, l = 1, r = ncol(gt)

)

}

for (i in 1:(ncol(gt) - 1)) {

gt <- gtable::gtable\_add\_grob(gt,

grobs = segmentsGrob(x0 = unit(1, "npc"),

x1 = unit(1, "npc"),

y0 = unit(0, "npc"),

y1 = unit(1, "npc"),

gp = gpar(lwd = 0.5, col = "black")),

t = 1, b = nrow(gt), l = i, r = i

)

}

# Add title

title <- textGrob("Adjusted Odds Ratios for Fair/Bad Health (Conditions)",

gp = gpar(fontface = "bold", fontsize = 14)

)

gt <- gtable\_add\_rows(gt, heights = grobHeight(title) + unit(0.5, "line"), pos = 0)

gt <- gtable\_add\_grob(gt, title, t = 1, l = 1, r = ncol(gt))

# Display the final plot

grid.newpage()

grid.draw(gt)

# Save the plot

ggsave("adjusted\_forest\_plot\_final\_v5.png", gt, width = 15, height = 10, dpi = 300)