DVT Data Analysis (Final)

Anaqi Amir

2024-01-22

Contents

1 Introduction

In this report, I conducted a multitude of bivariate analysis that mostly calculates the odds ratios and incidence rates of the different relevant features in the dataset. Sections 1,2, and 3 all focus on this analysis and the summary of the results can be in the "Findings" sub-section of the respective sections. Then, in Section 4, I conducted an odds ratio test between the different types of thromboprophylaxis and the bleeding results of the patients to see if there is a significant relationship between the two. I then filtered the patients to satisfy the requirements pertaining the hypothesis before finally finding the incidence rates of the patients developing VTE within 30 days of chemotherapy induction of hematopoietic stem cell transplantation.

This project was then in collaboration with Dr. Federico Carini (University of Toronto) and Professor Sotirious Damouras (University of Toronto, Scarborough).

2 Setup

2.1 Imports

```
library(readxl)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0     v purrr   1.0.1
## v tibble  3.1.8     v dplyr   1.1.0
## v tidyr   1.3.0     v stringr 1.5.0
## v readr   2.1.3     v forcats 1.0.0
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(epitools)
library(naniar)
```

2.2 Load Data

```
# Reading the excel file
df = read_excel('/Users/Anaqi_Amir/Downloads/DVTRetrospective22_DataCollection_CariniCurrent_DVTHemOncD

# Head of df
df
```

```
## # A tibble: 862 x 94
##    study_id first~1  Male   age hem_m~2 year_~3 dxtime dx1y     mv o2    dialy~4
##       <dbl>   <dbl> <dbl> <dbl>   <dbl> <chr>   <chr>  <chr> <dbl> <chr> <chr>
##  1        1       1     0  76.6       2 2014    0      0         1 0     0
##  2        2       1     1  33.0       2 2012    2      1         1 0     0
##  3        3       1     1  67.1       2 2007    9      1         2 2     0
##  4        4       1     0  69.1       7 2018    0      0         1 0     0
##  5        5       1     1  79.1       2 2020    0      0         1 0     0
##  6        6       1     1  59.3       9 2017    0      0         1 0     1
##  7        7       1     0  72.4       6 2015    0      0         1 1     1
##  8        8       1     1  64.0       2 2018    1      1         0 2     0
##  9        9       1     0  75.6       2 2021    0      0         0 3     0
## 10       10       1     0  67.4       2 2015    0      0         0 0     0
## # ... with 852 more rows, 83 more variables: pressors <chr>,
```

```
## #   hospital_admission_date <dttm>, icu_admission_date <dttm>,
## #   icu_discharge_date <dttm>, icu_disposition <dbl>,
## #   hospital_discharge_date <chr>, hospital_disposition <chr>, dnr_icu <dbl>,
## #   dnr_ward <chr>, pre_icu_los <dbl>, icu_los <dbl>, hospital_los <chr>,
## #   hospital_ad_dx <dbl>, icu_diag <dbl>, weight <chr>, height <chr>,
## #   bmi <chr>, covid <dbl>, cmbd_htn <dbl>, cmbd_cad <dbl>, cmbd_chf <dbl>, ...
```

3 Basic data exploration

3.1 Overview

```
# Shape of data
dim(df)
```

```
## [1] 862  94
```

```
# Summary of data
summary(df)
```

```
##     study_id      first_icu_ad         Male             age
##  Min.   :  1.0   Min.   :0.0000   Min.   :0.0000   Min.   :19.08
##  1st Qu.:216.2   1st Qu.:1.0000   1st Qu.:0.0000   1st Qu.:50.24
##  Median :431.5   Median :1.0000   Median :1.0000   Median :61.49
##  Mean   :431.8   Mean   :0.9466   Mean   :0.5882   Mean   :57.86
##  3rd Qu.:645.8   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:68.76
##  Max.   :864.0   Max.   :1.0000   Max.   :1.0000   Max.   :94.13
##    hem_malig     year_diagnosis        dxtime              dx1y
##  Min.   :1.000   Length:862         Length:862         Length:862
##  1st Qu.:2.000   Class :character   Class :character   Class :character
##  Median :2.000   Mode  :character   Mode  :character   Mode  :character
##  Mean   :3.423
##  3rd Qu.:6.000
##  Max.   :9.000
##        mv              o2              dialysis           pressors
##  Min.   :0.0000   Length:862         Length:862         Length:862
##  1st Qu.:0.0000   Class :character   Class :character   Class :character
##  Median :1.0000   Mode  :character   Mode  :character   Mode  :character
##  Mean   :0.6183
##  3rd Qu.:1.0000
##  Max.   :2.0000
##  hospital_admission_date          icu_admission_date
##  Min.   :2013-12-10 00:00:00.00   Min.   :2014-01-02 00:00:00.00
##  1st Qu.:2015-11-15 06:00:00.00   1st Qu.:2015-11-18 00:00:00.00
##  Median :2017-12-14 00:00:00.00   Median :2018-01-04 12:00:00.00
##  Mean   :2017-12-09 06:37:35.22   Mean   :2017-12-23 04:05:34.10
##  3rd Qu.:2020-01-06 06:00:00.00   3rd Qu.:2020-01-19 12:00:00.00
##  Max.   :2022-01-24 00:00:00.00   Max.   :2022-01-26 00:00:00.00
##  icu_discharge_date               icu_disposition hospital_discharge_date
##  Min.   :2014-01-08 00:00:00.00   Min.   :1.000   Length:862
##  1st Qu.:2015-11-23 06:00:00.00   1st Qu.:1.000   Class :character
##  Median :2018-01-13 00:00:00.00   Median :2.000   Mode  :character
##  Mean   :2017-12-29 00:03:20.46   Mean   :1.697
##  3rd Qu.:2020-01-22 18:00:00.00   3rd Qu.:2.000
##  Max.   :2022-01-28 00:00:00.00   Max.   :4.000
```

```
##  hospital_disposition    dnr_icu        dnr_ward          pre_icu_los
##  Length:862           Min.   :0.000   Length:862         Min.   :  0.00
##  Class :character     1st Qu.:0.000   Class :character   1st Qu.:  1.00
##  Mode  :character     Median :0.000   Mode  :character   Median :  7.00
##                       Mean   :0.326                      Mean   : 13.89
##                       3rd Qu.:1.000                      3rd Qu.: 19.00
##                       Max.   :1.000                      Max.   :408.00
##     icu_los        hospital_los       hospital_ad_dx     icu_diag
##  Min.   :  0.000   Length:862         Min.   :1.000   Min.   : 1.000
##  1st Qu.:  1.000   Class :character   1st Qu.:4.000   1st Qu.: 1.000
##  Median :  3.000   Mode  :character   Median :4.000   Median : 2.000
##  Mean   :  5.832                      Mean   :3.847   Mean   : 2.869
##  3rd Qu.:  6.000                      3rd Qu.:4.000   3rd Qu.: 4.000
##  Max.   :103.000                      Max.   :9.000   Max.   :11.000
##     weight             height             bmi               covid
##  Length:862         Length:862         Length:862         Min.   :0.00000
##  Class :character   Class :character   Class :character   1st Qu.:0.00000
##  Mode  :character   Mode  :character   Mode  :character   Median :0.00000
##                                                           Mean   :0.00464
##                                                           3rd Qu.:0.00000
##                                                           Max.   :1.00000
##     cmbd_htn         cmbd_cad          cmbd_chf          cmbd_dm
##  Min.   :0.0000   Min.   :0.00000   Min.   :0.00000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.0000
##  Median :0.0000   Median :0.00000   Median :0.00000   Median :0.0000
##  Mean   :0.3898   Mean   :0.08237   Mean   :0.09281   Mean   :0.1821
##  3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.00000   Max.   :1.00000   Max.   :2.0000
##    cmbd_copd       cmbd_malignancynonHem   cmbd_smok        cmbd_CKD
##  Min.   :0.00000   Min.   :0.00000       Min.   :0.000   Min.   :0.00000
##  1st Qu.:0.00000   1st Qu.:0.00000       1st Qu.:0.000   1st Qu.:0.00000
##  Median :0.00000   Median :0.00000       Median :0.000   Median :0.00000
##  Mean   :0.06265   Mean   :0.09629       Mean   :0.123   Mean   :0.05104
##  3rd Qu.:0.00000   3rd Qu.:0.00000       3rd Qu.:0.000   3rd Qu.:0.00000
##  Max.   :1.00000   Max.   :1.00000       Max.   :1.000   Max.   :1.00000
##  cmbd_cirrhosis      cmbd_AF          cmbd_etohl        cmbd_ob
##  Min.   :0.0000   Min.   :0.00000   Min.   :0.00000   Length:862
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000   Class :character
##  Median :0.0000   Median :0.00000   Median :0.00000   Mode  :character
##  Mean   :0.0116   Mean   :0.08817   Mean   :0.03016
##  3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0.00000
##  Max.   :1.0000   Max.   :1.00000   Max.   :1.00000
##      ecog              ecog2              sofa1              sofa3
##  Length:862         Length:862         Length:862         Length:862
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##     sofa7              apache2          pro_dvt_chemo      induction_date
##  Length:862         Length:862         Length:862         Length:862
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
```

```
##
##
##  induction_name         sct              sct_date         conditioning_date
##  Length:862         Length:862         Length:862         Length:862
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##  conditioning_name      gvhd             risk_fx           cvc_preicu
##  Length:862         Length:862         Length:862         Length:862
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##    cvc_icu           dvt_preicu         pe_preicu         dvtpe_preicu
##  Length:862         Length:862         Length:862         Length:862
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##     dvt__icu           pe_icu         dvtpe_icu        plt_onndvt
##  Min.   :0.00000   Min.   :0.0000   Min.   :0.00000   Length:862
##  1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.00000   Class :character
##  Median :0.00000   Median :0.0000   Median :0.00000   Mode  :character
##  Mean   :0.06497   Mean   :0.0174   Mean   :0.05104
##  3rd Qu.:0.00000   3rd Qu.:0.0000   3rd Qu.:0.00000
##  Max.   :3.00000   Max.   :1.0000   Max.   :1.00000
##    US_date            cta_date         timeicutousicu     timeicutocticu
##  Length:862         Length:862         Length:862         Length:862
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##     low_plt          vlow_plt       lowest_plt            low_hb
##  Min.   :0.0000   Min.   :0.0000   Length:862         Min.   :  0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   Class :character   1st Qu.:  0.0000
##  Median :0.0000   Median :1.0000   Mode  :character   Median :  1.0000
##  Mean   :0.2309   Mean   :0.5754                      Mean   :  0.6914
##  3rd Qu.:0.0000   3rd Qu.:1.0000                      3rd Qu.:  1.0000
##  Max.   :1.0000   Max.   :1.0000                      Max.   :100.0000
##     high_wbc         low_wbc       dvt_proph_preicu    ivc_filter
##  Min.   :0.0000   Min.   :0.0000   Length:862         Length:862
##  1st Qu.:0.0000   1st Qu.:0.0000   Class :character   Class :character
##  Median :0.0000   Median :1.0000   Mode  :character   Mode  :character
##  Mean   :0.1845   Mean   :0.5232
##  3rd Qu.:0.0000   3rd Qu.:1.0000
##  Max.   :9.0000   Max.   :1.0000
##  noproph_reason_preICU dvt.icu.proph      icuproph      dvt.icuproph_noreason
##  Length:862            Min.   :0.000   Min.   :0.0000   Min.   :0.000
##  Class :character      1st Qu.:1.000   1st Qu.:0.0000   1st Qu.:1.000
```

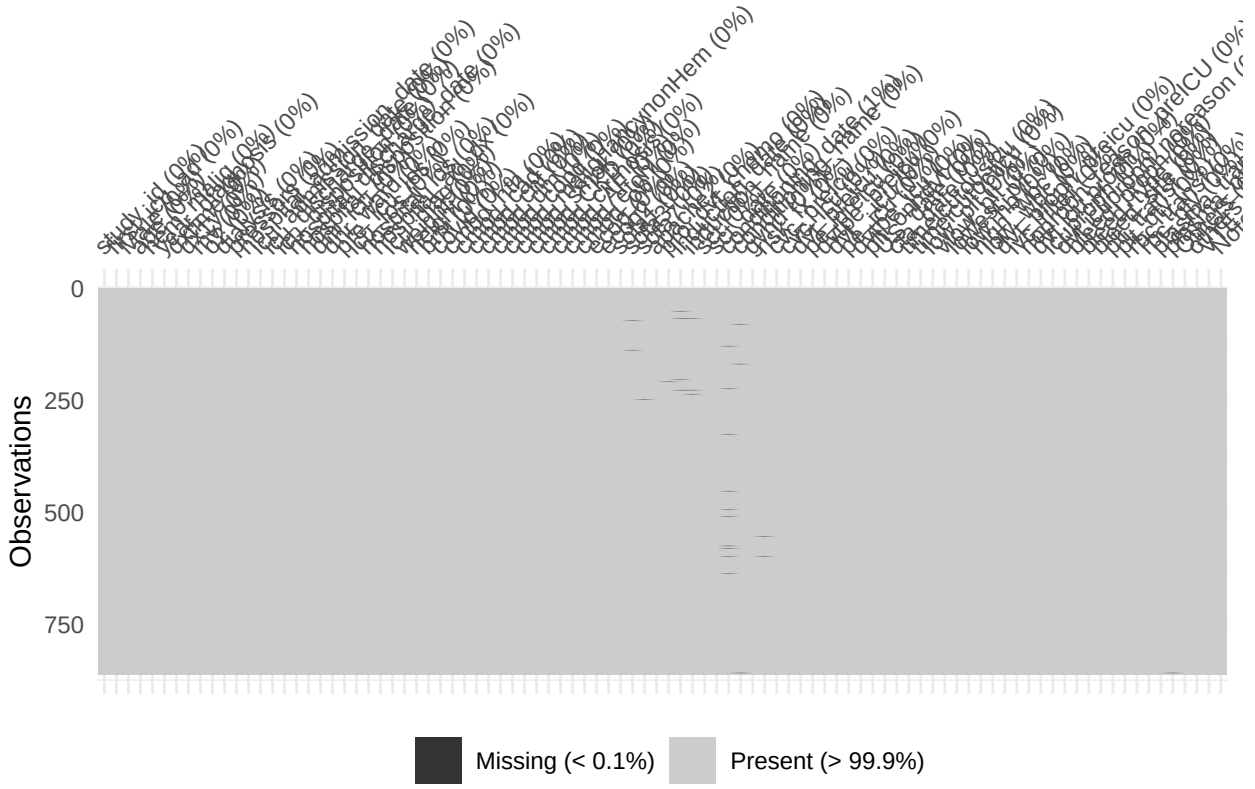
```
##  Mode  :character      Median :4.000   Median :0.0000   Median :1.000
##                       Mean   :2.858   Mean   :0.2227   Mean   :1.387
##                       3rd Qu.:4.000   3rd Qu.:0.0000   3rd Qu.:1.000
##                       Max.   :6.000   Max.   :1.0000   Max.   :5.000
##   bleed_proph        bleed_type      bleed_site       plt_transf
##  Min.   :0.00000   Min.   :0.000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.00000   1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.00000   Median :0.000   Median :0.0000   Median :1.0000
##  Mean   :0.01044   Mean   :0.355   Mean   :0.9606   Mean   :0.5986
##  3rd Qu.:0.00000   3rd Qu.:0.000   3rd Qu.:0.0000   3rd Qu.:1.0000
##  Max.   :1.00000   Max.   :2.000   Max.   :8.0000   Max.   :1.0000
##    plt_qty           rbc_transf       rbc_qty          plasma_transfusion
##  Length:862         Min.   :0.0000   Length:862         Min.   :0.0000
##  Class :character   1st Qu.:0.0000   Class :character   1st Qu.:0.0000
##  Mode  :character   Median :1.0000   Mode  :character   Median :0.0000
##                     Mean   :0.6752                      Mean   :0.1415
##                     3rd Qu.:1.0000                      3rd Qu.:0.0000
##                     Max.   :1.0000                      Max.   :1.0000
##   plasma_qty          Others           others_qty         VTE_compl
##  Length:862         Length:862         Length:862         Length:862
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##    Notes
##  Length:862
##  Class :character
##  Mode  :character
##
##
##
```

3.2 Missing values

The presence of missing values significantly influences the integrity of data analysis, potentially resulting in distorted findings. Accordingly, it is imperative to meticulously examine and address any missing values in the dataset prior to initiating the analysis, ensuring a proactive approach to managing such instances during the subsequent analytical procedures.

I will be using the `vis_miss` function from the `naniar` package to visualize missing data. Note that each dark line represent a missing value in its corresponding column.

```
vis_miss(df)
```



As we can see, there are minimal missing values in the dataset. Even though this is the case, I have decided to not remove these rows that contain missing values as the incidence rates of the factors that we would be observing are extremely low; therefore, removing rows might greatly affect the interpretation and analysis of our results in the future.

A modification that can be made to my report would be to fill in these missing values through the most appropriate technique (fill with mean, fill with a regression model, etc.), but I would not do that in this report.

4 Data Analysis

This section is divided into 3 sub-sections as requested by Dr. Carini. The purposes of each sub-section is simply to investigate the relationshop between various factors and how they affect one another. All analysis in this report are strictly bivariate, i.e., this report only looks at the relationship between two factors at a time.

4.1 Section 1: Prophylaxis

In this section, I looked at the relationship between various factors in a patient's health, and the presence of prophlyaxis in those patients. Here are the results:

```
prophylaxis_table = read.table("Prophylaxis analysis")
knitr::kable(prophylaxis_table, "simple")
```

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
Biologically Male vs SCDs	507	0.44	1.20 (0.91 - 1.58)	0.20
Biologically Male vs Pharmacological	507	0.19	0.67 (0.48 - 0.92)	0.01*
Biologically Male vs No Proph	507	0.20	0.94 (0.67 - 1.31)	0.70
Biologically Female vs SCDs	355	0.40	0.84 (0.63 - 1.10)	0.20
Biologically Female vs Pharmacological	355	0.26	1.50 (1.09 - 2.08)	0.01*

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
Biologically Female vs No Proph	355	0.21	1.09 (0.76 - 1.49)	0.70
Plt<=20 vs SCDs	496	0.54	3.23 (2.42 - 4.34)	4.44e-16*
Plt<=20 vs Pharmacological	496	0.09	0.15 (0.10 - 0.21)	2.13e-27 (Chi Square)*
Plt<=20 vs No Proph	496	0.23	1.30 (0.93 - 1.83)	0.13
20<Plt<50 vs SCDs	199	0.37	0.73 (0.52 - 1.01)	0.06
20<Plt<50 vs Pharamcological	199	0.26	1.32 (0.91 - 1.90)	0.14
20<Plt<50 vs No Proph	199	0.22	1.07 (0.72 - 1.57)	0.73
Low HB vs SCDs	457	0.49	1.71 (1.30 - 2.25)	1.25e-4 (Chi Square)*
Low HB vs Pharmacological	457	0.14	0.33 (0.24 - 0.47)	6.74e-11*
Low HB vs No Proph	457	0.22	1.22 (0.87 - 1.70)	0.25
High WBC vs SCDs	132	0.41	0.92 (0.63 - 1.34)	0.67
High WBC vs Pharmacological	132	0.23	1.08 (0.69 - 1.66)	0.73
High WBC vs No Proph	132	0.15	0.65 (0.38 - 1.05)	0.08
Low WBC vs SCDs	451	0.50	1.88 (1.43 - 2.48)	5.33e-6*
Low WBC vs Pharmacological	451	0.13	0.31 (0.22 - 0.43)	2.69e-12*
Low WBC vs No Proph	451	0.22	1.20 (0.86 - 1.67)	0.29
Pre-ICU proph vs SCDs	83	0.24	0.40 (0.23 - 0.66)	2.58e-4*
Pre-ICU proph vs Pharmacological	83	0.55	5.37 (3.36 - 8.64)	4.02e-12*
Pre-ICU proph vs No proph	83	0.14	0.63 (0.32 - 1.14)	0.13
> 61 years old vs SCDs	442	0.42	0.96 (0.73 - 1.26)	0.76
> 61 years old vs Pharmacological	442	0.24	1.22 (0.89 - 1.69)	0.22
> 61 years old vs No Proph	442	0.18	0.74 (0.53 - 1.03)	0.07

We can see that there are so surprising significant factors that determines the presence of prophylaxis in patients.

Next, we look at the effectiveness of prophylaxis on patients developing VTE (either DVT or PE). We would assume that the use of prophylaxis should lower the patients'odds of developing VTE. The table below also compares patiets' thrombocytopenia severity against the development of VTE. We would assume that the more severe the thrombocytopenia, the higher the odds of them developing VTE.

```
prophylaxis_effectiveness_table = read.table("Prophylaxis effectiveness")
knitr::kable(prophylaxis_effectiveness_table, "simple")
```

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
SCDs vs DVT	367	0.04	0.88 (0.42 - 1.78)	0.72
SCDs vs PE	367	0.01	0.34 (0.07 - 1.11)	0.08
Pharma Proph vs DVT	192	0.06	2.07 (0.96 - 4.23)	0.06
Pharma Proph vs PE	192	0.03	2.39 (0.77 - 6.81)	0.12
No prophylaxis vs DVT	179	0.02	0.53 (0.15 - 1.38)	0.21
No prophylaxis vs PE	179	0.03	2.62 (0.85 - 7.47)	0.09
Plt>=50 vs DVT	167	0.05	1.36 (0.56 - 2.97)	0.47
Plt>=50 vs PE	167	0.02	1.08 (0.23 - 3.50)	0.91
20<Plt<50 vs DVT	199	0.04	1.08 (0.45 - 2.35)	0.85
20<Plt<50 vs PE	199	0.03	1.71 (0.51 - 4.94)	0.36
Plt<=20 vs DVT	496	0.03	0.78 (0.38 - 1.58)	0.48
Plt<=20 vs PE	496	0.01	0.64 (0.22 - 1.84)	0.40

The table above has shown that none of these odds are significant, but we can still use it as a rough reference to get an idea on what is going on. For example, it is interesting to see that patients with pharmacological prophylaxis are at a higher odds of developing VTE compared to those on mechanical prophylaxis. This

is expected. However, the report will later show that those with mechanical prophylaxis are more likely to develop bleeding problem compared to their counterparts. So, there is a tradeoff that exists between these two facts.

Another thing that can be seen in the table is that those with sever thrombocytopenia (platletes<20) have lower odds of developing VTE compared to those with a higher platelet count. This is surprising. Though it can be seen that the p-values for these observation are extremely high, so the results cannot be deem significant.

4.2 Section 2: Catheters

In this section, I looked at the relationship between various factors in a patient's health, and the presence of catheters in those patients. Here are the results:

```
catheter_vte_table = read.table("Catheter vs VTE")
knitr::kable(catheter_vte_table, "simple")
```

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
Catheter vs DVT	251	0.03	1.42 (0.66 - 2.90)	0.36
Catheter vs Catheter-related DVT	10	0.50	1.00 (0.27 - 3.73)	1.00
Catheter vs Non catheter-related DVT	14	0.50	1.00 (0.33 - 3.00)	1.00
Catheter vs PE	251	0.02	1.65 (0.54 - 4.70)	0.36
Catheter vs DVT+PE	251	0.07	1.74 (0.92 - 3.23)	0.09

```
catheter_death_table = read.table("Catheter vs Death")
knitr::kable(catheter_death_table, "simple")
```

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
Catheter related DVT vs Death	12	0.08	0.19 (0.01 - 1.01)	0.05
Non-catheter related DVT vs Death	19	0.26	0.69 (0.22 - 1.86)	0.48

Both tables above have shown that the presence of catheters have no significant impact on the development of VTE and a patient's death. The high p-values in these analyses is primarily due to the fact that our sample size is very small.

4.3 Section 3: Others

In this section, I looked at a few others relationships between factors that I think would be of value.

```
others_table = read.table("Other findings")
knitr::kable(others_table, "simple")
```

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
Major bleeding vs Death	122	0.41	1.40 (0.94 - 2.07)	0.09
SCDs vs Death	367	0.35	1.05 (0.79 - 1.40)	0.73
DVT vs Death	33	0.18	0.42 (0.15 - 0.98)	0.04*
PE vs Death	15	0.40	1.30 (0.42 - 3.69)	0.63
DVT vs Major bleeding	33	0.15	1.11 (0.37 - 2.73)	0.83

The table shows that the only signicant relationship is between DVT and death where patients with DVT have a 58% decrease in odds to experience death. This is obviously surprising as one would expect that those

with the illness are at a higher rate of experiencing death compared to those without. The reason for this lower odds may be due to the fact that those with DVT are simply easier to recover from the illness and that death may come through other factors instead. More research and analysis is needed to determine the cause of this.

5 Hypothesis

In this section, I will be investigating the hypothesis proposed by Dr. Carini through 3 consecutive steps as follows:

Hypothesis:
In critically ill patients with hematologic malignancy (HM) who are thrombocytopenic, venous thromboembolism (VTE) is infrequent within the first 30 days following induction chemotherapy or hematopoietic cell transplant (HCT), rendering it possible to avoid the use of thromboprophylaxis (mechanical or pharmacological) and the associated risks (i.e. serious bleeding).

Step 1: Find the correlation between thromboprophylaxis and bleeding. This is to show that thromboprophylaxis does indeed lead to the patients developing bleeding symptoms.

Step 2: Filter the data to find the target demographic of critically ill patients with HM who are thrombocytopenic that either underwent chemo or HCT.

Step 3: Find the incidence rate of VTE of the patients from the filtered data within the first 30 days of the induction of either chemo or HCT.

5.1 Step 1

I will calculate the odd ratios of the following combinations:
- Mechanical prophylaxis (SCDs, Compression Stockings) vs Major Bleeding
- Mechanical prophylaxis (SCDs, Compression Stockings) vs Minor Bleeding
- Pharmological prophylaxis (LMWH, Unfractioned heparin, Oral anticoagulants) vs Major Bleeding
- Pharmological prophylaxis (LMWH, Unfractioned heparin, Oral anticoagulants) vs Minor Bleeding

```
step1_table = read.table("Step 1")
knitr::kable(step1_table, "simple")
```

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
Mechanical vs Major Bleeding	367	0.16	1.21 (0.83 - 1.79)	0.32
Pharma Proph vs Major bleeding	196	0.05	0.27 (0.13 - 0.50)	7.73e-06*
Mechanical vs Minor Bleeding	62	0.52	1.48 (0.88 - 2.50)	0.14
Pharma Proph vs Minor bleeding	39	0.23	0.42 (0.19 - 0.87)	0.02*

As we can see, if thromboprophylaxis were to be implemented, it is better to go with the pharmacological route rather that the mechanical route. This is because those who underwent pharmacological treatment display a significant decrease in developing serious bleeding while the opposite holds true for those who underwent the mechanical thromboprophylaxis route by either using SCDs or compression stocking.

5.2 Step 2

Filter data to find the intended demographic. We mange to reduce the rows from 862 to only 695.

5.3 Step 3

Finding the incidence rates. Here are the results:

```
step3_table = read.table("Step 3")
knitr::kable(step3_table, "simple")
```

	n	Incidence.Rate	Odds.Ratio..95..CI.	p.value
SCDs vs VTE within 30 days	171	0.92	0.57 (0.17 - 1.58)	0.29
Pharma Proph vs VTE within 30 days	65	0.98	4.37 (0.88 - 105.99)	0.08

6 Conclusion

This report has conducted a multitude of bivariate analyses to see how factors interact with one another. Mainly, the report was conducted with these 3 aims in mind:

1. Aim 1: Determine the use, type (mechanical or pharmacological) and timing of thromboprophylaxis.
2. Aim 2: Determine incidence of catheter and non catheter-related VTE (upper / lower extremity DVT and PE) and explore variables that are independently associated with it.
3. Aim 3: Determine incidence of bleeding, complications associated with SCDs, and complications associated with VTE, including death. Bleeding severity will be classified according to current accepted standards.

Alongside these 3 aims, the hypothesis laid out in Section 5 was also investigated as well.

Here are the conclusions obtained for each aim, as well as the hypothesis. Note that these conclusions are purely analytical, and lack the rigorous statistical material to be used as a final result in a proper research.

6.1 Aim 1

This aim was investigated in Section 4.1 and here are the significant results:

- Biologically male patients have a 33% decrease in odds to receive pharmacological prophylaxis.
- Biologically female patients have a 50% increase in odds to receive pharmacological prophylaxis.
- Patients with severe thrombocytopenia have a 232% increase in odds to receive SCDS.
- Patients with severe thrombocytopenia have a 85% decrease in odds to receive pharmacological prophylaxis.
- Patients with low hemoglobin levels (<70) have a 71% increase in odds to receive SCDs.
- Patients with low hemoglobin levels (<70) have a 67% decrease in odds to receive pharmacological prophylaxis.
- Patients with low white blood cells (<4) have a 88% increase in odds to receive SCDs.
- Patients with low white blood cells (<4) have a 87% decrease in odds to receive pharmacological prophylaxis.
- Patients who have received prophylaxis pre-ICU have a 60% decrease in odds to receive SCDs.
- Patients who have received prophylaxis pre-ICU have a 437% increase in odds to receiven pharmacological prophylaxis.

Note: I'm unsure on how the data was collected regarding pre-ICU and during ICU records of patients receiving prophylaxis. There is a chance that the prophylaxes were counted twice as the prophylaxis received during ICU is a mere continuation of the prophylaxis given pre-ICU. The converse might also be true.

Furthermore, these are the findings of how prophylaxis affects the development of VTE.

- Patients with SCDs have a 12% decrease in odds of developing DVT.
- Patients with SCDs have a 66% decrease in odds of developing PE.
- Patients with pharmacological prophlyaxis have a 201% increase in odds of developing DVT.
- Patients with pharmacological prophlyaxis have a 232% increase in odds of developing PE.

Therefore, we can see that those with mechanical prophylaxis are less likely to develop VTE compared to those on pharmacological prophylaxis.

However, as mentioned before, this comes at the expense of an increase in odds of major bleeding as:

- Patients with SCDS have a 21% increase in odds of developing major bleeding.
- Patients with pharmacological prophylaxis have a 74% decrease in odds od developing major bleeding.

So, there is a tradeoff between developing VTE and experience major bleeding in patients. This forces medical practitioners (with discussion of the patients) to consider whether mechanical prophylaxis is worth the risk for patients in attempts to prevent VTE.

6.2 Aim 2

Section 4.2 shows that the presence of catheters do not have affect the development of VTEs in patients. Though keep in mind that none of the values observed in Section 4.2 were significant as the sample size were too small. Therefore, it is worth conducting more research with a larger sample size to investigate this aim.

6.3 Aim 3

Section 4.3 explores this aim in a minimal setting. I investigated the relationship of 3 things:

1. VTE vs Death
2. Bleeding vs Death
3. VTE (DVT specifically) vs Death

From the table in Section 4.3, it is seen that the only significant result is that patients with DVT have lower odds of experiencing death compared to patients without DVT.

Other relationships that were explored are not significant and therefore, it is hard to access the results of this aim without a larger sample size in future research.

6.4 Hypothesis

For patients with HM and thrombocytopenia, here are the incidence rates for VTE for those who either received chemotherapy or stem-cell transplantation (SCT):

- 0.82 within first 30 days of chemotherapy induction
- 0.72 within first 30 days of SCT.
- 0.79 within first 30 days for either.
- Patients with SCDs have a 43% decrease in odds of developing VTE within 30 days of chemotherapy induction or SCT.
- Patients with LMWH (and others) have a 437% increase in odds of developing VTE within 30 days of chemotherapy induction or SCT.

Note:

- These numbers do **not** account for overlaps between chemotherapy and SCT.
- I did not conduct odds ratio on chemo or SCT vs VTE because I'm unsure how to filter those that did not receive chemo or SCT

7 Appendix

7.1 Description of variables in odds ratio tables

- `n`: The total number of patients that has both variables (e.g. Low HB & SCDs)
- `Incidence rate`: The percentage of patients that have the second variable over n.

- `Odds Ratio (95% Confidence Interval)`: The odds ratio of patients with the first variable (e.g. Low HB) having the second variable (e.g. DVT) compared to patients with the first variable not having the second variable.
- `p-value`: The p-value of this odds ratio (<0.05 is significant).