

# Supplement 2

## Sites of pain

Peter Kamerman

Last knitted: 01 July 2020

## Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Import and check data</b>  | <b>2</b>  |
| <b>2</b> | <b>Basic descriptive statistics</b>   | <b>3</b>  |
| 2.1      | Pain sites . . . . .  | 3         |
| 2.2      | Demographics . . . . .  | 4         |
| 2.3      | Add acute/chronic pain column to demographic data . . . . .                         | 5         |
| 2.4      | Bootstrap functions . . . . .   | 5         |
| <b>3</b> | <b>Proportion point estimates with 95% CIs</b>                                      | <b>5</b>  |
| 3.1      | Process data . . . . .  | 5         |
| 3.2      | Tabulated proportions (with 95% CIs), by body region . . . . .                      | 8         |
| 3.3      | Plotted proportions (with 95% CIs), by body region . . . . .                        | 9         |
| <b>4</b> | <b>By sex</b>   | <b>11</b> |
| 4.1      | Process data . . . . .  | 11        |
| 4.2      | Tabulated proportions (with 95% CIs), by age and body region . . . . .              | 13        |
| 4.3      | Plotted proportions (with 95% CIs), by age and body region . . . . .                | 14        |
| <b>5</b> | <b>By age</b>   | <b>19</b> |
| 5.1      | Process data . . . . .  | 19        |
| 5.2      | Tabulated proportions (with 95% CIs), by age group and body site . . . . .          | 22        |
| 5.3      | Plotted proportions (with 95% CIs), by age group and body site . . . . .            | 26        |
| <b>6</b> | <b>By most recent CD4 T-cell count</b>  | <b>44</b> |
| 6.1      | Process data . . . . .  | 44        |
| 6.2      | Tabulated proportions (with 95% CIs), by CD4 (recent) group and body site . . . . . | 47        |
| 6.3      | Plotted proportions (with 95% CIs), by CD4 (recent) group and body site . . . . .   | 51        |
| <b>7</b> | <b>By pain definition</b>   | <b>69</b> |
| 7.1      | Process data . . . . .  | 69        |
| 7.2      | Tabulated proportions (with 95% CIs), by age and body region . . . . .              | 72        |
| 7.3      | Plotted proportions (with 95% CIs), by age and body region . . . . .                | 73        |
| <b>8</b> | <b>Logistic regression</b>  | <b>78</b> |
| 8.1      | Process data . . . . .  | 78        |
| 8.2      | Family-wise error correction . . . . .  | 79        |
| 8.3      | Run <code>glmer</code> model for every body site to check for issues . . . . .      | 79        |
| 8.4      | Run <code>glm</code> for selected body sites to check for issues . . . . .          | 93        |
| 8.5      | GLMM data for tabulation and plotting . . . . .                                     | 101       |

|           |   |            |
|-----------|---|------------|
| 8.6       | GLM data for tabulation and plotting                  | 103        |
| 8.7       | Tabulated results (fixed effects)                     | 105        |
| 8.7.1     | GLMM tabulations                                      | 105        |
| 8.7.2     | GLM tabulations                                       | 107        |
| 8.8       | Plotted results (fixed effects)                       | 109        |
| <b>9</b>  | <b>Summary plots</b>                                  | <b>112</b> |
| 9.1       | Body sites with a point estimate pain proportion >10% | 112        |
| 9.2       | Body sites with a point estimate pain proportion <10% | 119        |
| <b>10</b> | <b>Session information</b>                            | <b>130</b> |

---

## 1 Import and check data

```
# Import
data <- read_rds('data-cleaned/data-pain-sites.rds')
demo <- read_rds('data-cleaned/data-demographics.rds')

# Check
## Pain sites
dim(data)

## [1] 596 21
names(data)

## [1] "ID" "Head" "Throat"
## [4] "Shoulder" "Arms" "Elbows"
## [7] "Wrists.Hands" "Chest" "Upper_back"
## [10] "Lower_back" "Abdomen" "Cervical_spine"
## [13] "Thoracic_spine" "Lumbosacral_spine" "Groin"
## [16] "Hips" "Legs" "Knees"
## [19] "Ankles.Feet" "Buttocks" "Site"

glimpse(data)

## Rows: 596
## Columns: 21
## $ ID <chr> "RPB73", "RPB74", "RPB75", "RPB76", "RPB77", "RPB...
## $ Head <chr> "No", "No", "No", "Yes", "Yes", "No", "No", "No", ...
## $ Throat <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Shoulder <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Arms <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Elbows <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Wrists.Hands <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Chest <chr> "No", "No", "No", "Yes", "No", "No", "No", "No", "..."
## $ Upper_back <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Lower_back <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Abdomen <chr> "No", "No", "Yes", "Yes", "No", "No", "Yes", "No", "..."
## $ Cervical_spine <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Thoracic_spine <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
## $ Lumbosacral_spine <chr> "No", "No", "No", "No", "No", "No", "No", "No", "..."
```

```
## $ Groin      <chr> "No", "No", "No", "No", "No", "No", "No", "No", "...
## $ Hips      <chr> "No", "No", "No", "No", "No", "No", "No", "Yes", ...
## $ Legs      <chr> "No", "No", "No", "No", "No", "Yes", "No", "Yes",...
## $ Knees     <chr> "No", "No", "No", "No", "No", "Yes", "No", "Yes",...
## $ Ankles.Feet <chr> "No", "No", "No", "No", "No", "Yes", "No", "Yes",...
## $ Buttocks  <chr> "No", "No", "No", "No", "No", "No", "No", "Yes", ...
## $ Site      <chr> "RP", "RP", "RP", "RP", "RP", "RP", "RP", "RP", "...

## Demographics
dim(demo)

## [1] 596      8

names(demo)

## [1] "ID"           "Site"           "Sex"
## [4] "Age"          "Employment_status" "CD4_recent"
## [7] "ART_currently" "Education"

glimpse(demo)

## Rows: 596
## Columns: 8
## $ ID      <chr> "RPB73", "RPB74", "RPB75", "RPB76", "RPB77", "RPB...
## $ Site    <chr> "RP", "RP", "RP", "RP", "RP", "RP", "RP", "RP", "...
## $ Sex     <chr> "Female", "Female", "Female", "Female", "Female",...
## $ Age     <dbl> 36, 27, 39, 36, 31, 32, 28, 37, 31, 25, 31, 24, 3...
## $ Employment_status <chr> "Other", "Unemployed", "Other", "Unemployed", "Un...
## $ CD4_recent <dbl> 391, 571, 591, 207, 126, 225, 543, 410, 74, 212, ...
## $ ART_currently <chr> "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", ...
## $ Education <chr> "Tertiary", "Secondary", "Secondary", "Primary", ...
```

## 2 Basic descriptive statistics

### 2.1 Pain sites

```
data %>%
  select(-ID, -Site) %>%
  mutate_if(is.character, factor) %>%
  skim()
```

Table 1: Data summary

| Name                   | Piped data |
|------------------------|------------|
| Number of rows         | 596        |
| Number of columns      | 19         |
| Column type frequency: |            |
| factor                 | 19         |
| Group variables        | None       |

### Variable type: factor

| skim_variable     | n_missing | complete_rate | n_unique | top_counts        |
|-------------------|-----------|---------------|----------|-------------------|
| Head              | 0         | 1             | 2        | No: 401, Yes: 195 |
| Throat            | 0         | 1             | 2        | No: 577, Yes: 19  |
| Shoulder          | 0         | 1             | 2        | No: 549, Yes: 47  |
| Arms              | 0         | 1             | 2        | No: 571, Yes: 25  |
| Elbows            | 0         | 1             | 2        | No: 574, Yes: 22  |
| Wrists.Hands      | 0         | 1             | 2        | No: 561, Yes: 35  |
| Chest             | 0         | 1             | 2        | No: 479, Yes: 117 |
| Upper_back        | 0         | 1             | 1        | No: 596           |
| Lower_back        | 0         | 1             | 2        | No: 554, Yes: 42  |
| Abdomen           | 0         | 1             | 2        | No: 437, Yes: 159 |
| Cervical_spine    | 0         | 1             | 2        | No: 566, Yes: 30  |
| Thoracic_spine    | 0         | 1             | 2        | No: 522, Yes: 74  |
| Lumbosacral_spine | 0         | 1             | 2        | No: 504, Yes: 92  |
| Groin             | 0         | 1             | 2        | No: 542, Yes: 54  |
| Hips              | 0         | 1             | 2        | No: 556, Yes: 40  |
| Legs              | 0         | 1             | 2        | No: 497, Yes: 99  |
| Knees             | 0         | 1             | 2        | No: 512, Yes: 84  |
| Ankles.Feet       | 0         | 1             | 2        | No: 412, Yes: 184 |
| Buttocks          | 0         | 1             | 2        | No: 577, Yes: 19  |

## 2.2 Demographics

```
demo %>%  
  select(-ID, -Site) %>%  
  mutate_if(is.character, factor) %>%  
  skim()
```

Table 3: Data summary

|                        |            |
|------------------------|------------|
| Name                   | Piped data |
| Number of rows         | 596        |
| Number of columns      | 6          |
| Column type frequency: |            |
| factor                 | 4          |
| numeric                | 2          |
| Group variables        | None       |

### Variable type: factor

| skim_variable     | n_missing | complete_rate | n_unique | top_counts                           |
|-------------------|-----------|---------------|----------|--------------------------------------|
| Sex               | 0         | 1.00          | 2        | Fem: 481, Mal: 115                   |
| Employment_status | 49        | 0.92          | 4        | Une: 330, Ful: 131, Par: 52, Oth: 34 |
| ART_currently     | 5         | 0.99          | 2        | Yes: 460, No: 131                    |
| Education         | 37        | 0.94          | 3        | Sec: 395, Pri: 99, Ter: 65           |

## Variable type: numeric

| skim_variable | n_missing | complete_rate | mean   | sd     | p0 | p25 | p50 | p75 | p100 |
|---------------|-----------|---------------|--------|--------|----|-----|-----|-----|------|
| Age           | 8         | 0.99          | 37.28  | 9.06   | 19 | 31  | 36  | 42  | 76   |
| CD4_recent    | 99        | 0.83          | 320.71 | 238.92 | 1  | 155 | 261 | 432 | 1232 |

## 2.3 Add acute/chronic pain column to demographic data

```
# Based on pain definition used in the study
## Acute assessing pain in the last 1 week or 1 month.
## Chronic assessing pain in the last 3 months
demo <- demo %>%
  mutate(Pain = case_when(
    Site == 'RP' |
      Site == 'NM' ~ 'Acute',
    TRUE ~ 'Chronic'
  ))
```

## 2.4 Bootstrap functions

```
# Proportion
prop_func <- function(d, i){
  dat <- d[i, ]
  dat_vec <- dat[[1]]
  dat_prop <- mean(dat_vec == 'Yes', na.rm = TRUE)
  dat_prop
}

# Median
median_func <- function(d, i){
  dat <- d[i, ]
  dat_vec <- dat[[1]]
  dat_median <- median(dat_vec, na.rm = TRUE)
  dat_median
}
```

---

## 3 Proportion point estimates with 95% CIs

### 3.1 Process data

```
# Set seed
set.seed(2020)

# Remove ID and upper_back (only one outcome -- no pain) columns
prop <- data[, !(names(data) %in% c('ID', 'Upper_back'))]
```

```

# Bootstrap CIs
prop_boot <- prop %>%
  # Remove Site column
  select(-Site) %>%
  # Pivot to long format
  pivot_longer(cols = everything(),
               names_to = 'body_site',
               values_to = 'pain_present') %>%
  # Add body regions
  mutate(region = case_when(
    body_site == 'Chest' |
    body_site == 'Head' |
    body_site == 'Throat' |
    body_site == 'Shoulder' ~ 'Head and upper torso',
    body_site == 'Lower_back' |
    body_site == 'Abdomen' |
    body_site == 'Hips' |
    body_site == 'Buttocks' |
    body_site == 'Groin' ~ 'Lower torso',
    body_site == 'Legs' |
    body_site == 'Knees' |
    body_site == 'Ankles.Feet' ~ 'Lower limbs',
    body_site == 'Arms' |
    body_site == 'Elbows' |
    body_site == 'Wrists.Hands' ~ 'Upper limbs',
    body_site == 'Cervical_spine' |
    body_site == 'Thoracic_spine' |
    body_site == 'Lumbosacral_spine' ~ 'Spinal column',
    TRUE ~ 'other'
  )) %>%
  # Nest by body region and body site
  group_by(region, body_site) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,
                    ~ boot(data = .x,
                           statistic = prop_func,
                           R = 999,
                           stype = 'i',
                           parallel = 'multicore',
                           ncpus = 4))) %>%
  # Get CI
  mutate(ci = map(.x = boot,
                  ~ boot.ci(.x, type = 'perc')))) %>%
  # Extract ci data
  mutate(point_est = map(.x = ci,
                        ~ .x$t0),
         lower_ci = map(.x = ci,
                        ~ .x$percent[[4]]),
         upper_ci = map(.x = ci,
                        ~ .x$percent[[5]])) %>%
  # Remove columns
  select(-data, -boot, -ci) %>%

```

```

# Unnest
unnest(cols = c(point_est, lower_ci, upper_ci))

# Re-nest by body region and generate figures and tables
prop_boot2 <- prop_boot %>%
  group_by(region) %>%
  nest() %>%
  # Fix site labels
  mutate(data = map(.x = data,
                    ~ .x %>%
                      mutate(body_site = str_replace_all(body_site,
                                                            pattern = '_',
                                                            replacement = ' '),
                             body_site = str_replace_all(body_site,
                                                            pattern = '\\.',
                                                            replacement = ' & '),
                             body_site = str_replace_all(body_site,
                                                            pattern = 'Lower back',
                                                            replacement = 'Flank')))) %>%

# Re-order sites by point_est
mutate(data = map(.x = data,
                  ~ .x %>%
                    mutate(body_site = fct_reorder(body_site,
                                                    point_est)))) %>%

# Plot data
mutate(plots = map2(.x = data,
                   .y = region,
                   ~ .x %>%
                     ggplot(data = .) +
                     aes(x = body_site,
                         y = point_est,
                         ymin = lower_ci,
                         ymax = upper_ci) +
                     geom_hline(yintercept = 0.1,
                                linetype = 2) +
                     geom_pointrange(size = 1) +
                     coord_flip() +
                     labs(title = .y,
                          subtitle = '(Point estimate with 95%CI)',
                          y = 'Proportion with pain') +
                     scale_y_continuous(limits = c(0, 0.4)) +
                     theme_minimal(base_size = 18) +
                     theme(plot.title = element_text(size = 18),
                           plot.subtitle = element_text(size = 12),
                           axis.title.y = element_blank(),
                           panel.grid = element_blank(),
                           axis.text = element_text(colour = '#000000'),
                           axis.line = element_line(size = 0.5),
                           axis.ticks = element_line(size = 0.5)))) %>%

# Tabulate data
mutate(tables = map2(.x = data,
                    .y = region,
                    ~ .x %>%

```

```
kable(caption = .y,
      digits = 2))
```

### 3.2 Tabulated proportions (with 95% CIs), by body region

```
walk(prop_boot2$tables, ~ print(.x))
```

Table 6: Head and upper torso

| body_site | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| Head      | 0.33      | 0.29     | 0.37     |
| Throat    | 0.03      | 0.02     | 0.05     |
| Shoulder  | 0.08      | 0.06     | 0.10     |
| Chest     | 0.20      | 0.17     | 0.23     |

Table 7: Upper limbs

| body_site      | point_est | lower_ci | upper_ci |
|----------------|-----------|----------|----------|
| Arms           | 0.04      | 0.03     | 0.06     |
| Elbows         | 0.04      | 0.02     | 0.05     |
| Wrists & Hands | 0.06      | 0.04     | 0.08     |

Table 8: Lower torso

| body_site | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| Flank     | 0.07      | 0.05     | 0.09     |
| Abdomen   | 0.27      | 0.23     | 0.30     |
| Groin     | 0.09      | 0.07     | 0.11     |
| Hips      | 0.07      | 0.05     | 0.09     |
| Buttocks  | 0.03      | 0.02     | 0.05     |

Table 9: Spinal column

| body_site         | point_est | lower_ci | upper_ci |
|-------------------|-----------|----------|----------|
| Cervical spine    | 0.05      | 0.03     | 0.07     |
| Thoracic spine    | 0.12      | 0.10     | 0.15     |
| Lumbosacral spine | 0.15      | 0.13     | 0.18     |

Table 10: Lower limbs

| body_site | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| Legs      | 0.17      | 0.14     | 0.20     |
| Knees     | 0.14      | 0.11     | 0.17     |



| body_site     | point_est | lower_ci | upper_ci |
|---------------|-----------|----------|----------|
| Ankles & Feet | 0.31      | 0.27     | 0.35     |

### 3.3 Plotted proportions (with 95% CIs), by body region

```
upper <- prop_boot2$plots[[1]] +
  theme(axis.title.x = element_blank())

arm <- prop_boot2$plots[[2]] +
  theme(axis.title.x = element_blank())

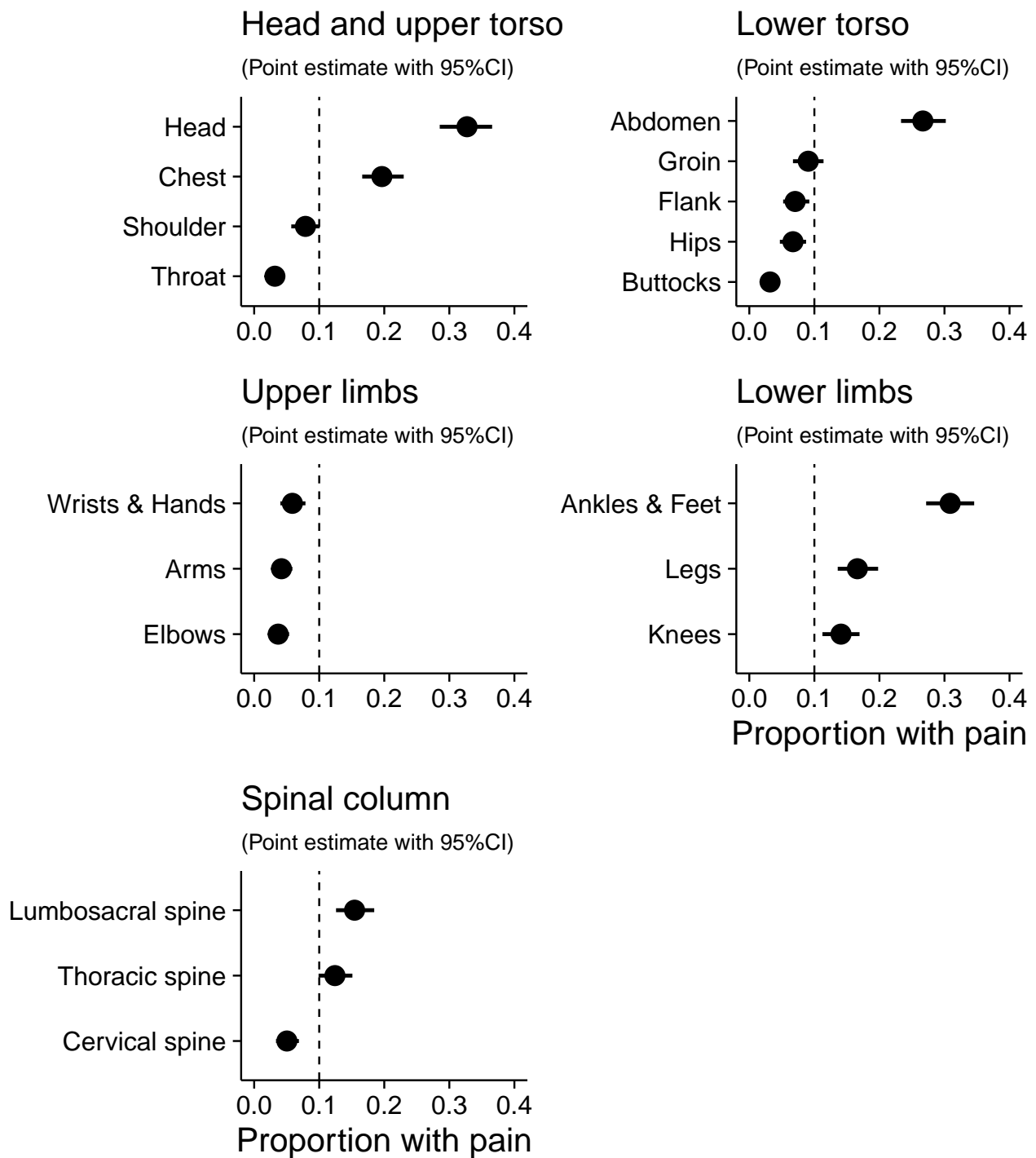
lower <- prop_boot2$plots[[3]] +
  theme(axis.title.x = element_blank())

spine <- prop_boot2$plots[[4]]

leg <- prop_boot2$plots[[5]]

# Patchwork
plot_prop <- upper + lower + arm + leg + spine +
  plot_layout(ncol = 2)

# Output
plot_prop
```



```
# Save
ggsave(filename = 'figures/figure_1.png',
        plot = plot_prop,
        width = 12,
        height = 12)
```

## 4 By sex

### 4.1 Process data

```
# Set seed
set.seed(2020)

# Select sex data
sex <- demo[, c('ID', 'Sex')]

# Join to boot_data & remove ID, site, and upper_back (only one outcome -- no pain)
sex <- left_join(data, sex) %>%
  select(-ID, -Site, -Upper_back)

# Bootstrap CIs
sex_boot <- sex %>%
  # Pivot to long format
  pivot_longer(cols = -Sex,
               names_to = 'body_site',
               values_to = 'pain_present') %>%
  # Add body regions
  mutate(region = case_when(
    body_site == 'Chest' |
    body_site == 'Head' |
    body_site == 'Throat' |
    body_site == 'Shoulder' ~ 'Head and upper torso',
    body_site == 'Lower_back' |
    body_site == 'Abdomen' |
    body_site == 'Hips' |
    body_site == 'Buttocks' |
    body_site == 'Groin' ~ 'Lower torso',
    body_site == 'Legs' |
    body_site == 'Knees' |
    body_site == 'Ankles.Feet' ~ 'Lower limbs',
    body_site == 'Arms' |
    body_site == 'Elbows' |
    body_site == 'Wrists.Hands' ~ 'Upper limbs',
    body_site == 'Cervical_spine' |
    body_site == 'Thoracic_spine' |
    body_site == 'Lumbosacral_spine' ~ 'Spinal column',
    TRUE ~ 'other'
  )) %>%
  # Nest by body region and body site
  group_by(Sex, region, body_site) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,
                    ~ boot(data = .x,
                          statistic = prop_func,
                          R = 999,
                          stype = 'i',
                          parallel = 'multicore',
                          ncpus = 4))) %>%
```

```

# Get CI
mutate(ci = map(.x = boot,
               ~ boot.ci(.x, type = 'perc')))) %>%
# Extract ci data
mutate(point_est = map(.x = ci,
                      ~ .x$t0),
       lower_ci = map(.x = ci,
                      ~ .x$percent[[4]]),
       upper_ci = map(.x = ci,
                      ~ .x$percent[[5]])) %>%
# Remove columns
select(-data, -boot, -ci) %>%
# Unnest
unnest(cols = c(point_est, lower_ci, upper_ci))

# Re-nest by body region and generate figures and tables
sex_boot2 <- sex_boot %>%
  group_by(region) %>%
  nest() %>%
  # Fix site labels
  mutate(data = map(.x = data,
                    ~ .x %>%
                      mutate(body_site = str_replace_all(body_site,
                                                            pattern = '_',
                                                            replacement = ' '),
                             body_site = str_replace_all(body_site,
                                                            pattern = '\\.',
                                                            replacement = ' & ')))) %>%
  # Re-order sites by point_est
  mutate(data = map(.x = data,
                    ~ .x %>%
                      mutate(body_site = fct_reorder(body_site,
                                                        point_est)))) %>%
# Plot data
mutate(plots = map2(.x = data,
                   .y = region,
                   ~ .x %>%
                     ggplot(data = .) +
                     aes(x = body_site,
                         y = point_est,
                         ymin = lower_ci,
                         ymax = upper_ci,
                         fill = Sex) +
                     geom_linerange(position = position_dodge2(width = 0.6),
                                     size = 1,
                                     colour = '#000000') +
                     geom_point(shape = 21,
                                colour = '#000000',
                                position = position_dodge2(width = 0.6),
                                size = 6,
                                stroke = 1) +
                     coord_flip() +
                     labs(title = .y,

```

```

        subtitle = '(Point estimate with 95%CI)',
        y = 'Proportion with pain') +
scale_y_continuous(limits = c(0, 1)) +
scale_fill_manual(values = c('#000000', '#FFFFFF')) +
theme_minimal(base_size = 18) +
theme(plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      legend.title = element_blank(),
      legend.position = 'top',
      axis.title.y = element_blank(),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))) %>%

# Tabulate data
mutate(tables = map2(.x = data,
                    .y = region,
                    ~ .x %>%
                      kable(caption = .y,
                            digits = 2)))

```

## 4.2 Tabulated proportions (with 95% CIs), by age and body region

```
walk(sex_boot2$tables, ~ print(.x))
```

Table 11: Head and upper torso

| Sex    | body_site | point_est | lower_ci | upper_ci |
|--------|-----------|-----------|----------|----------|
| Female | Head      | 0.38      | 0.33     | 0.42     |
| Female | Throat    | 0.03      | 0.01     | 0.04     |
| Female | Shoulder  | 0.07      | 0.05     | 0.10     |
| Female | Chest     | 0.20      | 0.16     | 0.23     |
| Male   | Head      | 0.12      | 0.06     | 0.18     |
| Male   | Throat    | 0.04      | 0.01     | 0.09     |
| Male   | Shoulder  | 0.11      | 0.06     | 0.18     |
| Male   | Chest     | 0.20      | 0.13     | 0.27     |

Table 12: Upper limbs

| Sex    | body_site      | point_est | lower_ci | upper_ci |
|--------|----------------|-----------|----------|----------|
| Female | Arms           | 0.04      | 0.03     | 0.06     |
| Female | Elbows         | 0.03      | 0.02     | 0.05     |
| Female | Wrists & Hands | 0.06      | 0.04     | 0.09     |
| Male   | Arms           | 0.03      | 0.01     | 0.07     |
| Male   | Elbows         | 0.05      | 0.02     | 0.10     |
| Male   | Wrists & Hands | 0.03      | 0.01     | 0.08     |

Table 13: Lower torso

| Sex    | body_site  | point_est | lower_ci | upper_ci |
|--------|------------|-----------|----------|----------|
| Female | Lower back | 0.07      | 0.05     | 0.10     |
| Female | Abdomen    | 0.28      | 0.24     | 0.32     |
| Female | Groin      | 0.10      | 0.07     | 0.12     |
| Female | Hips       | 0.06      | 0.04     | 0.09     |
| Female | Buttocks   | 0.03      | 0.01     | 0.04     |
| Male   | Lower back | 0.05      | 0.02     | 0.10     |
| Male   | Abdomen    | 0.19      | 0.12     | 0.27     |
| Male   | Groin      | 0.07      | 0.03     | 0.12     |
| Male   | Hips       | 0.10      | 0.05     | 0.15     |
| Male   | Buttocks   | 0.04      | 0.01     | 0.09     |

Table 14: Spinal column

| Sex    | body_site         | point_est | lower_ci | upper_ci |
|--------|-------------------|-----------|----------|----------|
| Female | Cervical spine    | 0.06      | 0.04     | 0.08     |
| Female | Thoracic spine    | 0.14      | 0.11     | 0.17     |
| Female | Lumbosacral spine | 0.16      | 0.12     | 0.19     |
| Male   | Cervical spine    | 0.03      | 0.00     | 0.06     |
| Male   | Thoracic spine    | 0.07      | 0.03     | 0.12     |
| Male   | Lumbosacral spine | 0.14      | 0.08     | 0.20     |

Table 15: Lower limbs

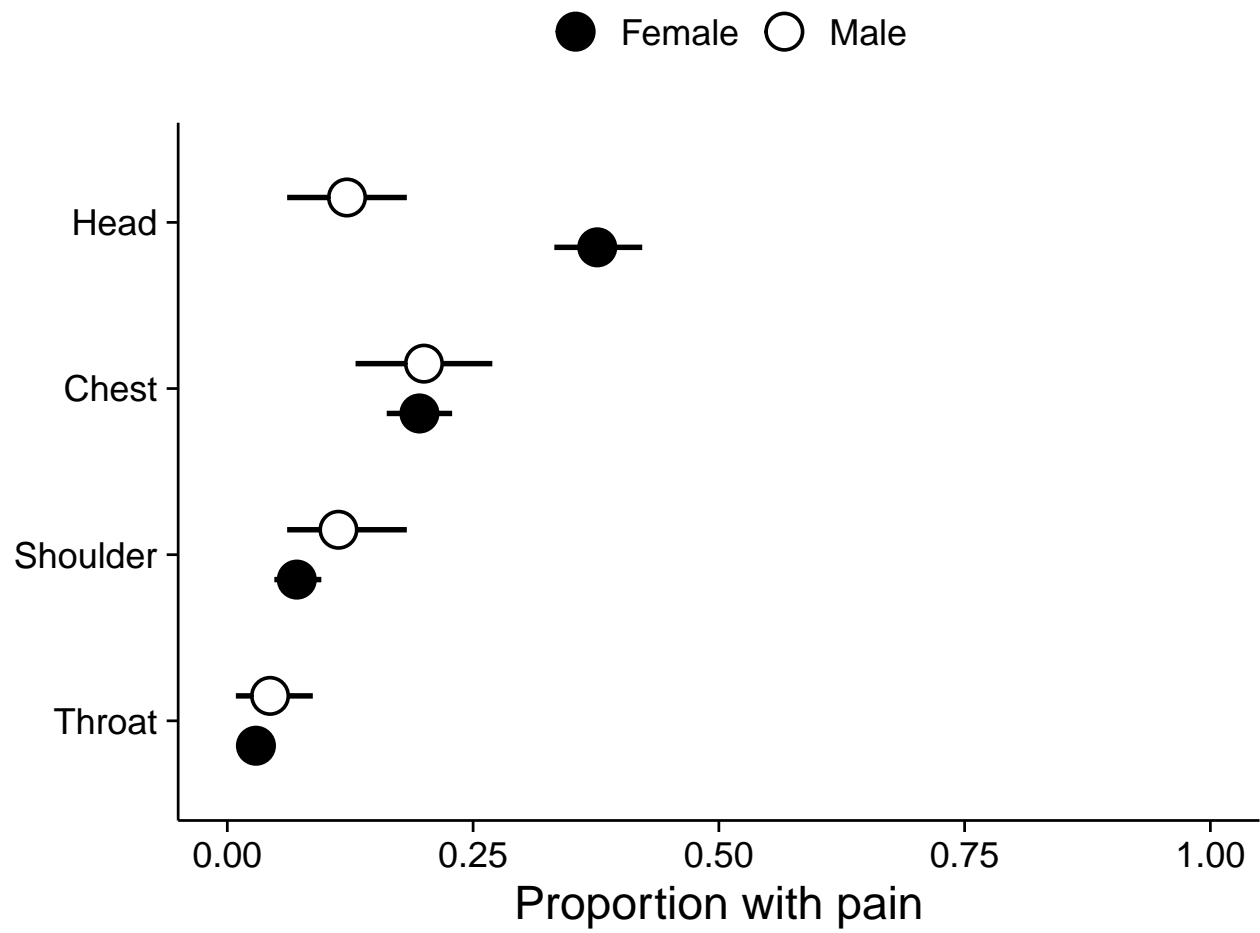
| Sex    | body_site     | point_est | lower_ci | upper_ci |
|--------|---------------|-----------|----------|----------|
| Female | Legs          | 0.15      | 0.12     | 0.18     |
| Female | Knees         | 0.13      | 0.10     | 0.16     |
| Female | Ankles & Feet | 0.28      | 0.25     | 0.32     |
| Male   | Legs          | 0.23      | 0.16     | 0.31     |
| Male   | Knees         | 0.19      | 0.12     | 0.26     |
| Male   | Ankles & Feet | 0.42      | 0.33     | 0.50     |

### 4.3 Plotted proportions (with 95% CIs), by age and body region

```
walk(sex_boot2$plots, ~ print(.x))
```

## Head and upper torso

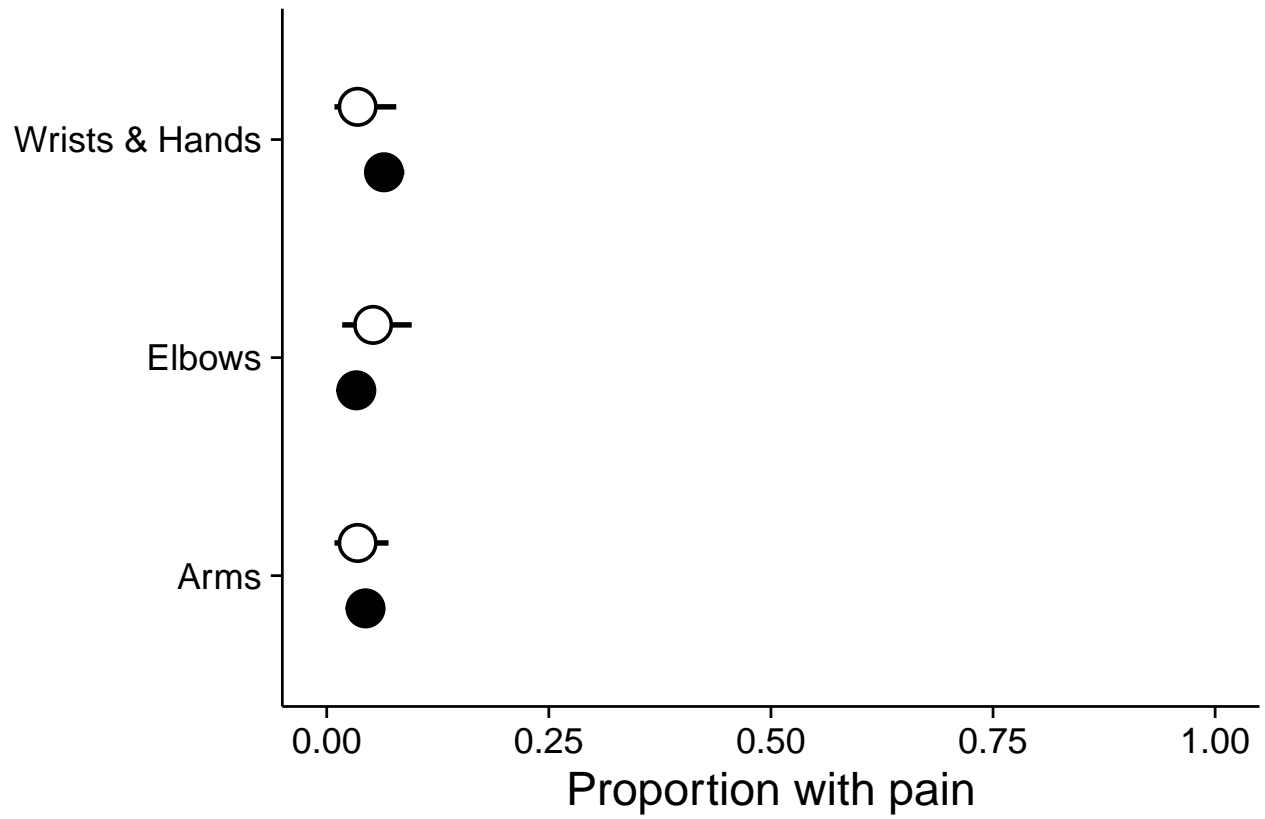
(Point estimate with 95%CI)



## Upper limbs

(Point estimate with 95%CI)

● Female ○ Male

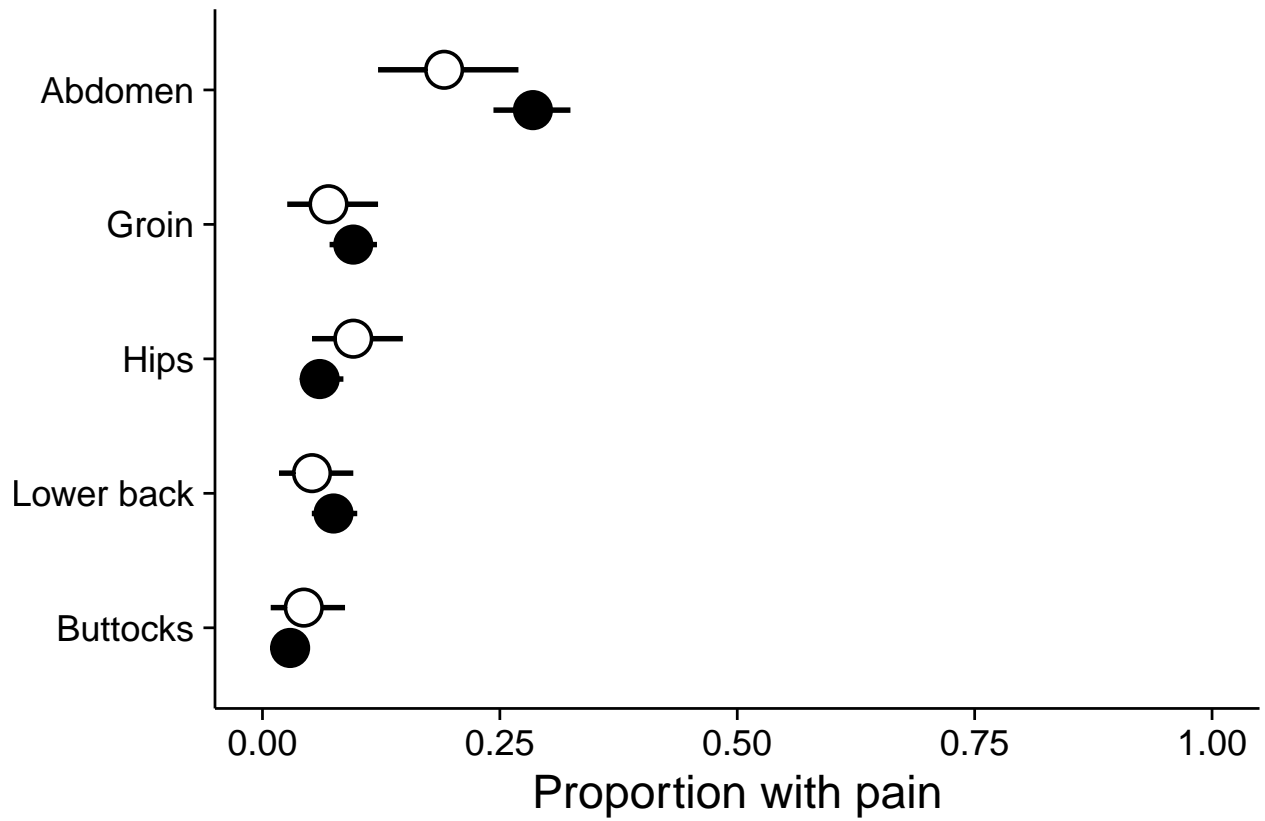




## Lower torso

(Point estimate with 95%CI)

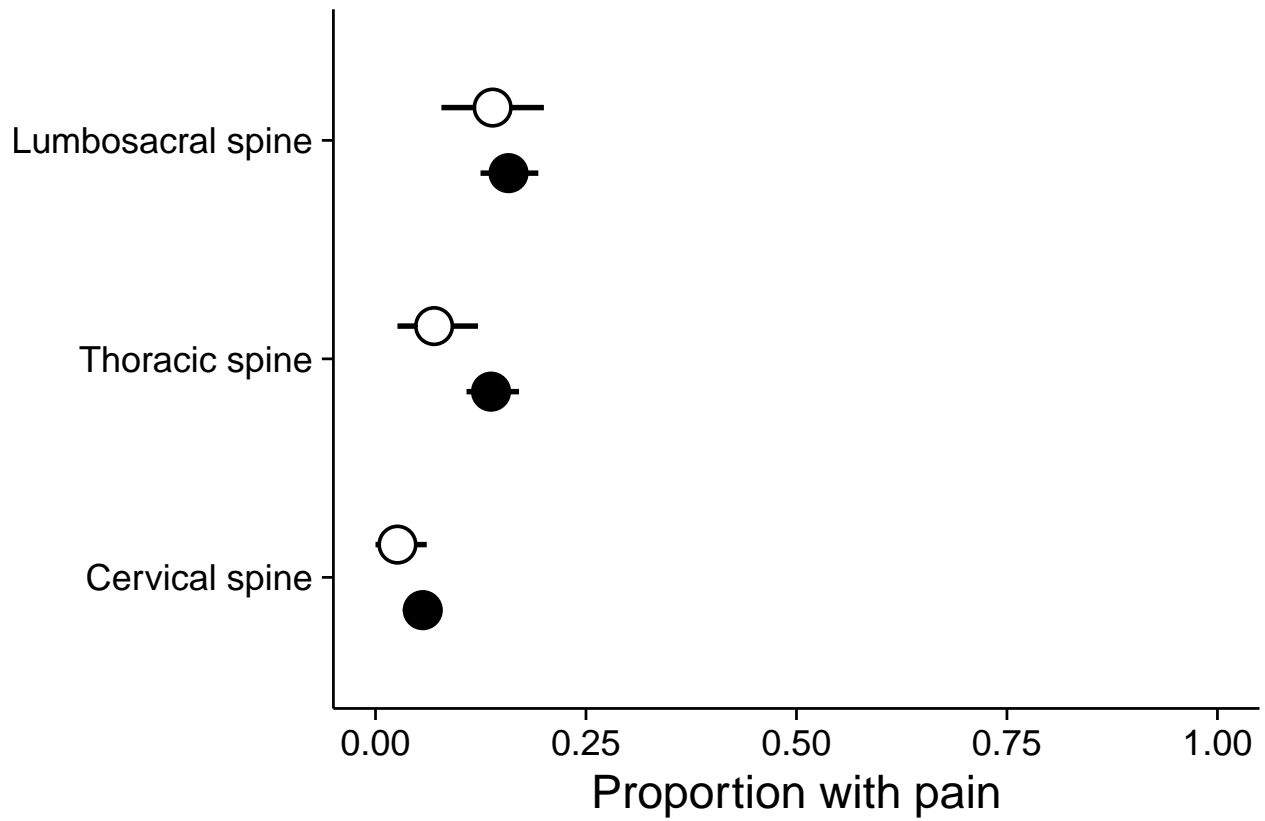
● Female ○ Male



## Spinal column

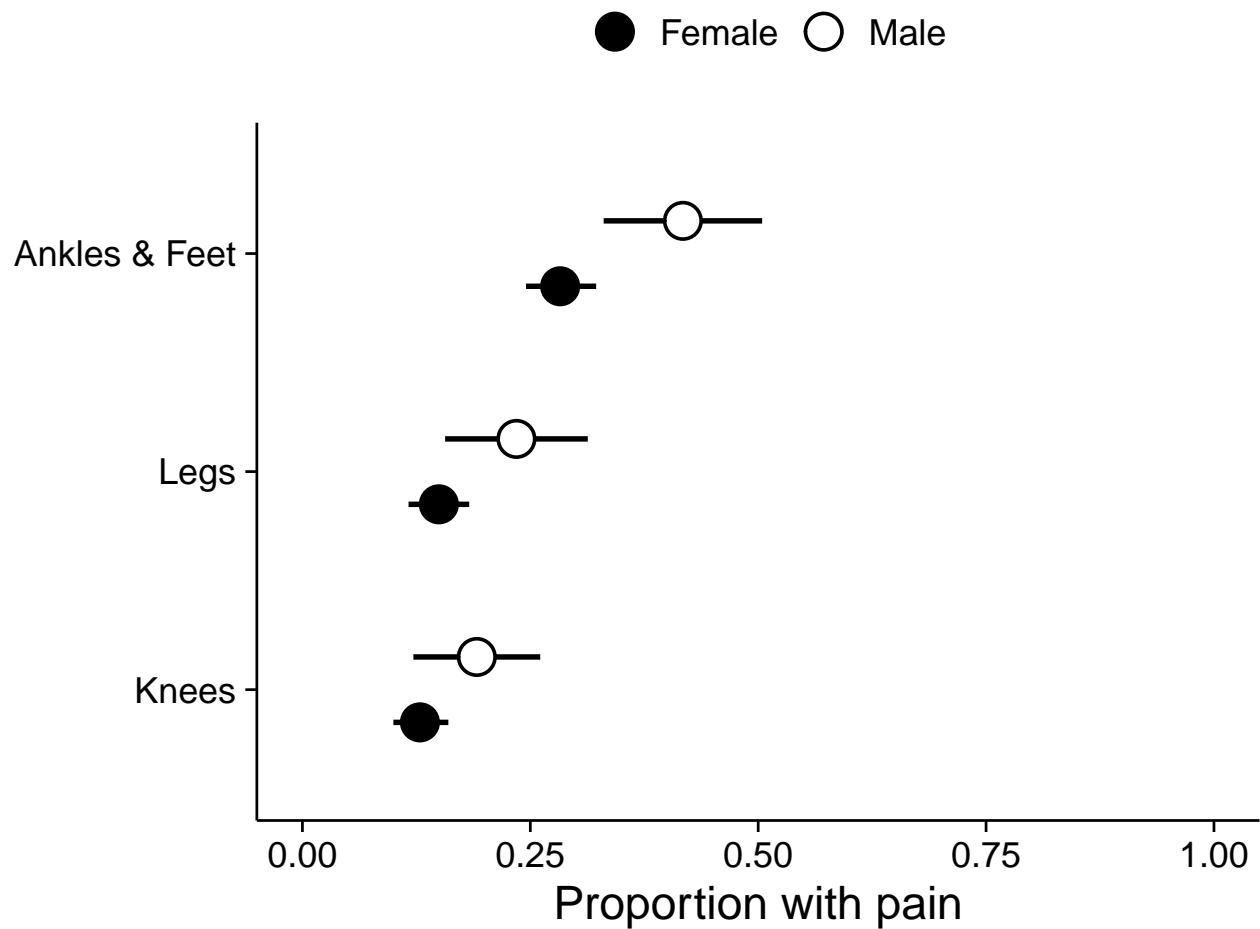
(Point estimate with 95%CI)

● Female ○ Male



## Lower limbs

(Point estimate with 95%CI)



## 5 By age

For ease of tabulation and plotting in this section of the supplement, I divided age into seven age categories. However for data analysis (see: Logistic regression, section 7), I analysed age as a continuous variable.

### 5.1 Process data

```
# Set seed
set.seed(2020)

# Select age data
age <- demo[, c('ID', 'Age')]

# Join to boot_data & remove ID, study site, and upper_back
```

```

# (only one outcome -- no pain)
age <- left_join(data, age) %>%
  select(-ID, -Site, -Upper_back)

# Get complete cases
age <- age[complete.cases(age), ]

# Pivot and add age group categories (10 year periods)
age_boot <- age %>%
  # Pivot to long format
  pivot_longer(cols = -Age,
               names_to = 'body_site',
               values_to = 'pain_present') %>%
  # Add age categories
  mutate(age_group = case_when(
    Age < 28 ~ '18-27',
    Age >= 28 & Age < 38 ~ '28-37',
    Age >= 38 & Age < 48 ~ '38-47',
    Age >= 48 & Age < 58 ~ '48-57',
    Age >= 58 & Age < 68 ~ '58-67',
    Age >= 68 & Age < 78 ~ '68-77',
    Age >= 78 & Age < 88 ~ '78-87'
  ))

# Print count per age group
age_boot %>%
  group_by(body_site, age_group) %>%
  summarise(count = n()) %>%
  filter(body_site == 'Abdomen') %>%
  ungroup() %>%
  select(-body_site) %>%
  kable(caption = 'Participant count per age group')

```

Table 16: Participant count per age group

| age_group | count |
|-----------|-------|
| 18-27     | 65    |
| 28-37     | 283   |
| 38-47     | 160   |
| 48-57     | 58    |
| 58-67     | 20    |
| 68-77     | 2     |

```

# Generate CIs
age_boot2 <- age_boot %>%
  # Remove age
  select(-Age) %>%
  # Remove categories with less than 20 counts
  filter(age_group != '68-77') %>%
  # Nest by age group and body site
  group_by(age_group, body_site) %>%
  nest() %>%

```

```

# Bootstrap data
mutate(boot = map(.x = data,
  ~ boot(data = .x,
    statistic = prop_func,
    R = 999,
    stype = 'i',
    parallel = 'multicore',
    ncpus = 4))) %>%

# Get CI
mutate(ci = map(.x = boot,
  ~ boot.ci(.x, type = 'perc')) %>%

# Extract ci data
mutate(point_est = map(.x = ci,
  ~ .x$t0),
  lower_ci = map(.x = ci,
    ~ .x$percent[[4]]),
  upper_ci = map(.x = ci,
    ~ .x$percent[[5]])) %>%

# Remove columns
select(-data, -boot, -ci) %>%

# Unnest
unnest(cols = c(point_est, lower_ci, upper_ci)) %>%
ungroup()

# Re-nest by body region and generate figures and tables
age_boot2 <- age_boot2 %>%

# Fix site labels
mutate(body_site = str_replace_all(body_site,
  pattern = '_',
  replacement = ' '),
  body_site = str_replace_all(body_site,
  pattern = '\\.',
  replacement = ' & ')) %>%

# Group and nest
group_by(body_site) %>%
nest() %>%

# Arrange age groups
# Plot data
mutate(plots = map2(.x = data,
  .y = body_site,
  ~ .x %>%
    ggplot(data = .) +
    aes(x = age_group,
      y = point_est,
      ymin = lower_ci,
      ymax = upper_ci) +
    geom_linerange(size = 1,
      colour = '#000000') +
    geom_point(colour = '#000000',
      size = 6) +
    labs(title = .y,
      subtitle = '(Point estimate with 95%CI)',
      caption = 'Age group 68-77 years removed because n = 2',

```

```

    x = 'Age group (Years)',
    y = 'Proportion with pain') +
scale_y_continuous(limits = c(0, 1)) +
theme_minimal(base_size = 18) +
theme(plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      plot.caption = element_text(size = 12),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))) %>%

# Tabulate data
mutate(
  tables = map2(.x = data,
                .y = body_site,
                ~ .x %>%
                  arrange(age_group) %>%
                  kable(caption = str_glue('{.y} (Age group 68-77 years removed because n = 2)',
                        digits = 2)))

```

## 5.2 Tabulated proportions (with 95% CIs), by age group and body site

```
walk(age_boot2$tables, ~ print(.x))
```

Table 17: Head (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.45      | 0.32     | 0.57     |
| 28-37     | 0.34      | 0.28     | 0.39     |
| 38-47     | 0.31      | 0.24     | 0.39     |
| 48-57     | 0.17      | 0.09     | 0.28     |
| 58-67     | 0.35      | 0.15     | 0.55     |

Table 18: Throat (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.05      | 0.00     | 0.11     |
| 28-37     | 0.03      | 0.01     | 0.05     |
| 38-47     | 0.03      | 0.01     | 0.06     |
| 48-57     | 0.05      | 0.00     | 0.12     |
| 58-67     | 0.00      | 0.00     | 0.00     |

Table 19: Shoulder (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.12      | 0.05     | 0.22     |
| 28-37     | 0.07      | 0.04     | 0.10     |

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 38-47     | 0.09      | 0.05     | 0.14     |
| 48-57     | 0.07      | 0.02     | 0.14     |
| 58-67     | 0.05      | 0.00     | 0.15     |

Table 20: Arms (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.02      | 0.00     | 0.05     |
| 28-37     | 0.04      | 0.02     | 0.07     |
| 38-47     | 0.06      | 0.03     | 0.09     |
| 48-57     | 0.05      | 0.00     | 0.12     |
| 58-67     | 0.00      | 0.00     | 0.00     |

Table 21: Elbows (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.00      | 0.00     | 0.00     |
| 28-37     | 0.03      | 0.01     | 0.06     |
| 38-47     | 0.04      | 0.01     | 0.07     |
| 48-57     | 0.05      | 0.00     | 0.10     |
| 58-67     | 0.15      | 0.00     | 0.30     |

Table 22: Wrists & Hands (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.02      | 0.00     | 0.05     |
| 28-37     | 0.07      | 0.04     | 0.10     |
| 38-47     | 0.04      | 0.01     | 0.07     |
| 48-57     | 0.07      | 0.02     | 0.14     |
| 58-67     | 0.15      | 0.00     | 0.30     |

Table 23: Chest (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.17      | 0.08     | 0.26     |
| 28-37     | 0.20      | 0.15     | 0.25     |
| 38-47     | 0.20      | 0.14     | 0.26     |
| 48-57     | 0.19      | 0.09     | 0.29     |
| 58-67     | 0.20      | 0.05     | 0.40     |

Table 24: Lower back (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.09      | 0.03     | 0.17     |
| 28-37     | 0.07      | 0.04     | 0.10     |
| 38-47     | 0.05      | 0.02     | 0.09     |
| 48-57     | 0.10      | 0.03     | 0.19     |
| 58-67     | 0.10      | 0.00     | 0.25     |

Table 25: Abdomen (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.23      | 0.14     | 0.34     |
| 28-37     | 0.31      | 0.26     | 0.37     |
| 38-47     | 0.29      | 0.22     | 0.37     |
| 48-57     | 0.10      | 0.03     | 0.19     |
| 58-67     | 0.05      | 0.00     | 0.15     |

Table 26: Cervical spine (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.12      | 0.05     | 0.20     |
| 28-37     | 0.05      | 0.02     | 0.08     |
| 38-47     | 0.04      | 0.02     | 0.08     |
| 48-57     | 0.00      | 0.00     | 0.00     |
| 58-67     | 0.00      | 0.00     | 0.00     |

Table 27: Thoracic spine (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.18      | 0.09     | 0.28     |
| 28-37     | 0.14      | 0.11     | 0.19     |
| 38-47     | 0.07      | 0.03     | 0.11     |
| 48-57     | 0.09      | 0.02     | 0.17     |
| 58-67     | 0.25      | 0.05     | 0.45     |

Table 28: Lumbosacral spine (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.18      | 0.09     | 0.29     |
| 28-37     | 0.15      | 0.11     | 0.19     |



| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 38-47     | 0.12      | 0.07     | 0.17     |
| 48-57     | 0.16      | 0.07     | 0.26     |
| 58-67     | 0.40      | 0.20     | 0.60     |

Table 29: Groin (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.09      | 0.03     | 0.17     |
| 28-37     | 0.10      | 0.07     | 0.14     |
| 38-47     | 0.09      | 0.05     | 0.13     |
| 48-57     | 0.05      | 0.00     | 0.12     |
| 58-67     | 0.05      | 0.00     | 0.15     |

Table 30: Hips (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.03      | 0.00     | 0.08     |
| 28-37     | 0.05      | 0.02     | 0.07     |
| 38-47     | 0.09      | 0.05     | 0.14     |
| 48-57     | 0.14      | 0.05     | 0.24     |
| 58-67     | 0.10      | 0.00     | 0.25     |

Table 31: Legs (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.12      | 0.05     | 0.22     |
| 28-37     | 0.13      | 0.09     | 0.17     |
| 38-47     | 0.21      | 0.15     | 0.28     |
| 48-57     | 0.29      | 0.17     | 0.41     |
| 58-67     | 0.20      | 0.05     | 0.40     |

Table 32: Knees (Age group 68-77 years removed because n = 2)

| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.11      | 0.05     | 0.18     |
| 28-37     | 0.10      | 0.07     | 0.14     |
| 38-47     | 0.17      | 0.11     | 0.23     |
| 48-57     | 0.22      | 0.12     | 0.34     |
| 58-67     | 0.30      | 0.10     | 0.50     |

Table 33: Ankles & Feet (Age group 68-77 years removed because n = 2)

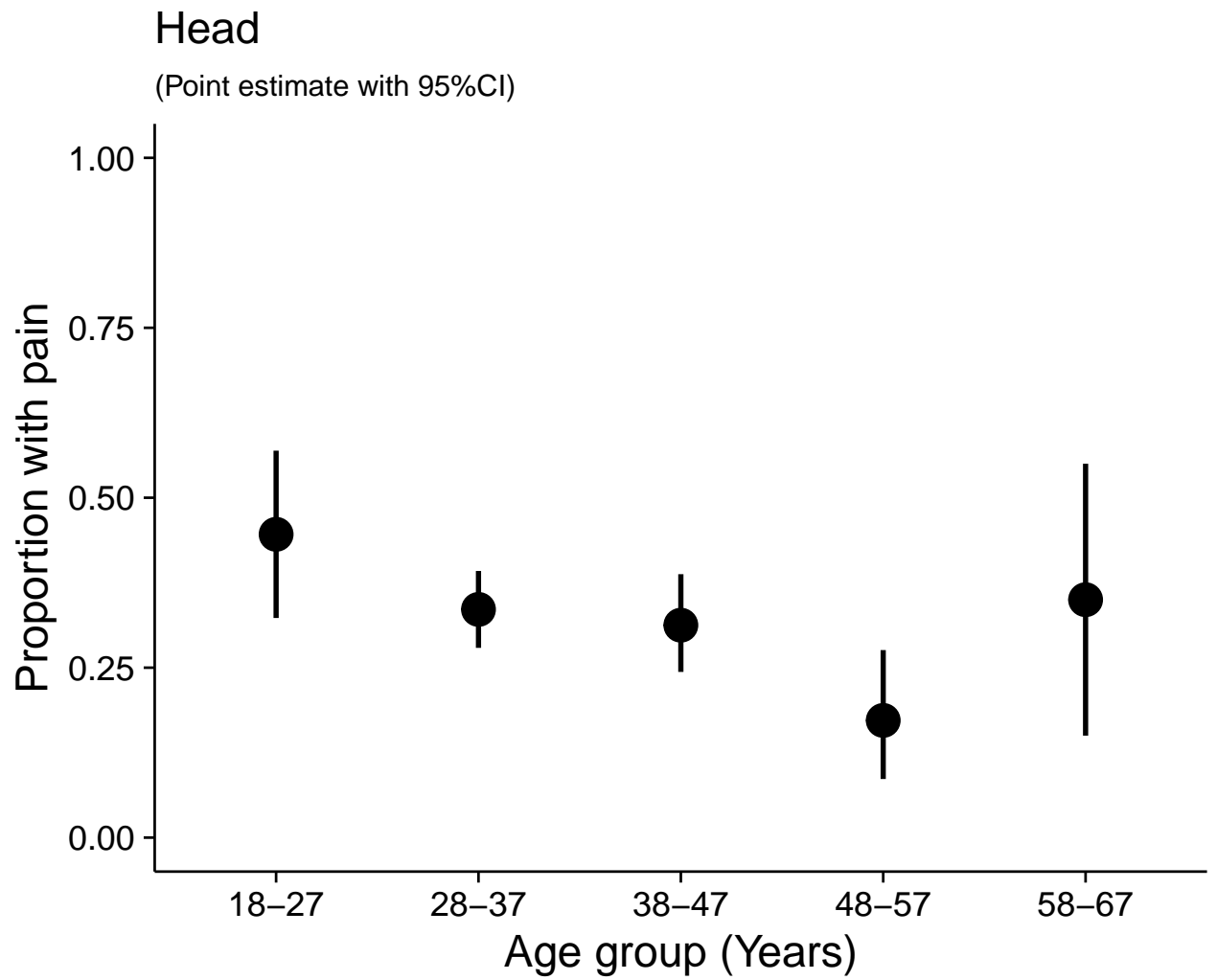
| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.22      | 0.12     | 0.32     |
| 28-37     | 0.24      | 0.19     | 0.29     |
| 38-47     | 0.37      | 0.29     | 0.44     |
| 48-57     | 0.52      | 0.40     | 0.64     |
| 58-67     | 0.55      | 0.35     | 0.75     |

Table 34: Buttocks (Age group 68-77 years removed because n = 2)

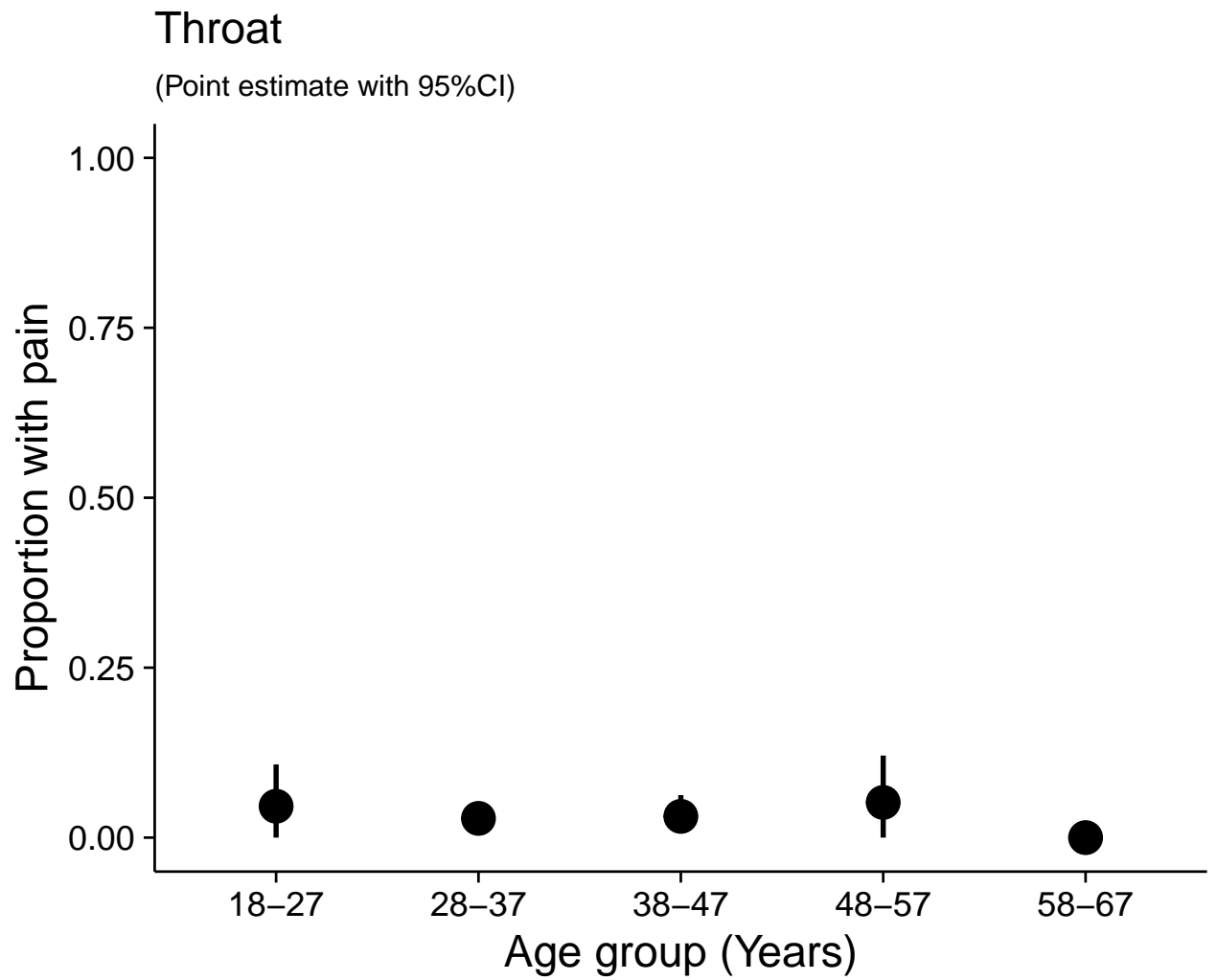
| age_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 18-27     | 0.03      | 0.00     | 0.08     |
| 28-37     | 0.04      | 0.02     | 0.06     |
| 38-47     | 0.03      | 0.01     | 0.05     |
| 48-57     | 0.02      | 0.00     | 0.05     |
| 58-67     | 0.05      | 0.00     | 0.15     |

### 5.3 Plotted proportions (with 95% CIs), by age group and body site

```
walk(age_boot2$plots, ~ print(.x))
```



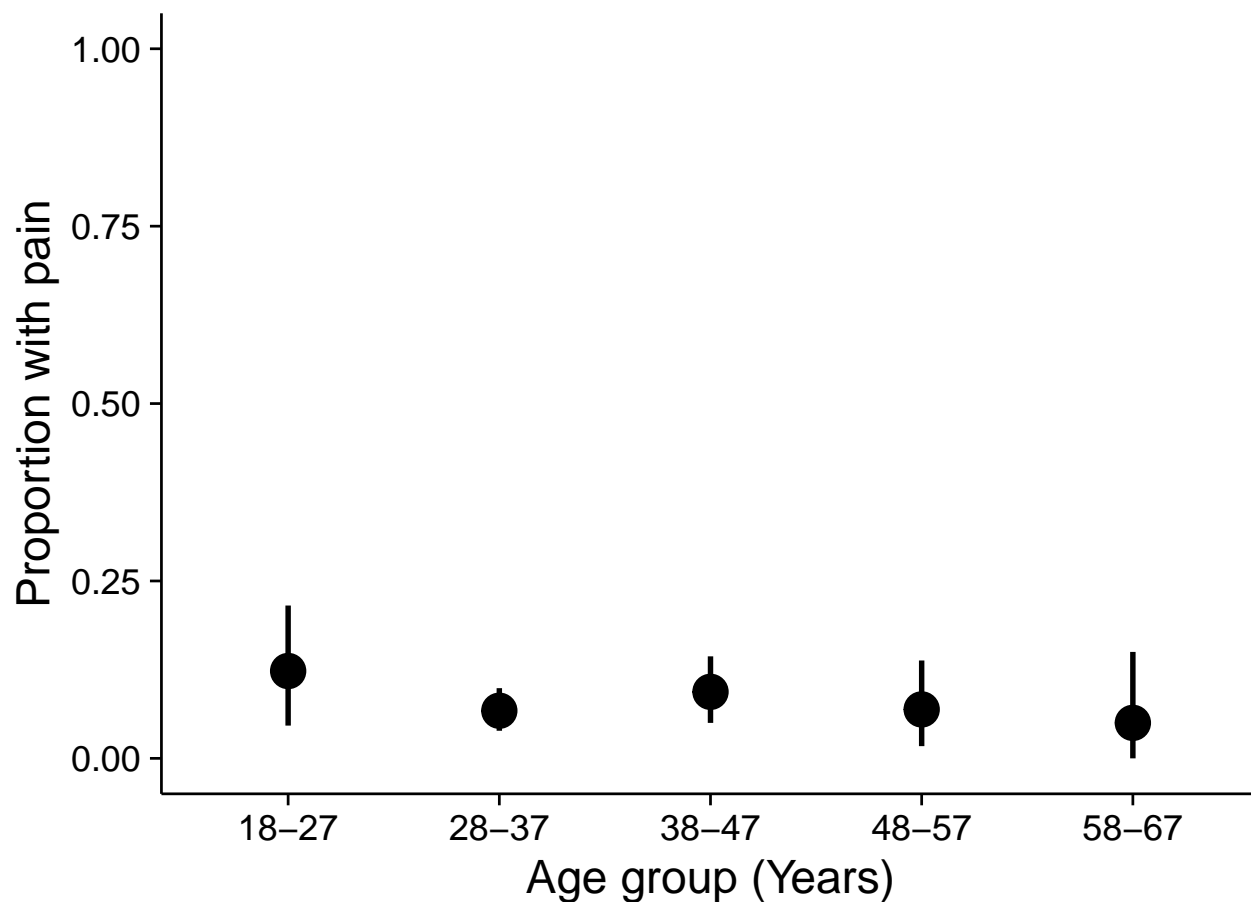
Age group 68-77 years removed because n = 2



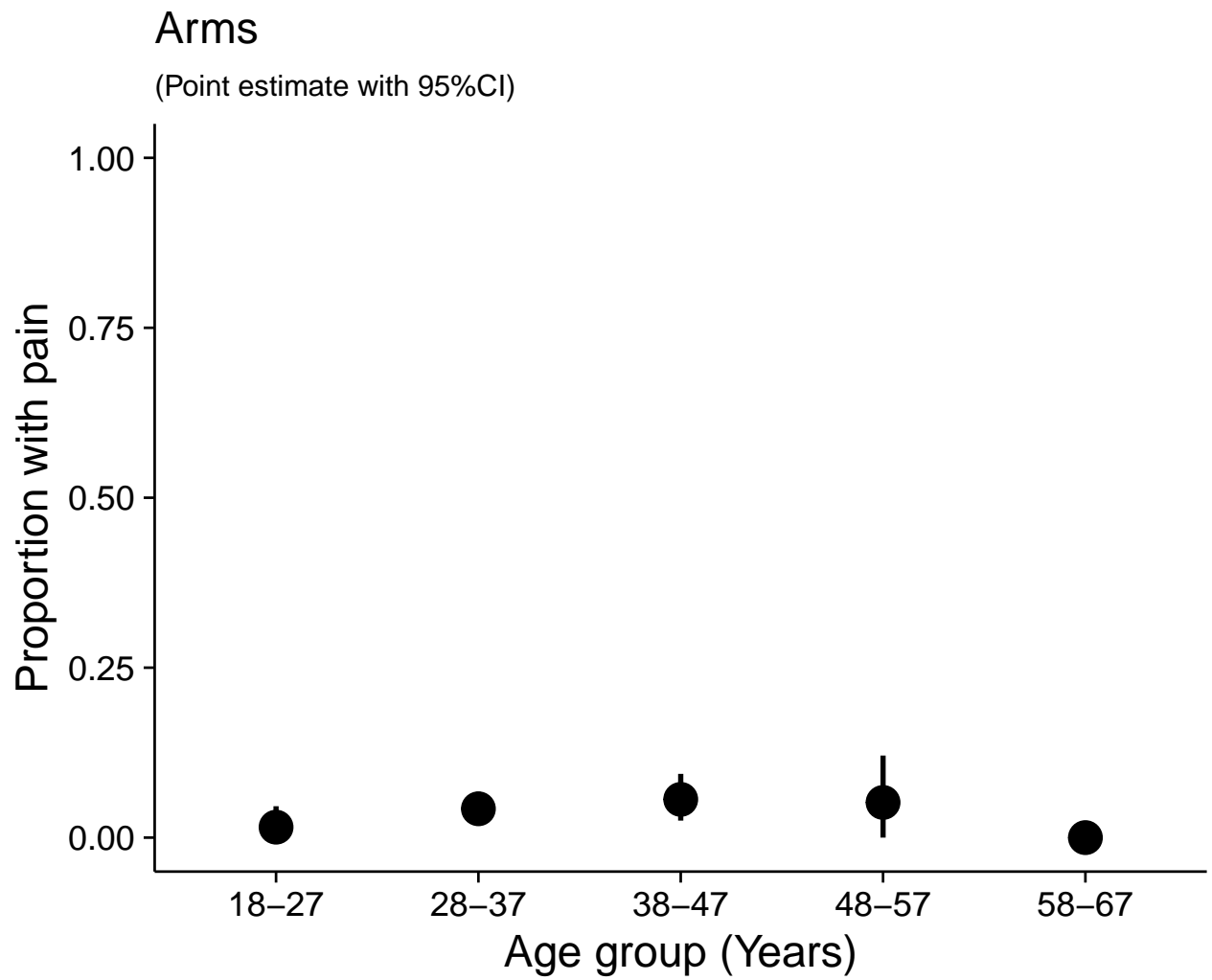
Age group 68-77 years removed because n = 2

## Shoulder

(Point estimate with 95%CI)



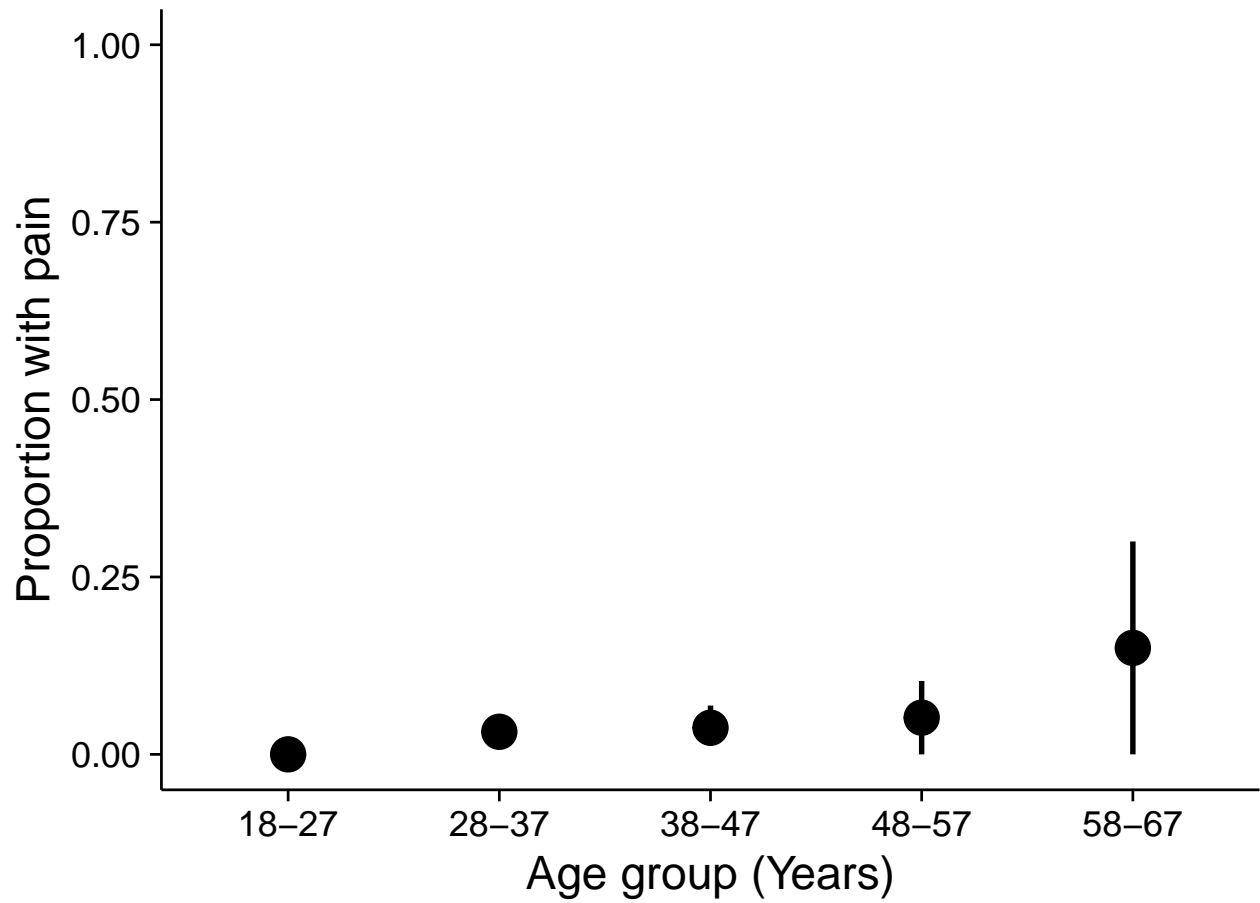
Age group 68-77 years removed because n = 2



Age group 68-77 years removed because n = 2

## Elbows

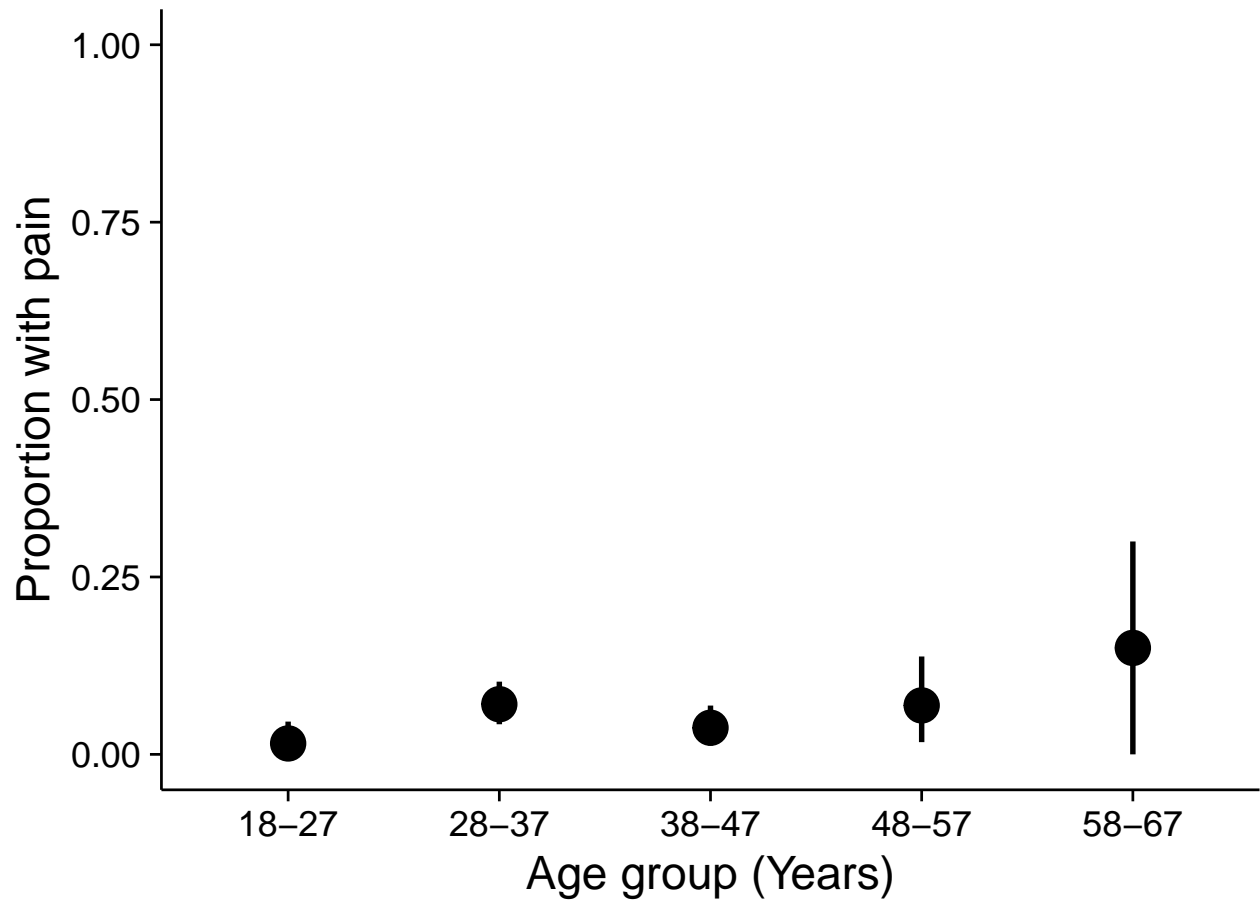
(Point estimate with 95%CI)



Age group 68-77 years removed because n = 2

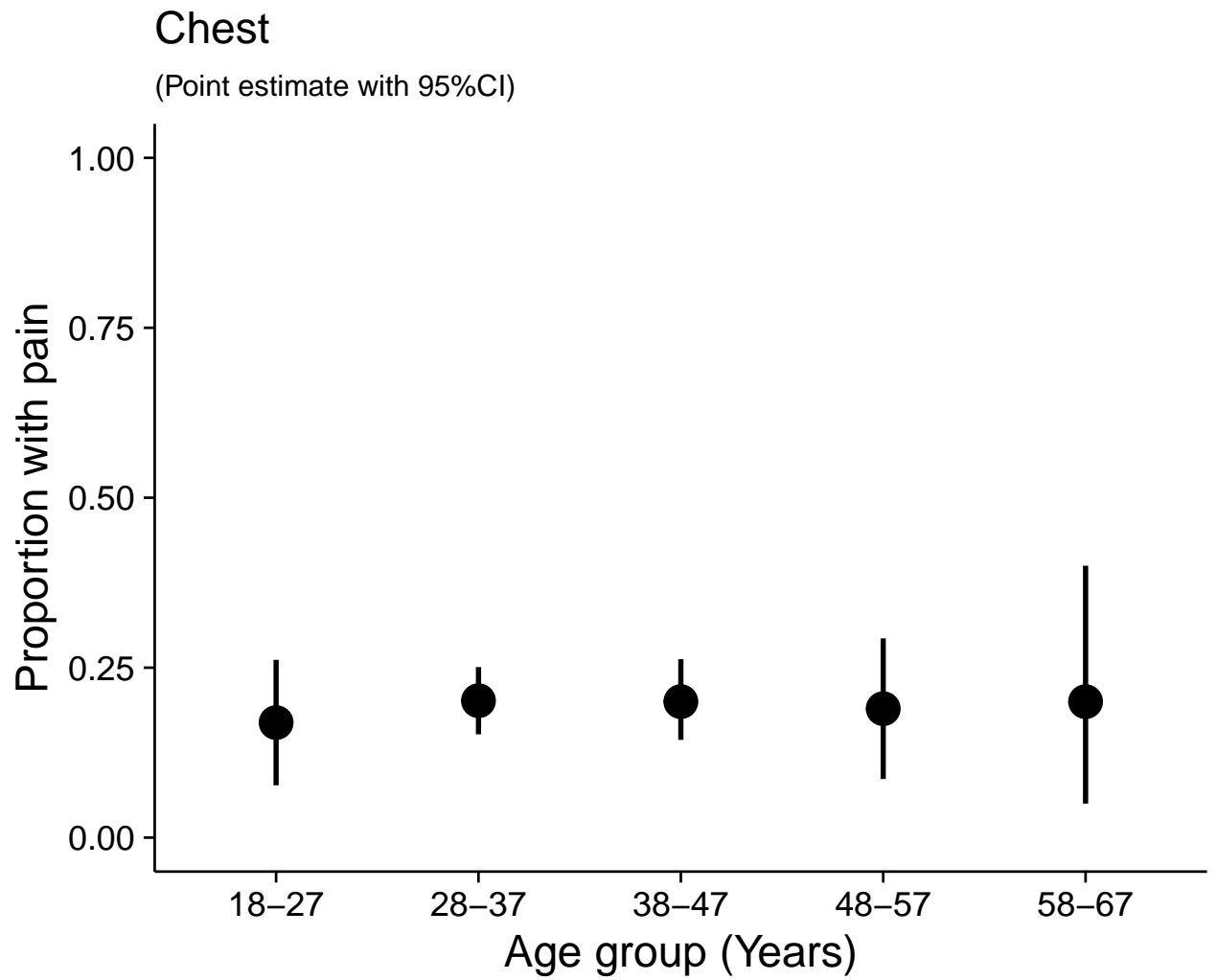
## Wrists & Hands

(Point estimate with 95%CI)

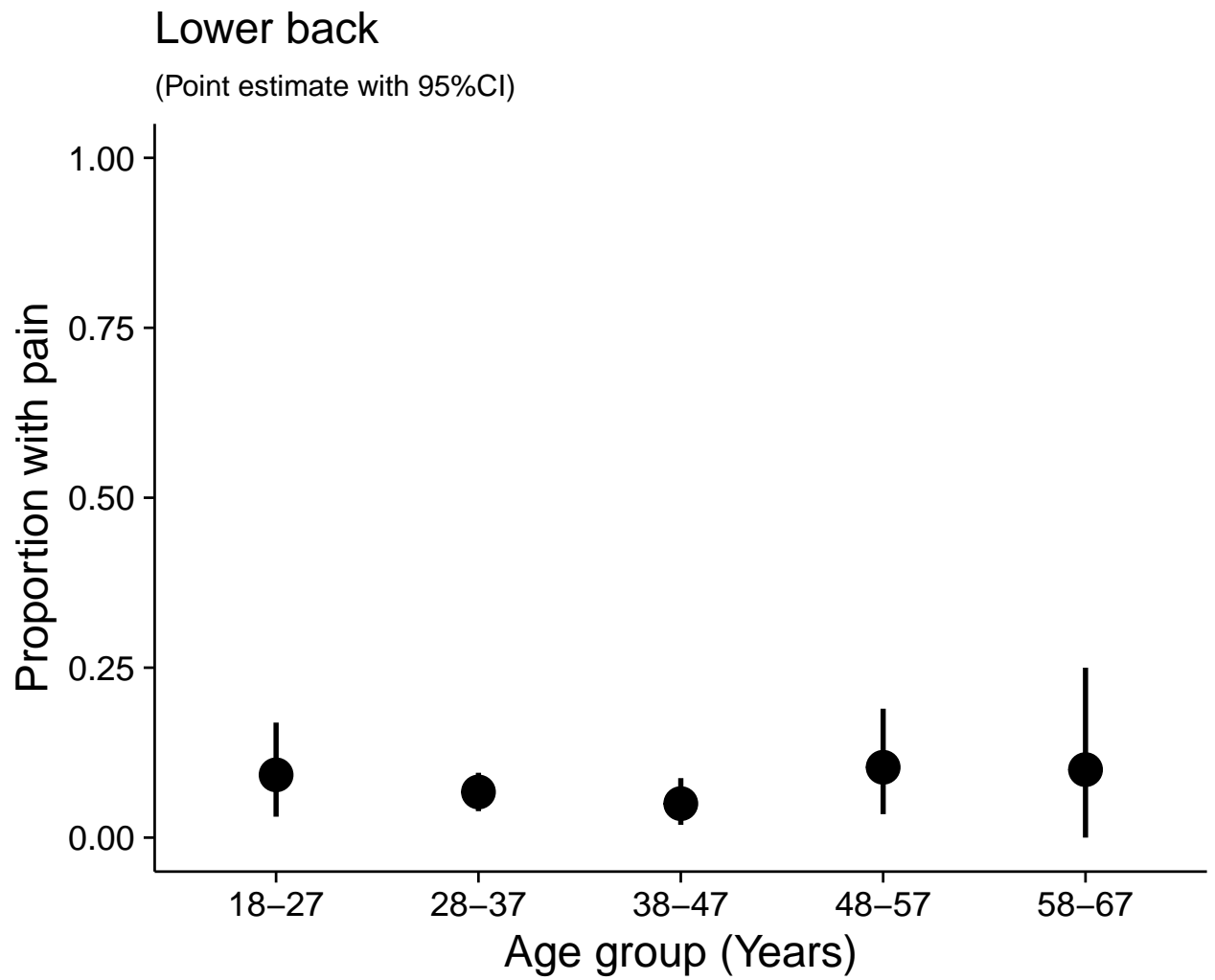


Age group 68-77 years removed because n = 2

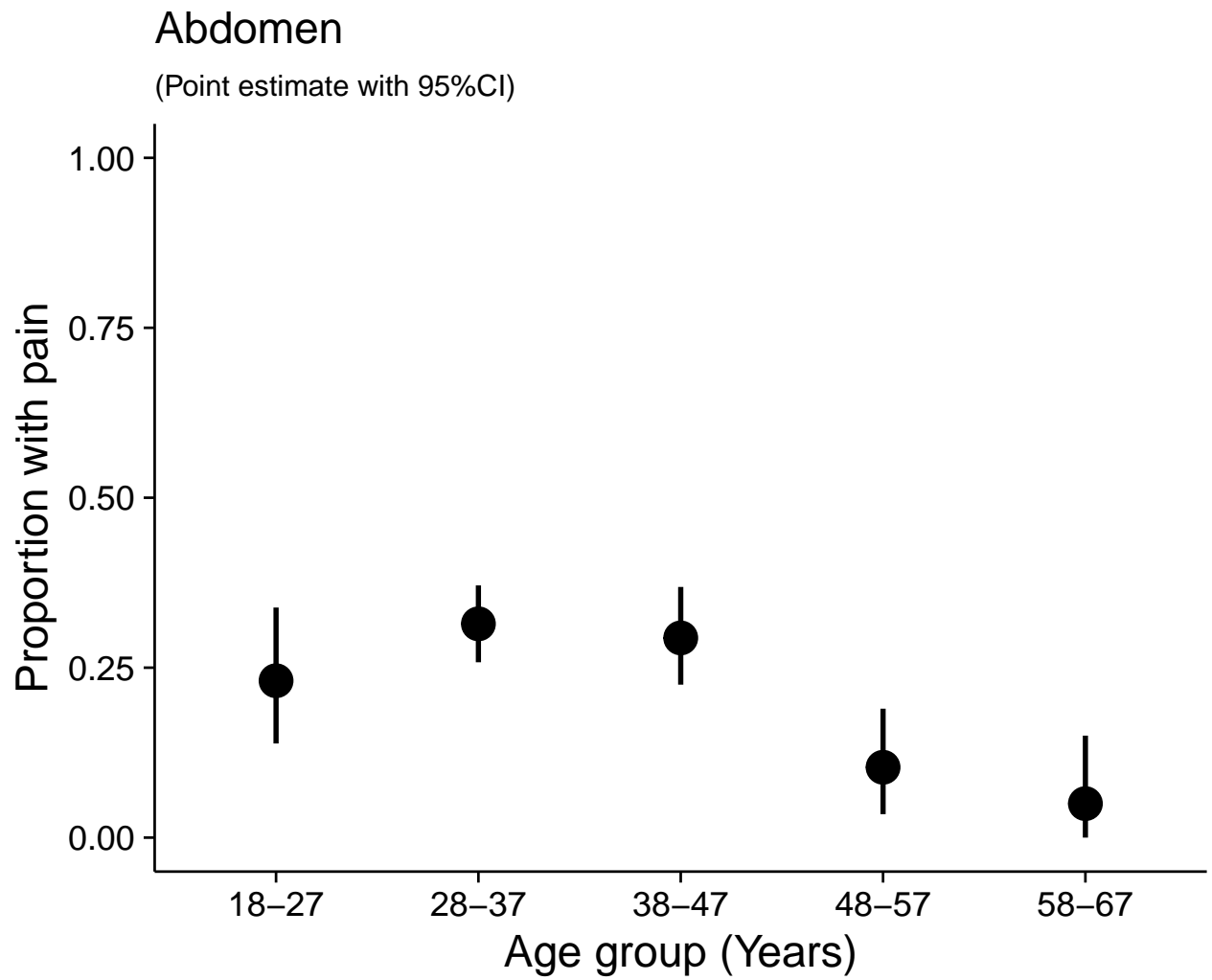




Age group 68-77 years removed because n = 2



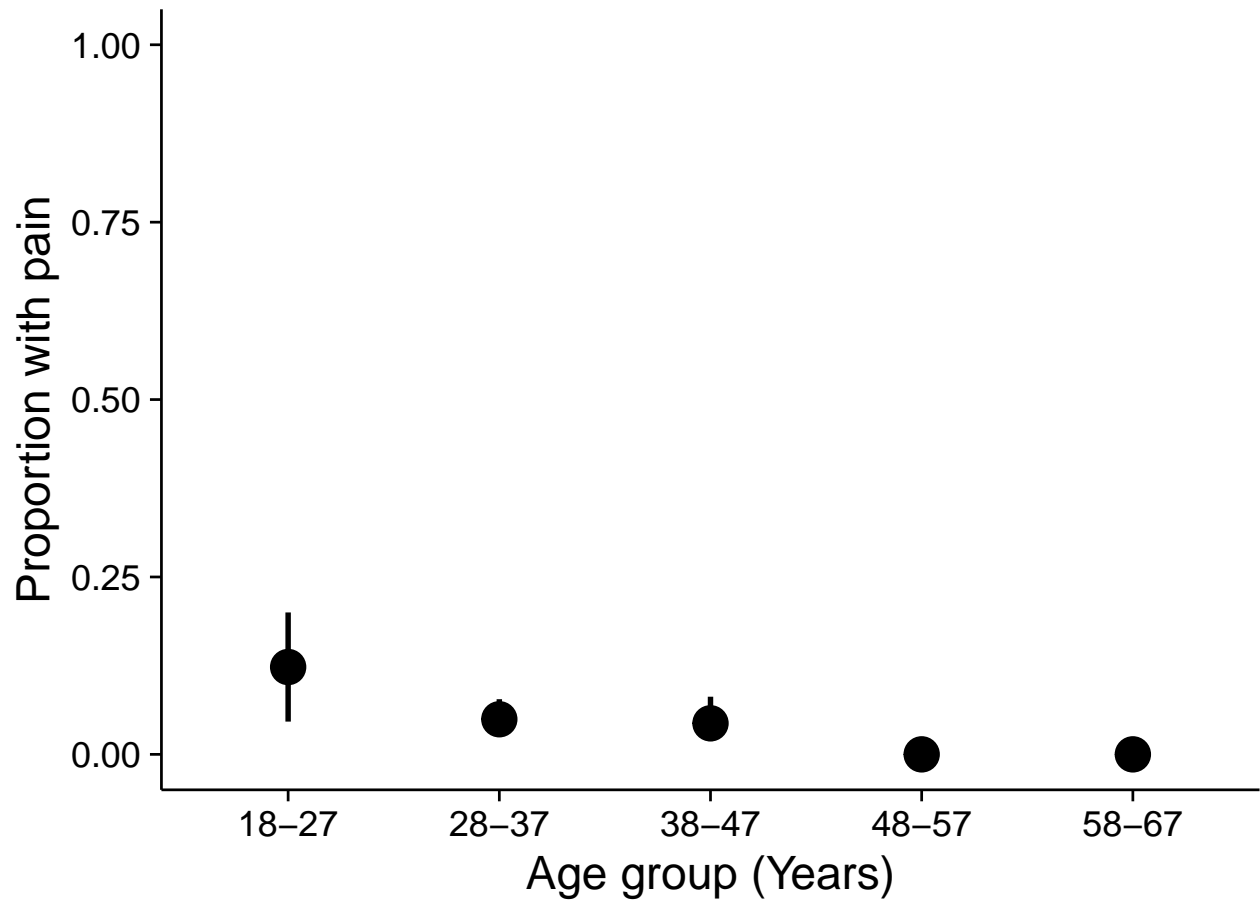
Age group 68-77 years removed because n = 2



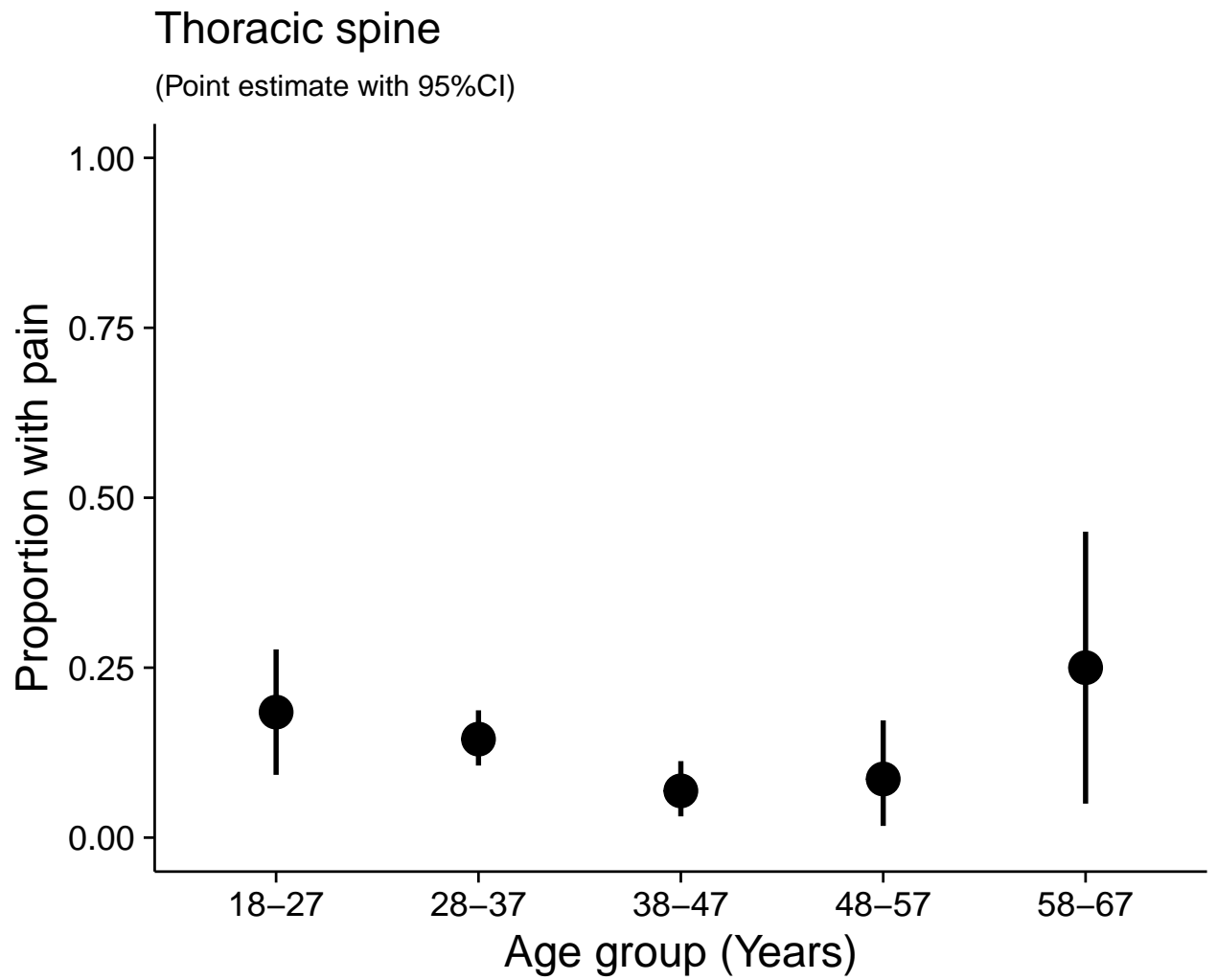
Age group 68-77 years removed because n = 2

## Cervical spine

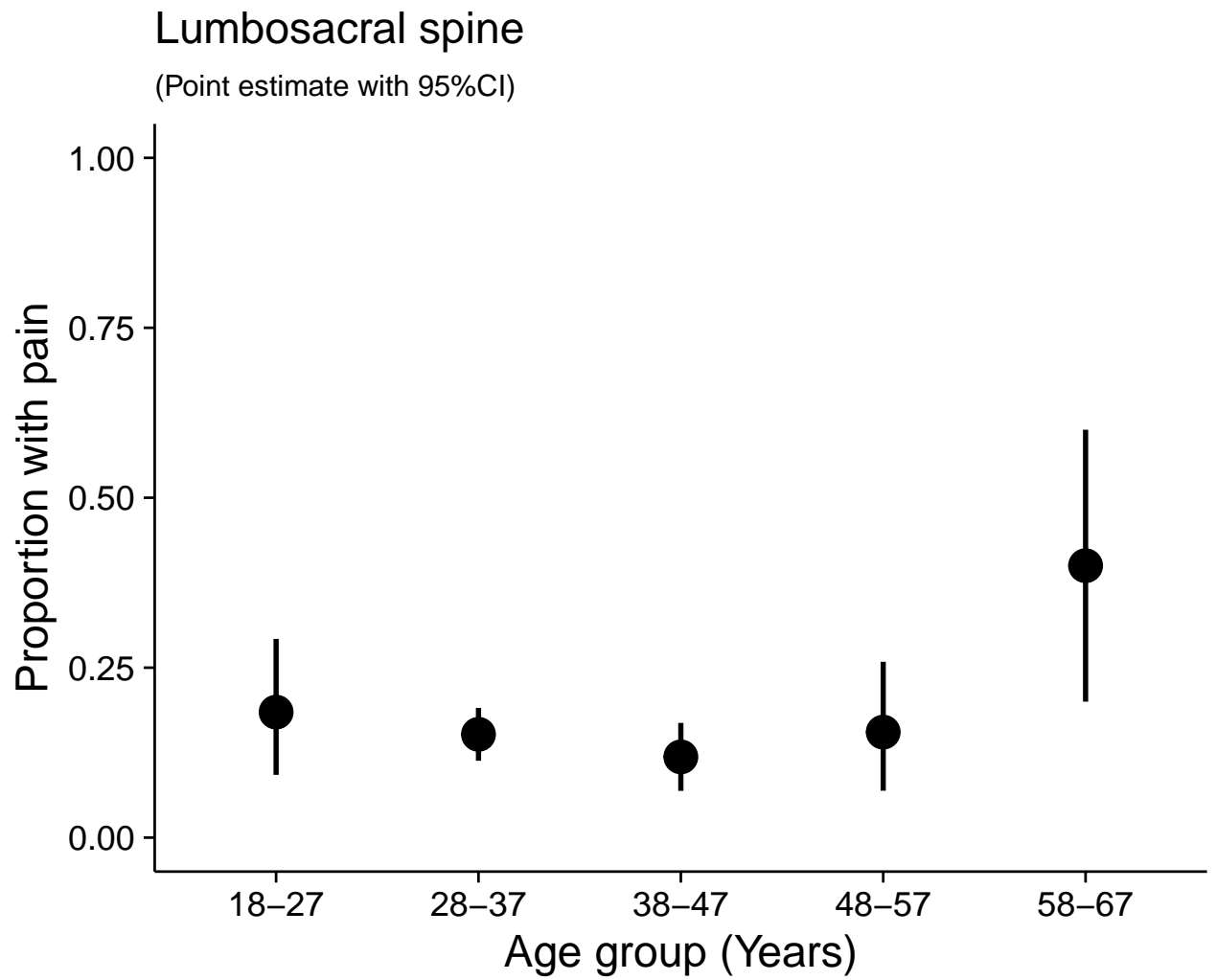
(Point estimate with 95%CI)



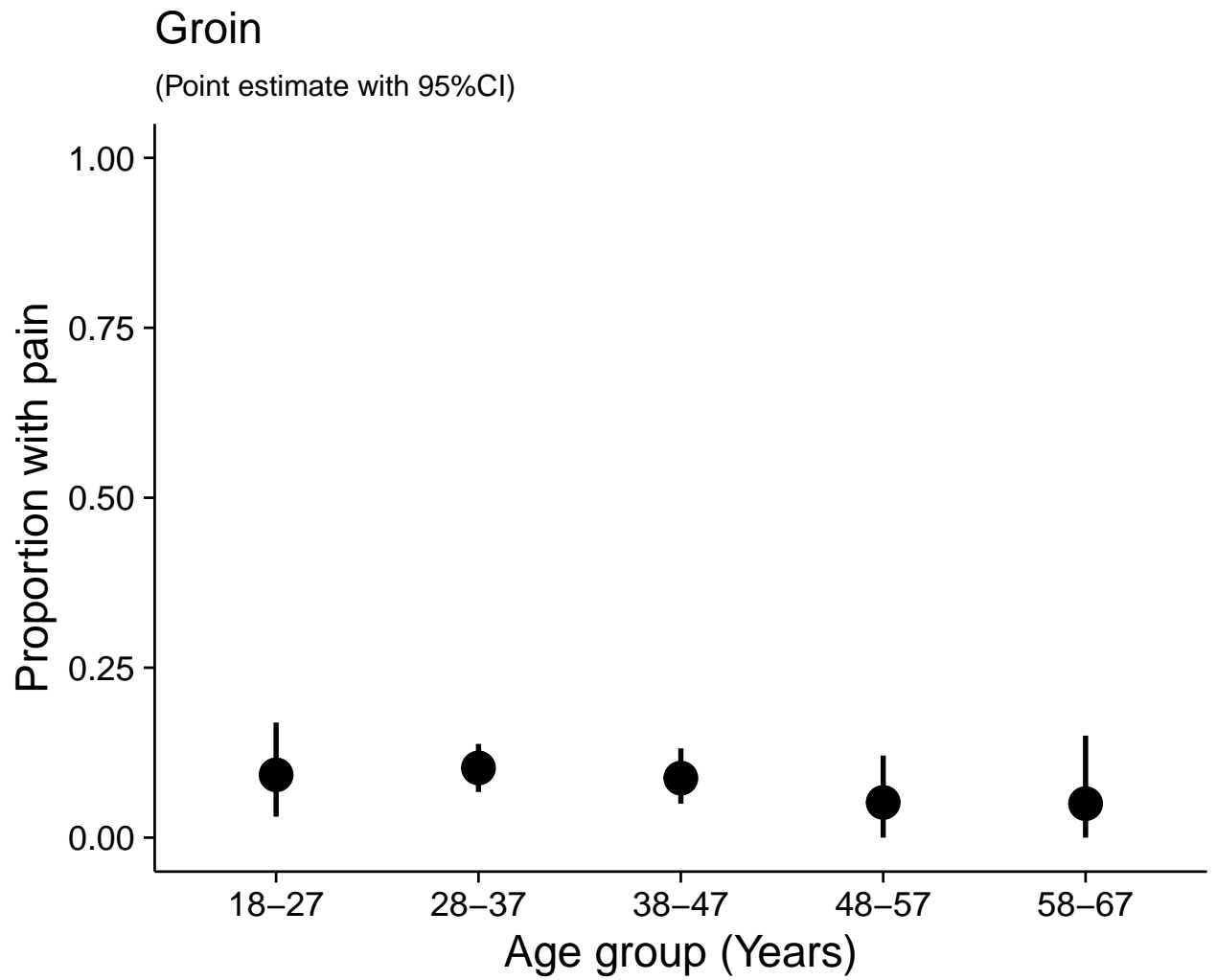
Age group 68-77 years removed because n = 2



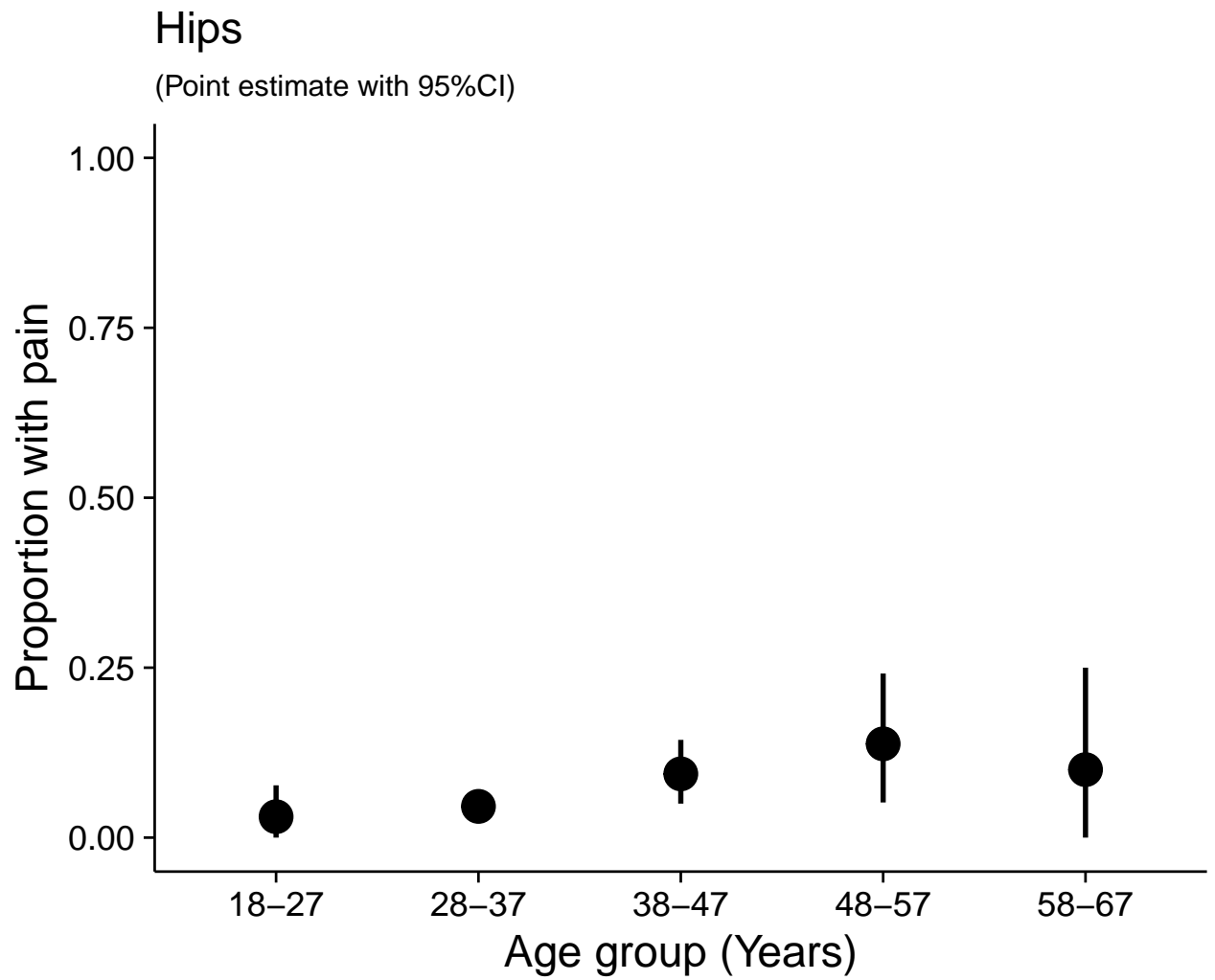
Age group 68-77 years removed because n = 2



Age group 68-77 years removed because n = 2

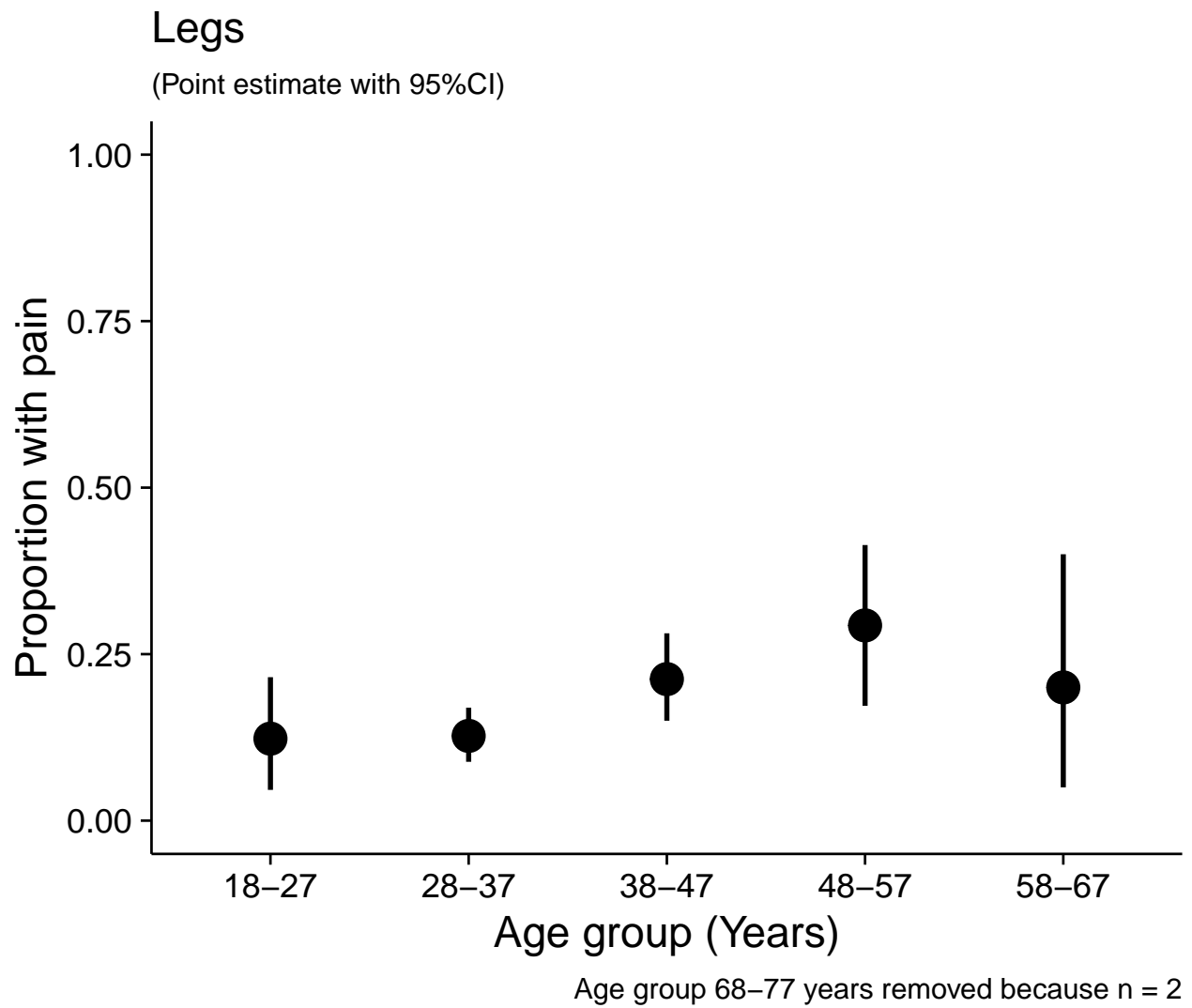


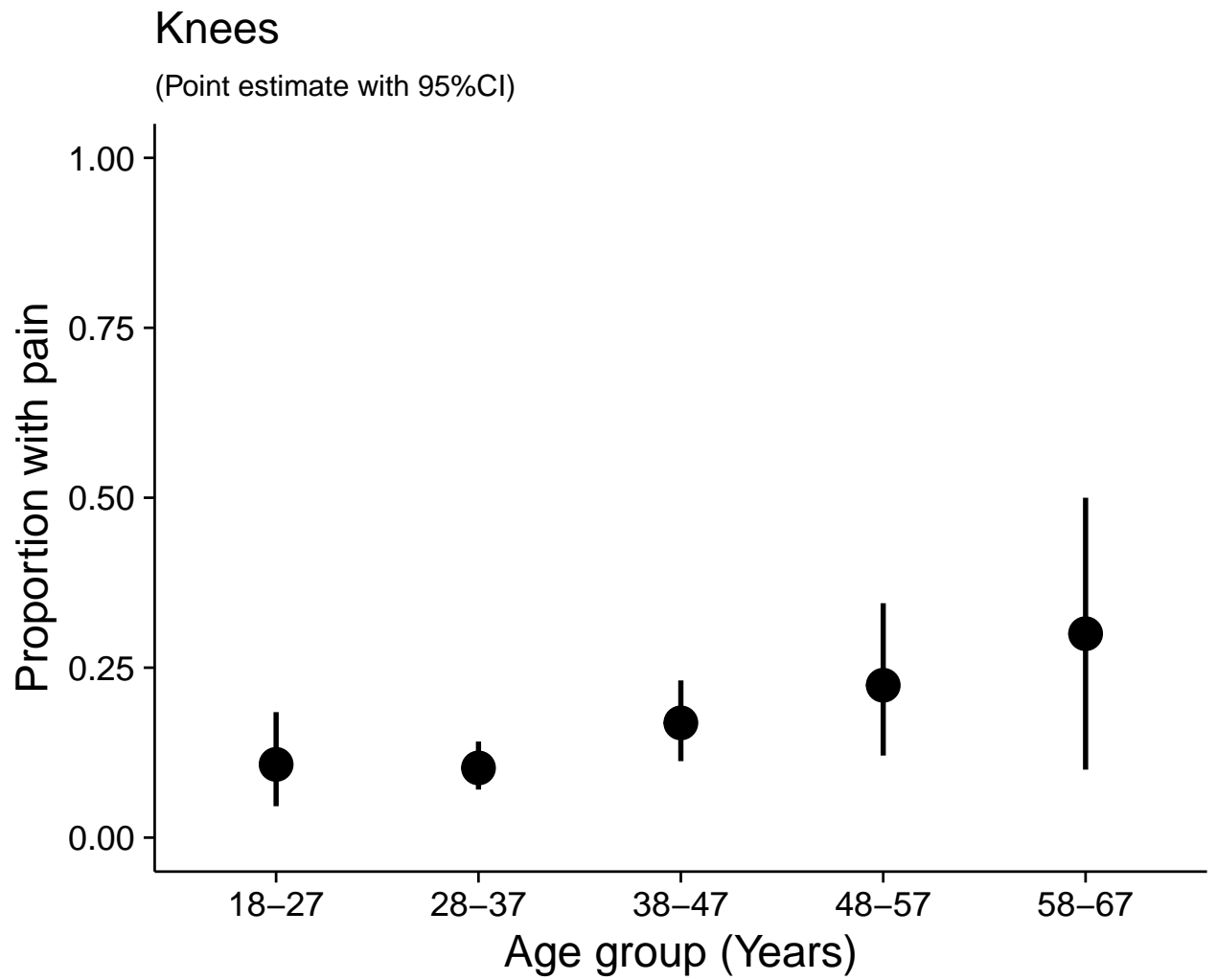
Age group 68-77 years removed because n = 2



Age group 68-77 years removed because n = 2



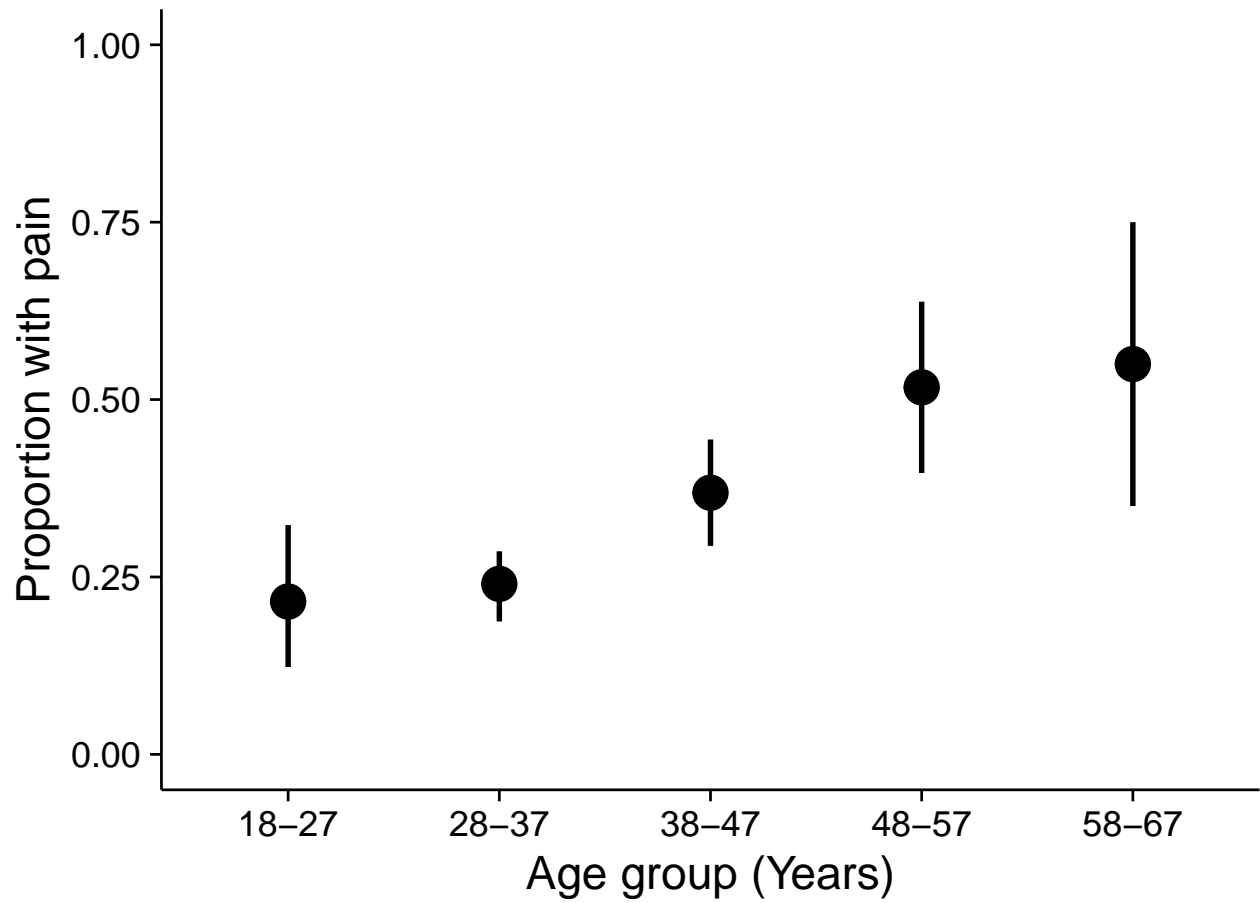




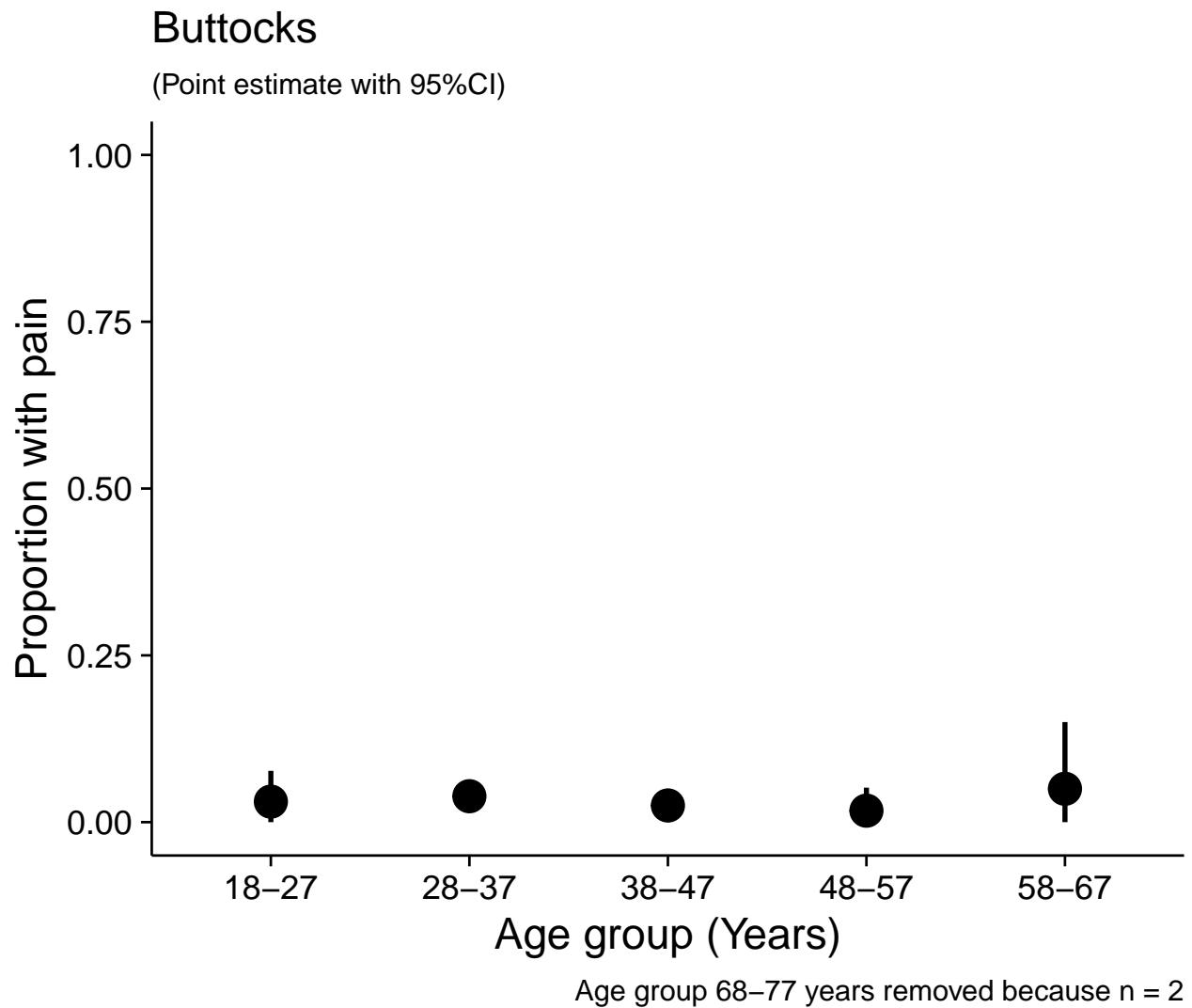
Age group 68-77 years removed because n = 2

## Ankles & Feet

(Point estimate with 95%CI)



Age group 68-77 years removed because n = 2



## 6 By most recent CD4 T-cell count

For ease of tabulation and plotting in this section of the supplement, I divided the most recent CD4 T-cell count into six categories. However for data analysis (see: Logistic regression, section 7), I analysed CD4 T-cell count as a continuous variable.

### 6.1 Process data

```
# Set seed
set.seed(2020)

# Select CD4 recent data
cd4 <- demo[, c('ID', 'CD4_recent')]
```

```

# Join to boot_data & remove ID, study site, and upper_back
# (only one outcome -- no pain)
cd4 <- left_join(data, cd4) %>%
  select(-ID, -Site, -Upper_back)

# Get complete cases
cd4 <- cd4[complete.cases(cd4), ]

# Pivot and add CD4 recent group categories (counts of 100)
cd4_boot <- cd4 %>%
  # Pivot to long format
  pivot_longer(cols = -CD4_recent,
               names_to = 'body_site',
               values_to = 'pain_present') %>%
  # Add CD4 recent categories
  mutate(cd4_group = case_when(
    CD4_recent < 100 ~ '0-99',
    CD4_recent >= 100 & CD4_recent < 200 ~ '100-199',
    CD4_recent >= 200 & CD4_recent < 300 ~ '200-299',
    CD4_recent >= 300 & CD4_recent < 400 ~ '300-399',
    CD4_recent >= 400 & CD4_recent < 500 ~ '400-499',
    CD4_recent >= 500 ~ '500+'
  ))

# Print count per CD4 recent group
cd4_boot %>%
  group_by(body_site, cd4_group) %>%
  summarise(count = n()) %>%
  filter(body_site == 'Abdomen') %>%
  ungroup() %>%
  select(-body_site) %>%
  kable(caption = 'Participant count per CD4 group')

```

Table 35: Participant count per CD4 group

| cd4_group | count |
|-----------|-------|
| 0-99      | 68    |
| 100-199   | 106   |
| 200-299   | 108   |
| 300-399   | 77    |
| 400-499   | 46    |
| 500+      | 92    |

```

# Generate CIs
cd4_boot2 <- cd4_boot %>%
  # Remove CD4 recent
  select(-CD4_recent) %>%
  # Nest by CD4 recent group and body site
  group_by(cd4_group, body_site) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,

```

```

      ~ boot(data = .x,
              statistic = prop_func,
              R = 999,
              stype = 'i',
              parallel = 'multicore',
              ncpus = 4))) %>%

# Get CI
mutate(ci = map(.x = boot,
               ~ boot.ci(.x, type = 'perc')))) %>%
# Extract ci data
mutate(point_est = map(.x = ci,
                      ~ .x$t0),
       lower_ci = map(.x = ci,
                      ~ .x$percent[[4]]),
       upper_ci = map(.x = ci,
                      ~ .x$percent[[5]])) %>%

# Remove columns
select(-data, -boot, -ci) %>%
# Unnest
unnest(cols = c(point_est, lower_ci, upper_ci)) %>%
ungroup()

# Re-nest by body region and generate figures and tables
cd4_boot2 <- cd4_boot2 %>%
# Fix site labels
mutate(body_site = str_replace_all(body_site,
                                   pattern = '_',
                                   replacement = ' '),
       body_site = str_replace_all(body_site,
                                   pattern = '\\\\.',
                                   replacement = ' & ')) %>%

# Group and nest
group_by(body_site) %>%
nest() %>%
# Arrange CD4 recent groups
# Plot data
mutate(plots = map2(.x = data,
                   .y = body_site,
                   ~ .x %>%
                     ggplot(data = .) +
                     aes(x = cd4_group,
                         y = point_est,
                         ymin = lower_ci,
                         ymax = upper_ci) +
                     geom_linerange(size = 1,
                                    colour = '#000000') +
                     geom_point(colour = '#000000',
                                size = 6) +
                     labs(title = .y,
                          subtitle = '(Point estimate with 95%CI)',
                          x = expression('CD4 group (cells.mm'^{-3*}')),
                          y = 'Proportion with pain') +
                     scale_y_continuous(limits = c(0, 1)) +

```

```

theme_minimal(base_size = 18) +
theme(plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))) %>%

# Tabulate data
mutate(
  tables = map2(.x = data,
                .y = body_site,
                ~ .x %>%
                  arrange(cd4_group) %>%
                  kable(caption = .y,
                        digits = 2)))

```

## 6.2 Tabulated proportions (with 95% CIs), by CD4 (recent) group and body site

```
walk(cd4_boot2$tables, ~ print(.x))
```

Table 36: Head

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.26      | 0.16     | 0.37     |
| 100-199   | 0.40      | 0.30     | 0.49     |
| 200-299   | 0.32      | 0.24     | 0.42     |
| 300-399   | 0.36      | 0.25     | 0.48     |
| 400-499   | 0.30      | 0.17     | 0.43     |
| 500+      | 0.25      | 0.16     | 0.34     |

Table 37: Throat

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.07      | 0.01     | 0.15     |
| 100-199   | 0.02      | 0.00     | 0.05     |
| 200-299   | 0.03      | 0.00     | 0.06     |
| 300-399   | 0.01      | 0.00     | 0.04     |
| 400-499   | 0.04      | 0.00     | 0.11     |
| 500+      | 0.00      | 0.00     | 0.00     |

Table 38: Shoulder

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.07      | 0.01     | 0.15     |
| 100-199   | 0.10      | 0.05     | 0.17     |
| 200-299   | 0.06      | 0.02     | 0.11     |
| 300-399   | 0.05      | 0.01     | 0.10     |
| 400-499   | 0.15      | 0.04     | 0.26     |

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 500+      | 0.08      | 0.02     | 0.13     |

Table 39: Arms

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.06      | 0.01     | 0.12     |
| 100-199   | 0.08      | 0.03     | 0.14     |
| 200-299   | 0.00      | 0.00     | 0.00     |
| 300-399   | 0.04      | 0.00     | 0.09     |
| 400-499   | 0.02      | 0.00     | 0.07     |
| 500+      | 0.05      | 0.01     | 0.11     |

Table 40: Elbows

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.01      | 0.00     | 0.04     |
| 100-199   | 0.05      | 0.01     | 0.09     |
| 200-299   | 0.02      | 0.00     | 0.06     |
| 300-399   | 0.04      | 0.00     | 0.08     |
| 400-499   | 0.07      | 0.00     | 0.15     |
| 500+      | 0.07      | 0.02     | 0.12     |

Table 41: Wrists & Hands

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.01      | 0.00     | 0.04     |
| 100-199   | 0.06      | 0.02     | 0.10     |
| 200-299   | 0.06      | 0.02     | 0.10     |
| 300-399   | 0.06      | 0.01     | 0.12     |
| 400-499   | 0.07      | 0.00     | 0.13     |
| 500+      | 0.14      | 0.08     | 0.22     |

Table 42: Chest

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.34      | 0.22     | 0.46     |
| 100-199   | 0.23      | 0.15     | 0.31     |
| 200-299   | 0.17      | 0.10     | 0.23     |
| 300-399   | 0.17      | 0.09     | 0.26     |
| 400-499   | 0.17      | 0.09     | 0.28     |
| 500+      | 0.11      | 0.05     | 0.17     |



Table 43: Lower back

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.06      | 0.01     | 0.12     |
| 100-199   | 0.08      | 0.04     | 0.13     |
| 200-299   | 0.06      | 0.02     | 0.10     |
| 300-399   | 0.06      | 0.01     | 0.13     |
| 400-499   | 0.11      | 0.02     | 0.20     |
| 500+      | 0.10      | 0.04     | 0.16     |

Table 44: Abdomen

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.34      | 0.22     | 0.46     |
| 100-199   | 0.29      | 0.21     | 0.38     |
| 200-299   | 0.27      | 0.19     | 0.36     |
| 300-399   | 0.23      | 0.14     | 0.32     |
| 400-499   | 0.30      | 0.17     | 0.46     |
| 500+      | 0.22      | 0.13     | 0.30     |

Table 45: Cervical spine

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.04      | 0.00     | 0.10     |
| 100-199   | 0.09      | 0.04     | 0.15     |
| 200-299   | 0.02      | 0.00     | 0.05     |
| 300-399   | 0.06      | 0.01     | 0.13     |
| 400-499   | 0.09      | 0.02     | 0.17     |
| 500+      | 0.03      | 0.00     | 0.08     |

Table 46: Thoracic spine

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.13      | 0.06     | 0.22     |
| 100-199   | 0.11      | 0.06     | 0.18     |
| 200-299   | 0.10      | 0.05     | 0.16     |
| 300-399   | 0.16      | 0.08     | 0.25     |
| 400-499   | 0.22      | 0.11     | 0.35     |
| 500+      | 0.14      | 0.08     | 0.22     |

Table 47: Lumbosacral spine

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.09      | 0.03     | 0.16     |
| 100-199   | 0.14      | 0.08     | 0.22     |
| 200-299   | 0.15      | 0.08     | 0.22     |

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 300-399   | 0.18      | 0.10     | 0.29     |
| 400-499   | 0.30      | 0.17     | 0.43     |
| 500+      | 0.21      | 0.13     | 0.28     |

Table 48: Groin

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.10      | 0.03     | 0.18     |
| 100-199   | 0.12      | 0.07     | 0.19     |
| 200-299   | 0.09      | 0.05     | 0.16     |
| 300-399   | 0.12      | 0.05     | 0.19     |
| 400-499   | 0.11      | 0.02     | 0.20     |
| 500+      | 0.08      | 0.02     | 0.13     |

Table 49: Hips

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.01      | 0.00     | 0.06     |
| 100-199   | 0.06      | 0.02     | 0.10     |
| 200-299   | 0.05      | 0.01     | 0.09     |
| 300-399   | 0.05      | 0.01     | 0.10     |
| 400-499   | 0.11      | 0.02     | 0.20     |
| 500+      | 0.14      | 0.08     | 0.22     |

Table 50: Legs

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.13      | 0.06     | 0.22     |
| 100-199   | 0.11      | 0.06     | 0.18     |
| 200-299   | 0.15      | 0.08     | 0.21     |
| 300-399   | 0.18      | 0.10     | 0.27     |
| 400-499   | 0.20      | 0.09     | 0.30     |
| 500+      | 0.23      | 0.15     | 0.32     |

Table 51: Knees

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.12      | 0.04     | 0.19     |
| 100-199   | 0.11      | 0.06     | 0.18     |
| 200-299   | 0.15      | 0.08     | 0.21     |
| 300-399   | 0.12      | 0.05     | 0.19     |
| 400-499   | 0.17      | 0.07     | 0.28     |
| 500+      | 0.22      | 0.14     | 0.30     |

Table 52: Ankles &amp; Feet

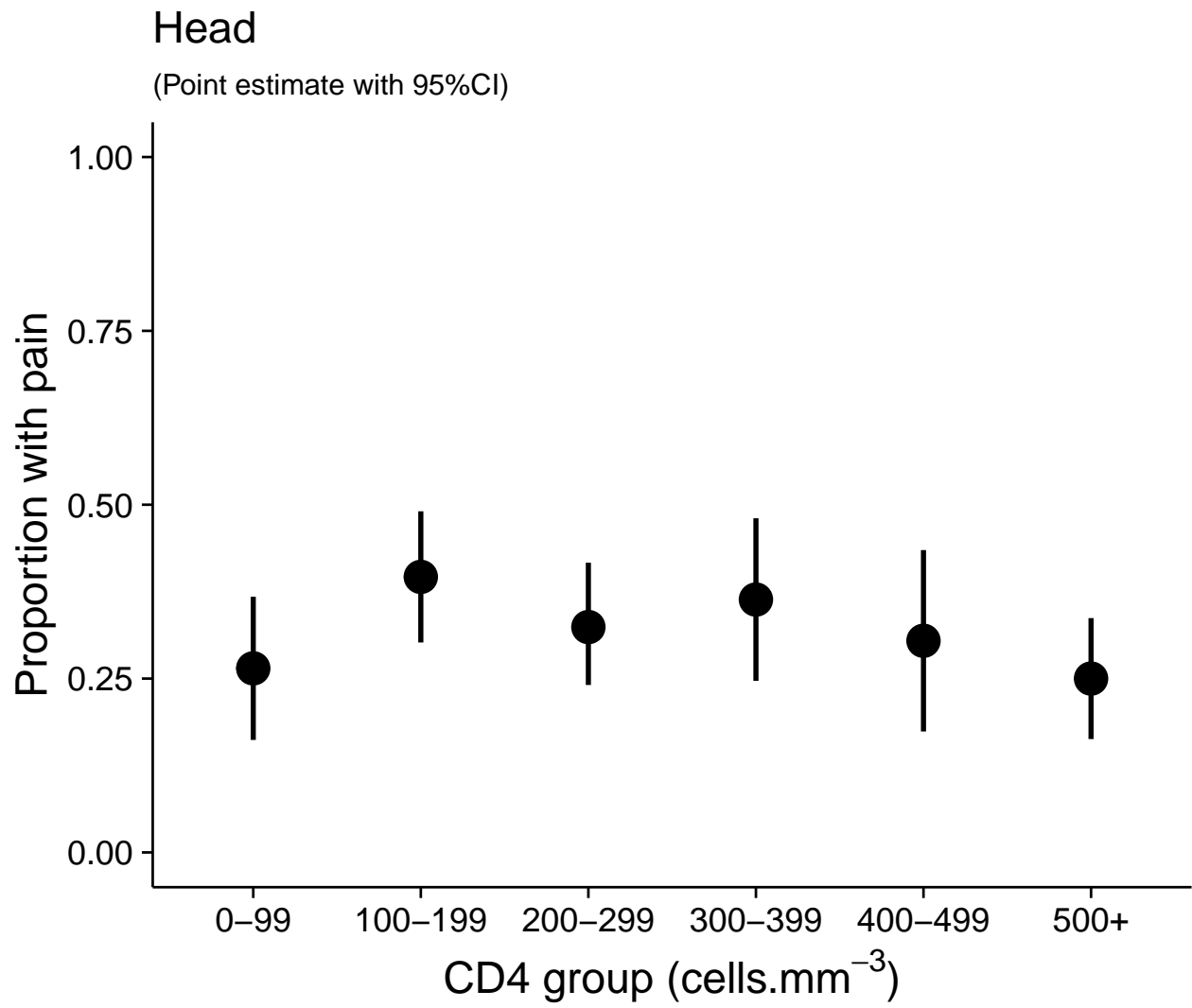
| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.26      | 0.16     | 0.37     |
| 100-199   | 0.30      | 0.22     | 0.40     |
| 200-299   | 0.31      | 0.23     | 0.41     |
| 300-399   | 0.25      | 0.16     | 0.34     |
| 400-499   | 0.26      | 0.15     | 0.39     |
| 500+      | 0.35      | 0.25     | 0.45     |

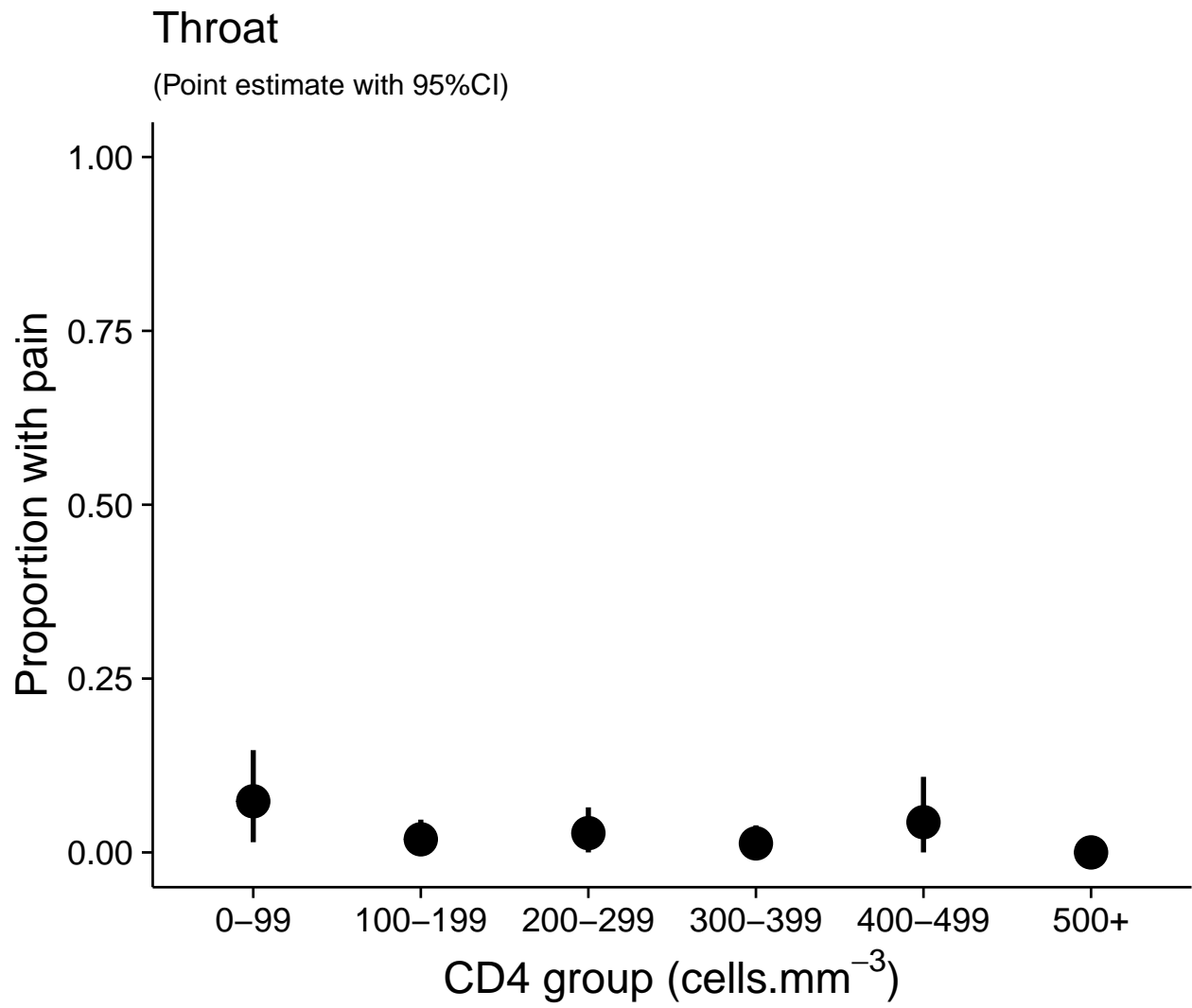
Table 53: Buttocks

| cd4_group | point_est | lower_ci | upper_ci |
|-----------|-----------|----------|----------|
| 0-99      | 0.04      | 0.00     | 0.10     |
| 100-199   | 0.01      | 0.00     | 0.03     |
| 200-299   | 0.03      | 0.00     | 0.06     |
| 300-399   | 0.05      | 0.01     | 0.10     |
| 400-499   | 0.04      | 0.00     | 0.11     |
| 500+      | 0.07      | 0.02     | 0.12     |

### 6.3 Plotted proportions (with 95% CIs), by CD4 (recent) group and body site

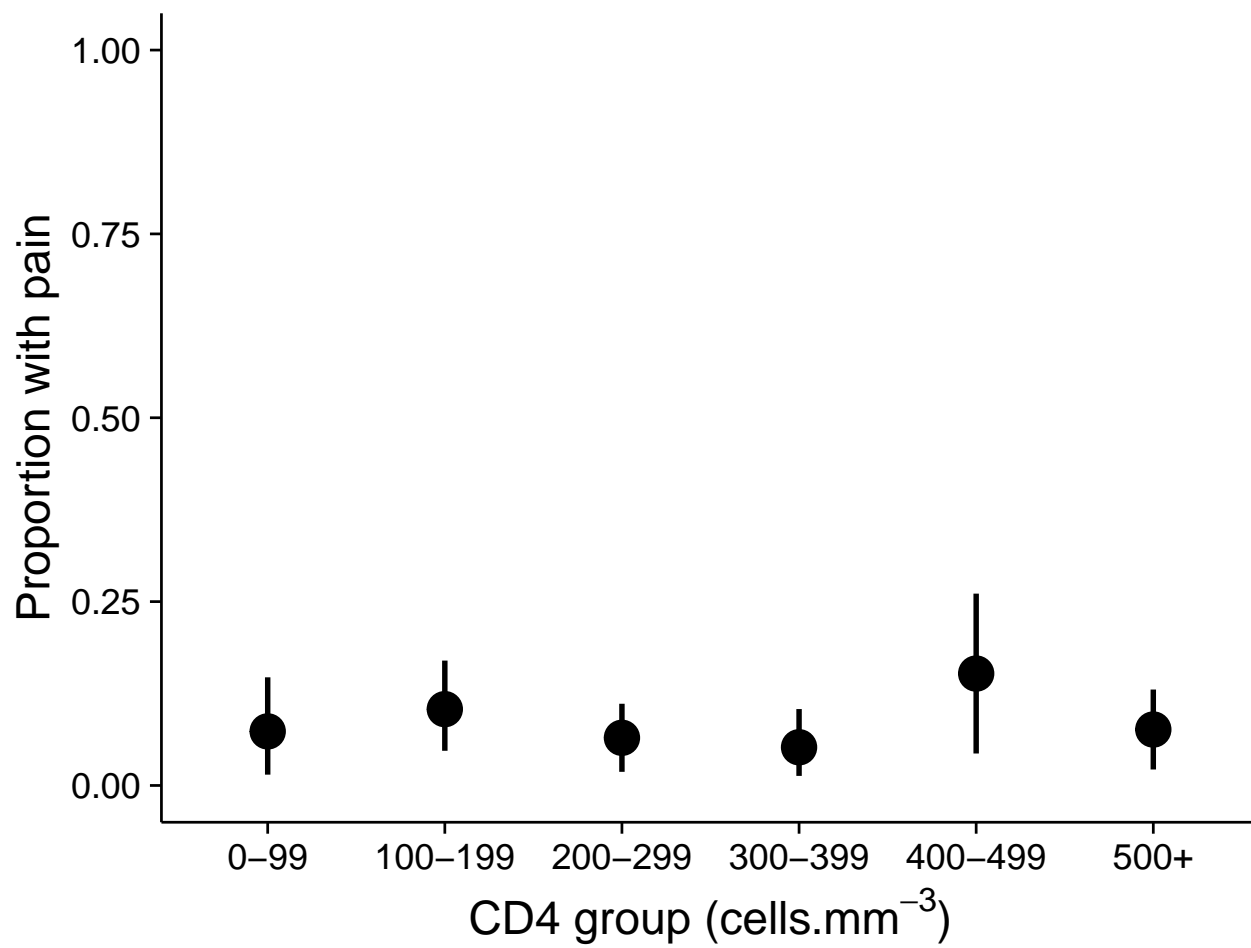
```
walk(cd4_boot2$plots, ~ print(.x))
```

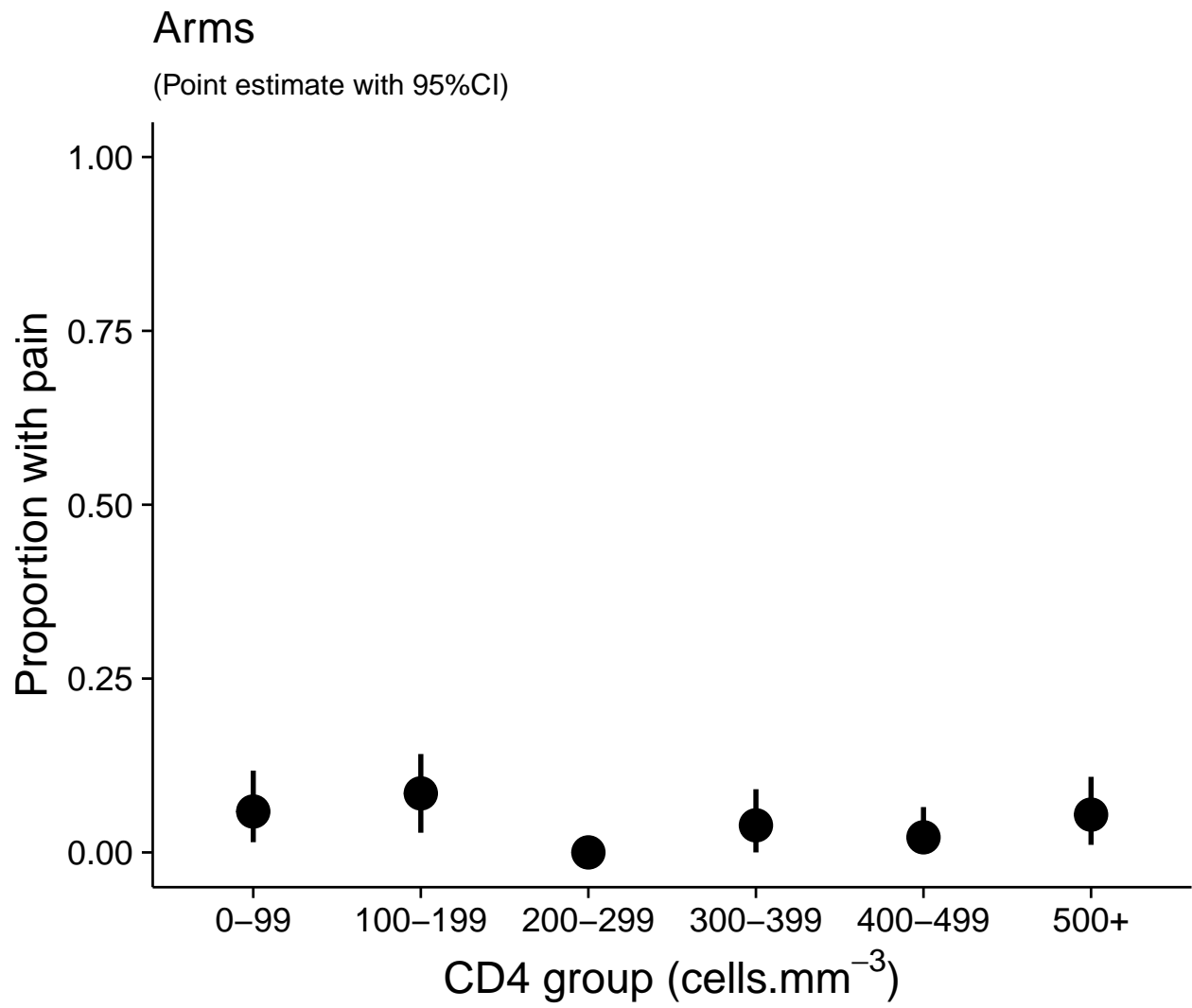




## Shoulder

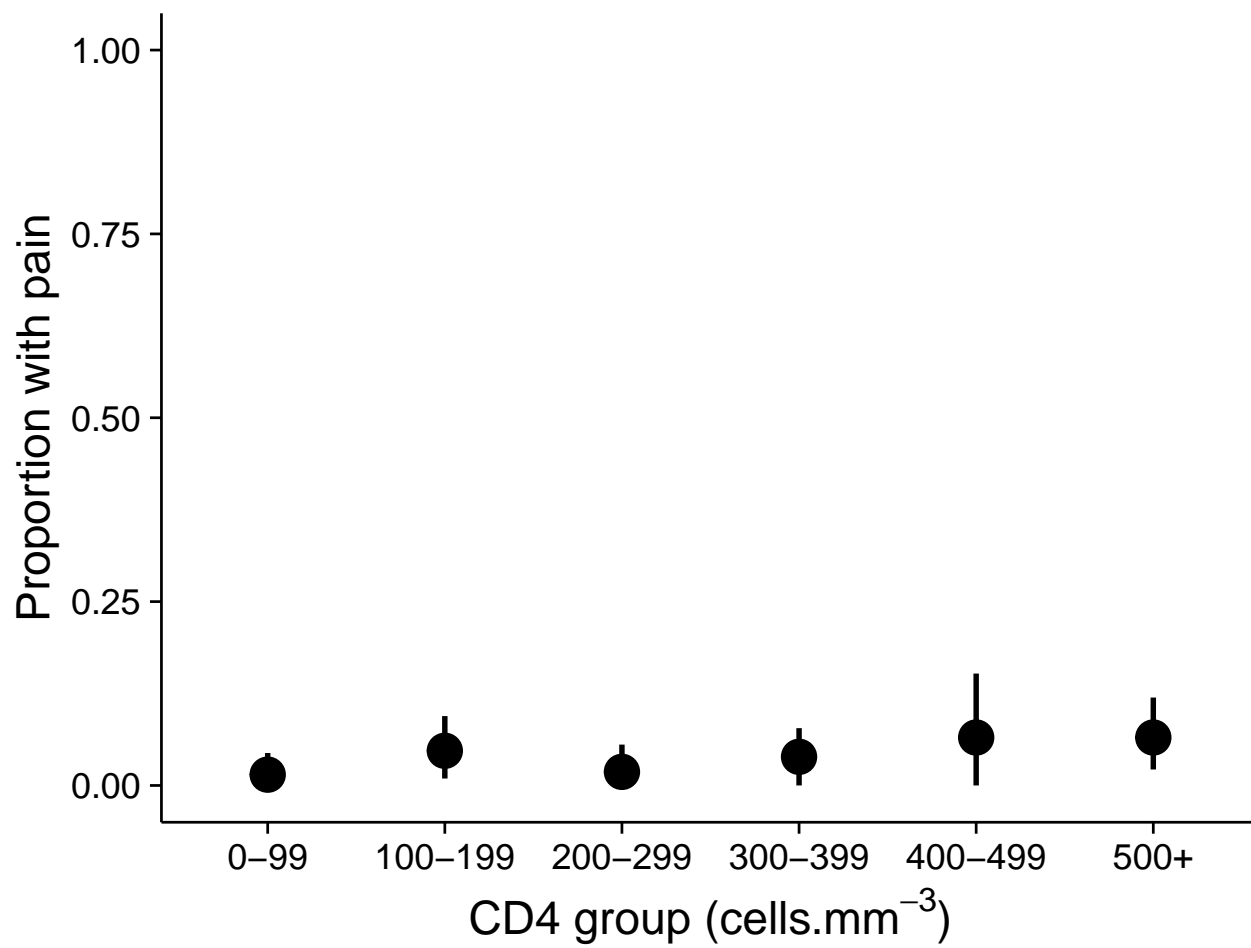
(Point estimate with 95%CI)





## Elbows

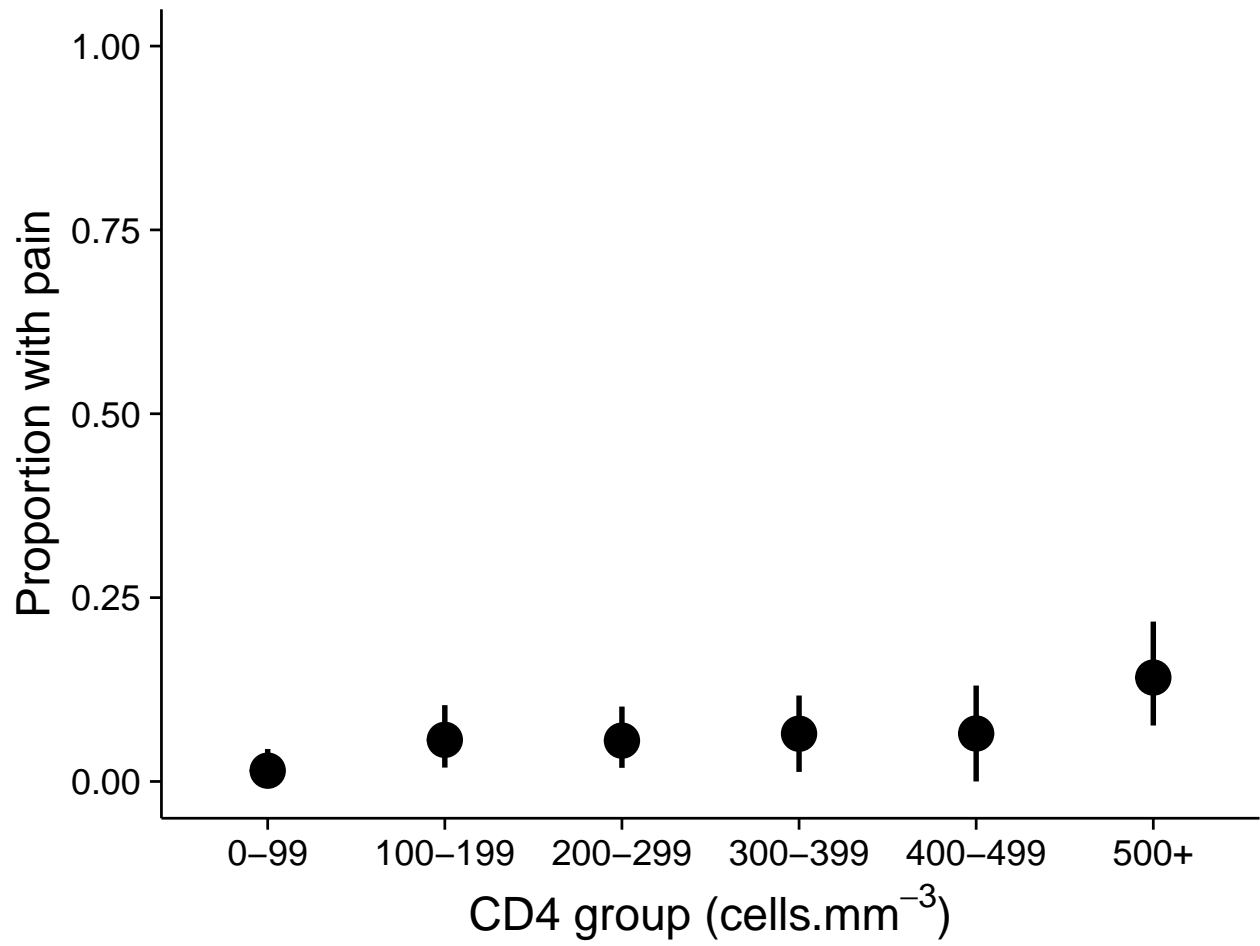
(Point estimate with 95%CI)

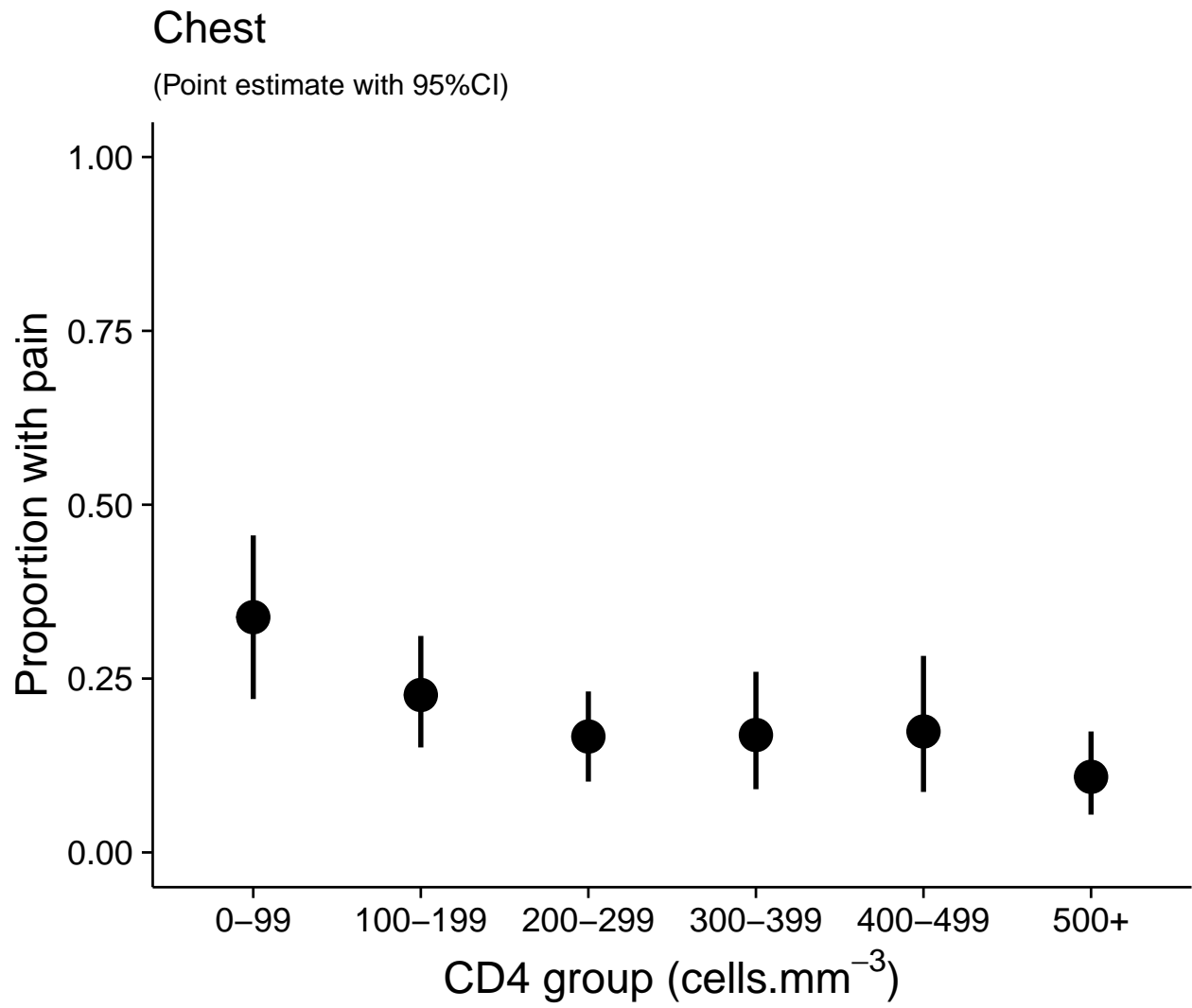


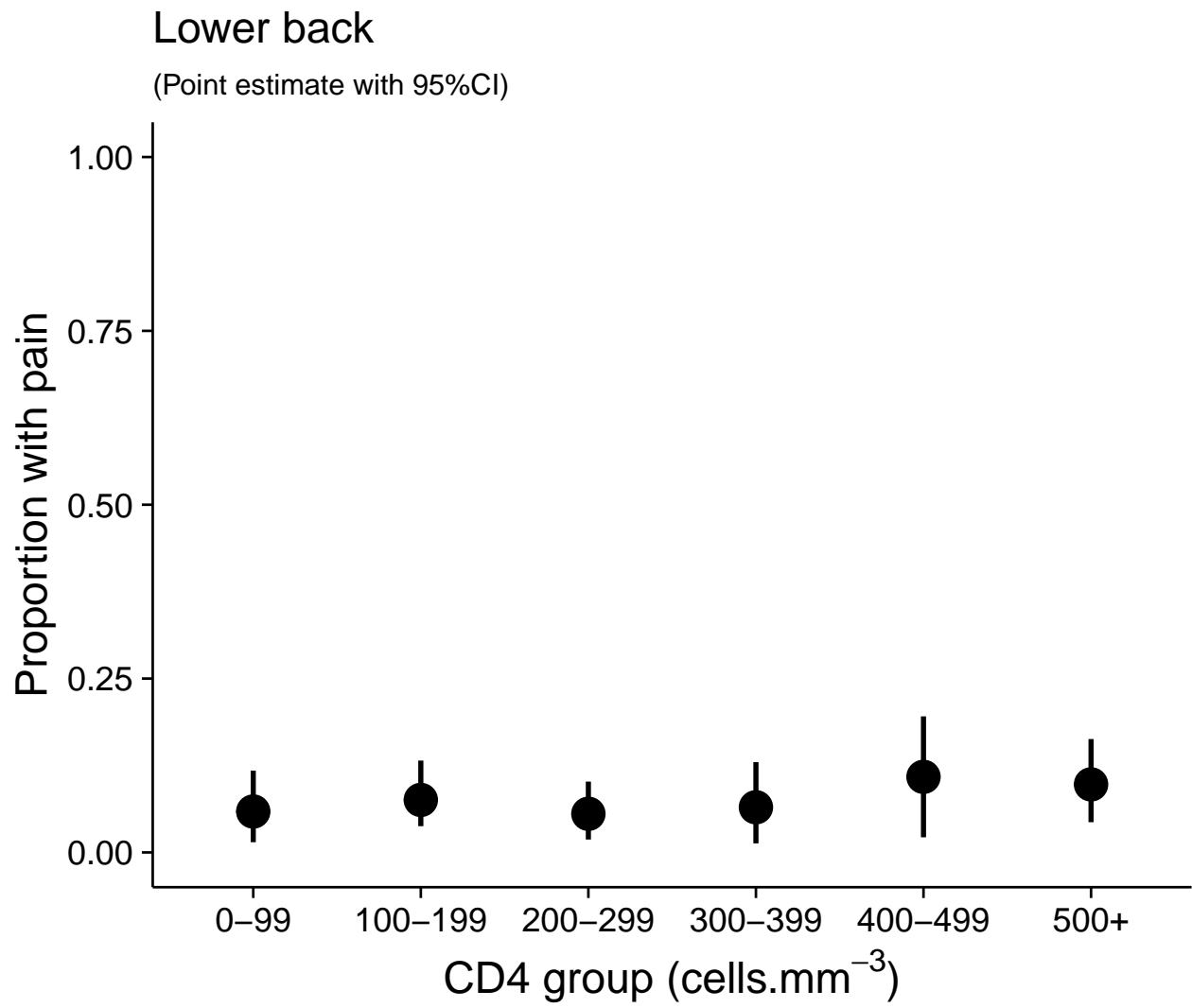


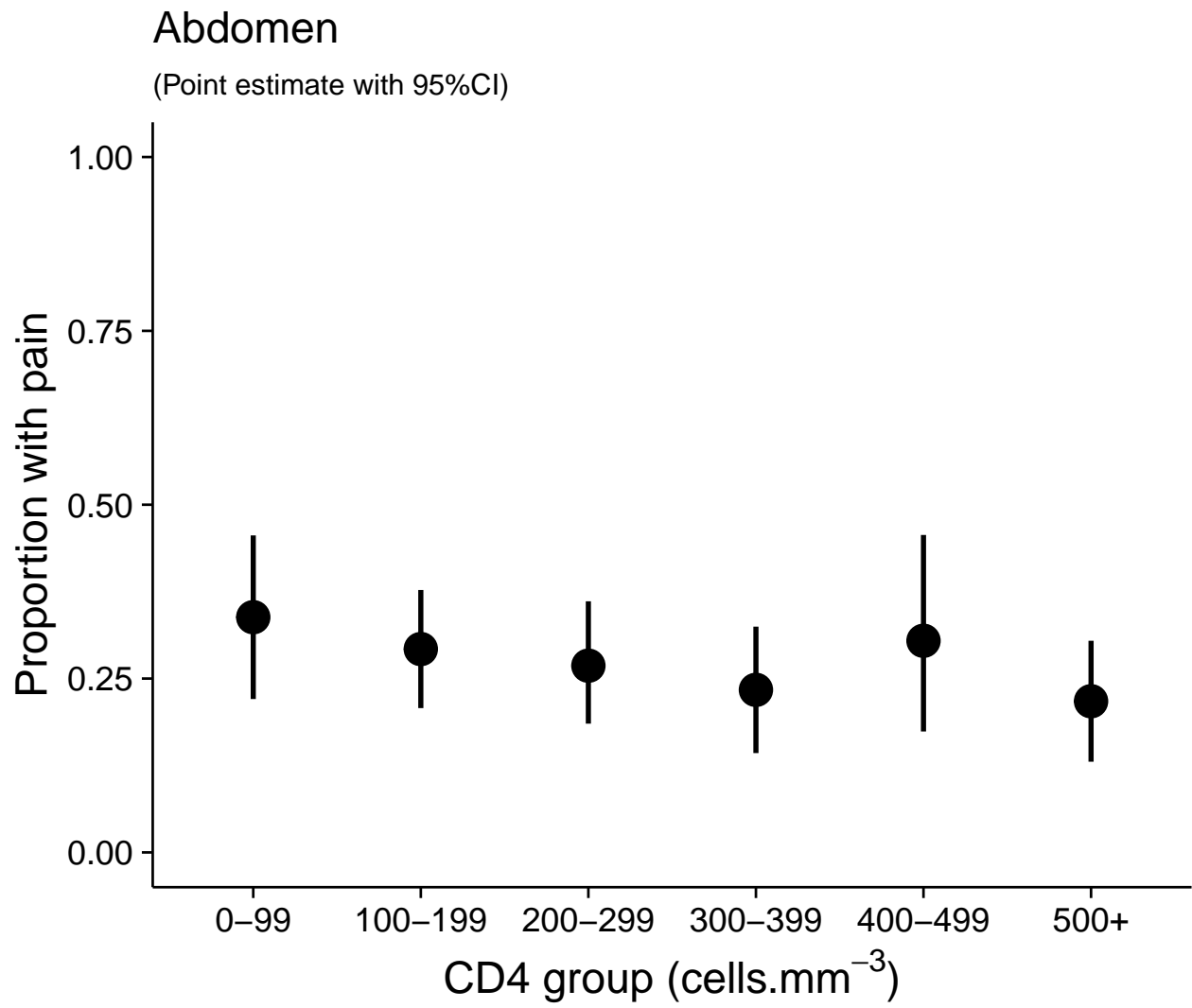
## Wrists & Hands

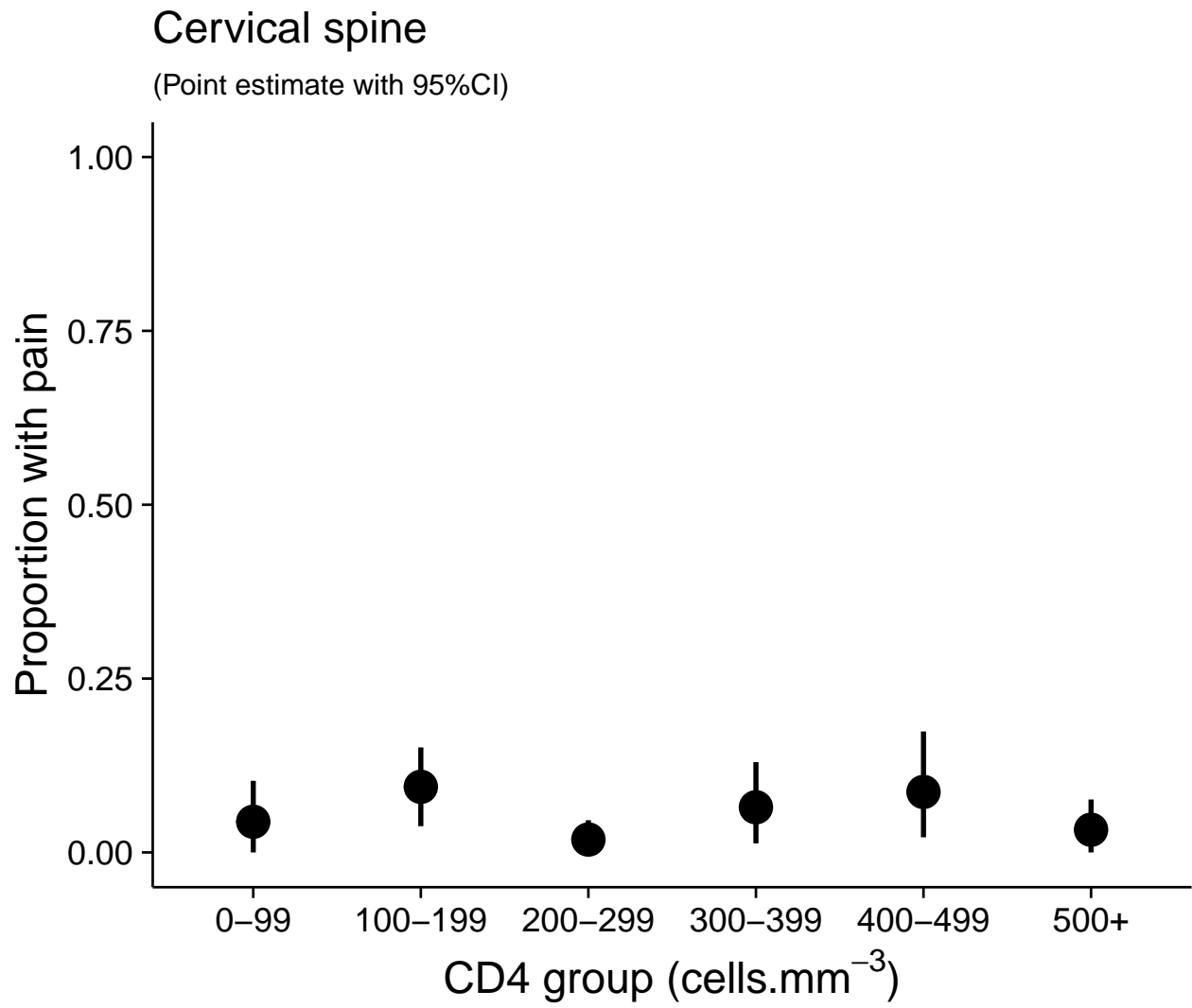
(Point estimate with 95%CI)

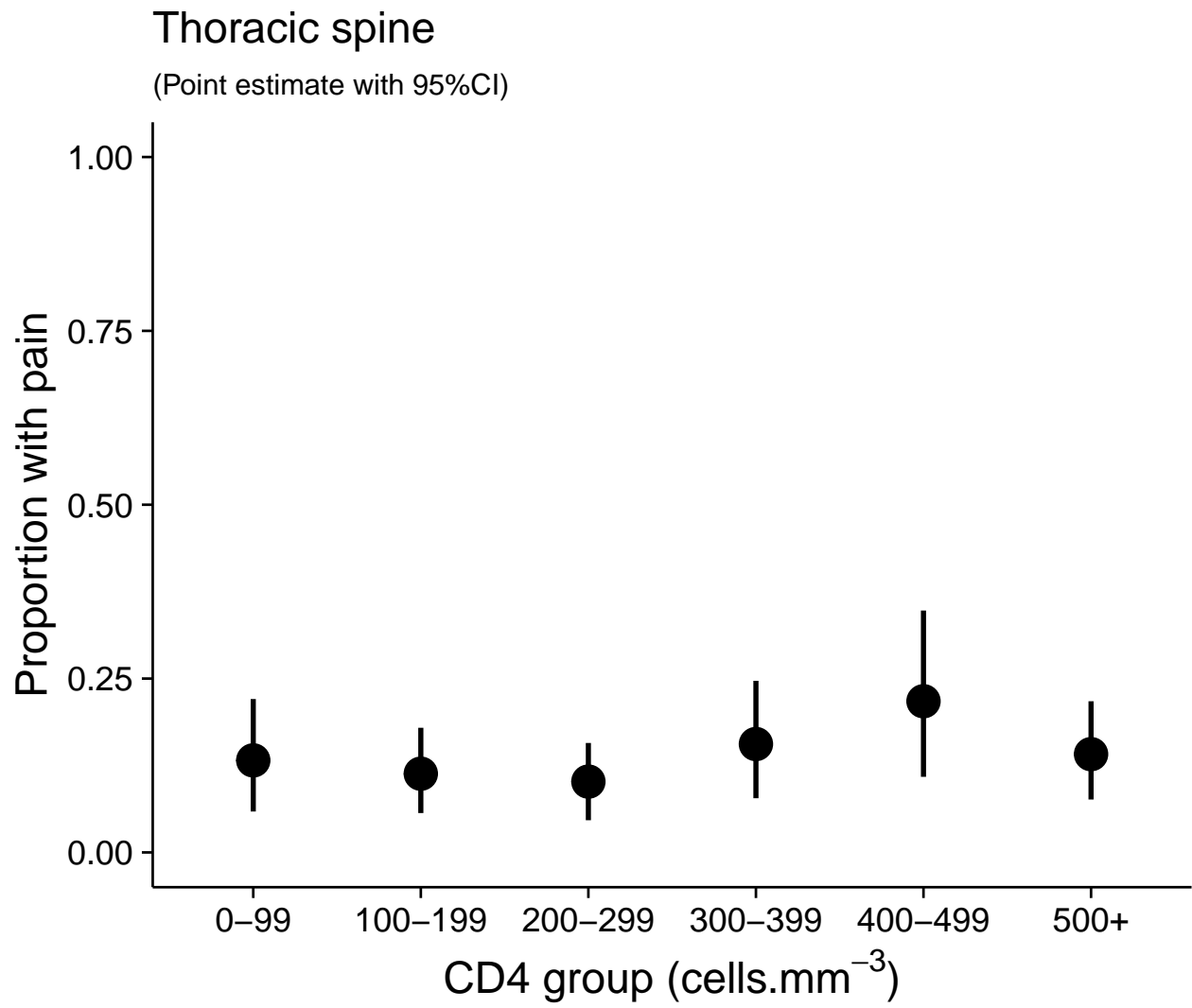


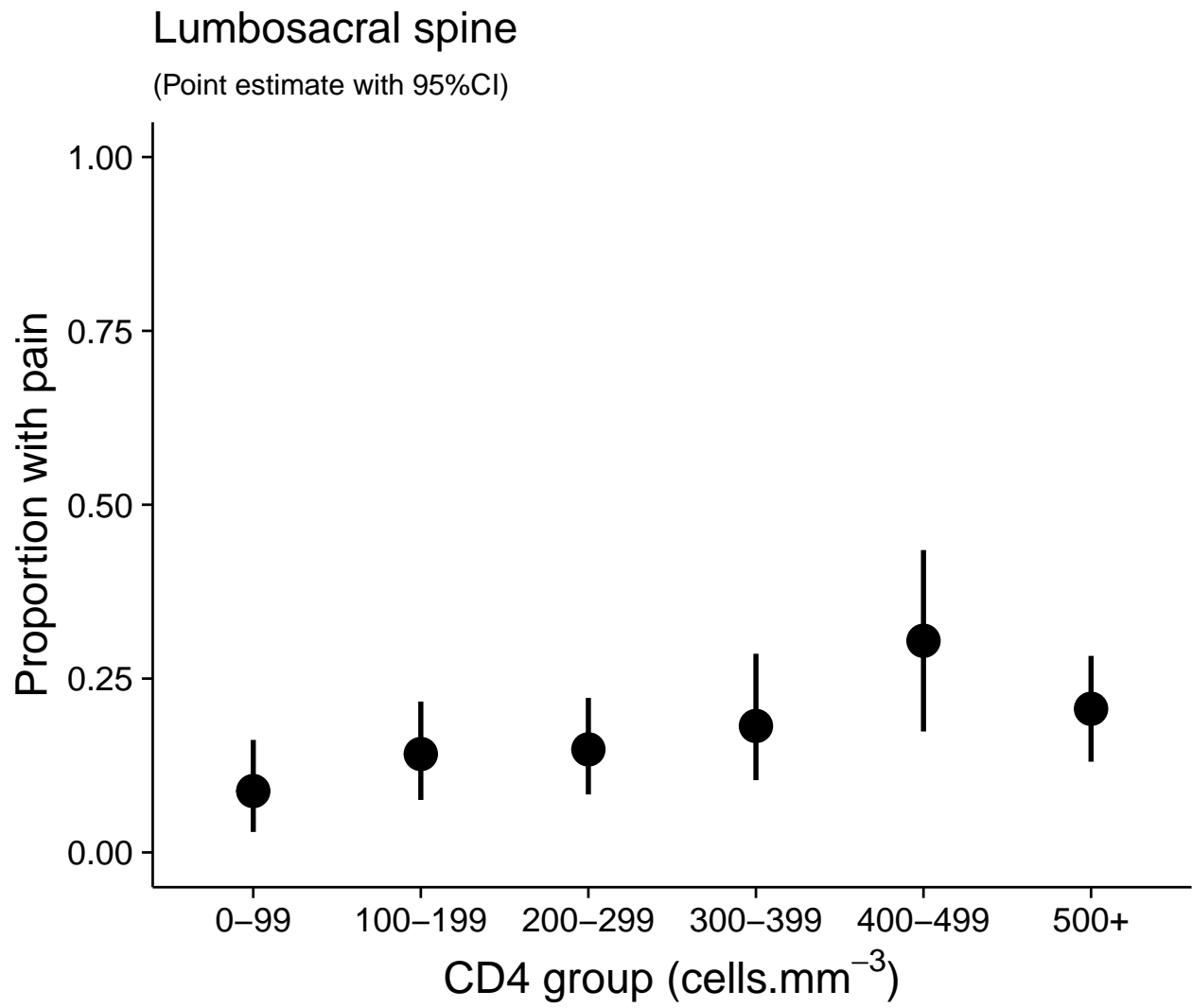


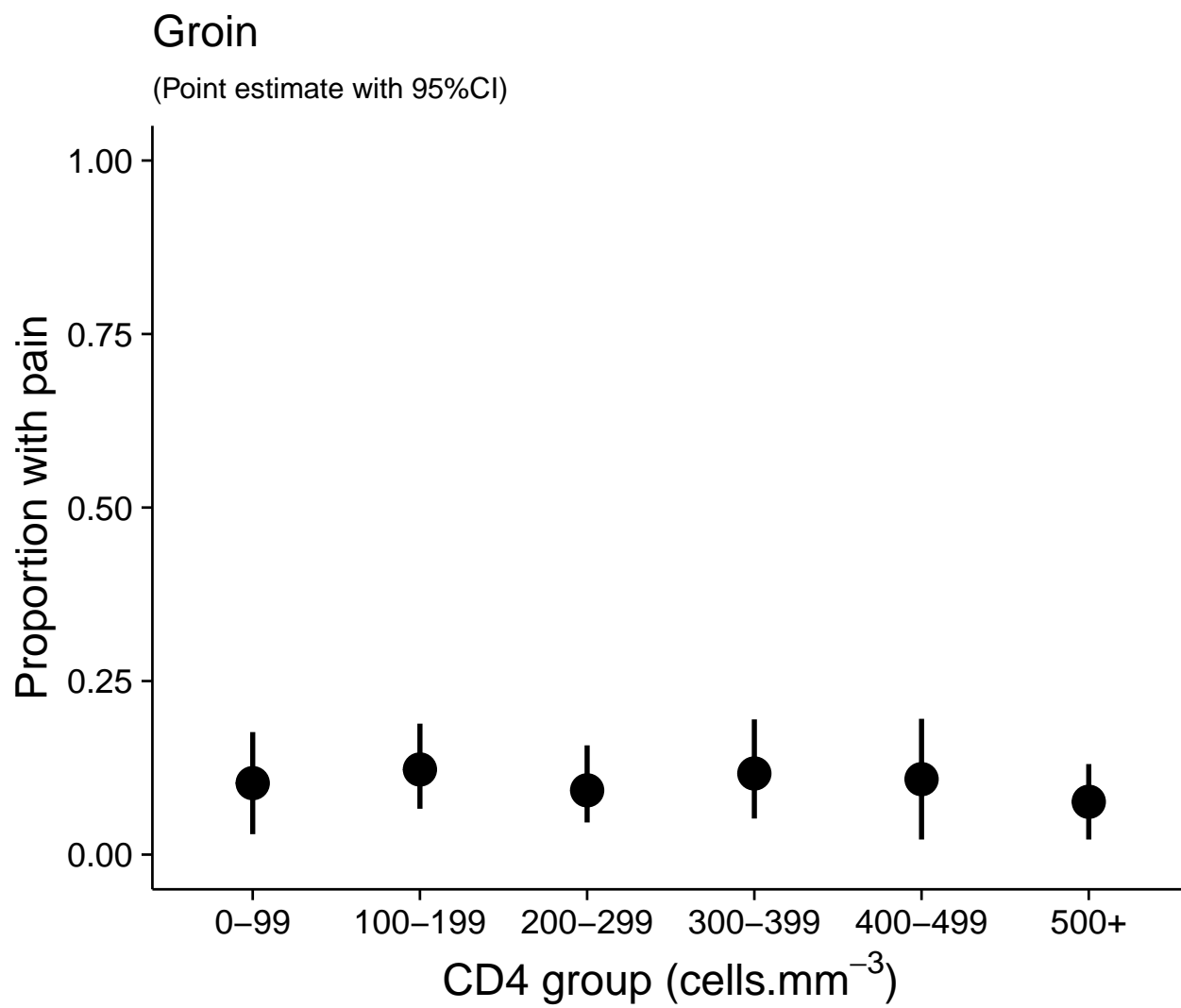




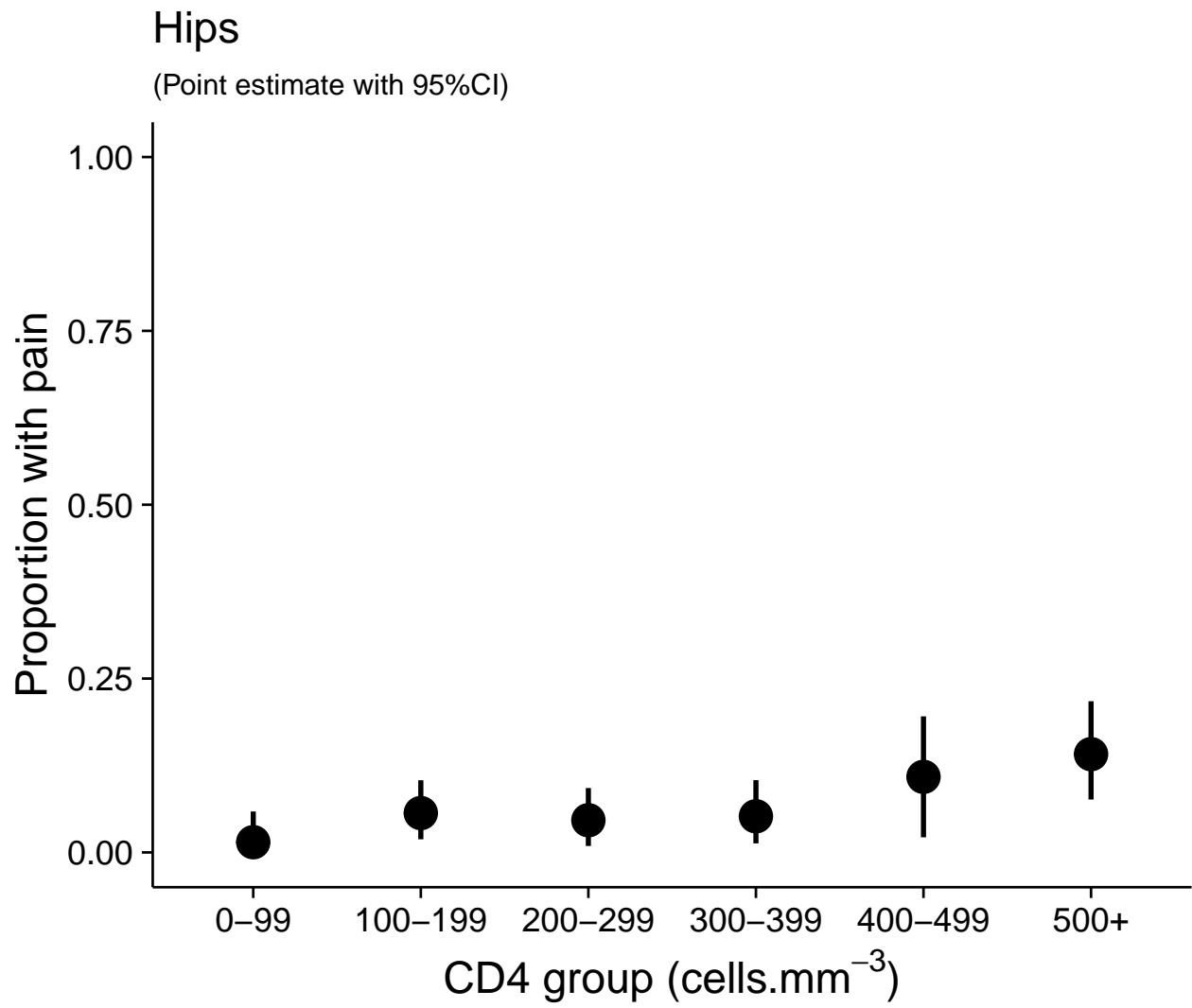


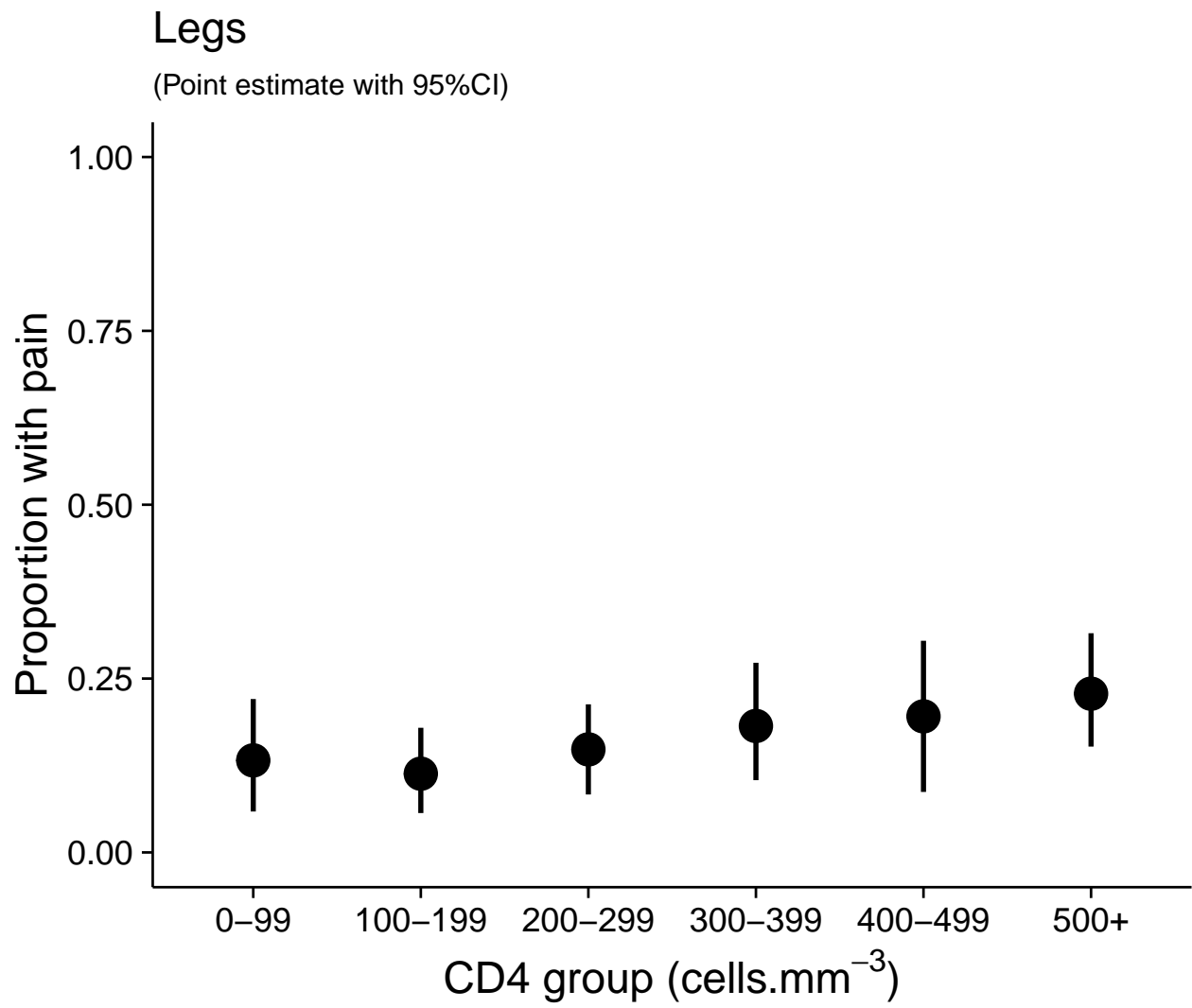






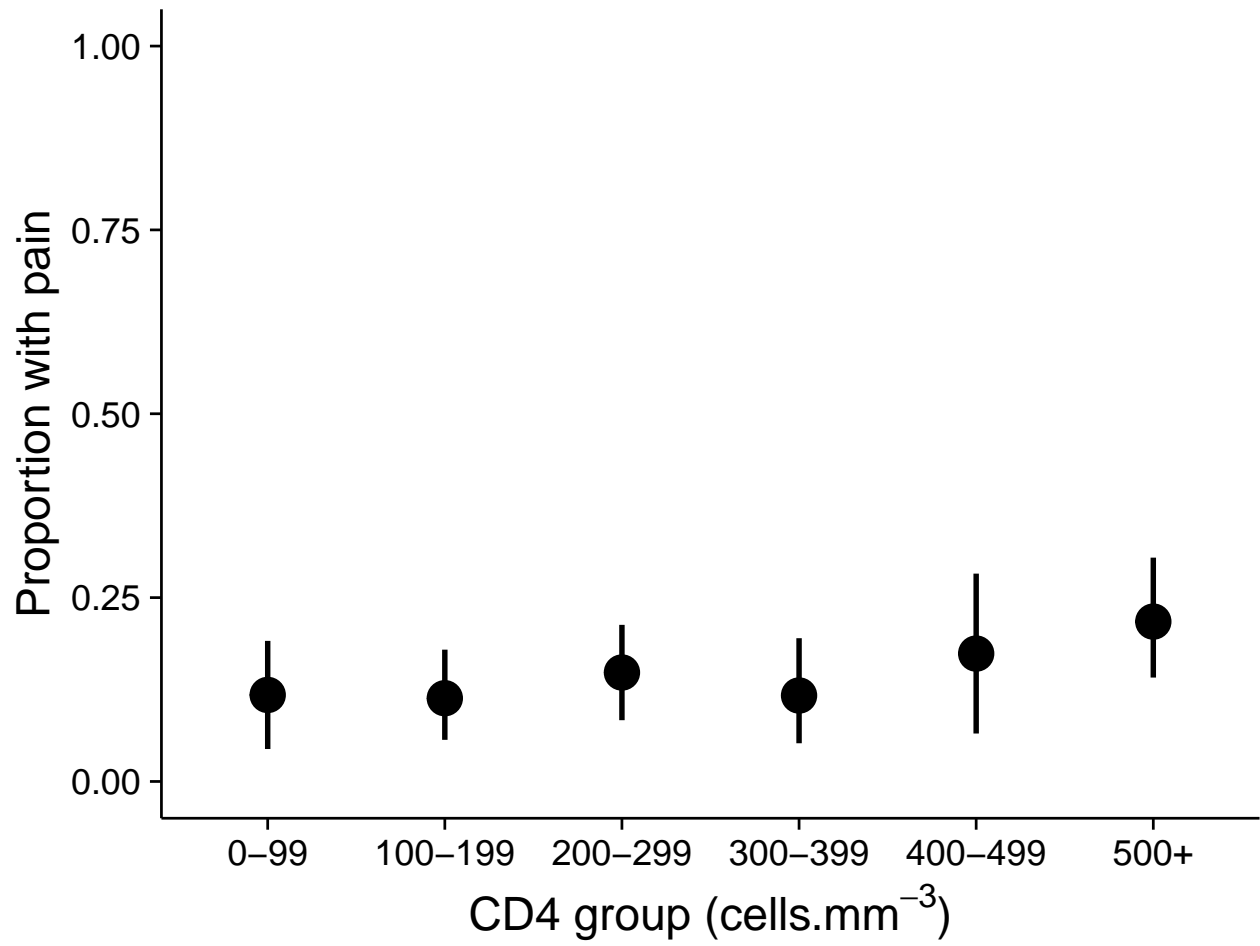


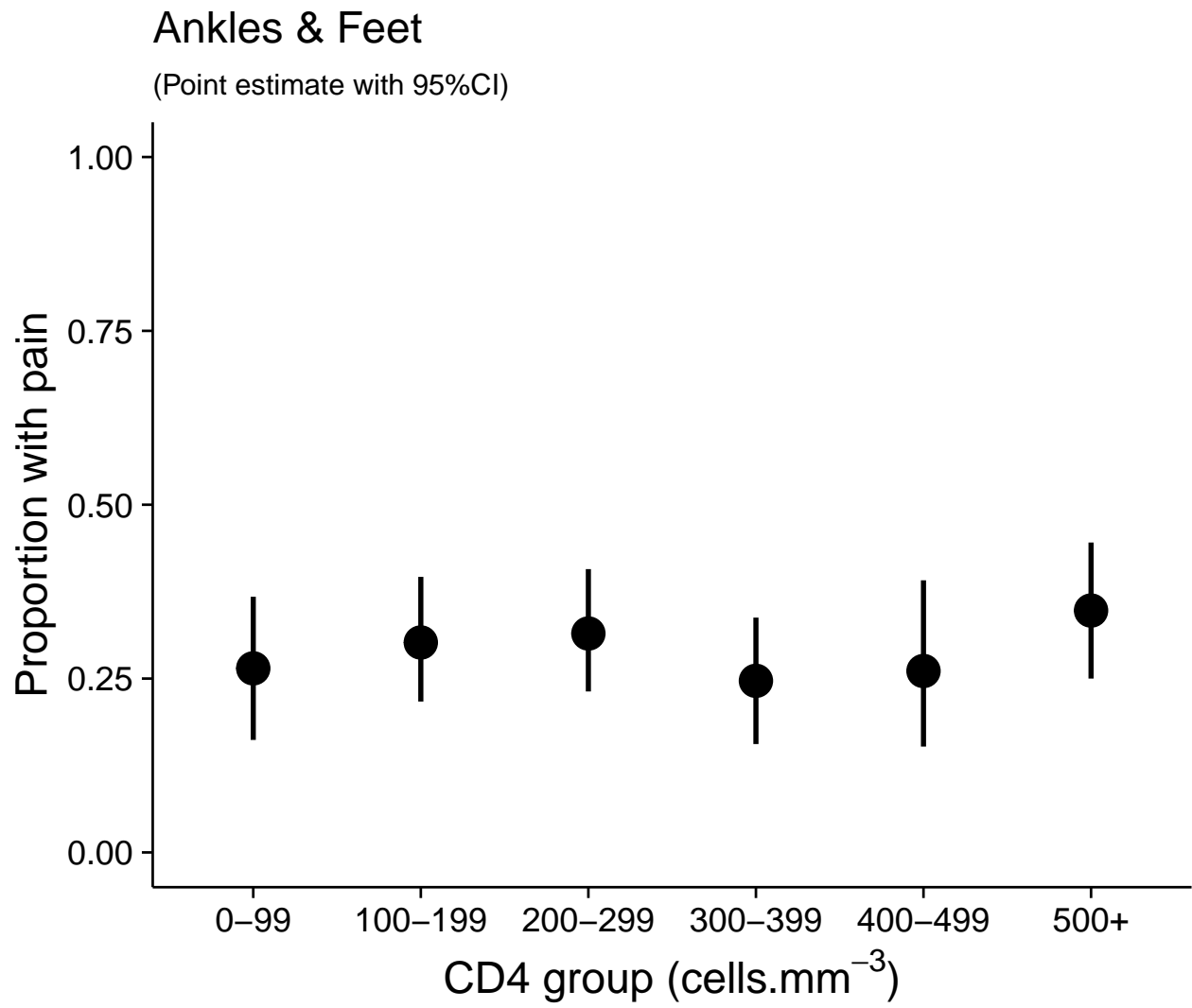


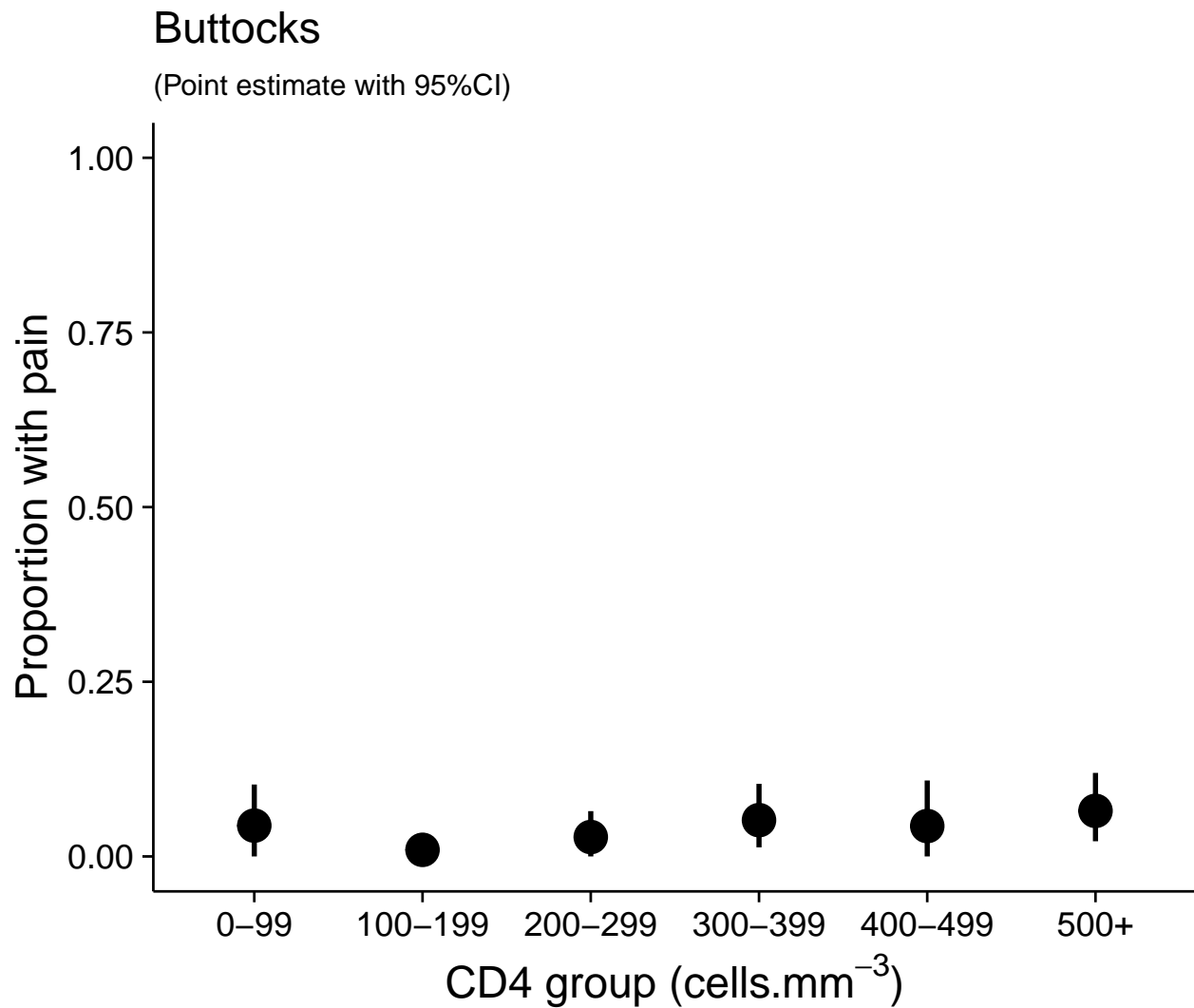


## Knees

(Point estimate with 95%CI)







## 7 By pain definition

Pain in the last week or month (acute) versus pain for the last 3 months (chronic)

### 7.1 Process data

```
# Set seed
set.seed(2020)

# Select pain data
pain <- demo[, c('ID', 'Pain')]

# Join to boot_data & remove ID, site, and upper_back (only one outcome -- no pain)
pain <- left_join(data, pain) %>%
```

```

select(-ID, -Site, -Upper_back)

# Bootstrap CIs
pain_boot <- pain %>%
  # Pivot to long format
  pivot_longer(cols = -Pain,
               names_to = 'body_site',
               values_to = 'pain_present') %>%
  # Add body regions
  mutate(region = case_when(
    body_site == 'Chest' |
    body_site == 'Head' |
    body_site == 'Throat' |
    body_site == 'Shoulder' ~ 'Head and upper torso',
    body_site == 'Lower_back' |
    body_site == 'Abdomen' |
    body_site == 'Hips' |
    body_site == 'Buttocks' |
    body_site == 'Groin' ~ 'Lower torso',
    body_site == 'Legs' |
    body_site == 'Knees' |
    body_site == 'Ankles.Feet' ~ 'Lower limbs',
    body_site == 'Arms' |
    body_site == 'Elbows' |
    body_site == 'Wrists.Hands' ~ 'Upper limbs',
    body_site == 'Cervical_spine' |
    body_site == 'Thoracic_spine' |
    body_site == 'Lumbosacral_spine' ~ 'Spinal column',
    TRUE ~ 'other'
  )) %>%
  # Nest by body region and body site
  group_by(Pain, region, body_site) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,
                    ~ boot(data = .x,
                           statistic = prop_func,
                           R = 999,
                           stype = 'i',
                           parallel = 'multicore',
                           ncpus = 4))) %>%
  # Get CI
  mutate(ci = map(.x = boot,
                  ~ boot.ci(.x, type = 'perc'))) %>%
  # Extract ci data
  mutate(point_est = map(.x = ci,
                          ~ .x$t0),
         lower_ci = map(.x = ci,
                          ~ .x$percent[[4]]),
         upper_ci = map(.x = ci,
                          ~ .x$percent[[5]])) %>%
  # Remove columns
  select(-data, -boot, -ci) %>%

```

```

# Unnest
unnest(cols = c(point_est, lower_ci, upper_ci))

# Re-nest by body region and generate figures and tables
pain_boot2 <- pain_boot %>%
  group_by(region) %>%
  nest() %>%
  # Fix site labels
  mutate(data = map(.x = data,
                    ~ .x %>%
                      mutate(body_site = str_replace_all(body_site,
                                                            pattern = '_',
                                                            replacement = ' '),
                            body_site = str_replace_all(body_site,
                                                            pattern = '\\.',
                                                            replacement = ' & ')))) %>%

# Re-order sites by point_est
mutate(data = map(.x = data,
                  ~ .x %>%
                    mutate(body_site = fct_reorder(body_site,
                                                    point_est)))) %>%

# Plot data
mutate(plots = map2(.x = data,
                   .y = region,
                   ~ .x %>%
                     ggplot(data = .) +
                     aes(x = body_site,
                         y = point_est,
                         ymin = lower_ci,
                         ymax = upper_ci,
                         fill = Pain) +
                     geom_linerange(position = position_dodge2(width = 0.6),
                                     size = 1,
                                     colour = '#000000') +
                     geom_point(shape = 21,
                                colour = '#000000',
                                position = position_dodge2(width = 0.6),
                                size = 6,
                                stroke = 1) +
                     coord_flip() +
                     labs(title = .y,
                          subtitle = '(Point estimate with 95%CI)',
                          y = 'Proportion with pain') +
                     scale_y_continuous(limits = c(0, 1)) +
                     scale_fill_manual(values = c('#000000', '#FFFFFF')) +
                     theme_minimal(base_size = 18) +
                     theme(plot.title = element_text(size = 18),
                           plot.subtitle = element_text(size = 12),
                           legend.title = element_blank(),
                           legend.position = 'top',
                           axis.title.y = element_blank(),
                           panel.grid = element_blank(),
                           axis.text = element_text(colour = '#000000')),

```

```

axis.line = element_line(size = 0.5),
axis.ticks = element_line(size = 0.5))) %>%

# Tabulate data
mutate(tables = map2(.x = data,
  .y = region,
  ~ .x %>%
    kable(caption = .y,
          digits = 2)))

```

## 7.2 Tabulated proportions (with 95% CIs), by age and body region

```
walk(pain_boot2$tables, ~ print(.x))
```

Table 54: Head and upper torso

| Pain    | body_site | point_est | lower_ci | upper_ci |
|---------|-----------|-----------|----------|----------|
| Acute   | Head      | 0.38      | 0.33     | 0.43     |
| Acute   | Throat    | 0.03      | 0.01     | 0.05     |
| Acute   | Shoulder  | 0.08      | 0.05     | 0.11     |
| Acute   | Chest     | 0.24      | 0.20     | 0.28     |
| Chronic | Head      | 0.23      | 0.18     | 0.30     |
| Chronic | Throat    | 0.03      | 0.01     | 0.06     |
| Chronic | Shoulder  | 0.08      | 0.04     | 0.11     |
| Chronic | Chest     | 0.12      | 0.08     | 0.17     |

Table 55: Upper limbs

| Pain    | body_site      | point_est | lower_ci | upper_ci |
|---------|----------------|-----------|----------|----------|
| Acute   | Arms           | 0.05      | 0.03     | 0.07     |
| Acute   | Elbows         | 0.03      | 0.02     | 0.05     |
| Acute   | Wrists & Hands | 0.05      | 0.03     | 0.07     |
| Chronic | Arms           | 0.02      | 0.00     | 0.05     |
| Chronic | Elbows         | 0.05      | 0.02     | 0.08     |
| Chronic | Wrists & Hands | 0.08      | 0.04     | 0.11     |

Table 56: Lower torso

| Pain    | body_site  | point_est | lower_ci | upper_ci |
|---------|------------|-----------|----------|----------|
| Acute   | Lower back | 0.07      | 0.04     | 0.09     |
| Acute   | Abdomen    | 0.30      | 0.26     | 0.35     |
| Acute   | Groin      | 0.11      | 0.08     | 0.15     |
| Acute   | Hips       | 0.02      | 0.01     | 0.04     |
| Acute   | Buttocks   | 0.03      | 0.02     | 0.05     |
| Chronic | Lower back | 0.08      | 0.04     | 0.11     |
| Chronic | Abdomen    | 0.20      | 0.15     | 0.26     |
| Chronic | Groin      | 0.05      | 0.02     | 0.08     |
| Chronic | Hips       | 0.15      | 0.11     | 0.20     |



| Pain    | body_site | point_est | lower_ci | upper_ci |
|---------|-----------|-----------|----------|----------|
| Chronic | Buttocks  | 0.03      | 0.01     | 0.06     |

Table 57: Spinal column

| Pain    | body_site         | point_est | lower_ci | upper_ci |
|---------|-------------------|-----------|----------|----------|
| Acute   | Cervical spine    | 0.07      | 0.05     | 0.10     |
| Acute   | Thoracic spine    | 0.14      | 0.11     | 0.18     |
| Acute   | Lumbosacral spine | 0.14      | 0.11     | 0.18     |
| Chronic | Cervical spine    | 0.01      | 0.00     | 0.02     |
| Chronic | Thoracic spine    | 0.10      | 0.06     | 0.13     |
| Chronic | Lumbosacral spine | 0.18      | 0.13     | 0.23     |

Table 58: Lower limbs

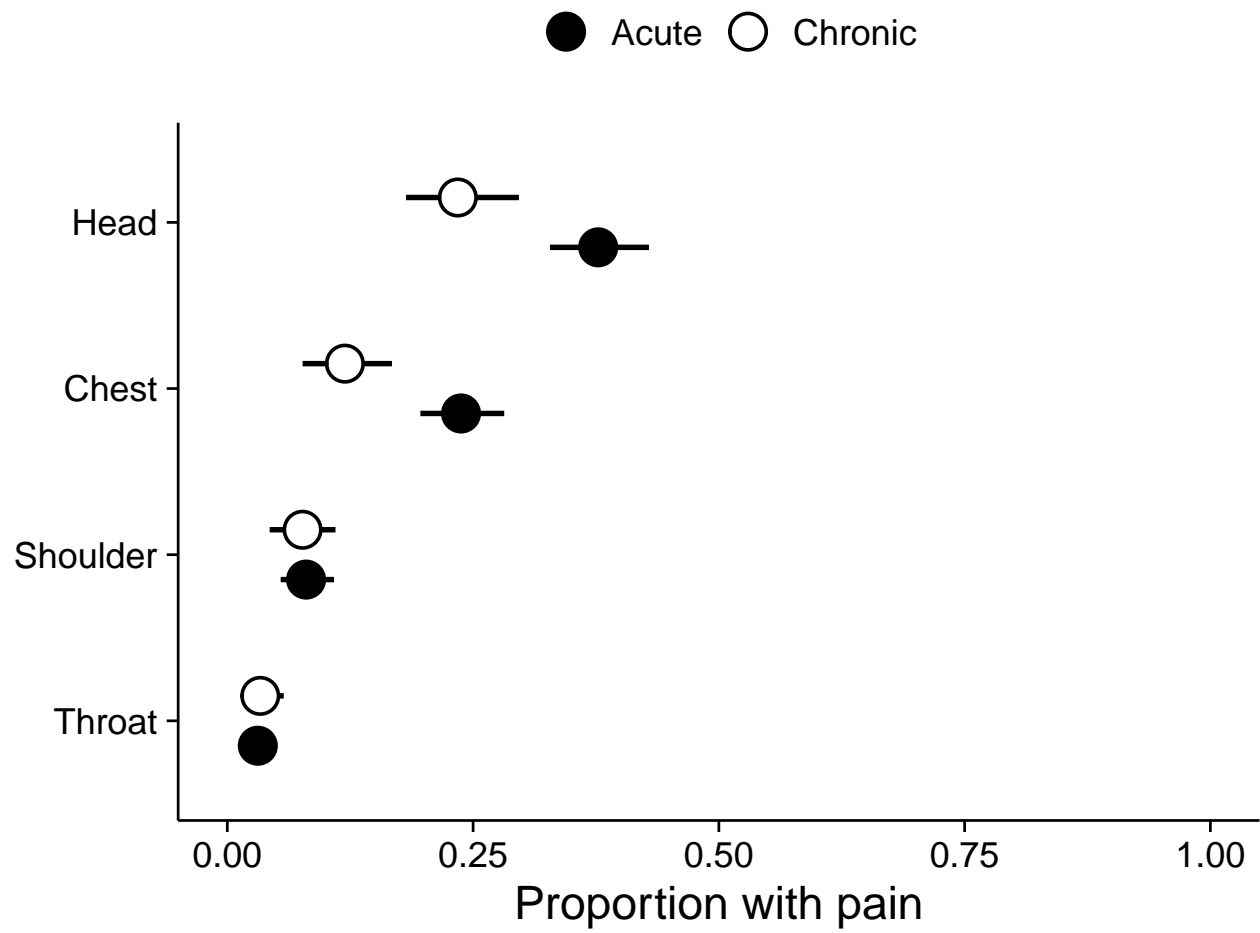
| Pain    | body_site     | point_est | lower_ci | upper_ci |
|---------|---------------|-----------|----------|----------|
| Acute   | Legs          | 0.12      | 0.09     | 0.15     |
| Acute   | Knees         | 0.10      | 0.07     | 0.13     |
| Acute   | Ankles & Feet | 0.23      | 0.19     | 0.27     |
| Chronic | Legs          | 0.25      | 0.19     | 0.32     |
| Chronic | Knees         | 0.22      | 0.16     | 0.27     |
| Chronic | Ankles & Feet | 0.46      | 0.39     | 0.53     |

### 7.3 Plotted proportions (with 95% CIs), by age and body region

```
walk(pain_boot2$plots, ~ print(.x))
```

## Head and upper torso

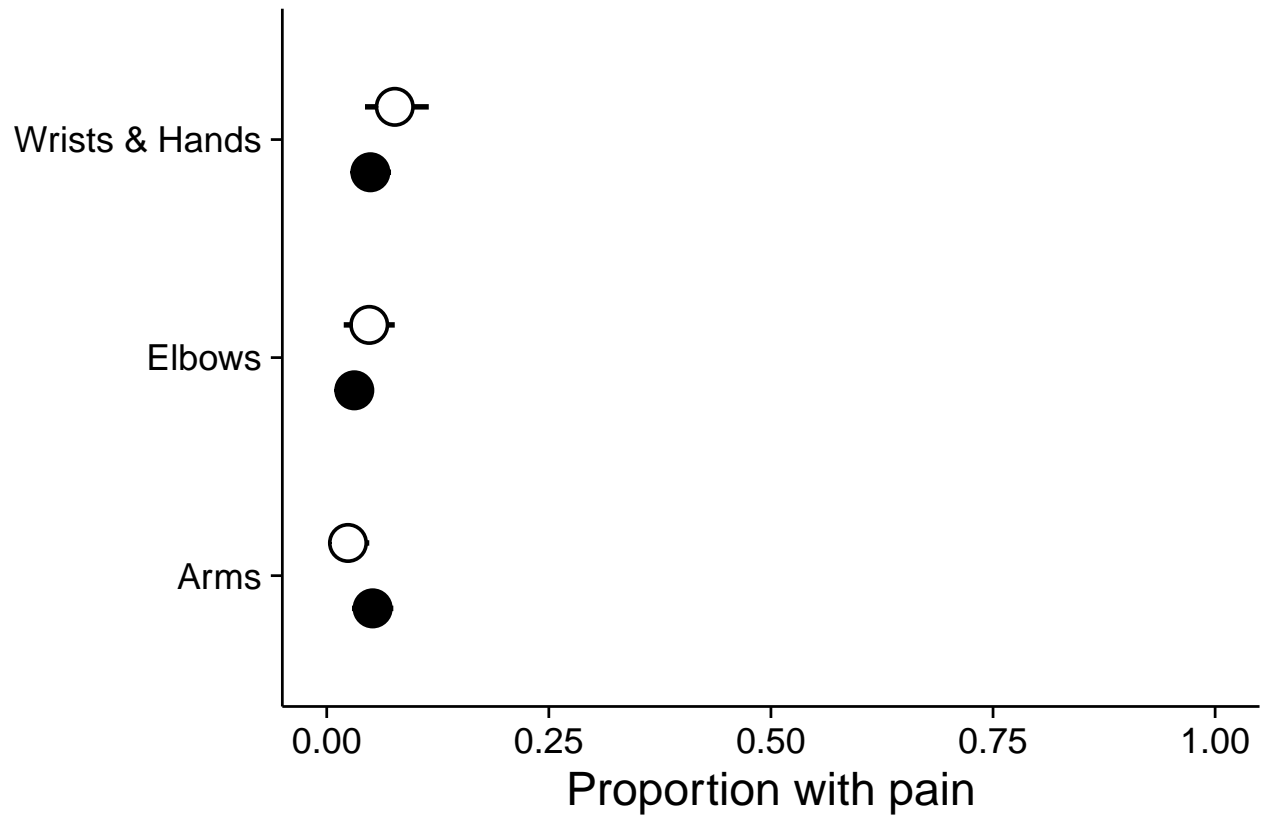
(Point estimate with 95%CI)



## Upper limbs

(Point estimate with 95%CI)

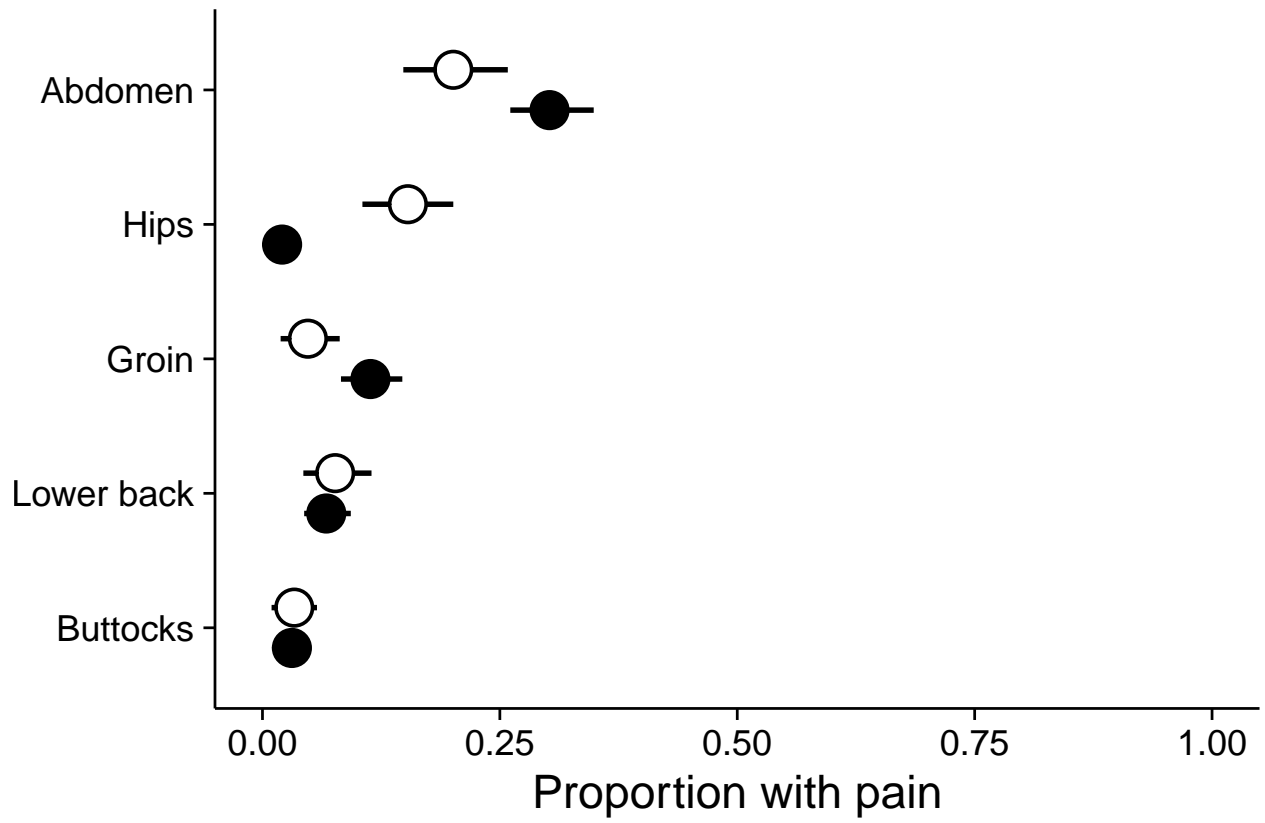
● Acute ○ Chronic



## Lower torso

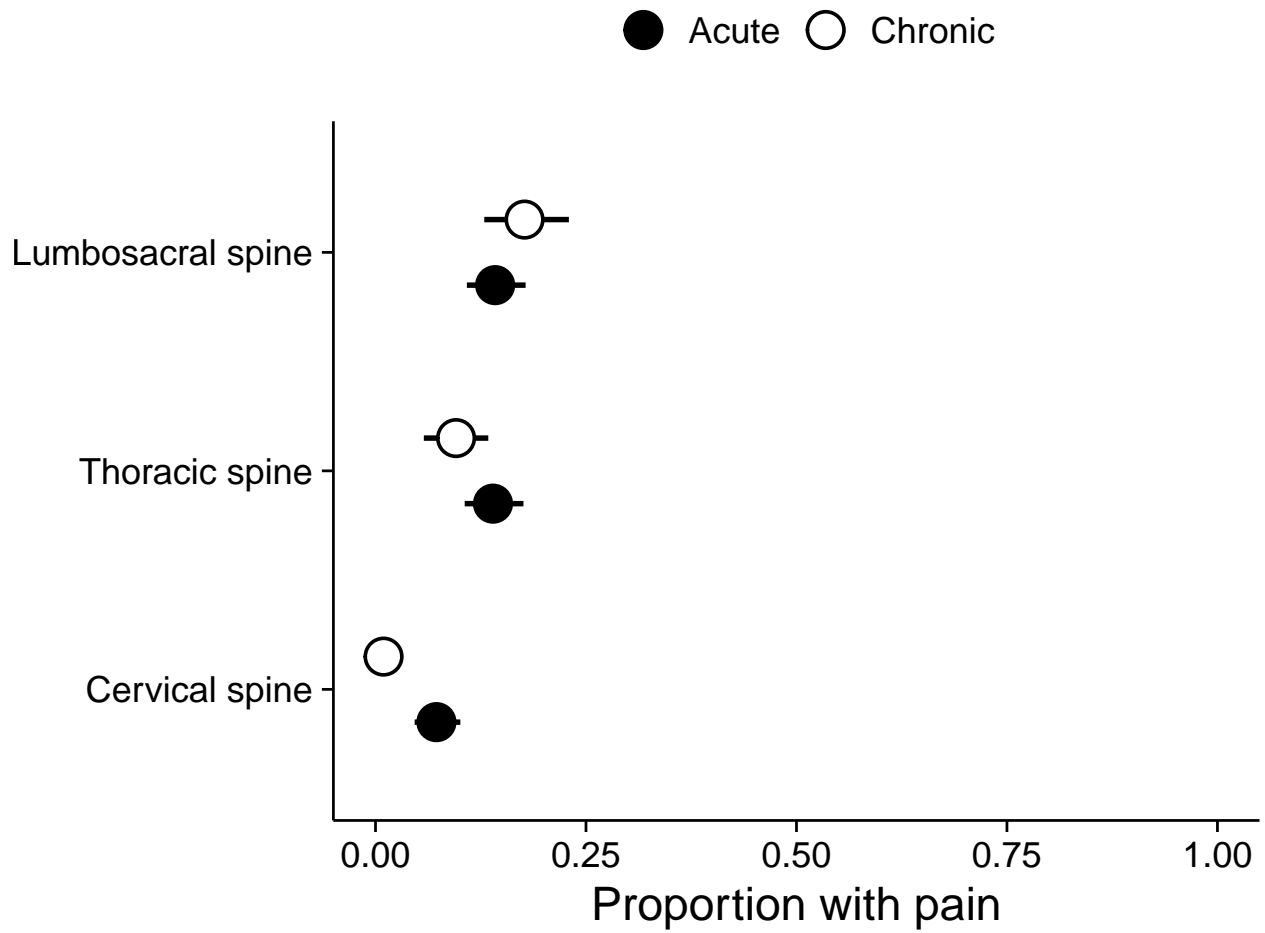
(Point estimate with 95%CI)

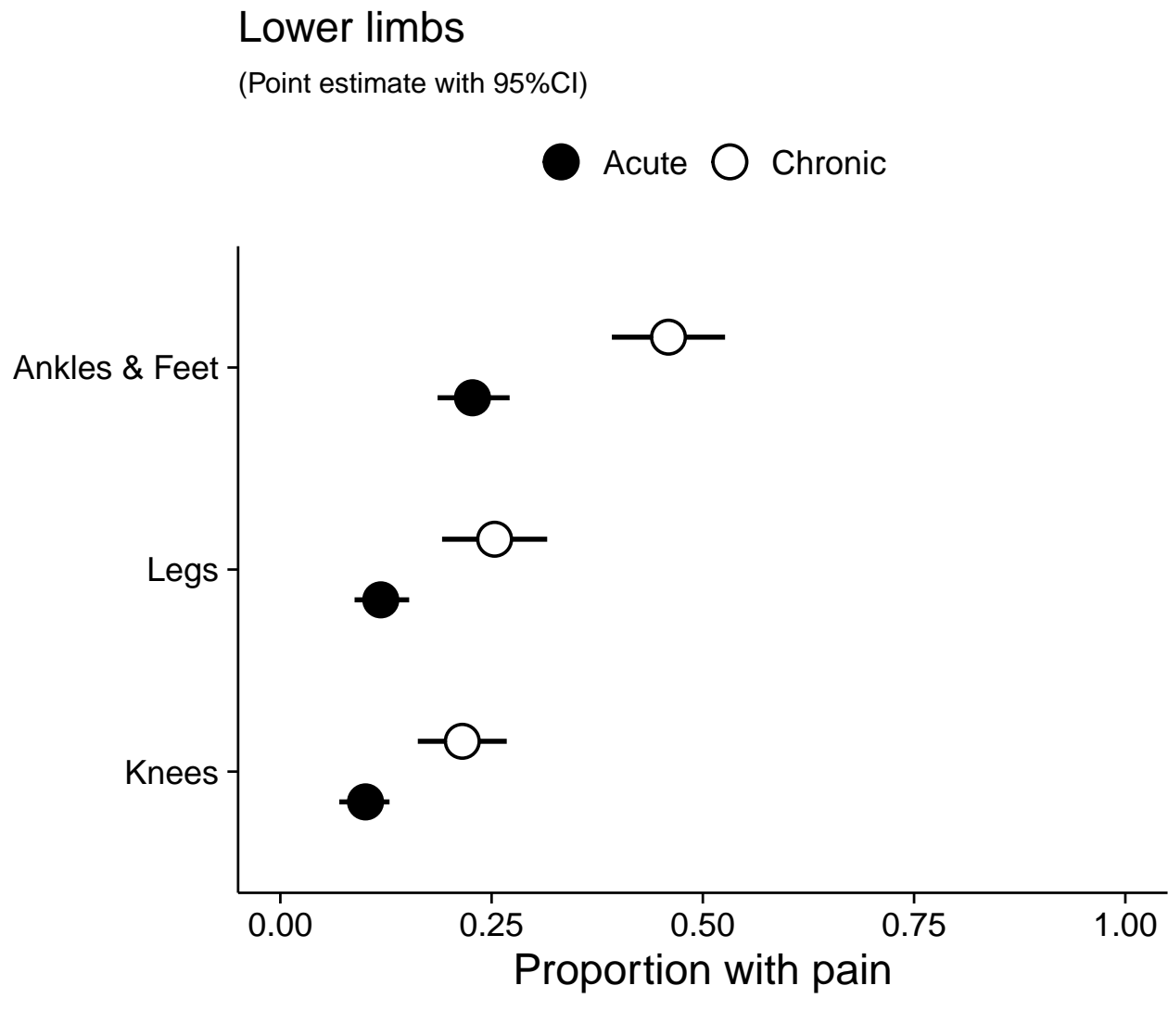
● Acute ○ Chronic



## Spinal column

(Point estimate with 95%CI)





## 8 Logistic regression

Note: Age and CD4 T-cell count treated as a continuous variable, not like the categories used in the figures shown above.

### 8.1 Process data

Extract and process data on study site, participant ID, CD4 T-cell count, age, sex, and pain chronicity.

```
# Select demographic data
demo_log <- demo[, c('ID', 'Site', 'CD4_recent', 'Sex', 'Age', 'Pain')]

# Process data
data_log <- data %>%
  # Join with extracted demographics data (demo_reduced)
```

```

left_join(demo_log) %>%
# Remove upper back (only one outcome -- no pain) and ID
select(-ID, -Upper_back)

# Convert data_glm to long format
data_log.long <- data_log %>%
  pivot_longer(cols = -c(CD4_recent, Age, Sex, Site, Pain),
               names_to = 'Pain_site',
               values_to = 'Pain_present')

# Scale Age and CD4_recent
data_log.long <- data_log.long %>%
  mutate_if(is.numeric, scale)

# Dummy code outcome variable
data_log.long <- data_log.long %>%
  mutate(Pain_present = ifelse(Pain_present == 'No',
                              yes = 0,
                              no = 1))

```

## 8.2 Family-wise error correction

Alpha threshold for significance was corrected to control the family-wise error rate. The correction was applied to the results of the likelihood ratio tests comparing null models to full models. We chose the conservative Dunn-Bonferroni correction.

## 8.3 Run glmer model for every body site to check for issues

P-values reported for the likelihood ratio tests below are uncorrected for multiple comparisons.

```

##-- Head --#
head <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Head') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
                  data = head,
                  family = binomial())

## Full model
head_mod <- glmer(Pain_present ~
                  Age +
                  Sex +
                  CD4_recent +
                  Pain +
                  (1|Site),
                  data = head,
                  family = binomial())

```

```

## Compare models
anova(null_mod, head_mod, test = 'LR')

## Data: head
## Models:
## null_mod: Pain_present ~ 1 + (1 | Site)
## head_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##      npar      AIC      BIC  logLik deviance Chisq Df Pr(>Chisq)
## null_mod    2 599.13 607.53 -297.56   595.13
## head_mod    6 589.86 615.06 -288.93   577.86 17.27  4   0.001713 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Significant, therefore, print summary
## (family-wise error correction, alpha = 0.0024)

summary(head_mod)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: head
##
##      AIC      BIC  logLik deviance df.resid
##    589.9    615.1  -288.9   577.9      487
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.0014 -0.7480 -0.4873  1.0605  2.6634
##
## Random effects:
## Groups Name             Variance Std.Dev.
## Site (Intercept) 0.1263  0.3553
## Number of obs: 493, groups: Site, 5
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.39255    0.28339  -1.385  0.16599
## Age          0.03543    0.12234   0.290  0.77212
## SexMale     -1.34687    0.38448  -3.503  0.00046 ***
## CD4_recent  -0.11133    0.11511  -0.967  0.33344
## PainChronic -0.54907    0.42960  -1.278  0.20121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Age    SexMal CD4_rc
## Age          0.186
## SexMale     -0.094 -0.103
## CD4_recent   0.068 -0.071  0.081
## PainChronic -0.685 -0.238 -0.042 -0.203

##-- Throat --#
throat <- data_log.long %>%

```



```

# Filter by pain site
filter(Pain_site == 'Throat') %>%
filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = throat,
  family = binomial())

## Full model
thrt_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = throat,
  family = binomial())

## Fit is singular, print model to check SD of random effect
thrt_mod

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: throat
##      AIC      BIC    logLik deviance df.resid
## 124.9796 150.1827 -56.4898 112.9796      487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
##   -4.12952    0.07201    0.91911   -0.78730    0.19658
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Shoulder --#
shoulder <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Shoulder') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = shoulder,
  family = binomial())

## Full model
shdr_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +

```

```

        Pain +
        (1|Site),
        data = shoulder,
        family = binomial())

## Compare models
anova(null_mod, shdr_mod, test = 'LR')

## Data: shoulder
## Models:
## null_mod: Pain_present ~ 1 + (1 | Site)
## shdr_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##      npar      AIC      BIC  logLik deviance  Chisq Df Pr(>Chisq)
## null_mod      2 286.42 294.82 -141.21   282.42
## shdr_mod      6 292.25 317.45 -140.12   280.25 2.1705  4      0.7044

## No significant difference between null and full model

##-- Arm --#
arms <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Arms') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = arms,
  family = binomial())

## Full model
arms_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = arms,
  family = binomial())

## Fit is singular, print model to check SD of random effect
arms_mod

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: arms
##      AIC      BIC  logLik deviance df.resid
## 189.1338 214.3369 -88.5669 177.1338      487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic

```

```

##      -2.89348      0.30866      -0.05672      -0.10059      -0.62336
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Elbow --#
elbow <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Elbows') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = elbow,
  family = binomial())

## Full model
elbw_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = elbow,
  family = binomial())

## Fit is singular, print model to check SD of random effect
elbw_mod

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: elbow
##      AIC      BIC    logLik deviance df.resid
## 172.1324 197.3355 -80.0662 160.1324      487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0.00000001534
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
##   -3.38921      0.42521     -0.16645      0.04229      0.45074
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Wrist & hand --#
hand <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Wrists.Hands') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = hand,
  family = binomial())

## Full model

```

```

hand_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = hand,
  family = binomial())

## Fit is singular, print model to check SD of random effect
hand_mod

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: hand
##      AIC      BIC   logLik deviance df.resid
## 243.296 268.499 -115.648 231.296      487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0.000000008689
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
##      -2.7754      0.2316      -0.6507      0.2766      0.4408
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Chest --#
chest <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Chest') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = chest,
  family = binomial())

## Full model
chst_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = chest,
  family = binomial())

## Compare models
anova(null_mod, chst_mod, test = 'LR')

## Data: chest
## Models:

```

```
## null_mod: Pain_present ~ 1 + (1 | Site)
## chst_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##          npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## null_mod    2 479.88 488.28 -237.94   475.88
## chst_mod    6 480.58 505.78 -234.29   468.58 7.3024  4    0.1207
```

*## No significant difference in null and full models*

*##-- Abdomen --#*

```
abdomen <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Abdomen') %>%
  filter(complete.cases(.))
```

*## Null*

```
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = abdomen,
  family = binomial())
```

*## Full model*

```
abdm_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = abdomen,
  family = binomial())
```

*## Fit is singular, print model to check SD of random effect*

abdm\_mod

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: abdomen
##      AIC      BIC    logLik deviance df.resid
## 581.1518 606.3548 -284.5759  569.1518     487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
##    -0.8813    -0.1613    -0.3072    -0.1145    -0.2542
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings
```

*##-- Lower back (flank) --#*

```
low_back <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Lower_back') %>%
  filter(complete.cases(.))
```

*## Null*

```

null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = low_back,
  family = binomial())

## Full model
back_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = low_back,
  family = binomial())

## Fit is singular, print model to check SD of random effect
back_mod

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: low_back
##      AIC      BIC    logLik deviance df.resid
## 272.1479 297.3510 -130.0739  260.1479     487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
##   -2.59625     0.10656    -0.23426     0.08667     0.33885
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Groin --#
groin <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Groin') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = groin,
  family = binomial())

## Full model
groin_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = groin,
  family = binomial())

```

```

## Compare models
anova(null_mod, groin_mod, test = 'LR')

## Data: groin
## Models:
## null_mod: Pain_present ~ 1 + (1 | Site)
## groin_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##           npar      AIC      BIC logLik deviance Chisq Df Pr(>Chisq)
## null_mod    2 304.54 312.94 -150.27   300.54
## groin_mod    6 308.23 333.43 -148.12   296.23 4.3039  4    0.3664
## No significant difference between null and full model

##-- Buttocks --#
buttocks <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Buttocks') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = buttocks,
  family = binomial())

## Full model
butt_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = buttocks,
  family = binomial())

## Fit is singular, print model to check SD of random effect
butt_mod

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: buttocks
##      AIC      BIC   logLik deviance df.resid
## 170.7258 195.9288 -79.3629 158.7258     487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0.000000003114
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
## -3.42677      -0.12939      0.65435      0.23689      0.09589
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Hips --#
hips <- data_log.long %>%

```

```

# Filter by pain site
filter(Pain_site == 'Hips') %>%
filter(complete.cases(.))

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = hips,
  family = binomial())

## Full model
hips_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = hips,
  family = binomial())

## Compare models
anova(null_mod, hips_mod, test = 'LR')

## Data: hips
## Models:
## null_mod: Pain_present ~ 1 + (1 | Site)
## hips_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##      npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## null_mod    2 224.25 232.65 -110.12   220.25
## hips_mod    6 225.28 250.49 -106.64   213.28 6.9633  4    0.1378

## No significant difference between null and full model

##-- Legs --#
legs <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Legs') %>%
  filter(complete.cases(.))

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = legs,
  family = binomial())

## Full model
legs_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = legs,
  family = binomial())

## Fit is singular, print model to check SD of random effect

```



```
legs_mod
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: legs
##      AIC      BIC    logLik deviance df.resid
## 431.0242 456.2273 -209.5121  419.0242     487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
##      -2.0238      0.1117      0.6088      0.1644      0.6567
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings
```

```
##-- Knees --#
```

```
knees <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Knees') %>%
  filter(complete.cases(.))
```

```
## Null
```

```
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = knees,
  family = binomial())
```

```
## Full model
```

```
knee_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = knees,
  family = binomial())
```

```
## Fit is singular, print model to check SD of random effect
```

```
knee_mod
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: knees
##      AIC      BIC    logLik deviance df.resid
## 400.1826 425.3857 -194.0913  388.1826     487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
```

```
##      -2.1860      0.2308      0.4132      0.0948      0.8155
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Ankles.Feet --#
ankles <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Ankles.Feet') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = ankles,
  family = binomial())

## Full model
ankl_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = ankles,
  family = binomial())

## Fit is singular, print model to check SD of random effect
ankl_mod

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
## Data: ankles
##      AIC      BIC    logLik deviance df.resid
## 570.1828 595.3858 -279.0914  558.1828     487
## Random effects:
## Groups Name      Std.Dev.
## Site (Intercept) 0.00000002105
## Number of obs: 493, groups: Site, 5
## Fixed Effects:
## (Intercept)      Age      SexMale  CD4_recent  PainChronic
##   -1.26753      0.27385      0.38509    -0.08338      0.96276
## convergence code 0; 0 optimizer warnings; 1 lme4 warnings

##-- Cervical spine --#
cervical <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Cervical_spine') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = cervical,
  family = binomial())

## Full model
```

```

neck_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = cervical,
  family = binomial())

## Compare models
anova(null_mod, neck_mod, test = 'LR')

## Data: cervical
## Models:
## null_mod: Pain_present ~ 1 + (1 | Site)
## neck_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##          npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## null_mod    2 205.43 213.83 -100.716   201.43
## neck_mod    6 206.19 231.40  -97.096   194.19 7.2402  4      0.1237

## No significant difference between null and full models

##-- Thoracic spine --#
thoracic <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Thoracic_spine') %>%
  filter(complete.cases(.))

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = thoracic,
  family = binomial())

## Full model
thrx_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = thoracic,
  family = binomial())

## Compare models
anova(null_mod, thrx_mod, test = 'LR')

## Data: thoracic
## Models:
## null_mod: Pain_present ~ 1 + (1 | Site)
## thrx_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##          npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## null_mod    2 392.73 401.13 -194.36   388.73
## thrx_mod    6 396.75 421.95 -192.38   384.75 3.9765  4      0.4092

```

```
## No significant difference between null and full models

# Lumbosacral spine
lumbar <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Lumbosacral_spine') %>%
  filter(complete.cases())

## Null
null_mod <- glmer(Pain_present ~ 1 + (1|Site),
  data = lumbar,
  family = binomial())

## Full model
lmbr_mod <- glmer(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain +
  (1|Site),
  data = lumbar,
  family = binomial())

## Compare models
anova(null_mod, lmbr_mod, test = 'LR')

## Data: lumbar
## Models:
## null_mod: Pain_present ~ 1 + (1 | Site)
## lmbr_mod: Pain_present ~ Age + Sex + CD4_recent + Pain + (1 | Site)
##      npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## null_mod    2 451.68 460.08 -223.84   447.68
## lmbr_mod    6 456.11 481.32 -222.06   444.11 3.567  4    0.4678

## No significant difference between null and full models
```

Body sites with models that generated an “is singular” error included:

- Throat
- Arms
- Elbows
- Wrists & hands
- Abdomen
- Lower back (flanks)
- Buttocks
- Legs
- Knees
- Ankles & feet

All these models had a random error variance of 0 (or tending to 0), indicating that **site** was not contributing to the model and could be removed. That is, there is not enough additional site-level variation to warrant

adding a random effect to explain all the observed variation. Analyzed these using `glm`, not `glmm`, and check for issues.

## 8.4 Run `glm` for selected body sites to check for issues

P-values reported for the likelihood ratio tests below are uncorrected for multiple comparisons.

```
##-- Throat --#
throat <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Throat') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
  data = throat,
  family = binomial())

## Full model
thrt_mod <- glm(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain,
  data = throat,
  family = binomial())

## Compare models
anova(null_mod, thrt_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         492      120.18
## 2         488      112.98  4    7.1991  0.1257

## No significant difference between the null and full models

##-- Arms --#
arms <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Arms') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
  data = arms,
  family = binomial())

## Full model
arms_mod <- glm(Pain_present ~
  Age +
```

```

        Sex +
        CD4_recent +
        Pain,
data = arms,
family = binomial())

## Compare models
anova(null_mod, arms_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         492      179.82
## 2         488      177.13  4    2.686   0.6117

## No significant difference between the null and full models

##-- Elbow --#
elbows <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Elbows') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
  data = elbows,
  family = binomial())

## Full model
elbw_mod <- glm(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain,
  data = elbows,
  family = binomial())

## Compare models
anova(null_mod, elbw_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         492      167.37
## 2         488      160.13  4    7.2361  0.1239

## No significant difference between the null and full models

##-- Wrist & hand --#
wrists <- data_log.long %>%
  # Filter by pain site

```

```

filter(Pain_site == 'Wrists.Hands') %>%
filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
               data = wrists,
               family = binomial())

## Full model
wrst_mod <- glm(Pain_present ~
               Age +
               Sex +
               CD4_recent +
               Pain,
               data = wrists,
               family = binomial())

## Compare models
anova(null_mod, wrst_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         492      242.2
## 2         488      231.3  4   10.908  0.02762 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## No significant difference between the null and full models
## (family-wise error correction, alpha = 0.0024)

##-- Abdomen --#
abdomen <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Abdomen') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
               data = abdomen,
               family = binomial())

## Full model
abdm_mod <- glm(Pain_present ~
               Age +
               Sex +
               CD4_recent +
               Pain,
               data = abdomen,
               family = binomial())

## Compare models

```

```

anova(null_mod, abdm_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         492      578.82
## 2         488      569.15  4   9.6645  0.04648 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## No significant difference between the null and full models
## (family-wise error correction, alpha = 0.0024)

##-- Lower back (flank) --#
low_back <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Lower_back') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
  data = low_back,
  family = binomial())

## Full model
back_mod <- glm(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain,
  data = low_back,
  family = binomial())

## Compare models
anova(null_mod, back_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         492      262.78
## 2         488      260.15  4   2.6328  0.621

## No significant difference between the null and full models

##-- Buttocks --#
buttocks <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Buttocks') %>%
  filter(complete.cases())

## Null

```



```

null_mod <- glm(Pain_present ~ 1,
               data = buttocks,
               family = binomial())

## Full model
butt_mod <- glm(Pain_present ~
               Age +
               Sex +
               CD4_recent +
               Pain,
               data = buttocks,
               family = binomial())

## Compare models
anova(null_mod, butt_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1         492      160.99
## 2         488      158.73  4   2.2631  0.6875

## No significant difference between the null and full models

##-- Legs --#
legs <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Legs') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
               data = legs,
               family = binomial())

## Full model
legs_mod <- glm(Pain_present ~
               Age +
               Sex +
               CD4_recent +
               Pain,
               data = legs,
               family = binomial())

## Compare models
anova(null_mod, legs_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance  Pr(>Chi)

```

```

## 1      492      440.48
## 2      488      419.02  4    21.454 0.0002574 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Significant difference between the null and full models
## (family-wise error correction, alpha = 0.0024)
summary(legs_mod)

##
## Call:
## glm(formula = Pain_present ~ Age + Sex + CD4_recent + Pain, family = binomial(),
##      data = legs)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0034  -0.6258  -0.4805  -0.4483   2.1998
##
## Coefficients:
##              Estimate Std. Error z value      Pr(>|z|)
## (Intercept)  -2.0238     0.1812 -11.170 <0.0000000000000002 ***
## Age           0.1117     0.1305   0.856     0.3920
## SexMale       0.6088     0.3013   2.020     0.0433 *
## CD4_recent    0.1644     0.1252   1.313     0.1892
## PainChronic  0.6567     0.2903   2.262     0.0237 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 440.48  on 492  degrees of freedom
## Residual deviance: 419.02  on 488  degrees of freedom
## AIC: 429.02
##
## Number of Fisher Scoring iterations: 4

##-- Knees --#
knees <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Knees') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
  data = knees,
  family = binomial())

## Full model
knee_mod <- glm(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain,
  data = knees,
  family = binomial())

```

```

## Compare models
anova(null_mod, knee_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance   Pr(>Chi)
## 1         492      413.48
## 2         488      388.18  4    25.298 0.00004382 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Significant difference between the null and full models
## (family-wise error correction, alpha = 0.0024)
summary(knee_mod)

##
## Call:
## glm(formula = Pain_present ~ Age + Sex + CD4_recent + Pain, family = binomial(),
##      data = knees)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0315  -0.5925  -0.4460  -0.4038   2.3429
##
## Coefficients:
##              Estimate Std. Error z value      Pr(>|z|)
## (Intercept)  -2.1860     0.1932 -11.314 < 0.0000000000000002 ***
## Age           0.2308     0.1323   1.745     0.08100 .
## SexMale       0.4132     0.3147   1.313     0.18928
## CD4_recent    0.0948     0.1320   0.718     0.47273
## PainChronic   0.8155     0.3015   2.705     0.00683 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 413.48  on 492  degrees of freedom
## Residual deviance: 388.18  on 488  degrees of freedom
## AIC: 398.18
##
## Number of Fisher Scoring iterations: 5

##-- Ankles/feet --#
ankles <- data_log.long %>%
  # Filter by pain site
  filter(Pain_site == 'Ankles.Feet') %>%
  filter(complete.cases())

## Null
null_mod <- glm(Pain_present ~ 1,
  data = ankles,
  family = binomial())

```

```

## Full model
ankl_mod <- glm(Pain_present ~
  Age +
  Sex +
  CD4_recent +
  Pain,
  data = ankles,
  family = binomial())

## Compare models
anova(null_mod, ankl_mod, test = 'LR')

## Analysis of Deviance Table
##
## Model 1: Pain_present ~ 1
## Model 2: Pain_present ~ Age + Sex + CD4_recent + Pain
##   Resid. Df Resid. Dev Df Deviance      Pr(>Chi)
## 1         492      600.78
## 2         488      558.18  4    42.596 0.00000001255 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Significant difference between the null and full models
## (family-wise error correction, alpha = 0.0024)
summary(ankl_mod)

##
## Call:
## glm(formula = Pain_present ~ Age + Sex + CD4_recent + Pain, family = binomial(),
##      data = ankles)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4967  -0.7619  -0.6565   1.1107   1.9524
##
## Coefficients:
##              Estimate Std. Error z value      Pr(>|z|)
## (Intercept) -1.26753    0.14377  -8.816 < 0.0000000000000002 ***
## Age          0.27385    0.11072   2.473    0.0134 *
## SexMale      0.38509    0.26289   1.465    0.1430
## CD4_recent  -0.08338    0.11100  -0.751    0.4526
## PainChronic  0.96276    0.24465   3.935    0.0000831 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 600.78  on 492  degrees of freedom
## Residual deviance: 558.18  on 488  degrees of freedom
## AIC: 568.18
##
## Number of Fisher Scoring iterations: 4

```

## 8.5 GLMM data for tabulation and plotting

Perform GLMM on selected body sites (head, shoulders, chest, groin, hips, cervical spine, thoracic spine, and lumbosacral spine) to extract fixed effect odds ratios (with 95% CIs) to tabulate and to produce forest plots.

```
# Perform GLMM on each selected body sites
glmm_mods <- data_log.long %>%
  filter(complete.cases(.)) %>%
  # remove body sites with glmm singular fits
  filter(!Pain_site %in% c("Throat", "Arms", "Elbows", "Lower_back",
                           "Wrists.Hands", "Abdomen", "Buttocks",
                           "Legs", "Knees", "Ankles.Feet")) %>%

  # Group and then nest by Pain_site
  group_by(Pain_site) %>%
  nest() %>%
  # Perform logistic regression
  mutate(logistic = map(.x = data,
                        ~ glmer(Pain_present ~
                                Age +
                                Sex +
                                CD4_recent +
                                Pain +
                                (1|Site),
                                data = .x,
                                family = binomial()))),
         null = map(.x = data,
                    ~glmer(Pain_present ~ 1 +
                            (1|Site),
                            data = .x,
                            family = binomial())))) %>%

  # Perform likelihood ratio test
  mutate(LRT = map2(.x = logistic,
                    .y = null,
                    ~ anova(.y, .x, test = 'LR')))) %>%

  # Extract LRT p-values and correct for multiple comparisons
  mutate(LRT_p = map(.x = LRT,
                     ~ .x[, 8][[2]]),
         LRT_p.corrected = map(.x = LRT_p,
                               ~ p.adjust(p = .x,
                                           method = 'bonferroni',
                                           n = 21))) %>%

  # Extract effect sizes
  mutate(OR = map(.x = logistic,
                  ~ exp(fixef(.x))),
         CI = map(.x = logistic,
                  ~ exp(confint.merMod(.x,
                                       method = 'Wald')[2:6, ]))) %>%

  ungroup() %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\\\.',
```

```

replacement = ' & ')') %>%
# Bind data into a dataframe and tabulate,
mutate(df = pmap(.l = list(OR, CI, LRT_p.corrected, Pain_site),
~ cbind(..1, ..2) %>%
.[2:5, ] %>%
kable(caption = str_glue('{..4} (GLMM; LRT corrected p-value = {round(..3, 3)})',
digits = 3,
col.names = c('OR', 'Wald lower 95%CI',
'Wald upper 95%CI')))) %>%

# Generate forest plot data
mutate(forest_data = map2(.x = OR,
.y = CI,
# Create the dataframe
~ as.data.frame(cbind(.x, .y)) %>%
set_names(nm = c('OR', 'lower_ci',
'upper_ci')) %>%
.[-1, ] %>%
rownames_to_column() %>%
mutate(rowname = case_when(
rowname == 'Age' ~ 'Age',
rowname == 'SexMale' ~ 'Sex (male)',
rowname == 'CD4_recent' ~ 'CD4 T-cells',
rowname == 'PainChronic' ~ 'Pain (chronic)'
)) %>%
# Round digits to 2 decimal places
mutate_if(is.numeric, round, 2) %>%
# Paste CI range
mutate(ci_range = paste0('(', lower_ci, ' to ',
upper_ci, ')')) %>%
# Cut short confidence interval for plotting
mutate(upper_ci = ifelse(upper_ci > 4,
yes = 4,
no = upper_ci)))) %>%

# Generate forest plots
mutate(forest_plots = pmap(.l = list(forest_data, Pain_site, LRT_p.corrected),
~ ..1 %>%
ggplot(data = .) +
aes(x = OR,
xmin = lower_ci,
xmax = upper_ci,
y = rowname,
label = ci_range) +
geom_pointrange(size = 1) +
geom_vline(xintercept = 1,
linetype = 2) +
scale_x_continuous(limits = c(-0.5, 8),
breaks = c(0, 1, 2, 3, 4)) +
geom_text(x = 4.5, hjust = 0, size = 5) +
annotate(geom = 'text',
label = '95% CI',
size = 5,
fontface = 2,
x = 4.5,

```

```

      y = 4.4,
      hjust = 0) +
labs(title = ..2,
      subtitle = str_glue('(GLMM | LRT corrected p-value = {round(..3
      x = 'Odds ratio') +
theme_minimal(base_size = 18) +
theme(plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 14),
      panel.grid = element_blank(),
      axis.title.y = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5)))

```

## 8.6 GLM data for tabulation and plotting

Perform GLM on selected body sites (throat, arms, elbows, wrists and hands, abdomen, buttocks, legs, knees, and ankles and feet) to extract fixed effect odds ratios (with 95% CIs) to tabulate and to produce forest plots.

```

# Perform GLM on each selected body sites
glm_mods <- data_log.long %>%
  filter(complete.cases()) %>%
  # Retain body sites with glmm singular fits
  filter(Pain_site %in% c("Throat", "Arms", "Elbows", "Lower_back",
    "Wrists.Hands", "Abdomen", "Buttocks",
    "Legs", "Knees", "Ankles.Feet")) %>%
  # Change lower_back to flank
  mutate(Pain_site = ifelse(Pain_site == 'Lower_back',
    yes = 'Lower back/flanks',
    no = Pain_site)) %>%
  # Group and then nest by Pain_site
  group_by(Pain_site) %>%
  nest() %>%
  # Perform logistic regression
  mutate(logistic = map(.x = data,
    ~ glm(Pain_present ~
      Age +
      Sex +
      CD4_recent +
      Pain,
      data = .x,
      family = binomial()))),
  null = map(.x = data,
    ~glm(Pain_present ~ 1,
      data = .x,
      family = binomial())) %>%
  # Perform likelihood ratio test
  mutate(LRT = map2(.x = logistic,
    .y = null,
    ~ anova(.y, .x, test = 'LR')) %>%
  # Extract LRT p-values and correct for multiple comparisons
  mutate(LRT_p = map(.x = LRT,

```

```

      ~ .x[, 5][[2]]),
    LRT_p.corrected = map(.x = LRT_p,
      ~ p.adjust(p = .x,
        method = 'bonferroni',
        n = 21))) %>%

# Extract effect sizes
mutate(OR = map(.x = logistic,
  ~ exp(coef(.x))),
  CI = map(.x = logistic,
    ~ exp(confint.default(.x)))) %>%

ungroup() %>%
mutate(Pain_site = str_replace(Pain_site,
  pattern = '_',
  replacement = ' ')) %>%
mutate(Pain_site = str_replace(Pain_site,
  pattern = '\\\\.',
  replacement = ' & ')) %>%

# Bind data into a dataframe and tabulate,
mutate(df = pmap(.l = list(OR, CI, LRT_p.corrected, Pain_site),
  ~ cbind(..1, ..2) %>%
    .[2:5, ] %>%
    kable(caption = str_glue('{..4} (GLM; LRT corrected p-value = {round(..3, 3)})'
      digits = 3,
      col.names = c('OR', 'Wald lower 95%CI',
        'Wald upper 95%CI')))) %>%

# Generate forest plot data
mutate(forest_data = map2(.x = OR,
  .y = CI,
  # Create the dataframe
  ~ as.data.frame(cbind(.x, .y)) %>%
    set_names(nm = c('OR', 'lower_ci',
      'upper_ci')) %>%
    .[-1, ] %>%
    rownames_to_column() %>%
    mutate(rowname = case_when(
      rowname == 'Age' ~ 'Age',
      rowname == 'SexMale' ~ 'Sex (male)',
      rowname == 'CD4_recent' ~ 'CD4 T-cells',
      rowname == 'PainChronic' ~ 'Pain (chronic)'
    )) %>%
    # Round digits to 2 decimal places
    mutate_if(is.numeric, round, 2) %>%
    # Paste CI range
    mutate(ci_range = paste0('(', lower_ci, ' to ',
      upper_ci, '))' %>%
    # Cut short confidence interval for plotting
    mutate(upper_ci = ifelse(upper_ci > 4,
      yes = 4,
      no = upper_ci)))) %>%

# Generate forest plots
mutate(forest_plots = pmap(.l = list(forest_data, Pain_site, LRT_p.corrected),
  ~ ..1 %>%
    ggplot(data = .) +

```



```

aes(x = OR,
    xmin = lower_ci,
    xmax = upper_ci,
    y = rowname,
    label = ci_range) +
geom_pointrange(size = 1) +
geom_vline(xintercept = 1,
            linetype = 2) +
scale_x_continuous(limits = c(-0.5, 8),
                    breaks = c(0, 1, 2, 3, 4)) +
geom_text(x = 4.5, hjust = 0, size = 5) +
annotate(geom = 'text',
          label = '95% CI',
          size = 5,
          fontface = 2,
          x = 4.5,
          y = 4.4,
          hjust = 0) +
labs(title = ..2,
     subtitle = str_glue('(GLM | LRT corrected p-value = {round(..3,
     x = 'Odds ratio')' +
theme_minimal(base_size = 18) +
theme(plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 14),
      panel.grid = element_blank(),
      axis.title.y = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5)))

```

## 8.7 Tabulated results (fixed effects)

### 8.7.1 GLMM tabulations

```

# Print OR for fixed effects with 95% CIs and p-values
walk(glmms_mods$df, ~print(.x))

```

Table 59: Head (GLMM; LRT corrected p-value = 0.036)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.036 | 0.815            | 1.317            |
| SexMale     | 0.260 | 0.122            | 0.553            |
| CD4_recent  | 0.895 | 0.714            | 1.121            |
| PainChronic | 0.577 | 0.249            | 1.340            |

Table 60: Shoulder (GLMM; LRT corrected p-value = 1)

|     | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-----|-------|------------------|------------------|
| Age | 0.861 | 0.578            | 1.282            |

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| SexMale     | 1.703 | 0.733            | 3.957            |
| CD4_recent  | 0.941 | 0.651            | 1.360            |
| PainChronic | 1.111 | 0.451            | 2.739            |

Table 61: Chest (GLMM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.099 | 0.838            | 1.442            |
| SexMale     | 1.044 | 0.551            | 1.979            |
| CD4_recent  | 0.757 | 0.564            | 1.015            |
| PainChronic | 0.470 | 0.219            | 1.008            |

Table 62: Cervical spine (GLMM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.159 | 0.690            | 1.945            |
| SexMale     | 0.546 | 0.113            | 2.641            |
| CD4_recent  | 1.068 | 0.666            | 1.713            |
| PainChronic | 0.071 | 0.008            | 0.622            |

Table 63: Thoracic spine (GLMM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 0.987 | 0.723            | 1.347            |
| SexMale     | 0.515 | 0.204            | 1.305            |
| CD4_recent  | 1.156 | 0.877            | 1.524            |
| PainChronic | 0.709 | 0.309            | 1.628            |

Table 64: Lumbosacral spine (GLMM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.030 | 0.781            | 1.359            |
| SexMale     | 1.033 | 0.524            | 2.038            |
| CD4_recent  | 1.254 | 0.984            | 1.599            |
| PainChronic | 0.943 | 0.449            | 1.980            |

Table 65: Groin (GLMM; LRT corrected p-value = 1)

|            | OR    | Wald lower 95%CI | Wald upper 95%CI |
|------------|-------|------------------|------------------|
| Age        | 0.947 | 0.680            | 1.320            |
| SexMale    | 0.532 | 0.224            | 1.264            |
| CD4_recent | 1.229 | 0.885            | 1.706            |

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| PainChronic | 0.661 | 0.084            | 5.217            |

Table 66: Hips (GLMM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.057 | 0.728            | 1.536            |
| SexMale     | 1.373 | 0.581            | 3.245            |
| CD4_recent  | 1.131 | 0.802            | 1.593            |
| PainChronic | 6.570 | 2.204            | 19.585           |

## 8.7.2 GLM tabulations

```
# Print OR for fixed effects with 95% CIs and p-values
walk(glm_mods$df, ~print(.x))
```

Table 67: Throat (GLM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.075 | 0.562            | 2.056            |
| SexMale     | 2.507 | 0.733            | 8.576            |
| CD4_recent  | 0.455 | 0.188            | 1.104            |
| PainChronic | 1.217 | 0.316            | 4.696            |

Table 68: Arms (GLM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.362 | 0.888            | 2.087            |
| SexMale     | 0.945 | 0.297            | 3.006            |
| CD4_recent  | 0.904 | 0.548            | 1.493            |
| PainChronic | 0.536 | 0.169            | 1.704            |

Table 69: Elbows (GLM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.530 | 1.013            | 2.310            |
| SexMale     | 0.847 | 0.266            | 2.699            |
| CD4_recent  | 1.043 | 0.657            | 1.656            |
| PainChronic | 1.569 | 0.548            | 4.491            |

Table 70: Wrists &amp; Hands (GLM; LRT corrected p-value = 0.58)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.261 | 0.885            | 1.795            |
| SexMale     | 0.522 | 0.172            | 1.583            |
| CD4_recent  | 1.319 | 0.943            | 1.844            |
| PainChronic | 1.554 | 0.659            | 3.663            |

Table 71: Lower back/flanks (GLM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.112 | 0.781            | 1.585            |
| SexMale     | 0.791 | 0.307            | 2.036            |
| CD4_recent  | 1.091 | 0.775            | 1.535            |
| PainChronic | 1.403 | 0.621            | 3.173            |

Table 72: Abdomen (GLM; LRT corrected p-value = 0.976)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 0.851 | 0.670            | 1.081            |
| SexMale     | 0.736 | 0.407            | 1.328            |
| CD4_recent  | 0.892 | 0.711            | 1.118            |
| PainChronic | 0.776 | 0.456            | 1.318            |

Table 73: Legs (GLM; LRT corrected p-value = 0.005)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.118 | 0.866            | 1.444            |
| SexMale     | 1.838 | 1.018            | 3.318            |
| CD4_recent  | 1.179 | 0.922            | 1.506            |
| PainChronic | 1.928 | 1.092            | 3.407            |

Table 74: Knees (GLM; LRT corrected p-value = 0.001)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 1.260 | 0.972            | 1.632            |
| SexMale     | 1.512 | 0.816            | 2.801            |
| CD4_recent  | 1.099 | 0.849            | 1.424            |
| PainChronic | 2.260 | 1.252            | 4.081            |

Table 75: Ankles &amp; Feet (GLM; LRT corrected p-value = 0)

|     | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-----|-------|------------------|------------------|
| Age | 1.315 | 1.058            | 1.634            |

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| SexMale     | 1.470 | 0.878            | 2.460            |
| CD4_recent  | 0.920 | 0.740            | 1.144            |
| PainChronic | 2.619 | 1.621            | 4.230            |

Table 76: Buttocks (GLM; LRT corrected p-value = 1)

|             | OR    | Wald lower 95%CI | Wald upper 95%CI |
|-------------|-------|------------------|------------------|
| Age         | 0.879 | 0.518            | 1.490            |
| SexMale     | 1.924 | 0.620            | 5.965            |
| CD4_recent  | 1.267 | 0.811            | 1.981            |
| PainChronic | 1.101 | 0.360            | 3.367            |

## 8.8 Plotted results (fixed effects)

Includes both GLM and GLMM results.

```
# Extract and process sub-plots from GLMM data
head <- glmm_mods$forest_plots[[1]] +
  theme(axis.title.x = element_blank())

shoulder <- glmm_mods$forest_plots[[2]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

chest <- glmm_mods$forest_plots[[3]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

cervical <- glmm_mods$forest_plots[[4]] +
  theme(axis.text.y = element_blank())

thoracic<- glmm_mods$forest_plots[[5]]

lumbar <- glmm_mods$forest_plots[[6]] +
  theme(axis.text.y = element_blank())

groin <- glmm_mods$forest_plots[[7]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

hips <- glmm_mods$forest_plots[[8]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

# Extract and process sub-plots from GLM data
throat <- glm_mods$forest_plots[[1]] +
  theme(axis.title.x = element_blank(),
        axis.text.y = element_blank())
```

```

arms <- glm_mods$forest_plots[[2]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

elbows <- glm_mods$forest_plots[[3]] +
  theme(axis.title.x = element_blank())

wrists.hands <- glm_mods$forest_plots[[4]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

flank <- glm_mods$forest_plots[[5]] +
  theme(axis.title.x = element_blank())

abdomen <- glm_mods$forest_plots[[6]] +
  theme(axis.title.x = element_blank(),
        axis.text.y = element_blank())

legs <- glm_mods$forest_plots[[7]] +
  theme(axis.title.x = element_blank())

knees <- glm_mods$forest_plots[[8]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

ankles.feet <- glm_mods$forest_plots[[9]] +
  theme(axis.text.y = element_blank())

buttocks <- glm_mods$forest_plots[[10]] +
  theme(axis.text.y = element_blank(),
        axis.title.x = element_blank())

# Patchwork
log_patch <- head + throat + chest + abdomen + flank + groin +
  shoulder + arms + elbows + wrists.hands +
  buttocks + hips + legs + knees + ankles.feet +
  cervical + thoracic + lumbar +
  plot_layout(ncol = 4)

# Save
ggsave(filename = 'figures/figure_2.png',
       plot = log_patch,
       width = 16,
       height = 18)

```

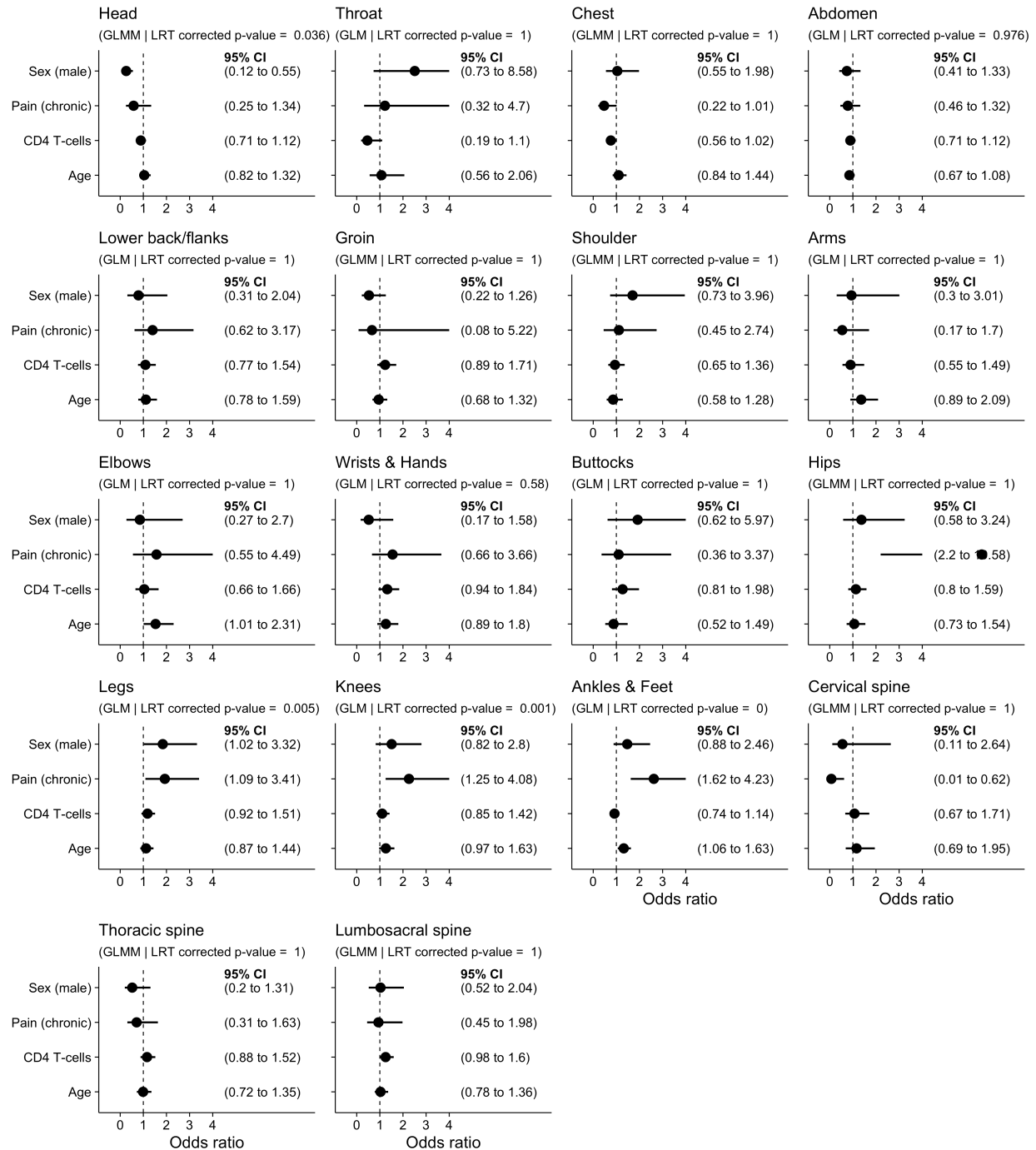


Figure 1: Plotted results (fixed effects)

## 9 Summary plots

### 9.1 Body sites with a point estimate pain proportion >10%

```
# Set seed
set.seed(2020)

# Generate a filter to extract sites with >10% pain prevalence
filter <- data_log %>%
  select(-Site, -CD4_recent, -Age, -Sex, -Pain) %>%
  pivot_longer(cols = everything(),
               names_to = 'Site',
               values_to = 'Response') %>%
  group_by(Site, Response) %>%
  summarise(count = n()) %>%
  mutate(total = sum(count),
         prop = count/total) %>%
  filter(Response == 'Yes')

filter_geq10 <- filter(filter, prop >= 0.1) %>%
  .$Site

# Generate new dataset
data_geq10 <- data_log %>%
  pivot_longer(cols = -c(Site, CD4_recent, Sex, Age, Pain),
               names_to = 'Pain_site',
               values_to = 'Pain_present') %>%
  # Filter by filter_geq10
  filter(Pain_site %in% filter_geq10)

pubs_nofacet <- data_geq10 %>%
  group_by(Pain_site) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,
                    ~ boot(data = .x[, 'Pain_present'],
                          statistic = prop_func,
                          R = 999,
                          stype = 'i',
                          parallel = 'multicore',
                          ncpus = 4))) %>%
  # Get CI
  mutate(ci = map(.x = boot,
                  ~ boot.ci(.x, type = 'perc')))) %>%
  # Extract ci data
  mutate(point_est = map(.x = ci,
                        ~ .x$t0),
         lower_ci = map(.x = ci,
                        ~ .x$percent[[4]]),
         upper_ci = map(.x = ci,
                        ~ .x$percent[[5]])) %>%
  # Remove columns
```



```

select(-boot, -ci) %>%
# Unnest
unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
ungroup()

plot_nofacet <- pubs_nofacet %>%
select(-CD4_recent, -Sex, -Age, -Pain_present, -Site, -Pain) %>%
distinct() %>%
mutate(Pain_site = str_replace(Pain_site,
                              pattern = '_',
                              replacement = ' ')) %>%
mutate(Pain_site = str_replace(Pain_site,
                              pattern = '\\\\.',
                              replacement = ' & ')) %>%
mutate(Pain_site = factor(Pain_site,
                          levels = c('Head', 'Chest', 'Abdomen',
                                      'Legs', 'Knees', 'Ankles & Feet',
                                      'Thoracic spine', 'Lumbosacral spine'),
                          ordered = TRUE)) %>%

ggplot(data = .) +
aes(x = fct_rev(Pain_site),
    y = point_est,
    ymin = lower_ci,
    ymax = upper_ci) +
geom_linerange(size = 1) +
geom_point(size = 5) +
coord_flip() +
labs(title = 'Body sites',
     subtitle = '(Point estimate with 95%CI)',
     y = 'Proportion of participants with pain') +
scale_y_continuous(limits = c(0, 0.6)) +
theme_minimal(base_size = 18) +
theme(plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      axis.title.y = element_blank(),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))

#-- Body sites by sex --#
pubs_nofacet.sex <- data_geq10 %>%
group_by(Pain_site, Sex) %>%
nest() %>%
# Bootstrap data
mutate(boot = map(.x = data,
                  ~ boot(data = .x[, 'Pain_present'],
                        statistic = prop_func,
                        R = 999,
                        stype = 'i',
                        parallel = 'multicore',
                        ncpus = 4))) %>%
# Get CI

```

```

mutate(ci = map(.x = boot,
               ~ boot.ci(.x, type = 'perc')))) %>%
# Extract ci data
mutate(point_est = map(.x = ci,
                      ~ .x$t0),
       lower_ci = map(.x = ci,
                      ~ .x$percent[[4]]),
       upper_ci = map(.x = ci,
                      ~ .x$percent[[5]])) %>%

# Remove columns
select(-boot, -ci) %>%
# Unnest
unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
ungroup()

plot_nofacet.sex <- pubs_nofacet.sex %>%
select(-CD4_recent, -Age, -Pain_present, -Site, -Pain) %>%
distinct() %>%
mutate(Pain_site = str_replace(Pain_site,
                              pattern = '_',
                              replacement = ' ')) %>%
mutate(Pain_site = str_replace(Pain_site,
                              pattern = '\\\\.',
                              replacement = ' & ')) %>%
mutate(Pain_site = factor(Pain_site,
                         levels = c('Head', 'Chest', 'Abdomen',
                                   'Legs', 'Knees', 'Ankles & Feet',
                                   'Thoracic spine', 'Lumbosacral spine'),
                         ordered = TRUE)) %>%

ggplot(data = .) +
aes(x = fct_rev(Pain_site),
   y = point_est,
   ymin = lower_ci,
   ymax = upper_ci,
   fill = Sex) +
geom_linerange(position = position_dodge2(width = 0.6),
               size = 1,
               colour = '#000000') +
geom_point(shape = 21,
            colour = '#000000',
            position = position_dodge2(width = 0.6),
            size = 4,
            stroke = 1) +
coord_flip() +
labs(title = 'Body sites by sex',
     subtitle = '(Point estimate with 95%CI)',
     y = 'Proportion of participants with pain') +
scale_y_continuous(limits = c(0, 0.6)) +
scale_fill_manual(values = c('#FFFFFF', '#000000')) +
theme_minimal(base_size = 18) +
theme(legend.position = c(0.8, 0.1),
      legend.title = element_blank(),
      plot.title = element_text(size = 18),

```

```

    plot.subtitle = element_text(size = 12),
    axis.title.y = element_blank(),
    axis.text.y = element_blank(),
    panel.grid = element_blank(),
    axis.text = element_text(colour = '#000000'),
    axis.line = element_line(size = 0.5),
    axis.ticks = element_line(size = 0.5))

#-- Body sites by pain chronicity --#
pubs_nofacet.pain <- data_geq10 %>%
  group_by(Pain_site, Pain) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,
                    ~ boot(data = .x[, 'Pain_present'],
                          statistic = prop_func,
                          R = 999,
                          stype = 'i',
                          parallel = 'multicore',
                          ncpus = 4))) %>%

  # Get CI
  mutate(ci = map(.x = boot,
                  ~ boot.ci(.x, type = 'perc')))) %>%
  # Extract ci data
  mutate(point_est = map(.x = ci,
                        ~ .x$t0),
         lower_ci = map(.x = ci,
                        ~ .x$percent[[4]]),
         upper_ci = map(.x = ci,
                        ~ .x$percent[[5]])) %>%
  # Remove columns
  select(-boot, -ci) %>%
  # Unnest
  unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
  ungroup()

plot_nofacet.pain <- pubs_nofacet.pain %>%
  select(-CD4_recent, -Age, -Pain_present, -Site, -Sex) %>%
  distinct() %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\\\.',
                                replacement = ' & ')) %>%
  mutate(Pain_site = factor(Pain_site,
                            levels = c('Head', 'Chest', 'Abdomen',
                                       'Legs', 'Knees', 'Ankles & Feet',
                                       'Thoracic spine', 'Lumbosacral spine'),
                            ordered = TRUE)) %>%

  ggplot(data = .) +
  aes(x = fct_rev(Pain_site),
      y = point_est,

```

```

    ymin = lower_ci,
    ymax = upper_ci,
    fill = Pain) +
geom_linerange(position = position_dodge2(width = 0.6),
               size = 1,
               colour = '#000000') +
geom_point(shape = 21,
           colour = '#000000',
           position = position_dodge2(width = 0.6),
           size = 4,
           stroke = 1) +
coord_flip() +
labs(title = 'Body sites by pain chronicity',
     subtitle = '(Point estimate with 95%CI)',
     y = 'Proportion of participants with pain') +
scale_y_continuous(limits = c(0, 0.6)) +
scale_fill_manual(values = c('#FFFFFF', '#000000')) +
theme_minimal(base_size = 18) +
theme(legend.position = c(0.8, 0.1),
      legend.title = element_blank(),
      plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      axis.title.y = element_blank(),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))

#-- Body sites by age --#
pubs_nofacet.age <- data_geq10 %>%
  select(-CD4_recent, -Site, -Sex, -Pain) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\\\.',
                                replacement = ' & ')) %>%
  group_by(Pain_site, Pain_present) %>%
  nest() %>%
  mutate(boot = map(.x = data,
                    ~ boot(data = .x,
                           statistic = median_func,
                           R = 999,
                           stype = 'i',
                           parallel = 'multicore',
                           ncpus = 4))) %>%

# Get CI
mutate(ci = map(.x = boot,
               ~ boot.ci(.x, type = 'perc')))) %>%

# Extract ci data
mutate(point_est = map(.x = ci,
                      ~ .x$t0),
       lower_ci = map(.x = ci,

```

```

        ~ .x$percent[[4]]),
    upper_ci = map(.x = ci,
        ~ .x$percent[[5]])) %>%

# Remove columns
select(-boot, -ci) %>%
# Unnest
unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
ungroup() %>%
mutate(Pain_site = factor(Pain_site,
    levels = c('Head', 'Chest', 'Abdomen', 'Groin',
        'Legs', 'Knees', 'Ankles & Feet',
        'Thoracic spine', 'Lumbosacral spine'),
    ordered = TRUE)) %>%

select(-Age) %>%
distinct()

plot_nofacet.age <- pubs_nofacet.age %>%
ggplot(data = .) +
  aes(x = fct_rev(Pain_site),
      y = point_est,
      ymin = lower_ci,
      ymax = upper_ci,
      fill = Pain_present) +
  geom_linerange(position = position_dodge2(width = 0.6),
      size = 1,
      colour = '#000000') +
  geom_point(shape = 21,
      colour = '#000000',
      position = position_dodge2(width = 0.6),
      size = 4,
      stroke = 1) +
  coord_flip() +
  labs(title = 'Body sites by median age',
      subtitle = '(Point estimate with 95%CI)',
      y = 'Age (years)') +
  scale_fill_manual(values = c('#FFFFFF', '#000000'),
      labels = c('No pain', 'Pain present')) +
  scale_y_continuous(limits = c(31, 45),
      breaks = c(31, 35, 39, 43)) +
  theme_minimal(base_size = 18) +
  theme(legend.position = c(0.8, 0.95),
      legend.title = element_blank(),
      plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      axis.title.y = element_blank(),
      axis.text.y = element_blank(),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))

#-- Body sites by CD4 --#
pubs_nofacet.cd4 <- data_geq10 %>%

```

```

select(-Age, -Site, -Sex, -Pain) %>%
mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\\\.',
                                replacement = ' & ')) %>%
group_by(Pain_site, Pain_present) %>%
nest() %>%
mutate(boot = map(.x = data,
                  ~ boot(data = .x,
                        statistic = median_func,
                        R = 999,
                        stype = 'i',
                        parallel = 'multicore',
                        ncpus = 4))) %>%

# Get CI
mutate(ci = map(.x = boot,
                ~ boot.ci(.x, type = 'perc')))) %>%
# Extract ci data
mutate(point_est = map(.x = ci,
                      ~ .x$t0),
       lower_ci = map(.x = ci,
                      ~ .x$percent[[4]]),
       upper_ci = map(.x = ci,
                      ~ .x$percent[[5]])) %>%

# Remove columns
select(-boot, -ci) %>%
# Unnest
unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
ungroup() %>%
mutate(Pain_site = factor(Pain_site,
                          levels = c('Head', 'Chest', 'Abdomen',
                                      'Legs', 'Knees', 'Ankles & Feet',
                                      'Thoracic spine', 'Lumbosacral spine'),
                          ordered = TRUE)) %>%

select(-CD4_recent) %>%
distinct()

plot_nofacet.cd4 <- pubs_nofacet.cd4 %>%
ggplot(data = .) +
aes(x = fct_rev(Pain_site),
    y = point_est,
    ymin = lower_ci,
    ymax = upper_ci,
    fill = Pain_present) +
geom_linerange(position = position_dodge2(width = 0.6),
               size = 1,
               colour = '#000000') +
geom_point(shape = 21,
            colour = '#000000',
            position = position_dodge2(width = 0.6),
            size = 4,

```

```

        stroke = 1) +
coord_flip() +
labs(title = 'Body sites by median CD4 T-cell count',
      subtitle = '(Point estimate with 95%CI)',
      y = expression('CD4 T-cell count (cell.mm'-3')) +
scale_fill_manual(values = c('#FFFFFF', '#000000'),
                  labels = c('No pain', 'Pain present')) +
scale_y_continuous(limits = c(100, 500),
                  breaks = c(100, 200, 300, 400, 500)) +
theme_minimal(base_size = 18) +
theme(legend.position = c(0.8, 0.95),
      legend.title = element_blank(),
      plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      axis.title.y = element_blank(),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))

# Patchwork
pub_plot <- plot_nofacet + plot_nofacet.sex + plot_nofacet.pain +
  plot_nofacet.age + plot_nofacet.cd4 +
  plot_layout(ncol = 2)

# Save
ggsave(filename = 'figures/figure_supp1.png',
       width = 14,
       height = 18)

```

## 9.2 Body sites with a point estimate pain proportion <10%

```

# Set seed
set.seed(2020)

# Generate filter
filter_less10 <- filter(filter, prop < 0.1) %>%
  .$Site

# Generate new dataset
data_less10 <- data_log %>%
  pivot_longer(cols = -c(Site, CD4_recent, Sex, Age, Pain),
               names_to = 'Pain_site',
               values_to = 'Pain_present') %>%
  # Filter by filter_geq10
  filter(Pain_site %in% filter_less10)

pubs_nofacet2 <- data_less10 %>%
  group_by(Pain_site) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,

```

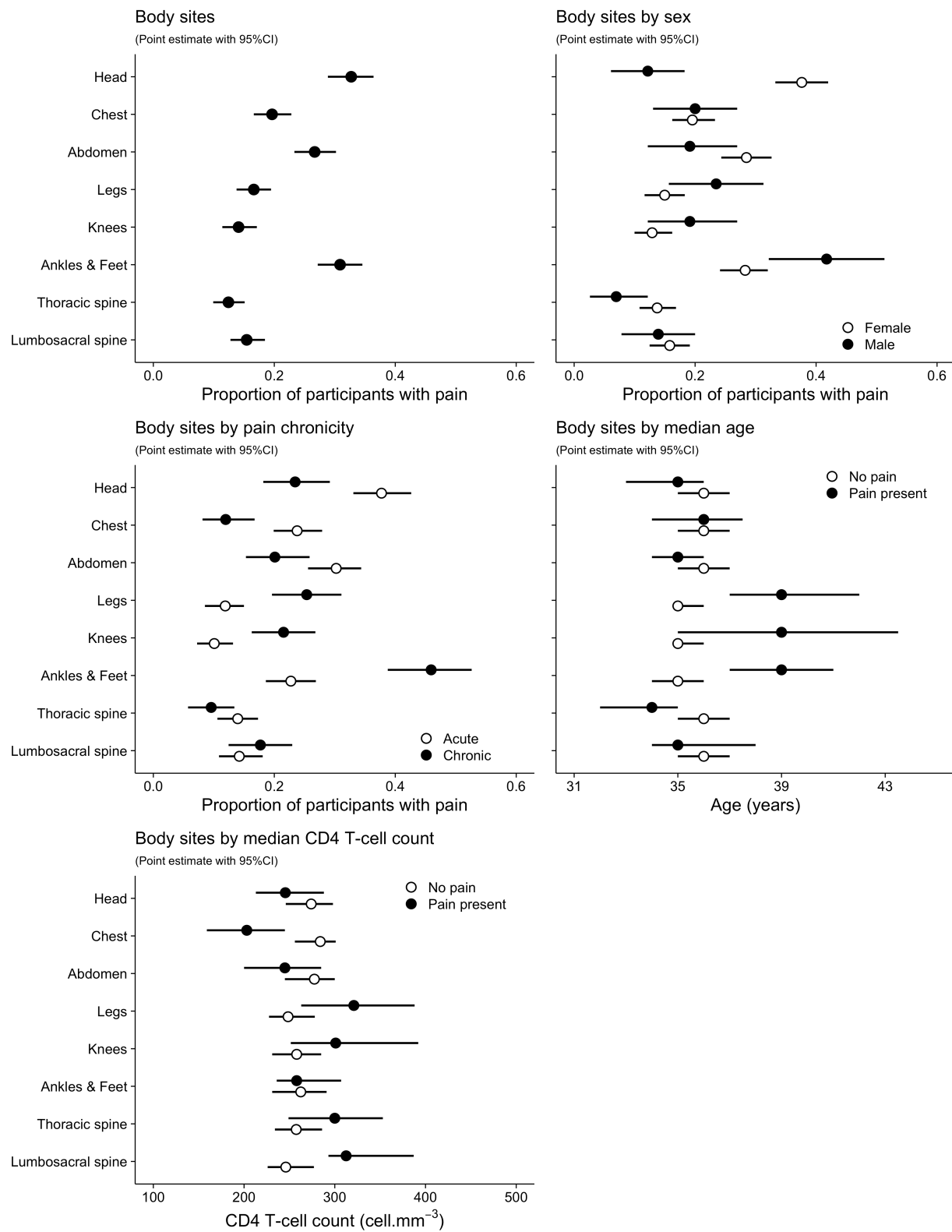


Figure 2: Body sites with a point estimate pain proportion >10%



```

      ~ boot(data = .x[, 'Pain_present'],
              statistic = prop_func,
              R = 999,
              stype = 'i',
              parallel = 'multicore',
              ncpus = 4))) %>%

# Get CI
mutate(ci = map(.x = boot,
                ~ boot.ci(.x, type = 'perc')))) %>%
# Extract ci data
mutate(point_est = map(.x = ci,
                      ~ .x$t0),
       lower_ci = map(.x = ci,
                      ~ .x$percent[[4]]),
       upper_ci = map(.x = ci,
                      ~ .x$percent[[5]])) %>%

# Remove columns
select(-boot, -ci) %>%
# Unnest
unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
ungroup()

plot_nofacet2 <- pubs_nofacet2 %>%
  select(-CD4_recent, -Sex, -Age, -Pain_present, -Site, -Pain) %>%
  distinct() %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\\\. ',
                                replacement = ' & ')) %>%
  mutate(Pain_site = ifelse(Pain_site == 'Lower back',
                            yes = 'Lower back/flanks',
                            no = Pain_site)) %>%
  mutate(Pain_site = factor(Pain_site,
                            levels = c('Throat', 'Shoulder', 'Arms',
                                       'Elbows', 'Wrists & Hands',
                                       'Lower back/flanks', 'Groin', 'Hips',
                                       'Buttocks', 'Cervical spine'),
                            ordered = TRUE)) %>%

ggplot(data = .) +
  aes(x = fct_rev(Pain_site),
      y = point_est,
      ymin = lower_ci,
      ymax = upper_ci) +
  geom_linerange(size = 1) +
  geom_point(size = 5) +
  coord_flip() +
  labs(title = 'Body sites',
       subtitle = '(Body sites with <10% pain; Point estimate with 95%CI)',
       y = 'Proportion of participants with pain') +
  scale_y_continuous(limits = c(0, 0.3)) +
  theme_minimal(base_size = 18) +

```

```

theme(plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      axis.title.y = element_blank(),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))

#-- Body sites by sex --#
pubs_nofacet.sex2 <- data_less10 %>%
  group_by(Pain_site, Sex) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,
                    ~ boot(data = .x[, 'Pain_present'],
                          statistic = prop_func,
                          R = 999,
                          stype = 'i',
                          parallel = 'multicore',
                          ncpus = 4))) %>%

  # Get CI
  mutate(ci = map(.x = boot,
                  ~ boot.ci(.x, type = 'perc')))) %>%
  # Extract ci data
  mutate(point_est = map(.x = ci,
                        ~ .x$t0),
         lower_ci = map(.x = ci,
                        ~ .x$percent[[4]]),
         upper_ci = map(.x = ci,
                        ~ .x$percent[[5]])) %>%
  # Remove columns
  select(-boot, -ci) %>%
  # Unnest
  unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
  ungroup()

plot_nofacet.sex2 <- pubs_nofacet.sex2 %>%
  select(-CD4_recent, -Age, -Pain_present, -Site, -Pain) %>%
  distinct() %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\\\.',
                                replacement = ' & ')) %>%
  mutate(Pain_site = ifelse(Pain_site == 'Lower back',
                            yes = 'Lower back/flanks',
                            no = Pain_site)) %>%
  mutate(Pain_site = factor(Pain_site,
                            levels = c('Throat', 'Shoulder', 'Arms',
                                         'Elbows', 'Wrists & Hands',
                                         'Lower back/flanks', 'Groin', 'Hips',
                                         'Buttocks', 'Cervical spine'),

```

```

                                ordered = TRUE)) %>%

ggplot(data = .) +
  aes(x = fct_rev(Pain_site),
      y = point_est,
      ymin = lower_ci,
      ymax = upper_ci,
      fill = Sex) +
  geom_linerange(position = position_dodge2(width = 0.6),
                size = 1,
                colour = '#000000') +
  geom_point(shape = 21,
            colour = '#000000',
            position = position_dodge2(width = 0.6),
            size = 4,
            stroke = 1) +
  coord_flip() +
  labs(title = 'Body sites by sex',
       subtitle = '(Body sites with <10% pain; Point estimate with 95%CI)',
       y = 'Proportion of participants with pain') +
  scale_y_continuous(limits = c(0, 0.3)) +
  scale_fill_manual(values = c('#FFFFFF', '#000000')) +
  theme_minimal(base_size = 18) +
  theme(legend.position = c(0.8, 0.1),
        legend.title = element_blank(),
        plot.title = element_text(size = 18),
        plot.subtitle = element_text(size = 12),
        axis.title.y = element_blank(),
        axis.text.y = element_blank(),
        panel.grid = element_blank(),
        axis.text = element_text(colour = '#000000'),
        axis.line = element_line(size = 0.5),
        axis.ticks = element_line(size = 0.5))

#-- Body sites by pain chronicity --#
pubs_nofacet.pain2 <- data_less10 %>%
  group_by(Pain_site, Pain) %>%
  nest() %>%
  # Bootstrap data
  mutate(boot = map(.x = data,
                    ~ boot(data = .x[, 'Pain_present'],
                          statistic = prop_func,
                          R = 999,
                          stype = 'i',
                          parallel = 'multicore',
                          ncpus = 4))) %>%

  # Get CI
  mutate(ci = map(.x = boot,
                  ~ boot.ci(.x, type = 'perc')))) %>%
  # Extract ci data
  mutate(point_est = map(.x = ci,
                        ~ .x$t0),
         lower_ci = map(.x = ci,
                        ~ .x$percent[[4]]),

```

```

    upper_ci = map(.x = ci,
                    ~ .x$percent[[5]]) %>%
# Remove columns
select(-boot, -ci) %>%
# Unnest
unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
ungroup()

plot_nofacet.pain2 <- pubs_nofacet.pain2 %>%
select(-CD4_recent, -Age, -Pain_present, -Site, -Sex) %>%
distinct() %>%
mutate(Pain_site = str_replace(Pain_site,
                              pattern = '_',
                              replacement = ' ')) %>%
mutate(Pain_site = str_replace(Pain_site,
                              pattern = '\\.',
                              replacement = ' & ')) %>%
mutate(Pain_site = ifelse(Pain_site == 'Lower back',
                          yes = 'Lower back/flanks',
                          no = Pain_site)) %>%
mutate(Pain_site = factor(Pain_site,
                          levels = c('Throat', 'Shoulder', 'Arms',
                                      'Elbows', 'Wrists & Hands',
                                      'Lower back/flanks', 'Groin', 'Hips',
                                      'Buttocks', 'Cervical spine'),
                          ordered = TRUE)) %>%

ggplot(data = .) +
aes(x = fct_rev(Pain_site),
    y = point_est,
    ymin = lower_ci,
    ymax = upper_ci,
    fill = Pain) +
geom_linerange(position = position_dodge2(width = 0.6),
               size = 1,
               colour = '#000000') +
geom_point(shape = 21,
            colour = '#000000',
            position = position_dodge2(width = 0.6),
            size = 4,
            stroke = 1) +
coord_flip() +
labs(title = 'Body sites by pain chronicity',
     subtitle = '(Body sites with <10% pain; Point estimate with 95%CI)',
     y = 'Proportion of participants with pain') +
scale_y_continuous(limits = c(0, 0.3)) +
scale_fill_manual(values = c('#FFFFFF', '#000000')) +
theme_minimal(base_size = 18) +
theme(legend.position = c(0.8, 0.1),
      legend.title = element_blank(),
      plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      axis.title.y = element_blank(),
      panel.grid = element_blank(),

```

```

    axis.text = element_text(colour = '#000000'),
    axis.line = element_line(size = 0.5),
    axis.ticks = element_line(size = 0.5))

#-- Body sites by age --#
pubs_nofacet.age2 <- data_less10 %>%
  select(-CD4_recent, -Site, -Sex, -Pain) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\\\.',
                                replacement = ' & ')) %>%
  group_by(Pain_site, Pain_present) %>%
  nest() %>%
  mutate(boot = map(.x = data,
                    ~ boot(data = .x,
                           statistic = median_func,
                           R = 999,
                           stype = 'i',
                           parallel = 'multicore',
                           ncpus = 4))) %>%

# Get CI
  mutate(ci = map(.x = boot,
                  ~ boot.ci(.x, type = 'perc')))) %>%

# Extract ci data
  mutate(point_est = map(.x = ci,
                        ~ .x$t0),
         lower_ci = map(.x = ci,
                        ~ .x$percent[[4]]),
         upper_ci = map(.x = ci,
                        ~ .x$percent[[5]])) %>%

# Remove columns
  select(-boot, -ci) %>%

# Unnest
  unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
  ungroup() %>%
  mutate(Pain_site = ifelse(Pain_site == 'Lower back',
                           yes = 'Lower back/flanks',
                           no = Pain_site)) %>%
  mutate(Pain_site = factor(Pain_site,
                           levels = c('Throat', 'Shoulder', 'Arms',
                                       'Elbows', 'Wrists & Hands',
                                       'Lower back/flanks', 'Groin', 'Hips',
                                       'Buttocks', 'Cervical spine'),
                           ordered = TRUE)) %>%

  select(-Age) %>%
  distinct()

plot_nofacet.age2 <- pubs_nofacet.age2 %>%
  ggplot(data = .) +
  aes(x = fct_rev(Pain_site),
      y = point_est,

```

```

    ymin = lower_ci,
    ymax = upper_ci,
    fill = Pain_present) +
geom_linerange(position = position_dodge2(width = 0.6),
               size = 1,
               colour = '#000000') +
geom_point(shape = 21,
           colour = '#000000',
           position = position_dodge2(width = 0.6),
           size = 4,
           stroke = 1) +
coord_flip() +
labs(title = 'Body sites by median age',
     subtitle = '(Body sites with <10% pain; Point estimate with 95%CI)',
     y = 'Age (years)') +
scale_fill_manual(values = c('#FFFFFF', '#000000'),
                  labels = c('No pain', 'Pain present')) +
scale_y_continuous(limits = c(27, 47),
                   breaks = c(27, 31, 35, 39, 43, 47)) +
theme_minimal(base_size = 18) +
theme(legend.position = c(0.8, 0.1),
      legend.title = element_blank(),
      plot.title = element_text(size = 18),
      plot.subtitle = element_text(size = 12),
      axis.title.y = element_blank(),
      axis.text.y = element_blank(),
      panel.grid = element_blank(),
      axis.text = element_text(colour = '#000000'),
      axis.line = element_line(size = 0.5),
      axis.ticks = element_line(size = 0.5))

#-- Body sites by CD4 --#
pubs_nofacet.cd42 <- data_less10 %>%
  select(-Age, -Site, -Sex, -Pain) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '_',
                                replacement = ' ')) %>%
  mutate(Pain_site = str_replace(Pain_site,
                                pattern = '\\.',
                                replacement = ' & ')) %>%
  group_by(Pain_site, Pain_present) %>%
  nest() %>%
  mutate(boot = map(.x = data,
                    ~ boot(data = .x,
                           statistic = median_func,
                           R = 999,
                           stype = 'i',
                           parallel = 'multicore',
                           ncpus = 4))) %>%

# Get CI
mutate(ci = map(.x = boot,
                ~ boot.ci(.x, type = 'perc')))) %>%
# Extract ci data

```

```

mutate(point_est = map(.x = ci,
  ~ .x$t0),
  lower_ci = map(.x = ci,
  ~ .x$percent[[4]]),
  upper_ci = map(.x = ci,
  ~ .x$percent[[5]])) %>%

# Remove columns
select(-boot, -ci) %>%
# Unnest
unnest(cols = c(data, point_est, lower_ci, upper_ci)) %>%
ungroup() %>%
mutate(Pain_site = ifelse(Pain_site == 'Lower back',
  yes = 'Lower back/flanks',
  no = Pain_site)) %>%
mutate(Pain_site = factor(Pain_site,
  levels = c('Throat', 'Shoulder', 'Arms',
    'Elbows', 'Wrists & Hands',
    'Lower back/flanks', 'Groin', 'Hips',
    'Buttocks', 'Cervical spine'),
  ordered = TRUE)) %>%

select(-CD4_recent) %>%
distinct()

plot_nofacet.cd42 <- pubs_nofacet.cd42 %>%
ggplot(data = .) +
aes(x = fct_rev(Pain_site),
  y = point_est,
  ymin = lower_ci,
  ymax = upper_ci,
  fill = Pain_present) +
geom_linerange(position = position_dodge2(width = 0.6),
  size = 1,
  colour = '#000000') +
geom_point(shape = 21,
  colour = '#000000',
  position = position_dodge2(width = 0.6),
  size = 4,
  stroke = 1) +
coord_flip() +
labs(title = 'Body sites by median CD4 T-cell count',
  subtitle = '(Body sites with <10% pain; Point estimate with 95%CI)',
  y = expression('CD4 T-cell count (cell.mm-3)')) +
scale_fill_manual(values = c('#FFFFFF', '#000000'),
  labels = c('No pain', 'Pain present')) +
scale_y_continuous(limits = c(50, 600),
  breaks = c(100, 200, 300, 400, 500, 600)) +
theme_minimal(base_size = 18) +
theme(legend.position = c(0.8, 0.95),
  legend.title = element_blank(),
  plot.title = element_text(size = 18),
  plot.subtitle = element_text(size = 12),
  axis.title.y = element_blank(),
  panel.grid = element_blank(),

```

```
axis.text = element_text(colour = '#000000'),  
axis.line = element_line(size = 0.5),  
axis.ticks = element_line(size = 0.5))
```

```
# Patchwork
```

```
pub_plot2 <- plot_nofacet2 + plot_nofacet.sex2 + plot_nofacet.pain2 +  
  plot_nofacet.age2 + plot_nofacet.cd42 +  
  plot_layout(ncol = 2)
```

```
# Save
```

```
ggsave(filename = 'figures/figure_supp2.png',  
  width = 14,  
  height = 18)
```



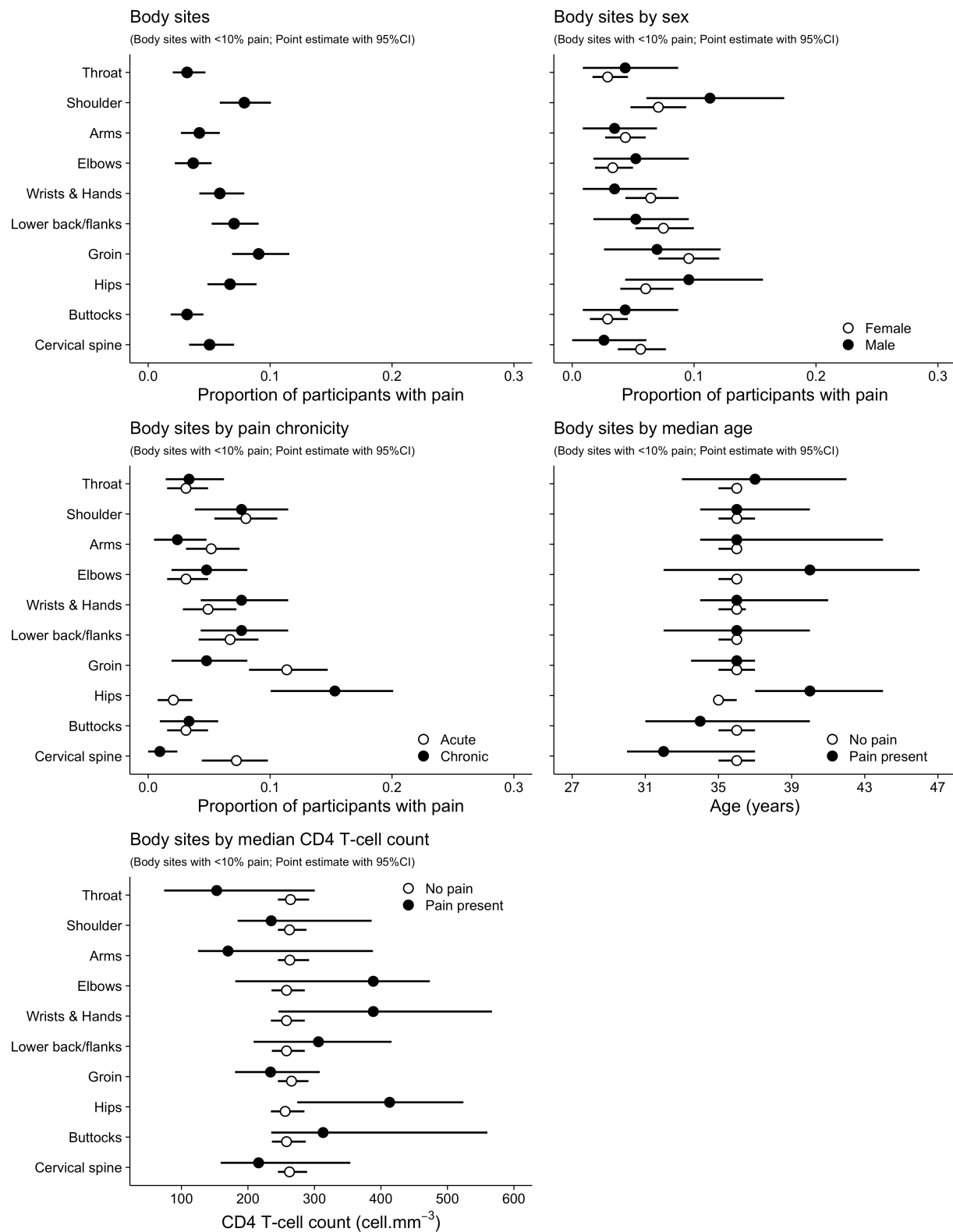


Figure 3: Body sites with a point estimate pain proportion <10%

## 10 Session information

```
sessionInfo()
```

```
## R version 4.0.2 (2020-06-22)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Catalina 10.15.5
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] patchwork_1.0.1 lme4_1.1-23      Matrix_1.2-18    boot_1.3-25
## [5] knitr_1.29       skimr_2.1.1      forcats_0.5.0    stringr_1.4.0
## [9] dplyr_1.0.0      purrr_0.3.4      readr_1.3.1      tidyr_1.1.0
## [13] tibble_3.0.1     ggplot2_3.3.2    tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
## [1] Rcpp_1.0.4.6      lubridate_1.7.9  lattice_0.20-41  utf8_1.1.4
## [5] assertthat_0.2.1  digest_0.6.25    R6_2.4.1         cellranger_1.1.0
## [9] repr_1.1.0        backports_1.1.8  reprex_0.3.0     evaluate_0.14
## [13] highr_0.8         httr_1.4.1       pillar_1.4.4     rlang_0.4.6
## [17] readxl_1.3.1      rstudioapi_0.11  minqa_1.2.4      nloptr_1.2.2.1
## [21] blob_1.2.1        rmarkdown_2.3    labeling_0.3     splines_4.0.2
## [25] statmod_1.4.34    munsell_0.5.0    broom_0.5.6      compiler_4.0.2
## [29] modelr_0.1.8      xfun_0.15        pkgconfig_2.0.3  base64enc_0.1-3
## [33] htmltools_0.5.0   tidyselect_1.1.0 fansi_0.4.1      crayon_1.3.4
## [37] dbplyr_1.4.4      withr_2.2.0      MASS_7.3-51.6    grid_4.0.2
## [41] nlme_3.1-148      jsonlite_1.6.1   gtable_0.3.0     lifecycle_0.2.0
## [45] DBI_1.1.0         magrittr_1.5     scales_1.1.1     cli_2.0.2
## [49] stringi_1.4.6     farver_2.0.3     fs_1.4.1         xml2_1.3.2
## [53] ellipsis_0.3.1    generics_0.0.2   vctrs_0.3.1      tools_4.0.2
## [57] glue_1.4.1        hms_0.5.3        parallel_4.0.2   yaml_2.2.1
## [61] colorspace_1.4-1  rvest_0.3.5      haven_2.3.1
```