

Propensity score (an application to COVID-19 maternal mortality data)

21/06/2022

About the database and R packages used

This research studies some demographic, clinical and mortality variables of COVID-19 pandemic aiming to compare pregnant and non-pregnant childbearing age (10 to 49 years) women living in Espírito Santo/BR. The search was limited to the first notified case of COVID-19 in February 2020 until April, 2021 because after this month the vaccination of pregnant and postpartum women started in Brazil. The dataset were obtained on July 19, 2021, on the site <https://coronavirus.es.gov.br/painel-covid-19-es>. The data are loaded next.

The data are analyzed using the free-software R (<https://www.R-project.org>) in version 4.0.3. Next, we present and load the libraries used in the data analysis process.

```
loadlibrary <- function(x) {  
  if (!require(x, character.only = TRUE)) {  
    install.packages(x, dependencies = T)  
    if (!require(x, character.only = TRUE))  
      stop("Package not found")  
  }  
}  
  
packages <-  
c(  
  "dplyr",  
  "lubridate",  
  "readr",  
  "readxl",  
  "ggplot2",  
  "kableExtra",  
  "tables",  
  "questionr",  
  "car",  
  "data.table",  
  "magrittr",  
  "readxl",  
  "summarytools",  
  "modelsummary",  
  "zoo",  
  "strex",  
  "WeightIt",  
  "MatchIt",  
  "cobalt",  
  "RIttools"
```

```

)
lapply(packages, loadlibrary)

### Data of Espírito Santo/BR #####
dados <- read_delim("MICRODADOS-19-07-2021.csv",
                    ";", quote = "\\\"", escape_double = FALSE,
                    locale = locale(encoding = "ISO-8859-1"),
                    trim_ws = TRUE)

#functions for summary measures
MEAN <- function(x) mean(x, na.rm = TRUE)
MEDIAN <- function(x) median(x, na.rm = TRUE)
SDev <- function(x) sd(x, na.rm = TRUE)
MIN <- function(x) base::min(x, na.rm = TRUE)
MAX <- function(x) base::max(x, na.rm = TRUE)
q25 <- function(x) stats::quantile(x, p=0.25, na.rm=TRUE)
q75 <- function(x) stats::quantile(x, p=0.75, na.rm=TRUE)
IQR <- function(x) round(q75(x)-q25(x), 2)
n <- function(x) sum(!is.na(x))

```

There are 1742179 observations in the database. Each line in this dataset corresponds to a data from February,2020 to July, 2021. Below, we present the variables considered in the study for each data.

```
## [1] Inf
```

Case selection

First, we select only confirmed cases.

```

dados <- dados %>%
  mutate(
    Classification = case_when(
      Classificacao == "Confirmados" ~ "Confirmed", #confirmed cases of COVID-19
      Classificacao == "Descartados" ~ "Discarded", #discarded cases of COVID-19
      Classificacao == "Suspeito" ~ "Suspect", #suspected cases of COVID-19
      TRUE ~ NA_character_
    )
  )

questionr::freq(
  dados$Classification,
  cum = FALSE,
  total = TRUE,
  na.last = TRUE,
  valid = FALSE
) %>%
  kable(caption = "Frequency table of classification",
        digits = 2) %>%
  kable_styling(latex_options = "HOLD_position")

```

Table 1: Frequency table of classification

	n	%
Confirmed	533846	30.6
Discarded	930071	53.4
Suspect	278262	16.0
Total	1742179	100.0

```
# Filtering only confirmed cases of COVID-19
dados1 <- dados %>%
  filter(Classification == 'Confirmed')

dados1 <- dados1 %>%
  mutate(
    PCR_results = case_when(
      ResultadoRT_PCR == "Inconclusivo" ~ "Inconclusive",
      ResultadoRT_PCR == "Não Informado" ~ "Not informed ",
      ResultadoRT_PCR == "Negativo" ~ "Negative",
      ResultadoRT_PCR == "Positivo" ~ "Positive",
      TRUE ~ NA_character_
    )
  )

questionr::freq(
  dados1$PCR_results,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
) %>%
  kable(caption = "Results of PCR test of the confirmed cases",
        digits = 2) %>%
  kable_styling(latex_options = "HOLD_position")
```

Table 2: Results of PCR test of the confirmed cases

	n	%
Inconclusive	960	0.2
Negative	5714	1.1
Not informed	229572	43.0
Positive	297600	55.7
Total	533846	100.0

```
dados1 <- dados1 %>%
  mutate(
    rapid_tests = case_when(
      TipoTesteRapido == "Ignorado" ~ "Unknown",
      TipoTesteRapido == "Não Informado" ~ "Not informed",
      TipoTesteRapido == "Teste rápido Antígeno" ~ "Antigen test",
      TipoTesteRapido == "Teste rápido IgM/IgG" ~ "IgM/IgG",
      TRUE ~ NA_character_
    )
  )
```

```

    )
  )

questionr::freq(
  dados1$rapid_tests,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
) %>%
kable(caption = "Types of rapid tests of the confirmed cases",
      digits = 2) %>%
kable_styling(latex_options = "HOLD_position")

```

Table 3: Types of rapid tests of the confirmed cases

	n	%
Antigen test	110440	20.7
IgM/IgG	40327	7.6
Not informed	378033	70.8
Unknown	5046	0.9
Total	533846	100.0

```

dados1 <- dados1 %>%
  mutate(
    rapid_results = case_when(
      ResultadoTesteRapido == "Inconclusivo" ~ "Inconclusive",
      ResultadoTesteRapido == "Não Informado" ~ "Not informed ",
      ResultadoTesteRapido == "Negativo" ~ "Negative",
      ResultadoTesteRapido == "Positivo" ~ "Positive",
      TRUE ~ NA_character_
    )
  )

questionr::freq(
  dados1$rapid_results,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
) %>%
kable(caption = "Results of rapid tests of the confirmed cases",
      digits = 2) %>%
kable_styling(latex_options = "HOLD_position")

```

Table 4: Results of rapid tests of the confirmed cases

	n	%
Inconclusive	242	0.0
Negative	12947	2.4
Not informed	325563	61.0
Positive	195094	36.5
Total	533846	100.0

In this search, we consider only observations with positive PCR or positive antigen test for COVID-19.

```
# Filtering only positive cases of COVID-19 detected by PCR or antigen test
dados1_1 <- dados1 %>%
  filter(PCR_results == 'Positive' | (rapid_tests == "Antigen test" & rapid_results == 'Positive'))
```

Considering only people with confirmed cases of COVID-19 by PCR or antigen test, the following table presents the distribution of the sex variable.

```
dados1_1 <- dados1_1 %>%
  mutate(
    Sex = case_when(
      Sexo == "F" ~ "Female",
      Sexo == "M" ~ "Male",
      Sexo == "I" ~ "Not informed",
      TRUE ~ NA_character_
    )
  )

questionr::freq(
  dados1_1$Sex,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
) %>%
  kable(caption = "Frequency table of sex variable",
        digits = 2) %>%
  kable_styling(latex_options = "HOLD_position")
```

Table 5: Frequency table of sex variable

	n	%
Female	215555	54.1
Male	182732	45.9
Not informed	81	0.0
Total	398368	100.0

The next selection consists of selecting females cases:

```
#Filtrando só as pessoas de sexo feminino
dados2 <- dados1_1 %>%
  filter(Sex == 'Female')
```

The next step is to identify pregnant and postpartum people (variable `classi_gesta_puerp`) in the female group.

```
#Creating the variable pregnant (any trimester or ignored gestational age) and non-pregnant
dados2 <- dados2 %>%
  mutate(pregnant_SN = case_when(
    Gestante == "1º trimestre" ~ "yes",
    Gestante == "2º trimestre" ~ "yes",
    Gestante == "3º trimestre" ~ "yes",
    Gestante == "Idade gestacional ignorada" ~ "yes",
    Gestante == "Não" ~ "no",
    TRUE ~ NA_character_
  ))
```

```
questionr::freq(
  dados2$pregnant_SN ,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
) %>%
  kable(caption = "Frequency table of pregnancy variable",
        digits = 2) %>%
  kable_styling(latex_options = "HOLD_position")
```

Table 6: Frequency table of pregnancy variable

	n	%
no	211836	98.3
yes	1922	0.9
NA	1797	0.8
Total	215555	100.0

Now, we only select women who are pregnant or not.

```
#Filtering only the cases that indicate whether the woman is pregnant or not:
dados3 <- dados2 %>%
  filter(!is.na(pregnant_SN)==TRUE)
```

Now we selected only women in childbearing (10-49).

```
#selecionando os casos de 10 a 49 anos:
dados4 <- dados3 %>%
  filter(FaixaEtaria != "0 a 4 anos" &
    FaixaEtaria != "05 a 9 anos" &
    FaixaEtaria != "50 a 59 anos" &
```

```

FaixaEtaria != "60 a 69 anos" &
FaixaEtaria != "70 a 79 anos" &
FaixaEtaria != "80 a 89 anos" &
FaixaEtaria != "90 anos ou mais" &
FaixaEtaria != "-")

```

```

dados4 <- dados4 %>%
  mutate(age_group = case_when(
    FaixaEtaria == "10 a 19 anos" ~ "10 - 19",
    FaixaEtaria == "20 a 29 anos" ~ "20 - 29",
    FaixaEtaria == "30 a 39 anos" ~ "30 - 39",
    FaixaEtaria == "40 a 49 anos" ~ "40 - 49",
    FaixaEtaria == "Não" ~ "no",
    TRUE ~ NA_character_)
  )

questionr::freq(
  dados4$age_group,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
) %>%
  kable(caption = "Frequency table of age group",
        digits = 2) %>%
  kable_styling(latex_options = "HOLD_position")

```

Table 7: Frequency table of age group

	n	%
10 - 19	13521	9.2
20 - 29	39539	26.9
30 - 39	51218	34.9
40 - 49	42565	29.0
Total	146843	100.0

Now we are going to select only the finalized cases (cases with information about cure or death).

```

#Filtering only finalized cases
dados5 <- dados4 %>%
  filter(Evolucao != "-" & Evolucao != "Ignorado")

dados5 <- dados5 %>%
  mutate(evol = case_when(
    Evolucao == "Cura" ~ "cure",
    Evolucao == "Óbito pelo COVID-19" ~ "death by COVID-19",
    Evolucao == "Óbito por outras causas" ~ "death from other causes",
    TRUE ~ NA_character_)
  )

with(dados5, ctable(pregnant_SN, evol))

```

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * evol
## Data Frame: dados5
##
## -----
##           evol           cure  death by COVID-19  death from other causes  Total
## pregnant_SN
##           no       141227 ( 99.6%)       571 (0.4%)       10 (0.007%)  141808 (100.0%)
##           yes        1870 ( 99.3%)        13 (0.7%)         0 (0.000%)   1883 (100.0%)
##           Total    143097 ( 99.6%)       584 (0.4%)       10 (0.007%)  143691 (100.0%)
## -----
```

Creating the evolution variable presenting death from other causes and death by COVID-19 as death.

```
dados5 <- dados5 %>%
  mutate(evolution = case_when(
    Evolucao == "Óbito por outras causas" ~ "death",
    Evolucao == "Óbito pelo COVID-19" ~ "death",
    Evolucao == "Cura" ~ "cure"
  ))

with(dados5, ctable(pregnant_SN, evolution, OR=TRUE, useNA = "no", chisq = TRUE))
```

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * evolution
## Data Frame: dados5
##
## -----
##           evolution           cure           death           Total
## pregnant_SN
##           no       141227 ( 99.6%)       581 (0.4%)       141808 (100.0%)
##           yes        1870 ( 99.3%)        13 (0.7%)        1883 (100.0%)
##           Total    143097 ( 99.6%)       594 (0.4%)       143691 (100.0%)
## -----
##
## -----
## Chi.squared   df   p.value
## -----
##      2.9071     1   0.0882
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##      1.69      0.97      2.93
## -----
```

The search was limited to the first notified case of COVID-19 in February 2020 until April, 2021 because after this month the vaccination of pregnant and postpartum women started in Brazil.


```

dados5 <- dados5 %>%
  dplyr::mutate(
    dt_notific = as.Date(DataNotificacao, format = "%d/%m/%Y"),
    dt_diag = as.Date(DataDiagnostico, format = "%d/%m/%Y"),
    year = lubridate::year(dt_diag),
    month = lubridate::month(dt_diag),
    dif_diag_notific = as.numeric(dt_notific - dt_diag)
  )

dados5 <- dados5 %>%
  dplyr::filter(
    (year == 2020) | (year == 2021 & month <= 4)
  )

with(dados5, table(month, year))

```

```

##      year
## month 2020 2021
##    1      2 12491
##    2      1  8735
##    3     95 20054
##    4    1735 13383
##    5    5205    0
##    6    6411    0
##    7    5572    0
##    8    3814    0
##    9    4581    0
##   10    8359    0
##   11   15303    0
##   12   15868    0

```

The data to be analyzed in this search consists of women, pregnant or not, with confirmed cases of COVID-19 by PCR and/or antigen test, aged 10 to 49, with finalized outcome (death/cure).

Demographic characteristics

Education

```

dados5 <- dados5 %>%
  mutate(esc_resumida = case_when(Escolaridade == "1ª a 4ª série incompleta do EF (antigo primário ou 1º grau)" ~ "Incompleta",
    Escolaridade == "4ª série completa do EF (antigo primário ou 1º grau)" ~ "Incompleta",
    Escolaridade == "5ª à 8ª série incompleta do EF (antigo ginásio ou 1º grau)" ~ "Incompleta",
    Escolaridade == "Analfabeto" ~ "No schooling",
    Escolaridade == "Educação superior completa" ~ "College",
    Escolaridade == "Educação superior incompleta" ~ "Middle school",
    Escolaridade == "Ensino fundamental completo (antigo ginásio ou 1º grau)" ~ "Elementary school",
    Escolaridade == "Ensino médio completo (antigo colegial ou 2º grau )" ~ "Middle school",
    Escolaridade == "Ensino médio incompleto (antigo colegial ou 2º grau )" ~ "Elementary school",
    Escolaridade == "Ignorado" ~ NA_character_,
    Escolaridade == "Não se aplica" ~ NA_character_,
  ))

```

```

TRUE ~ NA_character_))

dados5$esc_resumida <- factor(dados5$esc_resumida,
                             levels = c("No schooling", "Incomplete elementary", "Elementary", "Middle school",

## Ethnicity
dados5 <- dados5 %>%
  mutate(ethn = case_when(RacaCor == "Amarela" ~ "Yellow",
                           RacaCor == "Branca" ~ "White",
                           RacaCor == "Indigena" ~ "Indigenous",
                           RacaCor == "Parda" ~ "Brown",
                           RacaCor == "Preta" ~ "Black",
                           RacaCor == "Ignorado" ~ NA_character_,
                           TRUE ~ NA_character_))

```

Symptoms

```

## Fever
dados5 <- dados5 %>%
  mutate(fever = case_when(Febre == "Sim" ~ "yes",
                           Febre == "Não" ~ "no",
                           TRUE ~ NA_character_))

## Respiratory difficult
dados5 <- dados5 %>%
  mutate(RespDif = case_when(DificuldadeRespiratoria == "Sim" ~ "yes",
                              DificuldadeRespiratoria == "Não" ~ "no",
                              TRUE ~ NA_character_))

## Cough
dados5 <- dados5 %>%
  mutate(cough = case_when(Tosse == "Sim" ~ "yes",
                            Tosse == "Não" ~ "no",
                            TRUE ~ NA_character_))

# Coryza
dados5 <- dados5 %>%
  mutate(coryza = case_when(Coriza == "Sim" ~ "yes",
                             Coriza == "Não" ~ "no",
                             TRUE ~ NA_character_))

# Sore throat
dados5 <- dados5 %>%
  mutate(sore_throat = case_when(DorGarganta == "Sim" ~ "yes",
                                  DorGarganta == "Não" ~ "no",
                                  TRUE ~ NA_character_))

# Diarrhea
dados5 <- dados5 %>%
  mutate(diarrhea = case_when(Diarreia == "Sim" ~ "yes",
                               Diarreia == "Não" ~ "no",

```

```

TRUE ~ NA_character_))

# Headache
dados5 <- dados5 %>%
  mutate(headache = case_when(Cefaleia == "Sim" ~ "yes",
                              Cefaleia == "Não" ~ "no",
                              TRUE ~ NA_character_))

```

Comorbidities

```

# Pneumologic
dados5 <- dados5 %>%
  mutate(pneumo = case_when(ComorbidadePulmao == "Sim" ~ "yes",
                            ComorbidadePulmao == "Não" ~ "no",
                            TRUE ~ NA_character_))

# Cardiovascular
dados5 <- dados5 %>%
  mutate(cardio = case_when(ComorbidadeCardio == "Sim" ~ "yes",
                            ComorbidadeCardio == "Não" ~ "no",
                            TRUE ~ NA_character_))

# Renal
dados5 <- dados5 %>%
  mutate(renal = case_when(ComorbidadeRenal == "Sim" ~ "yes",
                           ComorbidadeRenal == "Não" ~ "no",
                           TRUE ~ NA_character_))

# Diabetes
dados5 <- dados5 %>%
  mutate(diabetes = case_when(ComorbidadeDiabetes == "Sim" ~ "yes",
                              ComorbidadeDiabetes == "Não" ~ "Não",
                              TRUE ~ NA_character_))

# Smoker
dados5 <- dados5 %>%
  mutate(smoker = case_when(ComorbidadeTabagismo == "Sim" ~ "yes",
                             ComorbidadeTabagismo == "Não" ~ "no",
                             TRUE ~ NA_character_))

# Obesity
dados5 <- dados5 %>%
  mutate(obesity = case_when(ComorbidadeObesidade == "Sim" ~ "yes",
                              ComorbidadeObesidade == "Não" ~ "no",
                              TRUE ~ NA_character_))

```

Outcomes

```
# Hospitalized
dados5 <- dados5 %>%
  mutate(hospiti = case_when(FicouInternado == "Sim" ~ "yes",
                             FicouInternado == "Não" ~ "no",
                             FicouInternado == "Ignorado" ~ NA_character_,
                             FicouInternado == "Não Informado" ~ NA_character_,
                             TRUE ~ NA_character_))

# Confirmation criterion
dados5 <- dados5 %>%
  mutate(confirmatio_crit = case_when(CriterioConfirmacao == "Clinico" ~ "Clinical",
                                       CriterioConfirmacao == "Clinico Epidemiologico" ~ "Epidemiological clinica",
                                       CriterioConfirmacao == "Laboratorial" ~ "Laboratorial",
                                       TRUE ~ NA_character_))
```

Analysis of the variables

```
with(dados5, ctable(age_group, pregnant_SN, prop = "c", useNA = "no", chisq = TRUE))
```

Age group

```
## Cross-Tabulation, Column Proportions
```

```
## age_group * pregnant_SN
```

```
## Data Frame: dados5
```

```
##
```

```
##
```

```
## -----
##      pregnant_SN      no      yes      Total
## age_group
## 10 - 19      10027 ( 8.4%)    131 ( 7.8%)    10158 ( 8.4%)
## 20 - 29      32105 ( 26.8%)    825 ( 49.1%)    32930 ( 27.1%)
## 30 - 39      42397 ( 35.4%)    654 ( 39.0%)    43051 ( 35.4%)
## 40 - 49      35401 ( 29.5%)     69 ( 4.1%)    35470 ( 29.2%)
## Total      119930 (100.0%)    1679 (100.0%)    121609 (100.0%)
## -----
```

```
##
```

```
## -----
```

```
## Chi.squared  df  p.value
```

```
## -----
```

```
## 679.0865    3      0
```

```
## -----
```

```
age <- dados5$IdadeNaDataNotificacao
age_years <- str_before_nth(age, "anos", 1)
```

```
age_years <- as.numeric(age_years)
dados5 <- cbind(dados5, age_years)

datasummary((pregnant_SN) ~ age_years*(n+MEAN+SD+MEDIAN+MIN+MAX+q25+q75+IQR), data = dados5, output = 't')
```

Age in the date of notification

	n	MEAN	SD	MEDIAN	MIN	MAX	q25	q75	IQR
no	119930.00	33.25	9.37	34.00	10.00	49.00	26.00	41.00	15.00
yes	1679.00	28.40	6.40	28.00	13.00	45.00	23.00	33.00	10.00

```
# Teste t
t.test(age_years ~ pregnant_SN, data = dados5)

##
## Welch Two Sample t-test
##
## data: age_years by pregnant_SN
## t = 30.624, df = 1780.2, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group no and group yes is not equal to 0
## 95 percent confidence interval:
##  4.542448 5.164095
## sample estimates:
## mean in group no mean in group yes
##      33.24875      28.39547
```

```
# Teste de Wilcoxon
wilcox.test(age_years ~ pregnant_SN, data = dados5)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: age_years by pregnant_SN
## W = 133719359, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
questionr::freq(
  dados5$ethn,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
) %>%
  kable(caption = "Frequency table of ethnicity", digits = 2) %>%
  kable_styling(latex_options = "HOLD_position")
```

Table 9: Frequency table of ethnicity

	n	%
Black	6319	5.2
Brown	40810	33.6
Indigenous	206	0.2
White	38701	31.8
Yellow	7263	6.0
NA	28310	23.3
Total	121609	100.0

```
with(dados5, ctable(ethn, pregnant_SN, prop = "c", useNA = "no", chisq = FALSE))
```

Ethnicity /Cor

```
## Cross-Tabulation, Column Proportions
```

```
## ethn * pregnant_SN
```

```
## Data Frame: dados5
```

```
##
```

```
## -----
##           pregnant_SN           no           yes           Total
##           ethn
##           Black           6197 ( 6.8%)           122 ( 8.11%)           6319 ( 6.8%)
##           Brown           40057 ( 43.6%)           753 ( 50.03%)           40810 ( 43.7%)
##           Indigenous           205 ( 0.2%)           1 ( 0.07%)           206 ( 0.2%)
##           White           38182 ( 41.6%)           519 ( 34.49%)           38701 ( 41.5%)
##           Yellow           7153 ( 7.8%)           110 ( 7.31%)           7263 ( 7.8%)
##           Total           91794 (100.0%)           1505 (100.00%)           93299 (100.0%)
## -----
```

```
#Exact's Fisher Test
```

```
fisher.test(dados5$ethn, dados5$pregnant_SN, simulate.p.value = TRUE)
```

```
##
```

```
## Fisher's Exact Test for Count Data with simulated p-value (based on
```

```
## 2000 replicates)
```

```
##
```

```
## data: dados5$ethn and dados5$pregnant_SN
```

```
## p-value = 0.0004998
```

```
## alternative hypothesis: two.sided
```

```
questionr::freq(
  dados5$esc_resumida,
  cum = FALSE,
  total = TRUE,
  na.last = FALSE,
  valid = FALSE
```

```
) %>%
  kable(caption = "Frequency table of education", digits = 2) %>%
  kable_styling(latex_options = "HOLD_position")
```

Table 10: Frequency table of education

	n	%
No schooling	192	0.2
Incomplete elementary	8519	7.0
Elementary	10864	8.9
Middle school	38001	31.2
College	20392	16.8
NA	43641	35.9
Total	121609	100.0

```
with(dados5, ctable(esc_resumida, pregnant_SN, prop = "c", useNA = "no", chisq = FALSE))
```

Education

```
## Cross-Tabulation, Column Proportions
## esc_resumida * pregnant_SN
## Data Frame: dados5
##
## -----
##               pregnant_SN              no              yes              Total
##   esc_resumida
##   No schooling              191 ( 0.2%)              1 ( 0.08%)              192 ( 0.2%)
##   Incomplete elementary      8362 ( 10.9%)             157 ( 12.40%)             8519 ( 10.9%)
##   Elementary                10594 ( 13.8%)             270 ( 21.33%)            10864 ( 13.9%)
##   Middle school             37398 ( 48.8%)             603 ( 47.63%)           38001 ( 48.7%)
##   College                   20157 ( 26.3%)             235 ( 18.56%)           20392 ( 26.2%)
##   Total                    76702 (100.0%)            1266 (100.00%)          77968 (100.0%)
## -----
```

```
#Exact's Fisher Test
fisher.test(dados5$esc_resumida, dados5$pregnant_SN, simulate.p.value = TRUE)
```

```
##
## Fisher's Exact Test for Count Data with simulated p-value (based on
## 2000 replicates)
##
## data: dados5$esc_resumida and dados5$pregnant_SN
## p-value = 0.0004998
## alternative hypothesis: two.sided
```

Symptoms variables

```
with(dados5, ctable(pregnant_SN, fever, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Febre

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * fever
## Data Frame: dados5
##
##
## -----
##          fever          no          yes          Total
## pregnant_SN
##          no      72674 (60.7%)  47060 (39.3%)  119734 (100.0%)
##          yes      1084 (64.6%)    595 (35.4%)   1679 (100.0%)
##          Total    73758 (60.7%)  47655 (39.3%)  121413 (100.0%)
## -----
##
## -----
## Chi.squared  df  p.value
## -----
##      10.2173    1  0.0014
## -----
##
## -----
## Odds Ratio  Lo - 95%  Hi - 95%
## -----
##      0.85      0.77    0.94
## -----
```

```
with(dados5, ctable(pregnant_SN, RespDif, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Respiratory difficult

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * RespDif
## Data Frame: dados5
##
##
## -----
##          RespDif          no          yes          Total
## pregnant_SN
##          no      102146 (85.3%)  17591 (14.7%)  119737 (100.0%)
##          yes      1351 (80.5%)    328 (19.5%)   1679 (100.0%)
##          Total    103497 (85.2%)  17919 (14.8%)  121416 (100.0%)
## -----
##
## -----
## Chi.squared  df  p.value
```



```
## -----
##      30.5002      1      0
## -----
##
## -----
## Odds Ratio      Lo - 95%      Hi - 95%
## -----
##      1.41      1.25      1.59
## -----
```

```
with(dados5, ctable(pregnant_SN, cough, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Cough

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * cough
## Data Frame: dados5
##
## -----
##      cough      no      yes      Total
## pregnant_SN
##      no      54473 (45.5%)      65268 (54.5%)      119741 (100.0%)
##      yes      773 (46.1%)      905 (53.9%)      1678 (100.0%)
##      Total      55246 (45.5%)      66173 (54.5%)      121419 (100.0%)
## -----
##
## -----
## Chi.squared      df      p.value
## -----
##      0.1976      1      0.6567
## -----
##
## -----
## Odds Ratio      Lo - 95%      Hi - 95%
## -----
##      0.98      0.89      1.08
## -----
```

```
with(dados5, ctable(pregnant_SN, coryza, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Coryza

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * coryza
## Data Frame: dados5
##
```

```
##
## -----
##           coryza           no           yes           Total
## pregnant_SN
##           no           70115 (58.6%)   49628 (41.4%)   119743 (100.0%)
##           yes           901 (53.7%)    778 (46.3%)    1679 (100.0%)
##           Total        71016 (58.5%)   50406 (41.5%)   121422 (100.0%)
## -----
##
## -----
## Chi.squared  df    p.value
## -----
##      16.1175    1    1e-04
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##      1.22      1.11     1.34
## -----
```

```
with(dados5, ctable(pregnant_SN, sore_throat, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Sore Throat

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * sore_throat
## Data Frame: dados5
##
##
## -----
##           sore_throat           no           yes           Total
## pregnant_SN
##           no           76350 (63.8%)   43393 (36.2%)   119743 (100.0%)
##           yes           1207 (71.9%)    472 (28.1%)    1679 (100.0%)
##           Total        77557 (63.9%)   43865 (36.1%)   121422 (100.0%)
## -----
##
## -----
## Chi.squared  df    p.value
## -----
##      47.0359    1      0
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##      0.69      0.62     0.77
## -----
```

```
with(dados5, ctable(pregnant_SN, diarrhea, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Diarrhea

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * diarrhea
## Data Frame: dados5
##
##
## -----
##           diarrhea           no           yes           Total
## pregnant_SN
##           no           101028 (84.4%)   18714 (15.6%)   119742 (100.0%)
##           yes           1485 (88.4%)    194 (11.6%)    1679 (100.0%)
##           Total        102513 (84.4%)   18908 (15.6%)   121421 (100.0%)
## -----
##
## -----
## Chi.squared  df    p.value
## -----
##      20.5953    1        0
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##      0.71      0.61      0.82
## -----
```

```
with(dados5, ctable(pregnant_SN, headache, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Headache

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * headache
## Data Frame: dados5
##
##
## -----
##           headache           no           yes           Total
## pregnant_SN
##           no           47302 (39.5%)   72440 (60.5%)   119742 (100.0%)
##           yes           736 (43.9%)    942 (56.1%)    1678 (100.0%)
##           Total        48038 (39.6%)   73382 (60.4%)   121420 (100.0%)
## -----
##
## -----
## Chi.squared  df    p.value
```

```
## -----
##      12.9652      1      3e-04
## -----
##
## -----
## Odds Ratio      Lo - 95%      Hi - 95%
## -----
##      0.84      0.76      0.92
## -----
```

Comorbidities variables

```
with(dados5, ctable(pregnant_SN, pneumo, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Pneumologic

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * pneumo
## Data Frame: dados5
##
## -----
##      pneumo      no      yes      Total
## pregnant_SN
##      no      116739 (97.5%)      2969 (2.5%)      119708 (100.0%)
##      yes      1644 (97.9%)      35 (2.1%)      1679 (100.0%)
##      Total      118383 (97.5%)      3004 (2.5%)      121387 (100.0%)
## -----
##
## -----
## Chi.squared      df      p.value
## -----
##      0.9161      1      0.3385
## -----
##
## -----
## Odds Ratio      Lo - 95%      Hi - 95%
## -----
##      0.84      0.60      1.17
## -----
```

```
with(dados5, ctable(pregnant_SN, cardio, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Cardiovascular

```
## Cross-Tabulation, Row Proportions
```

```
## pregnant_SN * cardio
## Data Frame: dados5
##
##
## -----
##           cardio           no           yes           Total
## pregnant_SN
##           no           111793 (93.4%)   7926 (6.6%)   119719 (100.0%)
##           yes           1600 (95.3%)    79 (4.7%)    1679 (100.0%)
##           Total         113393 (93.4%)   8005 (6.6%)   121398 (100.0%)
## -----
##
## -----
## Chi.squared   df   p.value
## -----
##      9.5534      1     0.002
## -----
##
## -----
## Odds Ratio    Lo - 95%    Hi - 95%
## -----
##      0.70      0.55      0.87
## -----
```

```
with(dados5, ctable(pregnant_SN, renal, chisq=FALSE, prop="r", OR=TRUE, useNA = "no"))
```

Renal

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * renal
## Data Frame: dados5
##
##
## -----
##           renal           no           yes           Total
## pregnant_SN
##           no           119468 ( 99.8%)   250 (0.2%)   119718 (100.0%)
##           yes           1677 ( 99.9%)    2 (0.1%)    1679 (100.0%)
##           Total         121145 ( 99.8%)   252 (0.2%)   121397 (100.0%)
## -----
##
## -----
## Odds Ratio    Lo - 95%    Hi - 95%
## -----
##      0.57      0.14      2.29
## -----
```

```
#Exact's Fisher Test
fisher.test(dados5$pregnant_SN, dados5$renal, simulate.p.value = TRUE)
```

```
##
```

```
## Fisher's Exact Test for Count Data
##
## data: dados5$pregnant_SN and dados5$renal
## p-value = 0.5924
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
## 0.06858209 2.08361404
## sample estimates:
## odds ratio
## 0.5699148
```

```
with(dados5, ctable(pregnant_SN, diabetes, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Diabetes

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * diabetes
## Data Frame: dados5
##
##
## -----
##      diabetes      Não      yes      Total
## pregnant_SN
##      no      117405 (98.1%)    2312 (1.9%)    119717 (100.0%)
##      yes      1629 (97.0%)      50 (3.0%)      1679 (100.0%)
##      Total    119034 (98.1%)    2362 (1.9%)    121396 (100.0%)
## -----
##
## -----
##      Chi.squared    df    p.value
## -----
##      8.9684         1      0.0027
## -----
##
## -----
##      Odds Ratio    Lo - 95%    Hi - 95%
## -----
##      1.56          1.17        2.07
## -----
```

```
with(dados5, ctable(pregnant_SN, smoker, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Smoker

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * smoker
## Data Frame: dados5
```

```
##
##
## -----
##          smoker          no          yes          Total
## pregnant_SN
##          no          118711 (99.2%)    1003 (0.8%)    119714 (100.0%)
##          yes          1668 (99.3%)     11 (0.7%)     1679 (100.0%)
##          Total        120379 (99.2%)    1014 (0.8%)    121393 (100.0%)
## -----
##
## -----
## Chi.squared  df  p.value
## -----
##    0.4648      1  0.4954
## -----
##
## -----
## Odds Ratio  Lo - 95%  Hi - 95%
## -----
##    0.78      0.43     1.42
## -----
```

```
with(dados5, ctable(pregnant_SN, obesity, chisq=TRUE, prop="r", OR=TRUE, useNA = "no"))
```

Obesity

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * obesity
## Data Frame: dados5
##
##
## -----
##          obesity          no          yes          Total
## pregnant_SN
##          no          116836 (97.6%)    2813 (2.4%)    119649 (100.0%)
##          yes          1653 (98.5%)     25 (1.5%)     1678 (100.0%)
##          Total        118489 (97.7%)    2838 (2.3%)    121327 (100.0%)
## -----
##
## -----
## Chi.squared  df  p.value
## -----
##    5.0018      1  0.0253
## -----
##
## -----
## Odds Ratio  Lo - 95%  Hi - 95%
## -----
##    0.63      0.42     0.93
## -----
```

Outcome analysis

```
with(dados5, ctable(pregnant_SN, hospi, prop = "r", useNA = "no", chisq = FALSE, OR=TRUE))
```

Hospitalized

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * hospi
## Data Frame: dados5
##
## -----
##           hospi           no           yes           Total
## pregnant_SN
##           no       53076 (98.8%)    645 (1.2%)    53721 (100.0%)
##           yes        974 (97.9%)     21 (2.1%)     995 (100.0%)
##           Total    54050 (98.8%)    666 (1.2%)    54716 (100.0%)
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##           1.77         1.14         2.75
## -----
```

```
fisher.test(dados5$hospi, dados5$pregnant_SN, simulate.p.value = TRUE)
```

```
##
## Fisher's Exact Test for Count Data
##
## data:  dados5$hospi and dados5$pregnant_SN
## p-value = 0.01794
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.085658 2.750818
## sample estimates:
## odds ratio
##  1.774155
```

```
with(dados5, ctable(pregnant_SN, evolution, prop = "r", useNA = "no", chisq = FALSE, OR=TRUE))
```

Evolution

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * evolution
## Data Frame: dados5
##
```



```
##
## -----
##           evolution           cure           death           Total
## pregnant_SN
##           no           119452 ( 99.6%)           478 (0.4%)           119930 (100.0%)
##           yes           1668 ( 99.3%)            11 (0.7%)            1679 (100.0%)
##           Total           121120 ( 99.6%)           489 (0.4%)           121609 (100.0%)
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##           1.65           0.90           3.00
## -----
```

#Exact's Fisher Test

```
fisher.test( dados5$pregnant_SN, dados5$evolution, simulate.p.value = TRUE)
```

```
##
## Fisher's Exact Test for Count Data
##
## data:  dados5$pregnant_SN and dados5$evolution
## p-value = 0.1146
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.8155371 2.9834308
## sample estimates:
## odds ratio
##  1.647998
```

Propensity Score Matching

Analyzing the balance between the groups pregnant and non-pregnant

```
bal.tab(pregnant_SN ~ age_group + esc_resumida + ethn + cardio + diabetes + obesity,
        data = dados5, estimand = "ATE", m.threshold = .05)
```

```
## Balance Measures
##
##           Type Diff.Un      M.Threshold.Un
## age_group_10 - 19      Binary -0.0056      Balanced, <0.05
## age_group_20 - 29      Binary  0.2237 Not Balanced, >0.05
## age_group_30 - 39      Binary  0.0360      Balanced, <0.05
## age_group_40 - 49      Binary -0.2541 Not Balanced, >0.05
## esc_resumida_No schooling      Binary -0.0017      Balanced, <0.05
## esc_resumida_Incomplete elementary      Binary  0.0150      Balanced, <0.05
## esc_resumida_Elementary      Binary  0.0752 Not Balanced, >0.05
## esc_resumida_Middle school      Binary -0.0113      Balanced, <0.05
## esc_resumida_College      Binary -0.0772 Not Balanced, >0.05
## esc_resumida:<NA>      Binary -0.1145 Not Balanced, >0.05
## ethn_Black      Binary  0.0136      Balanced, <0.05
## ethn_Brown      Binary  0.0640 Not Balanced, >0.05
## ethn_Indigenous      Binary -0.0016      Balanced, <0.05
```

```

## ethn_White          Binary -0.0711 Not Balanced, >0.05
## ethn_Yellow         Binary -0.0048      Balanced, <0.05
## ethn:<NA>           Binary -0.1310 Not Balanced, >0.05
## cardio_yes          Binary -0.0192      Balanced, <0.05
## cardio:<NA>         Binary -0.0018      Balanced, <0.05
## diabetes_yes        Binary  0.0105      Balanced, <0.05
## diabetes:<NA>       Binary -0.0018      Balanced, <0.05
## obesity_yes         Binary -0.0086      Balanced, <0.05
## obesity:<NA>        Binary -0.0017      Balanced, <0.05
##
## Balance tally for mean differences
##               count
## Balanced, <0.05      14
## Not Balanced, >0.05   8
##
## Variable with the greatest mean difference
##      Variable Diff.Un      M.Threshold.Un
## age_group_40 - 49 -0.2541 Not Balanced, >0.05
##
## Sample sizes
##      no  yes
## All 119930 1679

```

##Obteining the weights

```

ajuste <- weightit(pregnant_SN ~ age_group + esc_resumida + ethn + cardio + diabetes + obesity,
                  use.mlogit = FALSE, data = dados5, method = "ps", estimand = "ATE")
bal.tab(ajuste, m.threshold = 0.05, disp.v.ratio = TRUE, estimand = "ATE")

```

```

## Call
## weightit(formula = pregnant_SN ~ age_group + esc_resumida + ethn +
##      cardio + diabetes + obesity, data = dados5, method = "ps",
##      estimand = "ATE", use.mlogit = FALSE)
##
## Balance Measures
##
##      Type Diff.Adj      M.Threshold
## prop.score      Distance -0.0117 Balanced, <0.05
## age_group_10 - 19      Binary -0.0058 Balanced, <0.05
## age_group_20 - 29      Binary -0.0142 Balanced, <0.05
## age_group_30 - 39      Binary  0.0116 Balanced, <0.05
## age_group_40 - 49      Binary  0.0085 Balanced, <0.05
## esc_resumida_No schooling      Binary -0.0018 Balanced, <0.05
## esc_resumida_Incomplete elementary      Binary -0.0264 Balanced, <0.05
## esc_resumida_Elementary      Binary -0.0121 Balanced, <0.05
## esc_resumida_Middle school      Binary  0.0076 Balanced, <0.05
## esc_resumida_College      Binary  0.0327 Balanced, <0.05
## esc_resumida:<NA>      Binary -0.0252 Balanced, <0.05
## ethn_Black      Binary  0.0082 Balanced, <0.05
## ethn_Brown      Binary  0.0164 Balanced, <0.05
## ethn_Indigenous      Binary -0.0007 Balanced, <0.05
## ethn_White      Binary -0.0053 Balanced, <0.05
## ethn_Yellow      Binary -0.0185 Balanced, <0.05
## ethn:<NA>      Binary -0.0368 Balanced, <0.05

```

```

## cardio_yes          Binary -0.0237 Balanced, <0.05
## cardio:<NA>          Binary -0.0017 Balanced, <0.05
## diabetes_yes        Binary -0.0048 Balanced, <0.05
## diabetes:<NA>        Binary -0.0018 Balanced, <0.05
## obesity_yes         Binary  0.0107 Balanced, <0.05
## obesity:<NA>         Binary -0.0017 Balanced, <0.05
##                    V.Ratio.Adj
## prop.score          1.0473
## age_group_10 - 19    .
## age_group_20 - 29    .
## age_group_30 - 39    .
## age_group_40 - 49    .
## esc_resumida_No schooling .
## esc_resumida_Incomplete elementary .
## esc_resumida_Elementary .
## esc_resumida_Middle school .
## esc_resumida_College .
## esc_resumida:<NA> .
## ethn_Black .
## ethn_Brown .
## ethn_Indigenous .
## ethn_White .
## ethn_Yellow .
## ethn:<NA> .
## cardio_yes .
## cardio:<NA> .
## diabetes_yes .
## diabetes:<NA> .
## obesity_yes .
## obesity:<NA> .
##
## Balance tally for mean differences
##          count
## Balanced, <0.05      23
## Not Balanced, >0.05    0
##
## Variable with the greatest mean difference
##   Variable Diff.Adj    M.Threshold
## ethn:<NA> -0.0368 Balanced, <0.05
##
## Effective sample sizes
##          no      yes
## Unadjusted 119930 1679.
## Adjusted   119912 495.99

```

```
dados5$weight <- ajuste$weights
```

```
##After-PSM
```

```

with(dados5, ctable(pregnant_SN, hospi, prop = "r", useNA = "no",
                    chisq = FALSE, OR=TRUE, weights = weight))

```

```
## Cross-Tabulation, Row Proportions
```

```
## pregnant_SN * hospi
## Data Frame: dados5
##
##
## -----
##           hospi           no           yes           Total
## pregnant_SN
##           no           53880.5 (98.8%)           652.7 (1.2%)           54533.2 (100.0%)
##           yes           67473.0 (96.0%)           2794.0 (4.0%)           70267.0 (100.0%)
##           Total           121353.5 (97.2%)           3446.7 (2.8%)           124800.2 (100.0%)
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##           3.42           3.14           3.73
## -----
```

```
with(dados5, ctable(pregnant_SN, evolution, prop="r", OR=TRUE, useNA = "no",
                    chisq = TRUE, weights = weight))
```

```
## Cross-Tabulation, Row Proportions
## pregnant_SN * evolution
## Data Frame: dados5
##
##
## -----
##           evolution           cure           death           Total
## pregnant_SN
##           no           121125.1 ( 99.6%)           483.1 (0.4%)           121608.3 (100.0%)
##           yes           121513.4 ( 98.7%)           1599.1 (1.3%)           123112.5 (100.0%)
##           Total           242638.5 ( 99.1%)           2082.3 (0.9%)           244720.8 (100.0%)
## -----
##
## -----
## Chi.squared   df   p.value
## -----
##           588.4437           1           0
## -----
##
## -----
## Odds Ratio   Lo - 95%   Hi - 95%
## -----
##           3.30           2.98           3.65
## -----
```