Physical activity and mental health in Canadian seniors: A cross-sectional study using the Canadian Community Health Survey

Authors: *Reza Hosseini* and *Alexandra Roine*, MPH students at UBC
December 8, 2021

SYNOPSIS

The Canadian Physical Activity Guidelines (CPAG) suggest that seniors participate in 150 minutes of moderate to vigorous exercise each week; however, Canadian seniors have not been meeting these guidelines. Exercise improves physical and mental health, and maintenance of these throughout ageing is important. The impact of exercise participation on the mental health of Canadian seniors remains unknown. This study aimed to investigate the association between exercise participation and mental health in Canadian seniors using data from the 2017-2018 Canadian Community Health Survey (CCHS). Bivariate and logistic regression models were used to investigate the effect of exercise participation on perceived mental health. CCHS respondents over 60 years old (N=25,806) were included for analysis. Respondents physically active below CPAG were older, and less likely to be male, injured, or employed than respondents who met CPAG. The unadjusted odds of a respondent who met CPAG reporting Excellent-Good mental health were 1.38 (95% CI: 1.24, 1.54) times that of a respondent below CPAG. After adjusting for age, sex, race, injury, and employment, the odds ratio was 1.56 (95% CI: 1.37, 1.78). Sensitivity analysis showed missingness did not impact effect measures meaningfully. Employment status was an effect modifier, and the adjusted odds ratios for employed and unemployed respondents were 1.09 (95% CI: 0.85, 1.40) and 1.81 (95% CI: 1.55, 2.12), respectively. Findings suggest that exercise participation at CPAG levels by Canadians aged 60 years and above may improve perceived mental health.

LOAD LIBs & DATA

Libraries

```
library("checkpoint")
checkpoint("2021-01-01")  # checkpoint used to ensure reproducibility

library("dplyr")
```

```r
library("ggplot2")
library("tidyr")
library("forcats")
library("tableone")        # Used for creating table one
library("kableExtra")      # Used to present table one in markdown
library("Publish")         # Used for displaying regression outputs
```

Data

You can download the CCHS 2017-18 dataset **here**.

It has **113,290** observations (respondents) and **1051** variables:

```r
load("cchs1718_original.RData")

dim(cchs1718_original)
```

```
## [1] 113290    1051
```

CREATE ANALYT. DATASET

Selecting variables

Only 7 out of 1051 variables from the original CCHS dataset are selected.

```r
cchs_withExclusions <- cchs1718_original %>%
  select(GEN_015,  # Outcome: Perceived mental health
         PAADVACV, # Exposure: Physical activity indicator
         DHHGAGE,  # Age
         DHH_SEX,  # Sex
         SDCDGCGT, # Race
         INJDVSTT, # Injuries
         MAC_010,  # Employment status
         )
```

Redefining levels of variables

```r
cchs_withExclusions <- cchs_withExclusions %>%

  # Outcome: Perceived mental health
  mutate(pMentalHealth = case_when(GEN_015 %in% c("Fair",
                                                  "Poor") ~ "Fair-Poor",
                                   GEN_015 %in% c("Excellent",
```

```r
                                                "Very good",
                                                "Good") ~ "Excellent-Good")) %>%

  # Explanatory variable: Physical activity indicator
  mutate(physicalAct = case_when(
    PAADVACV == paste0("Physically active below ",
                       "recommended level from CPAG") ~ "below act. level",
    PAADVACV == paste0("Physically active at ",
                       "/ above recommended ",
                       "level from CPAG") ~ "above act. level")) %>%

  # Covariate: Age
  mutate(age = case_when(DHHGAGE == "Age between 12 and 14" ~ "12-14",
                         DHHGAGE == "Age between 15 and 17" ~ "15-17",
                         DHHGAGE == "Age between 18 and 19" ~ "18-19",
                         DHHGAGE == "Age between 20 and 24" ~ "20-24",
                         DHHGAGE == "Age between 25 and 29" ~ "25-29",
                         DHHGAGE == "Age between 30 and 34" ~ "30-34",
                         DHHGAGE == "Age between 35 and 39" ~ "35-39",
                         DHHGAGE == "Age between 40 and 44" ~ "40-44",
                         DHHGAGE == "Age between 45 and 49" ~ "45-49",
                         DHHGAGE == "Age between 50 and 54" ~ "50-54",
                         DHHGAGE == "Age between 55 and 59" ~ "55-59",
                         DHHGAGE == "Age between 60 and 64" ~ "60-64",
                         DHHGAGE == "Age between 65 and 69" ~ "65-69",
                         DHHGAGE == "Age between 70 and 74" ~ "70-74",
                         DHHGAGE == "Age between 75 and 79" ~ "75-79",
                         DHHGAGE == "Age 80 and older" ~ "80+")) %>%

  # Covariate: Sex
  rename(sex = DHH_SEX) %>%

  # Covariate: Race
  mutate(race = case_when(
    SDCDGCGT == "White" ~ "White",
    SDCDGCGT == "Non-white (Aboriginal or Other Visible Minority)" ~ "Non-white")) %>%

  # Covariate: Injuries:
  mutate(injury = case_when(
    INJDVSTT == "No injuries" ~ "No injuries",
    INJDVSTT == "Injury limiting activities only" ~ "Limiting activities only",
    INJDVSTT == "Treated injury (not limiting activities) only" ~
      "Treated injury only",
    INJDVSTT == "Injury limiting activities and treated injury" ~
      "Limiting activities and treated injury")) %>%

  # Covariate: Employment status
  mutate(employed = case_when(MAC_010 == "No" ~ "No",
                              MAC_010 == "Yes" ~ "Yes"))
```

Dropping unwanted vars.

We only keep the redefined variables:

```
cchs_withExclusions <- cchs_withExclusions %>%
  select(pMentalHealth, physicalAct, age, sex, race,
         injury, employed)
```

Making as factor

We convert all variables into factors to specify our desired reference levels.

```
cchs_withExclusions <- cchs_withExclusions %>%
  mutate(pMentalHealth = factor(pMentalHealth, levels = c("Fair-Poor",
                                                          "Excellent-Good"))) %>%
  mutate(physicalAct = factor(physicalAct,
                              levels = c("below act. level",
                                         "above act. level"))) %>%
  mutate(age = factor(age, levels = c("12-14", "15-17", "18-19",
                                      "20-24", "25-29", "30-34",
                                      "35-39", "40-44", "45-49",
                                      "50-54", "55-59", "60-64",
                                      "65-69", "70-74", "75-79",
                                      "80+"))) %>%
  mutate(sex = factor(sex, levels = c("Female", "Male"))) %>%
  mutate(race = factor(race, levels = c("White", "Non-white"))) %>%
  mutate(injury = factor(injury,
                         levels = c("No injuries",
                                    "Limiting activities only",
                                    "Treated injury only",
                                    "Limiting activities and treated injury"))) %>%
  mutate(employed = factor(employed, levels = c("No", "Yes")))
```

Inclusion/exclusion status

Only those respondents above 60 years old are included in our study. Additionally, all observations with missing values for either the explanatory variable (physical activity) or the outcome (perceived mental health) will be removed from the analytic sample.

Coding included/excluded

```
cchs_withExclusions <- cchs_withExclusions %>%
  mutate(inclusion = case_when((age %in% c("60-64", "65-69", "70-74",
                                           "75-79", "80+") &
                                  !is.na(physicalAct) &
                                  !is.na(pMentalHealth)) ~ "included",
                               TRUE ~ "excluded"))

table(cchs_withExclusions$inclusion, useNA = "ifany")
```

```
##
## excluded included
##    87484    25806
```

4

Included vs. excluded

Included and excluded groups were not different with regard to sex (included: 53.9% female; excluded: 53.6% female; p = 0.51); however, excluded respondents were more likely to be of non-white racial background (15.3% vs. 5.6%, p < 0.001), employed (77.9% vs. 39.3%, p < 0.001), and have activity-limiting injuries (15.1% vs. 10.5%, p < 0.001).

```
test_result <- chisq.test(table("Inclusion" = cchs_withExclusions$inclusion,
                                "Sex" = cchs_withExclusions$sex))
test_result
```

Sex

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table(Inclusion = cchs_withExclusions$inclusion, Sex = cchs_withExclusions$sex)
## X-squared = 0.42712, df = 1, p-value = 0.5134
```

```
test_result$expected
```

```
##           Sex
## Inclusion   Female     Male
##    excluded 47018.5 40465.5
##    included 13869.5 11936.5
```

```
prop.table(test_result$observed, 1)
```

```
##           Sex
## Inclusion      Female       Male
##    excluded 0.5369210 0.4630790
##    included 0.5392544 0.4607456
```

```
test_result <- chisq.test(table("Inclusion" = cchs_withExclusions$inclusion,
                                "Race" = cchs_withExclusions$race))
test_result
```

Race

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table(Inclusion = cchs_withExclusions$inclusion, Race = cchs_withExclusions$race)
## X-squared = 1596.3, df = 1, p-value < 2.2e-16
```

```
test_result$expected
```

```
##          Race
## Inclusion     White Non-white
##    excluded 69833.86 10511.136
##    included 21485.14  3233.864
```

```
prop.table(test_result$observed, 1)
```

```
##          Race
## Inclusion      White  Non-white
##    excluded 0.84611363 0.15388637
##    included 0.94413204 0.05586796
```

```
temp <- cchs_withExclusions %>%
  mutate(injury = case_when(
    injury %in% c("Limiting activities only",
                  "Limiting activities and treated injury") ~ "Limiting act.",
    injury %in% c("No injuries", "Treated injury only") ~ "Not limiting act."))

test_result <- chisq.test(table("Inclusion" = temp$inclusion,
                                "Race" = temp$injury))
test_result
```

Injury

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table(Inclusion = temp$inclusion, Race = temp$injury)
## X-squared = 354.63, df = 1, p-value < 2.2e-16
```

```
test_result$expected
```

```
##          Race
## Inclusion  Limiting act. Not limiting act.
##    excluded      12302.252          74984.75
##    included       3631.748          22136.25
```

```
prop.table(test_result$observed, 1)
```

```
##           Race
## Inclusion  Limiting act. Not limiting act.
##    excluded      0.1515346          0.8484654
##    included      0.1050528          0.8949472
```

```
test_result <- chisq.test(table("Inclusion" = cchs_withExclusions$inclusion,
                                 "Race" = cchs_withExclusions$employed))
test_result
```

Employment status

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table(Inclusion = cchs_withExclusions$inclusion, Race = cchs_withExclusions$employed)
## X-squared = 11073, df = 1, p-value < 2.2e-16
```

```
test_result$expected
```

```
##           Race
## Inclusion          No       Yes
##    excluded 22682.125 52360.88
##    included  6014.875 13885.12
```

```
prop.table(test_result$observed, 1)
```

```
##           Race
## Inclusion          No       Yes
##    excluded 0.2214863 0.7785137
##    included 0.6068342 0.3931658
```

Applying the inclusion criteria

```
cchs <- cchs_withExclusions %>%
  filter(inclusion == "included") %>%
  droplevels() %>%    # dropping extra levels
  select(-inclusion)  # dropping extra variables
```

Level counts

The number of respondents in each level of our variable is as below:

```
table(cchs$pMentalHealth, useNA = "ifany")      # Outcome: Perceived mental health
```

```
##
##      Fair-Poor Excellent-Good
##           1350          24456
```

```
table(cchs$physicalAct, useNA = "ifany")        # Explanatory var: Physical activity
```

```
##
## below act. level above act. level
##             9156            16650
```

```
table(cchs$age, useNA = "ifany")                # Age
```

```
##
## 60-64 65-69 70-74 75-79   80+
##  7426  7453  5117  3089  2721
```

```
table(cchs$sex, useNA = "ifany")                # Sex
```

```
##
## Female   Male
##  13916  11890
```

```
table(cchs$race, useNA = "ifany")               # Race
```

```
##
##      White Non-white      <NA>
##      23338      1381      1087
```

```
table(cchs$injury, useNA = "ifany")             # Injuries
```

```
##
##                     No injuries              Limiting activities only
##                           21966                                  2554
##             Treated injury only Limiting activities and treated injury
##                            1095                                   153
##                            <NA>
##                              38
```

```
table(cchs$employed, useNA = "ifany")           # Employment status
```

```
##
##    No   Yes  <NA>
## 12076  7824  5906
```

STUDY POPULATION

Of the 113,290 observations collected in the 2017-2018 CCHS, the final analytic sample included 25,806 (22.7%) observations from respondents over 60 years old who provided valid responses to explanatory and outcome variables. Excluded observations consisted of 70,715 (62.4%) observations from respondents under 60 years old, 16,030 (14.1%) invalid observations for whether a respondent met CPAG recommendations, which include participants who did not report physical activity minutes, and 739 invalid observations (0.7%) for perceived mental health status.

```
cchs_exclusionType <- cchs_withExclusions %>%
  mutate(inclusionType = case_when(age %in% c("60-64", "65-69", "70-74",
                                              "75-79", "80+") ~ "included",
                                   TRUE ~ "excluded_age"))

cchs_exclusionType <- cchs_exclusionType %>%
  mutate(inclusionType = case_when(inclusionType == "excluded_age" ~ "excluded_age",
                                   !is.na(physicalAct) ~ "included",
                                   TRUE ~ "excluded_physicalAct"))

cchs_exclusionType <- cchs_exclusionType %>%
  mutate(inclusionType = case_when(inclusionType == "excluded_age" ~ "excluded_age",
                                   inclusionType == "excluded_physicalAct" ~
                                                    "excluded_physicalAct",
                                   !is.na(pMentalHealth) ~ "included",
                                   TRUE ~ "excluded_pMentalHealth"))

cchs_exclusionType %>%
  count(inclusionType) %>%
  mutate(Percentage = round(n / sum(n) * 100, 1))
```

```
## # A tibble: 4 x 3
##   inclusionType              n Percentage
##   <chr>                  <int>      <dbl>
## 1 excluded_age           70715       62.4
## 2 excluded_physicalAct   16030       14.1
## 3 excluded_pMentalHealth   739        0.7
## 4 included               25806       22.8
```

Within the analytic sample, n = 9,156 (35.5%) respondents did not meet CPAG recommendations, while n = 16,650 (64.5%) were physically active at or above CPAG levels:

```
cchs %>%
  count(physicalAct) %>%
  mutate(Percentage = round(n / sum(n) * 100, 1))
```

```
## # A tibble: 2 x 3
##   physicalAct          n Percentage
##   <fct>            <int>      <dbl>
## 1 below act. level  9156       35.5
## 2 above act. level 16650       64.5
```

After re-coding CCHS perceived mental health answer options into two categories (Excellent-Good, including "excellent", "very good", and "good", and Fair-Poor, including "fair" and "poor"), 24,456 (94.8%) respondents reported Excellent-Good and 1,350 (5.2%) respondents reported Fair-Poor mental health:

```
cchs %>%
  count(pMentalHealth) %>%
  mutate(Percentage = round(n / sum(n) * 100, 1))
```

```
## # A tibble: 2 x 3
##   pMentalHealth      n Percentage
##   <fct>          <int>      <dbl>
## 1 Fair-Poor       1350        5.2
## 2 Excellent-Good 24456       94.8
```

The Fair-Poor perceived mental health group comprised 6.30% of those that did not meet CPAG recommendations, and 4.64% of those who participated in physical activity at or above recommended CPAG levels:

```
round(prop.table(table("Physical Activity" = cchs$physicalAct,
                       "Perceived Mental Health" = cchs$pMentalHealth,
                       useNA = "ifany"),
                 1) *100, 2)
```

```
##                   Perceived Mental Health
## Physical Activity  Fair-Poor Excellent-Good
##    below act. level      6.30          93.70
##    above act. level      4.64          95.36
```

TABLE ONE

Comparing Respondents who did and did not meet CPAG Recommendations:

Respondents who were physically active below CPAG recommendations were older (see **Table 1** below; p < 0.001), less likely to be male (40.0% vs. 49.4% male; p < 0.001), less likely to be injured (p < 0.001), and less likely to be employed (36.7% vs. 40.6%; p < 0.001) than were respondents who were physically active at or above CPAG recommendations. Respondents were similar in terms of racial background (5.5% vs. 5.7% non-white, including Aboriginal or other visible minority racial background; p = 0.509).

```
table_one_object <- CreateTableOne(vars = c("age", "sex", "race",
                                            "injury", "employed"),
                                   strata = "physicalAct",
                                   data = cchs,
                                   testApprox = chisq.test,
                                   argsApprox = list(correct = TRUE),
                                   testExact = fisher.test,
                                   testNormal = oneway.test,
                                   argsNormal = list(var.equal = FALSE),
                                   testNonNormal = kruskal.test)

table_one_printed <- print(table_one_object,
                           catDigits = 1,
                           contDigits = 1,
                           pDigits = 3,
                           # Supress printing in console:
                           printToggle = F)

# Creating a dataframe of the printed table one:
table_one <- as.data.frame(table_one_printed) %>%
  add_rownames(var = "Variable") %>%
  select(-test)
```

```
table_one %>%
  mutate(Variable = cell_spec(Variable,
                              italic = ifelse(row_number() %in% c(3:7, 11:14),
                                              T, F))) %>%
  kable(escape = FALSE,
        caption = paste(c("Table 1: Participant characteristics: Canadian",
                          "Community Health Survey participants aged 60",
                          "years and above who did and did not participate",
                          "in physical activity as recommended by the",
                          "Canadian Physical Activity Guidelines (CPAG)."),
                        collapse = "")) %>%
  add_indent(c(3:7, 11:14)) %>%
  kable_paper("hover", full_width = TRUE)
```

(UN)ADJUSTED ANALYSES

Here are our unadjusted and adjusted models based on complete cases. The unadjusted odds of a respondent who was physically active at or above recommended CPAG levels reporting Excellent-Good mental health were 1.38 (95% confidence interval, CI: 1.24, 1.54) times that of a respondent who was physically active below recommended CPAG levels. After adjusting for age, sex, racial background, injury, and employment status, the odds ratio (OR) increased to 1.56 (95% CI: 1.37, 1.78).

Table 1: Table 1: Participant characteristics: CanadianCommunity Health Survey participants aged 60years and above who did and did not participatein physical activity as recommended by theCanadian Physical Activity Guidelines (CPAG).

Variable	below act. level	above act. level	p
n	9156	16650	
age (%)			<0.001
60-64	2247 (24.5)	5179 (31.1)	
65-69	2399 (26.2)	5054 (30.4)	
70-74	1863 (20.3)	3254 (19.5)	
75-79	1231 (13.4)	1858 (11.2)	
80+	1416 (15.5)	1305 (7.8)	
sex = Male (%)	3660 (40.0)	8230 (49.4)	<0.001
race = Non-white (%)	477 (5.5)	904 (5.7)	0.509
injury (%)			<0.001
No injuries	7903 (86.4)	14063 (84.6)	
Limiting activities only	852 (9.3)	1702 (10.2)	
Treated injury only	349 (3.8)	746 (4.5)	
Limiting activities and treated injury	38 (0.4)	115 (0.7)	
employed = Yes (%)	2380 (36.7)	5444 (40.6)	<0.001

Bivariable Analysis

```
model <- glm(pMentalHealth ~ physicalAct,
         family = binomial(link = logit),
         data = cchs)
publish(model)
```

```
##    Variable             Units OddsRatio       CI.95 p-value
## physicalAct below act. level      Ref
##             above act. level     1.38 [1.24;1.54]  <1e-04
```

Multivariable Analysis

```
model <- glm(pMentalHealth ~ physicalAct + age + sex + race +
                      injury + employed,
         family = binomial(link = logit), data = cchs)
publish(model)
```

```
##    Variable                           Units OddsRatio       CI.95   p-value
## physicalAct              below act. level      Ref
##                          above act. level     1.56 [1.37;1.78]  < 1e-04
##      age                           60-64      Ref
##                                    65-69     1.81 [1.56;2.10]  < 1e-04
##                                    70-74     2.72 [2.25;3.27]  < 1e-04
##      sex                          Female      Ref
##                                     Male     1.10 [0.96;1.25]  0.167725
```

```
##        race                                      White       Ref
##                                              Non-white    0.82 [0.63;1.07]   0.145659
##      injury                                 No injuries     Ref
##                                 Limiting activities only    0.50 [0.42;0.59]    < 1e-04
##                                   Treated injury only     0.91 [0.66;1.24]   0.536891
##                      Limiting activities and treated injury  0.42 [0.23;0.75]   0.003639
##      employed                                        No     Ref
##                                                     Yes     1.97 [1.70;2.28]    < 1e-04
```

EVALUAT. MISSINGNESS

Amount of missingness

The number of missing values and the percentage of the dataset missing for each variable is:

```
missing_data <- data.frame(n_missing = colSums(is.na(cchs)),
                           nrow = nrow(cchs)) %>%
  add_rownames() %>%
  mutate(pct_missing = round((n_missing/nrow)*100, 1))

missing_data
```

```
## # A tibble: 7 x 4
##   rowname       n_missing  nrow pct_missing
##   <chr>             <dbl> <int>       <dbl>
## 1 pMentalHealth         0 25806           0
## 2 physicalAct           0 25806           0
## 3 age                   0 25806           0
## 4 sex                   0 25806           0
## 5 race               1087 25806         4.2
## 6 injury               38 25806         0.1
## 7 employed           5906 25806        22.9
```

Dummy-coded missingness

We added the missing values as another factor level to the dataset:

```
cchs_missing <- cchs %>%
  mutate_if(is.factor,
            fct_explicit_na,
            na_level = "Missing")
```

An example to show what happened after dummy coding was applied:

```
table(cchs$race, useNA="ifany")
```

```
##
##    White Non-white     <NA>
##    23338    1381     1087
```

```
table(cchs_missing$race, useNA="ifany")
```

```
##
##    White Non-white  Missing
##    23338    1381     1087
```

Sensitivity analysis

As confounding variables racial background (4.2% missing), injury (0.1% missing), and employment status (22.9% missing) contained missingness, a sensitivity analysis was performed to assess how the relationship between physical activity and perceived mental health status would change once missingness was included in the model with dummy coding:

```
model <- glm(pMentalHealth ~ physicalAct + age + sex + race +
                             injury + employed,
             family = binomial(link = logit), data = cchs_missing)
publish(model)
```

```
##    Variable                                Units OddsRatio        CI.95     p-value
## physicalAct                      below act. level       Ref
##                                  above act. level      1.42  [1.27;1.60]    < 1e-04
##      age                                   60-64       Ref
##                                            65-69      1.75  [1.52;2.02]    < 1e-04
##                                            70-74      2.56  [2.14;3.06]    < 1e-04
##                                            75-79      0.37  [0.05;2.68]   0.3252058
##                                              80+      0.24  [0.03;1.71]   0.1538558
##      sex                                  Female       Ref
##                                             Male      0.97  [0.87;1.09]   0.6169584
##     race                                  White       Ref
##                                       Non-white      0.79  [0.63;1.00]   0.0478628
##                                          Missing      0.60  [0.48;0.75]    < 1e-04
##    injury                           No injuries       Ref
##                            Limiting activities only   0.51  [0.44;0.60]    < 1e-04
##                              Treated injury only      0.79  [0.61;1.02]   0.0702304
##           Limiting activities and treated injury     0.40  [0.24;0.67]   0.0004708
##                                          Missing      0.23  [0.10;0.52]   0.0005065
##    employed                                  No       Ref
##                                              Yes      1.97  [1.71;2.27]    < 1e-04
##                                          Missing      7.54 [1.05;54.04]   0.0443673
```

Adjusting for age category, sex, racial background, injury, and employment status, a respondent who met CPAG recommendations was found to have 1.42 (95% CI: 1.27, 1.60) times the odds of reporting having Excellent-Good mental health status than those of a respondent who did not meet CPAG recommendations. The calculated effect measure stayed almost the same after considering missingness. Therefore, we continued with complete case analysis.

EFFECT MODIFICATION

Employment status was assessed for effect modification. If respondents were employed and met CPAG recommendations, they had 1.09 (95% CI: 0.85, 1.40) times the adjusted odds of reporting Excellent-Good mental health than those who did not meet recommendations. For those who were unemployed and who met CPAG recommendations, the OR was 1.81 (95% CI: 1.55, 2.12).

```
data_employedYes <- cchs %>%
  filter(employed == "Yes")

data_employedNo <- cchs %>%
  filter(employed == "No")



model_employedYes <- glm(pMentalHealth ~ physicalAct + age + sex + race +
                                      injury,
                    family = binomial(link = logit),
                    data = data_employedYes)
model_employedNo <- glm(pMentalHealth ~ physicalAct + age + sex + race +
                                      injury,
                    family = binomial(link = logit),
                    data = data_employedNo)
```

For those employed:

```
publish(model_employedYes)
```

```
##     Variable                                    Units OddsRatio       CI.95     p-value
## physicalAct                       below act. level      Ref
##                                   above act. level      1.09 [0.85;1.40]   0.5192917
##       age                                    60-64      Ref
##                                              65-69      1.58 [1.21;2.07]   0.0008293
##                                              70-74      2.38 [1.48;3.84]   0.0003619
##       sex                                   Female      Ref
##                                               Male      1.36 [1.08;1.72]   0.0092084
##      race                                    White      Ref
##                                          Non-white      1.08 [0.66;1.78]   0.7591004
##    injury                              No injuries      Ref
##                           Limiting activities only      0.52 [0.38;0.70]    < 1e-04
##                                Treated injury only      1.36 [0.71;2.58]   0.3543500
##             Limiting activities and treated injury      0.58 [0.21;1.63]   0.3031385
```

For those unemployed:

```
publish(model_employedNo)
```

```
##     Variable                                    Units OddsRatio        CI.95   p-value
##  physicalAct                            below act. level       Ref
##                                         above act. level      1.81 [1.55;2.12]   < 1e-04
##          age                                     60-64       Ref
##                                                  65-69      1.93 [1.62;2.31]   < 1e-04
##                                                  70-74      2.86 [2.33;3.51]   < 1e-04
##          sex                                    Female       Ref
##                                                   Male      0.99 [0.85;1.16]   0.91938
##         race                                     White       Ref
##                                               Non-white      0.72 [0.53;0.98]   0.03508
##       injury                              No injuries       Ref
##                               Limiting activities only      0.50 [0.40;0.62]   < 1e-04
##                                    Treated injury only      0.78 [0.55;1.12]   0.18021
##                  Limiting activities and treated injury      0.36 [0.17;0.74]   0.00566
```

FOREST PLOT

Odds ratios of the five logistic regression models resulting from exploration for missingness and effect modification are depicted in the figure below:

```r
label <- c("Unadjusted model (A)",
           "Adjusted - Complete case analysis (B)",
           "Adjusted - Dummy-coded missingness (C)",
           "EM: Employed (D)",
           "EM: Unemployed (E)")
mean  <- c(1.38, 1.56, 1.42, 1.09, 1.81)
lower <- c(1.24, 1.37, 1.27, 0.85, 1.55)
upper <- c(1.54, 1.78, 1.60, 1.40, 2.12)

df <- data.frame(label, mean, lower, upper)

# reverses the factor level ordering for labels after coord_flip()
df$label <- factor(df$label, levels=rev(df$label))

fp <- ggplot(data=df, aes(x=label, y=mean, ymin=lower, ymax=upper)) +
  geom_pointrange() +
  geom_text(aes(label=mean),
            position = position_dodge(width = .9),
            vjust = -1,
            size = 3) +
  geom_hline(yintercept=1, lty=2) +  # add a dotted line at x=1 after flip
  coord_flip() +                     # flip coordinates (puts labels on y axis)
  xlab("Logistic regression models") + ylab("Odds ratio (95% CI)") +
  theme_bw()
```
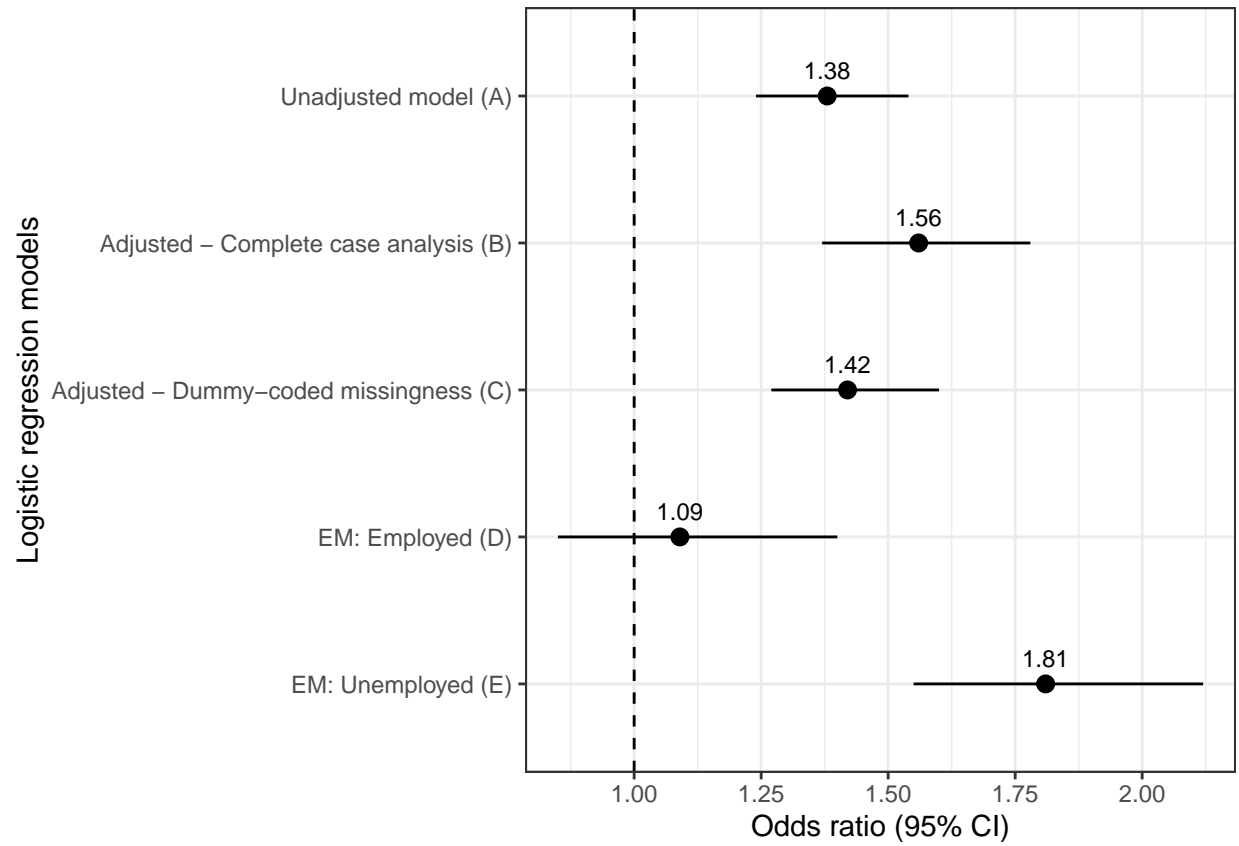
```
print(fp)
```


The end!