Plots

Eduardo Yuki Yada

Imports

```
library(tidyverse)
library(yaml)
```

Loading data

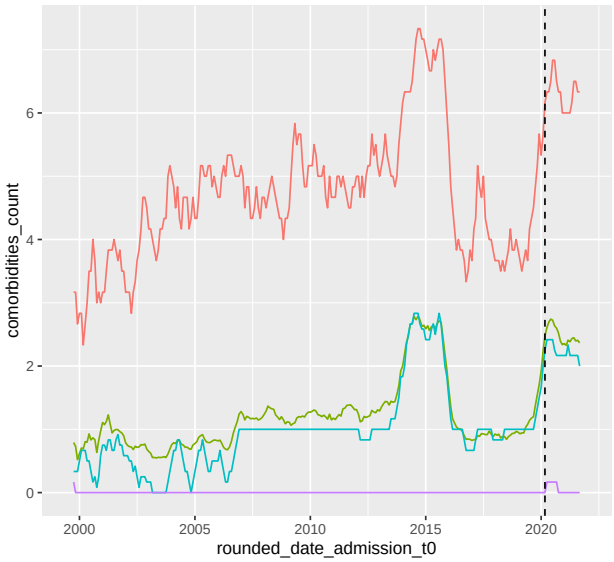
```
load('../dataset/processed_data.RData')
load('../dataset/processed_dictionary.RData')

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

outcome_column <- params$outcome_column
```

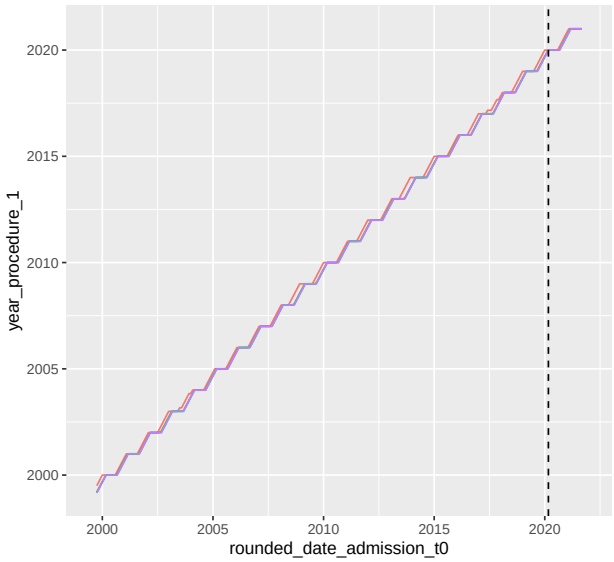
Plots

```
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
```

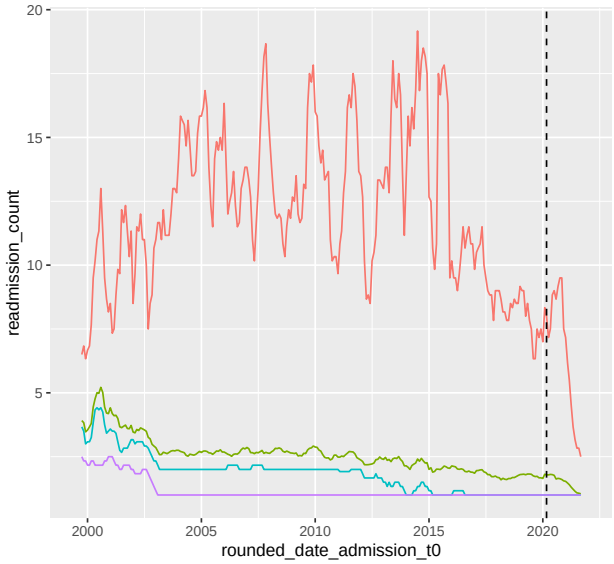
```
  print(p)
}
```

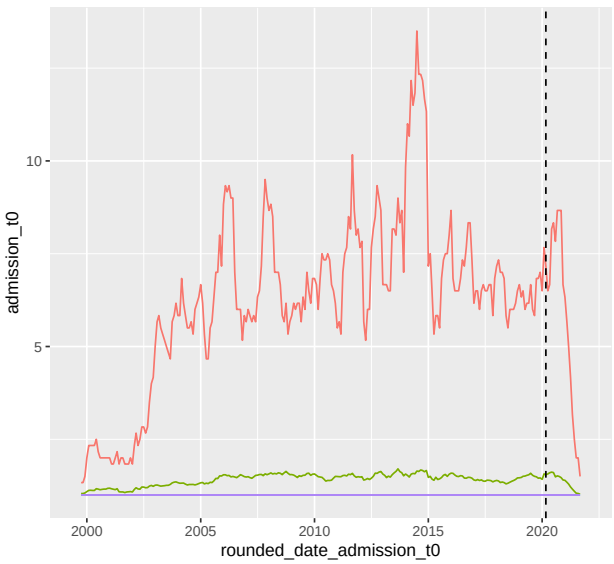
age
rounded_date_admission_t0
metric
Max
Mean
Median
Min

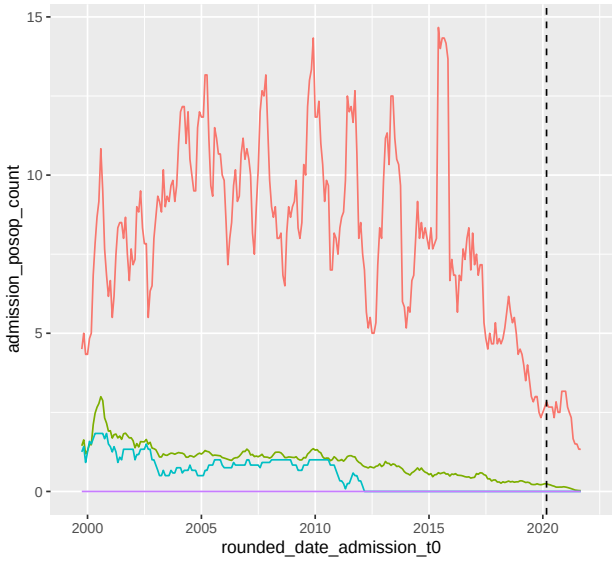

comorbidities_count
rounded_date_admission_t0
metric
Max
Mean
Median
Min

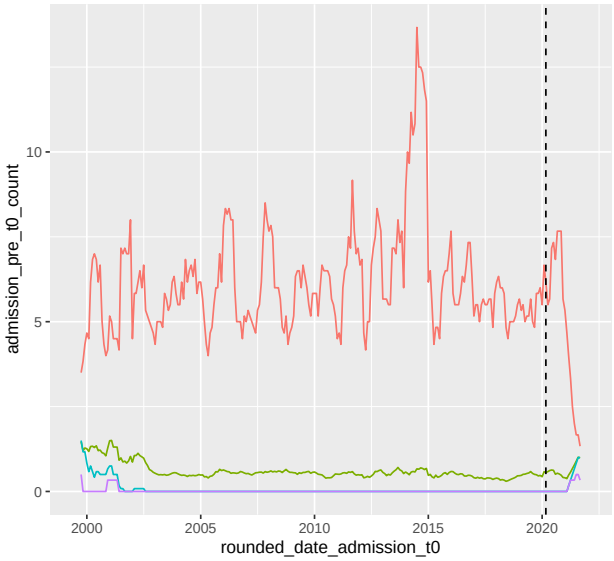

year_procedure_1
rounded_date_admission_t0
metric
Max
Mean
Median
Min

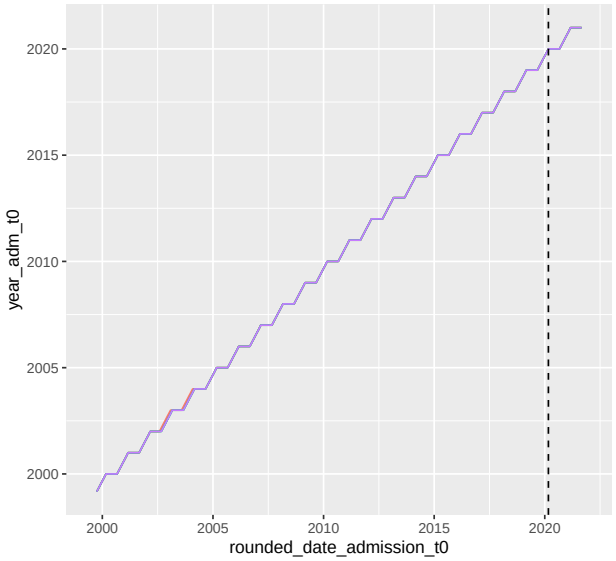
age_surgery_1
rounded_date_admission_t0
metric
Max
Mean
Median
Min

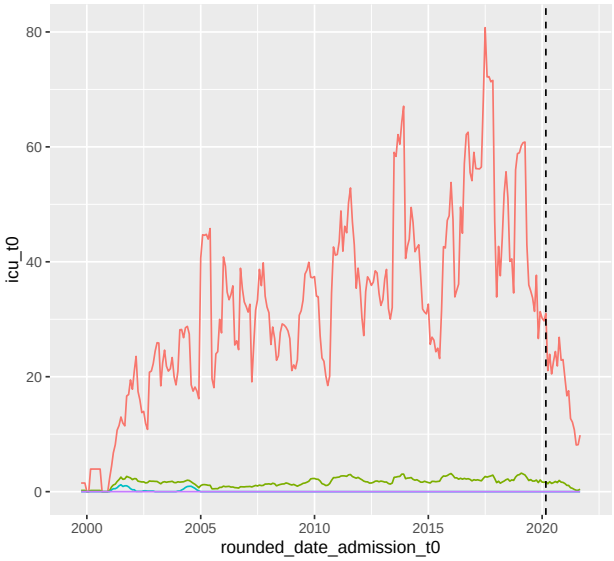

readmission_count
rounded_date_admission_t0
metric
Max
Mean
Median
Min

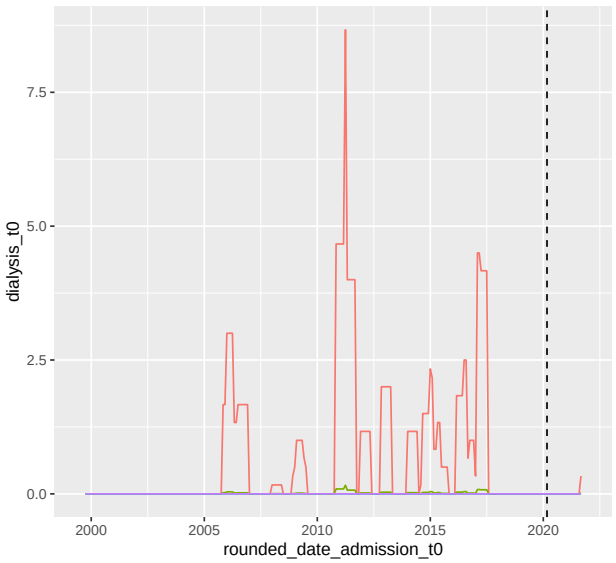

admission_t0
rounded_date_admission_t0
metric
Max
Mean
Median
Min

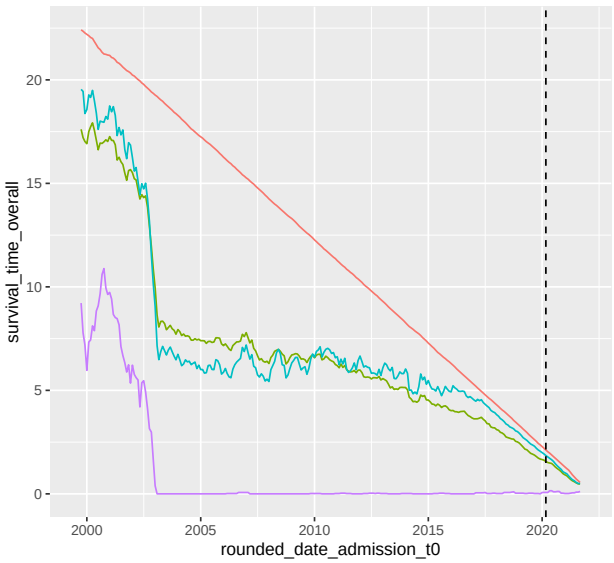

admission_posop_count
rounded_date_admission_t0
metric
Max
Mean
Median
Min

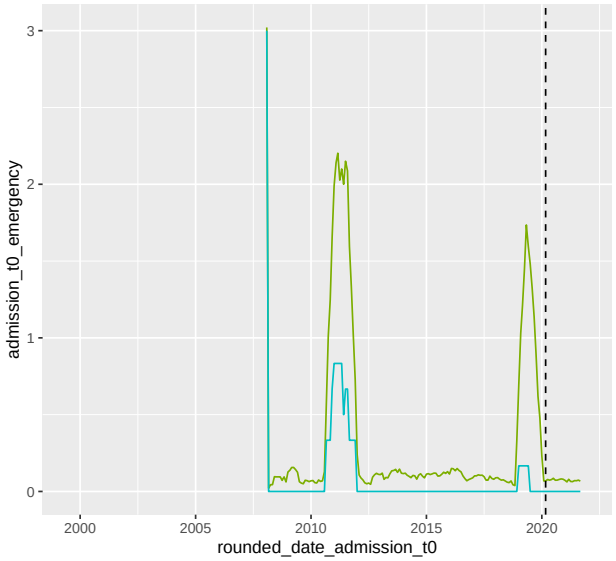

admission_pre_t0_count
rounded_date_admission_t0
metric
Max
Mean
Median
Min

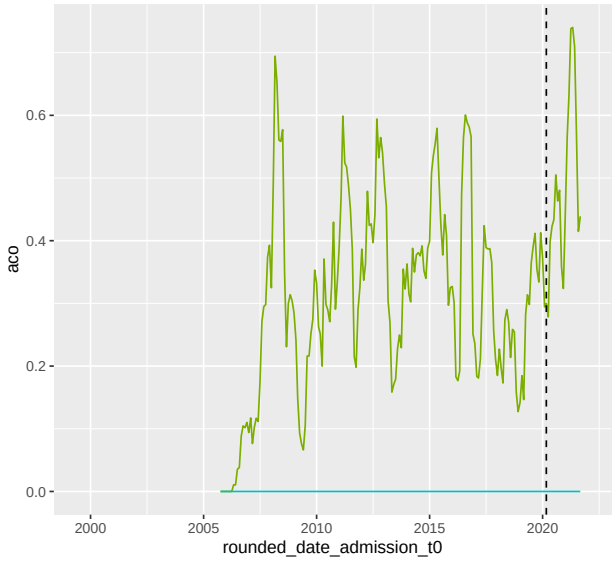

year_adm_t0
rounded_date_admission_t0
metric
Max
Mean
Median
Min

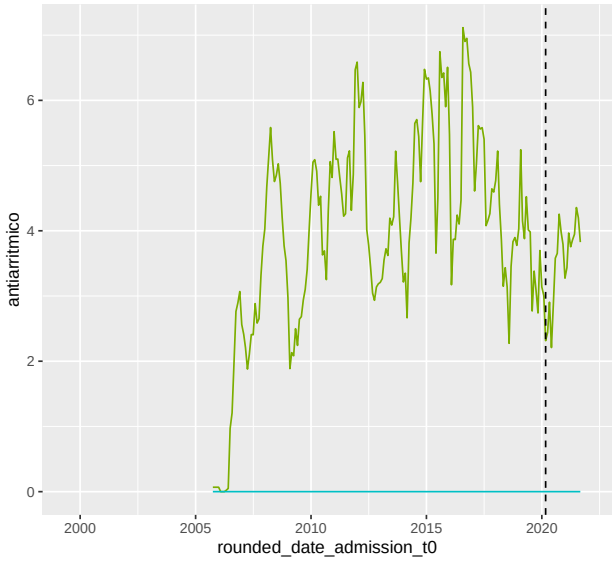

icu_t0
rounded_date_admission_t0
metric
Max
Mean
Median
Min

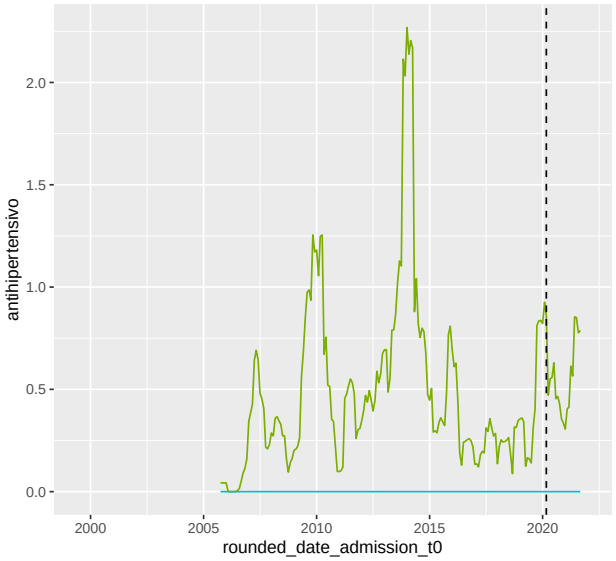

dialysis_t0
rounded_date_admission_t0
metric
Max
Mean
Median
Min

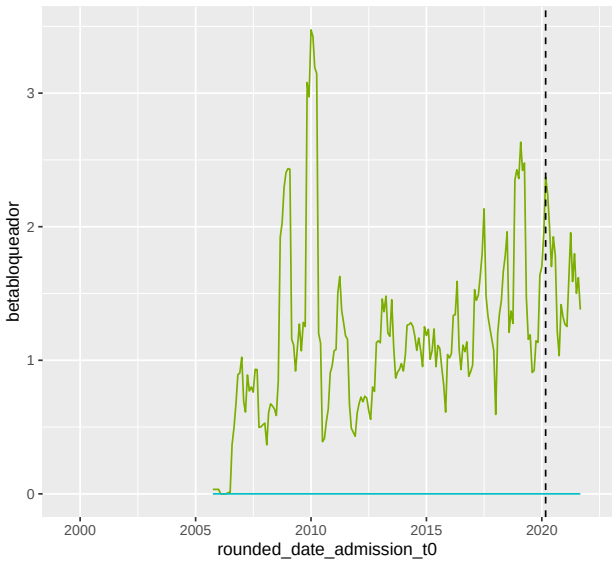

survival_time_overall
rounded_date_admission_t0
metric
Max
Mean
Median
Min

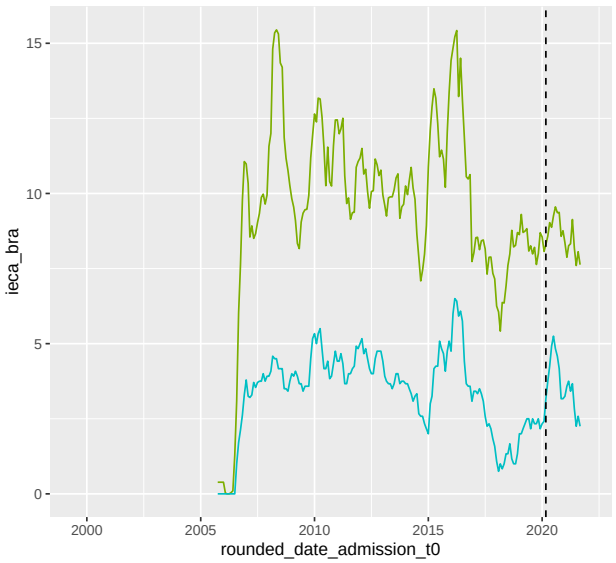

admission_t0_emergency
rounded_date_admission_t0
metric
Max
Mean
Median
Min


aco
rounded_date_admission_t0
metric
Max
Mean
Median
Min

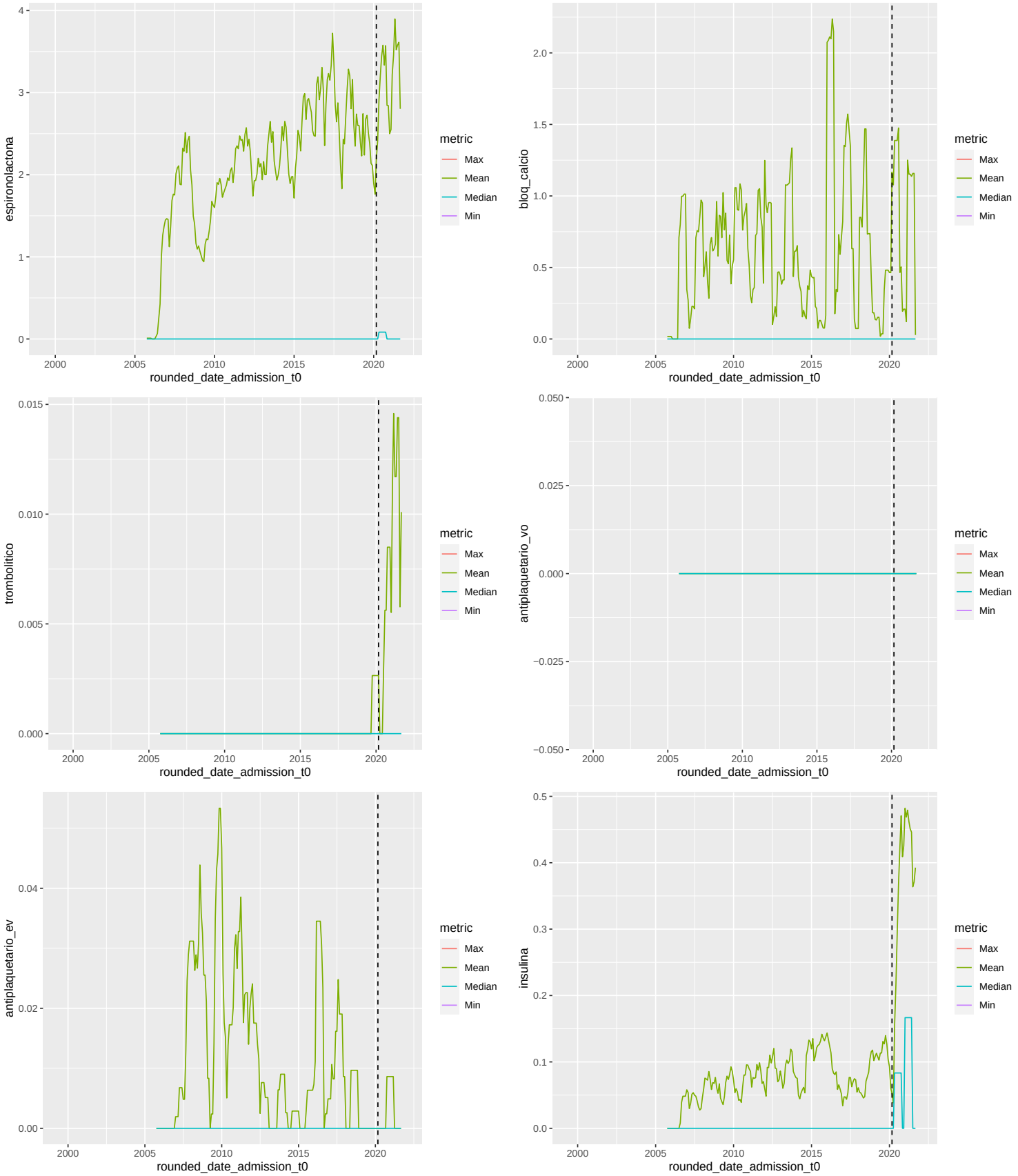

antiarritmico
rounded_date_admission_t0
metric
Max
Mean
Median
Min

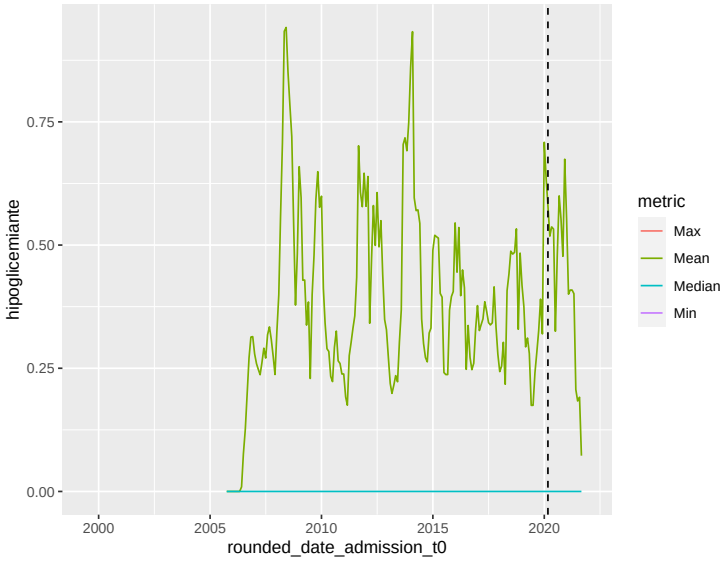

antihipertensivo
rounded_date_admission_t0
metric
Max
Mean
Median
Min

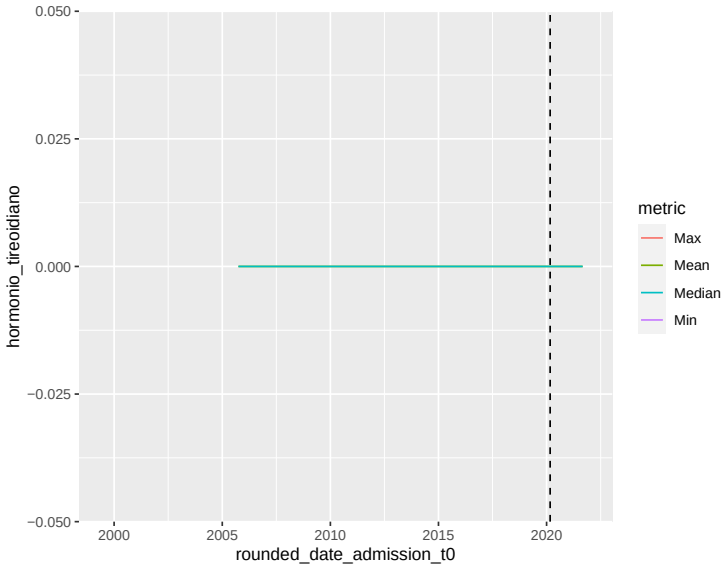

betabloqueador
rounded_date_admission_t0
metric
Max
Mean
Median
Min

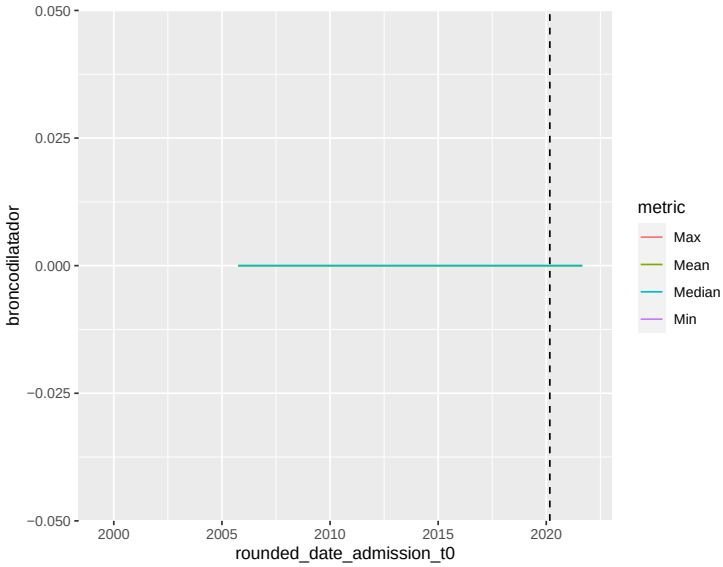

ieca_bra
rounded_date_admission_t0
metric
Max
Mean
Median
Min

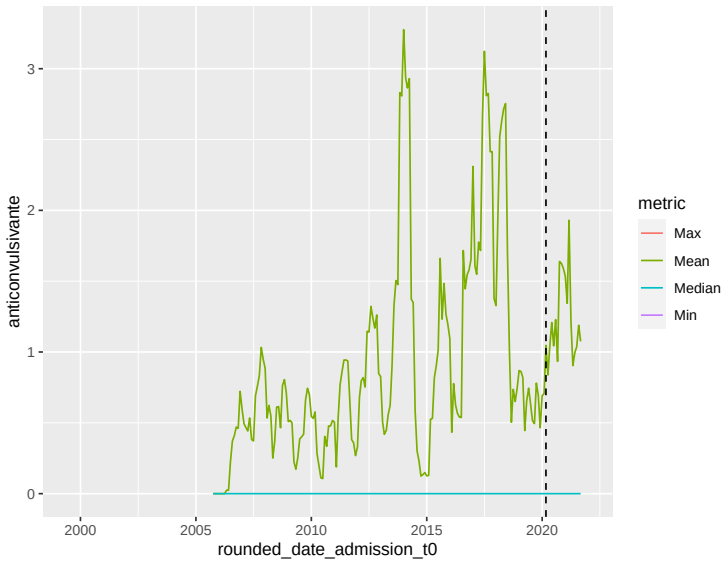
dva
digoxina
estatina
diuretico
vasodilatador
insuf_cardiaca
rounded_date_admission_t0
metric
Max
Mean
Median
Min

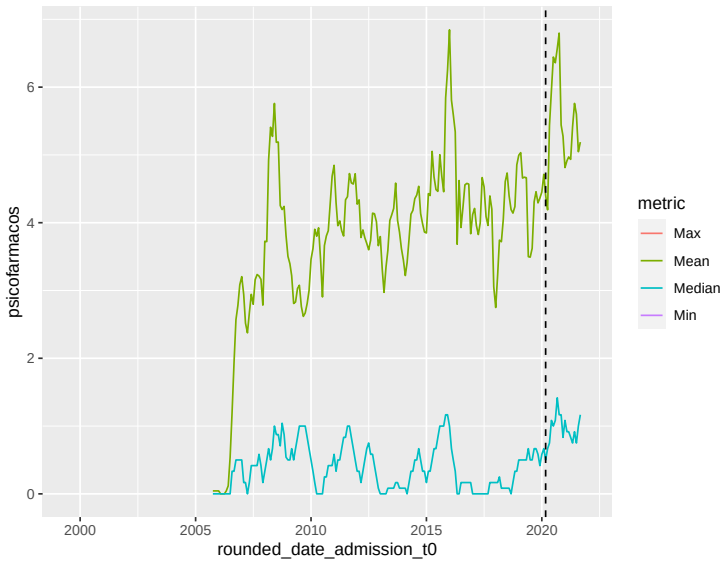

espironolactona
bloq_calcio
trombolitico
antiplaquetario_vo
antiplaquetario_ev
insulina
rounded_date_admission_t0
metric
Max
Mean
Median
Min

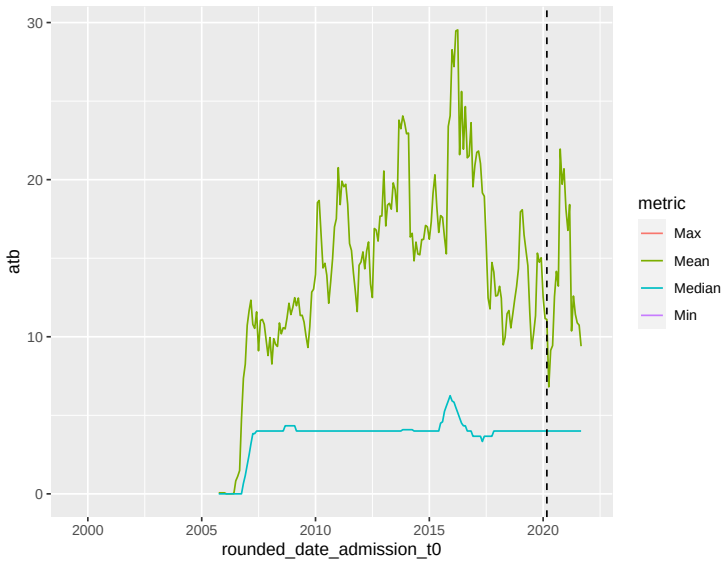

hipoglicemiante
0.75
0.50
0.25
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

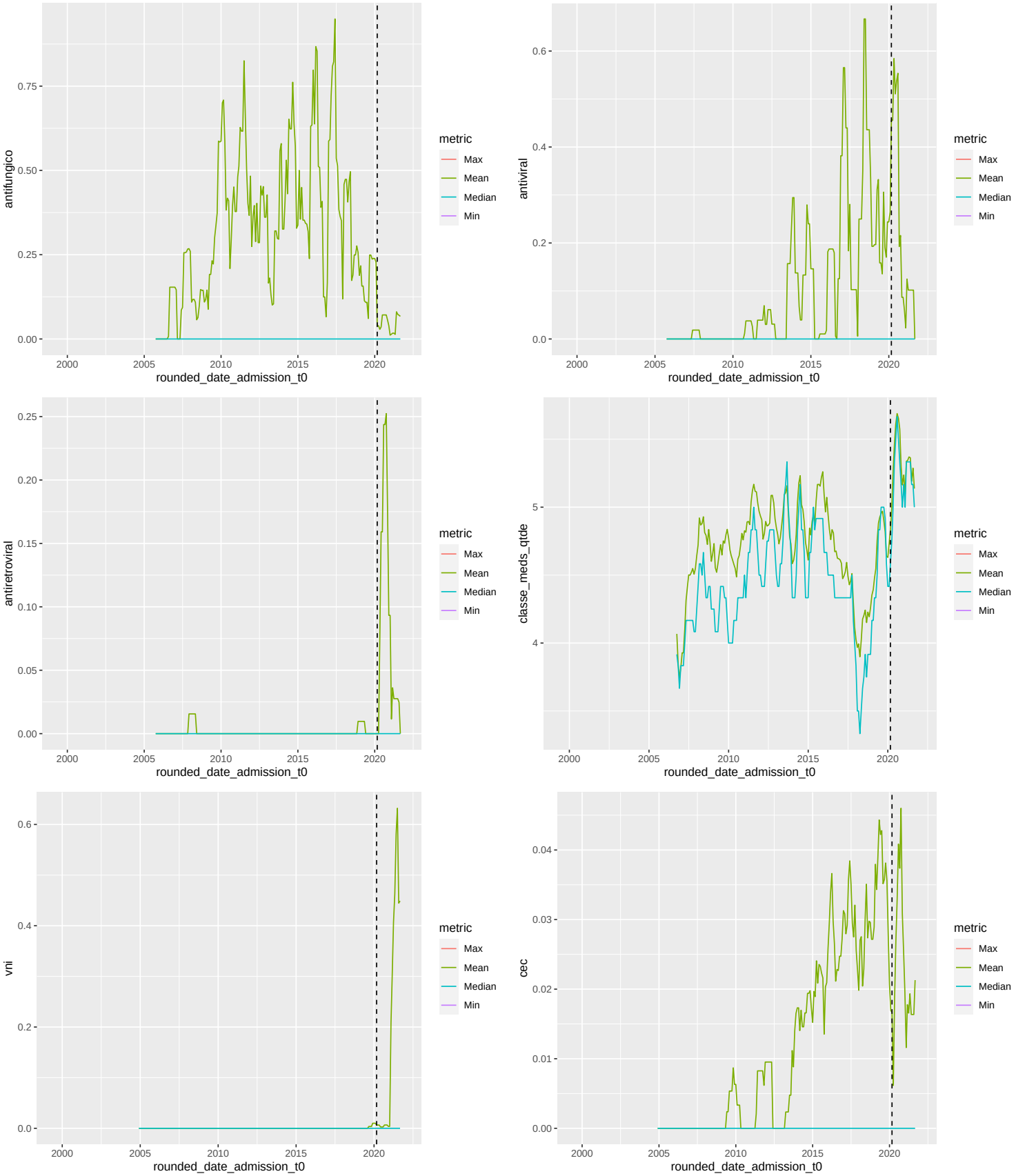

hormonio_tireoidiano
0.050
0.025
0.000
−0.025
−0.050
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

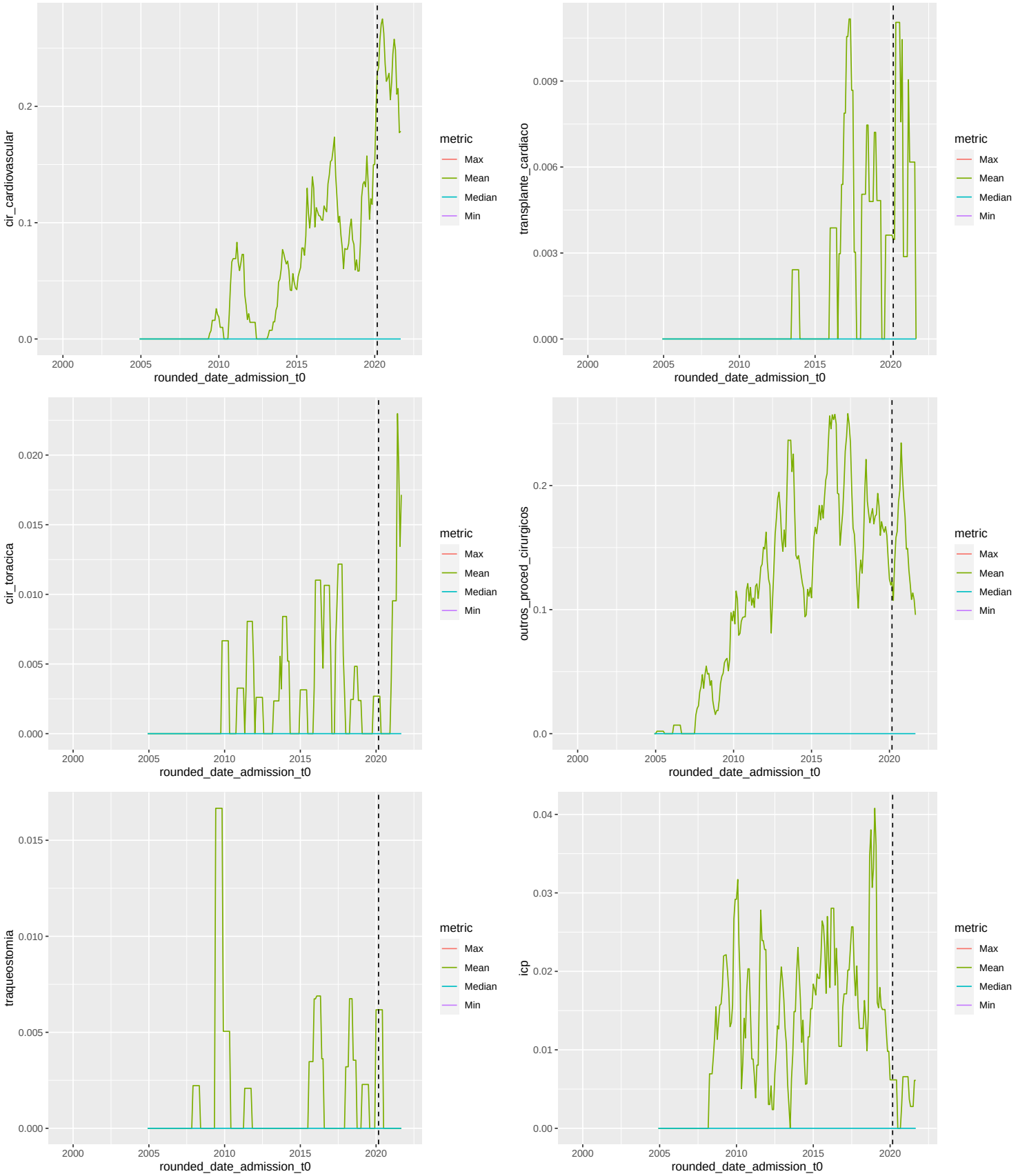

broncodilatador
0.050
0.025
0.000
−0.025
−0.050
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

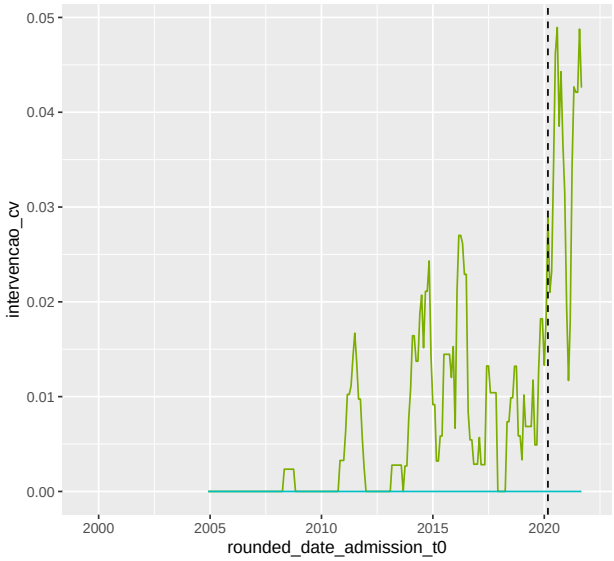

anticonvulsivante
3
2
1
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

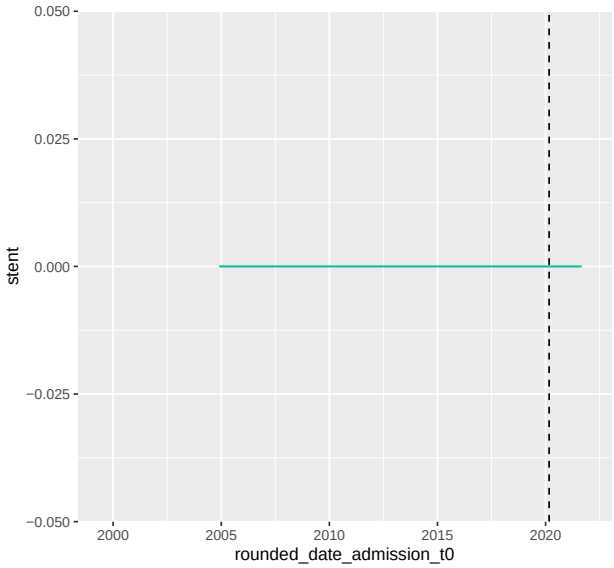

psicofarmacos
6
4
2
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

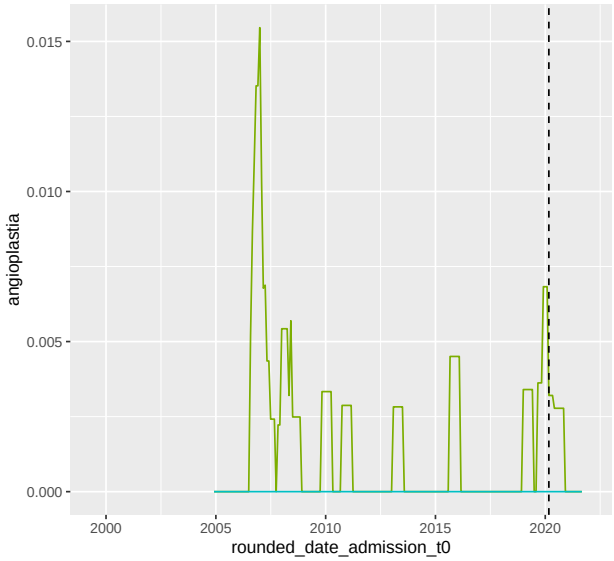

atb
30
20
10
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

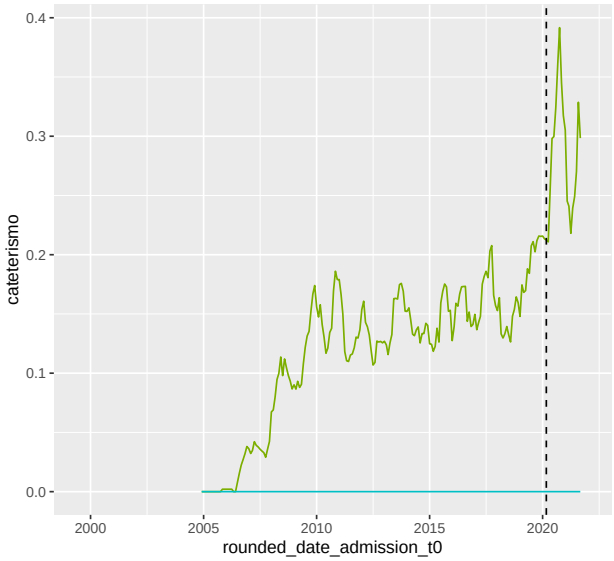

antifungico
0.00
0.25
0.50
0.75
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
antiviral
0.0
0.2
0.4
0.6
antiretroviral
0.00
0.05
0.10
0.15
0.20
0.25
classe_meds_qtde
4
5
vni
cec
0.00
0.01
0.02
0.03
0.04

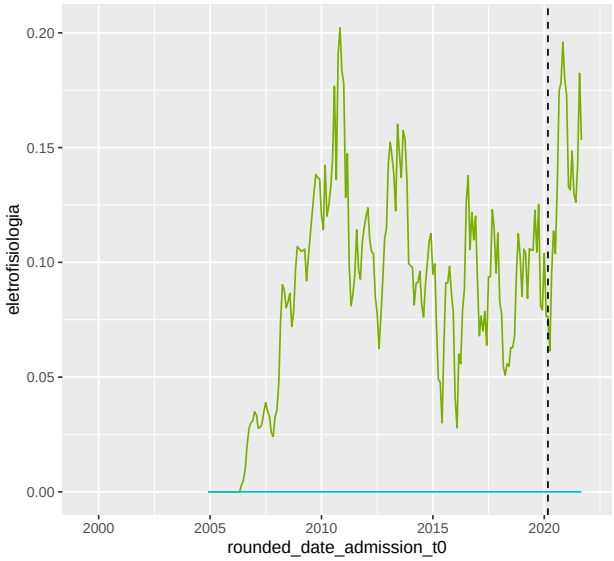

cir_cardiovascular
0.0
0.1
0.2
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
transplante_cardiaco
0.000
0.003
0.006
0.009
cir_toracica
0.000
0.005
0.010
0.015
0.020
outros_proced_cirurgicos
0.0
0.1
0.2
traqueostomia
0.000
0.005
0.010
0.015
icp
0.00
0.01
0.02
0.03
0.04

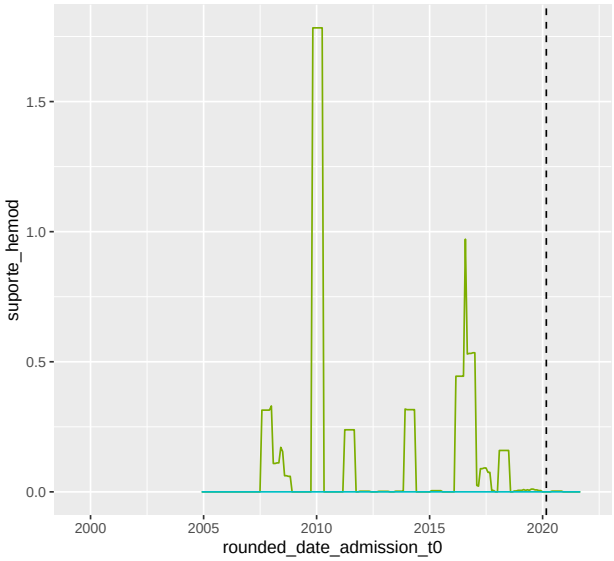

intervencao_cv
0.05
0.04
0.03
0.02
0.01
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

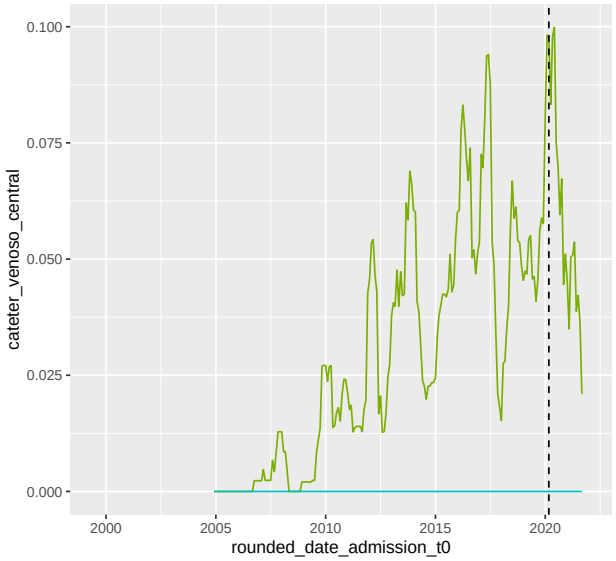

stent
0.050
0.025
0.000
−0.025
−0.050
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

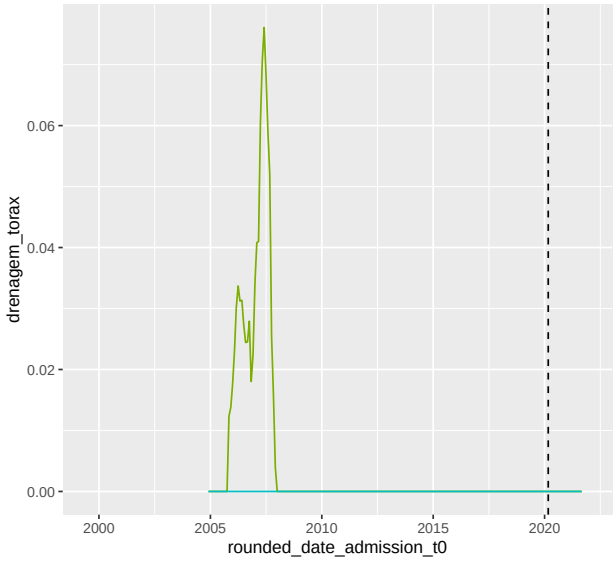

angioplastia
0.015
0.010
0.005
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

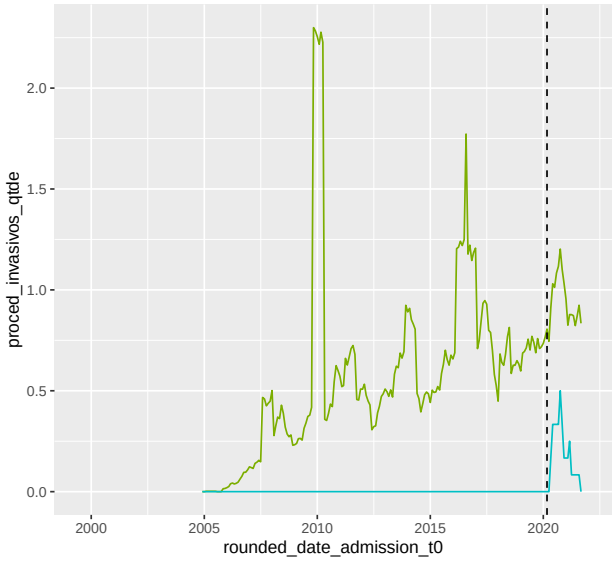

cateterismo
0.4
0.3
0.2
0.1
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

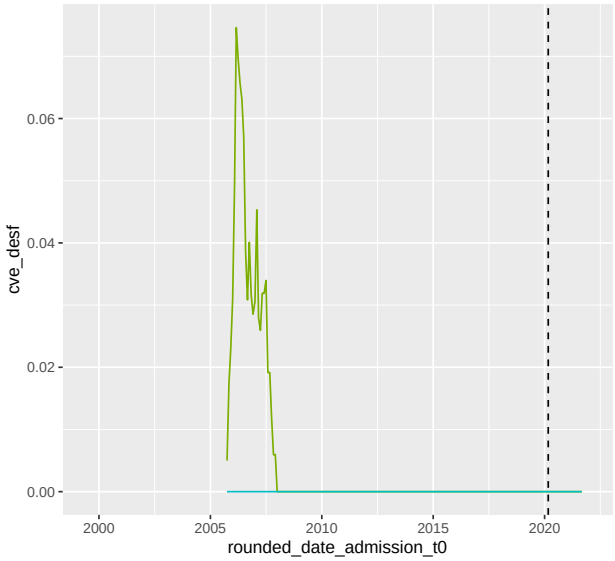

eletrofisiologia
0.20
0.15
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

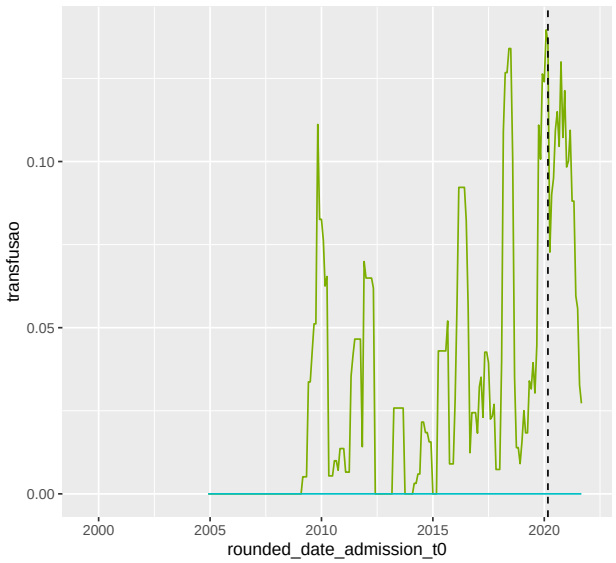

suporte_hemod
1.5
1.0
0.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

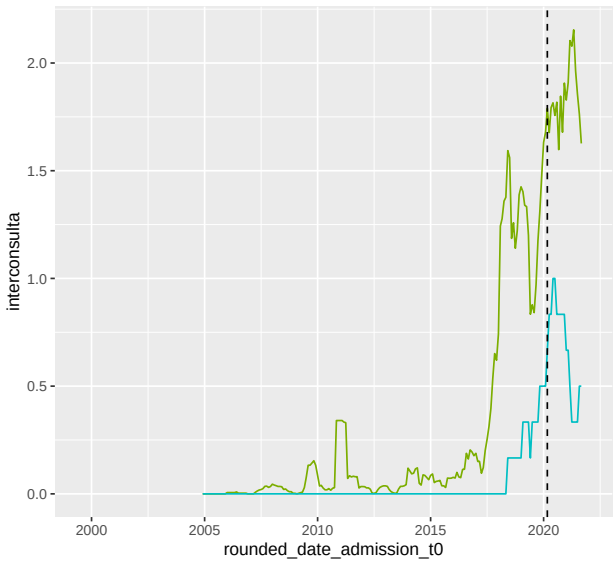

cateter_venoso_central
0.100
0.075
0.050
0.025
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

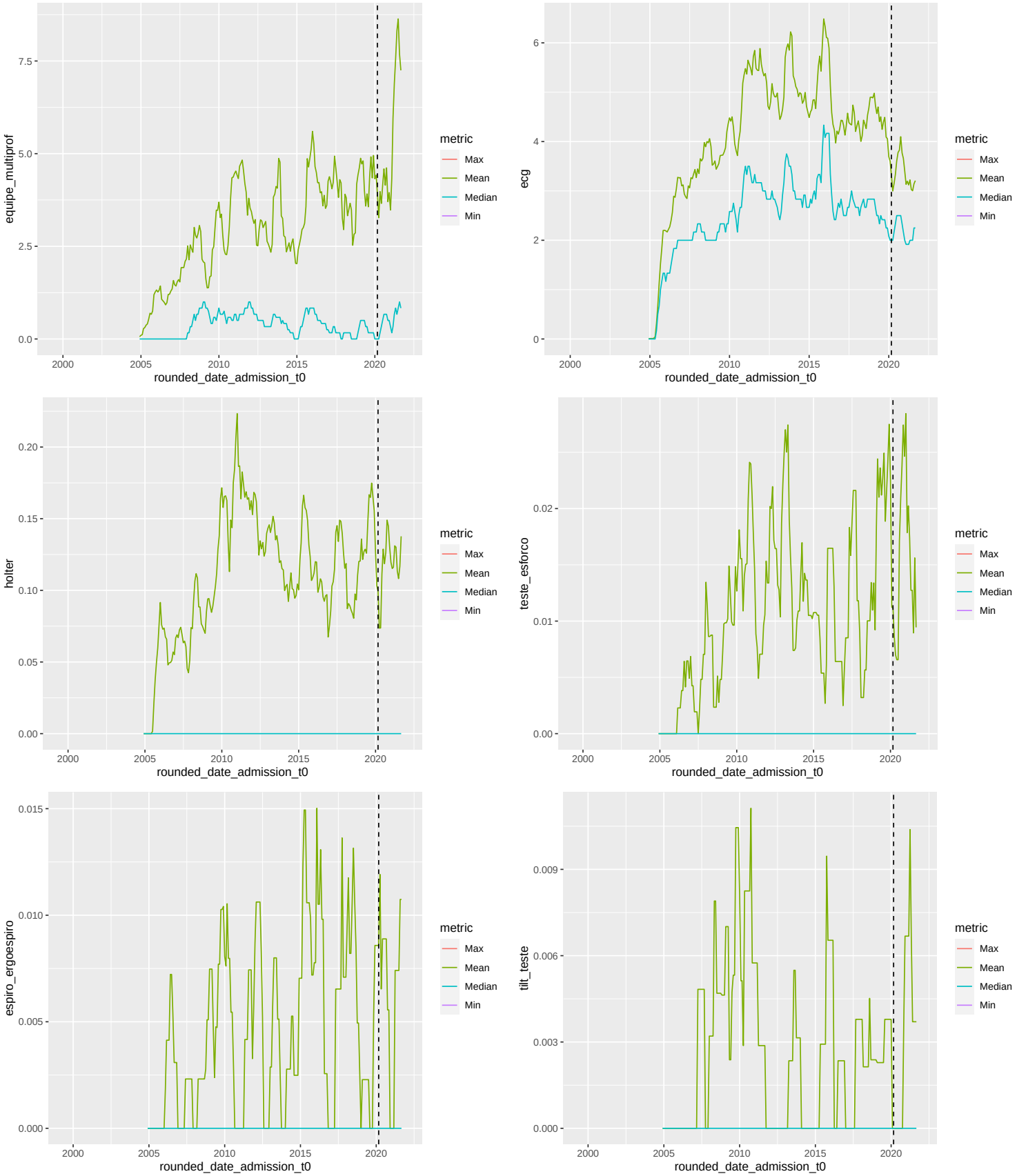

drenagem_torax
0.06
0.04
0.02
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

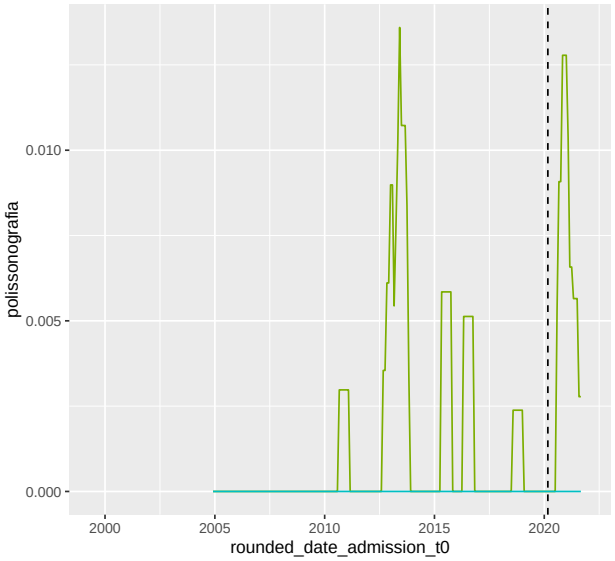

proced_invasivos_qtde
2.0
1.5
1.0
0.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

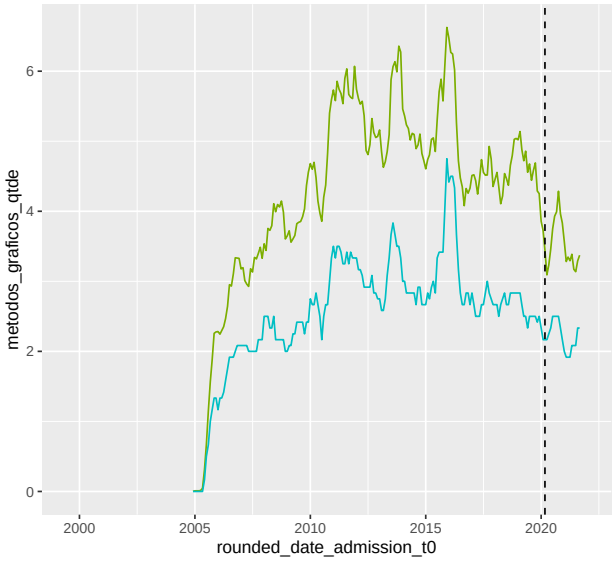

cve_desf
0.06
0.04
0.02
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

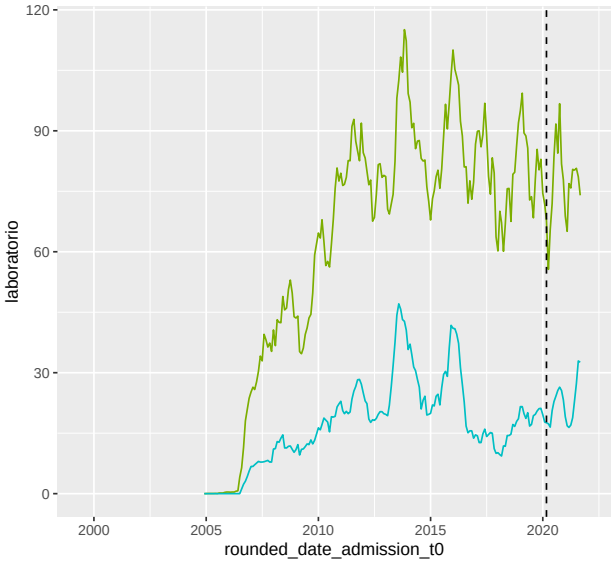

transfusao
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

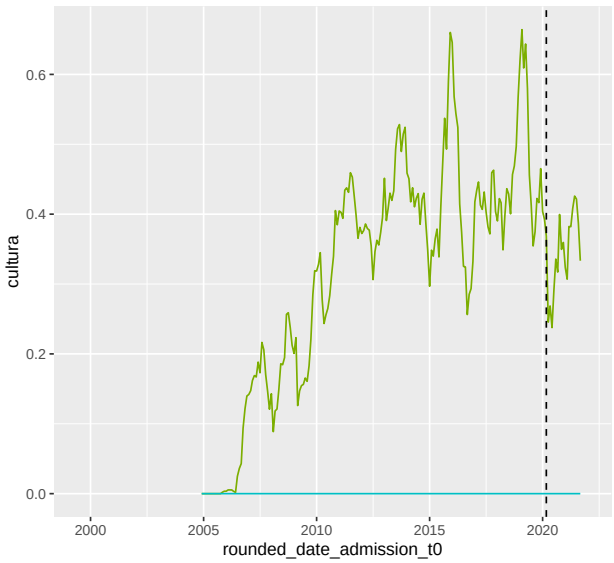

interconsulta
2.0
1.5
1.0
0.5
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

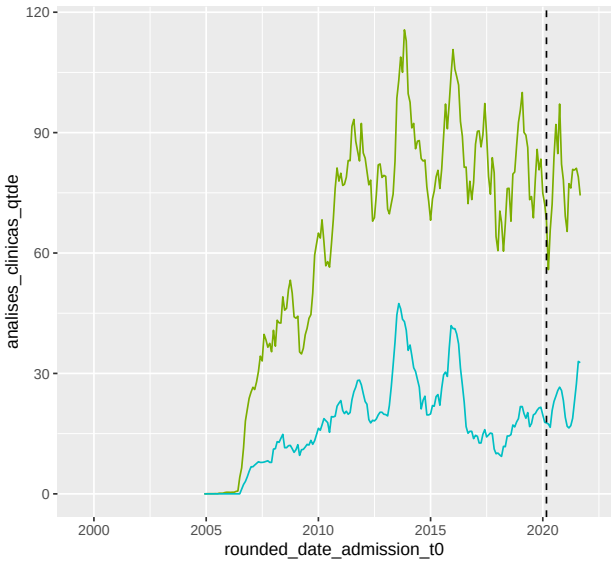

equipe_multiprof
rounded_date_admission_t0
ecg
holter
teste_esforco
espiro_ergoespiro
tilt_teste
metric
Max
Mean
Median
Min

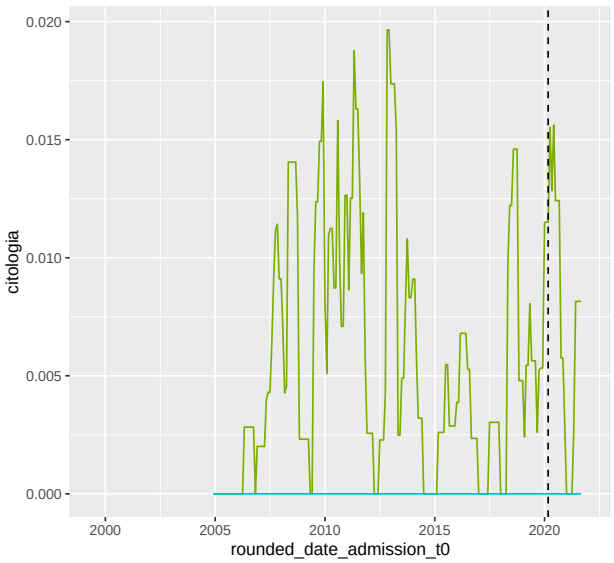

polissonografia
0.010
0.005
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

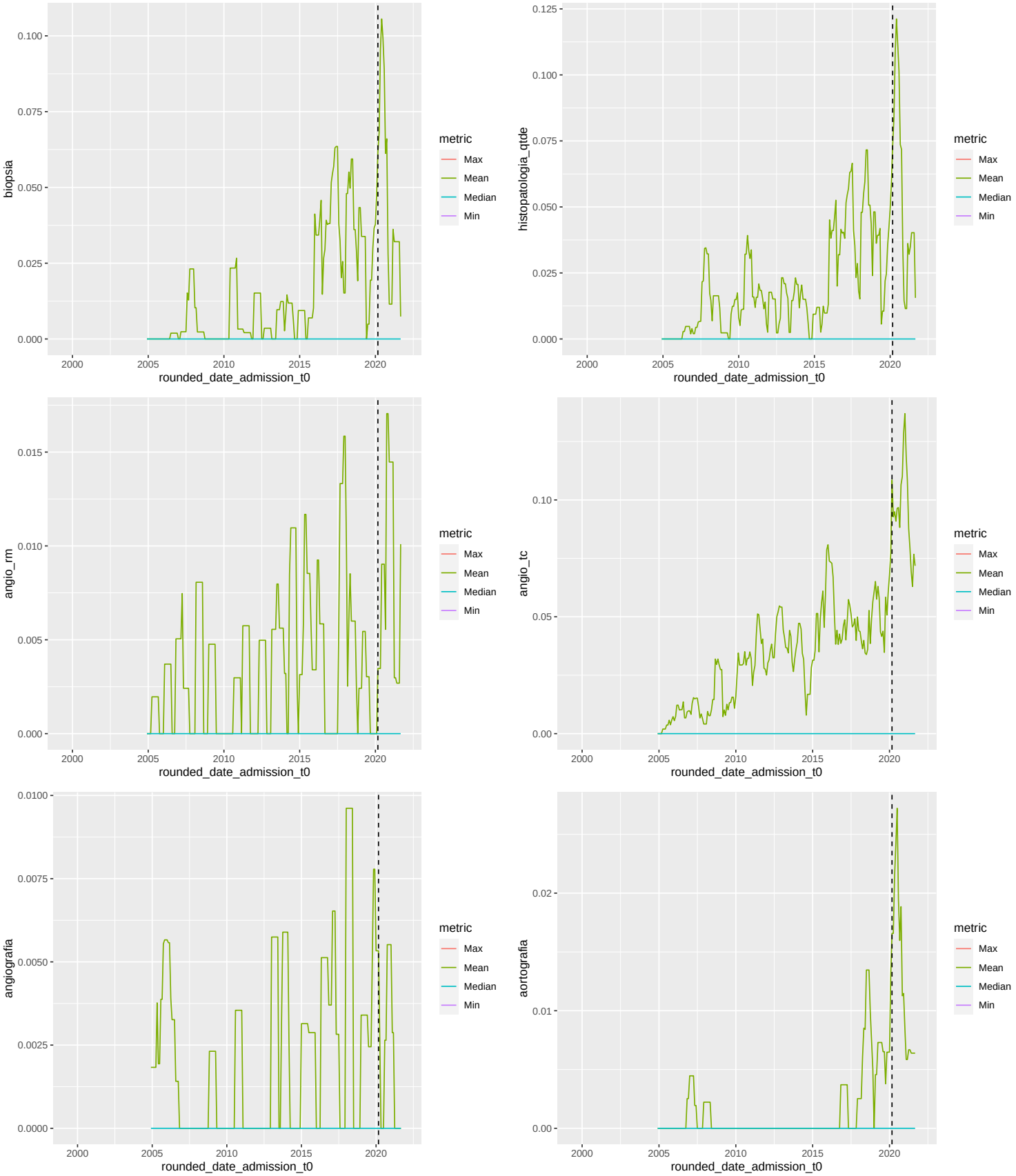

metodos_graficos_qtde
6
4
2
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

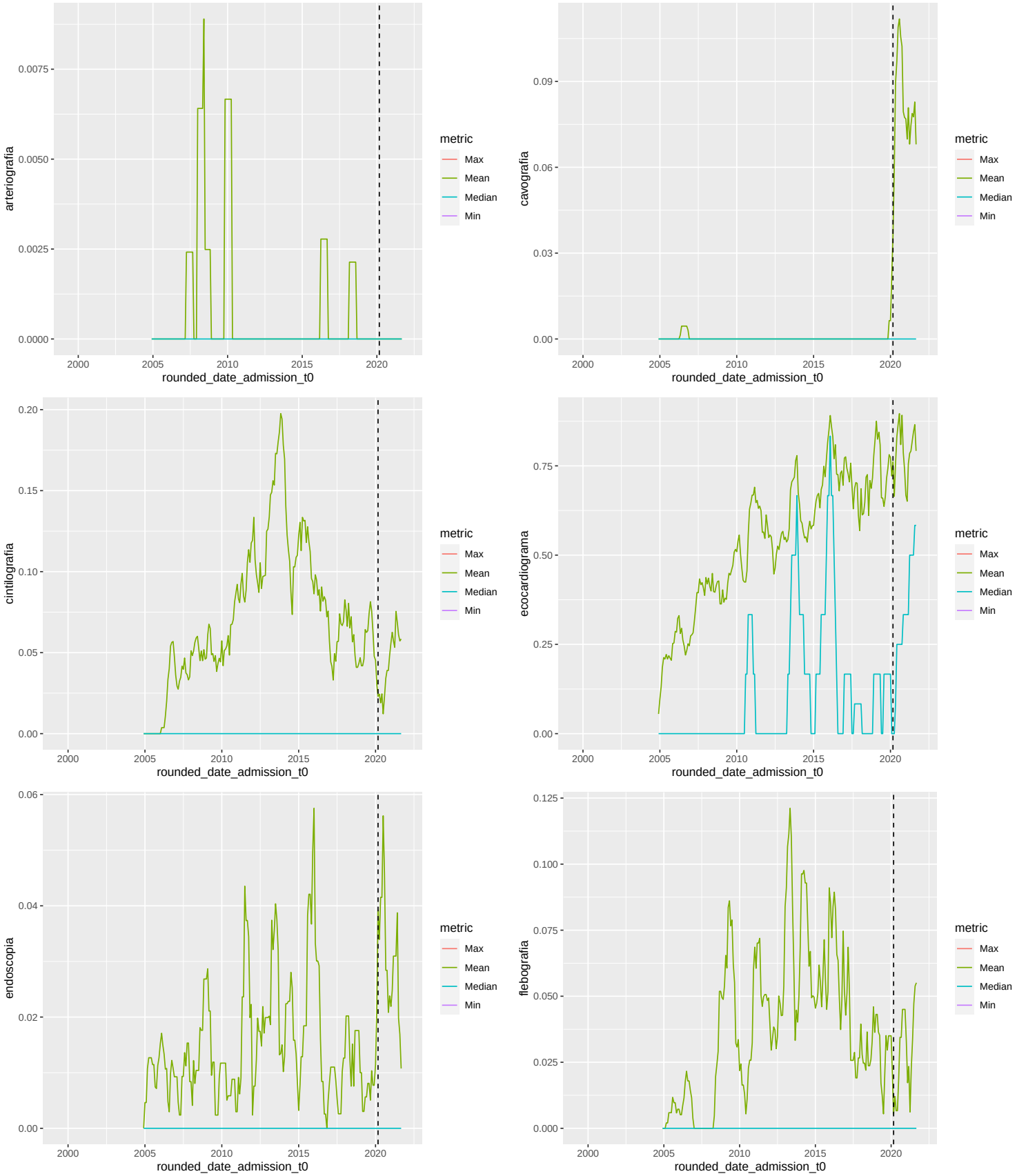

laboratorio
120
90
60
30
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

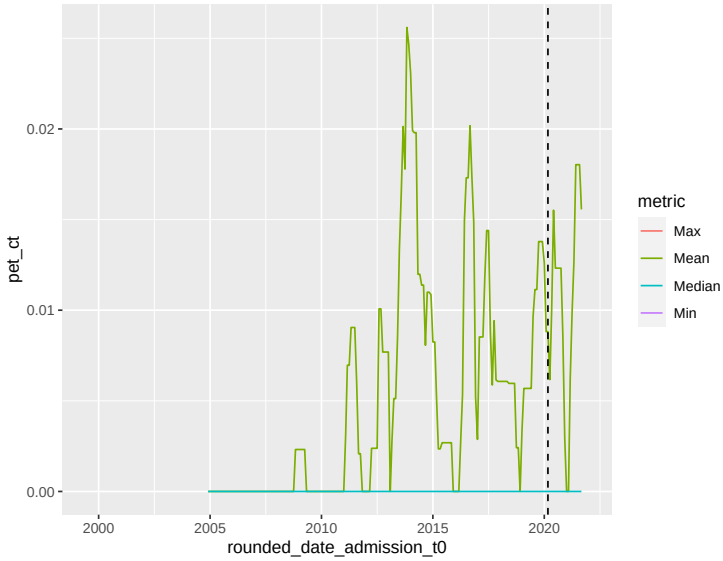

cultura
0.6
0.4
0.2
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

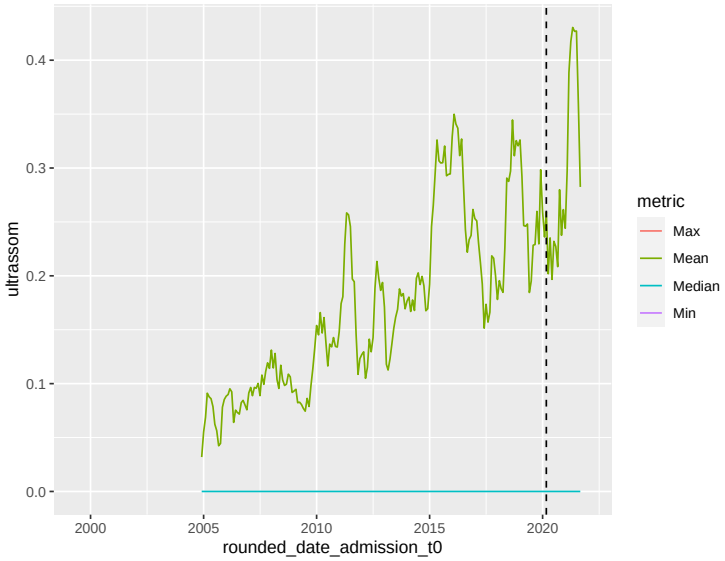

analises_clinicas_qtde
120
90
60
30
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

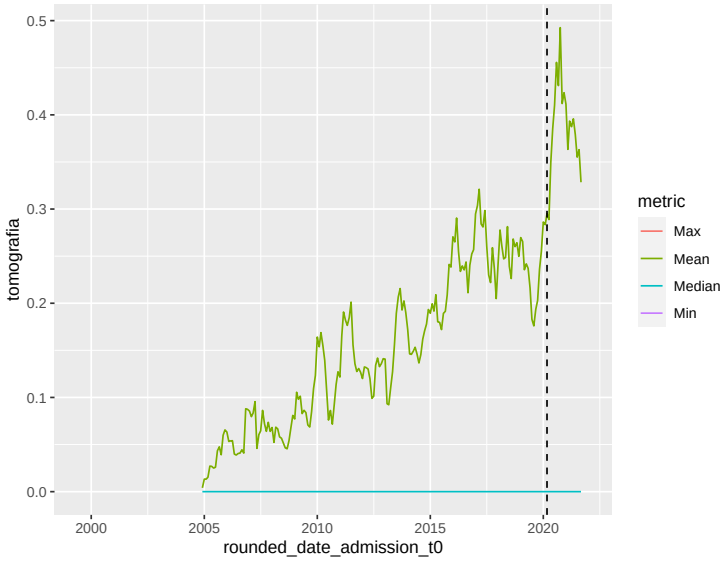

citologia
0.020
0.015
0.010
0.005
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

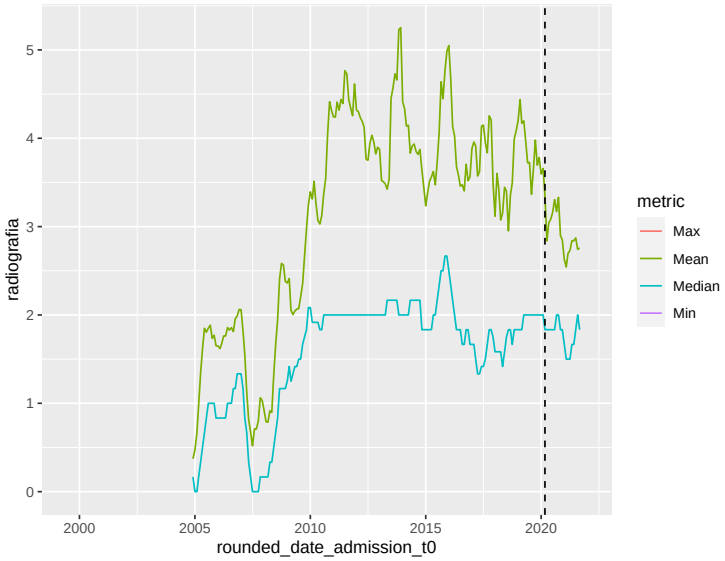

biopsia
0.100
0.075
0.050
0.025
0.000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
histopatologia_qtde
0.125
0.100
0.075
0.050
0.025
0.000
angio_rm
0.015
0.010
0.005
0.000
angio_tc
0.10
0.05
0.00
angiografia
0.0100
0.0075
0.0050
0.0025
0.0000
aortografia
0.02
0.01
0.00

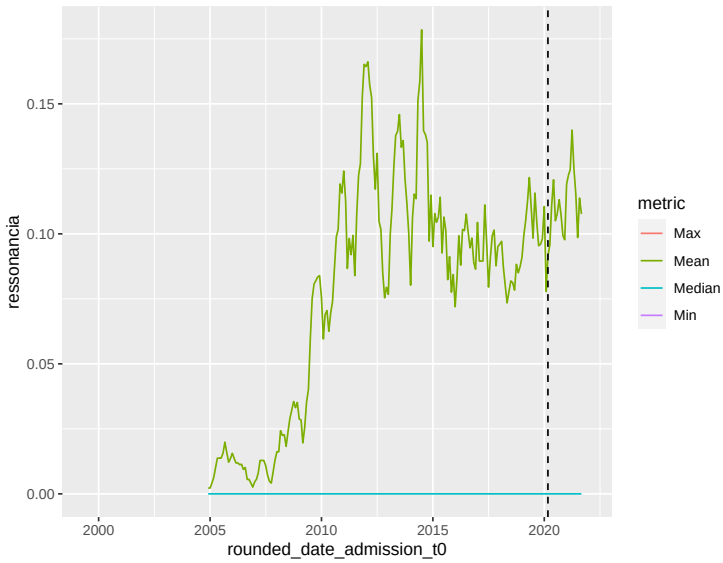

arteriografia
0.0075
0.0050
0.0025
0.0000
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
cavografia
0.09
0.06
0.03
0.00
cintilografia
0.20
0.15
0.10
0.05
0.00
ecocardiograma
0.75
0.50
0.25
endoscopia
0.06
0.04
0.02
flebografia
0.125
0.100
0.075
0.050
0.025
0.000

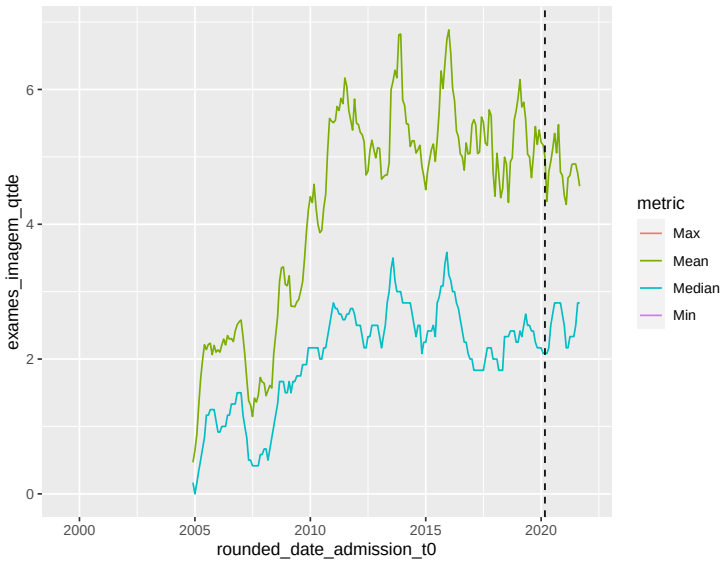

pet_ct
0.02
0.01
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

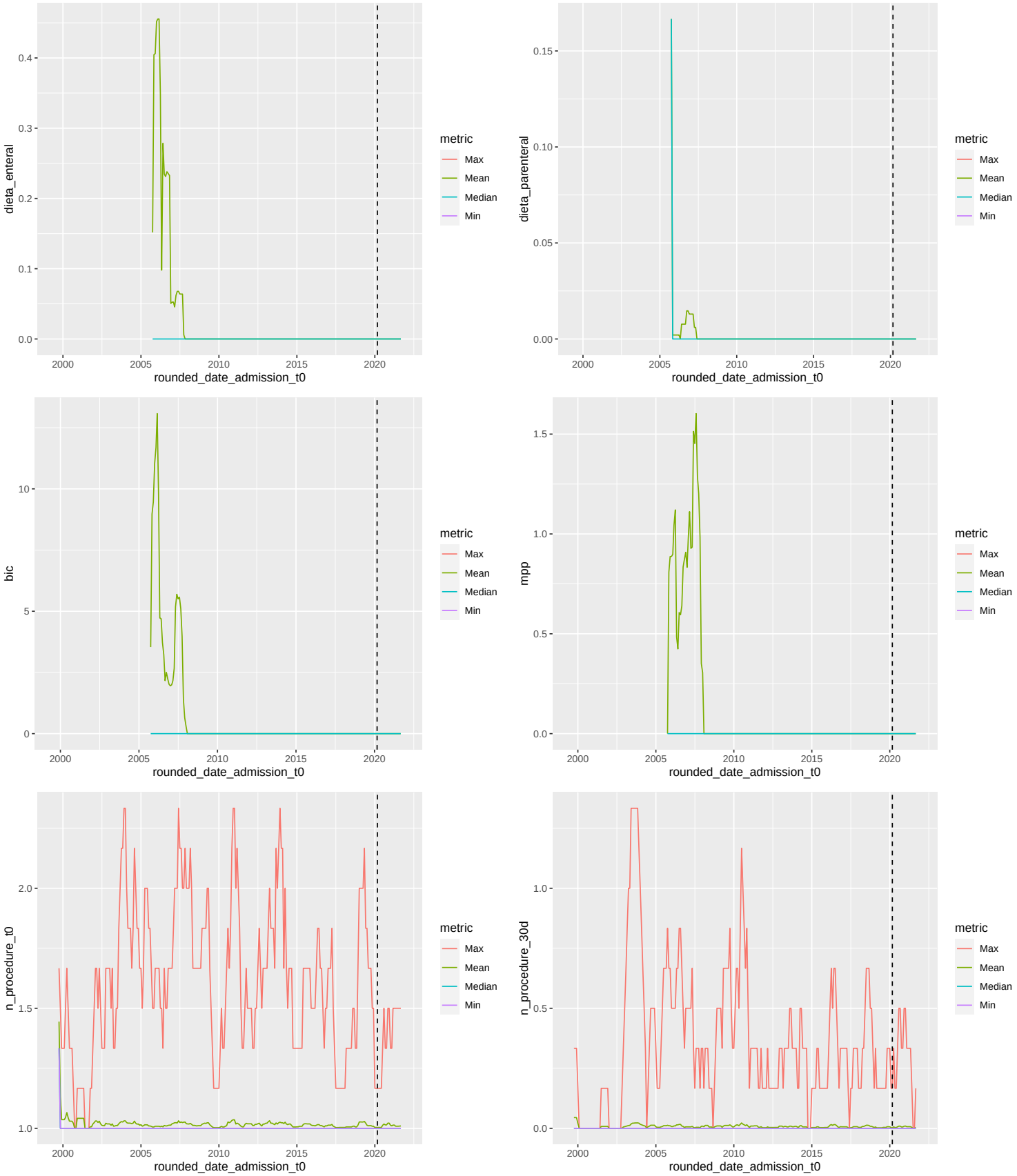

ultrassom
0.4
0.3
0.2
0.1
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

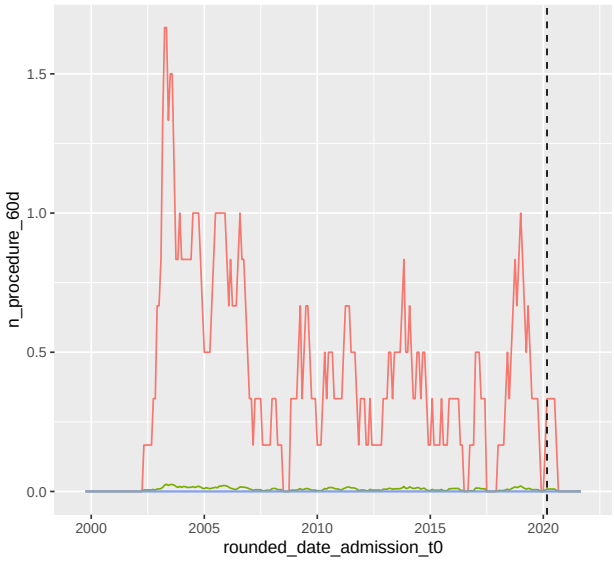

tomografia
0.5
0.4
0.3
0.2
0.1
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

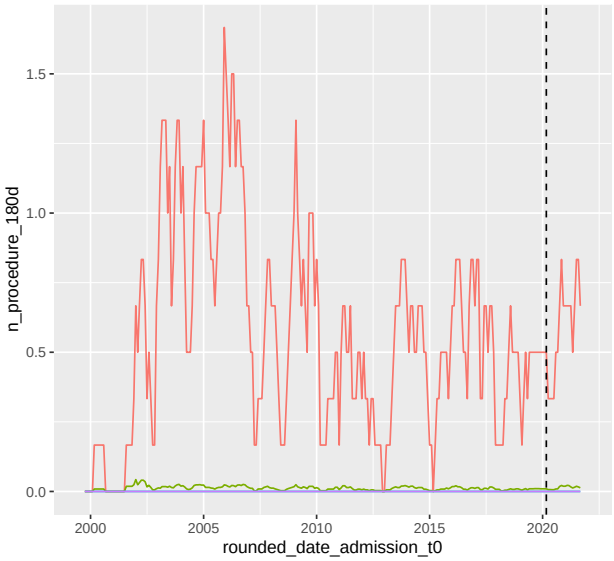

radiografia
5
4
3
2
1
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

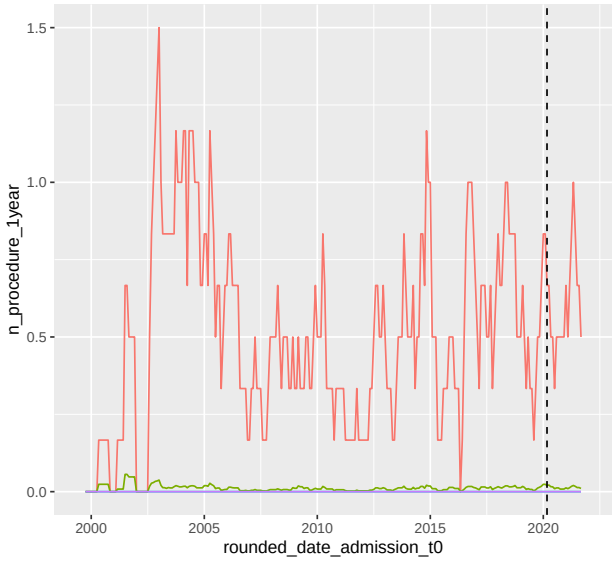

ressonancia
0.15
0.10
0.05
0.00
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

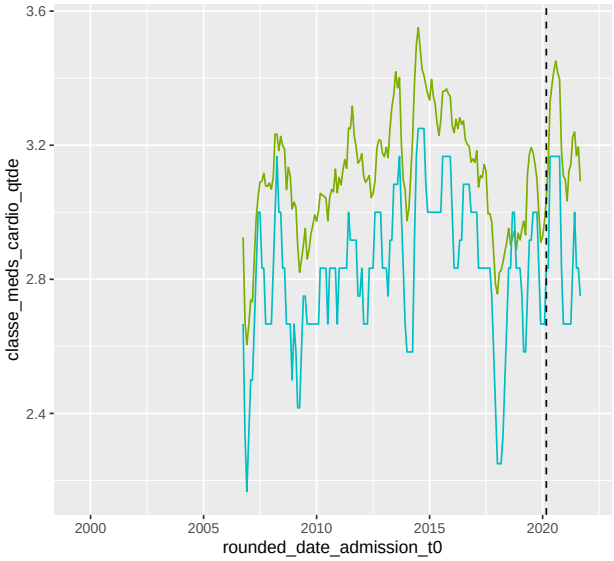

exames_imagem_qtde
6
4
2
0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min

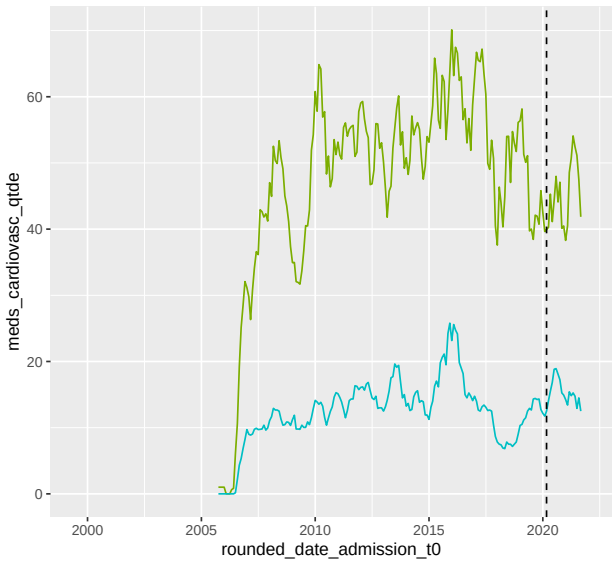

dieta_enteral
0.4
0.3
0.2
0.1
0.0
2000
2005
2010
2015
2020
rounded_date_admission_t0
metric
Max
Mean
Median
Min
dieta_parenteral
0.15
0.10
0.05
0.00
bic
10
5
0
mpp
1.5
1.0
0.5
0.0
n_procedure_t0
2.0
1.5
1.0
n_procedure_30d

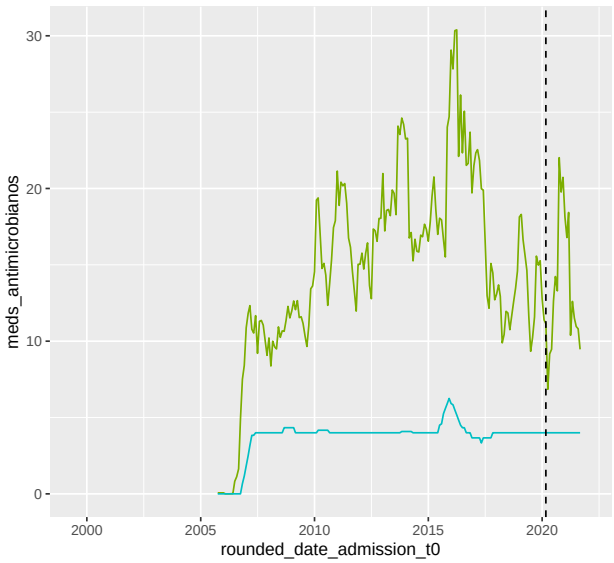








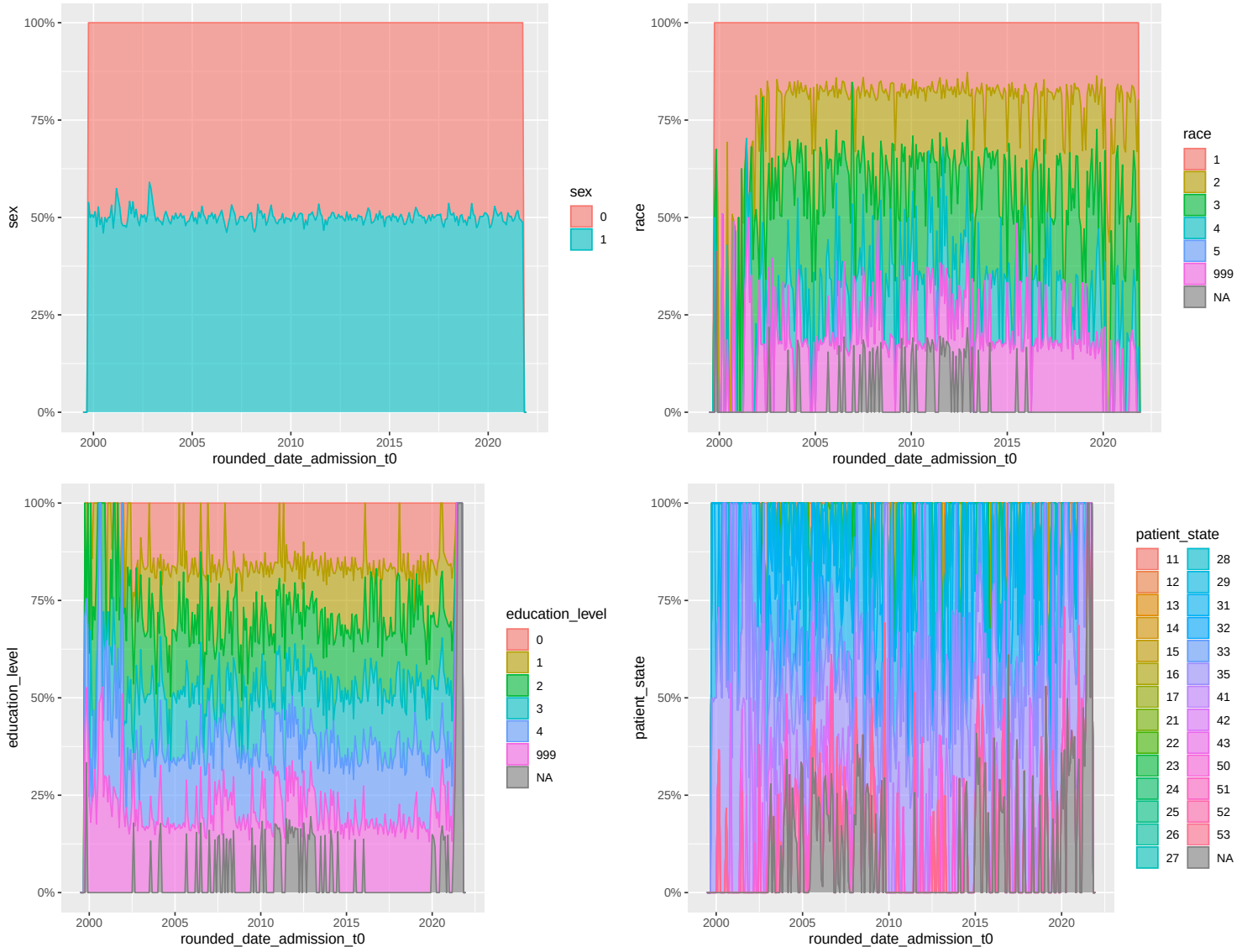


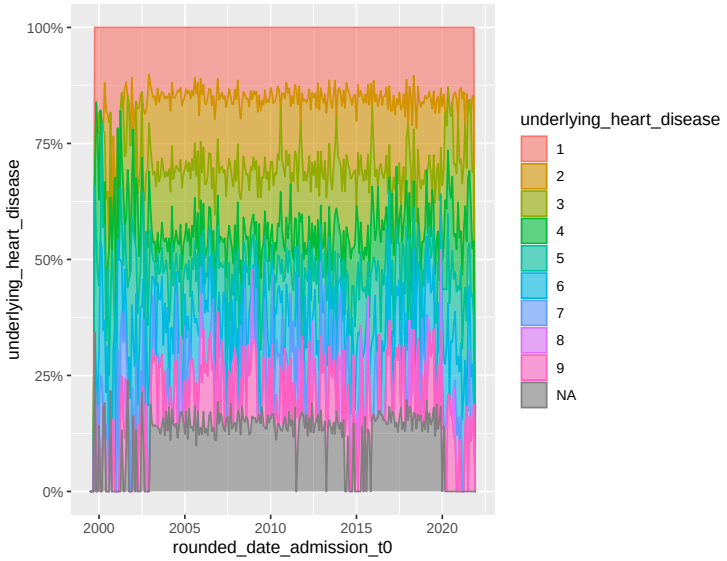


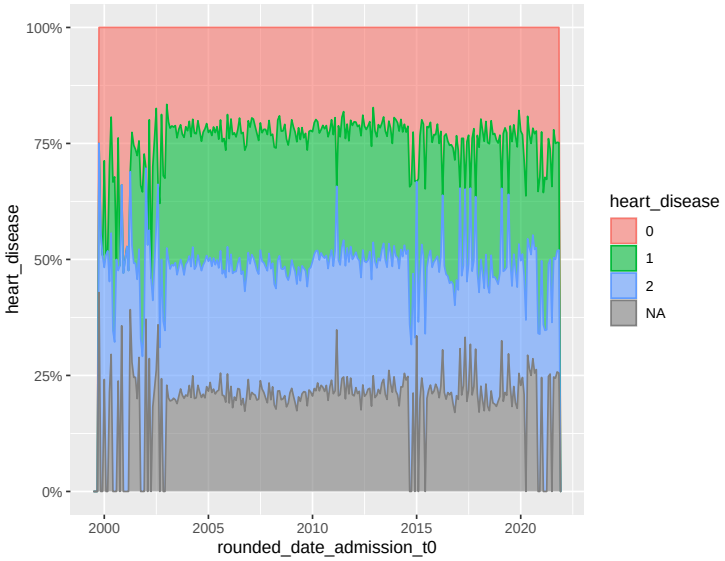

```
column <- 'death_intraop_2'

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)
```

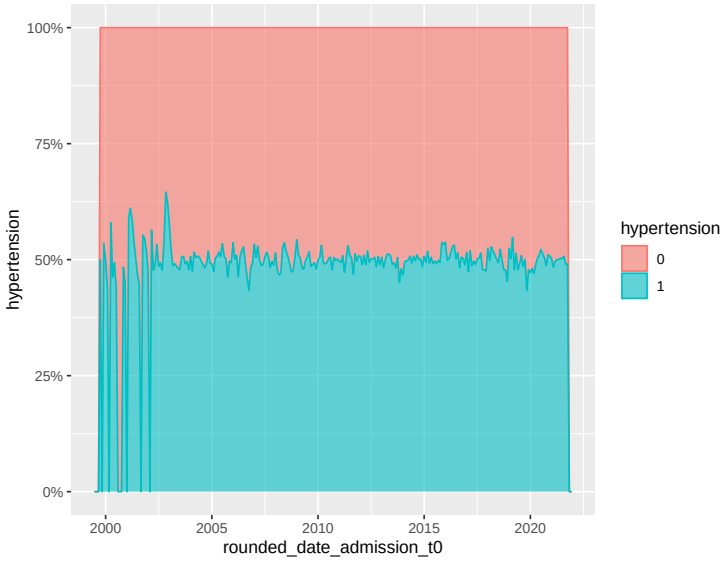
```
  df %>%
    mutate(rounded_date_admission_t0 = lubridate::floor_date(date_admission_t0,
                                                             'month')) %>%
    group_by(rounded_date_admission_t0, !!sym(column)) %>%
    summarise(N = n()) %>%
    ungroup %>%
    arrange(rounded_date_admission_t0) %>%
    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)
}
```

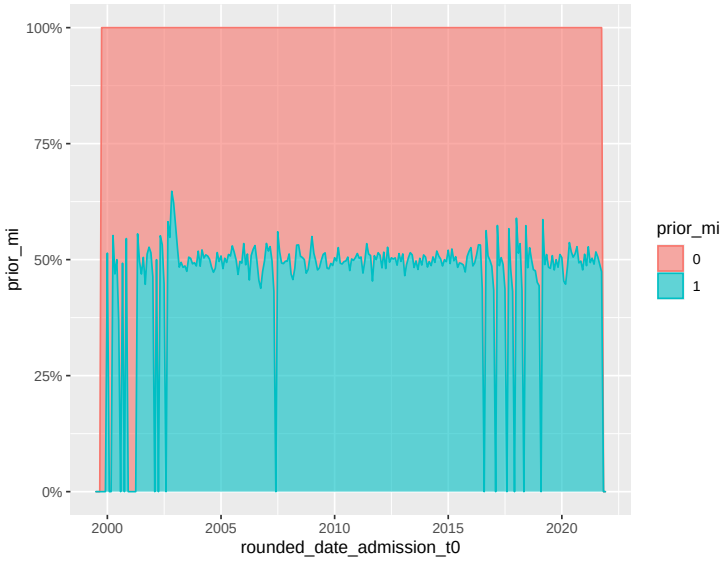
```
## [1] "patient_state"
```

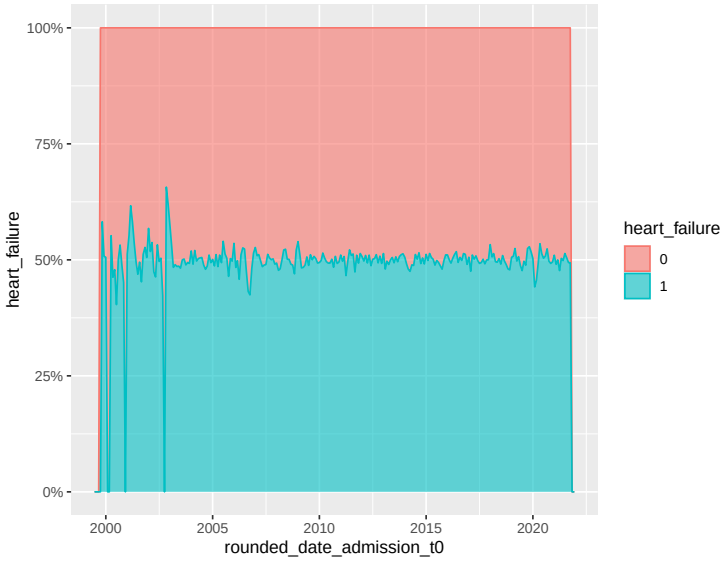


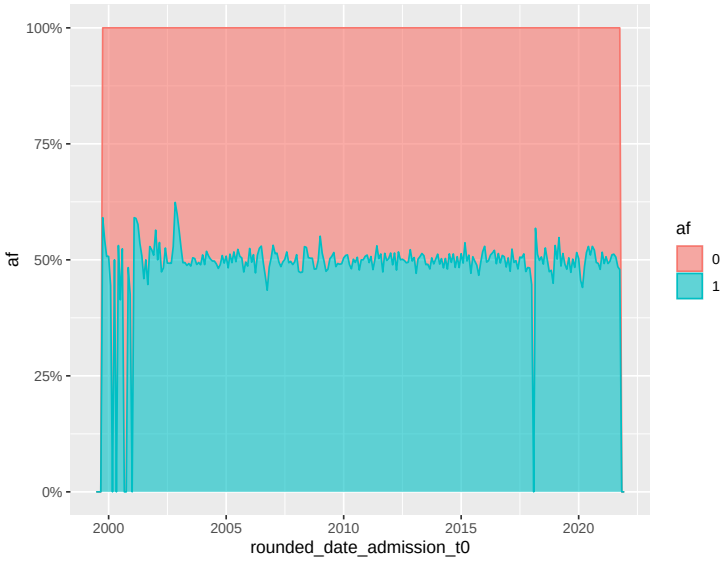

underlying_heart_disease
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
underlying_heart_disease
1
2
3
4
5
6
7
8
9
NA

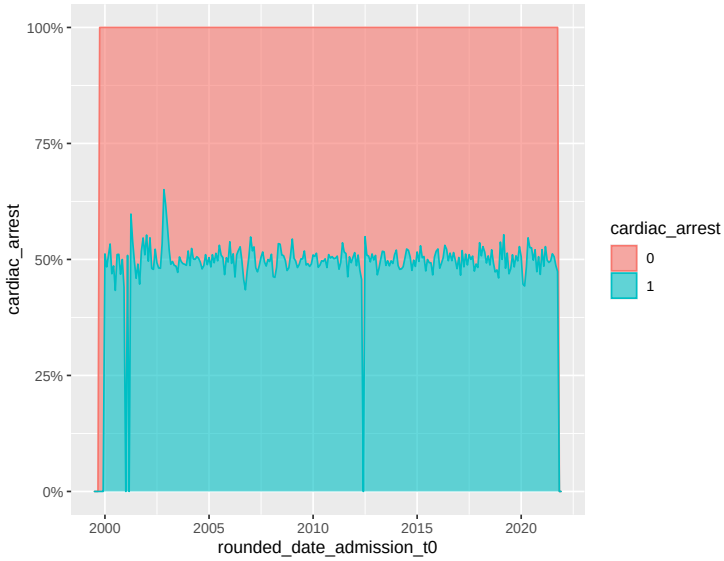

heart_disease
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
heart_disease
0
1
2
NA

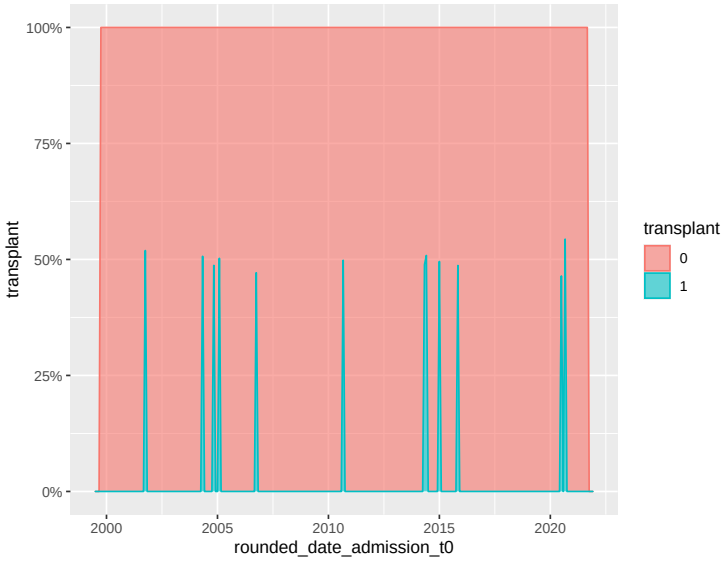

hypertension
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
hypertension
0
1

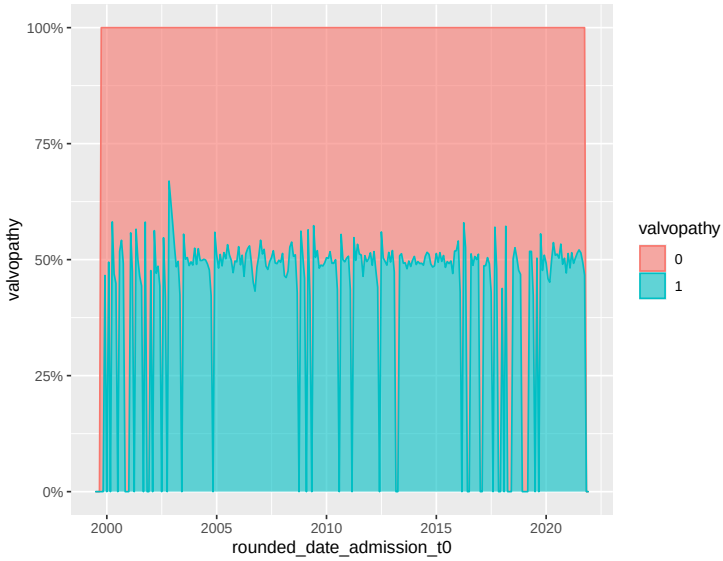

prior_mi
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
prior_mi
0
1

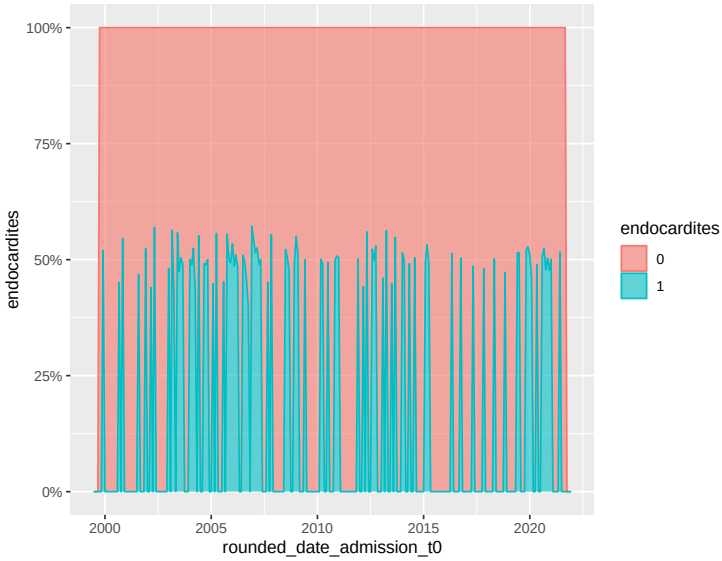

heart_failure
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
heart_failure
0
1

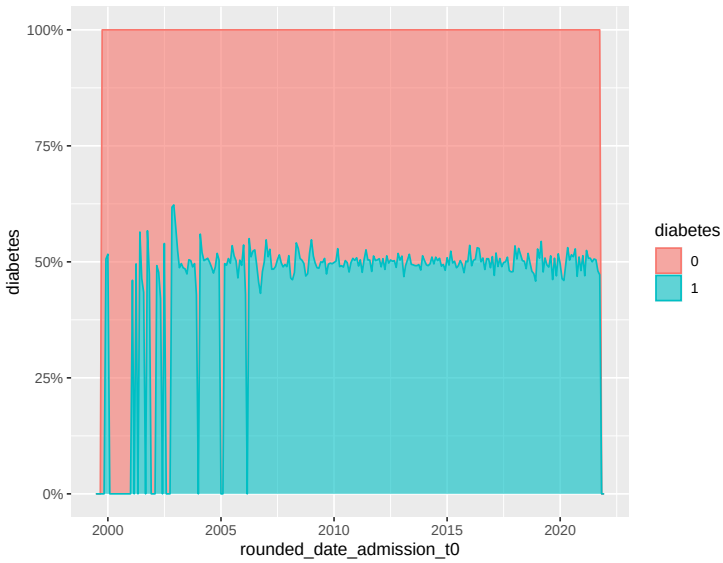

af
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
af
0
1

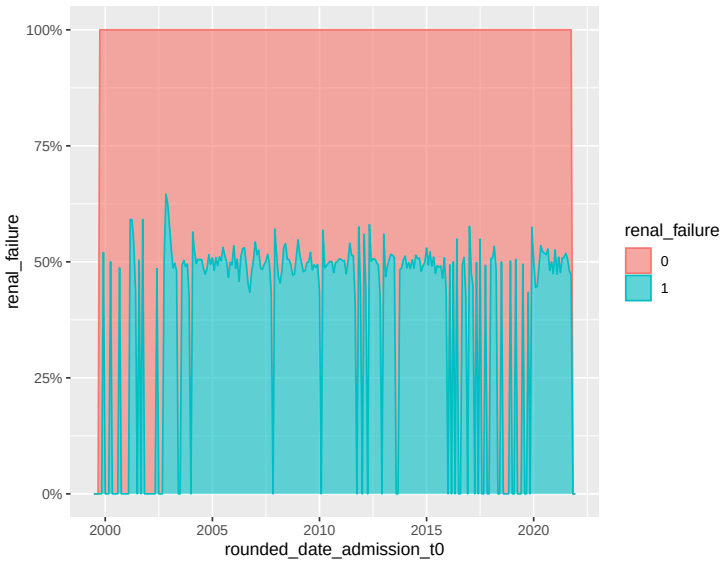

cardiac_arrest
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
cardiac_arrest
0
1

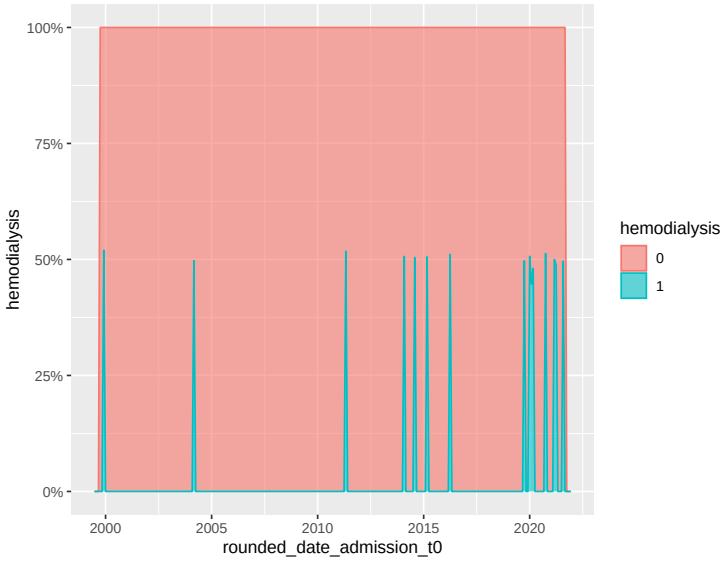

transplant
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
transplant
0
1

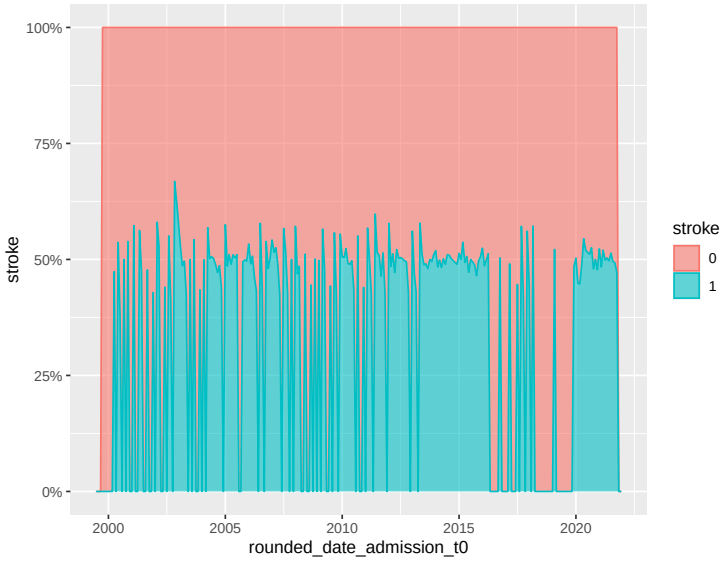

valvopathy
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
valvopathy
0
1

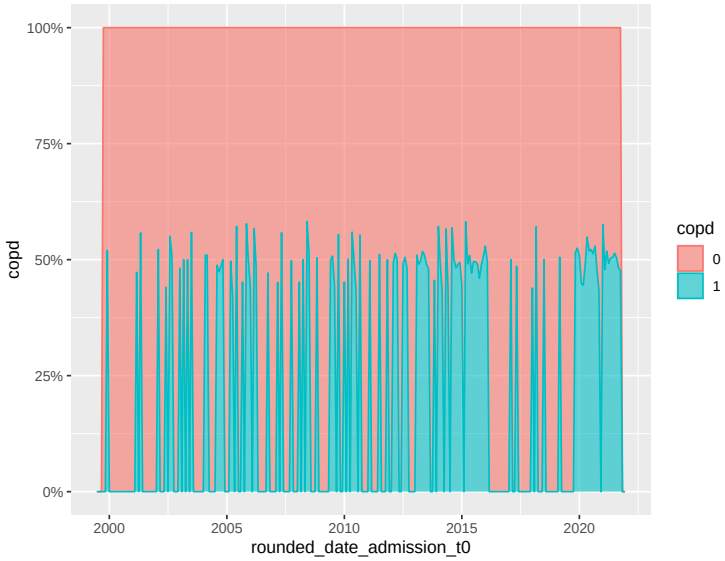

endocardites
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
endocardites
0
1

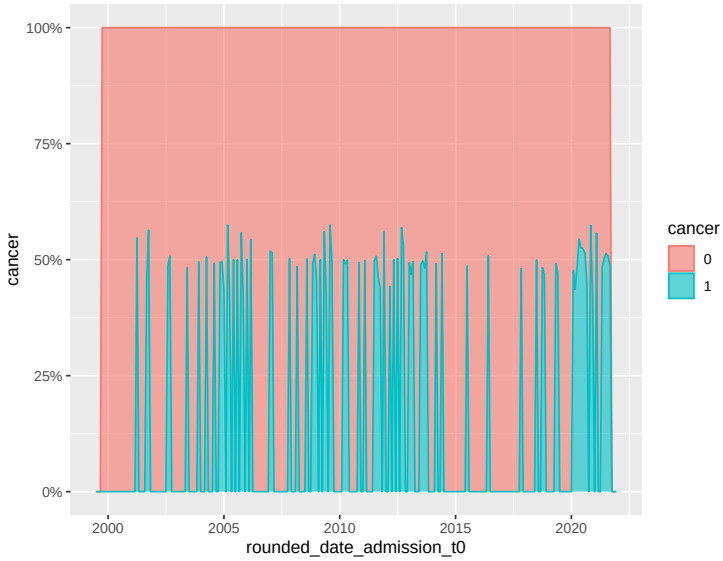

diabetes
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
diabetes
0
1

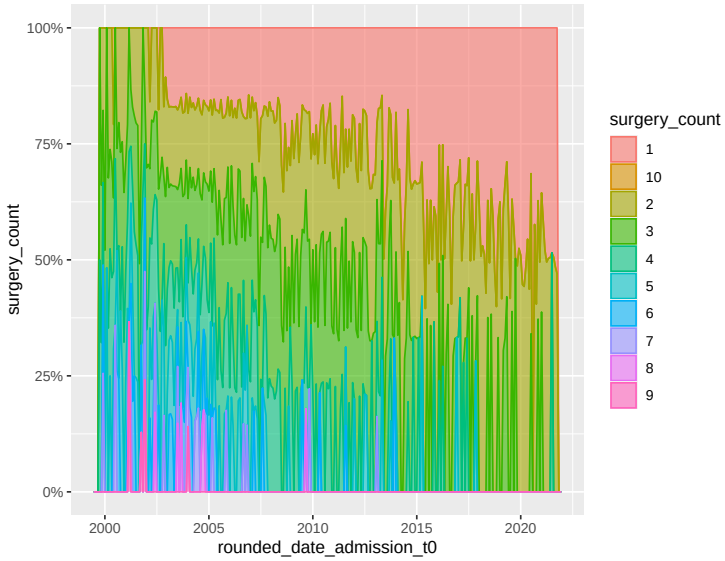

renal_failure
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
renal_failure
0
1

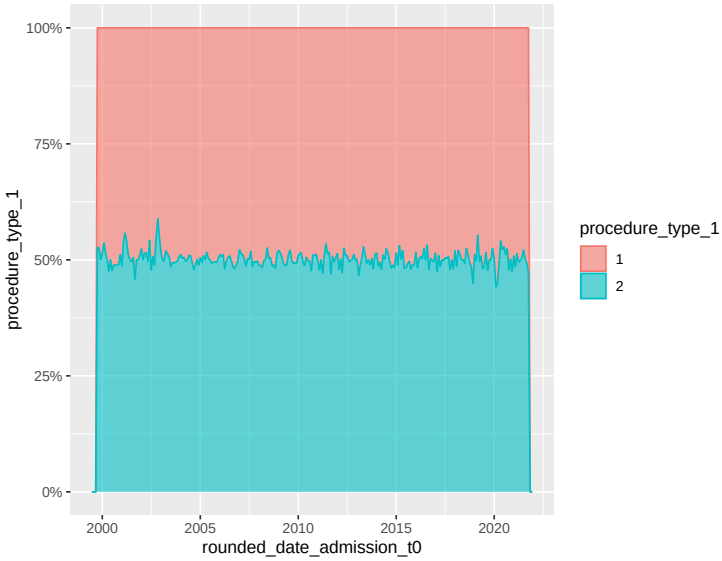

hemodialysis
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
hemodialysis
0
1

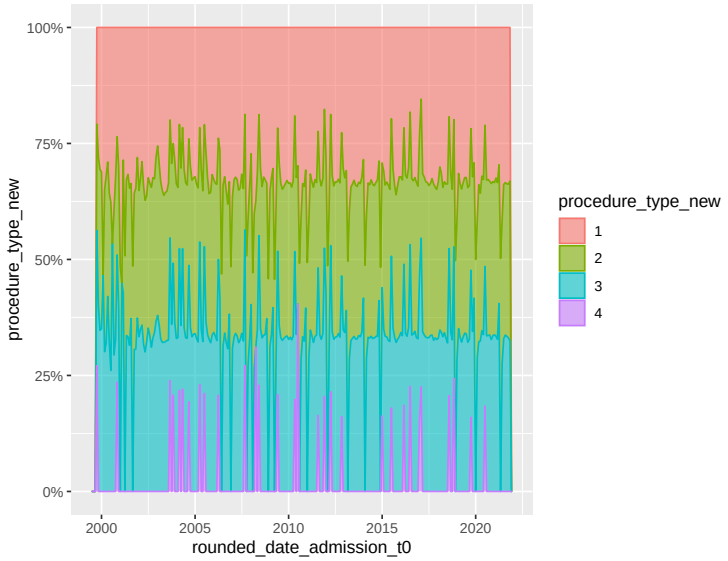

stroke
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
stroke
0
1

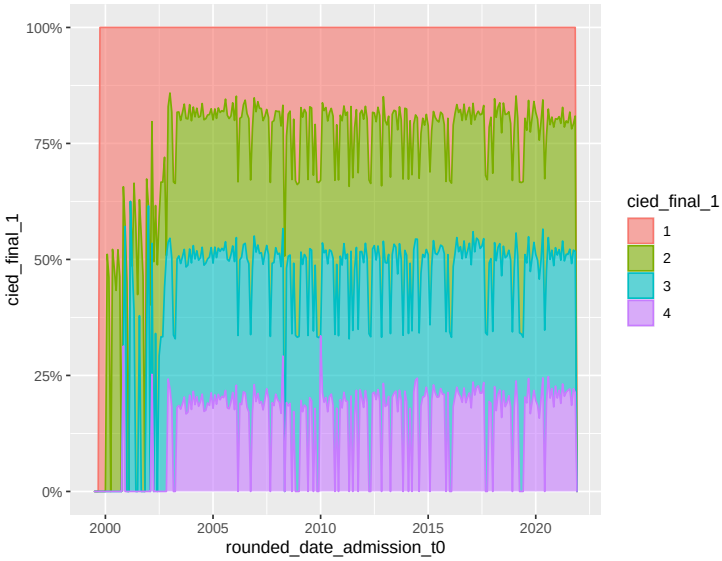

copd
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
copd
0
1

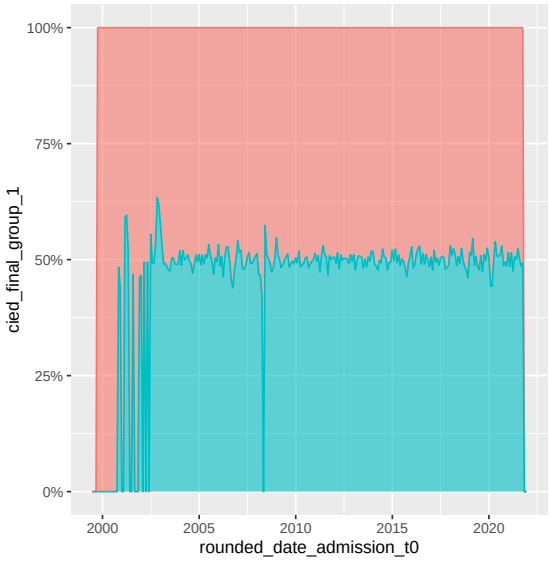

cancer
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
cancer
0
1

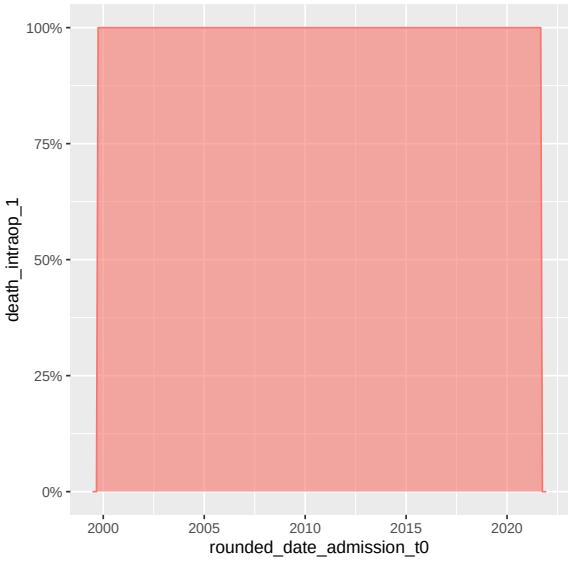

surgery_count
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
surgery_count
1
10
2
3
4
5
6
7
8
9

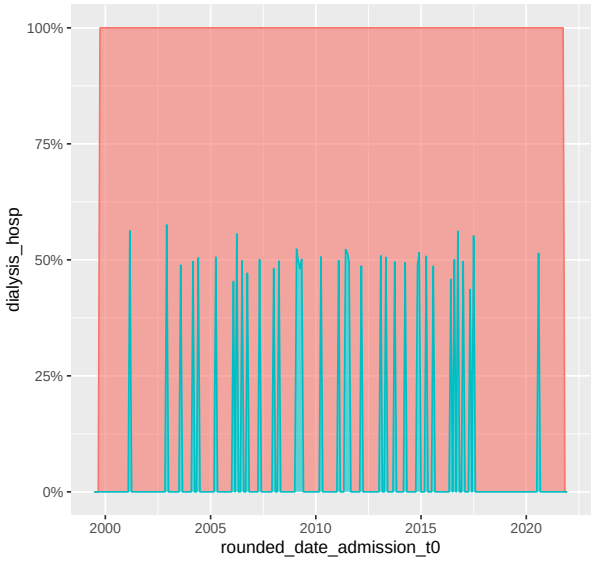

procedure_type_1
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
procedure_type_1
1
2

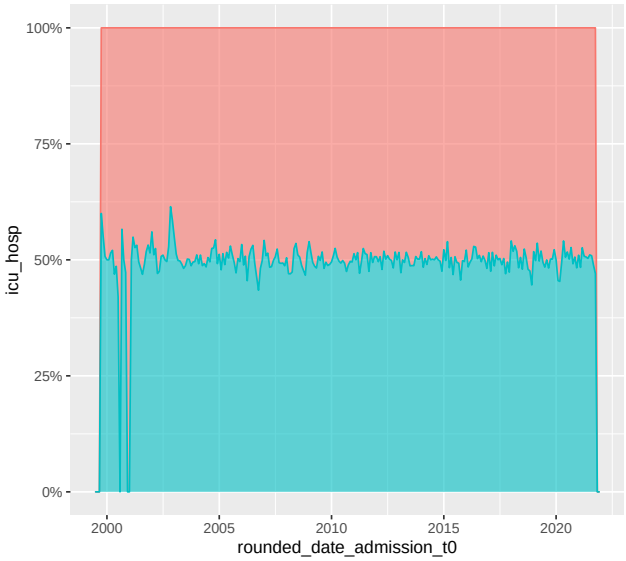

100%
75%
50%
25%
0%
procedure_type_new
2000
2005
2010
2015
2020
rounded_date_admission_t0
procedure_type_new
1
2
3
4

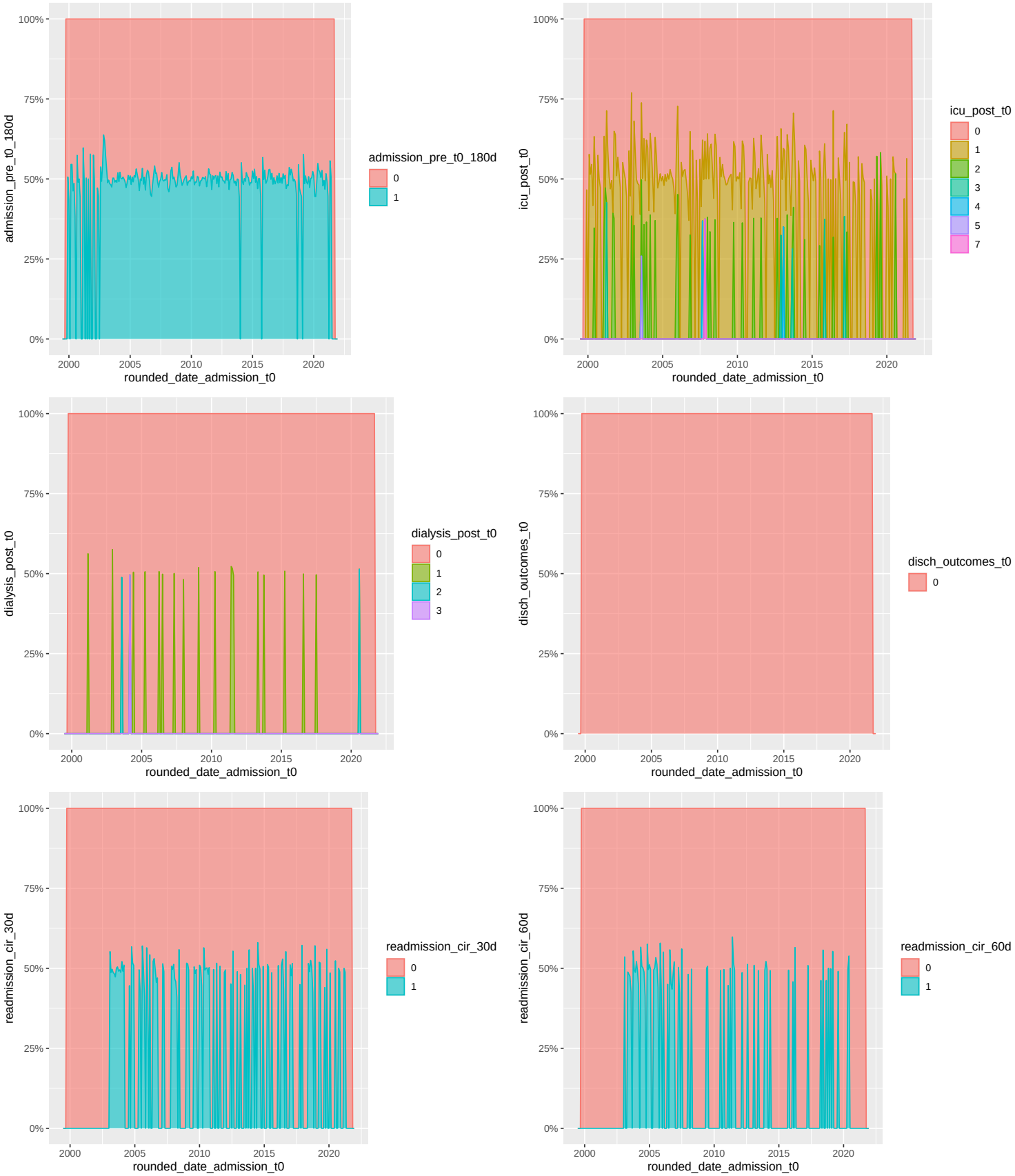

100%
75%
50%
25%
0%
cied_final_1
2000
2005
2010
2015
2020
rounded_date_admission_t0
cied_final_1
1
2
3
4

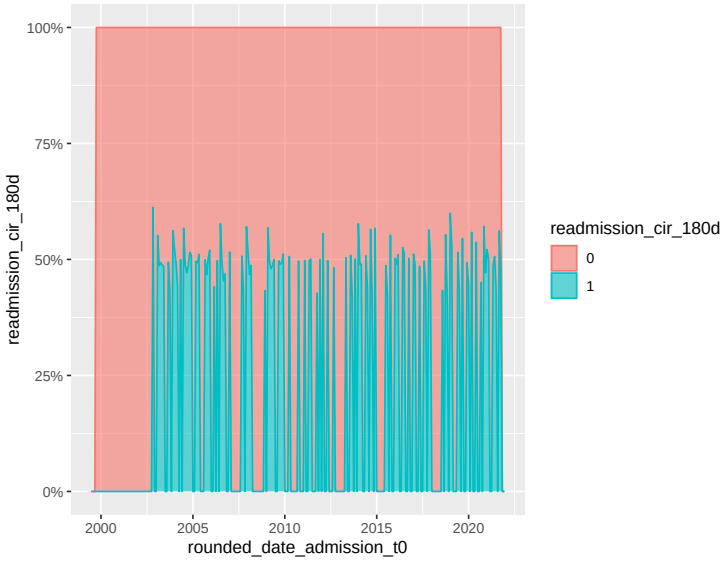

100%
75%
50%
25%
0%
cied_final_group_1
2000
2005
2010
2015
2020
rounded_date_admission_t0
cied_final_group_1
1
2

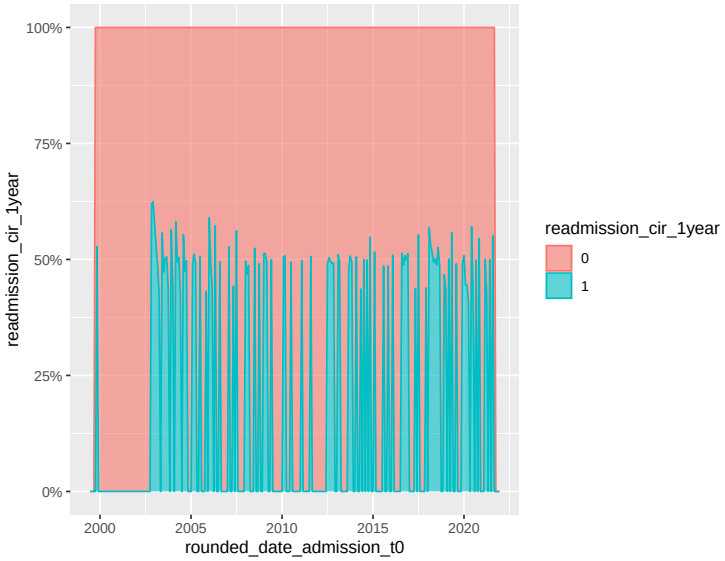

100%
75%
50%
25%
0%
death_intraop_1
2000
2005
2010
2015
2020
rounded_date_admission_t0
death_intraop_1
0

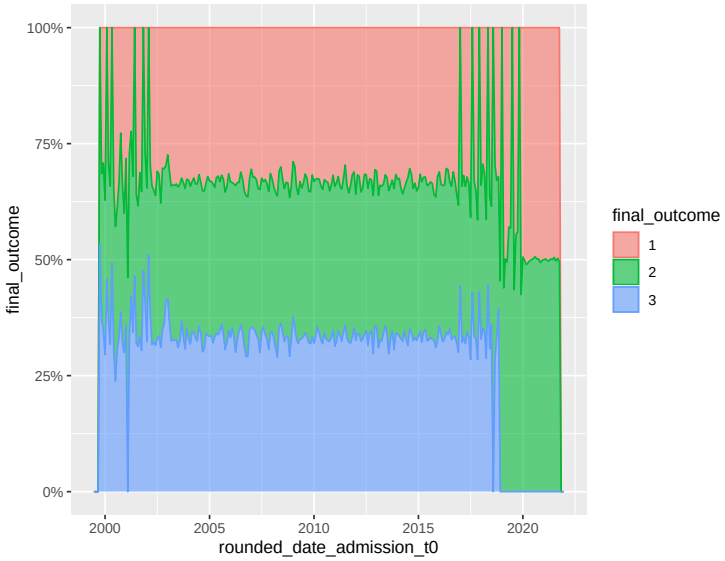

100%
75%
50%
25%
0%
dialysis_hosp
2000
2005
2010
2015
2020
rounded_date_admission_t0
dialysis_hosp
0
1

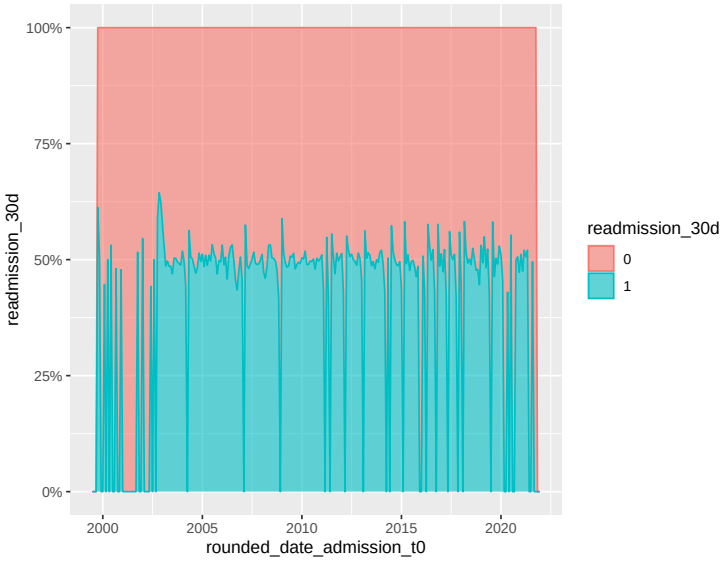

100%
75%
50%
25%
0%
icu_hosp
2000
2005
2010
2015
2020
rounded_date_admission_t0
icu_hosp
0
1

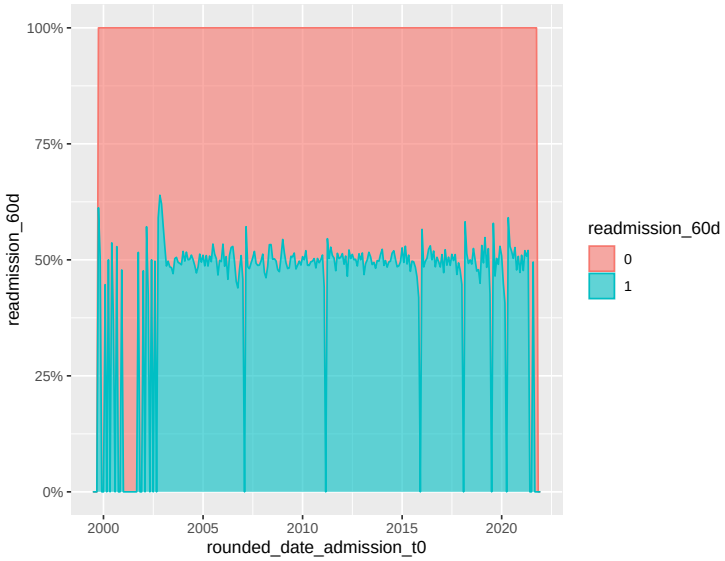

admission_pre_t0_180d
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
admission_pre_t0_180d
0
1
icu_post_t0
icu_post_t0
0
1
2
3
4
5
7
dialysis_post_t0
dialysis_post_t0
0
1
2
3
disch_outcomes_t0
disch_outcomes_t0
0
readmission_cir_30d
readmission_cir_30d
0
1
readmission_cir_60d
readmission_cir_60d
0
1

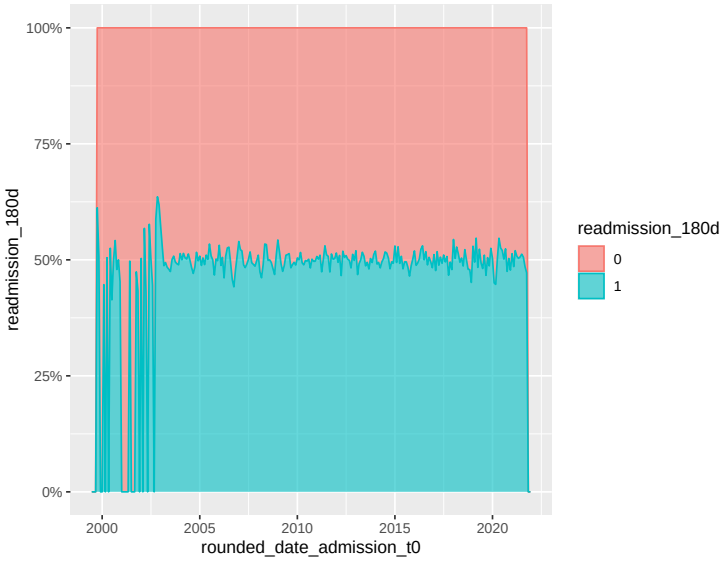

100%
75%
50%
25%
0%
readmission_cir_180d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_cir_180d
0
1

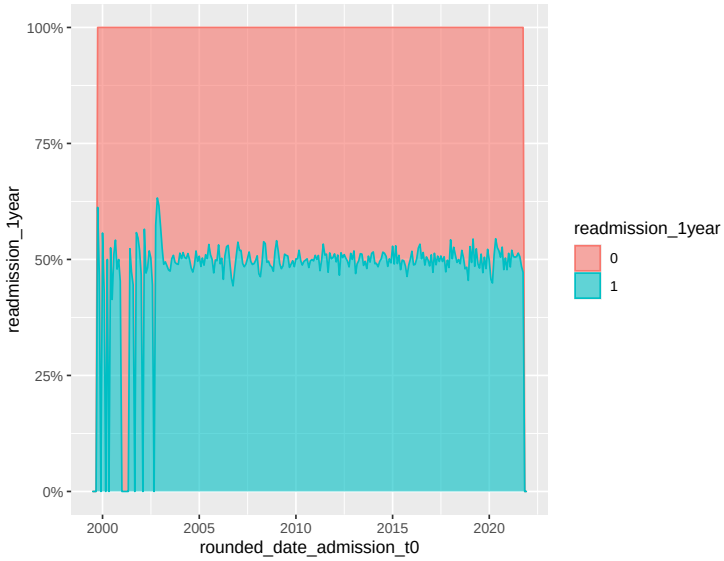

100%
75%
50%
25%
0%
readmission_cir_1year
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_cir_1year
0
1

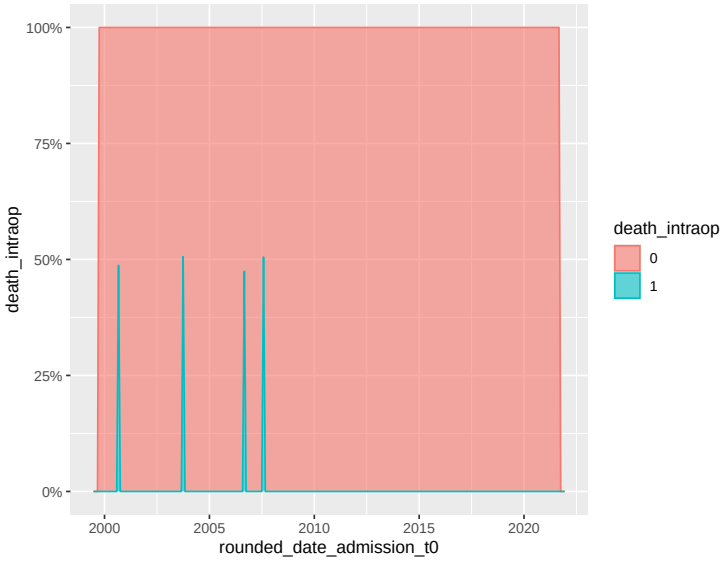

100%
75%
50%
25%
0%
final_outcome
2000
2005
2010
2015
2020
rounded_date_admission_t0
final_outcome
1
2
3

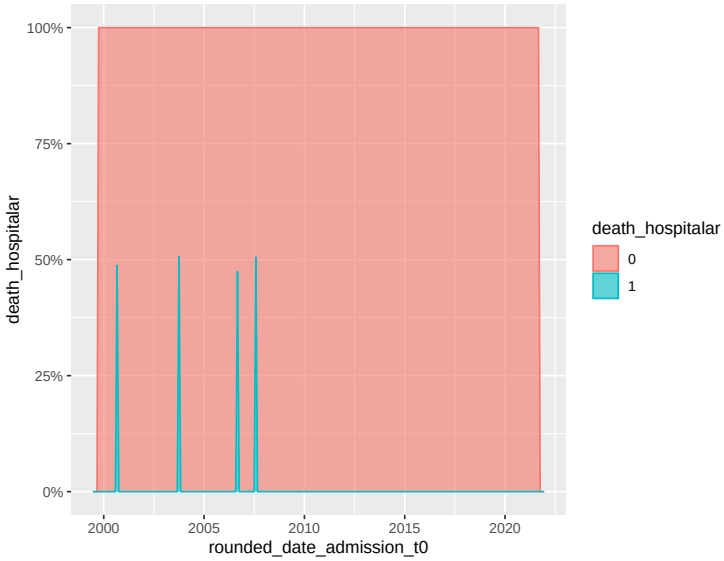

100%
75%
50%
25%
0%
readmission_30d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_30d
0
1

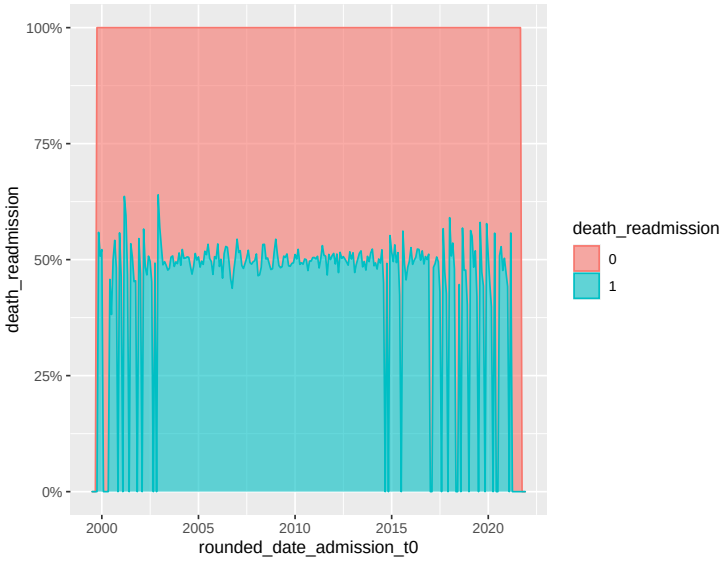

100%
75%
50%
25%
0%
readmission_60d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_60d
0
1

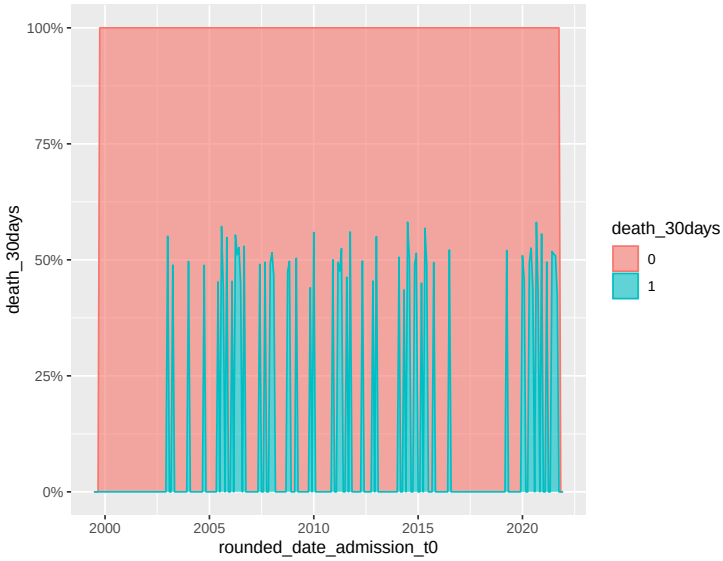

100%
75%
50%
25%
0%
readmission_180d
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_180d
0
1

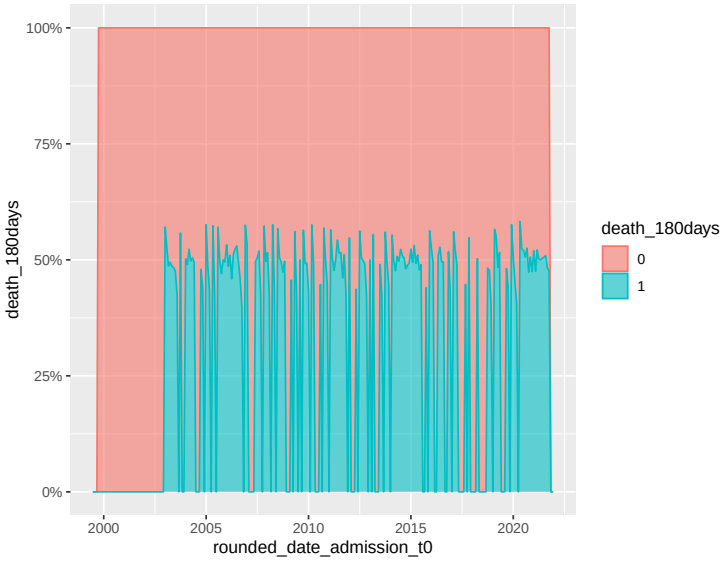

readmission_1year
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
readmission_1year
0
1

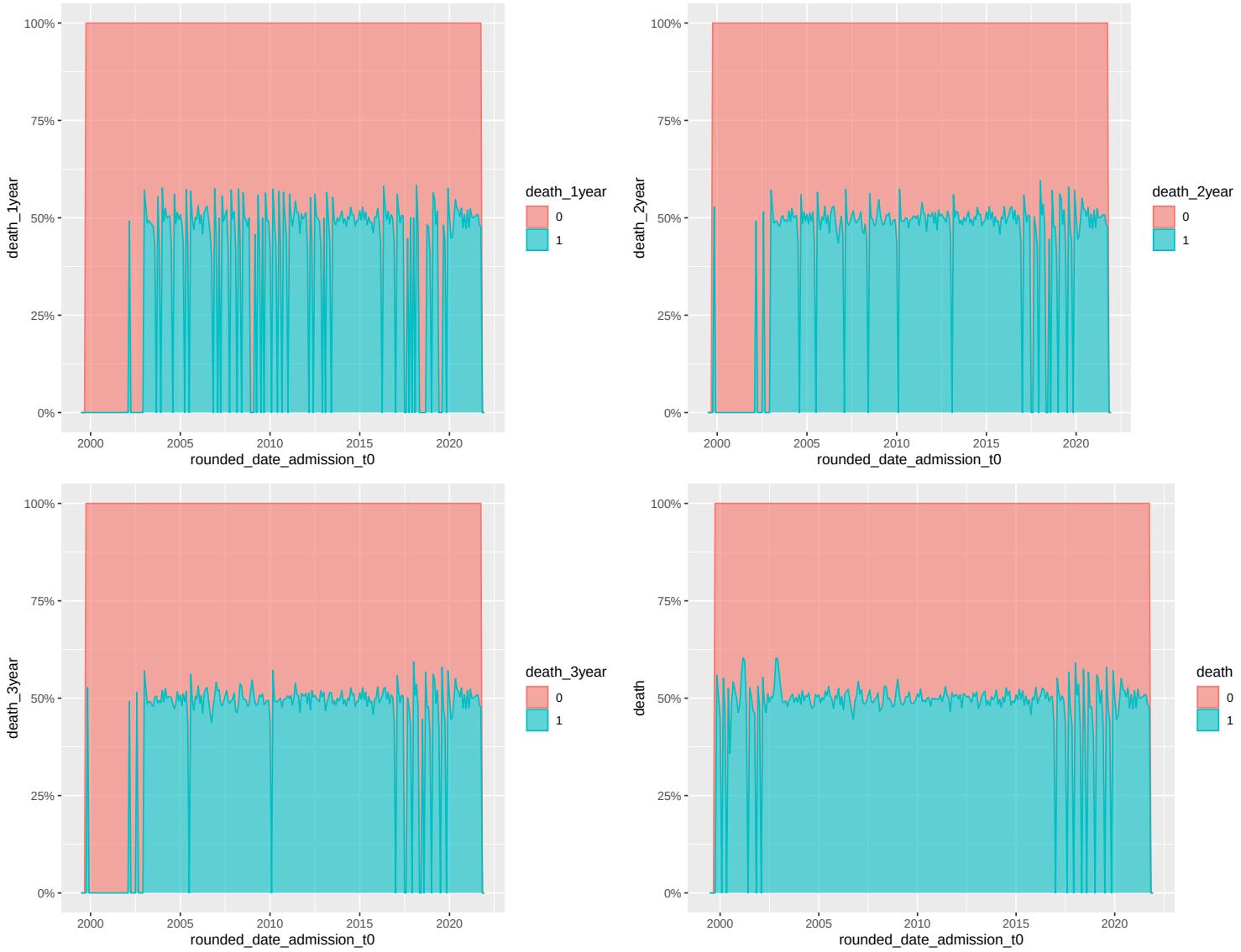

death_intraop
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
death_intraop
0
1


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


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


death_30days
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
death_30days
0
1


death_180days
100%
75%
50%
25%
0%
2000
2005
2010
2015
2020
rounded_date_admission_t0
death_180days
0
1




```
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
```