Gasto cardíaco por ECOTT vs Fick

Parte 0: Limpieza de datos e impresión inicial

J. Mancilla Galindo, E. Garza Santiago, L.A. Fernández Urrutia

2024-12-10

Session and package dependencies

```
R version 4.4.0 (2024-04-24 ucrt)
Platform: x86_64-w64-mingw32/x64
Running under: Windows 11 x64 (build 22631)

Matrix products: default


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

time zone: Europe/Amsterdam
tzcode source: internal

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

other attached packages:
 [1] report_0.5.9     styler_1.10.3    haven_2.5.4      ggrepel_0.9.6
 [5] gt_0.11.0        readxl_1.4.3     overviewR_0.0.13 lubridate_1.9.3
 [9] forcats_1.0.0    stringr_1.5.1    dplyr_1.1.4      purrr_1.0.2
[13] readr_2.1.5      tidyr_1.3.1      tibble_3.2.1     ggplot2_3.5.1
[17] tidyverse_2.0.0  pacman_0.5.1
```

Load data and import attributes from dataset dictionary

```
# Load data
raw_data <- read_excel(
  file.path(inputfolder,"GC X FICK vs ECOTT.xlsm"),
  sheet = "Sheet1",
  range = "A6:AZ91",
  na = "NA"
  )

# Load variable dictionary
dictionary <- read_excel(
  file.path(inputfolder,"Variables dictionary ventilación mecánica corregido.xlsx"),
  range = cell_cols("A:I"),
  na = "."
  )
```

Manual inspection of the raw data and the dictionary revealed that the variable names in the raw data are not the same as the variable names in the dictionary. These were checked with the clinical team to update the dataset dictionary for the correct variable names.

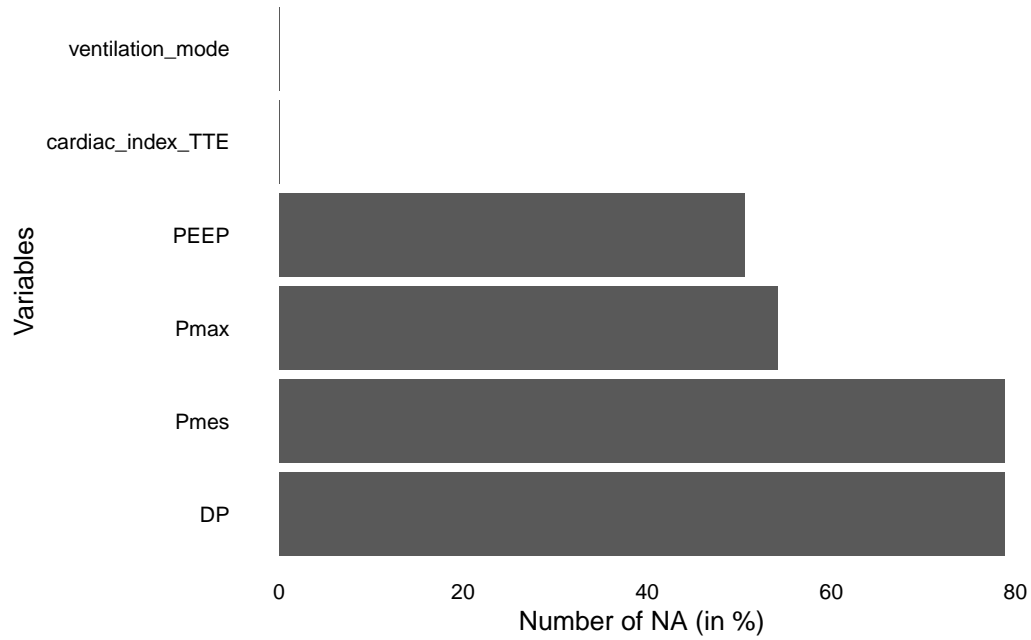The number of unique participants in the study is:

$$\frac{n}{56}$$

Of these, the number of unique participants and participants with more than one measurement is as follows:

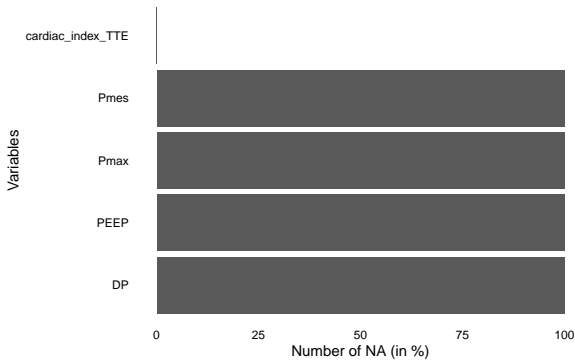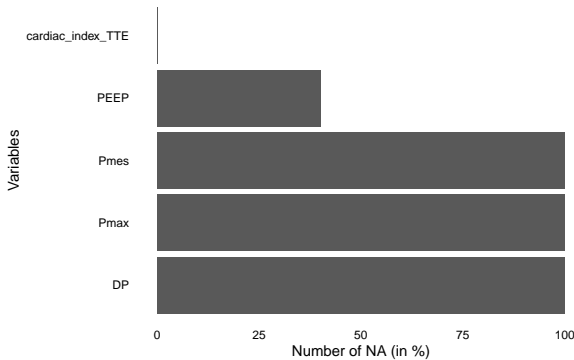| Measurements	n
1	41
2	8
3	5
4	1
9	1

Explore data

Ventilation Parameters

Missing data likely corresponding to participants without mechanical ventilation. I will filter and check.
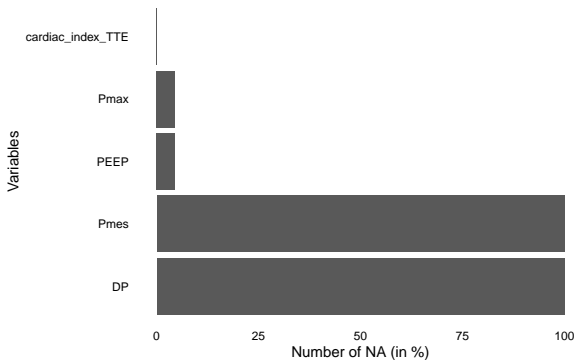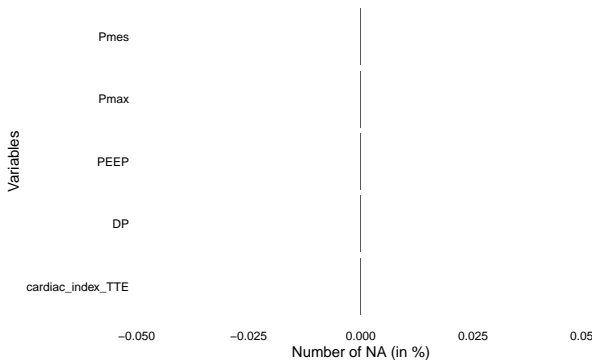
Missing data per ventilation mode

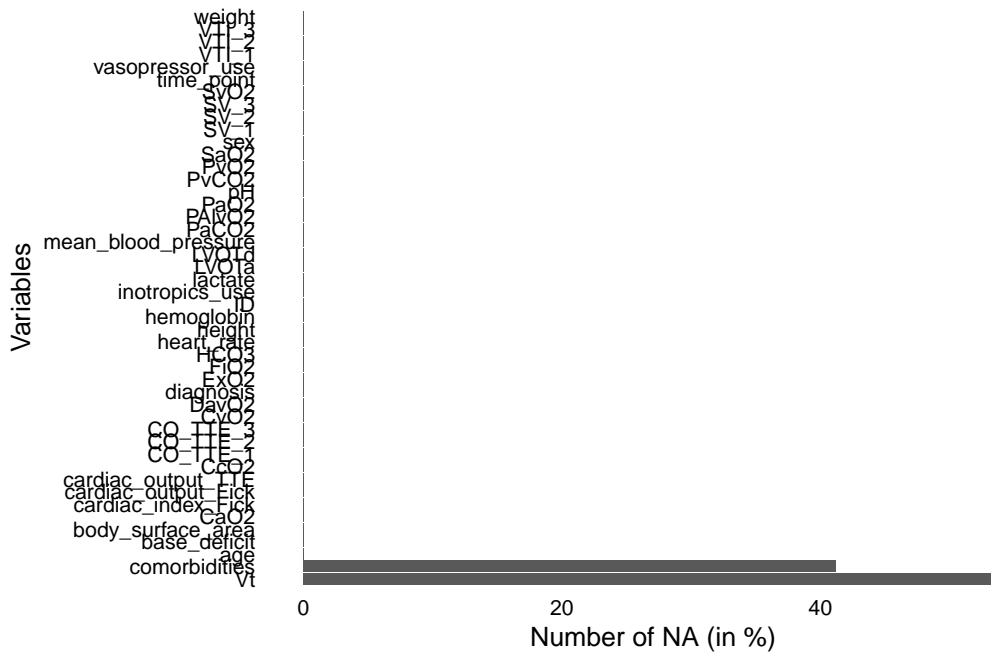(a) MR, Without MV, Spontaneous

(a) HFNC

(a) CPAP|PS

(a) ACV

Other variables

There are only missing values in the comorbidities column, corresponding to participants with no comorbidities reported.

The following function was used to check if diagnosis, comorbidity, and other baseline values matched for all rows with the same ID. Any unmatches were corroborated with the medical team and corrected.

```
check_non_unique <- function(data, id_col, x_col) {
  # Group by ID and summarize to count unique values of x for each ID
  non_unique_ids <- data %>%
    group_by(!!sym(id_col)) %>% # Group by the ID column
    summarize(unique_x_count = n_distinct(!!sym(x_col))) %>% # Count unique values of x
    filter(unique_x_count > 1) %>% # Filter IDs where x is not unique
    pull(!!sym(id_col)) # Pull the IDs where x is not unique

  # Return the non-unique IDs
  return(non_unique_ids)
}
```

Weight and BMI

There are slight variations in `weight` for values reported in the same individual. When consulted with the clinical team, such variations were said to be due to weight loss during hospital stay. It is desirable to use the exact calculations for body surface area and cardiac output calculations. However, for the reporting of baseline BMI, I will use the higher weight value at admission.

```r
data <- data %>%
  group_by(ID) %>%
  mutate(BMI = round((max(weight)/(height^2)),2)) %>%
  ungroup() %>%
  relocate(BMI, .after = height)

attr(data$BMI, "label") <- "Body Mass Index"
attr(data$BMI, "unit") <- "kg/m²"
```

```r
data <- data %>%
  mutate(
    obesity = factor(case_when(
      BMI < 18.5 ~ "Underweight",
      BMI >= 18.5 & BMI < 25 ~ "Normal",
      BMI >= 25 & BMI < 30 ~ "Overweight",
      BMI >= 30 & BMI < 35 ~ "Obesity class I",
      BMI >= 35 & BMI < 40 ~ "Obesity class II",
      BMI >= 40 ~ "Obesity class III"
    )) %>%
    fct_relevel(
      c("Underweight", "Normal", "Overweight",
        "Obesity class I", "Obesity class II", "Obesity class III")
    )) %>%
  relocate(obesity, .after = BMI)

attr(data$obesity, "label") <- "Obesity status"
```
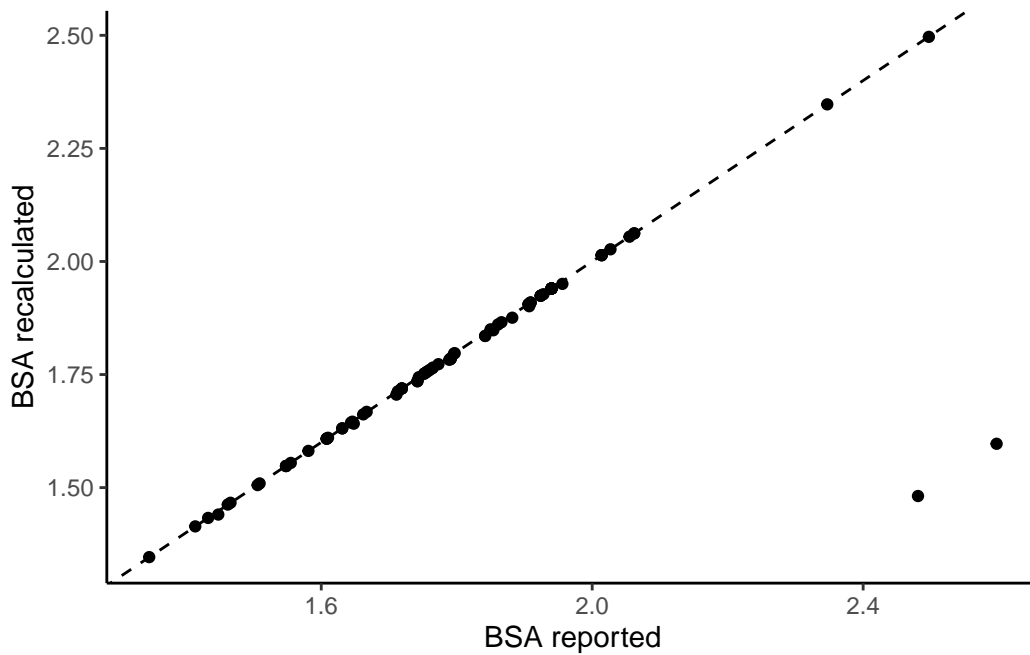
Body surface area

Recalculation of body surface area to corroborate values in data:

$$BSA = \sqrt{\frac{height \times weight}{3600}}$$

Check if recalculated BSA values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Body surface area	body_surface_area	Numeric	1.0-3.0	square meters (m²)

min	max
1.35	2.5

I will thus replace the BSA values in the dataset with the recalculated values.

7

Cardiac output (CO)

I will recalculate cardiac output using the echocardiographic (ECOTT) and Fick methods to corroborate.

ECOTT

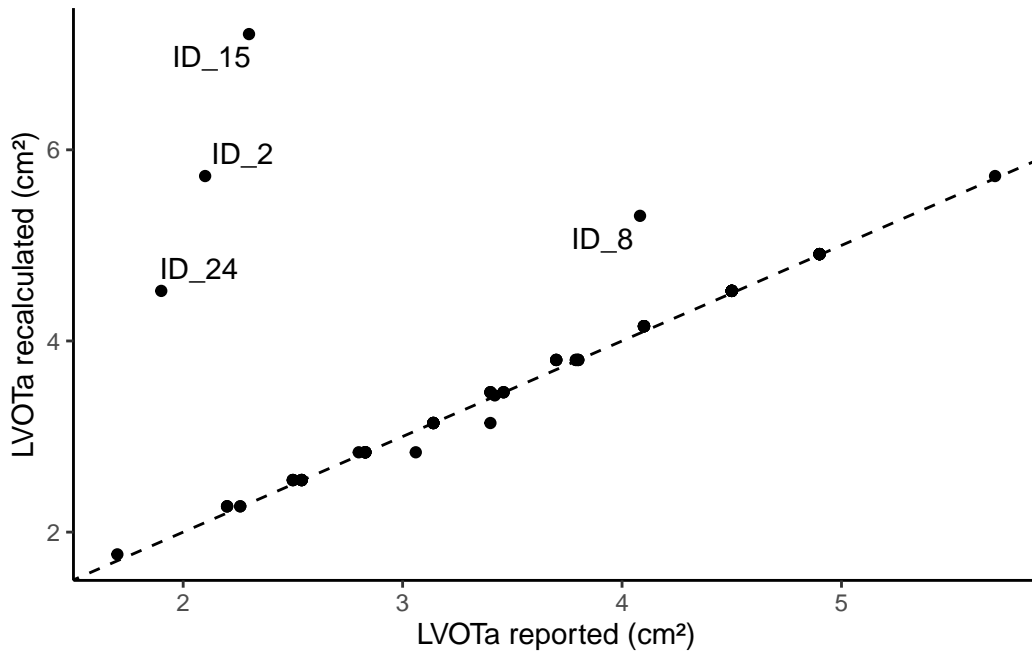### LVOTa

The left ventricle outflow tract area (LVOTa) is calculated using the formula:

$$LVOTa = \pi(d/2)^2$$

where d = Left ventricle outflow tract diameter (LVOTd)

The highlighted IDs in the plot were selected for manual inspection of the LVOTd and LVOTa values due to the large differences in recalculated values which cannot be explained by rounding error. The clinical team was unable to trace back which of the two values was incorrect (LVOTd or LVOTa). Thus, these participants are excluded from subsequent analyses.

After exclusion, the number of unique participants in the study is:

n
52
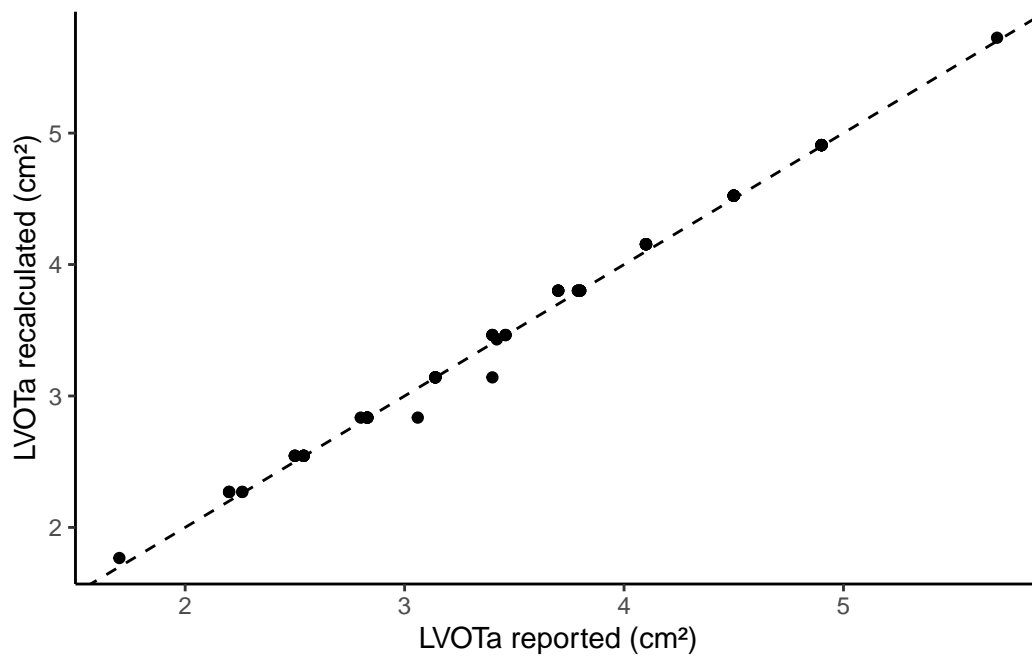
I will use the recalculated LVOTa values for subsequent analyses. However, I will provide plots of reported values versus recalculated values for the record.

Check if the recalculated LVOTa values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Area Left Ventricle Outflow	LVOTa	Numeric	1.7-6.0	cm/m²

min	max
1.77	5.73

Stroke Volume (SV)

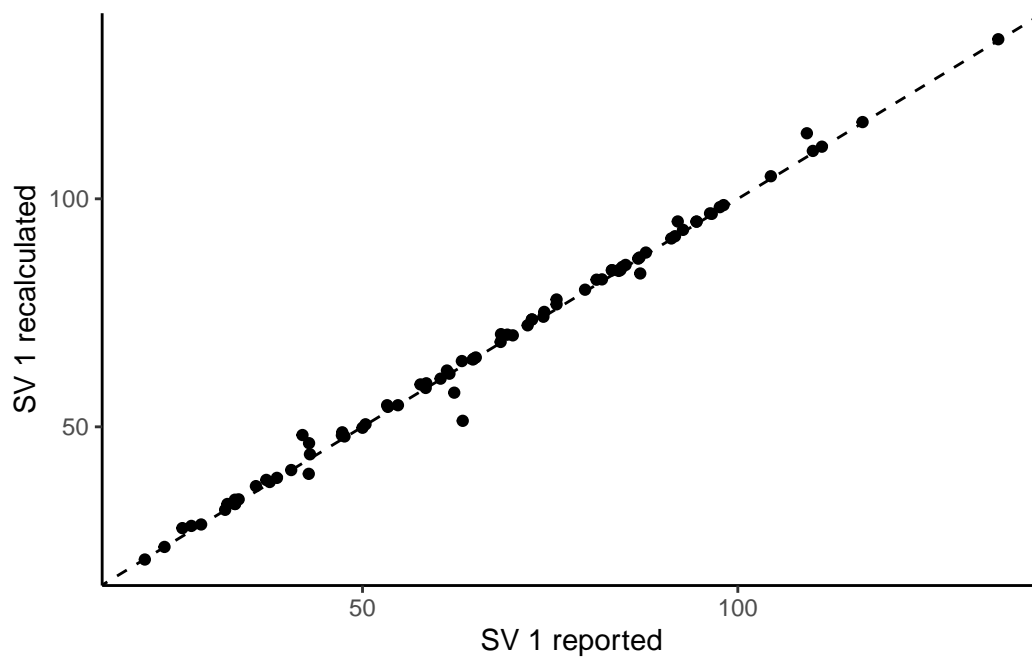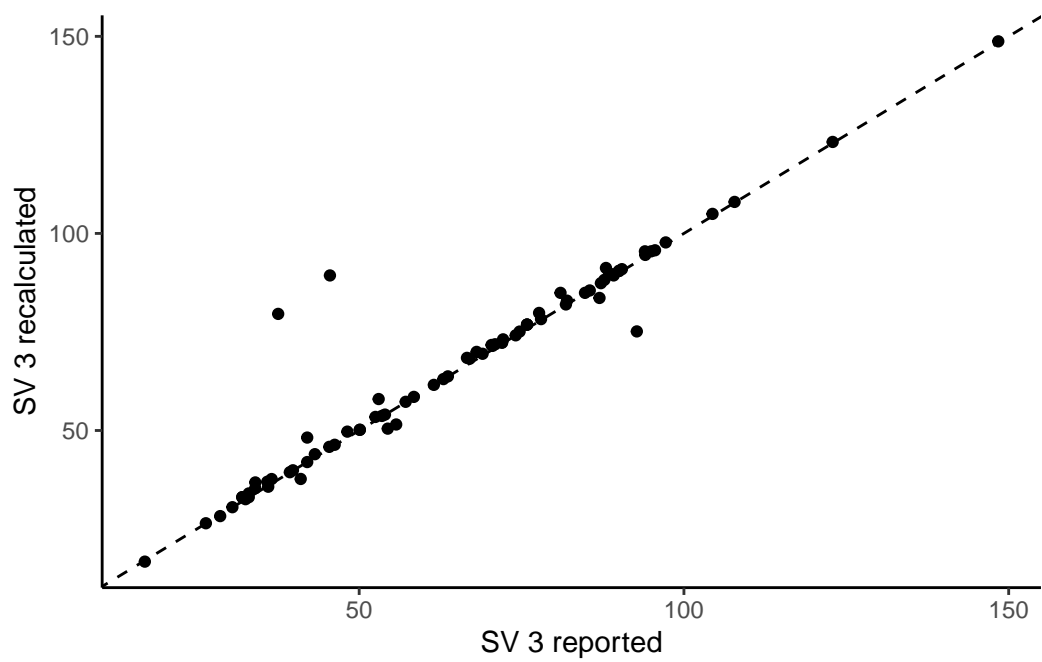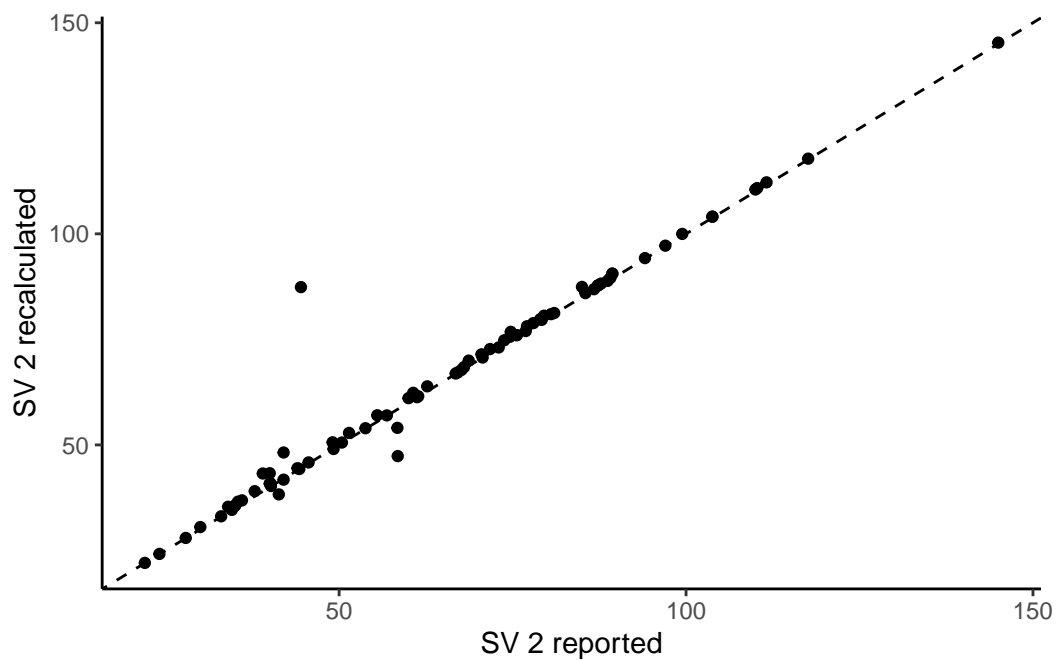The stroke volume (SV) is calculated using the formula:

$$SV = LVOTa \times VTI$$

where VTI is the velocity time integral.

Check if recalculated SV values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Systolic Volume (1)	SV_1	Numeric	10.0-150.0	ml

min	max
16.73	148.73

Cardiac Output (CO)

Cardiac output is calculated as:

$$CO = SV \times HR$$
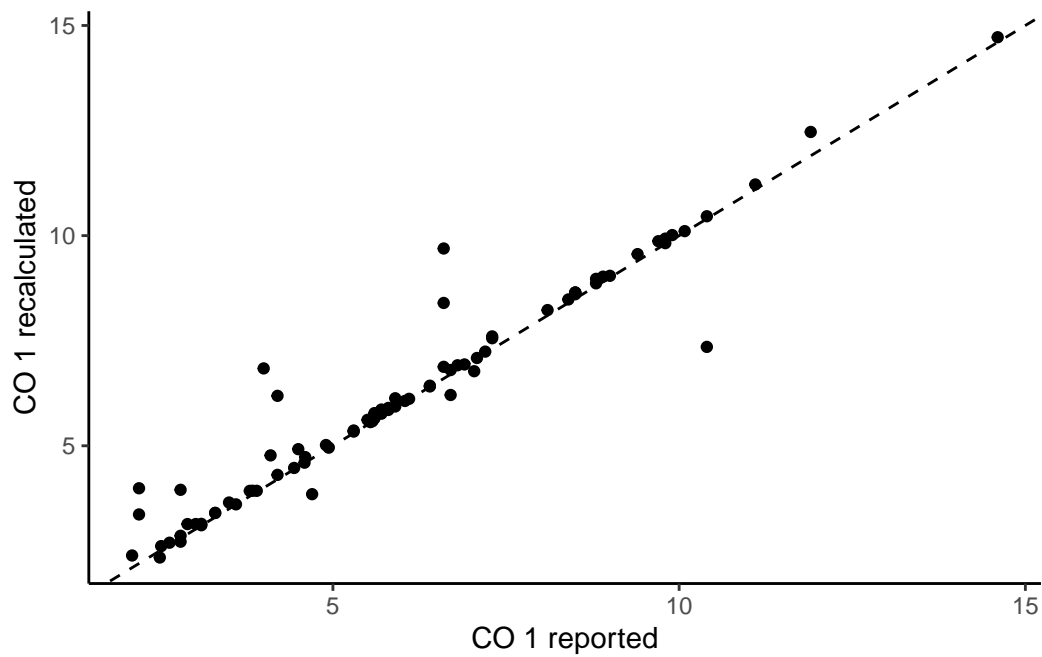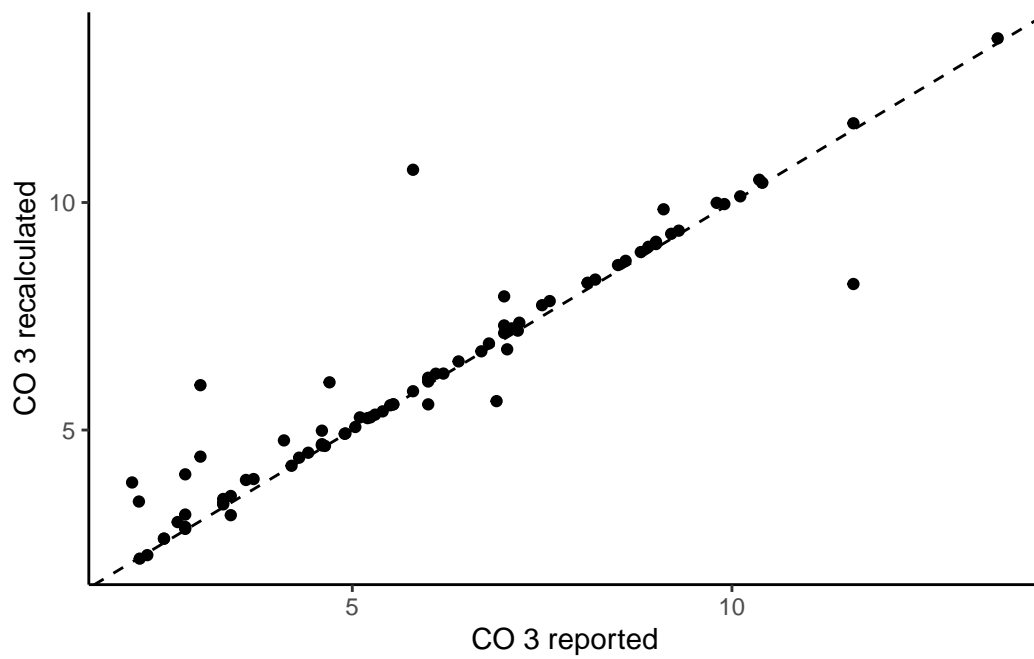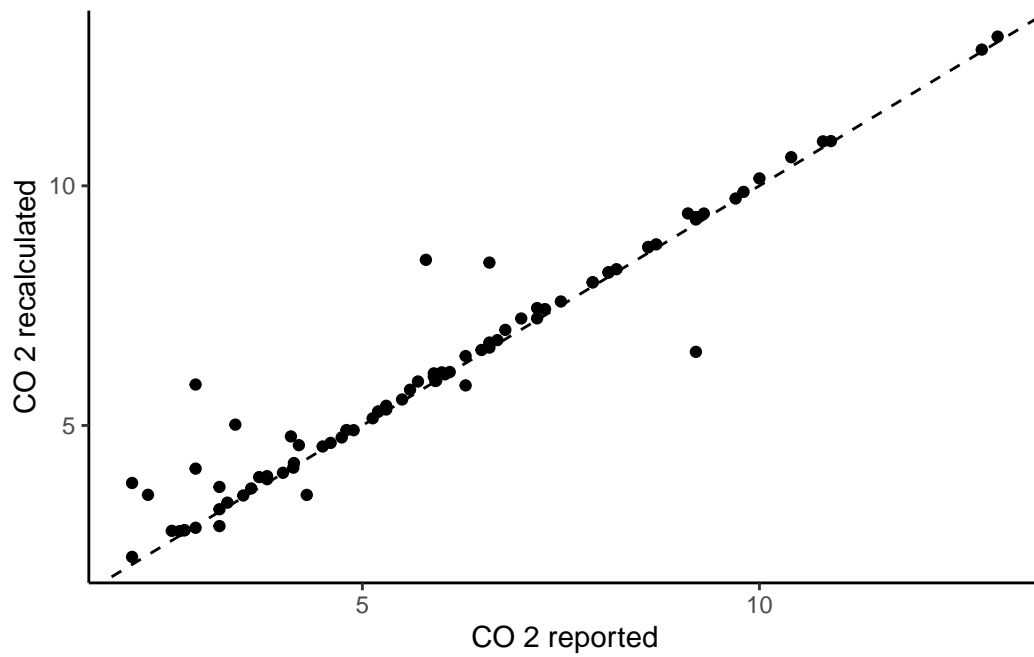
where HR is the heart rate.

Check if recalculated CO values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Cardiac output, TTE	cardiac_output_TTE	Numeric	2.0-35.0	L/min

min	max
2.17	14.72

The final value of cardiac output will be the average of the three measurements.
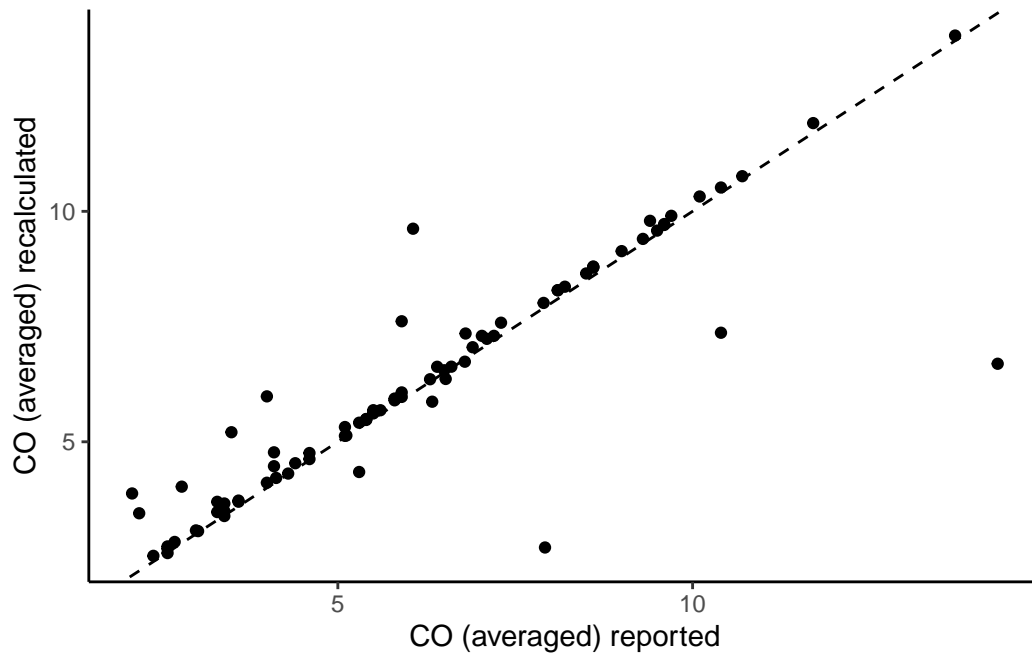
Cardiac index

Cardiac index is calculated as:

$$CI = CO/BSA$$

where BSA is the body surface area.

```
# Substitute values of LVOTa, SV, and CO with recalculated values
data <- data %>%
  mutate(
    LVOTa = LVOTa_recalculated,
    SV_1 = SV_1_recalculated,
    SV_2 = SV_2_recalculated,
    SV_3 = SV_3_recalculated,
    CO_TTE_1 = CO_1_recalculated,
    CO_TTE_2 = CO_2_recalculated,
    CO_TTE_3 = CO_3_recalculated,
    cardiac_output_TTE = CO_recalculated,
    cardiac_index_TTE = CO_recalculated / body_surface_area,
    CI_TTE_1 = CO_1_recalculated / body_surface_area,
    CI_TTE_2 = CO_2_recalculated / body_surface_area,
    CI_TTE_3 = CO_3_recalculated / body_surface_area
  ) %>%
  select(-contains("recalculated"))
```

Check if recalculated CI values is within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Cardiac index, TTE	cardiac_index_TTE	Numeric	1.0-9.0	L/min/m²

min	max
1.47	6.86

Fick method

Cardiac output with the Fick method is calculated as:

$$CO = \frac{(DavO_2 \times 100)/CaO_2}{DavO_2}$$

where: DavO2 = Arteriovenous oxygen difference, CaO2 = Arterial oxygen content.

Before recalculating CO, I will recalculate DavO2 and CaO2.

CaO2

Arterial oxygen content (CaO2) is calculated as:

$$CaO_2 = (Hb \times 1.34 \times SaO_2) + (0.0031 \times PaO_2)$$

where: Hb = Hemoglobin, SaO2 = arterial oxygen saturation, PaO2 = partial pressure of oxygen.

Check if recalculated CaO2 values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Arterial oxygen content	CaO2	Numeric	6.0-23.0	ml/dl

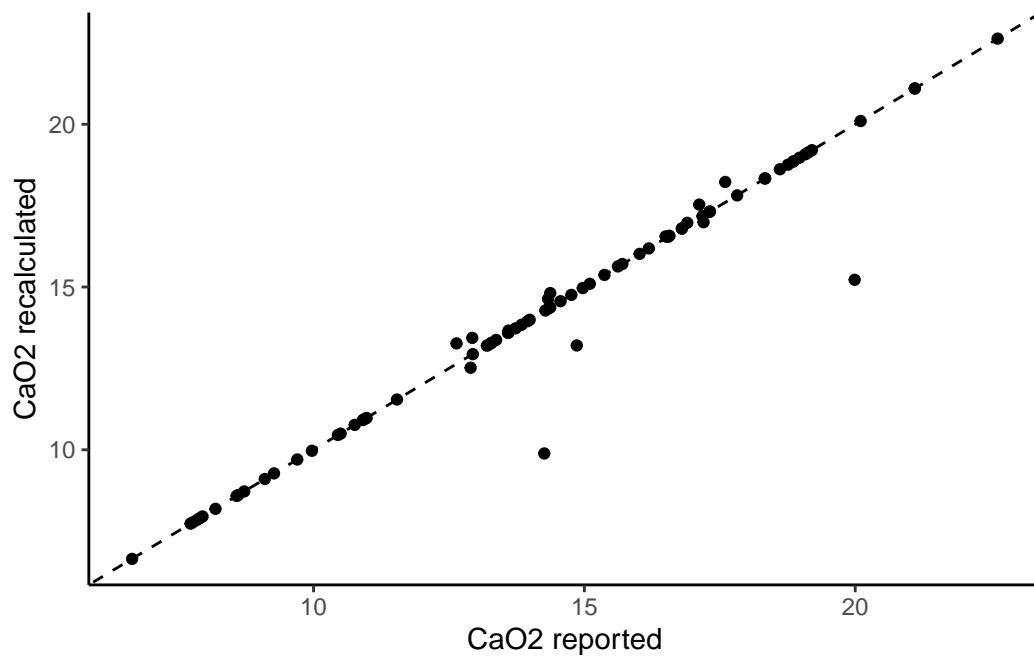| min	max
6.65	22.63

CvO2

Venous oxygen content (CvO2) is calculated as:

$$CvO_2 = (Hb \times 1.34 \times SvO_2) + (0.0031 \times PvO_2)$$

where:Hb = Hemoglobin, SvO2 = venous oxygen saturation, PvO2 = Partial pressure of oxygen in venous blood.

Check if recalculated CaO2 values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Venous oxygen content	CvO2	Numeric	4.0-20.0	ml/dl

min	max
4.06	17.49
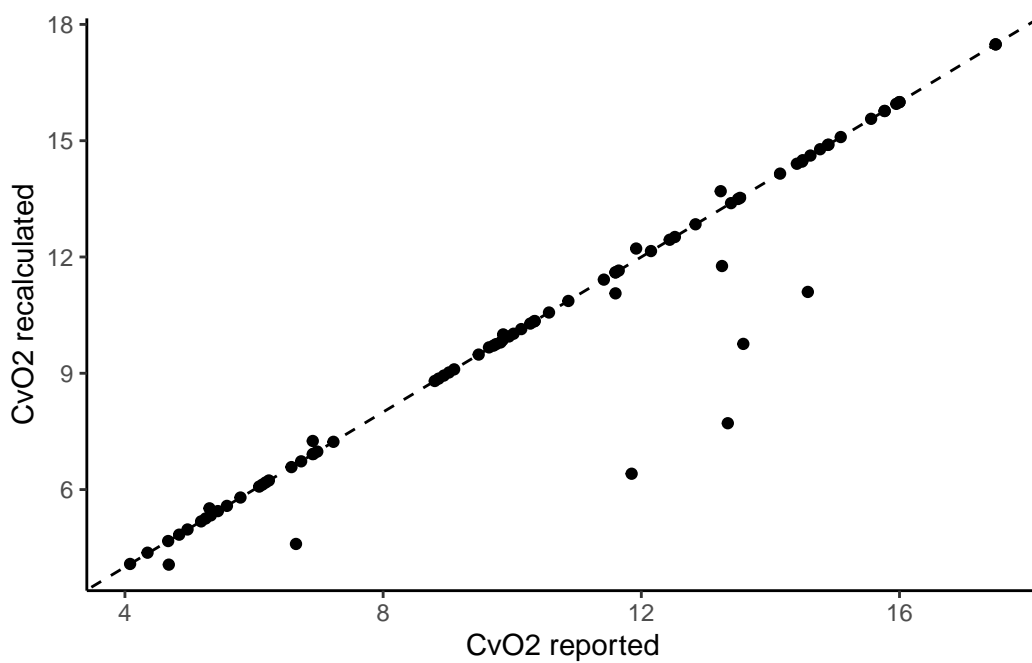
DavO2

Arteriovenous oxygen difference is calculated as:

$$DavO_2 = CaO_2 - CvO_2$$

Check if recalculated DavO2 values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Arteriovenous oxygen difference	DavO2	Numeric	0.5-12.0	ml/dl

min	max
0.85	11.35
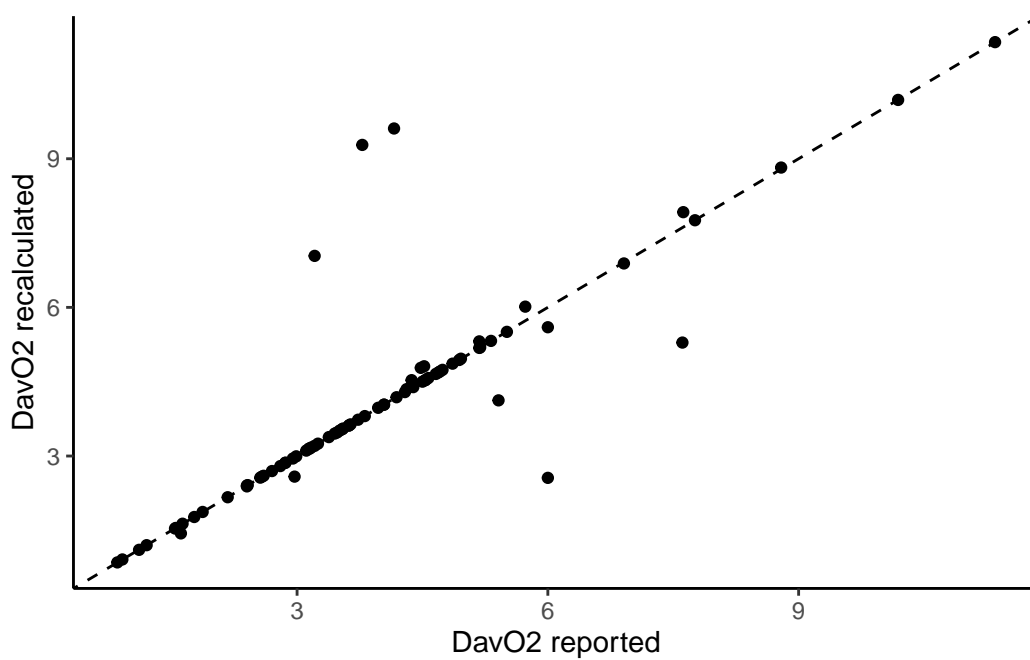
ExO2

The Oxygen extraction (ExO2) is calculated as:

$$ExO_2 = \frac{(SaO_2 - SvO_2)}{SaO_2}$$

Check if recalculated DavO2 values are within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Oxygen extraction	ExO2	Numeric	0.05-0.80	%

min	max
0.05	0.63
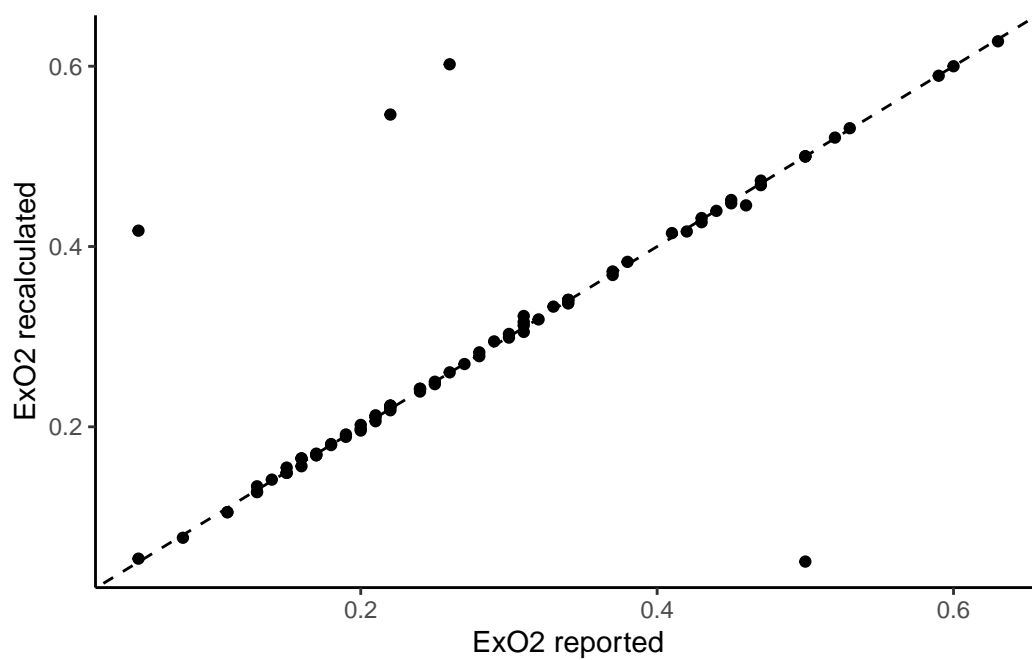
CO Fick

Check if recalculated CI values is within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Cardiac output, Fick	cardiac_output_Fick	Numeric	2.0-35.0	L/min

min	max
4.42	15.04

Cardiac index Fick

Cardiac index is calculated as:

$$CI = CO/BSA$$

where BSA is the body surface area.

```
# Substitute values of LVOTa, SV, and CO with recalculated values
data <- data %>%
  mutate(
    CaO2 = CaO2_recalculated,
    CvO2 = CvO2_recalculated,
    DavO2 = DavO2_recalculated,
    ExO2 = ExO2_recalculated,
    cardiac_output_Fick = CO_Fick_recalculated,
    cardiac_index_Fick = cardiac_output_Fick / body_surface_area
  ) %>%
  select(-contains("recalculated"))
```
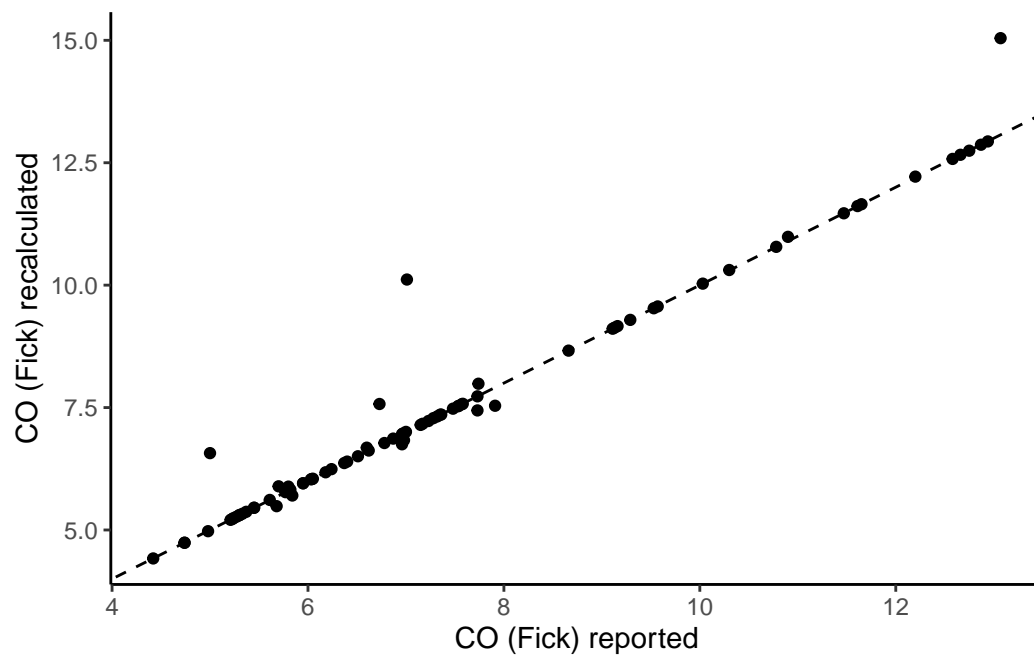
Check if recalculated CI values is within the range indicated in dictionary.

Label	Variable	Measurement_unit	Allowed values	Units
Cardiac index, Fick	cardiac_index_Fick	Numeric	1.0-9.0	L/min/m²

min	max
1.47	6.86

Diagnosis

diagnosis	n
Acute respiratory distress syndrome	11
Traumatic brain injury	10
Sepsis	9
Acute kidney injury	8
COVID-19	7
Shock	5
Community-acquired pneumonia	4
Pancreatitis	4
Ventilation-associated pneumonia	4
Diabetic ketoacidosis	3
Polytrauma	3
Postsurgical	3
Cardiac arrest	2
Urinary tract infection	2
Abdominal trauma	1
Acute myocardial infarction	1
Burns	1
Endocarditis	1
Hospital-acquired pneumonia	1
Hypertensive emergency	1
Influenza	1
Peritonitis	1
Pulmonary thromboembolism	1
Pyelonephritis	1
Soft-tissue infection	1
Strangulation	1
Tuberculous meningitis	1
Upper gastrointestinal bleeding	1

Comorbidities

comorbidities	n
Overweight	17
Obesity	16
Type 2 diabetes mellitus	11
NA	10
Systemic arterial hypertension	7
Alcoholism	3
Chronic obstructive pulmonary disease	3
Chronic kidney disease	2
Type 1 diabetes mellitus	2
Human immunodeficiency virus infection	1
Malnutrition	1
Parkinson's disease	1
Pulmonary arterial hypertension	1
Status epilepticus	1

Save data

The following sourced script assigns random values of ID for each participant. The script is hidden to preserve confidentiality and prevent re-identification of participants.

```r
source("scripts/random_ID_generator.R")
```

Save as R data file.

```r
save(data, file = file.path(psfolder,"cardiac_output_Fick_vs_ECOTT.RData"))
```

Save as sav file:

```r
write_sav(data, file.path(psfolder,"cardiac_output_Fick_vs_ECOTT.sav"))
```

Package References

- 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/.
- Iannone R, Cheng J, Schloerke B, Hughes E, Lauer A, Seo J, Brevoort K, Roy O (2024). *gt: Easily Create Presentation-Ready Display Tables.* R package version 0.11.0, https://CRAN.R-project.org/package=gt.
- 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/.
- Meyer C, Hammerschmidt D (2023). *overviewR: Easily Extracting Information About Your Data.* R package version 0.0.13, https://CRAN.R-project.org/package=overviewR.
- Müller K, Walthert L (2024). *styler: Non-Invasive Pretty Printing of R Code.* R package version 1.10.3, https://CRAN.R-project.org/package=styler.
- Müller K, Wickham H (2023). *tibble: Simple Data Frames.* R package version 3.2.1, https://CRAN.R-project.org/package=tibble.
- R Core Team (2024). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
- Rinker TW, Kurkiewicz D (2018). *pacman: Package Management for R.* version 0.5.0, http://github.com/trinker/pacman.
- Slowikowski K (2024). *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'.* R package version 0.9.6, https://CRAN.R-project.org/package=ggrepel.
- 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, Bryan J (2023). *readxl: Read Excel Files.* R package version 1.4.3, https://CRAN.R-project.org/package=readxl.
- 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, Miller E, Smith D (2023). *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files.* R package version 2.5.4, https://CRAN.R-project.org/package=haven.
- Wickham H, Vaughan D, Girlich M (2024). *tidyr: Tidy Messy Data.* R package version 1.3.1, https://CRAN.R-project.org/package=tidyr.