

Team Project Multivariate Analysis

Adrian White, Cesar Conejo, Xavier Bryant

11/14/2020

Team members

- Adrian White
- Cesar Conejo: 100443596 (Representative member)
- Xavier Bryant: 100445659

Introduction data set

Adrian resume about the CRASH2 trial in general:

Some interesting links:

- <https://pubmed.ncbi.nlm.nih.gov/23477634/>
- [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(10\)60835-5/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)60835-5/fulltext)

info available in Vanderbilt Biostatistics

Summary variables in the data set

Adrian/xavier. You can send me the columns with this info. By the moment, I am using this ones:

- entryid: Unique Numbers for Entry Forms
- sex:
- ageP
- injurytime: Hours Since Injury
- injurytype:
- sbp: Systolic Blood Pressure
- rr: Respiratory Rate
- cc: Central Capillary Refill Time
- hr: Heart Rate
- btransf: Blood Products Transfusion
- ncell: Number of Units of Red Cell Products Transfused
- nplasma: Number of Units of Fresh Frozen Plasma Transfused
- nplatelets: Number of Units of Platelets Transfused
- ncryo: Number of Units of Cryoprecipitate Transfused
- bvii: Recombinant Factor VIIa Given

- boxid: Treatment Box Number
- packnum: Treatment Pack Number

Summary and Graphical display

A review of the structure of the dataset is the following:

```
## 'data.frame': 9543 obs. of 17 variables:
## $ entryid : int 1 3 4 6 7 8 9 11 12 14 ...
## $ sex : Factor w/ 2 levels "male","female": 1 1 1 1 1 1 1 1 1 2 ...
## $ age : int 50 30 40 19 27 16 29 41 56 37 ...
## $ injurytime: num 1 1 2 3 0.5 1 1 0.5 0.5 8 ...
## $ injurytype: Factor w/ 3 levels "blunt","penetrating",...: 1 1 2 2 2 2 1 2 1 2 ...
## $ sbp : int 75 70 60 90 90 90 116 120 60 104 ...
## $ rr : int 28 26 20 30 26 28 15 15 9 23 ...
## $ cc : int 5 6 5 5 5 2 3 3 3 5 ...
## $ hr : int 120 130 120 90 96 118 118 70 100 92 ...
## $ btransf : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...
## $ ncell : num 1 2 4 2 1 1 16 8 4 4 ...
## $ nplasma : int 0 0 0 0 0 0 9 11 9 0 ...
## $ nplatelets: int 0 0 0 0 0 0 22 10 0 0 ...
## $ ncryo : int 0 0 0 0 0 0 0 0 0 0 ...
## $ bvii : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
## $ boxid : int 2001 2011 2011 2011 2011 2011 4035 4036 4036 8001 ...
## $ packnum : int 25 21 22 24 25 28 41 42 43 82 ...
```

A summary of the values in the data set are:

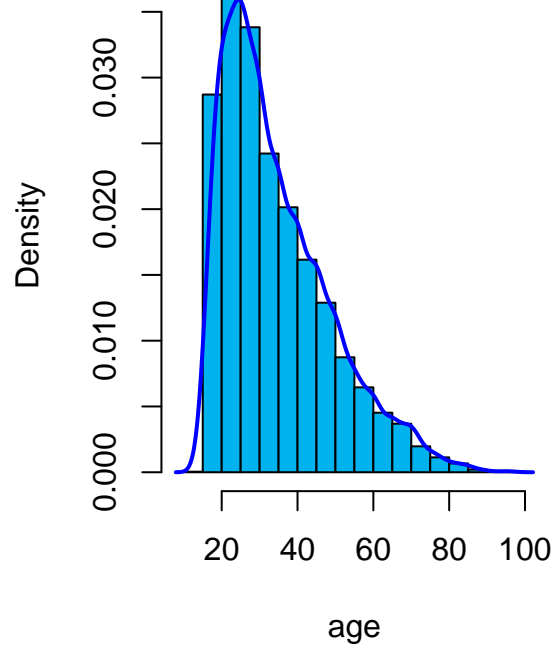
```
##      entryid      sex      age      injurytime
## Min.   :    1  male :7952  Min.   :14.00  Min.   : 0.100
## 1st Qu.: 4734  female:1591 1st Qu.:24.00 1st Qu.: 1.000
## Median : 9345                Median :31.00 Median : 3.000
## Mean   : 9667                Mean   :34.65 Mean   : 3.093
## 3rd Qu.:14599              3rd Qu.:43.00 3rd Qu.: 4.500
## Max.   :20270              Max.   :96.00 Max.   :48.000
##
##      injurytype      sbp      rr      cc
## blunt      :5219  Min.   : 4.00  Min.   : 2.00  Min.   : 1.000
## penetrating :2970 1st Qu.: 80.00 1st Qu.:20.00 1st Qu.: 2.000
## blunt and penetrating:1354 Median : 90.00 Median :22.00 Median : 3.000
##
##      Mean   : 93.12 Mean   :23.47 Mean   : 3.439
##      3rd Qu.:104.00 3rd Qu.:28.00 3rd Qu.: 4.000
##      Max.   :225.00 Max.   :91.00 Max.   :20.000
##
##      hr      btransf      ncell      nplasma      nplatelets
## Min.   : 3  0: 12  Min.   : 0.000  Min.   : 0.000  Min.   : 0.000
## 1st Qu.: 96 1:9531 1st Qu.: 2.000 1st Qu.: 0.000 1st Qu.: 0.000
## Median :110                Median : 3.000 Median : 0.000 Median : 0.000
## Mean   :108                Mean   : 3.907 Mean   : 1.432 Mean   : 0.534
## 3rd Qu.:120                3rd Qu.: 5.000 3rd Qu.: 1.000 3rd Qu.: 0.000
## Max.   :220                Max.   :60.000 Max.   :60.000 Max.   :87.000
##
##      ncryo      bvii      boxid      packnum
## Min.   : 0.0000  0:9502  Min.   :2001  Min.   :21.00
## 1st Qu.: 0.0000  1: 41  1st Qu.:2686 1st Qu.:27.00
## Median : 0.0000                Median :4431 Median :44.00
## Mean   : 0.2565                Mean   :5090 Mean   :51.09
## 3rd Qu.: 0.0000                3rd Qu.:8200 3rd Qu.:82.00
## Max.   :61.0000                Max.   :9065 Max.   :98.00
```

Finally, the list of different values by column is the following:

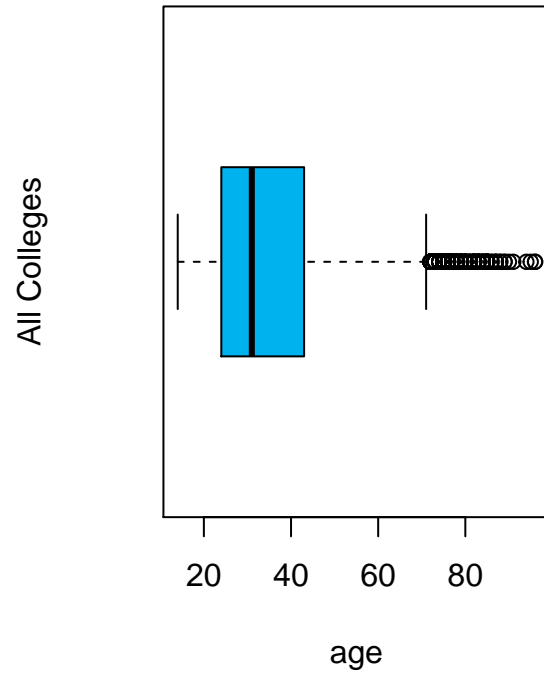
```
##      entryid      sex      age      injurytime      injurytype      sbp      rr
##      9543      2      81      78      3      153      58
##      cc      hr      btransf      ncell      nplasma      nplatelets      ncryo
##      16      155      2      47      45      39      28
##      bvii      boxid      packnum
##      2      2441      48
```

Some visualizations of the variables is:

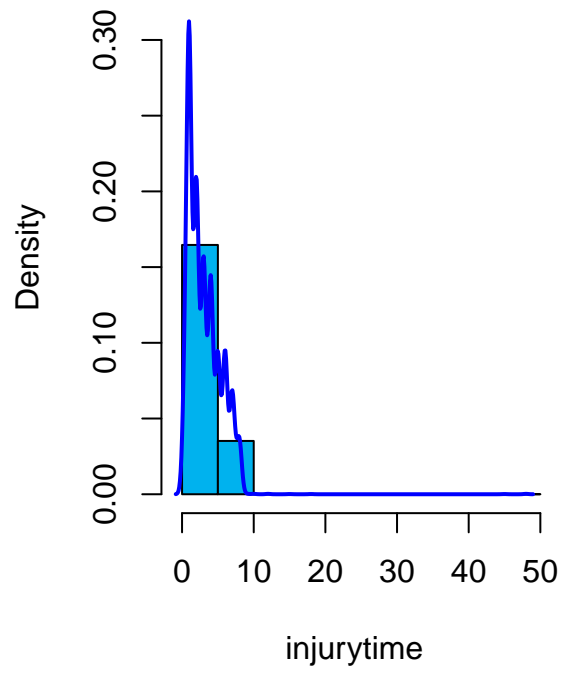
Histogram of age



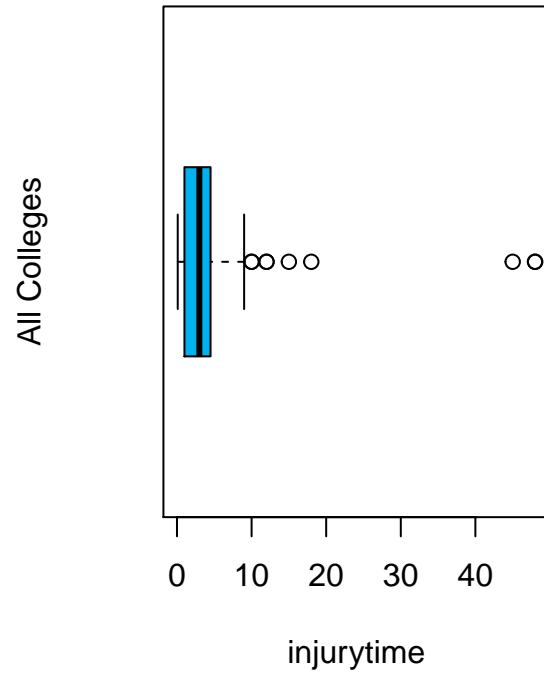
Boxplot of age



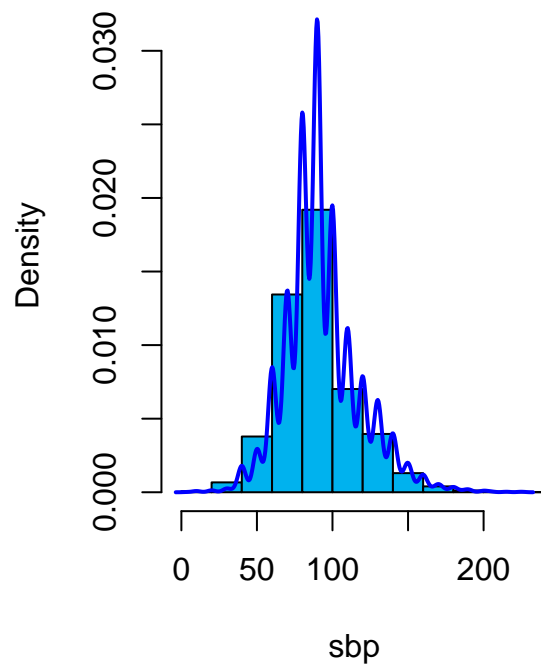
Histogram of injurytime



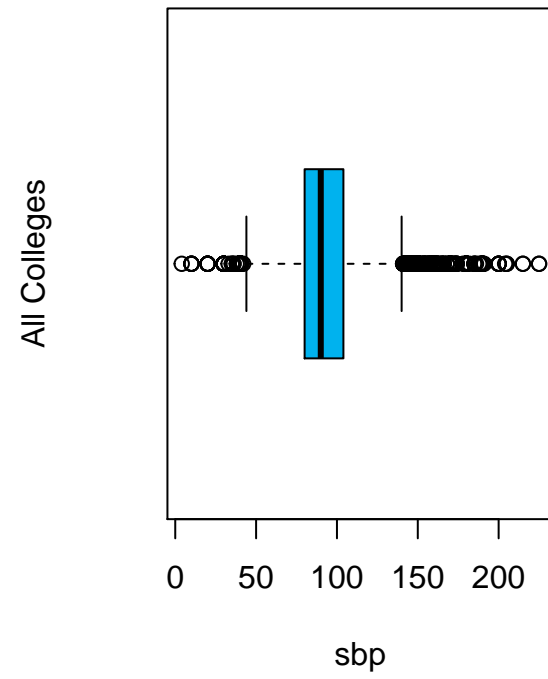
Boxplot of injurytime



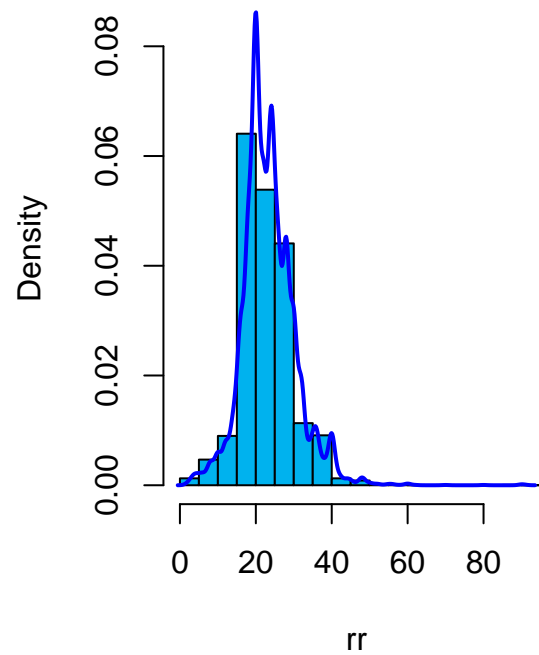
Histogram of sbp



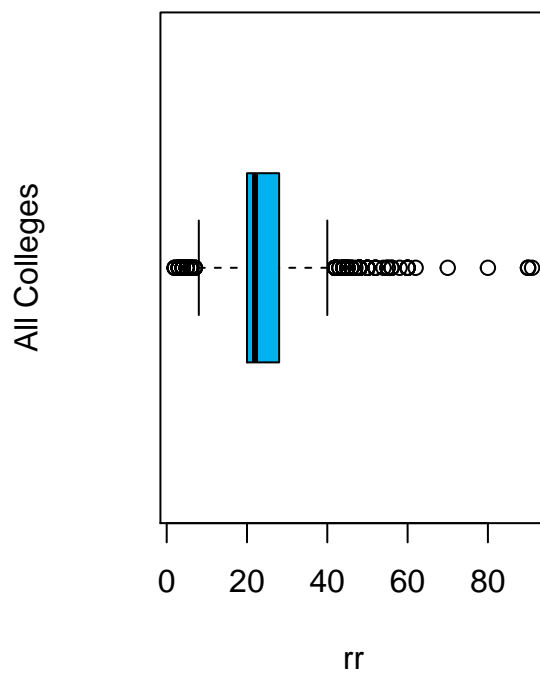
Boxplot of sbp



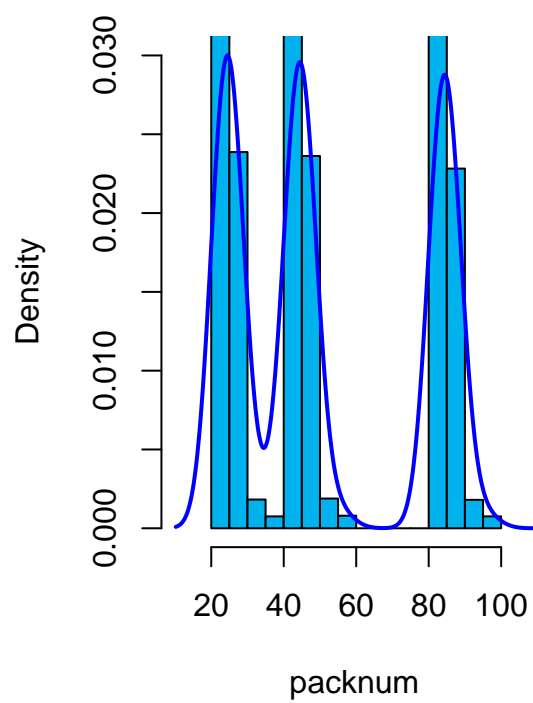
Histogram of rr



Boxplot of rr



Histogram of packnum



Boxplot of packnum

