Preoperative Atelectasis

Part 2: Descriptive characteristics and map

Javier Mancilla Galindo

2024-04-08

Table of contents

Setup

Packages used

```
if (!require("pacman", quietly = TRUE)) {
  install.packages("pacman")
}


pacman::p_load(
  tidyverse, # Used for basic data handling and visualization.
  table1, #Used to create table of descriptive characteristics of sample.
  RColorBrewer, #Color palettes for data visualization.
  gridExtra, #Used to arrange multiple ggplots in a grid.
  grid, #Used to arrange multiple ggplots in a grid.
```

```
    rnaturalearth, #Used to extract geographical data to create maps.
    rnaturalearthhires, #Used together with the prior package to create map.
    sf, #Used together with the prior package to create map.
    plotly, #Used together with prior two packages to create map.
    reticulate, #Required together with plotly
    flextable, #Used to export tables.
    officer,  #Used to export tables.
    report, #Used to cite packages used in this session.
    webshot2 #Required to render this document as pdf.
  )
```

Session and package dependencies

```
R version 4.3.3 (2024-02-29 ucrt)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 11 x64 (build 22631)

Matrix products: default


locale:
[1] LC_COLLATE=Spanish_Mexico.utf8  LC_CTYPE=Spanish_Mexico.utf8
[3] LC_MONETARY=Spanish_Mexico.utf8 LC_NUMERIC=C
[5] LC_TIME=Spanish_Mexico.utf8

time zone: Europe/Berlin
tzcode source: internal

attached base packages:
[1] grid      stats     graphics  grDevices datasets  utils     methods
[8] base

other attached packages:
 [1] remotes_2.5.0                  webshot2_0.1.1
 [3] report_0.5.8                   officer_0.6.5
 [5] flextable_0.9.5                reticulate_1.35.0
 [7] plotly_4.10.4                  sf_1.0-15
 [9] rnaturalearthhires_1.0.0.9000  rnaturalearth_1.0.1
[11] gridExtra_2.3                  RColorBrewer_1.1-3
[13] table1_1.4.3                   lubridate_1.9.3
[15] forcats_1.0.0                  stringr_1.5.1
```

```
[17] dplyr_1.1.4               purrr_1.0.2
[19] readr_2.1.5               tidyr_1.3.1
[21] tibble_3.2.1              ggplot2_3.5.0
[23] tidyverse_2.0.0           pacman_0.5.1
```

State of residence of participants

Map generated with the accompanying script ***Map_USA_Canada.R***

This map was built by partly using code adapted from contribution by cpsievert.

State	n
Texas	81
Washington	39
Utah	27
Alberta	21
Florida	20
California	16
Louisiana	16
Arkansas	7
Missouri	5
Nevada	4

Number of patients

80
70
60
50
40
30
20
10
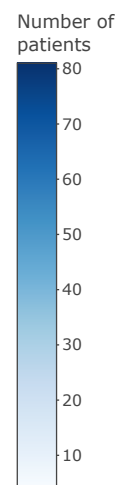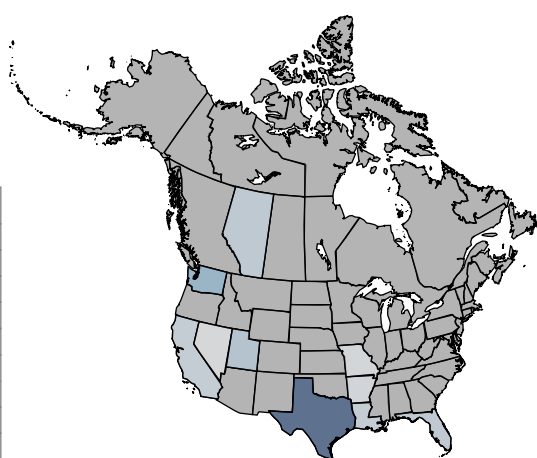
Distribution of numerical variables

Distributions were examined with the accompanying sourced function ***distribution__numerical__variables.R***

Near normal distribution:
- Age: light tails
- height: heavy right tail, 4 outliers right
- hb: heavy tails, bilateral outliers
- hct: heavy tails, bilateral outliers
- leu: near normal, bilateral outliers
- neu_absolute: heavy right tail, two right outliers
- linf_absolute: heavy right tail, bilateral outliers (more right)
- mon_absolute: heavy right tail, bilateral outliers (more right)
- platelets: two right outliers
- urea: four right outliers
- creatinine: three right outliers

Distribution not normal:
- Weight: right-skewed, outliers are verified observations of extreme weight.
- BMI: right-skewed, outliers are verified observations of extreme BMI.
- spo2_VPO: Left-skewed
- neu_percent: left-skewed
- linf_percent: right-skewed
- glucose: right-skewed
- mon_percent: observations around only 5 data points. Will not use this variable, only absolute monocytes will be used.
- altitude: distribution not clear as values are quite apart an concentrate around single states with diferring mean altitudes. Will attempt to model a smooth term or categorical term in subsequent analyses.

Outcome variable:
- atelectasis_percent: Zero-inflated. Would be difficult to manage as categorical ordinal due to low number of patients in some categories. Will re-assess alongside subsequent analyses to decide.

Characteristics of participants

Table 1 generated with the accompanying sourced script ***table1__arguments.R***

Characteristics of participants are shown for the total sample and by obesity class category as defined by the World Health Organization:

- Class 1, BMI (30-35) kg/m2

- Class 2, BMI (35-40) kg/m2

- Class 3, BMI >40 kg/m2

Characteristics of participants according to BMI class are shown in **Table 1**.

Table 1

	Total	Class 1 Obesity	Class 2 Obesity	Class 3 Obesity
	(N=236)	(N=63)	(N=53)	(N=120)
Sex				
Woman	214 (90.7%)	60 (95.2%)	48 (90.6%)	106 (88.3%)
Man	22 (9.3%)	3 (4.8%)	5 (9.4%)	14 (11.7%)
Age (years)				
Mean (SD)	40.3 (9.87)	42.1 (10.2)	40.8 (9.25)	39.1 (9.85)
Weight (kilograms (kg))				
Median [Q1, Q3]	111 [97.4, 130]	89.2 [84.5, 95.9]	107 [102, 112]	128 [114, 142]
Height (meters (m))				
Mean (SD)	1.67 (0.0822)	1.67 (0.0674)	1.69 (0.0856)	1.67 (0.0876)
Body mass index (kg/m²)				
Median [Q1, Q3]	40.3 [34.6, 46.0]	33.0 [31.5, 33.8]	38.3 [36.6, 39.1]	45.8 [42.4, 51.2]
Surgical procedure				
LBGS	31 (13.1%)	5 (7.9%)	9 (17.0%)	17 (14.2%)
OAGB	5 (2.1%)	1 (1.6%)	1 (1.9%)	3 (2.5%)
RYGB	6 (2.5%)	1 (1.6%)	1 (1.9%)	4 (3.3%)
SG	189 (80.1%)	53 (84.1%)	41 (77.4%)	95 (79.2%)
ARISCAT risk group				
Intermediate Risk	61 (25.8%)	18 (28.6%)	12 (22.6%)	31 (25.8%)
Low Risk	175 (74.2%)	45 (71.4%)	41 (77.4%)	89 (74.2%)
CO-RADS				

	Total	Class 1 Obesity	Class 2 Obesity	Class 3 Obesity
CO-RADS 1	230 (97.5%)	62 (98.4%)	51 (96.2%)	117 (97.5%)
CO-RADS 2	6 (2.5%)	1 (1.6%)	2 (3.8%)	3 (2.5%)
Oxygen saturation (SpO2) (%)				
Median [Q1, Q3]	96.0 [93.0, 97.0]	97.0 [95.0, 97.5]	96.0 [94.0, 97.0]	94.0 [92.0, 97.0]
Mean altitude (meters)				
Median [Q1, Q3]	519 [519, 806]	519 [382, 806]	519 [519, 885]	519 [519, 806]
Acute Myocardial Infarction				
No	210 (89.0%)	62 (98.4%)	52 (98.1%)	96 (80.0%)
Yes	26 (11.0%)	1 (1.6%)	1 (1.9%)	24 (20.0%)
Hypertension				
No	177 (75.0%)	53 (84.1%)	40 (75.5%)	84 (70.0%)
Yes	59 (25.0%)	10 (15.9%)	13 (24.5%)	36 (30.0%)
Diabetes				
No	211 (89.4%)	58 (92.1%)	48 (90.6%)	105 (87.5%)
Yes	25 (10.6%)	5 (7.9%)	5 (9.4%)	15 (12.5%)
Asthma				
No	216 (91.5%)	56 (88.9%)	46 (86.8%)	114 (95.0%)
Yes	20 (8.5%)	7 (11.1%)	7 (13.2%)	6 (5.0%)
COPD				
No	228 (96.6%)	62 (98.4%)	53 (100%)	113 (94.2%)
Yes	8 (3.4%)	1 (1.6%)	0 (0%)	7 (5.8%)
Obstructive sleep apnea				
No	203 (86.0%)	60 (95.2%)	50 (94.3%)	93 (77.5%)
Yes	33 (14.0%)	3 (4.8%)	3 (5.7%)	27 (22.5%)
Oxygen use				
No	206 (87.3%)	60 (95.2%)	50 (94.3%)	96 (80.0%)
Yes	30 (12.7%)	3 (4.8%)	3 (5.7%)	24 (20.0%)
CPAP				
No	203 (86.0%)	60 (95.2%)	50 (94.3%)	93 (77.5%)
Yes	33 (14.0%)	3 (4.8%)	3 (5.7%)	27 (22.5%)
Hypothyroidism				
No	213 (90.3%)	56 (88.9%)	50 (94.3%)	107 (89.2%)
Yes	23 (9.7%)	7 (11.1%)	3 (5.7%)	13 (10.8%)
Dyslipidemia				
No	218 (92.4%)	59 (93.7%)	48 (90.6%)	111 (92.5%)

	Total	Class 1 Obesity	Class 2 Obesity	Class 3 Obesity
Yes	18 (7.6%)	4 (6.3%)	5 (9.4%)	9 (7.5%)
Antidepressants use				
No	142 (60.2%)	37 (58.7%)	33 (62.3%)	72 (60.0%)
Yes	94 (39.8%)	26 (41.3%)	20 (37.7%)	48 (40.0%)

NOTE: The **ASA** physical status variable has not been included in analyses since the updated version of ASA consulted in October 2023 classifies obesity (30<BMI<40) as ASA 2 and obesity (BMI 40) as ASA 3. The distribution of frequencies of ASA~obesity class in this dataset does not match such definition. This occurred since an outdated version of ASA that did not include obesity was likely used by clinicians when writing the preoperative assessment medical note:

```
      Class 1 Obesity Class 2 Obesity Class 3 Obesity
ASA 1              31              18               3
ASA 2              30              34              84
ASA 3               0               0              32
```

Package references

- Auguie B (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics.* R package version 2.3, https://CRAN.R-project.org/package=gridExtra.
- Chang W (2023). *webshot2: Take Screenshots of Web Pages.* R package version 0.1.1, https://CRAN.R-project.org/package=webshot2.
- Gohel D, Moog S (2024). *officer: Manipulation of Microsoft Word and PowerPoint Documents.* R package version 0.6.5, https://CRAN.R-project.org/package=officer.
- Gohel D, Skintzos P (2024). *flextable: Functions for Tabular Reporting.* R package version 0.9.5, https://CRAN.R-project.org/package=flextable.
- Grolemund G, Wickham H (2011). "Dates and Times Made Easy with lubridate." *Journal of Statistical Software*, *40*(3), 1-25. https://www.jstatsoft.org/v40/i03/.
- Makowski D, Lüdecke D, Patil I, Thériault R, Ben-Shachar M, Wiernik B (2023). "Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption." *CRAN*. https://easystats.github.io/report/.
- Massicotte P, South A (2023). *rnaturalearth: World Map Data from Natural Earth.* R package version 1.0.1, https://CRAN.R-project.org/package=rnaturalearth.
- Müller K, Wickham H (2023). *tibble: Simple Data Frames.* R package version 3.2.1, https://CRAN.R-project.org/package=tibble.
- Neuwirth E (2022). *RColorBrewer: ColorBrewer Palettes.* R package version 1.1-3, https://CRAN.R-project.org/package=RColorBrewer.

- Pebesma E, Bivand R (2023). *Spatial Data Science: With applications in R.* Chapman and Hall/CRC. doi:10.1201/9780429459016 https://doi.org/10.1201/9780429459016, https://r-spatial.org/book/. Pebesma E (2018). "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal*, *10*(1), 439-446. doi:10.32614/RJ-2018-009 https://doi.org/10.32614/RJ-2018-009, https://doi.org/10.32614/RJ-2018-009.
- R Core Team (2024). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
- Rich B (2023). *table1: Tables of Descriptive Statistics in HTML.* R package version 1.4.3, https://CRAN.R-project.org/package=table1.
- Rinker TW, Kurkiewicz D (2018). *pacman: Package Management for R.* version 0.5.0, http://github.com/trinker/pacman.
- Sievert C (2020). *Interactive Web-Based Data Visualization with R, plotly, and shiny.* Chapman and Hall/CRC. ISBN 9781138331457, https://plotly-r.com.
- South A, Michael S, Massicotte P (????). *rnaturalearthhires: High Resolution World Vector Map Data from Natural Earth used in rnaturalearth.* R package version 1.0.0.9000, https://github.com/ropensci/rnaturalearthhires, https://docs.ropensci.org/rnaturalearthhires.
- Ushey K, Allaire J, Tang Y (2024). *reticulate: Interface to 'Python'.* R package version 1.35.0, https://CRAN.R-project.org/package=reticulate.
- Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.
- Wickham H (2023). *forcats: Tools for Working with Categorical Variables (Factors).* R package version 1.0.0, https://CRAN.R-project.org/package=forcats.
- Wickham H (2023). *stringr: Simple, Consistent Wrappers for Common String Operations.* R package version 1.5.1, https://CRAN.R-project.org/package=stringr.
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." *Journal of Open Source Software*, *4*(43), 1686. doi:10.21105/joss.01686 https://doi.org/10.21105/joss.01686.
- Wickham H, François R, Henry L, Müller K, Vaughan D (2023). *dplyr: A Grammar of Data Manipulation.* R package version 1.1.4, https://CRAN.R-project.org/package=dplyr.
- Wickham H, Henry L (2023). *purrr: Functional Programming Tools.* R package version 1.0.2, https://CRAN.R-project.org/package=purrr.
- Wickham H, Hester J, Bryan J (2024). *readr: Read Rectangular Text Data.* R package version 2.1.5, https://CRAN.R-project.org/package=readr.
- Wickham H, Vaughan D, Girlich M (2024). *tidyr: Tidy Messy Data.* R package version 1.3.1, https://CRAN.R-project.org/package=tidyr.