

Seleção dos casos de H1N1 para paper de comparação com COVID-19

Gestantes

29/05/2021

Sobre a base de dados e pacotes do R utilizados

A seguir são carregados os pacotes do R (<https://www.r-project.org>) utilizados para filtragem e tratamento dos dados.

```
#carregar pacotes
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",
    "ggplot2",
    "kableExtra",
    "tables",
    "questionr",
    "car",
    "data.table",
    "magrittr",
    "tidyverse",
    "readxl",
    "summarytools",
    "zoo",
    "grid",
    "gridExtra",
    "cowplot"
  )
lapply(packages, loadlibrary)
```

A base de dados SIVEP-Gripe (Sistema de Informação da Vigilância Epidemiológica da Gripe) tem os registros dos casos e óbitos de SRAG (Síndrome Respiratória Aguda Grave). A notificação é compulsória para síndrome gripal (caracterizado por pelo menos dois dos seguintes sinais e sintomas: febre, mesmo que

referida, calafrios, dor de garganta, dor de cabeça, tosse, coriza, distúrbios olfatórios ou de paladar) e que tem dispneia / desconforto respiratório ou pressão persistente no peito ou Saturação de O₂ menor que 95% no ar ambiente ou cor azulada dos lábios ou rosto. Indivíduos assintomáticos com confirmação laboratorial por biologia molecular ou exame imunológico para infecção por COVID-19 também são relatados.

Para notificações no Sivep-Gripe, os casos hospitalizados em hospitais públicos e privados e todas as mortes devido a infecções respiratórias agudas graves, independentemente da hospitalização, devem ser considerados.

A vigilância da SRAG no Brasil é desenvolvida pelo Ministério da Saúde (MS), por meio da Secretaria de Vigilância em Saúde (SVS), desde a pandemia de Influenza A (H1N1) em 2009. Mais informações em <https://coronavirus.saude.gov.br/definicao-de-caso-e-notificacao>.

O período analisado compreende de dados epidemiológicos de 2009 e 2010, com banco de dados obtido em 05/05/2021 no site <https://opendatasus.saude.gov.br/dataset/bd-srag-2009-a-2012>. Os dados de 2009 e de 2010 são carregados e combinados abaixo:

```
##### carregando as bases de dados #####
#2009 e 2010
dados2009 <- read_delim(
  "influd09_limpo-final.csv",
  ";",
  escape_double = FALSE,
  locale = locale(encoding = "ISO-8859-2"),
  trim_ws = TRUE
)

dados2010 <- read_delim(
  "influd10_limpo-final.csv",
  ";",
  escape_double = FALSE,
  locale = locale(encoding = "ISO-8859-2"),
  trim_ws = TRUE
)

dados_2009 <- dados2009 %>%
  rename(SRAGFINAL = SRAG2009FINAL)

dados_2010 <- dados2010 %>%
  rename(SRAGFINAL = SRAG2010FINAL)

dados <- full_join(dados_2009, dados_2010)

sem <- 19

memory.limit(999999)

#Criar variavel de ano do caso
dados <- dados %>%
  dplyr::mutate(
    dt_sint = as.Date(DT_SIN_PRI, format = "%d/%m/%Y"),
    ano = lubridate::year(dt_sint),
    mes = lubridate::month(dt_sint)
  )
```

Há atualmente 217167 observações na base de dados e são as variáveis:

```
names(dados)
```

```
## [1] "DT_NOTIFIC" "ID_MUNICIP" "SEM_NOT" "NU_ANO" "SG_UF_NOT"
## [6] "DT_SIN_PRI" "DT_NASC" "NU_IDADE_N" "CS_SEXO" "CS_GESTANT"
## [11] "CS_RACA" "CS_ESCOL_N" "SG_UF" "ID_MN_RESI" "ID_OCUPA_N"
## [16] "VACINA" "FEBRE" "TOSSE" "CALAFRIO" "DISPNEIA"
## [21] "GARGANTA" "ARTRALGIA" "MIALGIA" "CONJUNTIV" "CORIZA"
## [26] "DIARREIA" "OUTRO_SIN" "OUTRO_DES" "CARDIOPATI" "PNEUMOPATI"
## [31] "RENAL" "HEMOGLOBI" "IMUNODEPRE" "TABAGISMO" "METABOLICA"
## [36] "OUT_MORBI" "MORB_DESC" "HOSPITAL" "DT_INTERNA" "CO_UF_INTE"
## [41] "CO_MU_INTE" "DT_PCR" "PCR_AMOSTR" "PCR_OUT" "PCR_RES"
## [46] "PCR_ETIOL" "PCR_TIPO_H" "PCR_TIPO_N" "DT_CULTURA" "CULT_AMOST"
## [51] "CULT_OUT" "CULT_RES" "DT_HEMAGLU" "HEMA_RES" "HEMA_ETIOL"
## [56] "HEM_TIPO_H" "HEM_TIPO_N" "DT_RAIOX" "RAIOX_RES" "RAIOX_OUT"
## [61] "CLASSI_FIN" "CLASSI_OUT" "CRITERIO" "TPAUTOCTO" "DOENCA_TRA"
## [66] "EVOLUCAO" "DT_OBITO" "DT_ENCERRA" "DT_DIGITA" "MONITORA"
## [71] "SRAGFINAL" "OBES_IMC" "OUT_AMOST" "DS_OAGEETI" "DS_OUTMET"
## [76] "DS_OUTSUB" "OUT_ANTIV" "DT_COLETA" "DT_ENTUTI" "DT_ANTIVIR"
## [81] "DT_IFI" "DT_OUTMET" "DT_SAIDUTI" "RES_ADNO" "AMOSTRA"
## [86] "HEPATICA" "NEUROLOGIC" "OBESIDADE" "PUERPERA" "SIND_DOWN"
## [91] "RES_FLUA" "RES_FLUB" "UTI" "IFI" "PCR"
## [96] "RES_OUTRO" "OUT_METODO" "RES_PARA1" "RES_PARA2" "RES_PARA3"
## [101] "DESC_RESP" "SATURACAO" "ST_TIPOFI" "TIPO_PCR" "ANTIVIRAL"
## [106] "SUPORT_VEN" "RES_VSR" "RES_FLUASU" "AVE_10_DIA" "dt_sint"
## [111] "ano" "mes"
```

Para ver o dicionário das variáveis, acesse: https://opendatasus.saude.gov.br/dataset/c9a8f286-44bc-444e-94b4-f4ceded3af2c/resource/8e4ee33a-a7bd-42d1-9505-bbf0eb7e6141/download/dic_dados_influenza-pandemica_antigo.pdf-2009-a-set-2012.pdf

Filtragem e tratamento dos dados para projeto

A primeira filtragem consiste em selecionar os casos de maio de 2009 até maio de 2010.

```
#filtrando só os casos de maio de 2009 até maio de 2010
dados1 <- filter(dados,
                 (mes >=5 & ano == 2009) | (mes <=5 & ano == 2010))
```

Há 212584 observações na base de dados.

A próxima seleção de pessoas do sexo feminino:

```
#filtrando F
dados2 <- filter(dados1, CS_SEXO == "F")
```

Há 115357 observações na base de dados.

O próximo passo é filtrar só as mulheres entre 10 e 49 anos.

```
# Filtrando pela faixa de idade de interesse
dados3 <- dados2 %>%
  filter(NU_IDADE_N > 4009 & NU_IDADE_N < 4050)
```

Há 75717 observações na base de dados.

Agora vamos selecionar só as pessoas gestantes ou não gestante.

```
with(dados3, freq(CS_GESTANT))
```

```
## Frequencies
## dados3$CS_GESTANT
## Type: Numeric
##
##          Freq  % Valid  % Valid Cum.  % Total  % Total Cum.
## -----
##          1   2576     3.40         3.40    3.40    3.40
##          2   4782     6.32         9.72    6.32    9.72
##          3   4284     5.66        15.38    5.66   15.38
##          4    434     0.57        15.95    0.57   15.95
##          5  49663    65.59        81.54   65.59   81.54
##          6   5348     7.06        88.60    7.06   88.60
##          9   8630    11.40       100.00   11.40  100.00
##         <NA>     0         0.00    0.00  100.00
##        Total 75717   100.00       100.00  100.00  100.00
```

```
dados3 <- dados3 %>%
  mutate(
    classi_gesta = case_when(
      CS_GESTANT == 1 ~ "1tri",
      CS_GESTANT == 2 ~ "2tri",
      CS_GESTANT == 3 ~ "3tri",
      CS_GESTANT == 4 ~ "IG_ig",
      CS_GESTANT == 5 ~ "não",
      TRUE ~ NA_character_
    )
  )

#filtrando só gestante ou não gestante
dados4 <- dados3 %>%
  filter(!is.na(classi_gesta))

# Criando a variável gestante_SN
dados4 <- dados4 %>%
  mutate(gestante_SN = ifelse(CS_GESTANT == 5, "não", "sim"))
```

Há 61739 observações na base de dados.

A próxima seleção são os casos de covid indicado pela variável CLASSI_FIN.

```
with(dados4, freq(CLASSI_FIN))
```

```
## Frequencies
## dados4$CLASSI_FIN
## Type: Numeric
##
```

##		Freq	% Valid	% Valid Cum.	% Total	% Total Cum.
##	-----	-----	-----	-----	-----	-----
##	1	35214	59.842	59.842	57.037	57.037
##	2	1230	2.090	61.932	1.992	59.029
##	3	22351	37.983	99.915	36.202	95.232
##	4	24	0.041	99.956	0.039	95.270
##	9	26	0.044	100.000	0.042	95.313
##	<NA>	2894			4.687	100.000
##	Total	61739	100.000	100.000	100.000	100.000

```
dados5 <- dados4 %>%
  filter(CLASSI_FIN == 1)
```

Há 35214 observações na base de dados.

Agora vamos criar a variável se CLASSI_FIN==1 por PCR ou outro tipo de diagnóstico.

Essa variável é `pcr_test`, com as categorias: `pcr_pos` se PCR positivo (`CRITERIO == 1` e `PCR_RES == 1`) e `não` caso contrário.

```
#criar a variavel de pcr_test pelas variaveis de pcr
dados5 <- dados5 %>%
  mutate(pcr_test = case_when(PCR_RES == 1 & #pcr positivo
                                PCR_ETIOL== 1 ~ "pcr_pos",
                                TRUE ~ "não"
                              ))
```

Agora iremos filtrar os casos com `pcr_test!=não` que são os casos onde podem ser covid-19 apenas por PCR :

```
dados6 <- dados5 %>%
  filter(pcr_test == "pcr_pos")
```

Há 8854 observações na base de dados.