

Covid Staging Data

Steph Reynolds (Stephanie.Reynolds@ucsf.edu)

Dec 14, 2021 14:07 PM

Background

Project

MINDSCAPE: Modeling of infectious network dynamics for surveillance, control and prevention enhancement

Description

This file imports `bin_clin_scen_df_11.08.2021.csv` which contains all binary clinical scenarios and O2 devices for all patients, and runs a function to generate WHO Clinical Stages of Severity (4-10).

SourceData

- Binary Clinical Scenarios and O2 Devices (`bin_clin_scen_df_11.08.2021.csv`)
 - This file contains data on patients' O2 flow rates, clinical scenarios, receipt of O2 devices and their respective category of respiratory support.

Load required packages

```
library(here)
library(tidyverse)
library(tableone)
```

Import and preview data

```
bin_clin_scen_df <- read_csv(here("data", "binary_clin_scen_df_11.08.21.csv"))

## Rows: 15741 Columns: 33

## -- Column specification -----
## Delimiter: ","
## chr   (7): ID, sex, zip, race, ethnicity, smoking, end_in_death
## dbl   (24): age, BMI, LOS, VP, ECMO, CRRT, NIV, NIV_per_day, HD, INTUB, SF_LT...
## date  (2): date, death_date
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
# Look at structure of df and variable types
# str(bin_clin_scen_df)
head(bin_clin_scen_df, n=40)
```

```
## # A tibble: 40 x 33
##   ID      date      age sex   zip   race ethnicity smoking  BMI end_in_death
##   <chr> <date>    <dbl> <chr> <chr> <chr> <chr>      <chr>  <dbl> <chr>
## 1 2db2~ 2021-05-06    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 2 2db2~ 2021-05-07    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 3 2db2~ 2021-05-08    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 4 2db2~ 2021-05-09    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 5 2db2~ 2021-05-10    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 6 2db2~ 2021-05-11    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 7 2db2~ 2021-05-12    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 8 2db2~ 2021-05-13    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 9 2db2~ 2021-05-14    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## 10 2db2~ 2021-05-15    70 Male 95490 Whit~ Not Hisp~ Not Cu~ 23.6 No
## # ... with 30 more rows, and 23 more variables: death_date <date>, LOS <dbl>,
## #   VP <dbl>, ECMO <dbl>, CRRT <dbl>, NIV <dbl>, NIV_per_day <dbl>, HD <dbl>,
## #   INTUB <dbl>, SF_LT_200 <dbl>, O2 <dbl>, LowO2 <dbl>, HighO2 <dbl>,
## #   NODEV <dbl>, SIMPLEDEV <dbl>, SIMPLE_PER_DAY <dbl>, NIVDEV <dbl>,
## #   IVDEV <dbl>, CPAPDEV <dbl>, NCDEV <dbl>, NC_PER_DAY <dbl>, NC_GT_12 <dbl>,
## #   DEATH <dbl>
```

Clean bin_clin_scen_df – add new variable STAGE and set values to NA

```
# Create new variable `STAGE` and set values to NA
bin_clin_scen_df$STAGE <- NA
```

Create function to get COVID stages based on patients' binary clinical scenarios and usage of O2 devices (and their respective level of respiratory support)

```
stages_df <- bin_clin_scen_df %>%
  mutate(STAGE = case_when(DEATH==1 ~ 10,
                           ECMO==1 |
                             ((INTUB==1 | IVDEV==1) & SF_LT_200==1 & (VP==1 | CRRT==1)) ~ 9,      # SHO
                             ((INTUB==1 | IVDEV==1) & SF_LT_200==1) |      # SHOULD I JUST USE 'IVDEV' OR (
                             VP==1 |
                             CRRT==1 ~ 8,
                             (INTUB==1 | IVDEV==1) &
```

```

        SF_LT_200==0 ~ 7,
(NIVDEV==1 & CPAPDEV==0) &
    IVDEV==0 &
        ((NIV_per_day > NC_PER_DAY) | NC_GT_12==1 | (NC_PER_DAY>0 & NC_GT_12==1)) ~
SIMPLEDEV==1 & (LowO2==1 | HighO2 == 1 | SIMPLE_PER_DAY > 1) ~ 5,
NC_PER_DAY>0 & NC_GT_12==0 ~ 5,
(NIVDEV==0 & CPAPDEV==0 & INTUB==0 & IVDEV==0 & SIMPLEDEV==0) & LowO2==1 ~ 5,
TRUE ~ 4))

```

Save and export file as .csv file

```

write_csv(stages_df, here("data", "stages_df_11.08.21.csv"))

```

End of Document