Plots

Eduardo Yuki Yada

Imports

```
library(tidyverse)
library(yaml)

options(dplyr.summarise.inform = FALSE)
```

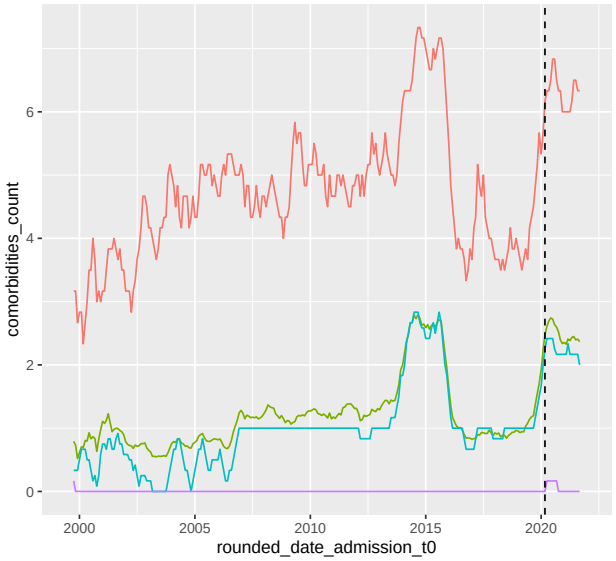
Loading data

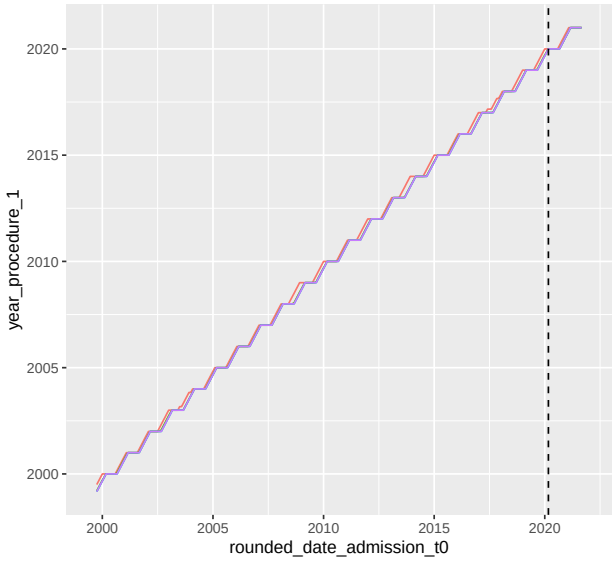
```
load('dataset/processed_data.RData')
load('dataset/processed_dictionary.RData')

columns_list <- yaml.load_file("./auxiliar/columns_list.yaml")

outcome_column <- params$outcome_column

df[columns_list$outcome_columns] <- lapply(df[columns_list$outcome_columns], as.character)
df[columns_list$outcome_columns] <- lapply(df[columns_list$outcome_columns], as.integer)
```

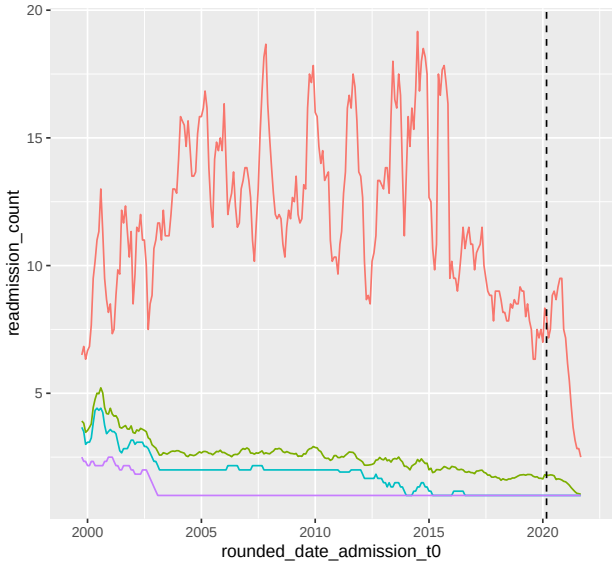
Plots

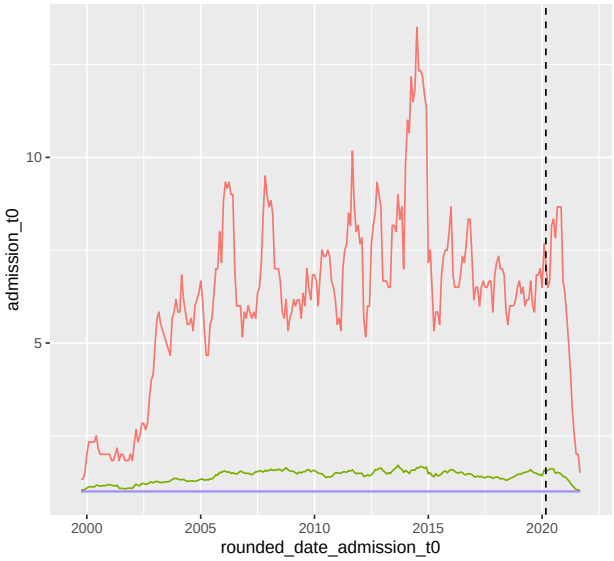
```
library(lubridate)

medianWithoutNA <- function(x) {
   median(x[which(!is.na(x))])
}

k = 6

df[columns_list$outcome_columns] <- lapply(df[columns_list$outcome_columns], as.numeric)

for (column in setdiff(columns_list$numerical_columns, columns_list$outcome_columns)){
  if (mean(is.na(df[[column]])) > 0.5) next

  df %>%
    mutate(rounded_date_admission_t0 = lubridate::floor_date(date_admission_t0,
                                                             'month')) %>%
    group_by(rounded_date_admission_t0) %>%
    summarise('Mean' = mean(!!sym(column), na.rm = T),
              'Min' = min(!!sym(column), na.rm = T),
              'Median' = medianWithoutNA(!!sym(column)),
              'Max' = max(!!sym(column), na.rm = T)) %>%
    ungroup %>%
    arrange(rounded_date_admission_t0) %>%
    mutate(Mean = zoo::rollmean(Mean, k = k, fill = NA),
           Min = zoo::rollmean(Min, k = k, fill = NA),
           Median = zoo::rollmean(Median, k = k, fill = NA),
           Max = zoo::rollmean(Max, k = k, fill = NA)) %>%
```

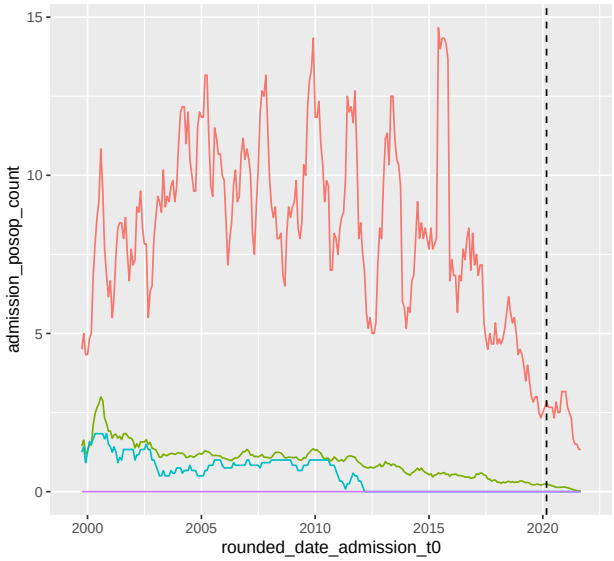
```
    pivot_longer(!rounded_date_admission_t0, names_to = 'metric') %>%
    ggplot(aes(x = rounded_date_admission_t0, y = value, color = metric)) +
      geom_line() +
      geom_vline(xintercept = lubridate::ymd('2020-03-01'), linetype = "dashed") +
      labs(y = column)-> p

  print(p)
}
```

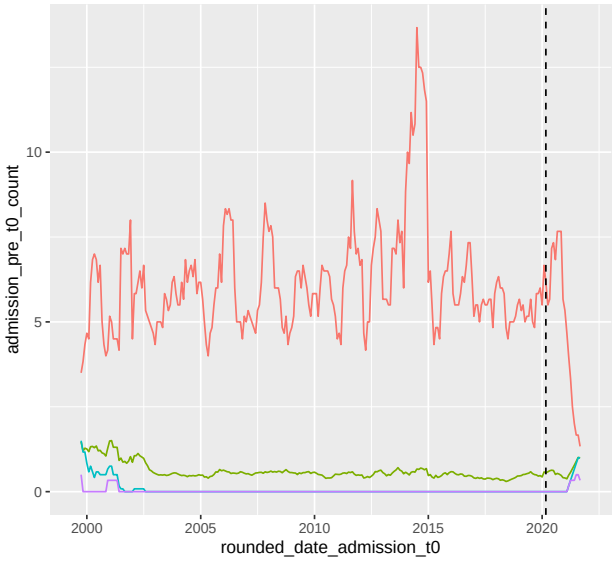


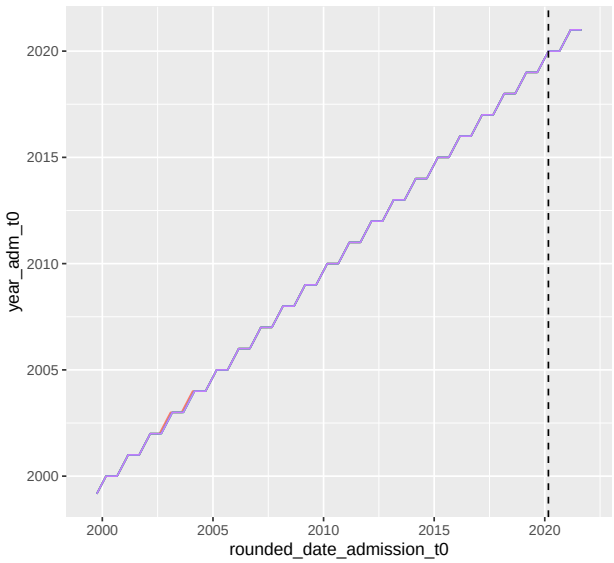


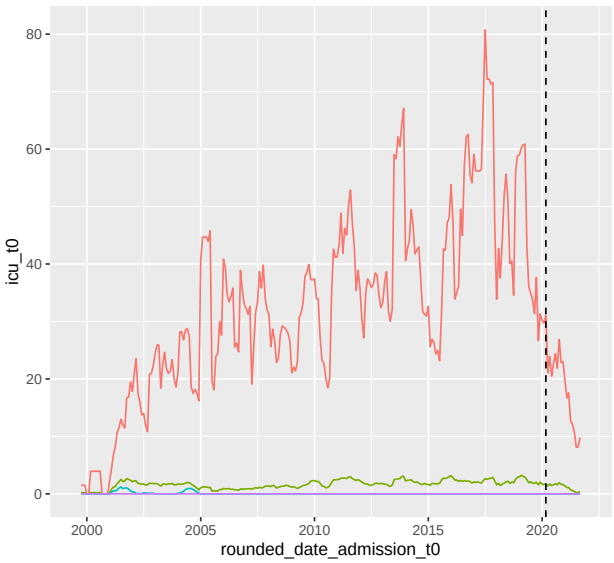

readmission_count
rounded_date_admission_t0
metric
Max
Mean
Median
Min

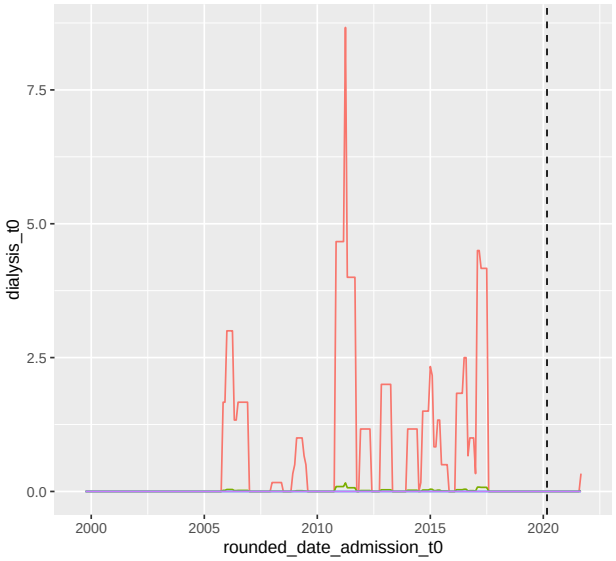

admission_t0
rounded_date_admission_t0
metric
Max
Mean
Median
Min

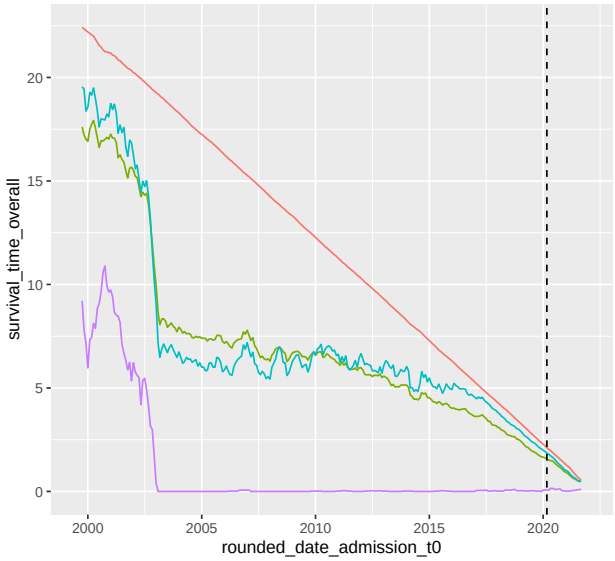

admission_posop_count
rounded_date_admission_t0
metric
Max
Mean
Median
Min

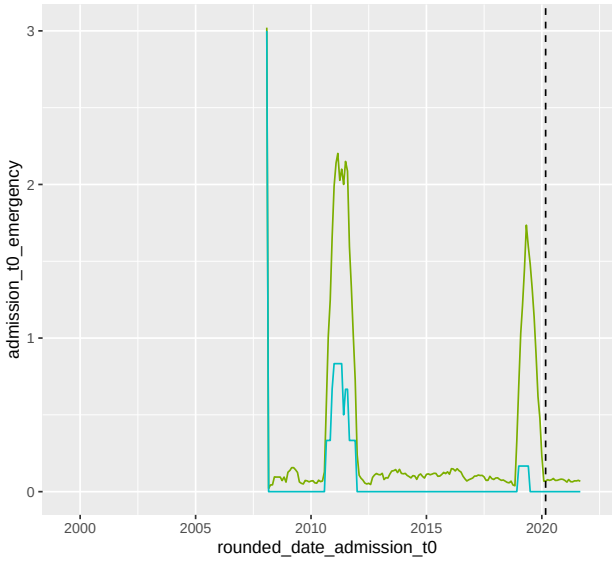

admission_pre_t0_count
rounded_date_admission_t0
metric
Max
Mean
Median
Min

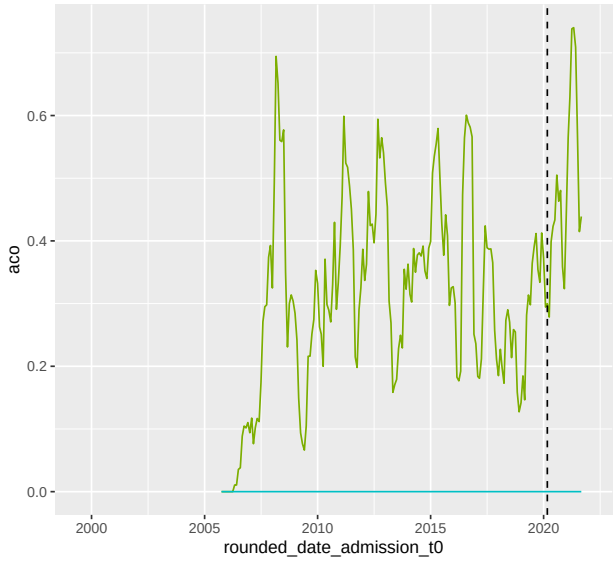

year_adm_t0
rounded_date_admission_t0
metric
Max
Mean
Median
Min

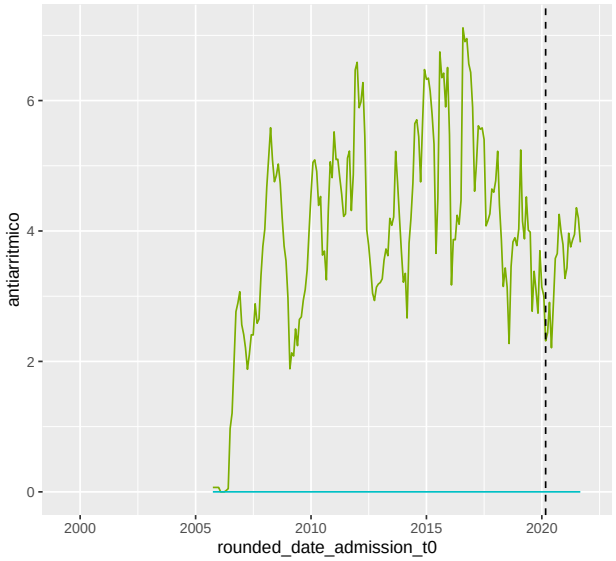

icu_t0
rounded_date_admission_t0
metric
Max
Mean
Median
Min

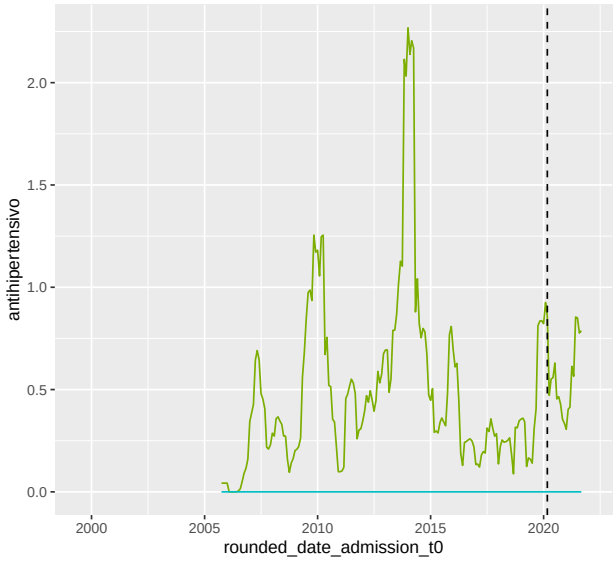

dialysis_t0
7.5
5.0
2.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min


survival_time_overall
20
15
10
5
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

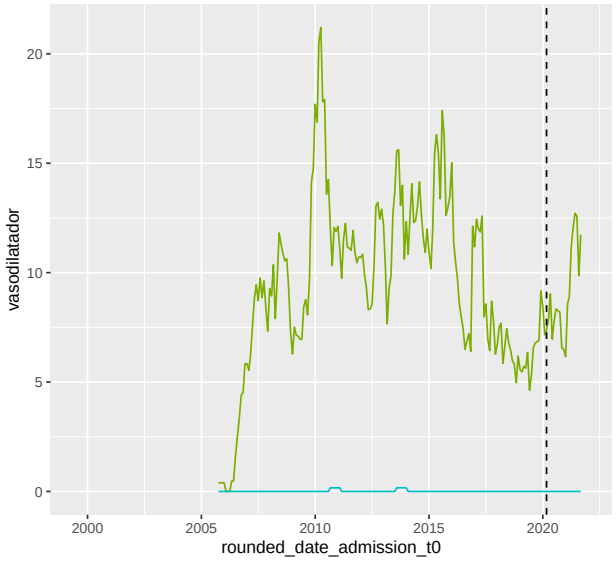

admission_t0_emergency
3
2
1
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

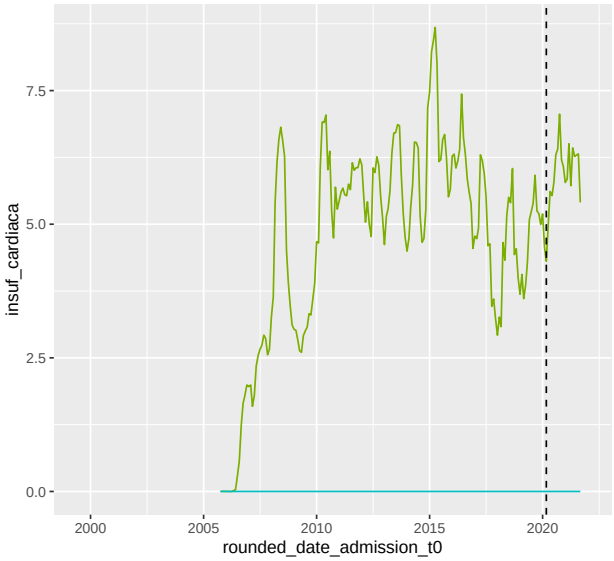

aco
0.6
0.4
0.2
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

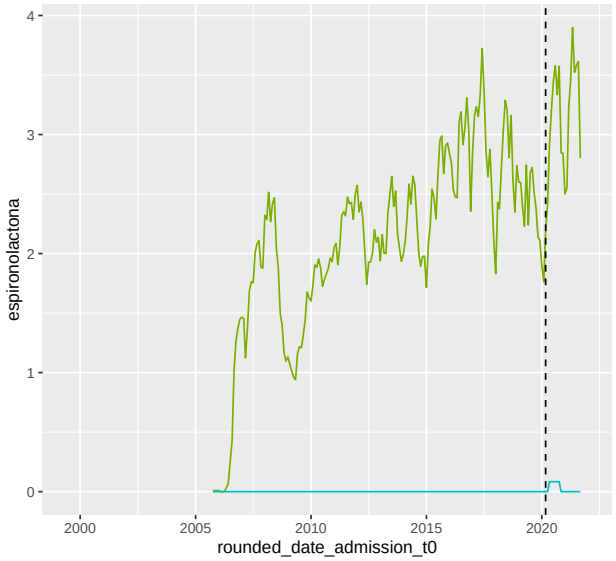

antiarritmico
6
4
2
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

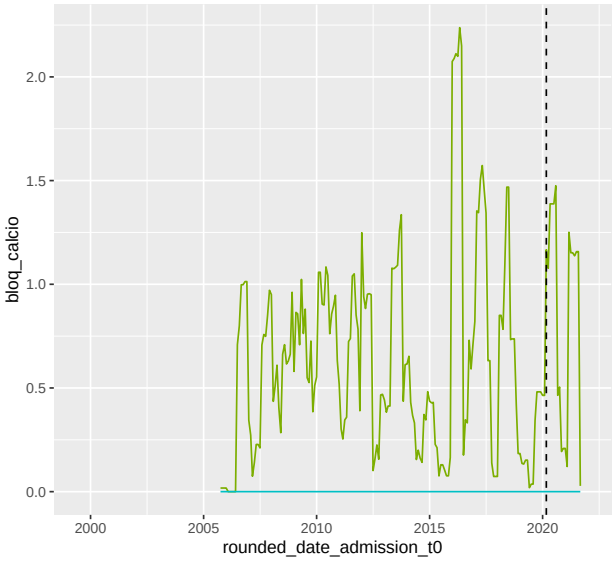

antihipertensivo
2.0
1.5
1.0
0.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

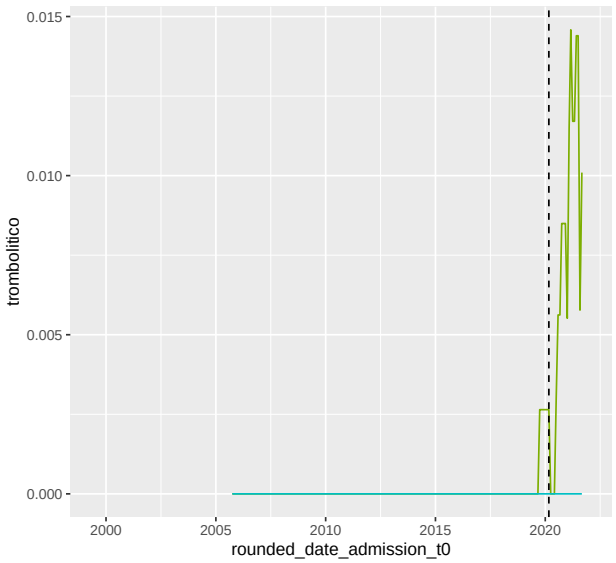

betabloqueador
ieca_bra
dva
digoxina
estatina
diuretico
rounded_date_admission_t0
metric
Max
Mean
Median
Min

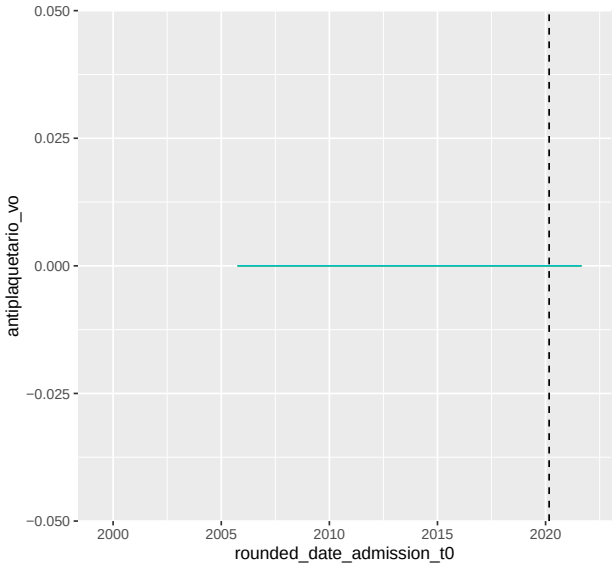

vasodilatador
rounded_date_admission_t0
metric
Max
Mean
Median
Min

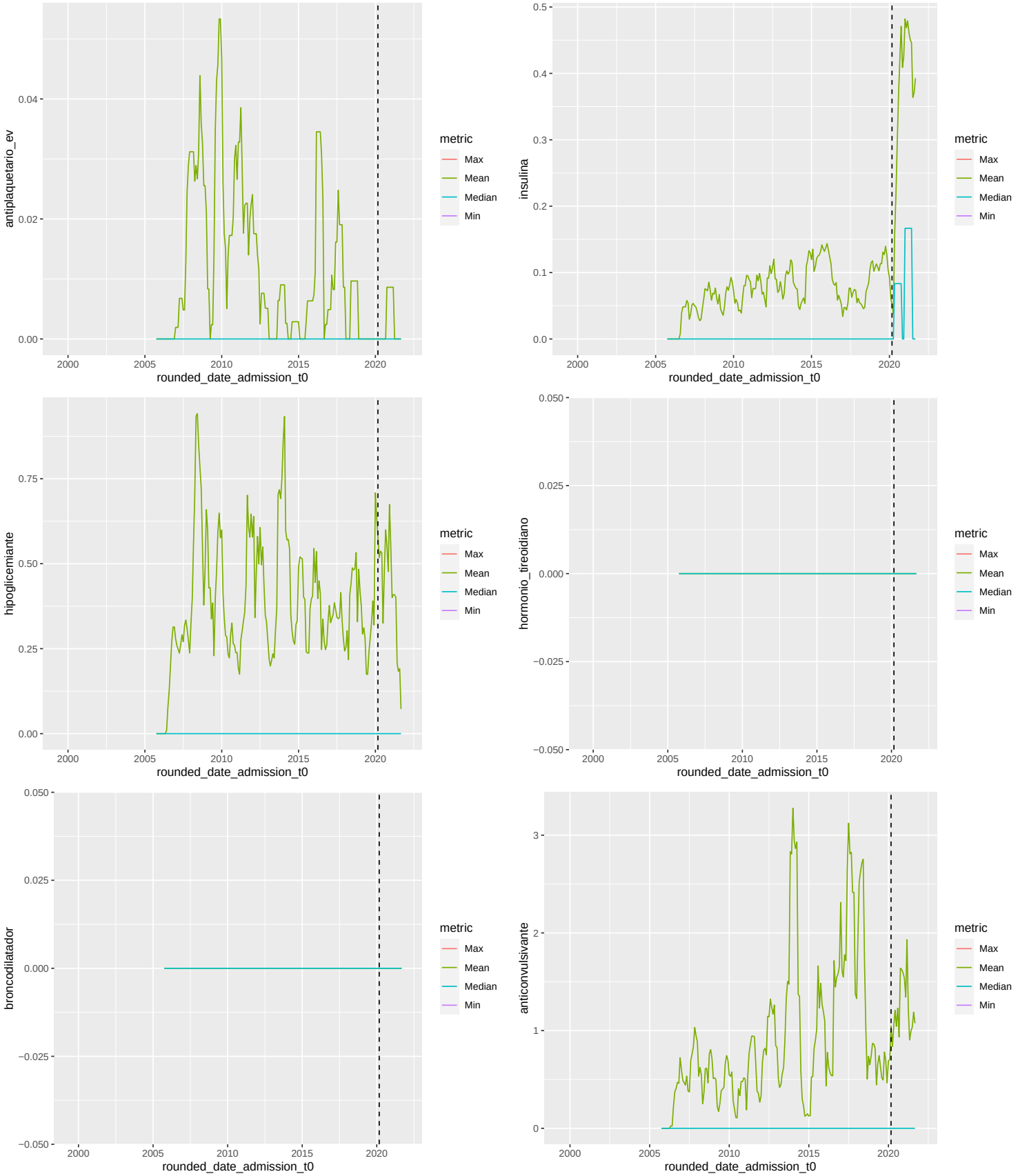

insuf_cardiaca
rounded_date_admission_t0
metric
Max
Mean
Median
Min

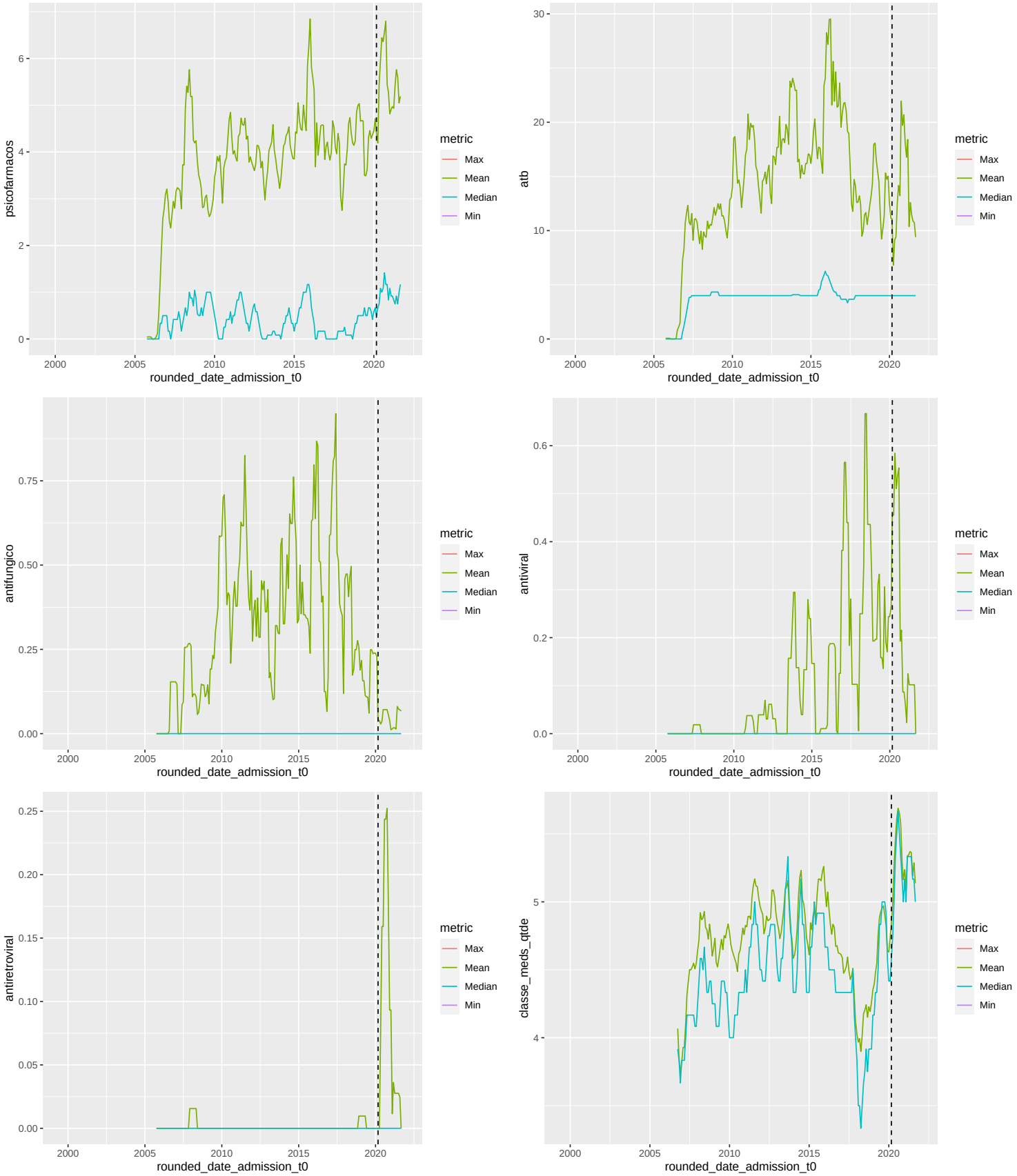

espironolactona
rounded_date_admission_t0
metric
Max
Mean
Median
Min

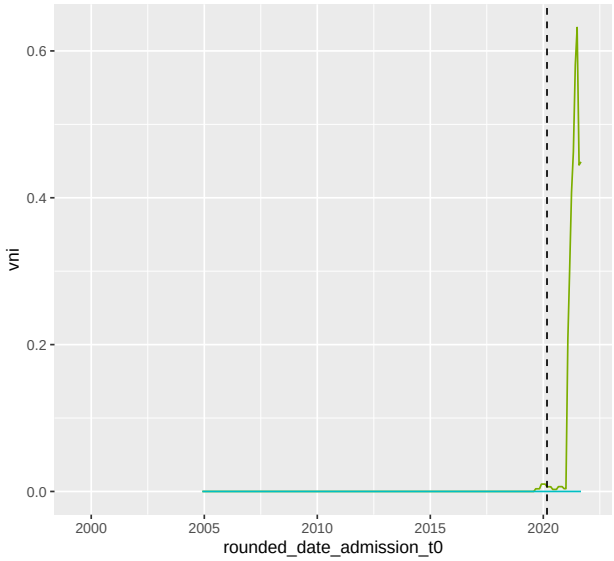

bloq_calcio
rounded_date_admission_t0
metric
Max
Mean
Median
Min

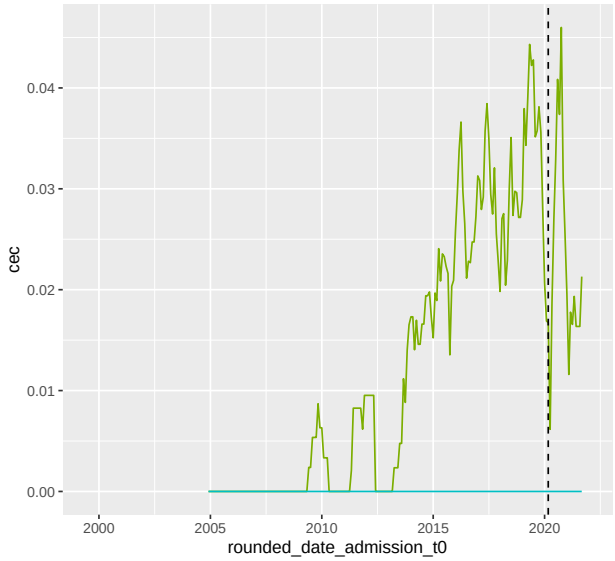

trombolitico
rounded_date_admission_t0
metric
Max
Mean
Median
Min

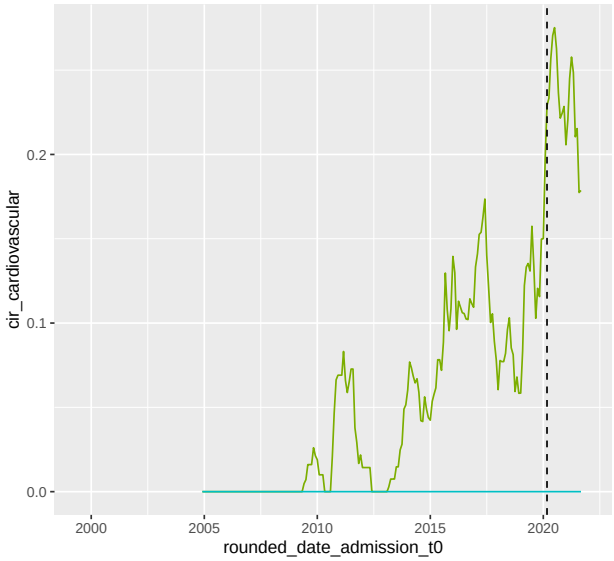

antiplaquetario_vo
rounded_date_admission_t0
metric
Max
Mean
Median
Min

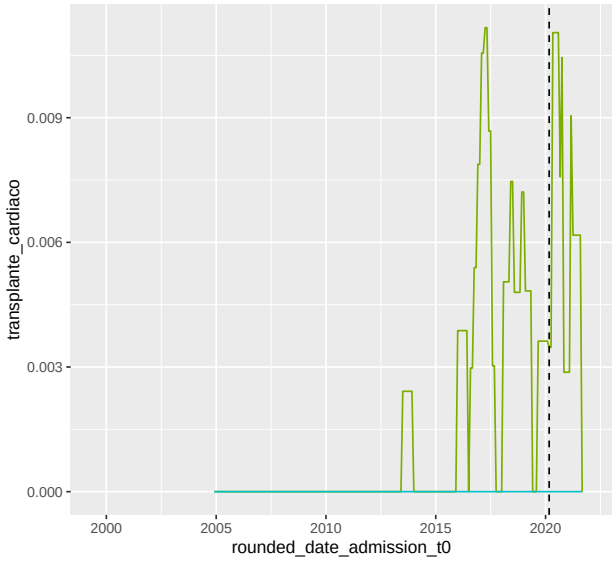

antiplaquetario_ev
0.04
0.02
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
insulina
0.5
0.4
0.3
0.2
0.1
0.0
hipoglicemiante
0.75
0.50
0.25
0.00
hormonio_tireoidiano
0.050
0.025
0.000
−0.025
−0.050
broncodilatador
anticonvulsivante
3
2
1
0

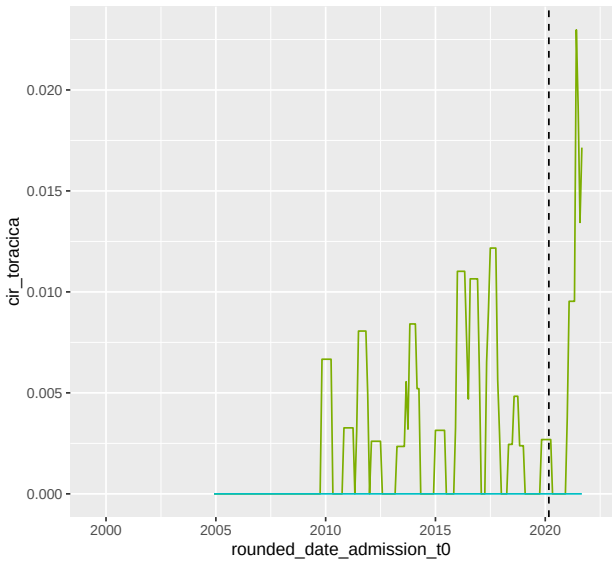

psicofarmacos
rounded_date_admission_t0
atb
antifungico
antiviral
antiretroviral
classe_meds_qtde
metric
Max
Mean
Median
Min

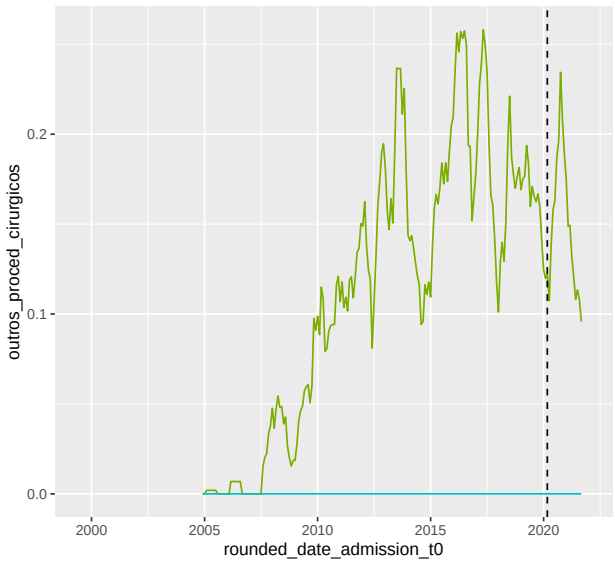

vni
0.6
0.4
0.2
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

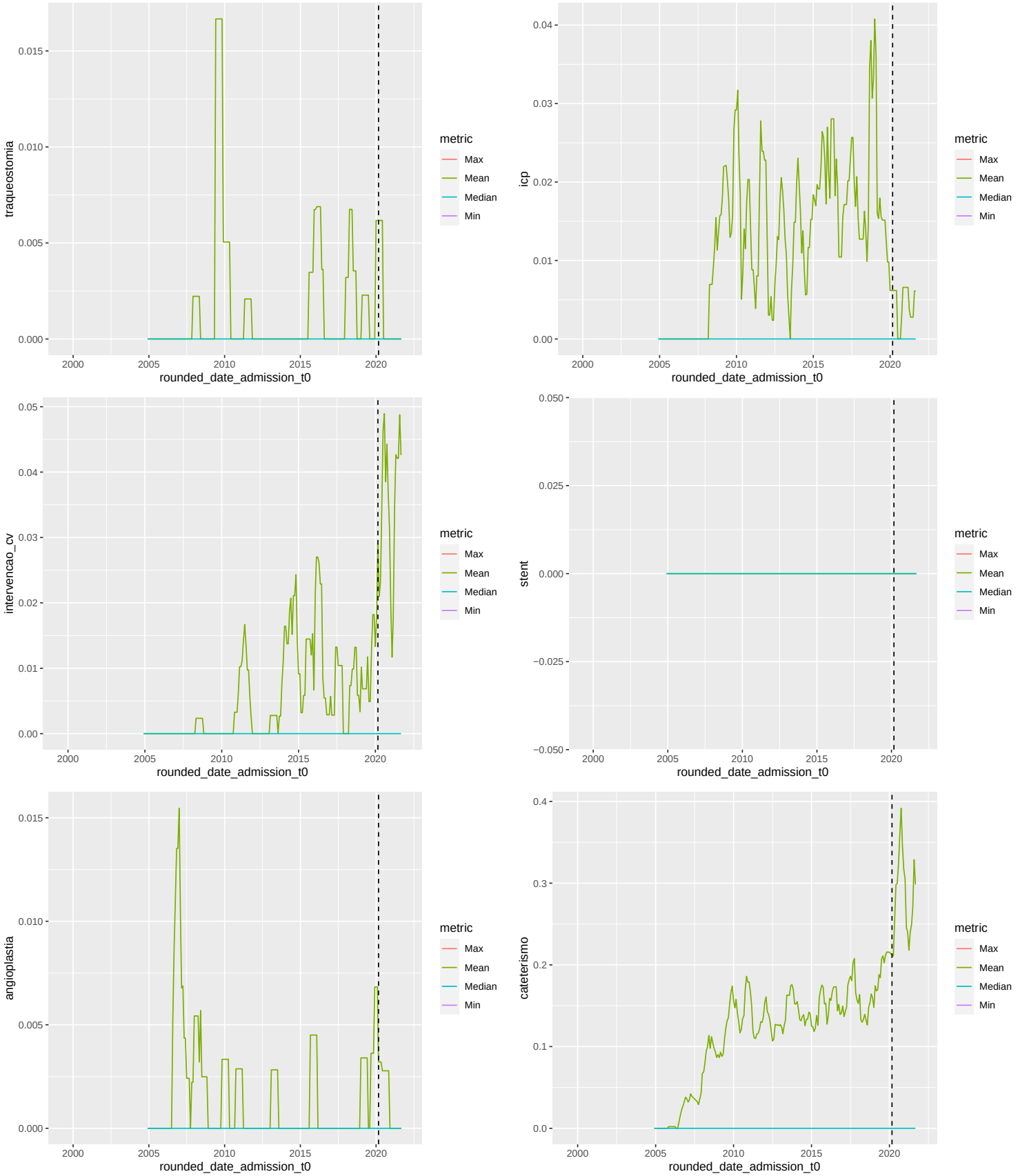

cec
0.04
0.03
0.02
0.01
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

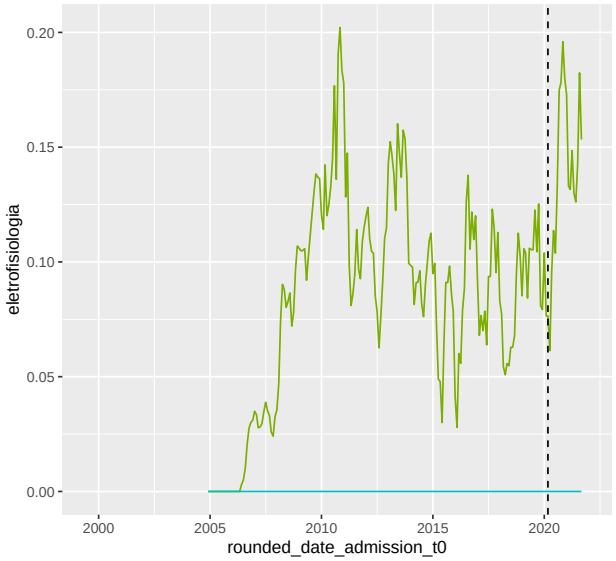

cir_cardiovascular
0.2
0.1
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

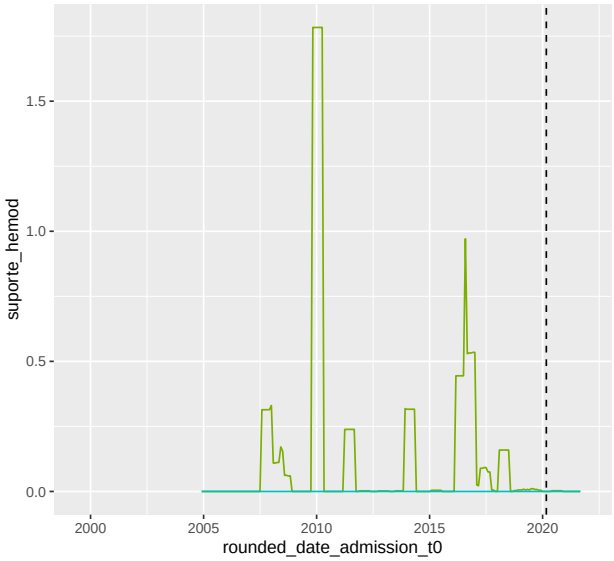

transplante_cardiaco
0.009
0.006
0.003
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

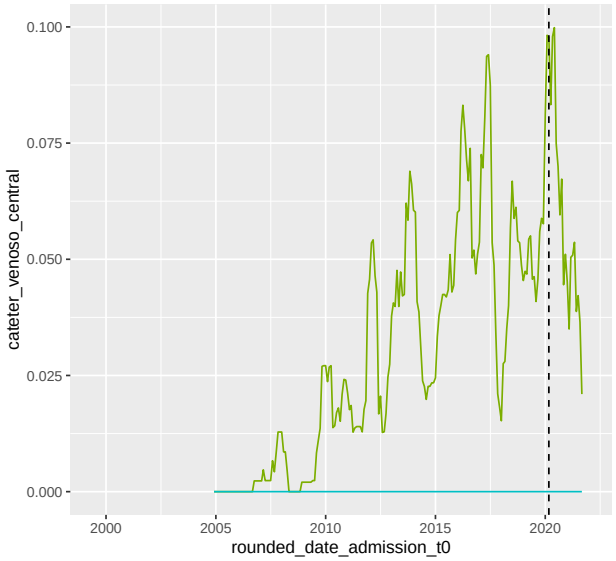

cir_toracica
0.020
0.015
0.010
0.005
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

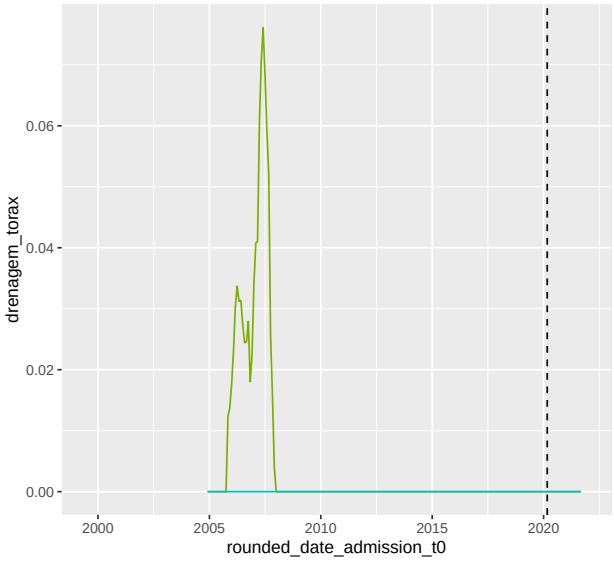

outros_proced_cirurgicos
0.2
0.1
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

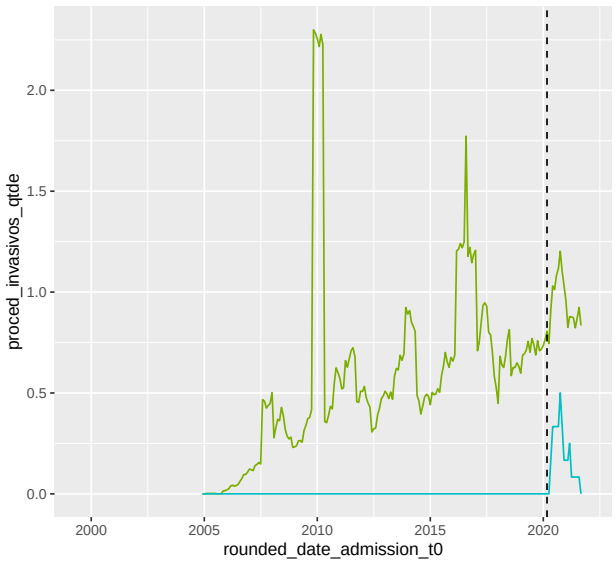

traqueostomia
0.015
0.010
0.005
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
icp
0.04
0.03
0.02
0.01
0.00
intervencao_cv
0.05
0.04
0.03
0.02
0.01
0.00
stent
0.050
0.025
0.000
−0.025
−0.050
angioplastia
0.015
0.010
0.005
0.000
cateterismo
0.4
0.3
0.2
0.1
0.0

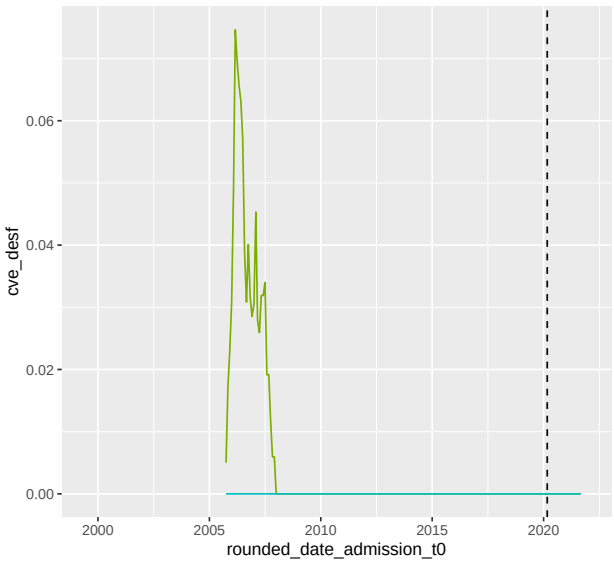

eletrofisiologia
0.20
0.15
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

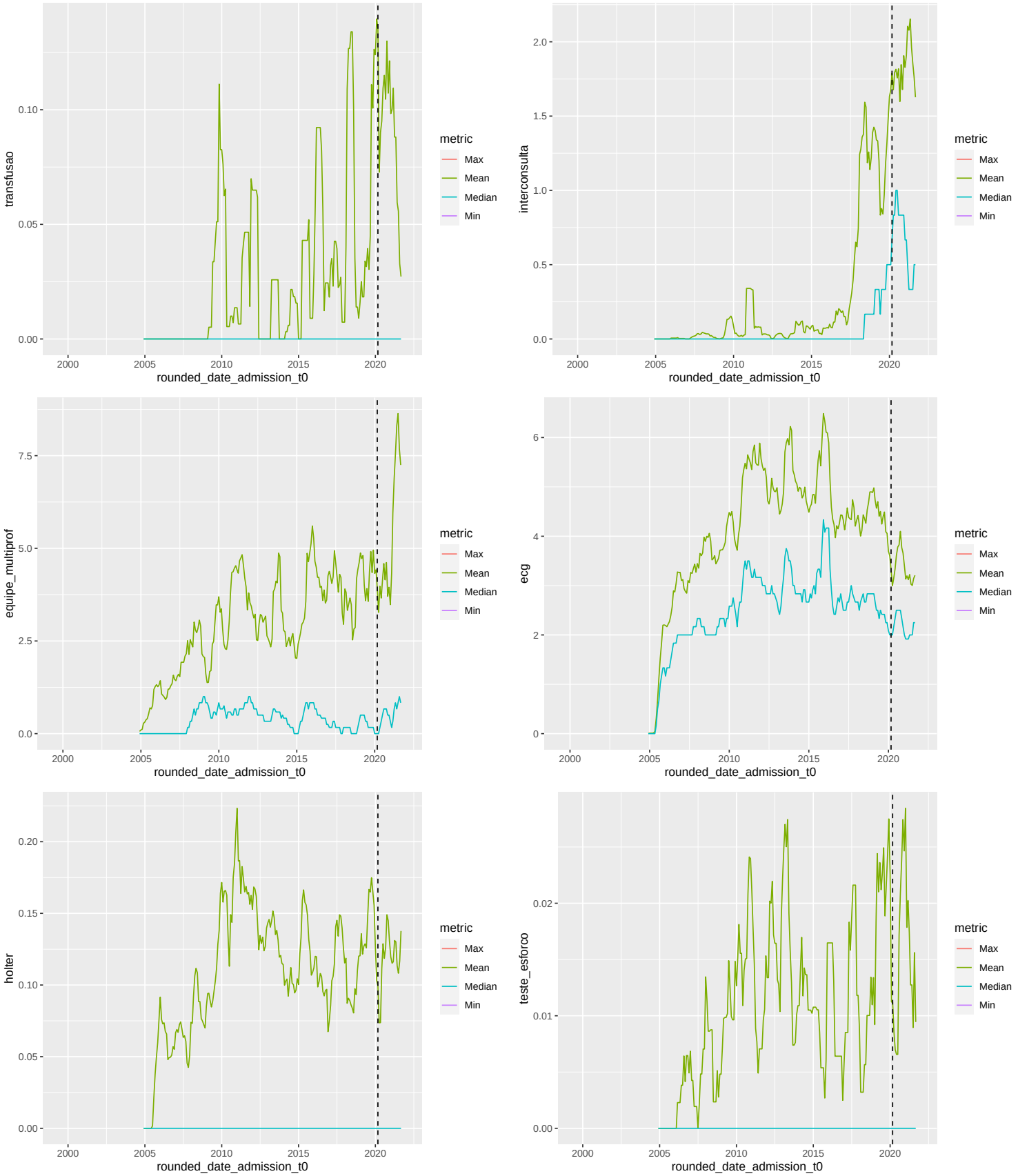

suporte_hemod
1.5
1.0
0.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

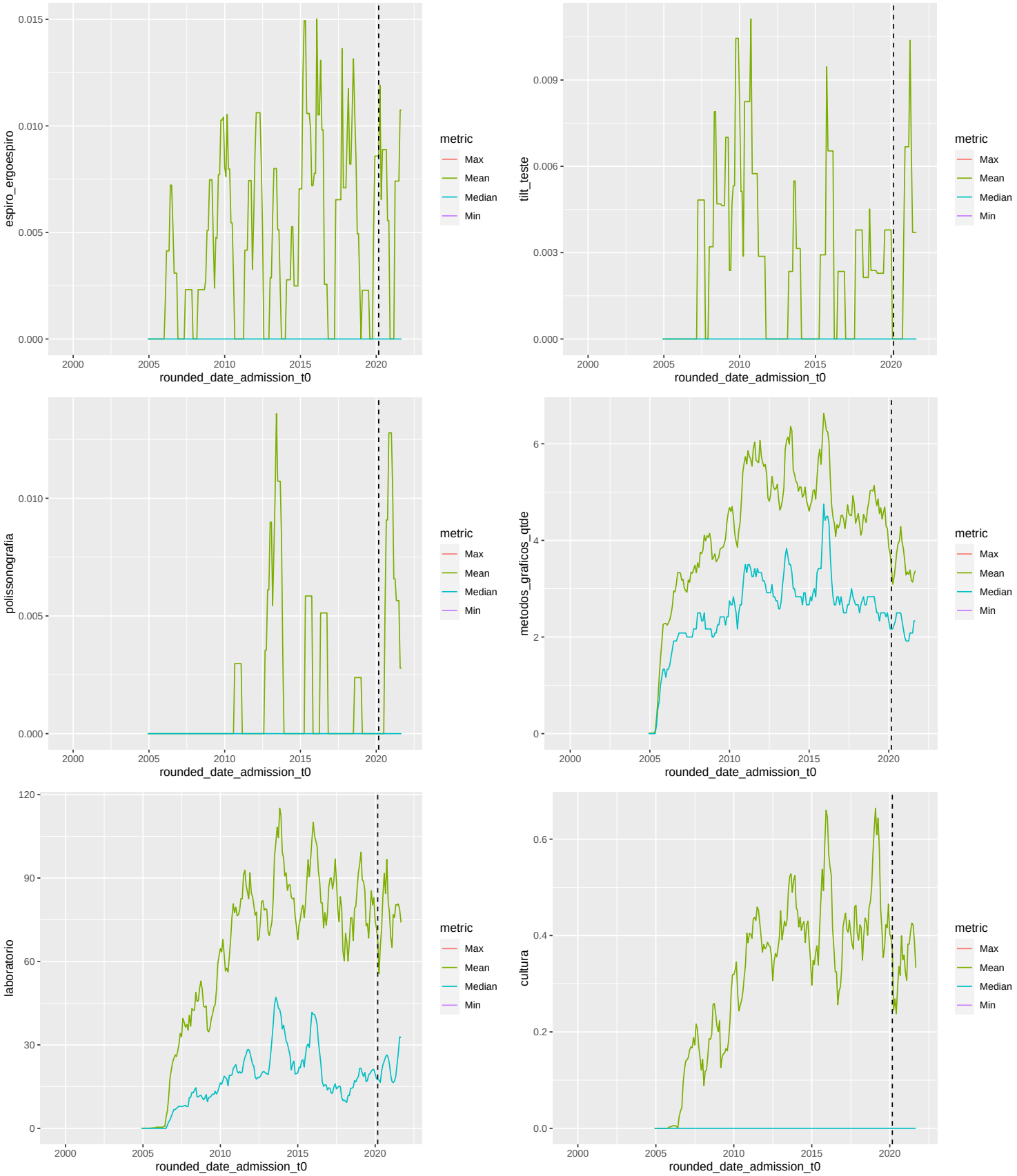

cateter_venoso_central
0.100
0.075
0.050
0.025
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min


drenagem_torax
0.06
0.04
0.02
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

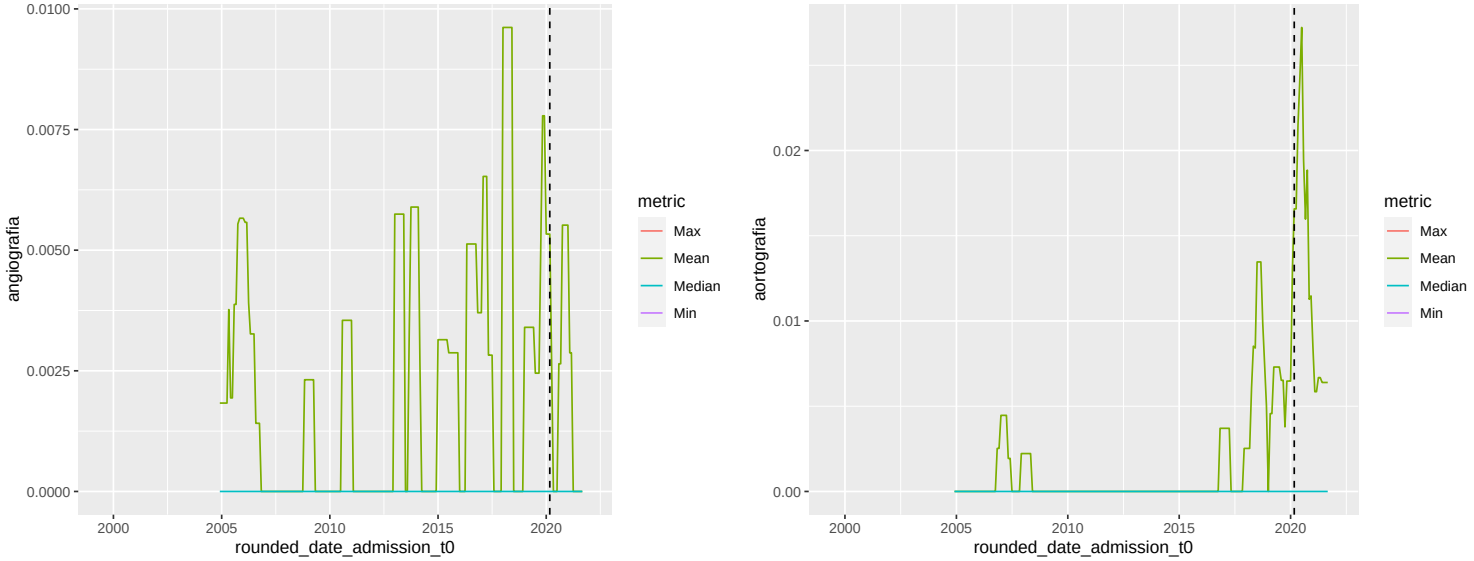

proced_invasivos_qtde
2.0
1.5
1.0
0.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min


cve_desf
0.06
0.04
0.02
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

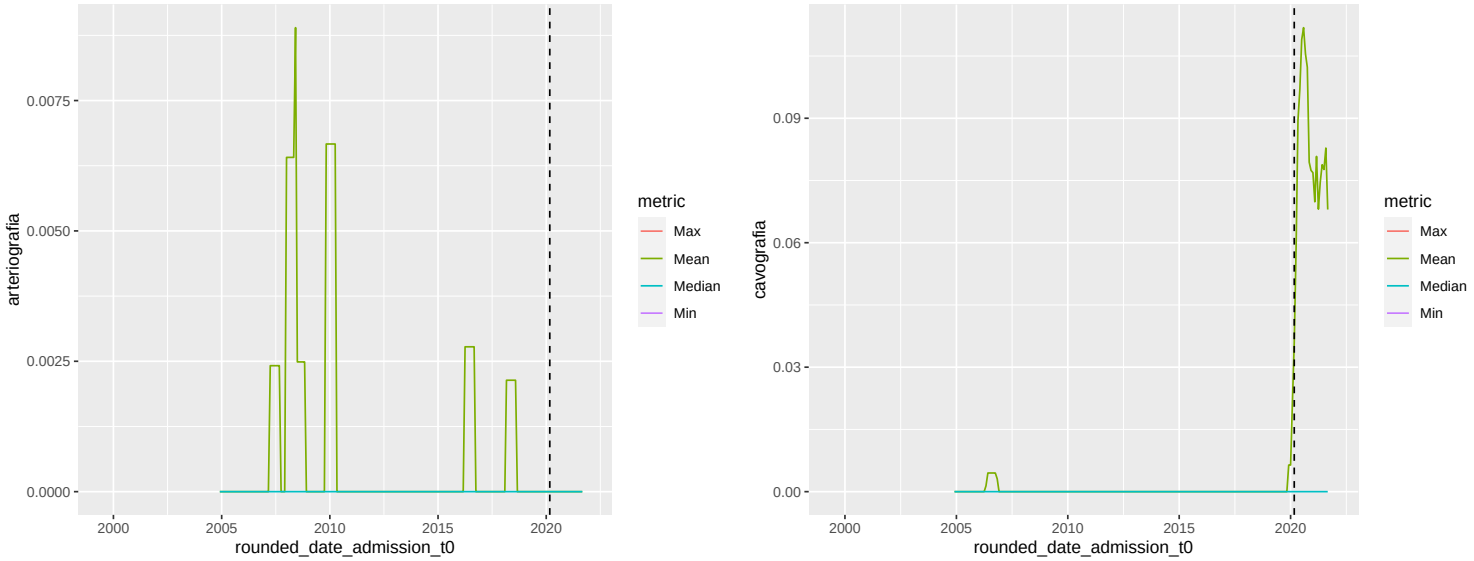

transfusao
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
interconsulta
2.0
1.5
1.0
0.5
0.0
equipe_multiprof
7.5
5.0
2.5
ecg
6
4
2
holter
0.20
0.15
teste_esforco
0.02
0.01


espiro_ergoespiro
0.015
0.010
0.005
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
tilt_teste
0.009
0.006
0.003
0.000
polissonografia
metodos_graficos_qtde
6
4
2
0
laboratorio
120
90
60
30
cultura
0.6
0.4
0.2
0.0

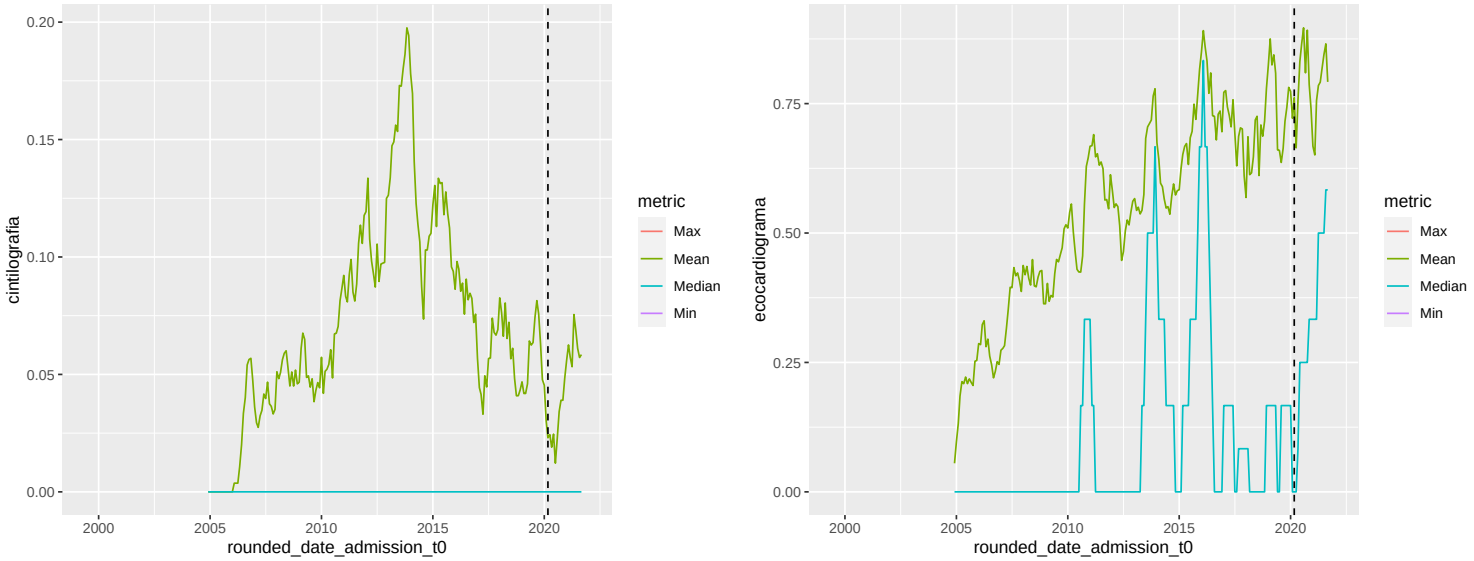

analises_clinicas_qtde
120
90
60
30
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
citologia
0.020
0.015
0.010
0.005
0.000
biopsia
0.100
0.075
0.050
0.025
histopatologia_qtde
0.125
angio_rm
0.015
0.010
0.005
angio_tc
0.10
0.05
0.00


angiografia
0.0100
0.0075
0.0050
0.0025
0.0000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

aortografia
0.02
0.01
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

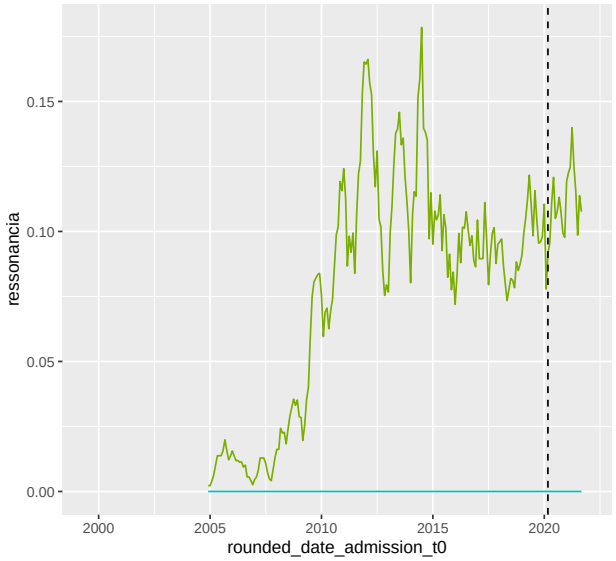

arteriografia
0.0075
0.0050
0.0025
0.0000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

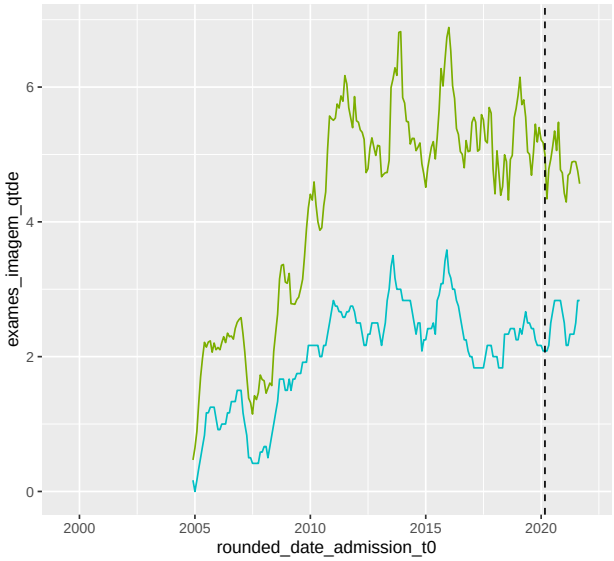
cavografia
0.09
0.06
0.03
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

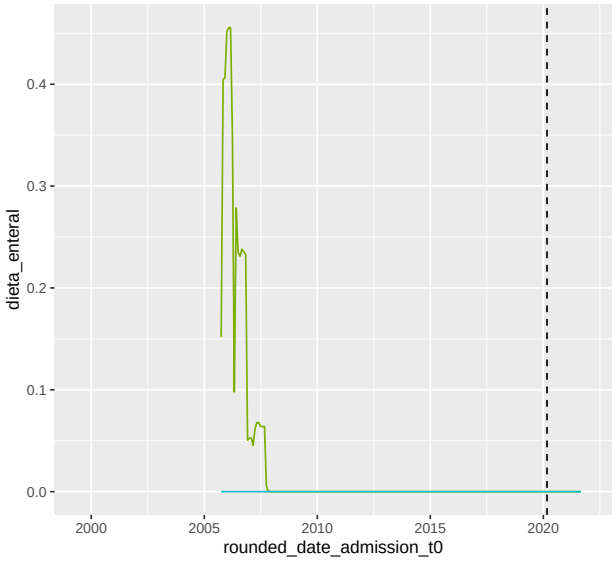

cintilografia
0.20
0.15
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

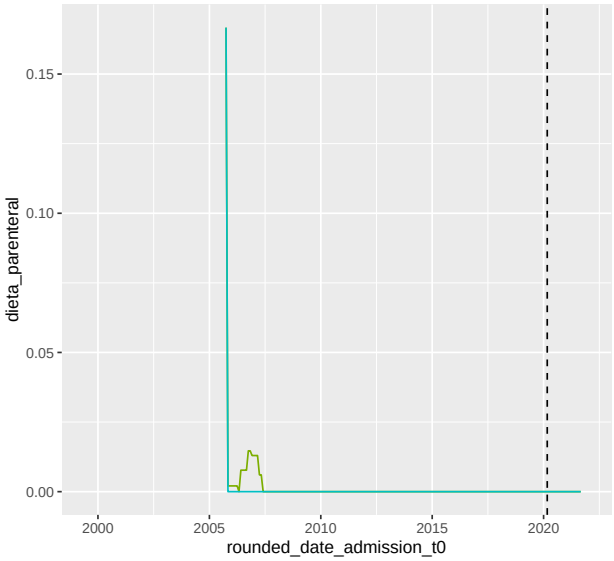
ecocardiograma
0.75
0.50
0.25
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

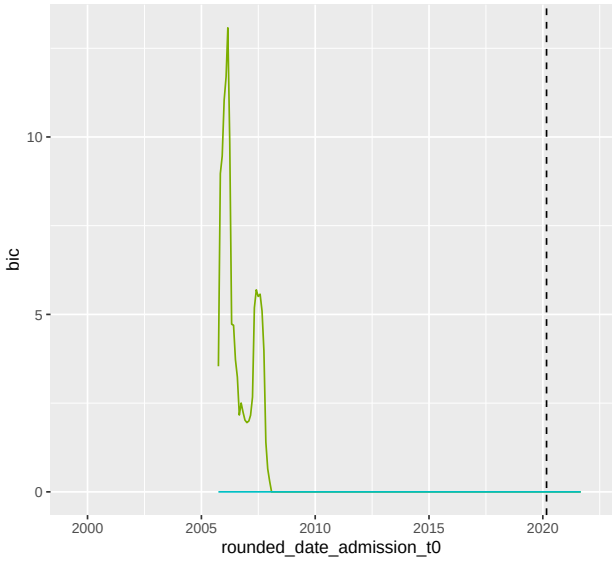

endoscopia
0.06
0.04
0.02
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
flebografia
0.125
0.100
0.075
0.050
0.025
0.000
pet_ct
0.02
0.01
0.00
ultrassom
0.4
0.3
0.2
0.1
0.0
tomografia
0.5
radiografia
5
4
3
2
1
0

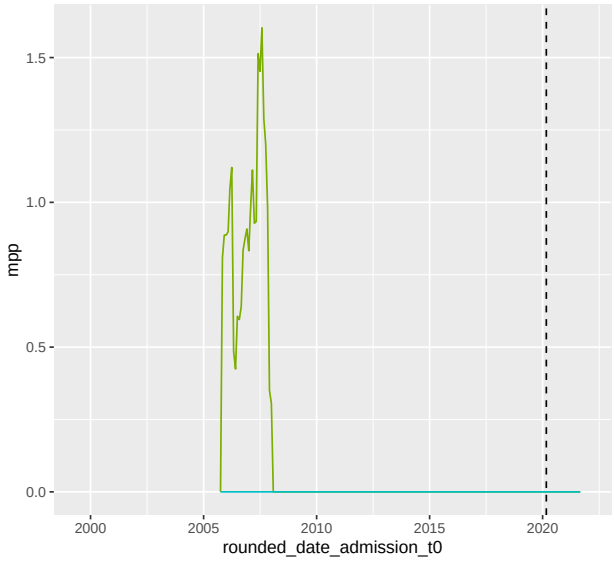

ressonancia
0.15
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

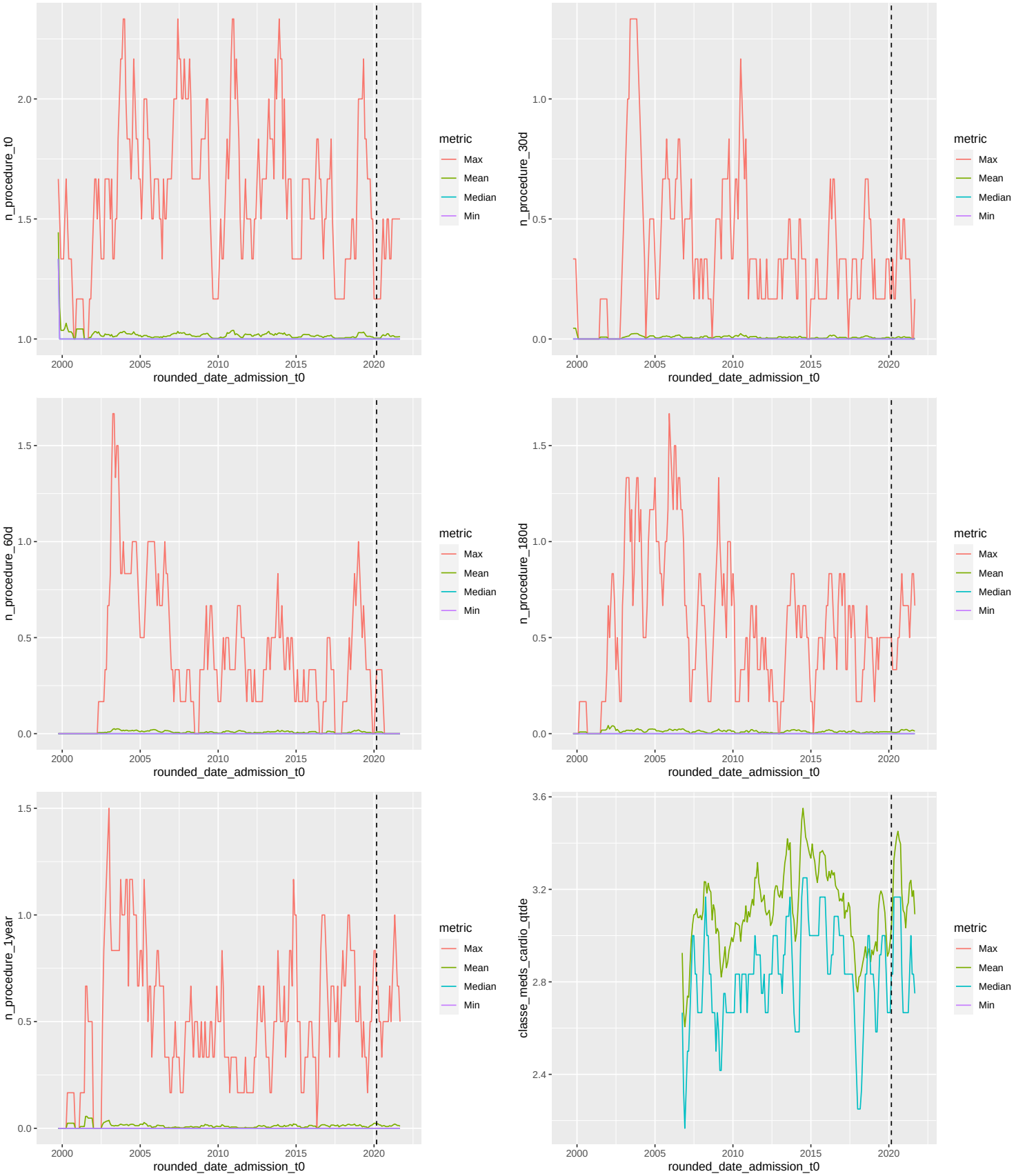

exames_imagem_qtde
6
4
2
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

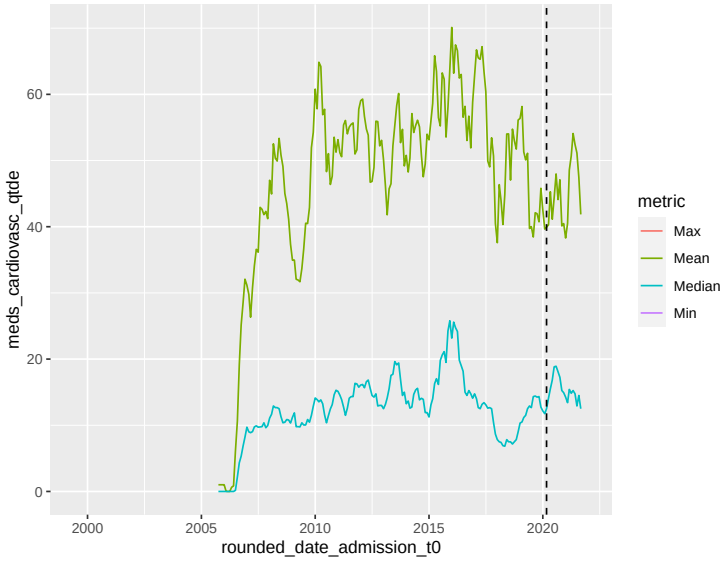

dieta_enteral
0.4
0.3
0.2
0.1
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

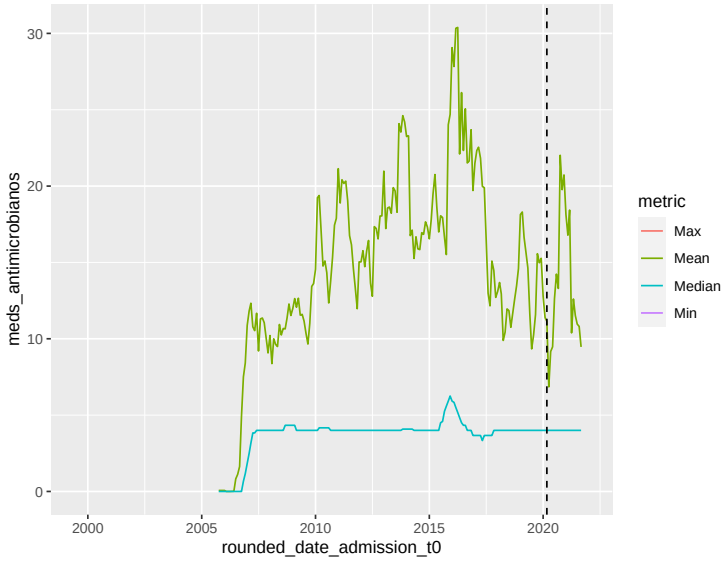

dieta_parenteral
0.15
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min


bic
10
5
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min


mpp
1.5
1.0
0.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

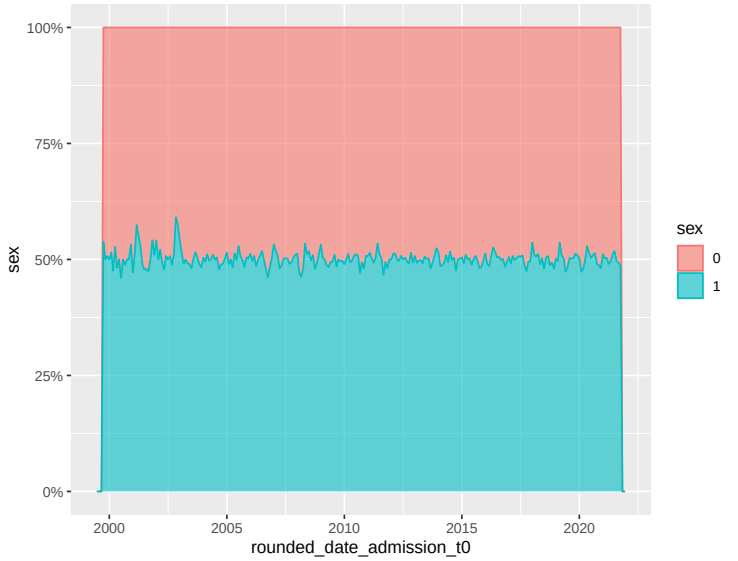

n_procedure_t0
n_procedure_30d
n_procedure_60d
n_procedure_180d
n_procedure_1year
classe_meds_cardio_qtde
rounded_date_admission_t0
metric
Max
Mean
Median
Min

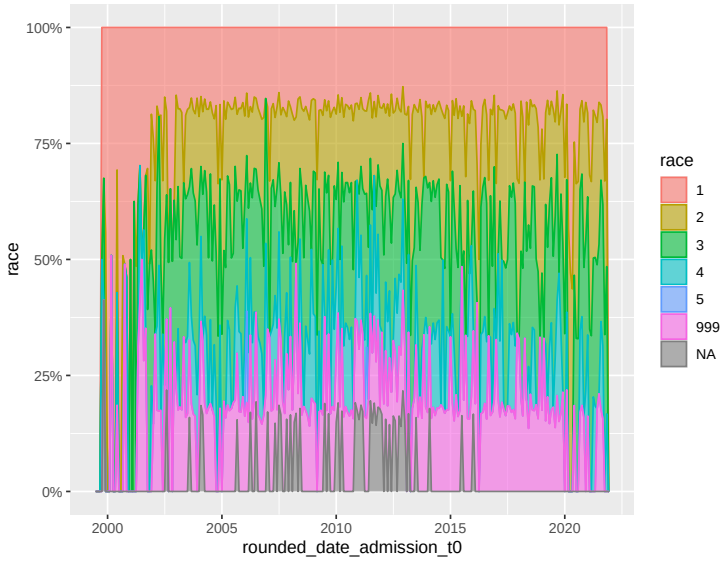


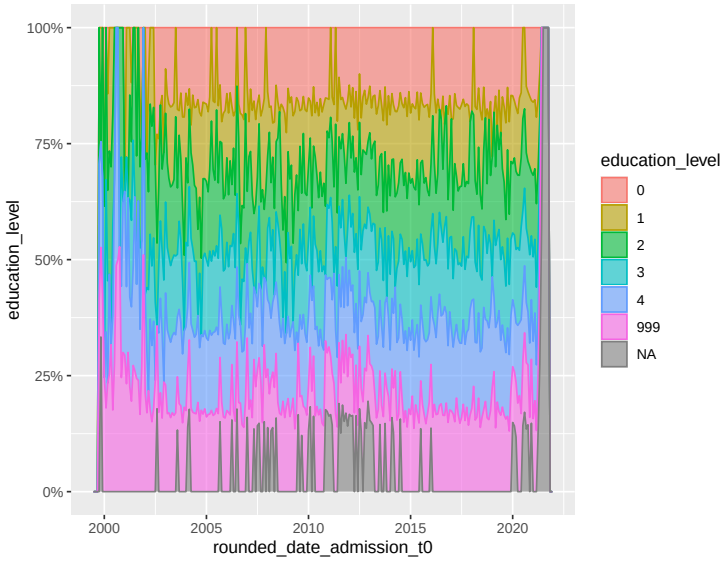


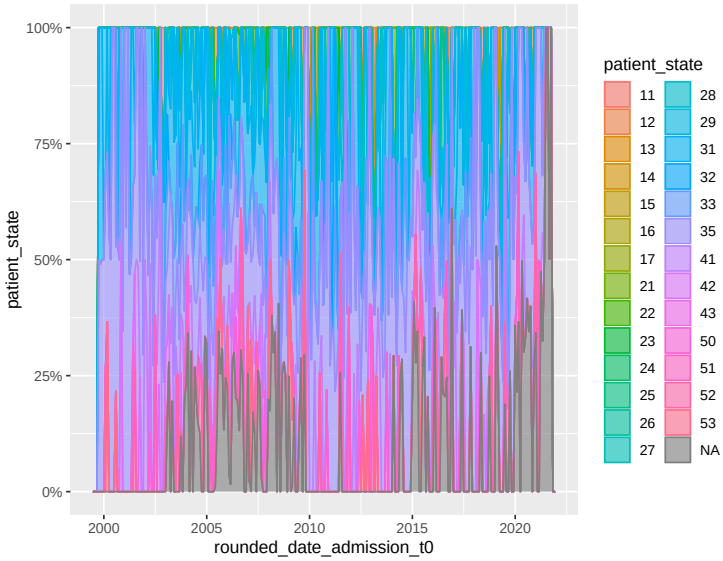

```
column <- 'death_intraop_2'

df[columns_list$outcome_columns] <- lapply(df[columns_list$outcome_columns], factor)

for (column in c(columns_list$categorical_columns, columns_list$outcome_columns)){
  if (mean(is.na(df[[column]])) > 0.5) next
  if (length(unique(df[[column]])) > 10) print(column)

  df %>%
    mutate(rounded_date_admission_t0 = lubridate::floor_date(date_admission_t0,
                                                             'month')) %>%
    group_by(rounded_date_admission_t0, !!sym(column)) %>%
    summarise(N = n()) %>%
    ungroup %>%
    arrange(rounded_date_admission_t0) %>%
    mutate(N = zoo::rollmean(N, k = k, fill = NA)) %>%
    group_by(rounded_date_admission_t0) %>%
    mutate(percentage = N / sum(N)) %>%
    ungroup %>%
    tidyr::complete(rounded_date_admission_t0, !!sym(column),
                    fill = list(N = 0, percentage = 0)) %>%
    ggplot(aes(x = rounded_date_admission_t0, y = percentage,
               color = !!sym(column), fill = !!sym(column))) +
      geom_area(alpha=0.6) +
      scale_y_continuous(labels = scales::percent) +
      labs(y = column)-> p

  print(p)
}
```

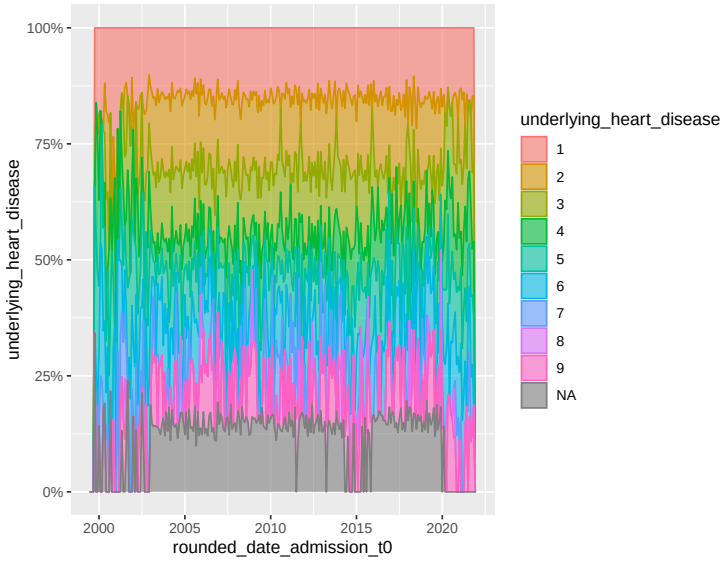
```
## [1] "patient_state"
```

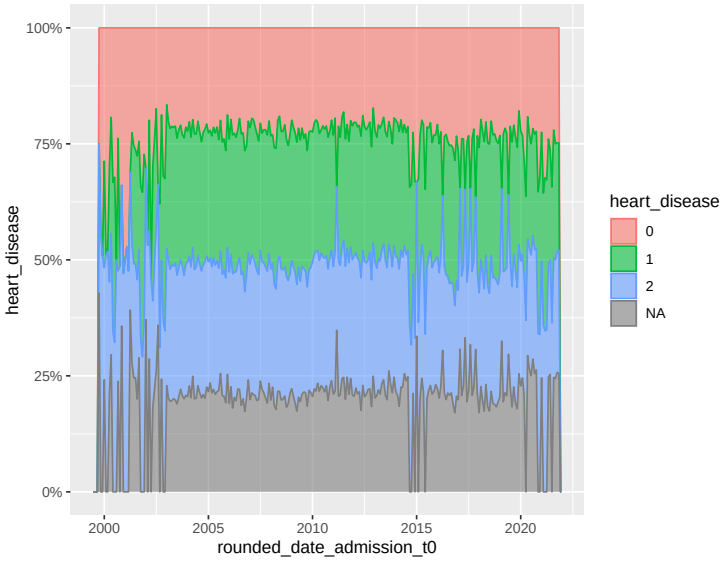

100%
75%
50%
25%
0%
sex
2000
2005
2010
2015
2020
rounded_date_admission_t0
sex
0
1

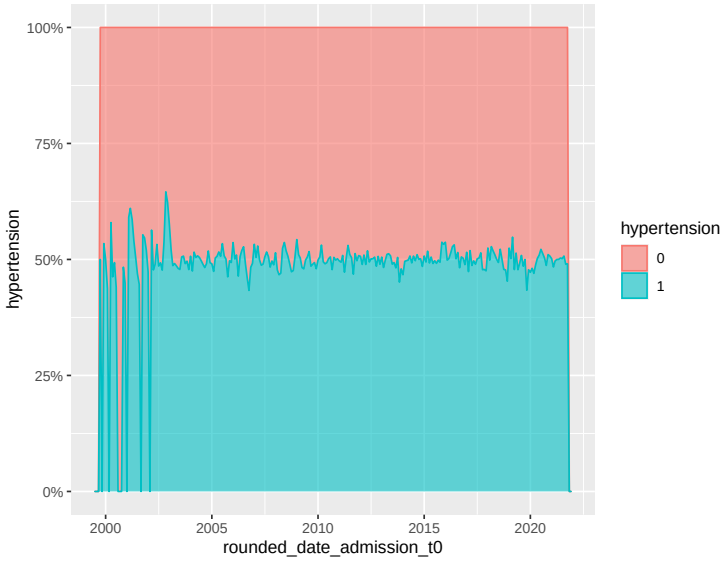

100%
75%
50%
25%
0%
race
2000
2005
2010
2015
2020
rounded_date_admission_t0
race
1
2
3
4
5
999
NA

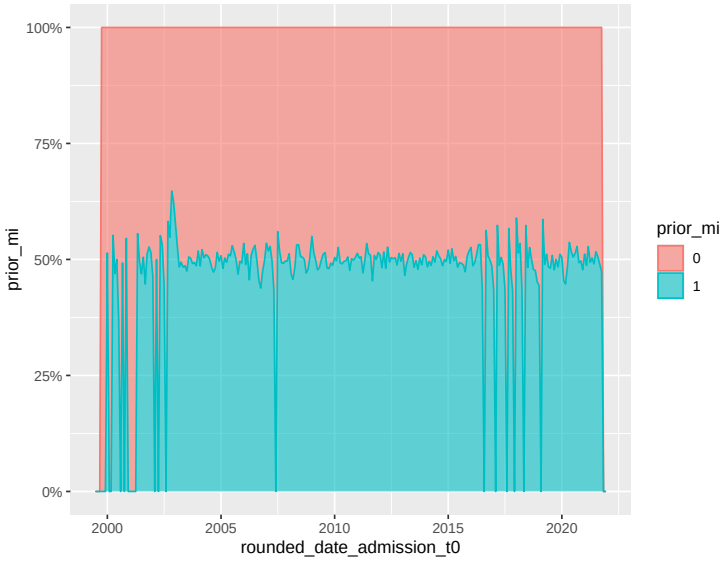

100%
75%
50%
25%
0%
education_level
2000
2005
2010
2015
2020
rounded_date_admission_t0
education_level
0
1
2
3
4
999
NA

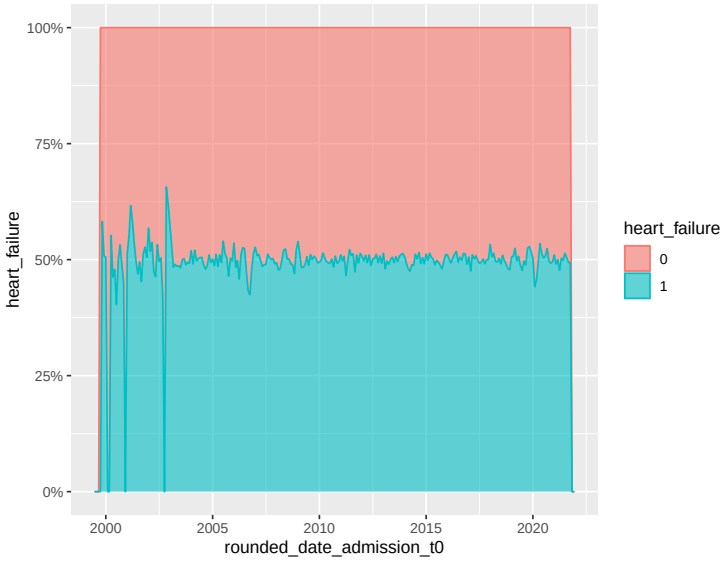

100%
75%
50%
25%
0%
patient_state
2000
2005
2010
2015
2020
rounded_date_admission_t0
patient_state
11
12
13
14
15
16
17
21
22
23
24
25
26
27
28
29
31
32
33
35
41
42
43
50
51
52
53
NA

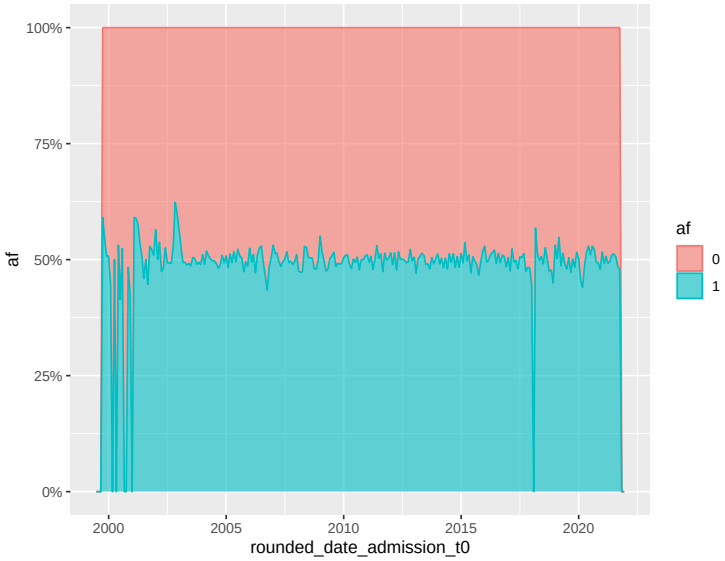

100%
75%
50%
25%
0%
underlying_heart_disease
2000
2005
2010
2015
2020
rounded_date_admission_t0
underlying_heart_disease
1
2
3
4
5
6
7
8
9
NA

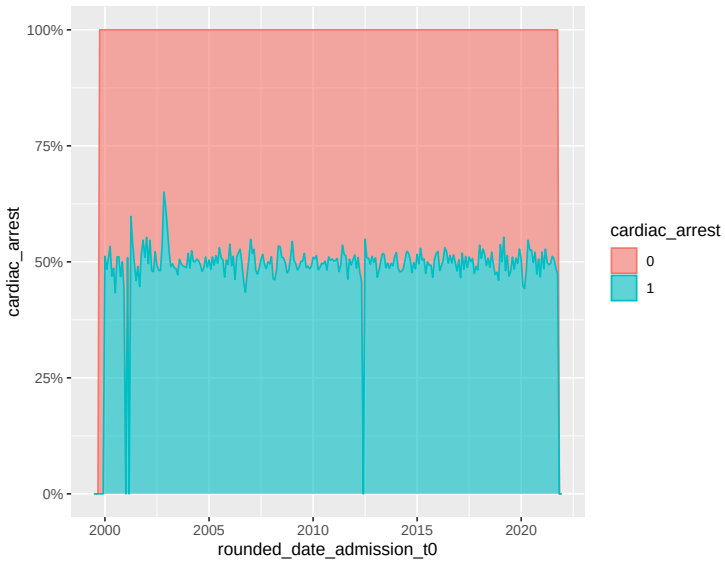

100%
75%
50%
25%
0%
heart_disease
2000
2005
2010
2015
2020
rounded_date_admission_t0
heart_disease
0
1
2
NA

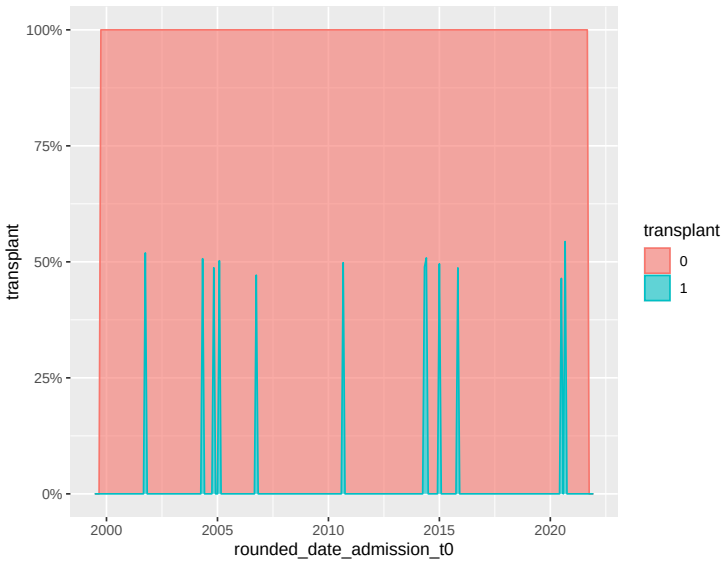

hypertension
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
hypertension
0
1

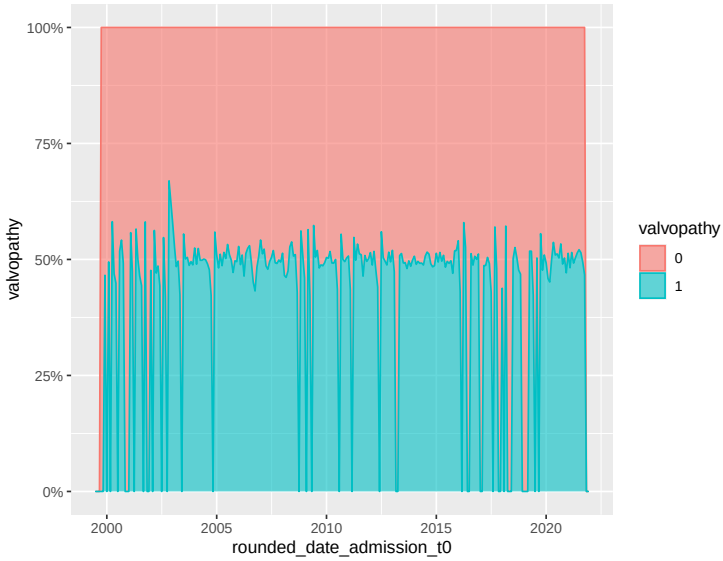

prior_mi
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
prior_mi
0
1

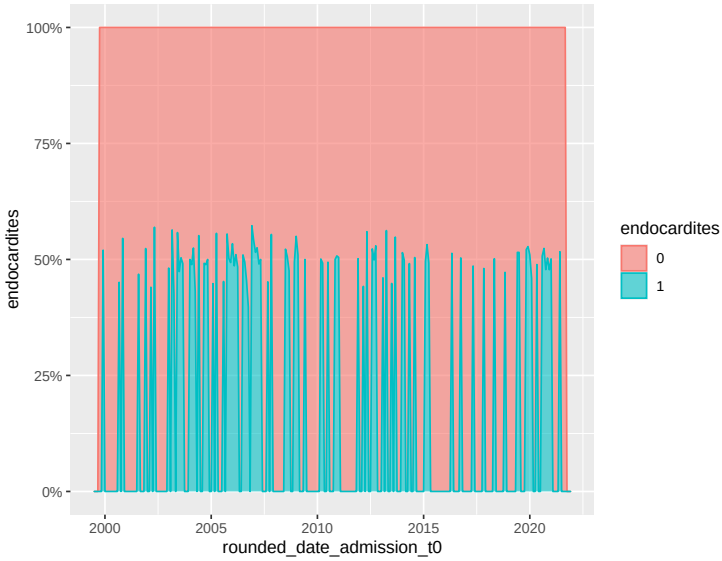

heart_failure
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
heart_failure
0
1

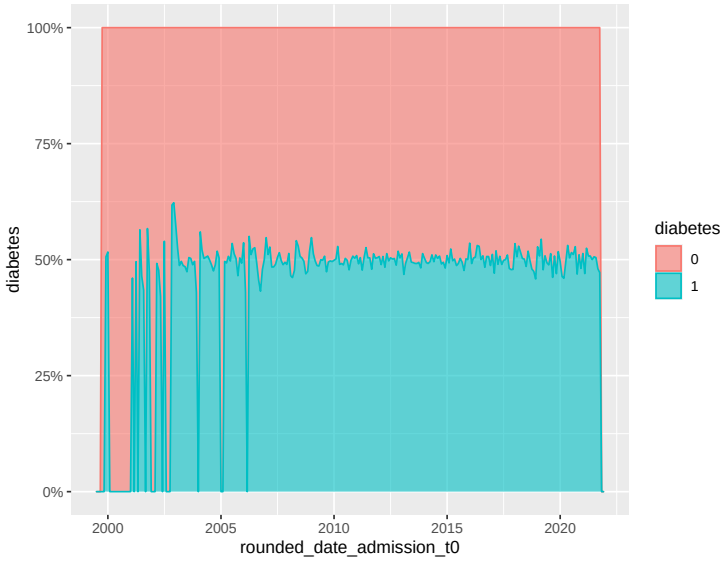

af
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
af
0
1

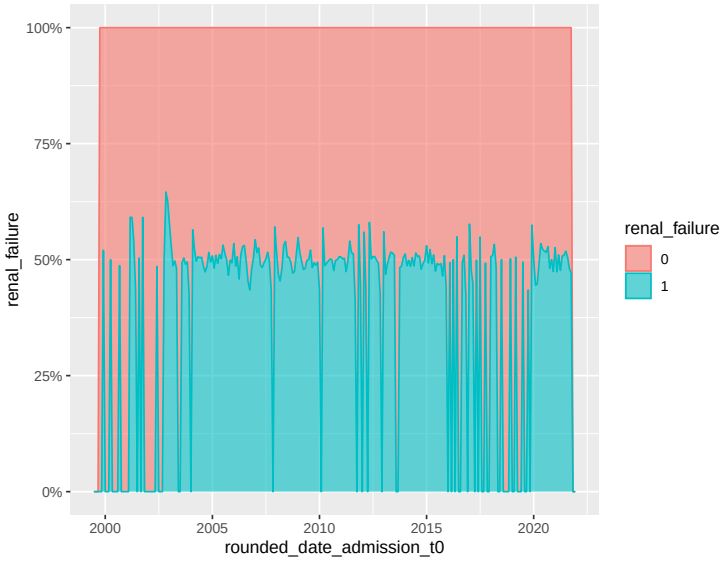

cardiac_arrest
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
cardiac_arrest
0
1

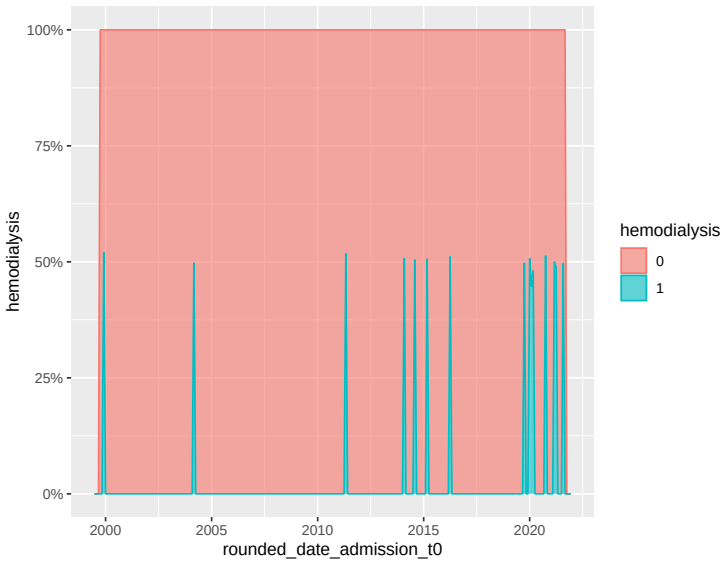

transplant
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
transplant
0
1

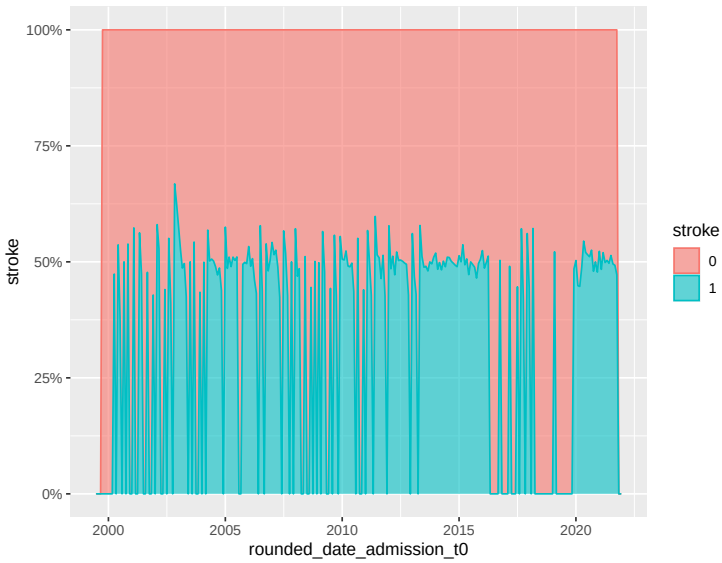

valvopathy
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
valvopathy
0
1

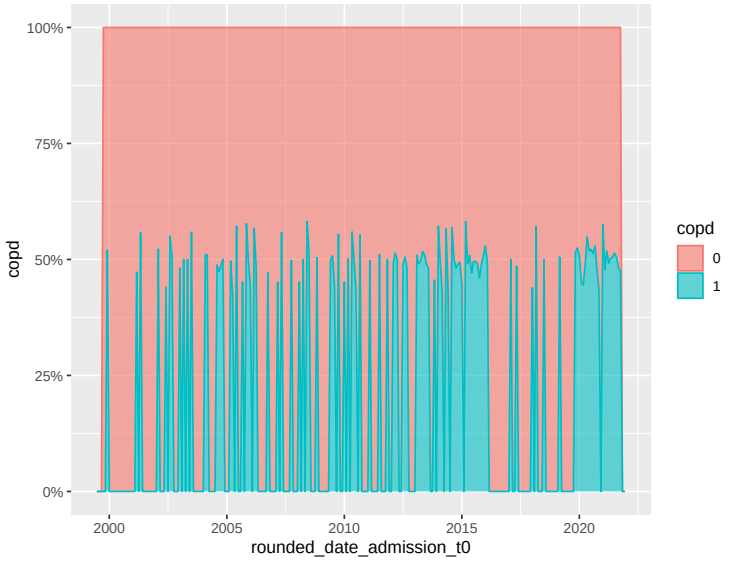

endocardites
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
endocardites
0
1

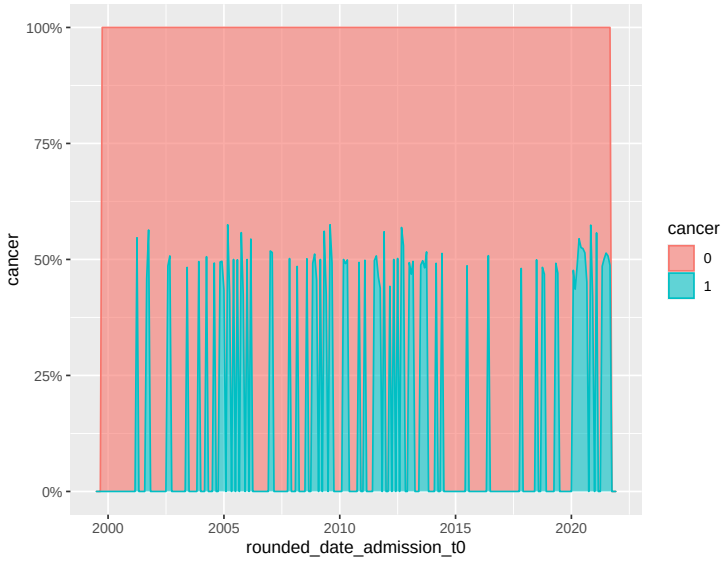

diabetes
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
diabetes
0
1

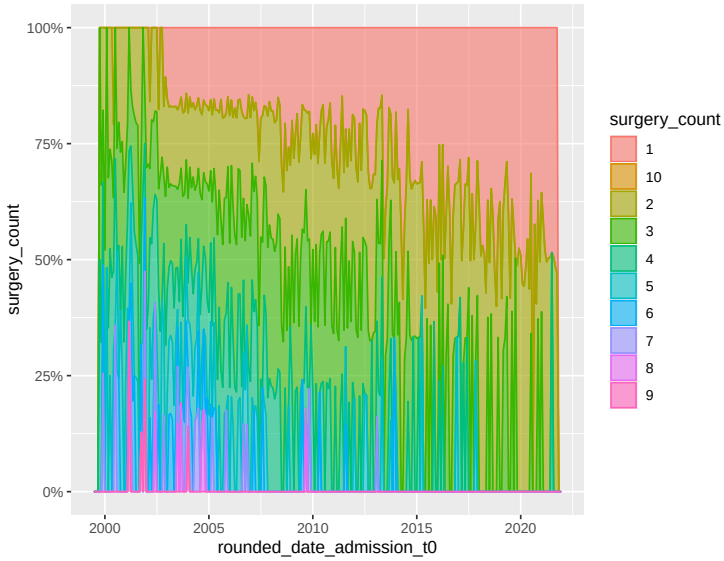

renal_failure
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
renal_failure
0
1

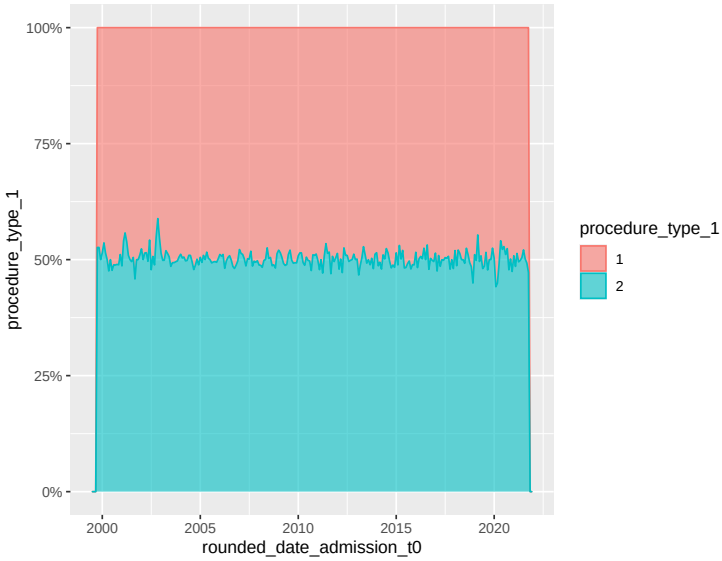

hemodialysis
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
hemodialysis
0
1

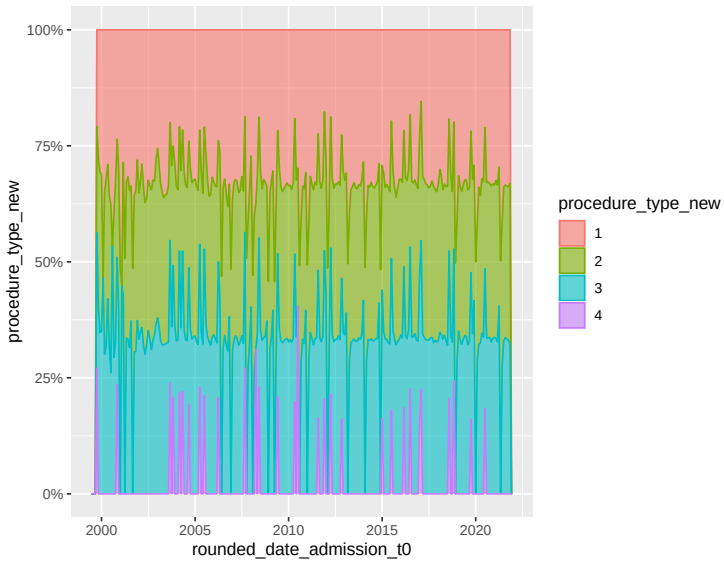

stroke
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
stroke
0
1

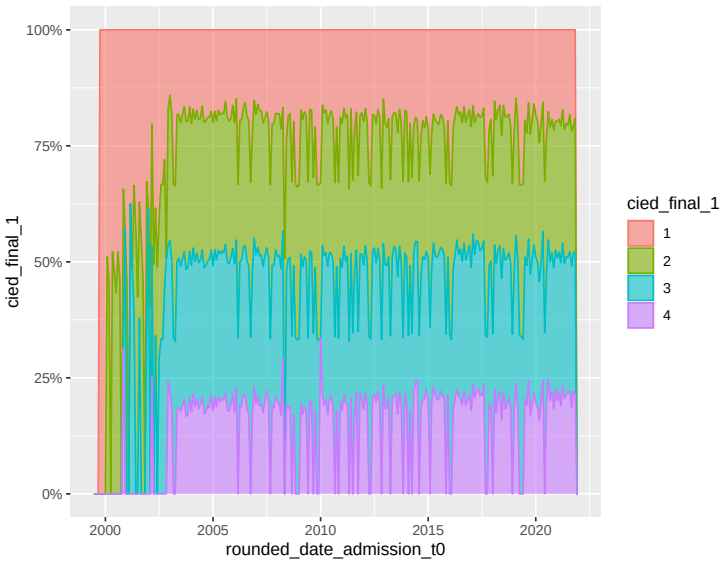

100%
75%
50%
25%
0%
copd
2000
2005
2010
2015
2020
rounded_date_admission_t0
copd
0
1

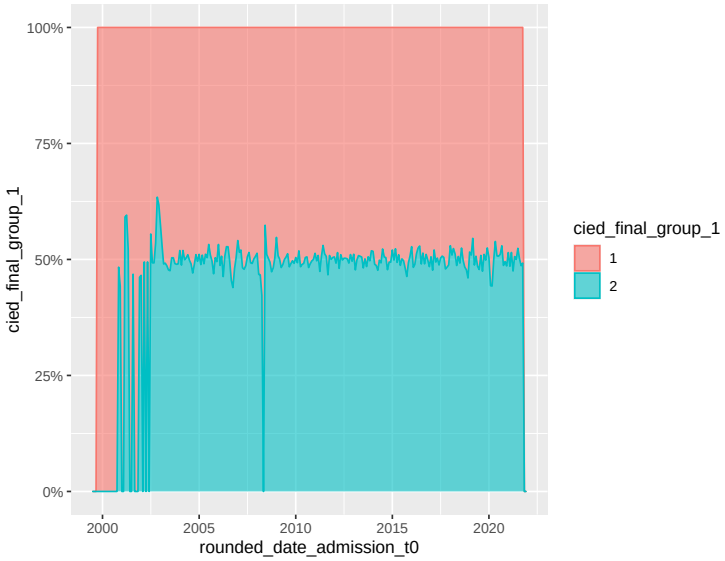

100%
75%
50%
25%
0%
cancer
2000
2005
2010
2015
2020
rounded_date_admission_t0
cancer
0
1

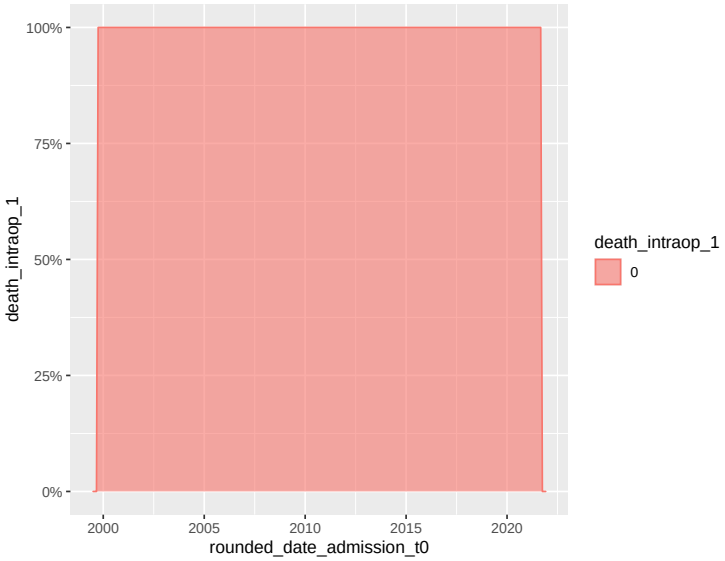

100%
75%
50%
25%
0%
surgery_count
2000
2005
2010
2015
2020
rounded_date_admission_t0
surgery_count
1
10
2
3
4
5
6
7
8
9

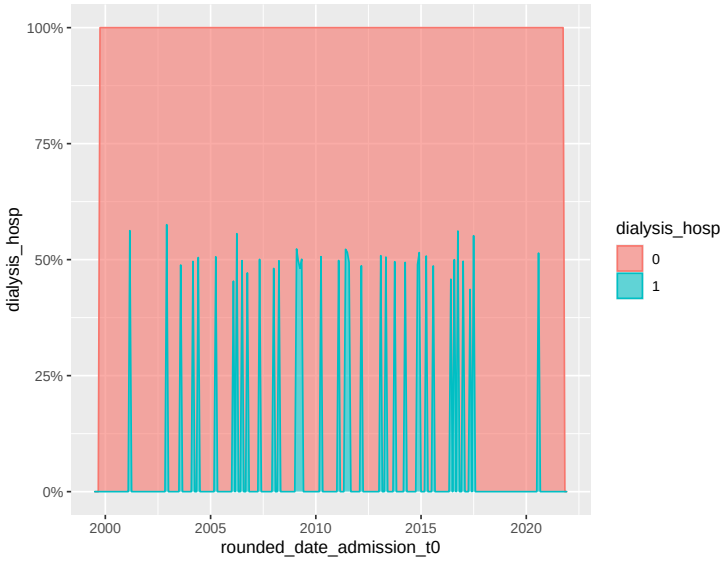

100%
75%
50%
25%
0%
procedure_type_1
2000
2005
2010
2015
2020
rounded_date_admission_t0
procedure_type_1
1
2

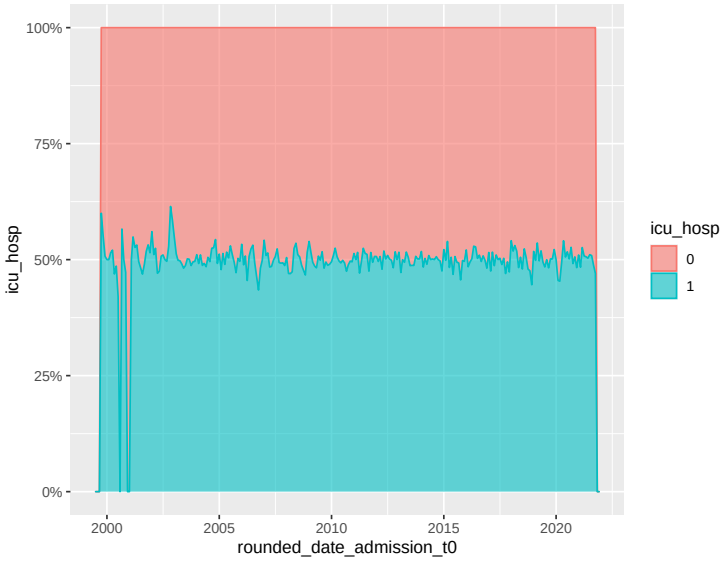

100%
75%
50%
25%
0%
procedure_type_new
2000
2005
2010
2015
2020
rounded_date_admission_t0
procedure_type_new
1
2
3
4

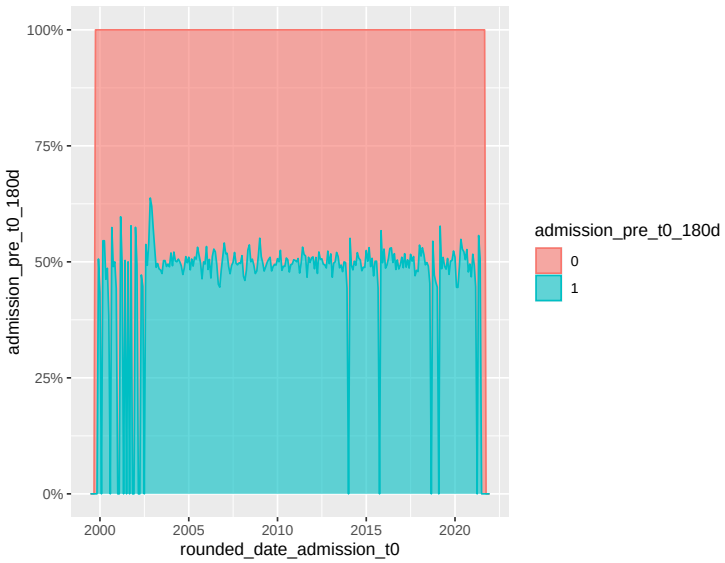

100%
75%
50%
25%
0%
cied_final_1
2000
2005
2010
2015
2020
rounded_date_admission_t0
cied_final_1
1
2
3
4

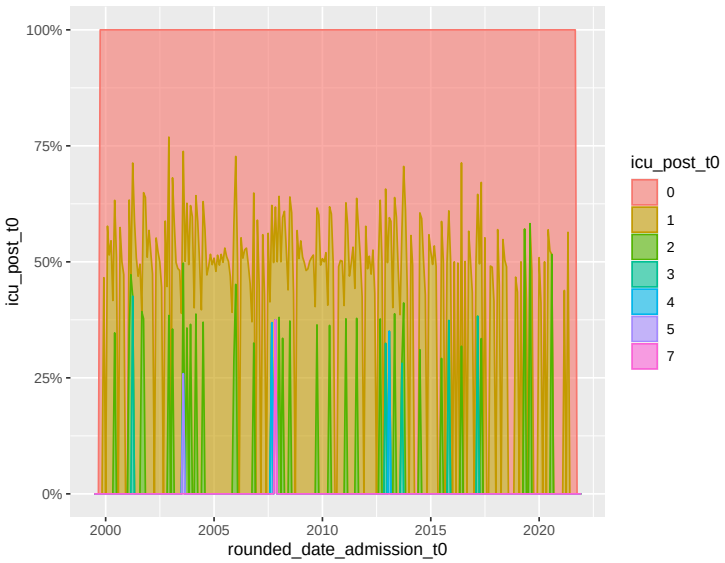

100%
75%
50%
25%
0%
cied_final_group_1
2000
2005
2010
2015
2020
rounded_date_admission_t0
cied_final_group_1
1
2

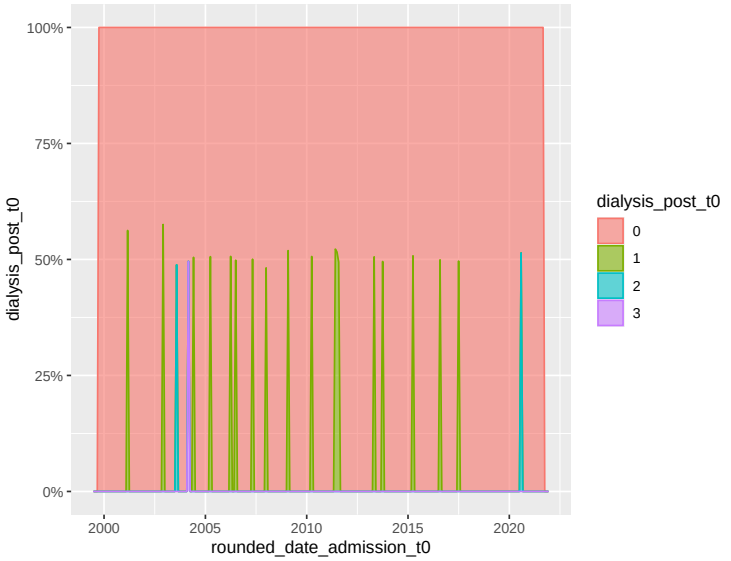

100%
75%
50%
25%
0%
death_intraop_1
2000
2005
2010
2015
2020
rounded_date_admission_t0
death_intraop_1
0

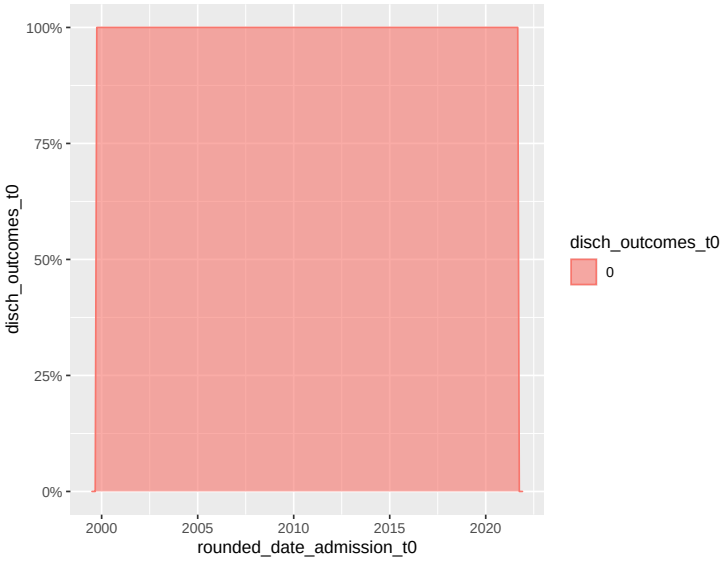

100%
75%
50%
25%
0%
dialysis_hosp
2000
2005
2010
2015
2020
rounded_date_admission_t0
dialysis_hosp
0
1

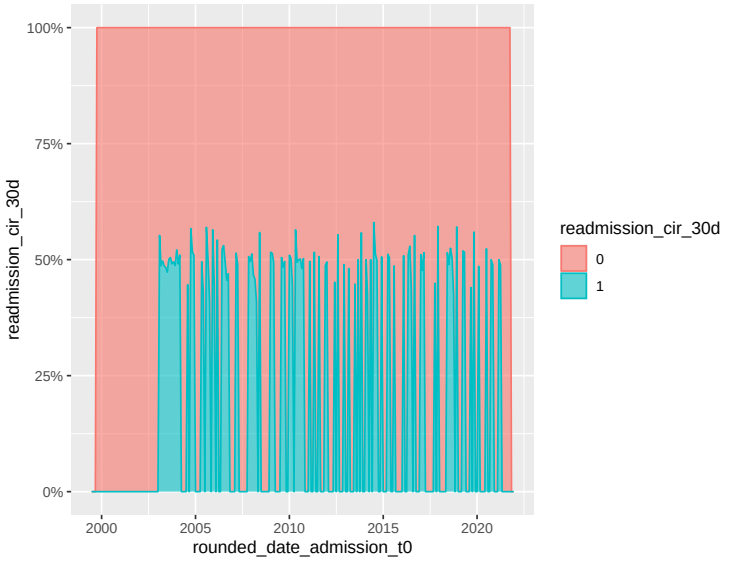

100%
75%
50%
25%
0%
icu_hosp
2000
2005
2010
2015
2020
rounded_date_admission_t0
icu_hosp
0
1

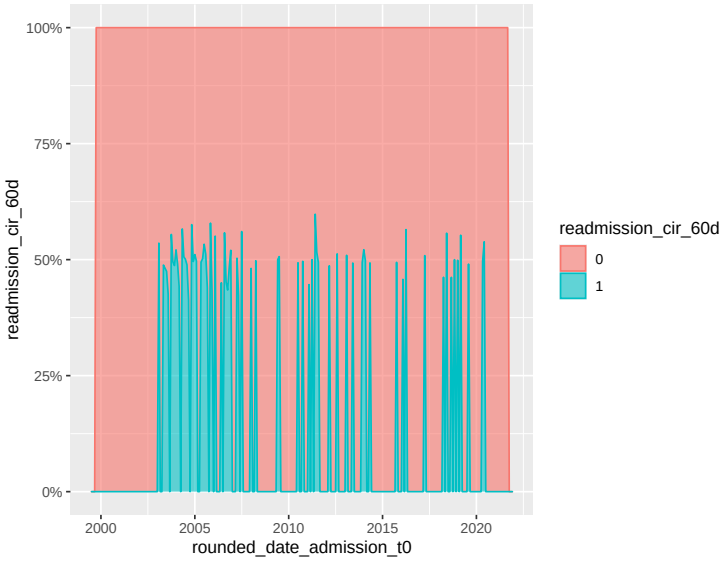

100%
75%
50%
25%
0%
admission_pre_t0_180d
2000
2005
2010
2015
2020
rounded_date_admission_t0
admission_pre_t0_180d
0
1

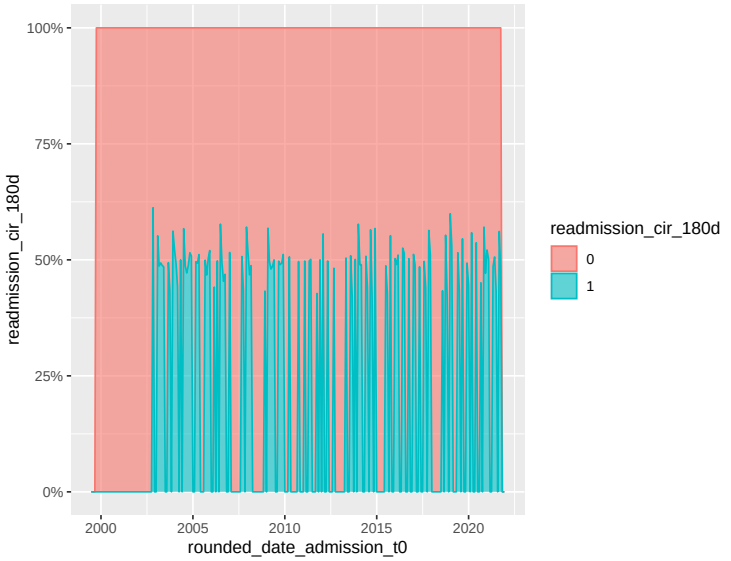

100%
75%
50%
25%
0%
icu_post_t0
2000
2005
2010
2015
2020
rounded_date_admission_t0
icu_post_t0
0
1
2
3
4
5
7

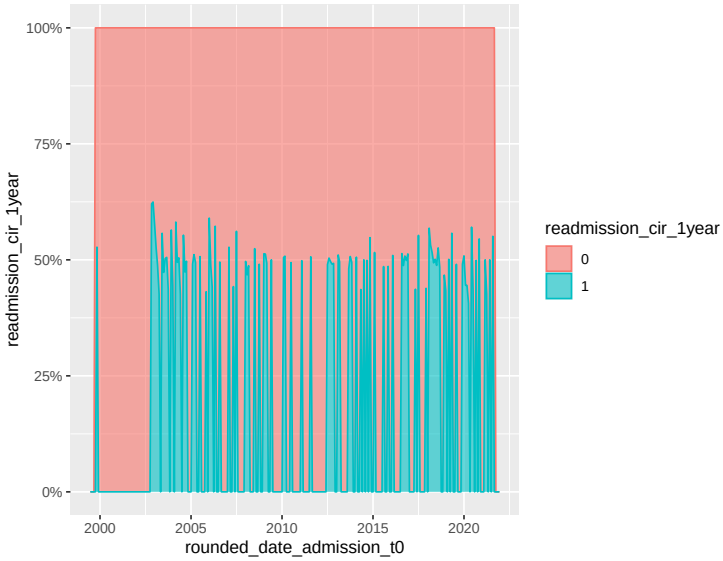

100%
75%
50%
25%
0%
dialysis_post_t0
2000
2005
2010
2015
2020
rounded_date_admission_t0
dialysis_post_t0
0
1
2
3

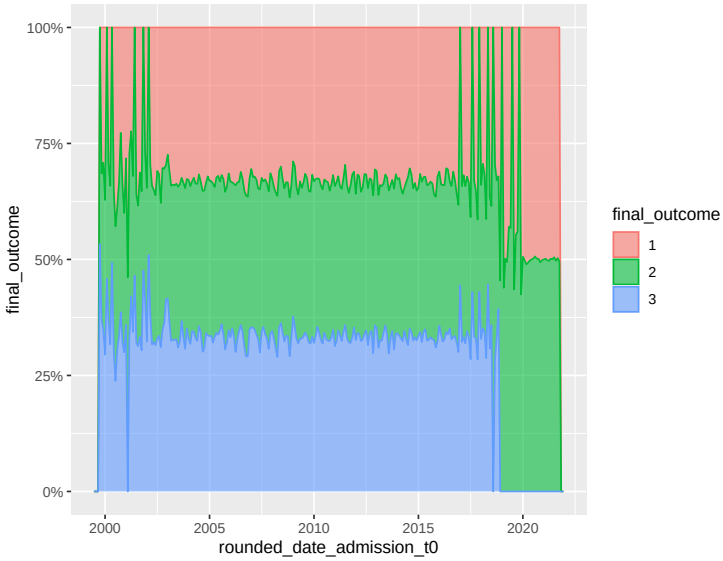

100%
75%
50%
25%
0%
disch_outcomes_t0
2000
2005
2010
2015
2020
rounded_date_admission_t0
disch_outcomes_t0
0

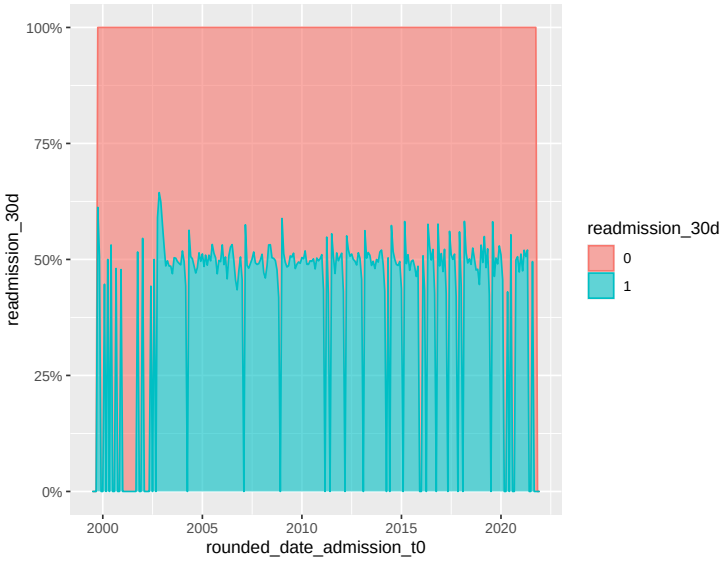

100%
75%
50%
25%
0%
readmission_cir_30d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_cir_30d
0
1

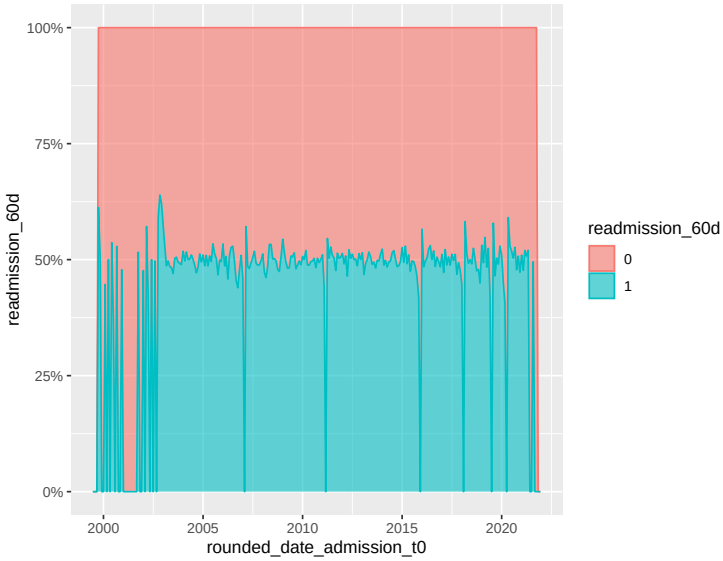

100%
75%
50%
25%
0%
readmission_cir_60d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_cir_60d
0
1

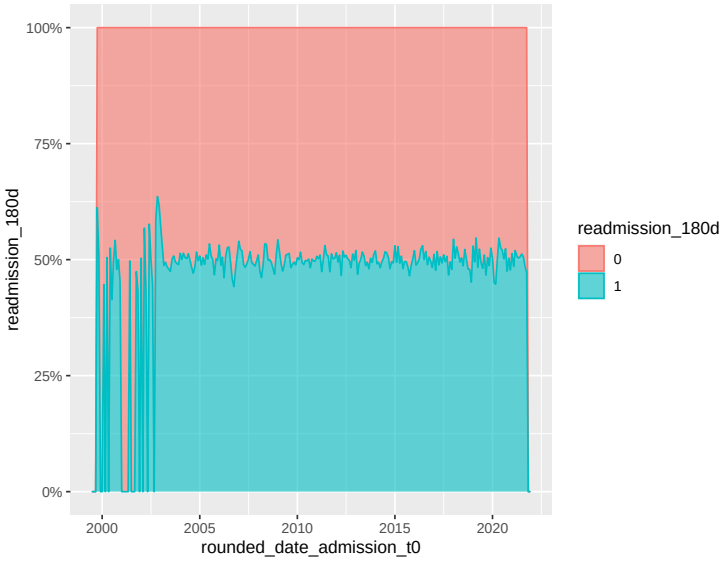

100%
75%
50%
25%
0%
readmission_cir_180d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_cir_180d
0
1

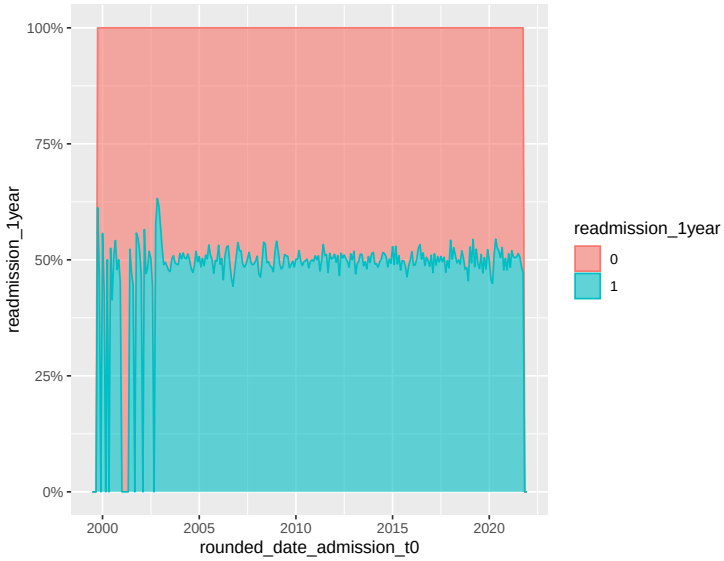

100%
75%
50%
25%
0%
readmission_cir_1year
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_cir_1year
0
1

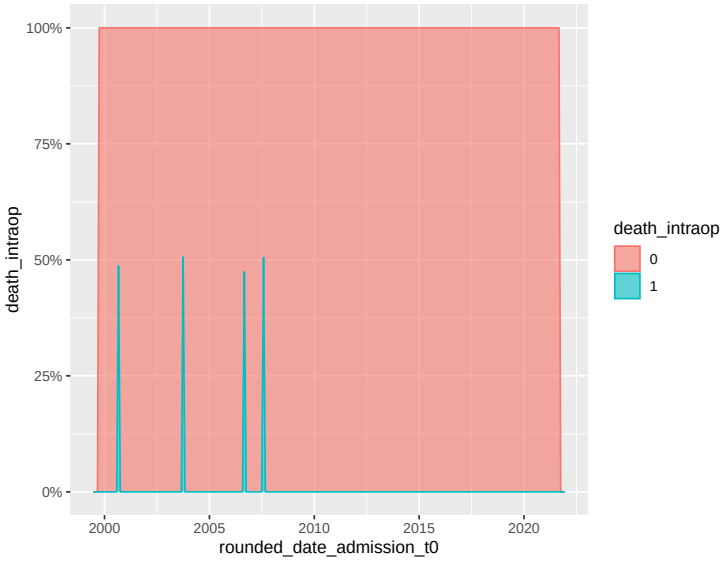

100%
75%
50%
25%
0%
final_outcome
2000
2005
2010
2015
2020
rounded_date_admission_t0
final_outcome
1
2
3

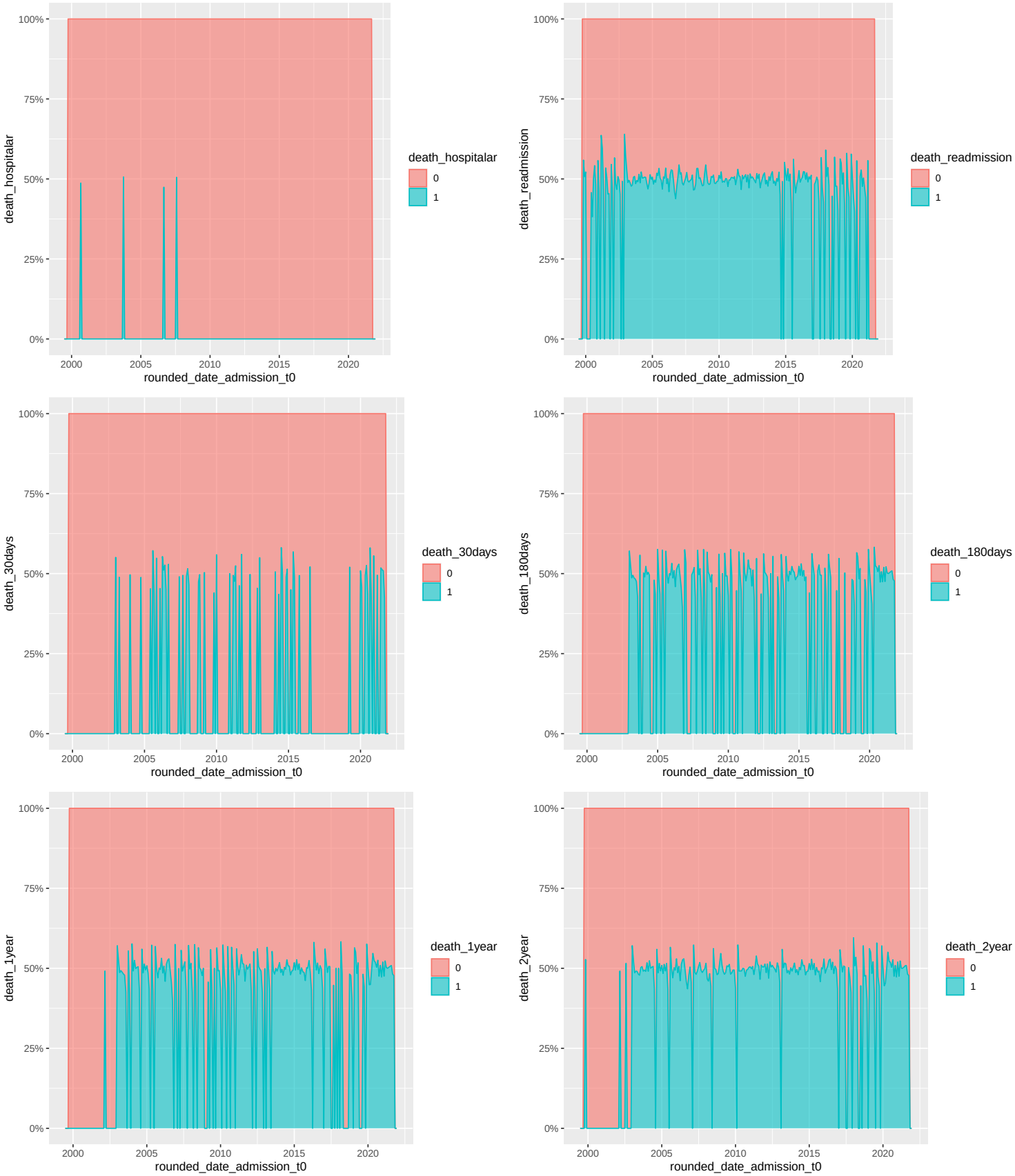

100%
75%
50%
25%
0%
readmission_30d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_30d
0
1

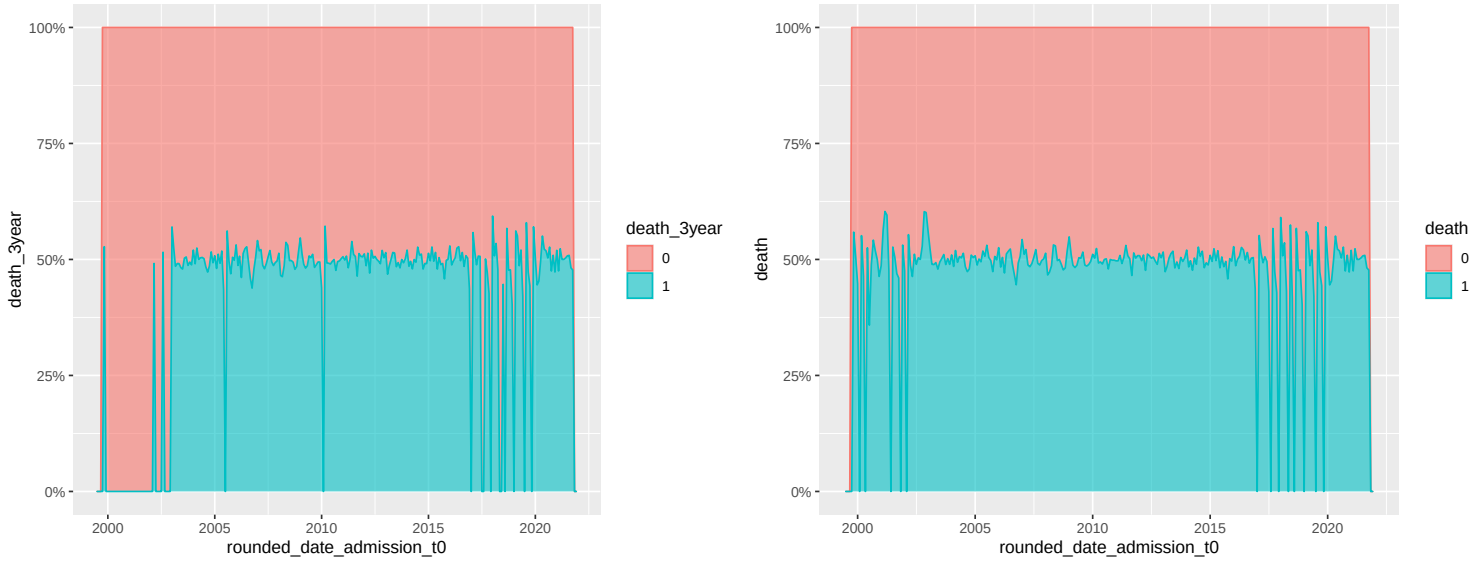

100%
75%
50%
25%
0%
readmission_60d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_60d
0
1


100%
75%
50%
25%
0%
readmission_180d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_180d
0
1


100%
75%
50%
25%
0%
readmission_1year
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_1year
0
1


100%
75%
50%
25%
0%
death_intraop
2000
2005
2010
2015
2020
rounded_date_admission_t0
death_intraop
0
1


death_hospitalar
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
death_hospitalar
0
1
death_readmission
death_readmission
death_30days
death_30days
death_180days
death_180days
death_1year
death_1year
death_2year
death_2year




```
df %>%
  mutate(rounded_date_admission_t0 = lubridate::floor_date(date_admission_t0,
                                                           'month')) %>%
  filter(rounded_date_admission_t0 <= lubridate::ymd('2021-08-01'),
         rounded_date_admission_t0 <= lubridate::ymd('2021-01-01')) %>%
  .$aco %>%
  max(na.rm=T)
```

```
## [1] 98
```

```
column <- 'aco'
df %>%
    mutate(rounded_date_admission_t0 = lubridate::floor_date(date_admission_t0,
                                                             'month')) %>%
    group_by(rounded_date_admission_t0) %>%
    summarise('Mean' = mean(!!sym(column), na.rm = T),
              'Min' = min(!!sym(column), na.rm = T),
              'Median' = medianWithoutNA(!!sym(column)),
              'Max' = max(!!sym(column), na.rm = T)) %>%
    ungroup %>%
    arrange(rounded_date_admission_t0) %>%
    mutate(Mean = zoo::rollmean(Mean, k = k, fill = NA),
           Min = zoo::rollmean(Min, k = k, fill = NA),
           Median = zoo::rollmean(Median, k = k, fill = NA),
           Max = zoo::rollmean(Max, k = k, fill = NA)) %>%
  filter(rounded_date_admission_t0 <= lubridate::ymd('2021-12-01'),
         rounded_date_admission_t0 >= lubridate::ymd('2021-01-01'))
```

```
## # A tibble: 12 x 5
##    rounded_date_admission_t0   Mean   Min Median   Max
##    <date>                     <dbl> <dbl>  <dbl> <dbl>
##  1 2021-01-01                 0.441   NaN      0   NaN
##  2 2021-02-01                 0.565   NaN      0   NaN
##  3 2021-03-01                 0.630   NaN      0   NaN
##  4 2021-04-01                 0.738   NaN      0   NaN
##  5 2021-05-01                 0.740   NaN      0   NaN
##  6 2021-06-01                 0.709   NaN      0   NaN
##  7 2021-07-01                 0.569   NaN      0   NaN
##  8 2021-08-01                 0.415   NaN      0   NaN
##  9 2021-09-01                 0.439   NaN      0   NaN
## 10 2021-10-01                NA        NA     NA    NA
## 11 2021-11-01                NA        NA     NA    NA
## 12 2021-12-01                NA        NA     NA    NA
```