

# Diet

2025-11-05

## Table of contents

<b>Setting up in R</b>	<b>2</b>
Load the data . . . . .	2
Data cleaning . . . . .	3
Total meat protein . . . . .	6
Crohn's disease . . . . .	6
Patient reported flare . . . . .	6
Objective flare . . . . .	8
Ulcerative colitis . . . . .	10
Patient reported flare . . . . .	10
Objective flare . . . . .	12
Dietary Fibre . . . . .	14
Crohn's disease . . . . .	14
Patient reported flare . . . . .	14
Objective flare . . . . .	16
Ulcerative colitis . . . . .	18
Patient reported flare . . . . .	18
Objective flare . . . . .	20
Polyunsaturated fatty acids . . . . .	22
Crohn's disease . . . . .	22
Patient reported flare . . . . .	22
Objective flare . . . . .	24
Ulcerative colitis . . . . .	26
Patient reported flare . . . . .	26
Objective flare . . . . .	28
Ultra-processed food . . . . .	30
Crohn's disease . . . . .	30
Patient reported flare . . . . .	30
Objective flare . . . . .	32

Ulcerative colitis . . . . .	34
Patient reported flare . . . . .	34
Objective flare . . . . .	36

## Setting up in R

### Load the data

```

library(splines)
library(patchwork)

source("Survival/utils.R")

# Setup analysis environment
analysis_setup <- setup_analysis()
paths <- analysis_setup$paths
demo <- analysis_setup$demo

flare.df <- readRDS(paste0(paths$outdir, "flares-biochem.RDS"))
flare.cd.df <- readRDS(paste0(paths$outdir, "flares-biochem-cd.RDS"))
flare.uc.df <- readRDS(paste0(paths$outdir, "flares-biochem-uc.RDS"))

summon_plot_broom_hr <- function(data){
  data %>%
    # Remove frailty
    dplyr::filter(!stringr::str_detect(term, "frailty")) %>%
    # Rename confidence intervals only if they don't already exist
    {if(!"conf.low" %in% names(.)) dplyr::rename(., conf.low = `2.5 %`, conf.high = `97.5 %`)}
    # Significance flag
    dplyr::mutate(
      significant = (p.value <= 0.05)
    ) %>%
    ggplot(aes(
      y =forcats::as_factor(term),
      x = estimate,
      xmin = conf.low,
      xmax = conf.high,
      colour = significant)) +
    geom_point() +
    geom_errorbarh() +
  
```

```

    geom_vline(xintercept = 1, linetype = "dashed") +
  xlab("Hazard Ratio (95% CI)") +
  ylab("") +
  scale_colour_manual(values = c("TRUE" = "red", "FALSE" = "black")) +
  theme_minimal()
}

```

## Data cleaning

```

# FC has been logged twice - reverse
flare.df %<>%
  dplyr::mutate(FC = exp(FC))

flare.cd.df %<>%
  dplyr::mutate(FC = exp(FC))

flare.uc.df %<>%
  dplyr::mutate(FC = exp(FC))
# So now FC is the log of the FC measurement

# Transform age variable to decades
flare.df %<>%
  dplyr::mutate(
    age_decade = Age/10
  )

flare.cd.df %<>%
  dplyr::mutate(
    age_decade = Age/10
  )

flare.uc.df %<>%
  dplyr::mutate(
    age_decade = Age/10
  )

# Categorised continuous variables

# Categorize meat protein by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "Meat_sum", reference_data = flare.df)

```

```

flare.uc.df <- categorize_by_quantiles(flare.uc.df, "Meat_sum", reference_data = flare.df)

# Categorize dietary fibre by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "fibre", reference_data = flare.df)
flare.uc.df <- categorize_by_quantiles(flare.uc.df, "fibre", reference_data = flare.df)

# Categorize PUFA by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "PUFA_percEng", reference_data = flare.df)
flare.uc.df <- categorize_by_quantiles(flare.uc.df, "PUFA_percEng", reference_data = flare.df)

# Categorize UPF percentage by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "UPF_perc", reference_data = flare.df)
flare.uc.df <- categorize_by_quantiles(flare.uc.df, "UPF_perc", reference_data = flare.df)

# Smoking
# Reorder Smoke factor levels
flare.cd.df %<>%
  dplyr::mutate(
    Smoke = forcats::fct_relevel(Smoke, "Never", "Previous", "Current")
  )

flare.uc.df %<>%
  dplyr::mutate(
    Smoke = forcats::fct_relevel(Smoke, "Never", "Previous", "Current")
  )

# Missingness in smoking
# CD
flare.cd.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                Meat_sum_cat,
                Smoke,
                SiteNo) %>%
  # Remove missing on all columns except smoke

```

```

dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
dplyr::pull(Smoke) %>%
forcats::fct_count(prop = TRUE)

```

```

# A tibble: 4 x 3
  f          n      p
  <fct>    <int>  <dbl>
1 Never     268  0.559
2 Previous   170  0.355
3 Current    31  0.0647
4 <NA>       10  0.0209

```

```

# UC
flare.uc.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                Meat_sum_cat,
                Smoke,
                SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
dplyr::pull(Smoke) %>%
forcats::fct_count(prop = TRUE)

```

```

# A tibble: 4 x 3
  f          n      p
  <fct>    <int>  <dbl>
1 Never     268  0.541
2 Previous   184  0.372
3 Current    29  0.0586
4 <NA>       14  0.0283

```

## Total meat protein

### Crohns' disease

#### Patient reported flare

```
# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others
# Calculate cumulative hazard

# Soft flare
data_impute_meat_cd_soft <- flare.cd.df %>%
  dplyr::select(softflare_time,
                softflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                Meat_sum_cat,
                Smoke,
                SiteNo) %>%
  # Remove missing on all columns except smoke
  dplyr::filter(!dplyr::if_any(
    .cols = -Smoke,
    .fns = is.na
  )) %>%
  # Calculate Cumulative hazard
  dplyr::mutate(
    cumhaz = mice::nelsonalen(
      data =.,
      timevar = softflare_time,
      statusvar = softflare
    )
  )

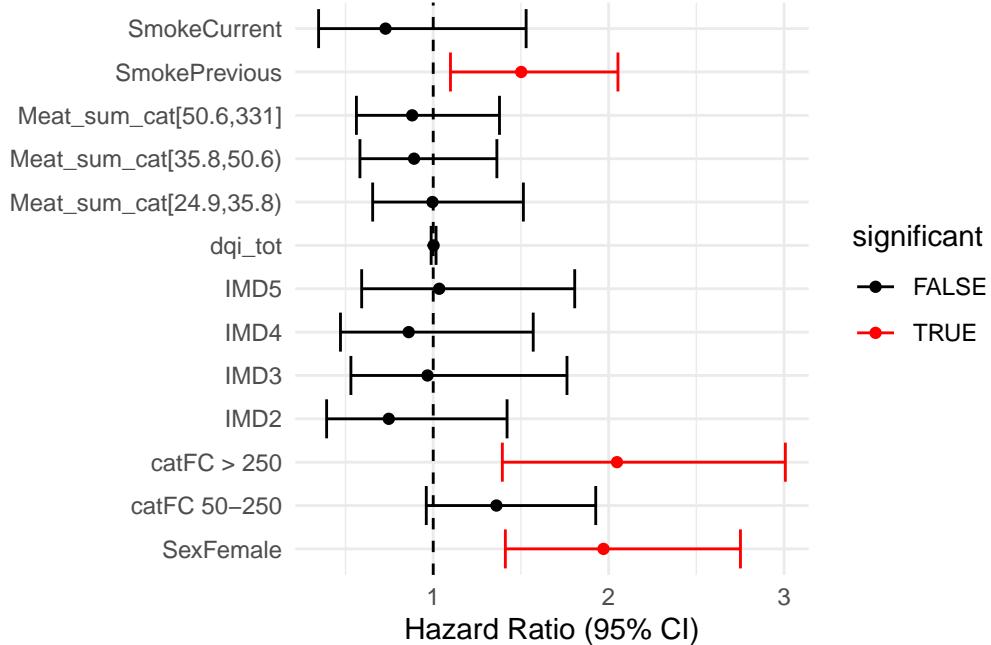
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_meat_cd_soft)

pred_matrix[, 'softflare_time'] <- 0

# MICE with 10 imputations
mice_meat_cd_soft <- mice::mice(
  data = data_impute_meat_cd_soft,
```

```
predictorMatrix = pred_matrix,
m = 10,
maxit = 20,
seed = 73,
print = FALSE
)

# Fit pooled Cox model
with(
  mice_meat_cd_soft,
  coxph(
    Surv(softflare_time, softflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      Meat_sum_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()
```



## Objective flare

```
# Hard flare
data_impute_meat_cd_hard <- flare.cd.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                Meat_sum_cat,
                Smoke,
                SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
```

```

    timevar = hardflare_time,
    statusvar = hardflare
))

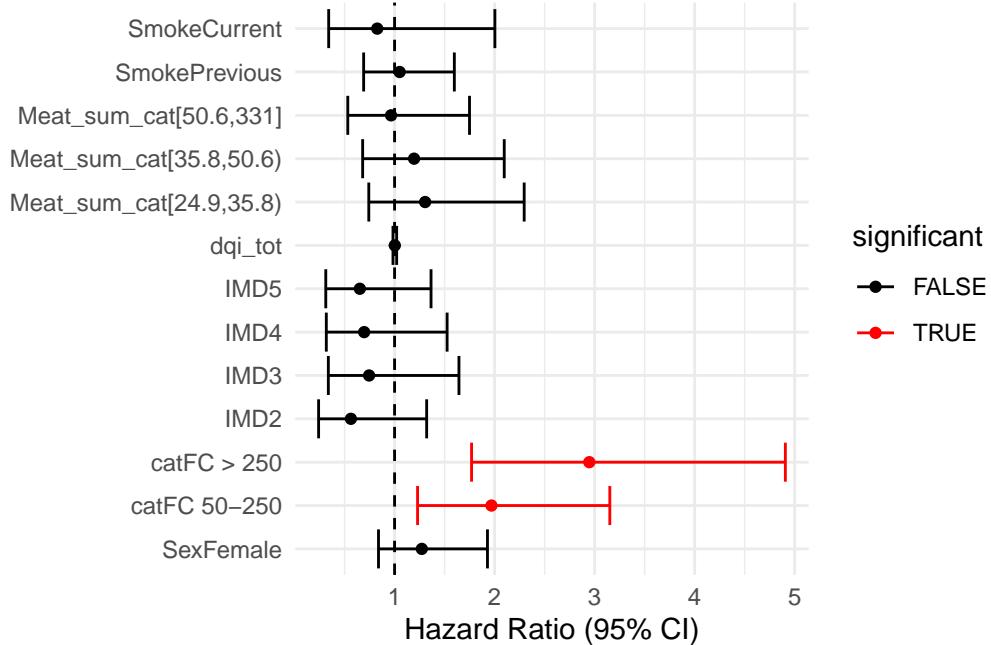
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_meat_cd_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_meat_cd_hard <- mice::mice(
  data = data_impute_meat_cd_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

# Fit pooled Cox model
with(
  mice_meat_cd_hard,
  coxph(
    Surv(hardflare_time, hardflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      Meat_sum_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()

```



## Ulcerative colitis

### Patient reported flare

```
# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others
# Calculate cumulative hazard

# Soft flare
data_impute_meat_uc_soft <- flare.uc.df %>%
  dplyr::select(softflare_time,
                softflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                Meat_sum_cat,
                Smoke,
                SiteNo) %>%
  
  # Remove missing on all columns except smoke
  dplyr::filter(!dplyr::if_any(
    .cols = -Smoke,
```

```

.fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = softflare_time,
    statusvar = softflare
  ))
}

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_meat_uc_soft)

pred_matrix[, 'softflare_time'] <- 0

# MICE with 10 imputations
mice_meat_uc_soft <- mice::mice(
  data = data_impute_meat_uc_soft,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)
}

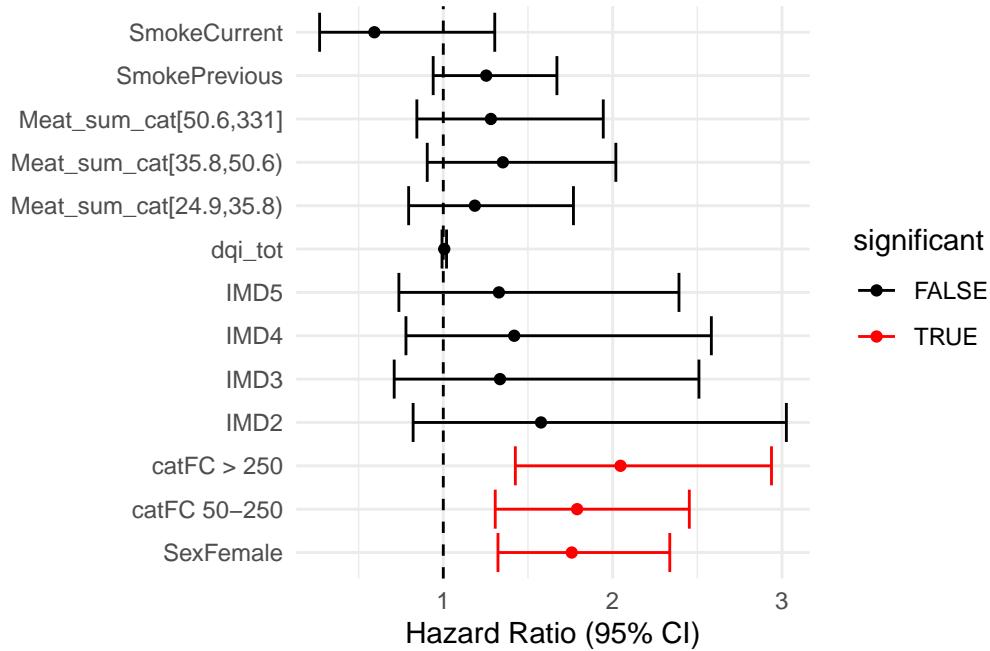
# Fit pooled Cox model
with(
  mice_meat_uc_soft,
  coxph(
    Surv(softflare_time, softflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      Meat_sum_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,

```

```

exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Objective flare

```

# Hard flare
data_impute_meat_uc_hard <- flare.uc.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                Meat_sum_cat,
                Smoke,
                SiteNo) %>%
  # Remove missing on all columns except smoke
  dplyr::filter(!dplyr::if_any(
    .cols = -Smoke,
    .fns = is.na
  )) %>%

```

```

# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data =.,
    timevar = hardflare_time,
    statusvar = hardflare
  ))
}

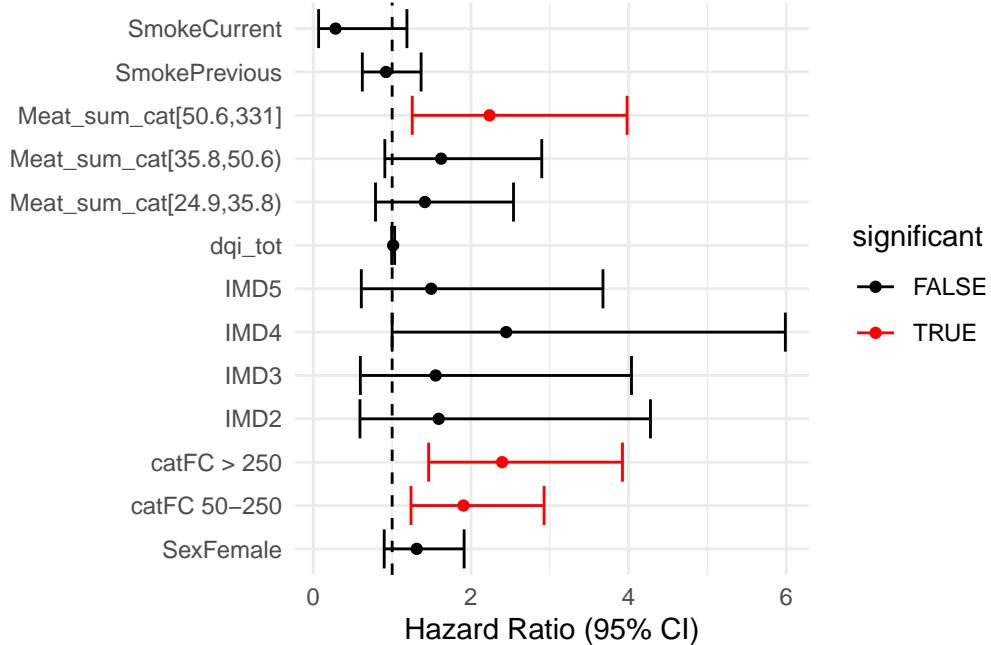
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_meat_uc_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_meat_uc_hard <- mice::mice(
  data = data_impute_meat_uc_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

# Fit pooled Cox model
with(
  mice_meat_uc_hard,
  coxph(
    Surv(hardflare_time, hardflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      Meat_sum_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()

```



## Dietary Fibre

### Crohn's disease

#### Patient reported flare

```
# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others
# Calculate cumulative hazard

# Soft flare
data_impute_fibre_cd_soft <- flare.cd.df %>%
  dplyr::select(softflare_time,
                softflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                fibre_cat,
                Smoke,
                SiteNo) %>%
  # Remove missing on all columns except smoke
```

```

dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = softflare_time,
    statusvar = softflare
  )
)

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_fibre_cd_soft)

pred_matrix[, 'softflare_time'] <- 0

# MICE with 10 imputations
mice_fibre_cd_soft <- mice::mice(
  data = data_impute_fibre_cd_soft,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

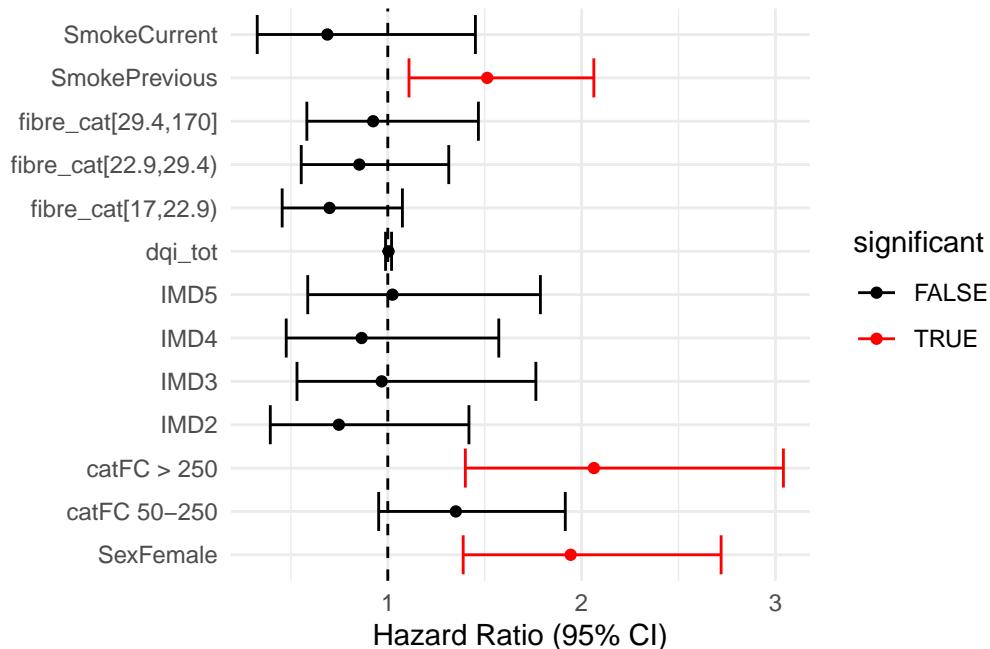
# Fit pooled Cox model
with(
  mice_fibre_cd_soft,
  coxph(
    Surv(softflare_time, softflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      fibre_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%

```

```

summary(conf.int = TRUE,
       conf.level = 0.95,
       exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Objective flare

```

# Hard flare
data_impute_fibre_cd_hard <- flare.cd.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                fibre_cat,
                Smoke,
                SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,

```

```

.fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = hardflare_time,
    statusvar = hardflare
  ))
}

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_fibre_cd_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_fibre_cd_hard <- mice::mice(
  data = data_impute_fibre_cd_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

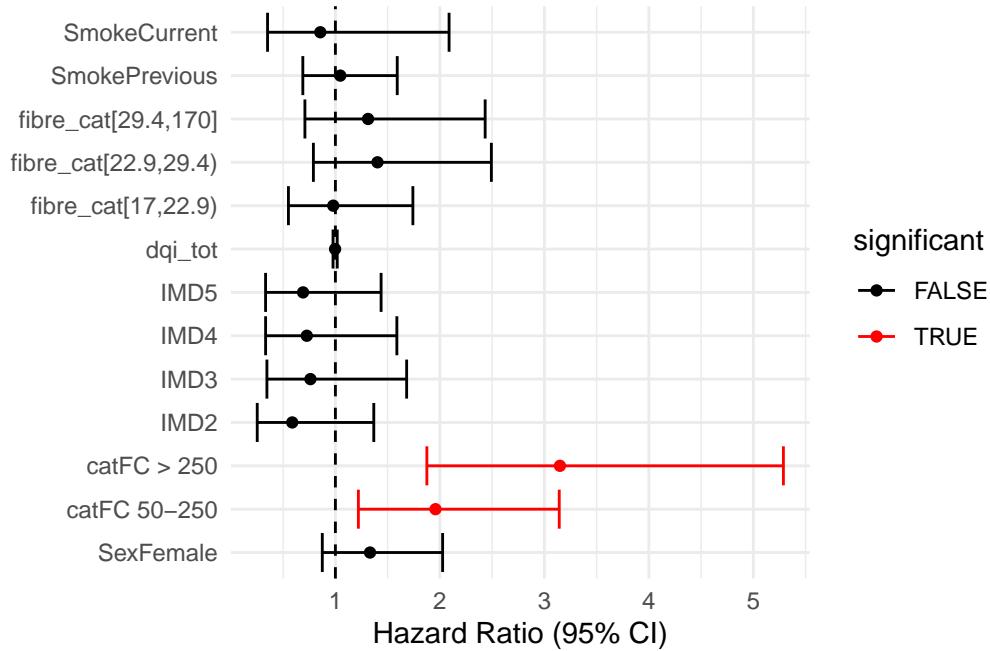
# Fit pooled Cox model
with(
  mice_fibre_cd_hard,
  coxph(
    Surv(hardflare_time, hardflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      fibre_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,

```

```

exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Ulcerative colitis

### Patient reported flare

```

# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others
# Calculate cumulative hazard

# Soft flare
data_impute_fibre_uc_soft <- flare.uc.df %>%
  dplyr::select(softflare_time,
                softflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                fibre_cat,
                Smoke,

```

```

    SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = softflare_time,
    statusvar = softflare
  )
)

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_fibre_uc_soft)

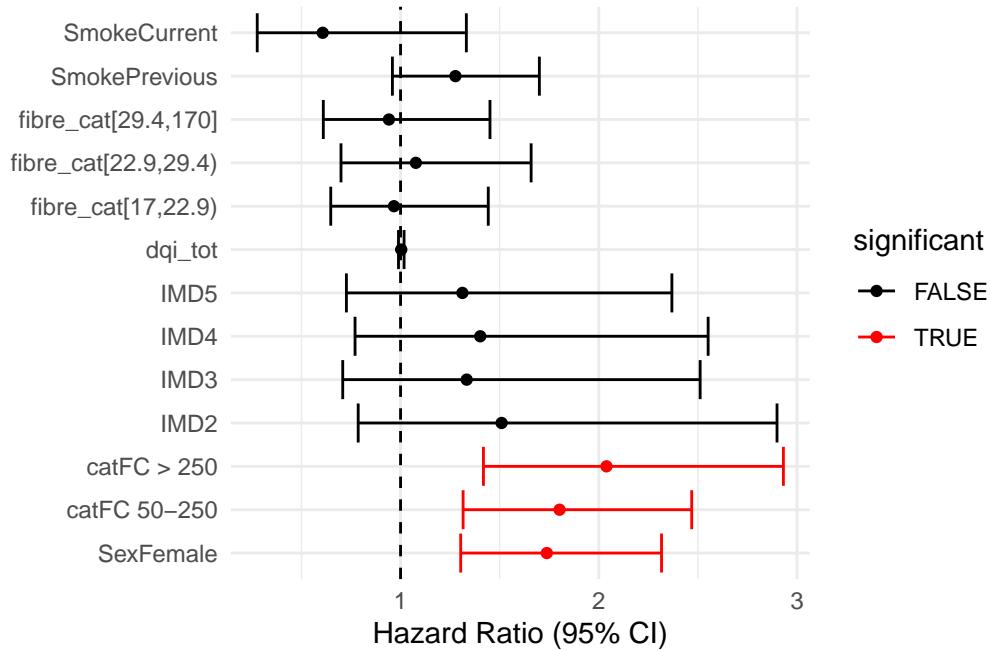
pred_matrix[, 'softflare_time'] <- 0

# MICE with 10 imputations
mice_fibre_uc_soft <- mice::mice(
  data = data_impute_fibre_uc_soft,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

# Fit pooled Cox model
with(
  mice_fibre_uc_soft,
  coxph(
    Surv(softflare_time, softflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      fibre_cat +
      Smoke +
      frailty(SiteNo)
  )
)

```

```
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()
```



## Objective flare

```
# Hard flare
data_impute_fibre_uc_hard <- flare.uc.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                fibre_cat,
                Smoke,
                SiteNo) %>%
# Remove missing on all columns except smoke
```

```

dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = hardflare_time,
    statusvar = hardflare
  )
)

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_fibre_uc_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_fibre_uc_hard <- mice::mice(
  data = data_impute_fibre_uc_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

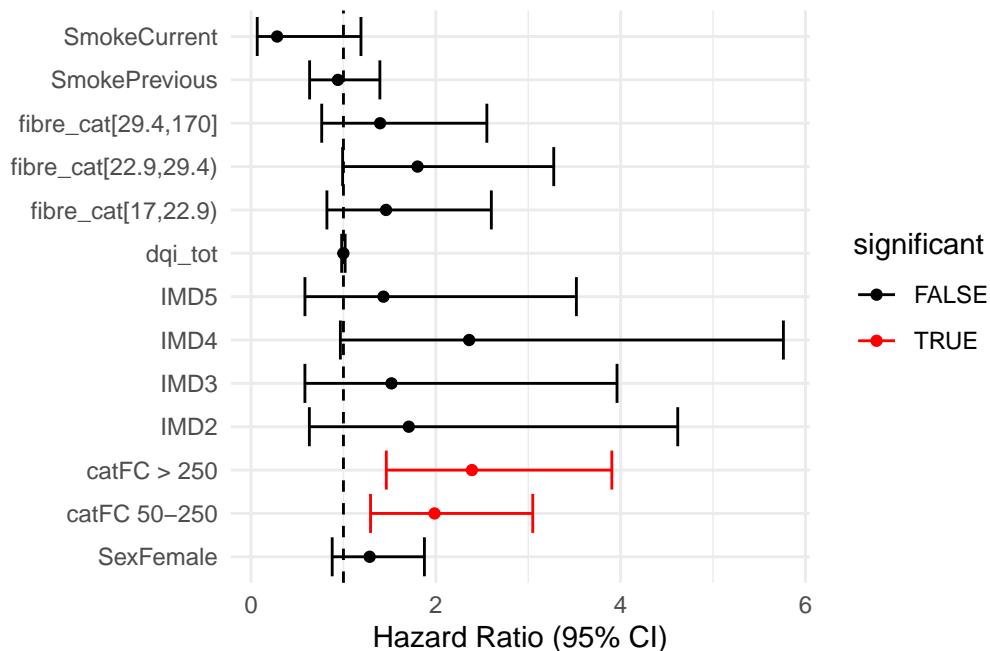
# Fit pooled Cox model
with(
  mice_fibre_uc_hard,
  coxph(
    Surv(hardflare_time, hardflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      fibre_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%

```

```

summary(conf.int = TRUE,
        conf.level = 0.95,
        exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Polyunsaturated fatty acids

### Crohn's disease

#### Patient reported flare

```

# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others
# Calculate cumulative hazard

# Soft flare
data_impute_pufa_cd_soft <- flare.cd.df %>%
  dplyr::select(softflare_time,
                softflare,
                Sex,
                cat,

```

```

    IMD,
    dqi_tot,
    PUFA_percEng_cat,
    Smoke,
    SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = softflare_time,
    statusvar = softflare
  ))
}

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_pufa_cd_soft)

pred_matrix[, 'softflare_time'] <- 0

# MICE with 10 imputations
mice_pufa_cd_soft <- mice::mice(
  data = data_impute_pufa_cd_soft,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

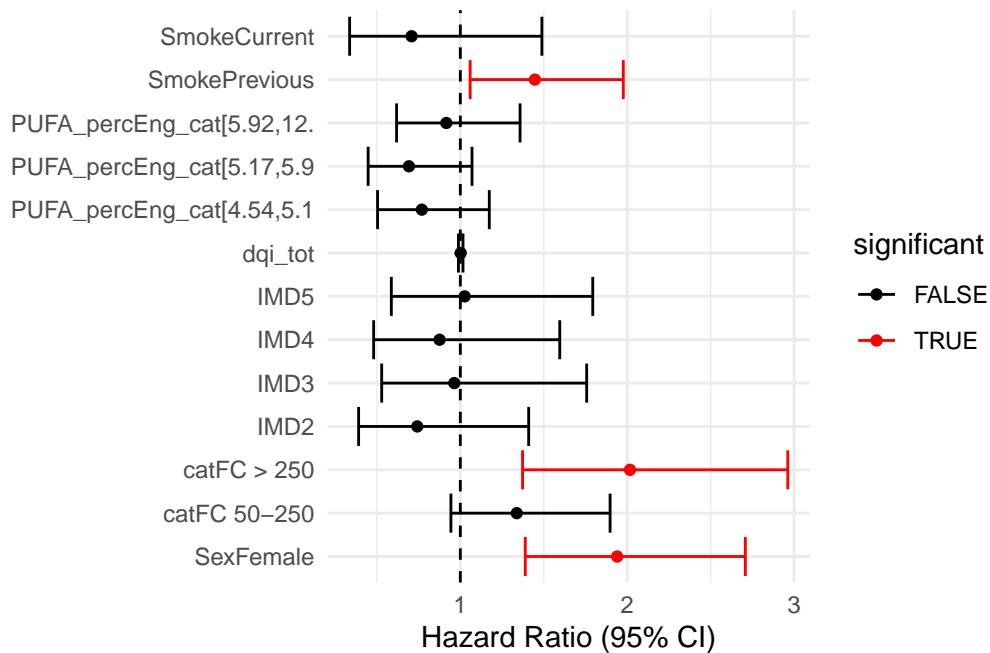
# Fit pooled Cox model
with(
  mice_pufa_cd_soft,
  coxph(
    Surv(softflare_time, softflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +

```

```

PUFA_percEng_cat +
Smoke +
frailty(SiteNo)
)
) %>%
mice::pool() %>%
summary(conf.int = TRUE,
conf.level = 0.95,
exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Objective flare

```

# Hard flare
data_impute_pufa_cd_hard <- flare.cd.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,

```

```

    PUFA_percEng_cat,
    Smoke,
    SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonaalen(
    data = .,
    timevar = hardflare_time,
    statusvar = hardflare
  ))
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_pufa_cd_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_pufa_cd_hard <- mice::mice(
  data = data_impute_pufa_cd_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

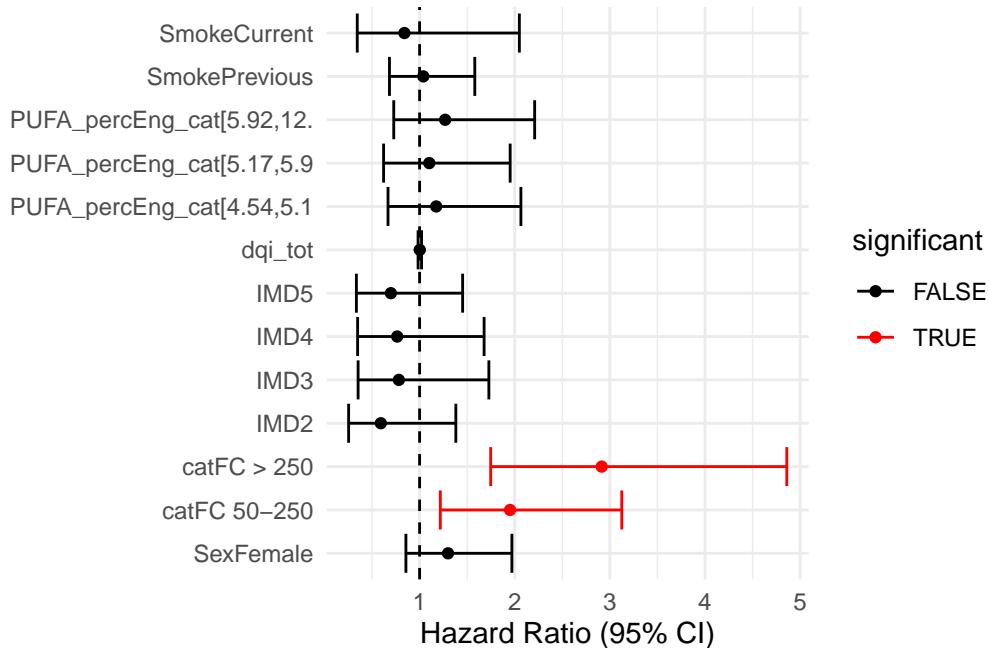
# Fit pooled Cox model
with(
  mice_pufa_cd_hard,
  coxph(
    Surv(hardflare_time, hardflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      PUFA_percEng_cat +
      Smoke +

```

```

    frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()

```



## Ulcerative colitis

### Patient reported flare

```

# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others
# Calculate cumulative hazard

# Soft flare
data_impute_pufa_uc_soft <- flare.uc.df %>%
  dplyr::select(softflare_time,
                softflare,

```

```

    Sex,
    cat,
    IMD,
    dqi_tot,
    PUFA_percEng_cat,
    Smoke,
    SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = softflare_time,
    statusvar = softflare
))
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_pufa_uc_soft)

pred_matrix[, 'softflare_time'] <- 0

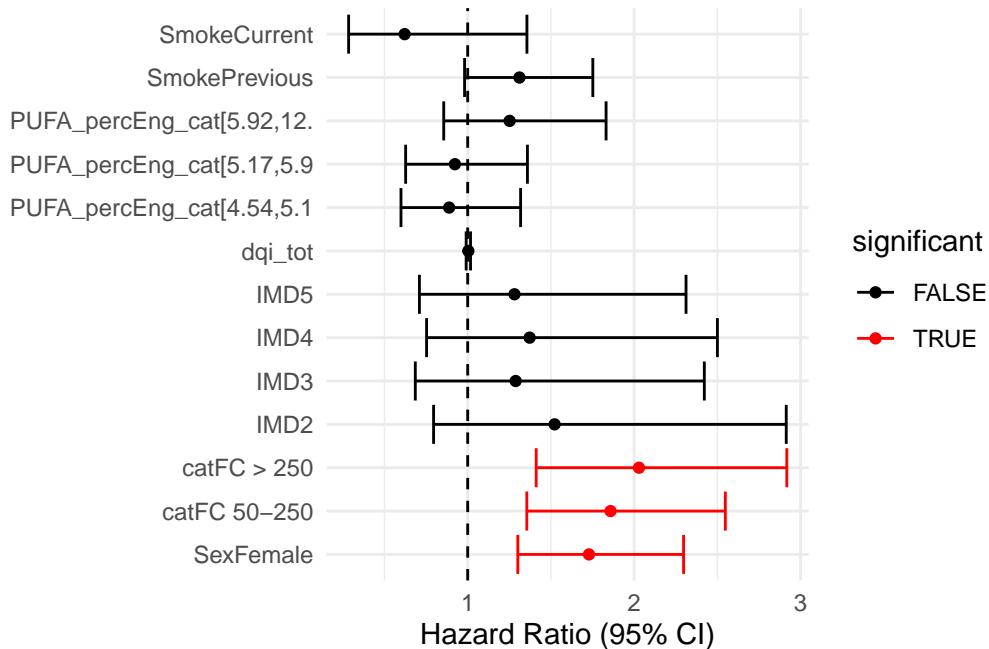
# MICE with 10 imputations
mice_pufa_uc_soft <- mice::mice(
  data = data_impute_pufa_uc_soft,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)
# Fit pooled Cox model
with(
  mice_pufa_uc_soft,
  coxph(
    Surv(softflare_time, softflare) ~
      Sex +
      cat +

```

```

IMD +
dqi_tot +
PUFA_percEng_cat +
Smoke +
frailty(SiteNo)
)
) %>%
mice::pool() %>%
summary(conf.int = TRUE,
conf.level = 0.95,
exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Objective flare

```

# Hard flare
data_impute_pufa_uc_hard <- flare.uc.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,

```

```

    IMD,
    dqi_tot,
    PUFA_percEng_cat,
    Smoke,
    SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = hardflare_time,
    statusvar = hardflare
  ))
}

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_pufa_uc_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_pufa_uc_hard <- mice::mice(
  data = data_impute_pufa_uc_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

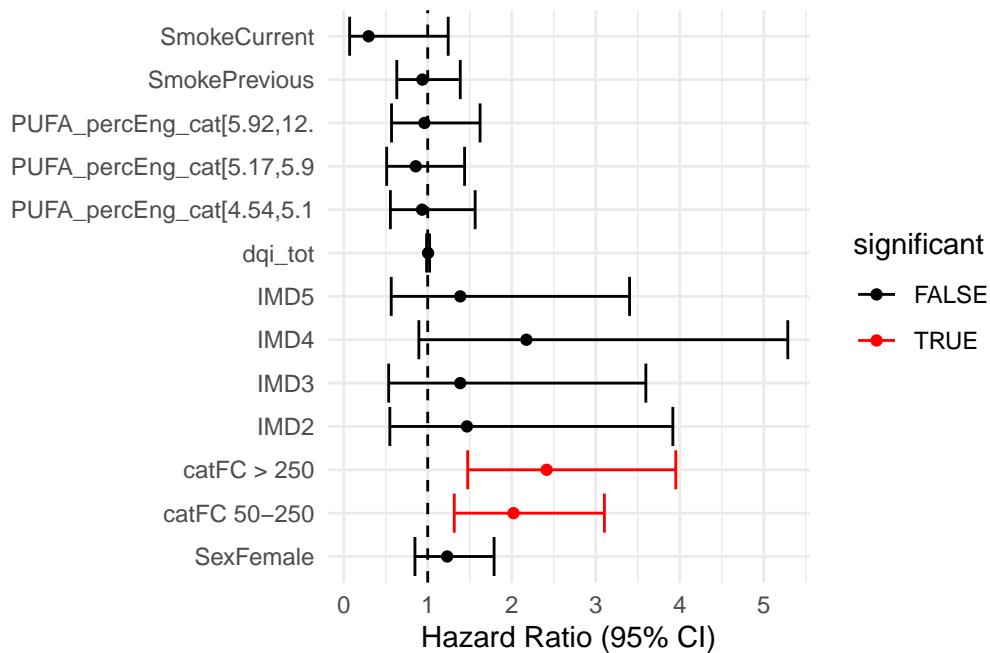
# Fit pooled Cox model
with(
  mice_pufa_uc_hard,
  coxph(
    Surv(hardflare_time, hardflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +

```

```

    PUFA_percEng_cat +
    Smoke +
    frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Ultra-processed food

### Crohn's disease

#### Patient reported flare

```

# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others
# Calculate cumulative hazard

```

```

# Soft flare
data_impute_upf_cd_soft <- flare.cd.df %>%
  dplyr::select(softflare_time,
                softflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                UPF_perc_cat,
                Smoke,
                SiteNo) %>%
  
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data =.,
    timevar = softflare_time,
    statusvar = softflare
  ))
  
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_upf_cd_soft)

pred_matrix[, 'softflare_time'] <- 0

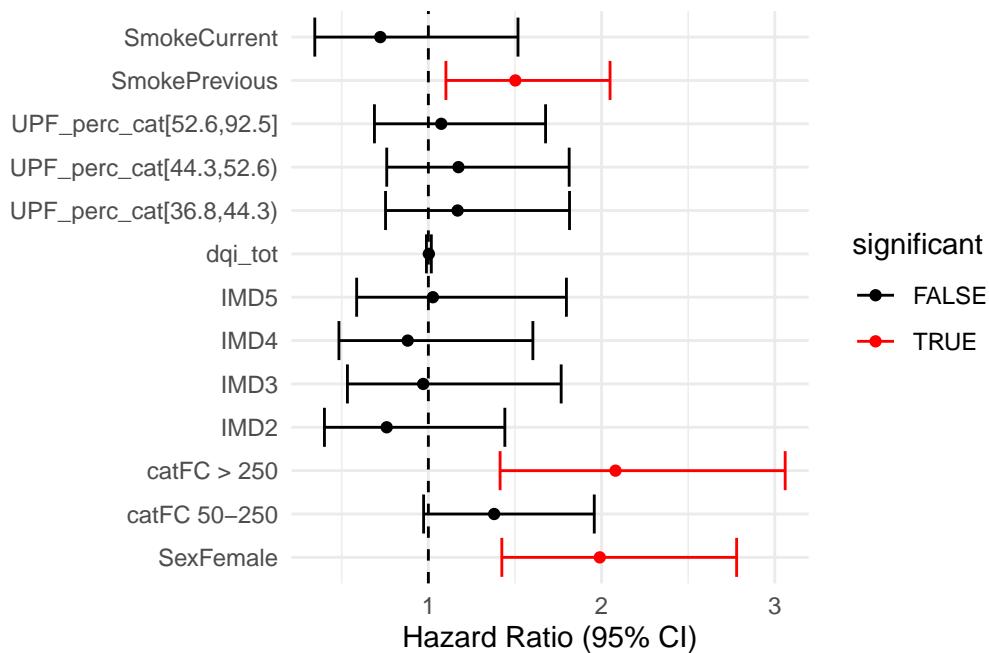
# MICE with 10 imputations
mice_upf_cd_soft <- mice::mice(
  data = data_impute_upf_cd_soft,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)
  
# Fit pooled Cox model
with(
  mice_upf_cd_soft,

```

```

coxph(
  Surv(softflare_time, softflare) ~
    Sex +
    cat +
    IMD +
    dqi_tot +
    UPF_perc_cat +
    Smoke +
    frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()

```



## Objective flare

```

# Hard flare
data_impute_upf_cd_hard <- flare.cd.df %>%

```

```

dplyr::select(hardflare_time,
              hardflare,
              Sex,
              cat,
              IMD,
              dqi_tot,
              UPF_perc_cat,
              Smoke,
              SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonaalen(
    data =.,
    timevar = hardflare_time,
    statusvar = hardflare
))
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_upf_cd_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_upf_cd_hard <- mice::mice(
  data = data_impute_upf_cd_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

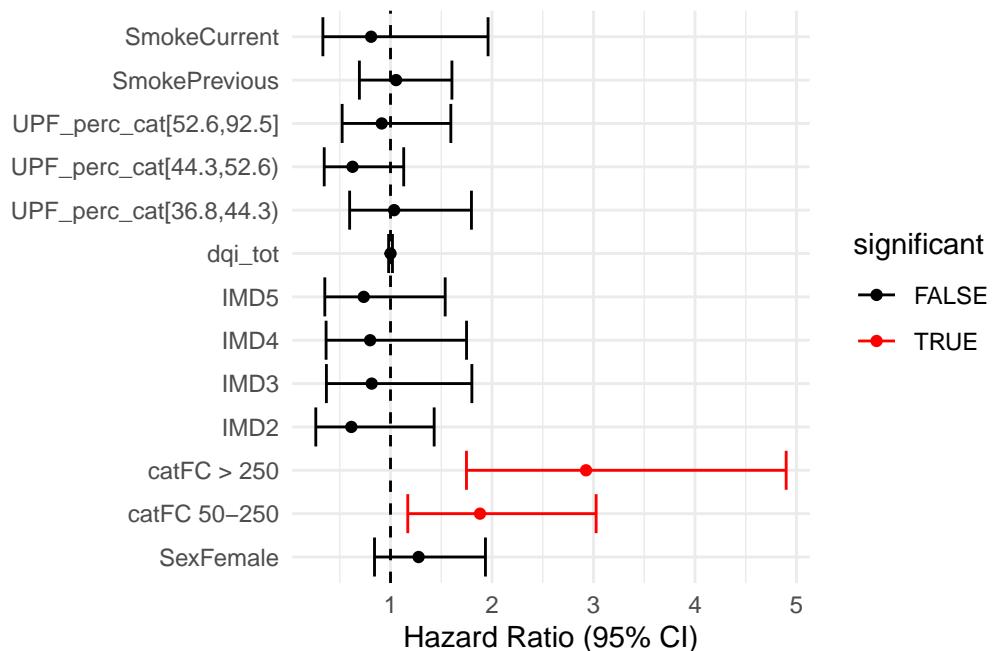
# Fit pooled Cox model
with(
  mice_upf_cd_hard,
  coxph(
    Surv(hardflare_time, hardflare) ~

```

```

Sex +
cat +
IMD +
dqi_tot +
UPF_perc_cat +
Smoke +
frailty(SiteNo)
)
) %>%
mice::pool() %>%
summary(conf.int = TRUE,
conf.level = 0.95,
exponentiate = TRUE) %>%
summon_plot_broom_hr()

```



## Ulcerative colitis

### Patient reported flare

```

# Only select relevant columns for the imputation model
# Only imputing smoking so remove missing others

```

```

# Calculate cumulative hazard

# Soft flare
data_impute_upf_uc_soft <- flare.uc.df %>%
  dplyr::select(softflare_time,
                softflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                UPF_perc_cat,
                Smoke,
                SiteNo) %>%
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data = .,
    timevar = softflare_time,
    statusvar = softflare
))

# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_upf_uc_soft)

pred_matrix[, 'softflare_time'] <- 0

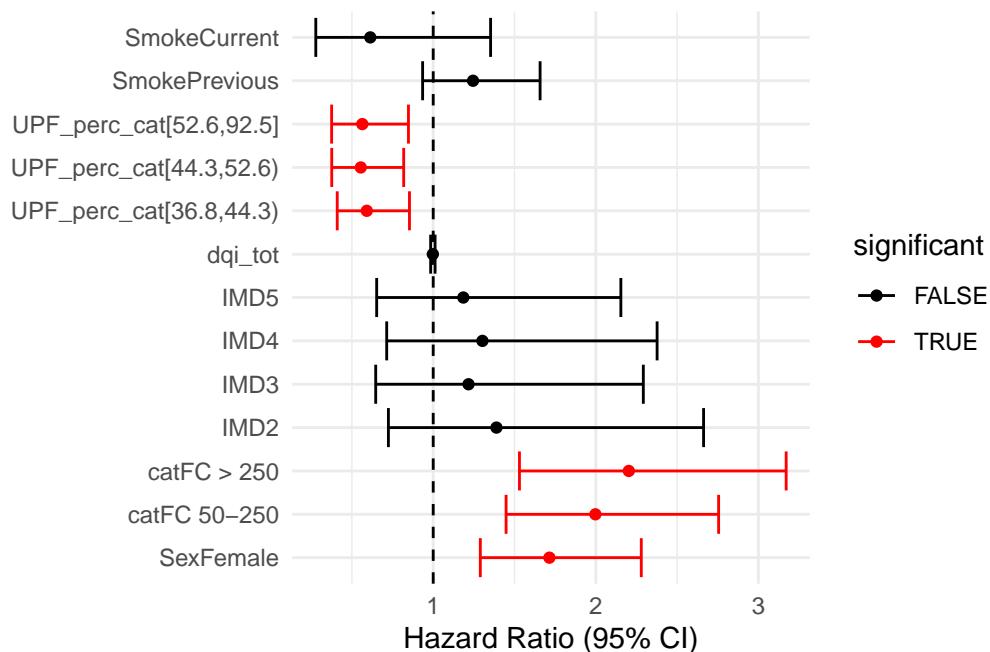
# MICE with 10 imputations
mice_upf_uc_soft <- mice::mice(
  data = data_impute_upf_uc_soft,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)

```

```

# Fit pooled Cox model
with(
  mice_upf_uc_soft,
  coxph(
    Surv(softflare_time, softflare) ~
      Sex +
      cat +
      IMD +
      dqi_tot +
      UPF_perc_cat +
      Smoke +
      frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()

```



## Objective flare

```

# Hard flare
data_impute_upf_uc_hard <- flare.uc.df %>%
  dplyr::select(hardflare_time,
                hardflare,
                Sex,
                cat,
                IMD,
                dqi_tot,
                UPF_perc_cat,
                Smoke,
                SiteNo) %>%
  
# Remove missing on all columns except smoke
dplyr::filter(!dplyr::if_any(
  .cols = -Smoke,
  .fns = is.na
)) %>%
# Calculate Cumulative hazard
dplyr::mutate(
  cumhaz = mice::nelsonalen(
    data =.,
    timevar = hardflare_time,
    statusvar = hardflare
  ))
  
# Predictor matrix - need to exclude time from the model
pred_matrix <- mice::make.predictorMatrix(data_impute_upf_uc_hard)

pred_matrix[, 'hardflare_time'] <- 0

# MICE with 10 imputations
mice_upf_uc_hard <- mice::mice(
  data = data_impute_upf_uc_hard,
  predictorMatrix = pred_matrix,
  m = 10,
  maxit = 20,
  seed = 73,
  print = FALSE
)
  
# Fit pooled Cox model
with(
  mice_upf_uc_hard,

```

```

coxph(
  Surv(hardflare_time, hardflare) ~
    Sex +
    cat +
    IMD +
    dqi_tot +
    UPF_perc_cat +
    Smoke +
    frailty(SiteNo)
  )
) %>%
  mice::pool() %>%
  summary(conf.int = TRUE,
         conf.level = 0.95,
         exponentiate = TRUE) %>%
  summon_plot_broom_hr()

```

