Predicting mortality of patients based on the first 24 hours and last 24 hours
- we will compare which model performs better
- discuss does feature importance change between the two time periods?

Feature selection:
- we use three different models
  - logistic
  - random forest
  - RFE
  - to find which indicators we should set on

	feature	importance
1	Mean_GCS.x	0.07223031787977341
20	Mean_PaCO2.x	0.04315522047425926
2	Mean_HR.x	0.040680473761705443
23	Mean_FiO2.x	0.03861268980106999
12	Mean_HCO3.x	0.03798905880402715
7	Mean_Temp.x	0.03610220991557654
15	Mean_Platelets.x	0.034831225040007754
18	Mean_WBC.x	0.03456678688811546
8	Mean_Urine.x	0.0343379909899835
11	Mean_Glucose.x	0.033223630573684326

RF

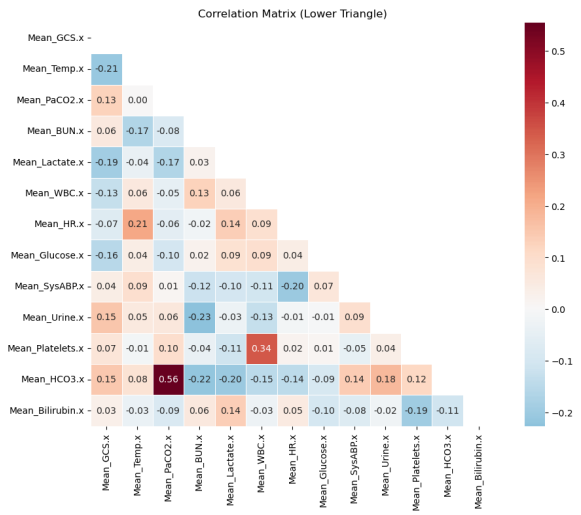
	Feature	Coefficient	Abs Importance
32	Mean_Bilirubin.x	1.825013	1.825013
1	Mean_GCS.x	-1.656659	1.656659
9	Mean_BUN.x	1.263250	1.263250
23	Mean_FiO2.x	1.174630	1.174630
2	Mean_HR.x	1.035356	1.035356
28	Mean_Albumin.x	-0.882943	0.882943
11	Mean_Glucose.x	0.845737	0.845737
18	Mean_WBC.x	0.832326	0.832326
33	Mean_Lactate.x	0.823838	0.823838
27	Mean_SaO2.x	-0.794865	0.794865
4	Mean_NIMAP.x	-0.793791	0.793791
22	Mean_DiasABP.x	-0.758433	0.758433
0	Mean_Weight.x	-0.643891	0.643891
20	Mean_PaCO2.x	-0.598986	0.598986
19	Mean_pH.x	0.573889	0.573889

After comparing between each model we settled with these 13 indicators...

colums_keep = ["Mean_GCS.x", "Mean_Temp.x", "Mean_PaCO2.x", "Mean_BUN.x", "Mean_Lactate.x", "Mean_WBC.x", "Mean_HR.x", "Mean_Glucose.x", "Mean_SysABP.x", "Mean_Urine.x", "Mean_Platelets.x", "Mean_HCO3.x", "Mean_Bilirubin.x"]

- 48 hour model produced the same results... with same important indicators
- Building the model and what is important
  - RECALL and PRECISION
  - because in the medical field FN where we predict survival when death is bad.
  - FP, means surive but predicted as dead... this means that we could overload staff with work
  - there needs to be some form of balance between the two
  - we however will prioritise Recall as getting treatement is important.

- settle with RF. Why? no linear relationship between the dataset



FIRST TRIAL 24 HR

Accuracy 0.647
Recall 0.38

0.0789658139634718

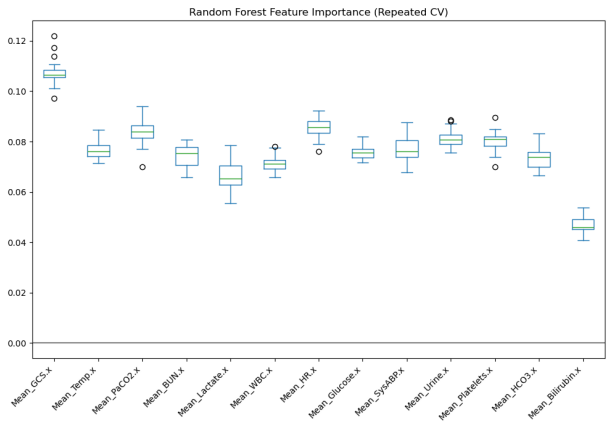
	feature	# importance
0	Mean_GCS.x	0.10744106806480182
6	Mean_HR.x	0.08917761716604443
2	Mean_PaCO2.x	0.0850450583745996
10	Mean_Platelets.x	0.08175684840848065
1	Mean_Temp.x	0.0789658139634718
3	Mean_BUN.x	0.07774996977740713
9	Mean_Urine.x	0.07763896685192358
11	Mean_HCO3.x	0.07470639268376662
7	Mean_Glucose.x	0.07283139895252985
8	Mean_SysABP.x	0.07017110042004564

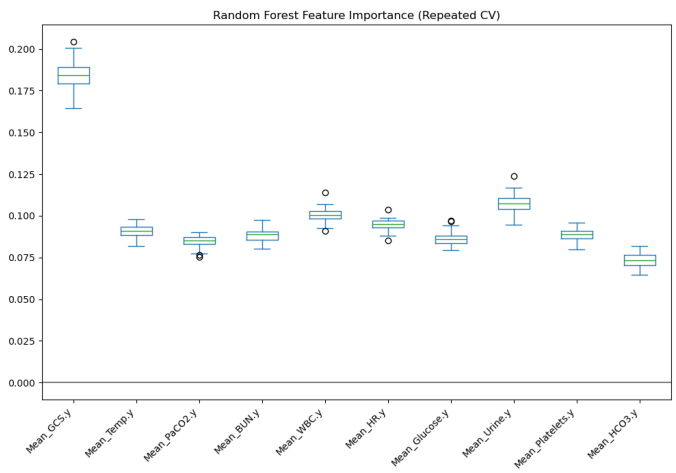
FIRST TRIAL last 24HR

Accuracy 0.742
Recall 0.538

	feature	importance
0	Mean_GCS.y	0.155882
9	Mean_Urine.y	0.088689
5	Mean_WBC.y	0.084760
2	Mean_PaCO2.y	0.074638
1	Mean_Temp.y	0.074378
6	Mean_HR.y	0.073811
7	Mean_Glucose.y	0.072888
10	Mean_Platelets.y	0.069516
3	Mean_BUN.y	0.065193
8	Mean_SysABP.y	0.062780
11	Mean_HCO3.y	0.062682
4	Mean_Lactate.y	0.060919
12	Mean_Bilirubin.y	0.053863

CV_COMPARISON





-talk about features and differences and similarities

```
Recall   : 0.436
Precision: 0.686
F1 Score : 0.533
```

	feature	importance
0	Mean_GCS.x	0.11757822828632197
5	Mean_HR.x	0.10541856021809944
2	Mean_PaCO2.x	0.10341894843174691
8	Mean_Urine.x	0.09009489464415893
9	Mean_Platelets.x	0.08745184426110812
1	Mean_Temp.x	0.08434134970400803
7	Mean_SysABP.x	0.08424943773904622

```
Recall   : 0.486
Precision: 0.609
F1 Score : 0.541
```

	feature	importance
0	Mean_GCS.y	0.19832784549633817
7	Mean_Urine.y	0.10259396036682171
4	Mean_WBC.y	0.09632603779195371
1	Mean_Temp.y	0.09293719472873034
2	Mean_PaCO2.y	0.09017992189780619
5	Mean_HR.y	0.08773558097159688
6	Mean_Glucose.y	0.08665061592569286

AFTER REMOVING SOME VARIABLES

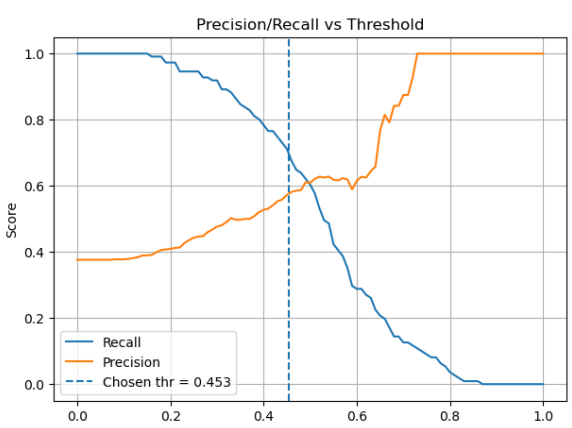
```
Using threshold: 0.4534867988694888
Accuracy : 0.698
Recall   : 0.712
Precision: 0.581
F1 Score : 0.64
ROC AUC  : 0.749
```

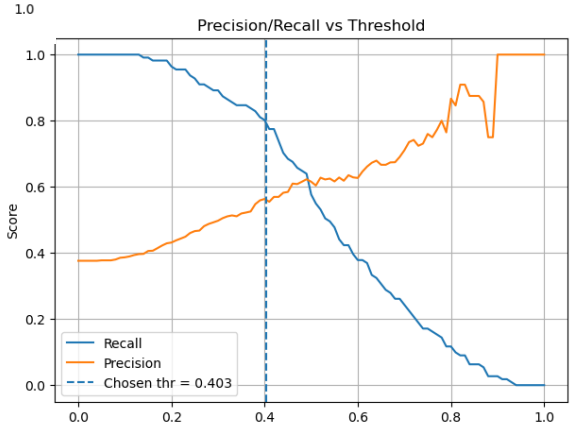
Improving the model... using GridSearch to find best hyperparamaters and some finetuning

```
Using threshold: 0.4025322696760549
Accuracy : 0.692
Recall   : 0.802
Precision: 0.563
F1 Score : 0.662
ROC AUC  : 0.763
```

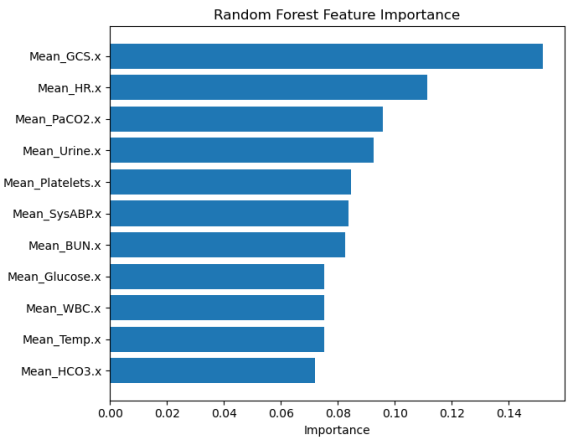
Clear to say that 48 hr is better indicator. why?

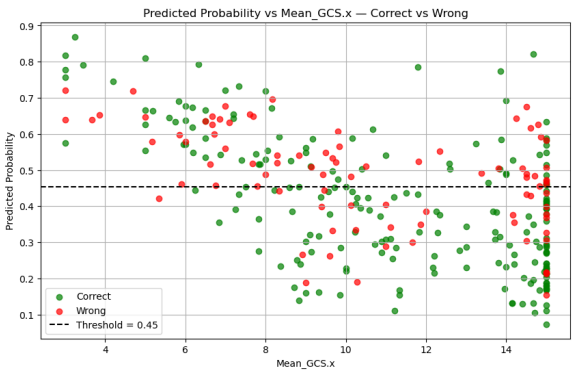
Some graphs to visualise tradeoff

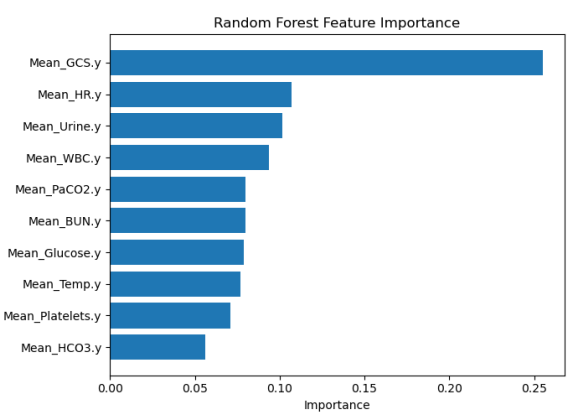


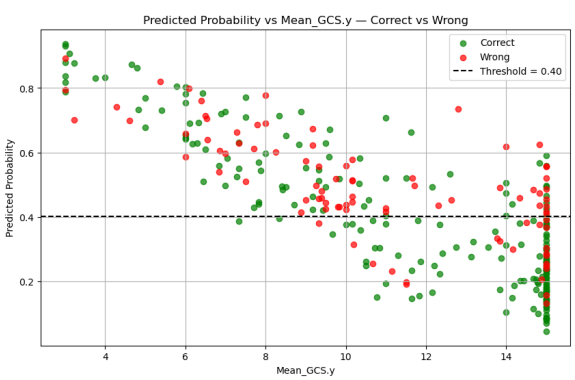


Interpreting the model









- comment on feature importance
- 2nd graph showcases some trend/correlation between GCS and predicted prob for Random Forest
- around the threshold random forest is not as accurate in predicting mortality or not
- GCS 3-8 is critical
- still a very strong indicator though
- 24hr features hold similar importance vs 48hr which shows GCS as most important by alot

What are the applicatons of this model and limitations?

- discussion or part of mine?
- black box hard to understand
- DNR
- unbalanced datasets can affect model performance
- what form of medical faciltiies are being used

Conclusion

- insights and possible things to be improved on
- more information
- use different models for future
- interpreting RF
- consider doing time series and predict probability of death over time
- feature engineering
- based on indicators coming in and final hours. does this spread significantly impact probabilty of death