PingDoctors

WELCOME TO COVID 19 EARLY WARNING SYSTEM FROM BENCHPEDIA

- COUNTRY TO REGION TO HOSPITAL BREAKDOWN FOR DAILY CASES AND DEATHS

- AI MODELS ARE USING COMMONLY AVAILABLE LABORATORY DATA TO FIND COVID-19 PATIENTS NEEDING MOST ATTENTION

- Predicts admission to general ward, semi-intensive unit or intensive care unit among confirmed COVID-19 cases.

- RESOURCE AND COST OPTIMIZATION FOR PROVIDERS TO HANDLE PATIENTS

- DATA REFRESHES EVERY DAY , AND ALLOWS YOU TO BUILD NEW PREDICTIONS AND USE DIFFERENT VARIABLES TO TAKE BEST MODEL

- CAN BE LINKED TO HOSPITAL / CLINICAL SYSTEMS TO PING PATIENT, DOCTORS, HEALTHCARE PROVIDERS IN REAL TIME USING BENCHPEDIA ECOSYSTEM OF SERVICES.

CUSTOMIZE AND DEVELOP YOUR END TO END EARLY WARNING SYSTEM , CARE MARKERS AND INTEGRATED INTO BENCHPEDIA LLC SERVICES FOR GENERAL AVAILABILITY

GO TO WWW.PINGDOCTORS.COM FOR MORE DETAILS.

COMPLETE MODEL DETAILS AND OUTPUTS ARE OPEN-SOURCED BY BENCHPEDIA LLC

DISCLAIMERS - WE HAVE USED INPUTS FROM MICROSOFT BI FOR COUNTRY AND COUNTY VIEW AND KAGGLE COVID-19 FRAMEWORKS FOR DATA AND MODELS BASELINES

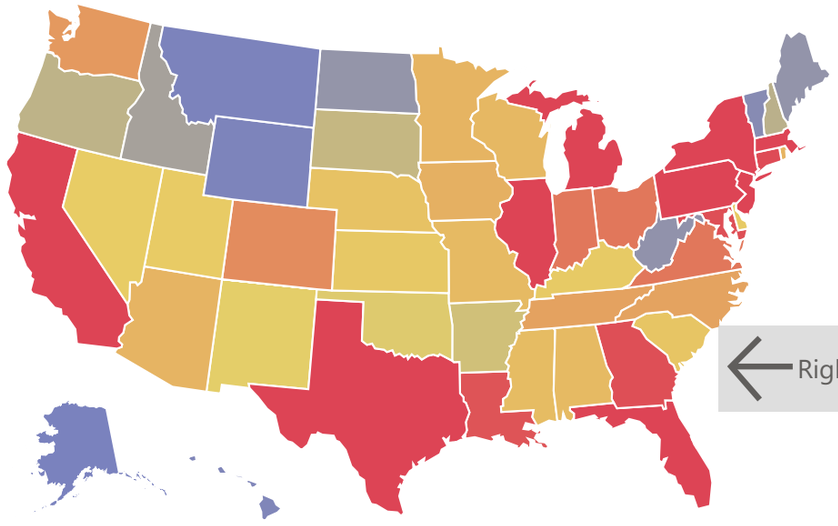
The dataset used here for Demo contains anonymized data from patients seen at the Hospital Israelita Albert Einstein, at São Paulo, Brazil, and who had samples collected to perform the SARS-CoV-2 RT-PCR and additional laboratory tests during a visit to the hospital. All data were anonymized following the best international practices and recommendations. All clinical data were standardized to have a mean of zero and a unit standard deviation. Limitations for the test data set - small number of data samples; imbalanced dataset; and high sparsity in the predictors

CASE BREAKDOWN BY STATE WITH DRILL THRU TO COUNTY LEVEL

Benchpedia – A New Standard Of Patient Care

USA	**1,321,813**	**78,723**	**6.0%**
Wyoming	504	7	1.4%
Wisconsin	10,219	400	3.9%
West Virginia	1,362	54	4.0%
Washington	16,891	931	5.5%
Virginia	24,088	711	3.0%
Vermont	927	53	5.7%
Utah	6,254	66	1.1%
Texas	38,869	1,088	2.8%
Tennessee	14,985	243	1.6%
South Dakota	3,517	34	1.0%
South Carolina	7,653	331	4.3%
Rhode Island	11,274	422	3.7%
Pennsylvania	56,611	3,707	6.5%
Oregon	3,228	127	3.9%
Oklahoma	4,589	272	5.9%
Ohio	24,081	1,341	5.6%
North Dakota	1,491	9	0.6%
North Carolina	14,765	547	3.7%
New York	335,464	26,656	7.9%
New Mexico	4,844	198	4.1%
New Jersey	138,532	9,255	6.7%
New Hampshire	3,071	133	4.3%
Nevada	6,135	322	5.2%
Nebraska	8,173	103	1.3%

Confirmed Cases | Total Deaths



Right click to drill thru

Click on a State to view by County

1/22/2020 5/10/2020

Methodology

This interactive feature aggregates data from the Centers for Disease Control and Prevention (CDC), state- and local-level public health agencies. County-level data is confirmed by referencing state and local agencies directly (link).

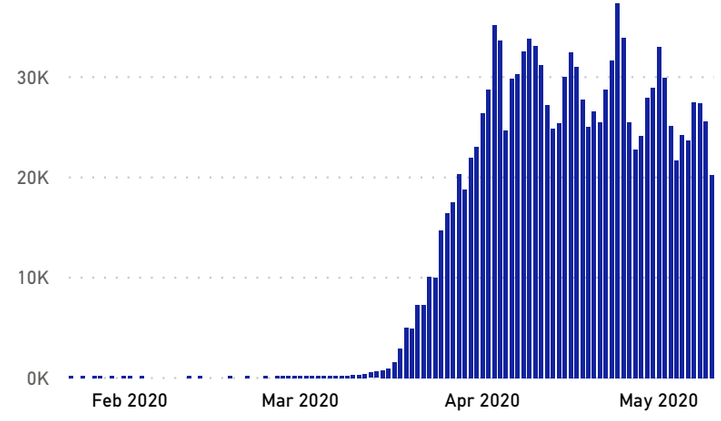
Data Source

Data provided by USAFacts. Because of the frequency of data updates, they may not reflect the exact numbers reported by government organizations or the news media. For more information or to download the data, please click the logo below. Data updated through May 10, 2020.

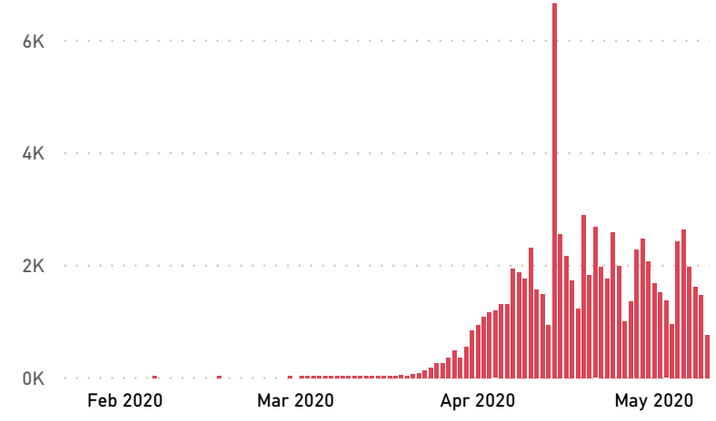
USAFACTS

Daily increments | Cumulative

Daily Cases



Daily Deaths



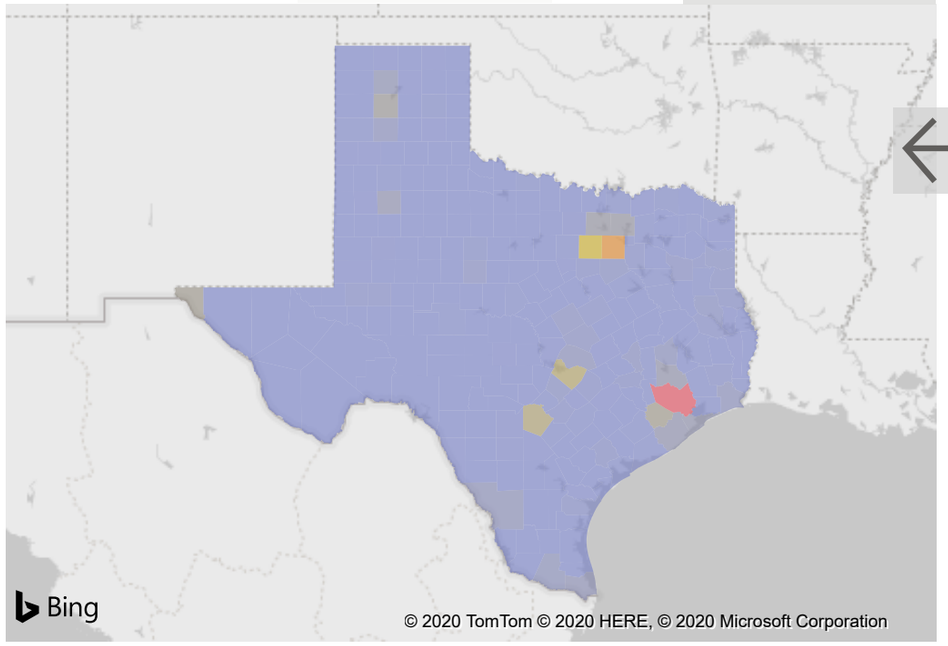
CASE BREAKDOWN BY COUNTY WITH DRILL THRU TO PATIENT LEVEL

Benchpedia – A New Standard Of Patient Care

Anderson County	41	0	0.0%
Andrews County	21	0	0.0%
Angelina County	100	0	0.0%
Aransas County	2	0	0.0%
Archer County	0	0	
Armstrong County	2	0	0.0%
Atascosa County	22	1	4.5%
Austin County	15	0	0.0%
Bailey County	3	0	0.0%
Bandera County	6	0	0.0%
Bastrop County	109	2	1.8%
Baylor County	0	0	
Bee County	6	0	0.0%
Bell County	220	3	1.4%
Bexar County	1,901	56	2.9%
Blanco County	6	0	0.0%
Borden County	0	0	
Bosque County	5	0	0.0%
Bowie County	76	2	2.6%
Brazoria County	656	9	1.4%
Brazos County	265	18	6.8%
Brewster County	1	0	0.0%
Briscoe County	1	0	0.0%
Brooks County	1	0	0.0%
Brown County	38	6	15.8%
Total	**38,869**	**1,088**	**2.8%**

Confirmed Cases | Total Deaths

Back Country/States



Right click to drill thru

1/22/2020 5/10/2020

Back Country/States

Methodology

This interactive feature aggregates data from the Centers for Disease Control and Prevention (CDC), state- and local-level public health agencies. County-level data is confirmed by referencing state and local agencies directly (link).

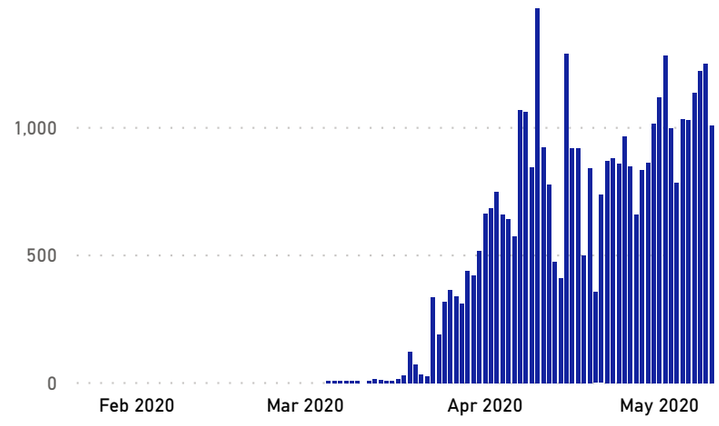
Data Source

Data provided by USAFacts. Because of the frequency of data upates, they may not reflect the exact numbers reported by government organizations or the news media. For more information or to download the data, please click the logo below. Data updated through May 10, 2020.

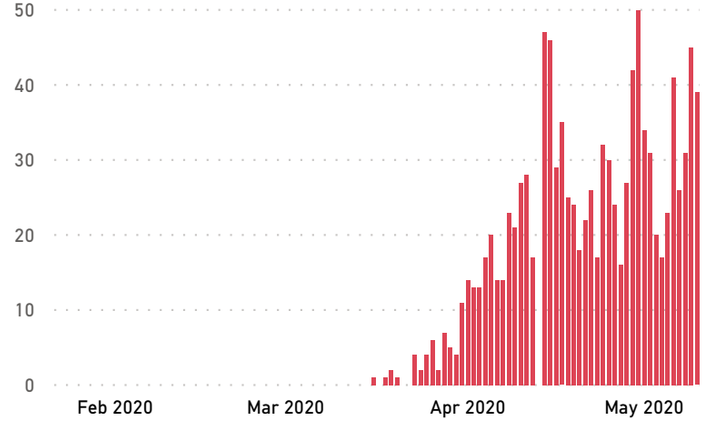
USAFACTS

Daily increments | Cumulative

Daily Cases



Daily Deaths



COUNTY BREAKDOWN OF PATIENTS NEEDING CARE

Benchpedia – A New Standard Of Patient Care

Back Country/States

County

Lee County, TX

Total Patients in County / State (this is real data from a given region for modelling)

5644

Total Patients in Test Model

668

FULLY LOADED MODEL - NO OPTIMIZATION

ROC	Sens	Spec
92%	96%	42%

Patient needing attention ASAP - Move to ICUs !!

182

False Positive - High potential

278

False Negative - General Ward

6

True Negative - Can go home!!

202

TRUE POSITIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
010d0ffe8b83059	positive	positive	95%	0.05
0123beb6c71cae3	positive	positive	95%	0.05

Right click to drill thru

FALSE POSITIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
0057f527d9e0d1f	negative	positive	0.04	0.96
039227fa5739595	negative	positive	0.14	0.86

Right click to drill thru

FALSE NEGATIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
03b1a146884942a	positive	negative	0.01	0.99
17319c86ff7db04	positive	negative	0.00	1.00

Right click to drill thru

TRUE NEGATIVE

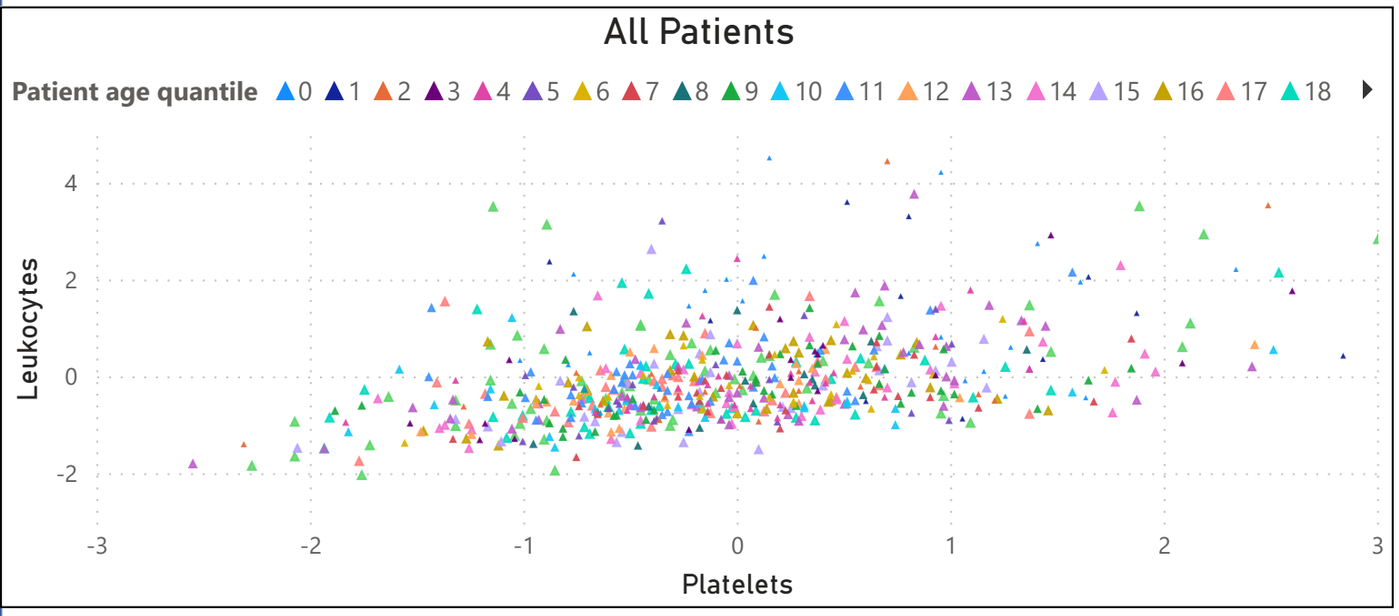
SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
02a6e340419afaa	negative	negative	0.00	1.00
02ff56fe934f3b0	negative	negative	0.00	1.00

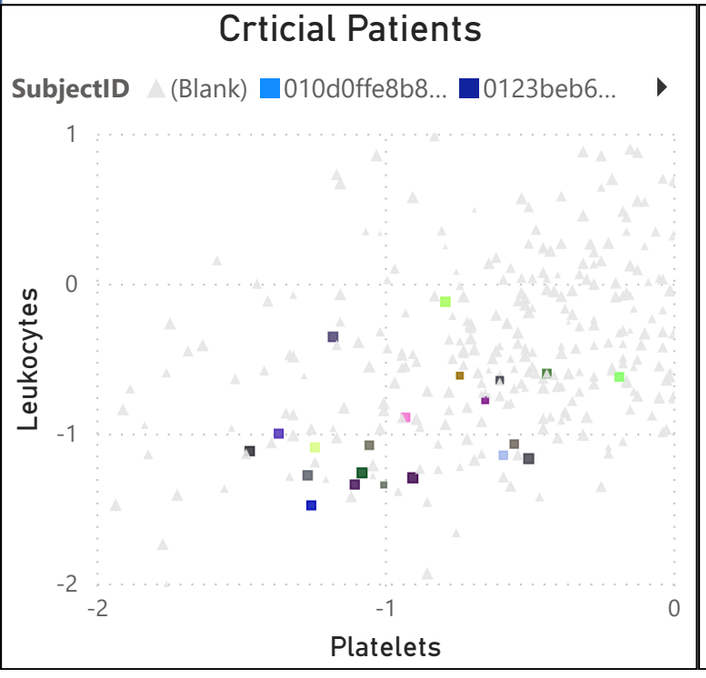
Right click to drill thru

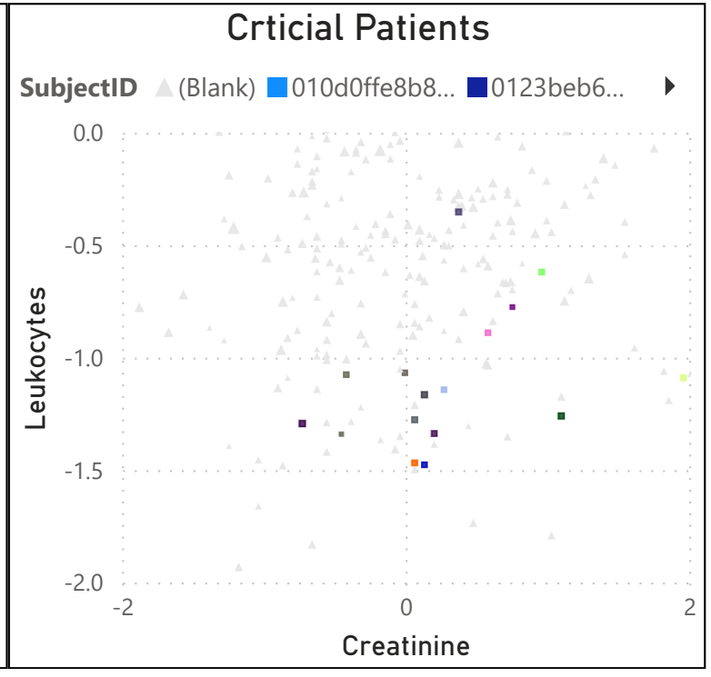
INDIVIDUAL BREAKDOWN OF PATIENTS NEEDING CARE

Benchpedia – A New Standard Of Patient Care

Back Patient Splits







Total Patients

5644

Total Patients in Test

668

Critical Patients

182

Patients for ICU

Model.positive

4% 100%

SubjectID	Patient age quantile	Model. positive	Creatinine	Leukocytes	Platelets
1a51e30b49fe00c	15	79%	0.06	-1.46	-2.06
a7bc77ac8caf300	16	8%		-1.11	-1.47
794346399cac402	14	94%		-0.99	-1.37
482a960eb07d3cd	16	92%	0.06	-1.27	-1.27
6efc63ca6d085c1	14	95%	0.13	-1.47	-1.26
88baf02b9f4c1b3	14	42%	1.95	-1.09	-1.25
Total	**52593**		**0.00**	**0.00**	**0.00**

ADD MORE LAB DATA FOR PATIENT

PATIENT SPLITS - WHEN YOU HAVE NO LIMITS ON RESOURCES

Maximum Resources

ICU bound Patients !!

182

Patients to watch !!

278

TRUE POSITIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
010d0ffe8b83059	positive	positive	95%	0.05
0123beb6c71cae3	positive	positive	95%	0.05
01596f72ec049fa	positive	positive	95%	0.05
0161b6af50624fd	positive	positive	86%	0.14
02a743f3bb71d35	positive	positive	17%	0.83
0461443399b27d5	positive	positive	95%	0.05
057e8ce3c6f2f31	positive	positive	95%	0.05
05f0703db4a9723	positive	positive	95%	0.05
07b49072f096ad7	positive	positive	95%	0.05
07edb2faf4d7d4b	positive	positive	90%	0.10

FALSE POSITIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
0057f527d9e0d1f	negative	positive	0.04	0.96
039227fa5739595	negative	positive	0.14	0.86
05448eec447b12e	negative	positive	0.08	0.92
062808aacc476ba	negative	positive	0.04	0.96
06b933398e6e9e6	negative	positive	0.07	0.93
0835169097d077d	negative	positive	0.14	0.86
0879012c5f10c0e	negative	positive	0.03	0.97
08e04abcd2027c8	negative	positive	0.04	0.96
0b1eca05ec635a2	negative	positive	0.14	0.86
0b3f4b44ffdc600	negative	positive	0.07	0.93

General Ward Patients !!

33

Can go Home !!

461

FALSE NEGATIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
03b1a146884942a	positive	negative	0.01	0.99
17319c86ff7db04	positive	negative	0.00	1.00
4382f5ea05e60c4	positive	negative	0.02	0.98
6c66c5e197e6b13	positive	negative	0.00	1.00
b388e20effbbf0b	positive	negative	0.00	1.00
c3c08e6b5fb20d7	positive	negative	0.00	1.00

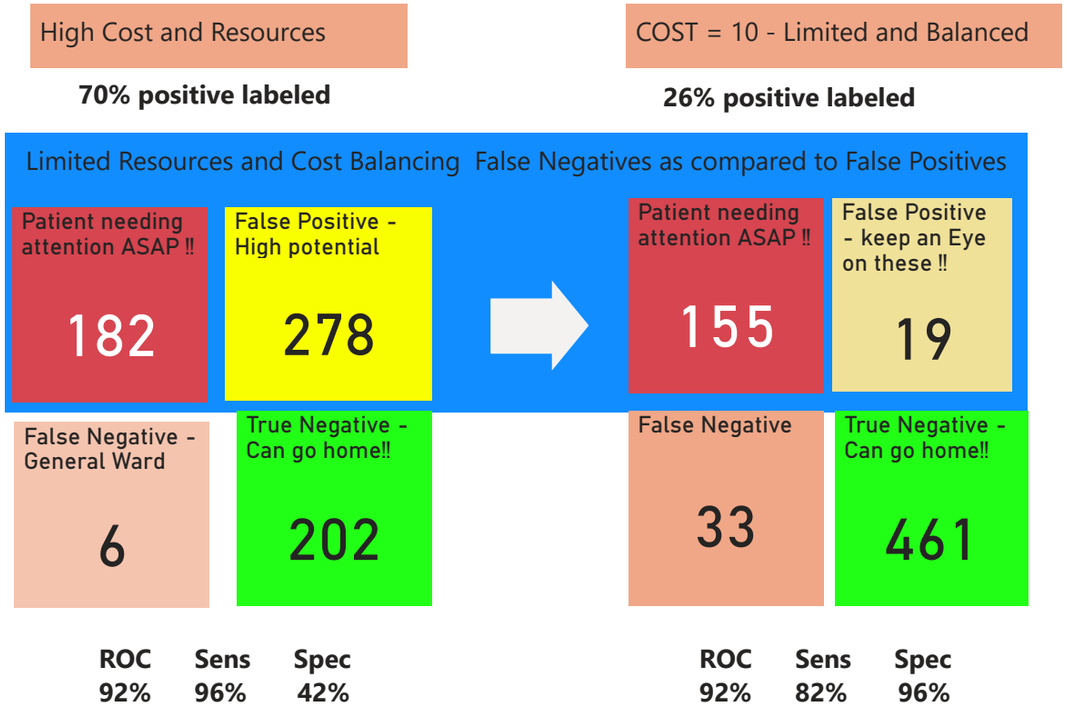
TRUE NEGATIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
02a6e340419afaa	negative	negative	0.00	1.00
02ff56fe934f3b0	negative	negative	0.00	1.00
03fd6282deaf293	negative	negative	0.02	0.98
04235e8a80d92ed	negative	negative	0.00	1.00
04ffc6b0eeddb93	negative	negative	0.00	1.00
06f185cb0ce00d0	negative	negative	0.00	1.00

IDEAL BALANCED SCENARIOS - Patient Splits for Optimized Resources

Parameterized model with a cost function that allows doctors to over- or under-weight false negative classification as compared with a false positive classification depending on a target policy as well as input prevalence as input parameter.

Scenario - Limited resources, different policies with a COST of 10
COST can range from 1 (min resources) to 200 (max resource availability) for all 182 True Positive patients in best case scenario



TRUE POSITIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
010d0ffe8b83059	positive	positive	0.95	0.05
0123beb6c71cae3	positive	positive	0.95	0.05
01596f72ec049fa	positive	positive	0.95	0.05
0161b6af50624fd	positive	positive	0.86	0.14
0461443399b27d5	positive	positive	0.95	0.05
057e8ce3c6f2f31	positive	positive	0.95	0.05
05f0703db4a9723	positive	positive	0.95	0.05

Different countries and regions can have different prevalence,
while different hospitals might have different policies to define the cost parameter.
objective function is - $\max(\text{sensitivity}+r*\text{specificity})$
where
$r=1-\text{prevalence}/\text{cost}*\text{prevalence}$ and r is the relative cost of a false negative classification (as compared with a false positive classification).

Resource Optimizer

IDEAL BALANCED SCENARIOS - APPROACH FOR COST AND CARE RESOURCE MANAGEMENT

Benchpedia – A New Standard Of Patient Care

COST = 10

26% positive labeled

ICU Bound Patients !!

155

TRUE POSITIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
010d0ffe8b83059	positive	positive	0.95	0.05
0123beb6c71cae3	positive	positive	0.95	0.05
01596f72ec049fa	positive	positive	0.95	0.05
0161b6af50624fd	positive	positive	0.86	0.14
0461443399b27d5	positive	positive	0.95	0.05
057e8ce3c6f2f31	positive	positive	0.95	0.05
05f0703db4a9723	positive	positive	0.95	0.05
07b49072f096ad7	positive	positive	0.95	0.05
07edb2faf4d7d4b	positive	positive	0.90	0.10
094ad06472341b2	positive	positive	0.96	0.04

Patients to Watch !!

19

FALSE POSITIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
0c6359f77661e20	negative	positive	0.67	0.33
15c78aa6cb0c7e4	negative	positive	0.37	0.63
1f9ef4c4c5cc30f	negative	positive	1.00	0.00
2f901095f53518e	negative	positive	0.97	0.03
3c78fb09a857ddc	negative	positive	0.99	0.01
49b52b44625292a	negative	positive	0.61	0.39
5624291313277ad	negative	positive	0.82	0.18
5c56087861a9b73	negative	positive	0.50	0.50
68935693477d7fe	negative	positive	0.98	0.02
78e377b1635f8de	negative	positive	0.71	0.29

General Ward Patients

33

FALSE NEGATIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
02a743f3bb71d35	positive	negative	0.17	0.83
03b1a146884942a	positive	negative	0.01	0.99
14acdd505a4d3e4	positive	negative	0.07	0.93
17319c86ff7db04	positive	negative	0.00	1.00
2081f4ac7e7ee9b	positive	negative	0.25	0.75
35c9659c7a3be84	positive	negative	0.08	0.92

Can Go Home !!

461

TRUE NEGATIVE

SubjectID	Model.obs	Model.pred	Model.positive	Model.negative
0057f527d9e0d1f	negative	negative	0.04	0.96
02a6e340419afaa	negative	negative	0.00	1.00
02ff56fe934f3b0	negative	negative	0.00	1.00
039227fa5739595	negative	negative	0.14	0.86
03fd6282deaf293	negative	negative	0.02	0.98
04235e8a80d92ed	negative	negative	0.00	1.00
04ffc6b0eeddb93	negative	negative	0.00	1.00

Go back to Critical patients

Patient ID	SARS-Cov-2 exam result	Patient age quantile	Platelets	Leukocytes	Adenovirus	Albumin	Alanine transaminase	Alkaline phosphatase	Arteiral Fio2	Arteri
d9c2385bfe97417	negative	19	9.53	1.65	not_detected		-0.50	0.14		
f8484cba3fae34f	negative	0	3.38	2.81	not_detected		-0.39	2.18		
f1cef41f035ade9	negative	0	3.28	2.87					2.110163927	-1.09
57c883b4579eea9	negative	0	3.04	1.76	not_detected		0.10			
a027e45b25b31f2	negative	19	3.00	2.85	not_detected					
9724cde45cc9fa4	negative	1	2.84	0.43		0.211694866	0.63	2.22		
2c97006c5a4e6ae	negative	3	2.60	1.78	not_detected					
d88bc973d6012be	negative	18	2.54	2.16	not_detected					
b26b04c1e6de8e7	negative	10	2.51	0.56	not_detected					
4ebd8587504c1e3	negative	2	2.48	3.55	not_detected					
9488e19a8c49e4a	negative	12	2.42	0.67						
ab53eb2988092a0	negative	13	2.41	0.21						
c88564b2564a8c1	negative	0	2.33	2.22	not_detected		-0.56	2.33		
b64bc1acd95fd62	negative	19	2.18	2.95	not_detected		-0.28	-0.17		
Total		**52593**	**0.00**	**0.00**			**0.00**	**0.00**		

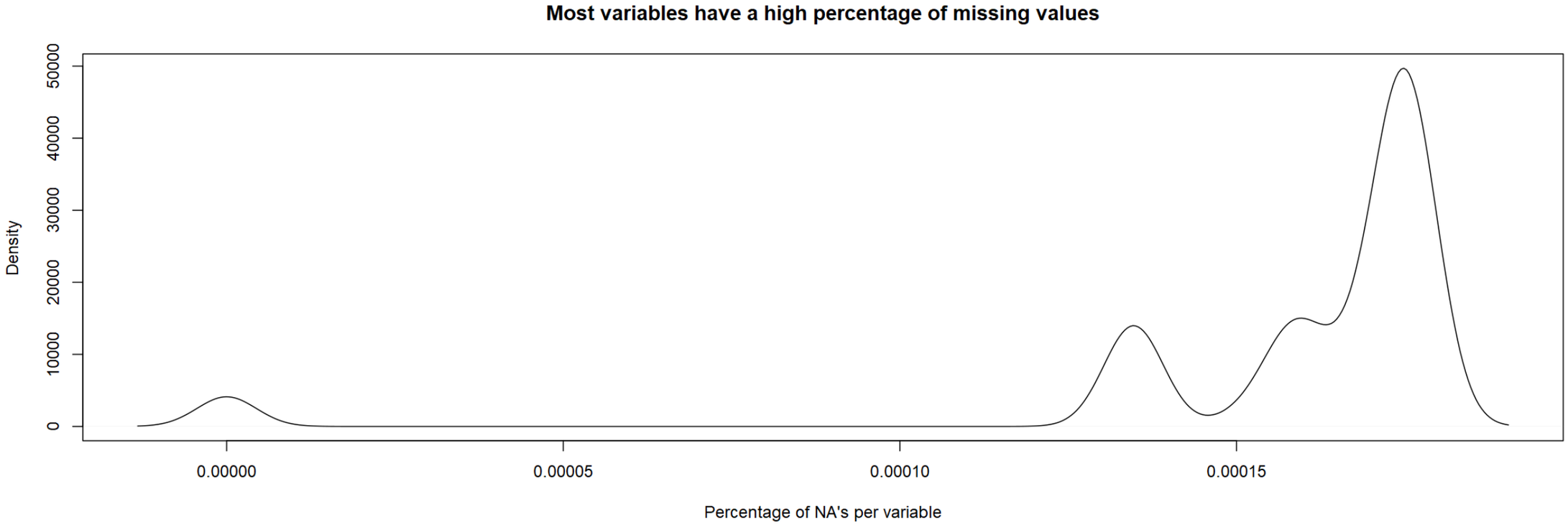
The dataset contains 109 variables (predictors), a Patient ID and one target outcome variable, which indicates whether the patient tested positive/negative for SARS-Cov-2. There are 5644 samples available with 558 positive cases, which constitutes 10% of the dataset.

Patient ID

33bbe2fb8df7a5b

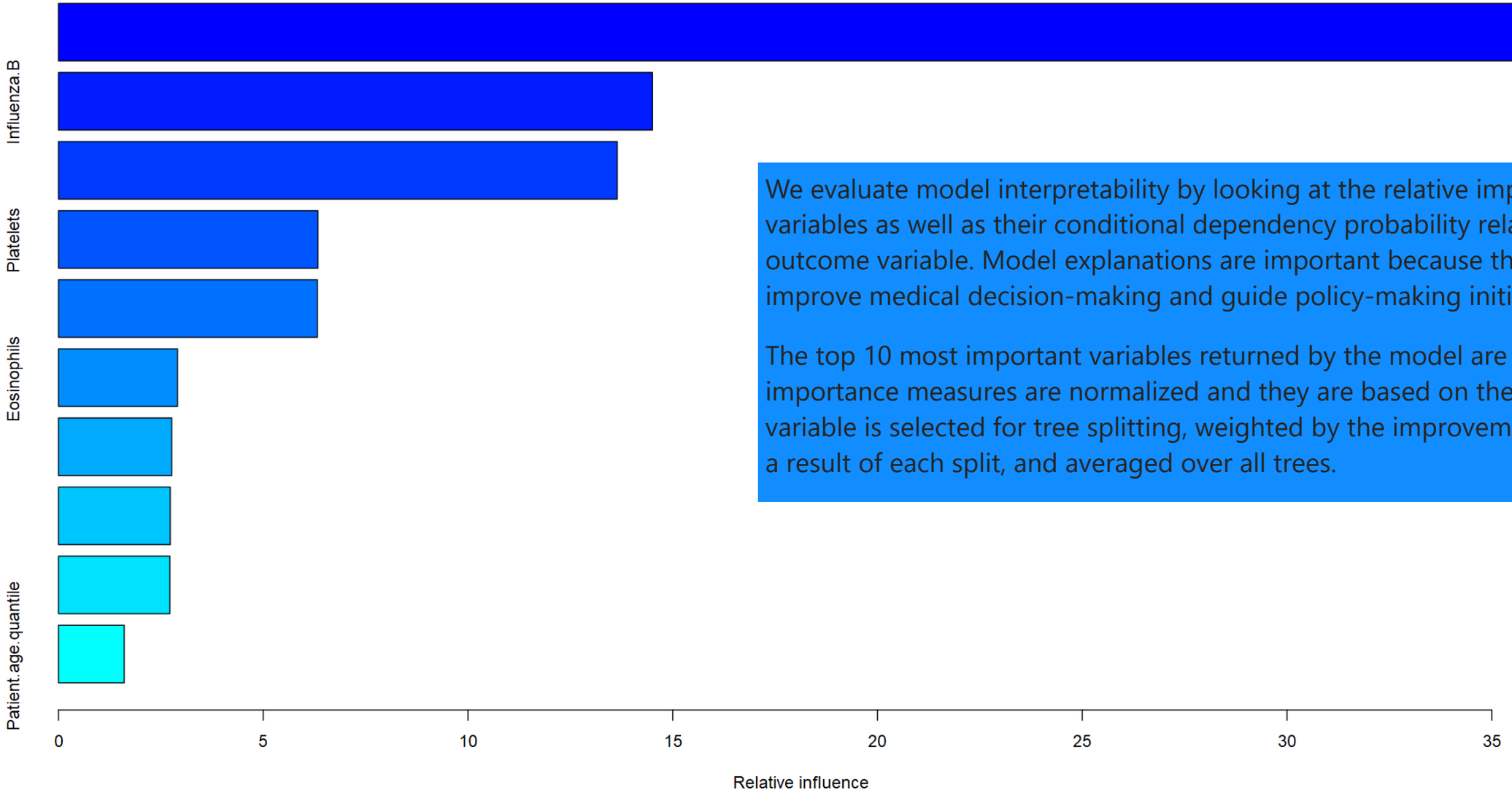
33bbe2fb8df7a5b

Data Missing % of NA



MOST VARIABLES HAVE NO DATA - we remove variables that have too many missing data points (>= 95%). We also remove samples that are too sparse in laboratory data, we choose to keep negative samples that have at least 10 variables with data points available.
This is performed to avoid an overfit scenario where a few samples (sparse but positive) may have an undue influence on the predictive model.

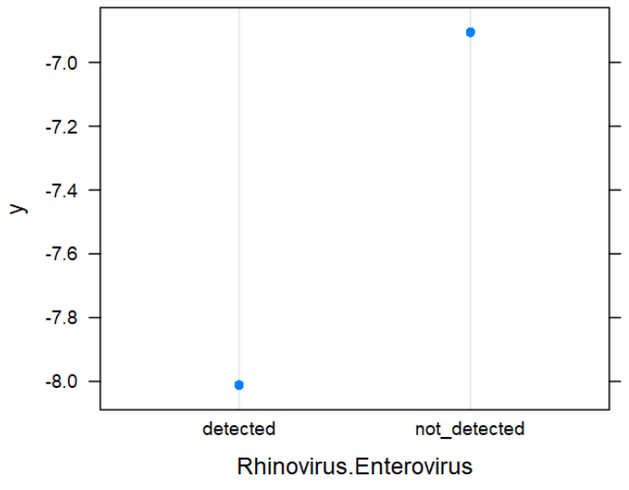
Top 10 variables



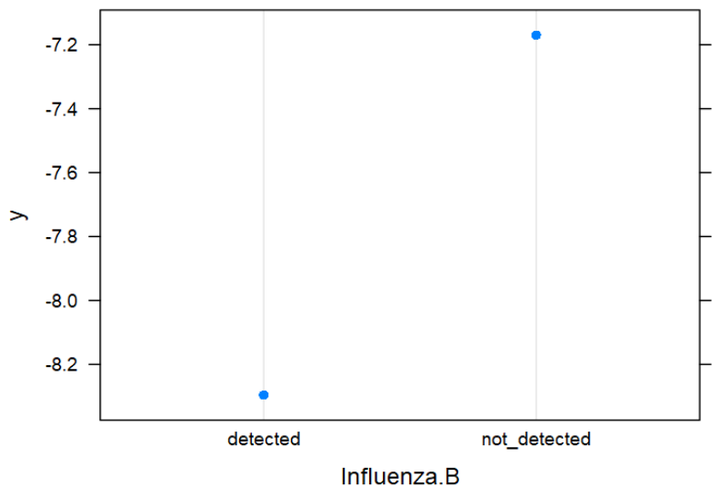
We evaluate model interpretability by looking at the relative importance of the variables as well as their conditional dependency probability relative to the outcome variable. Model explanations are important because they can be used to improve medical decision-making and guide policy-making initiatives.

The top 10 most important variables returned by the model are given below. The importance measures are normalized and they are based on the number of times a variable is selected for tree splitting, weighted by the improvement to the model as a result of each split, and averaged over all trees.

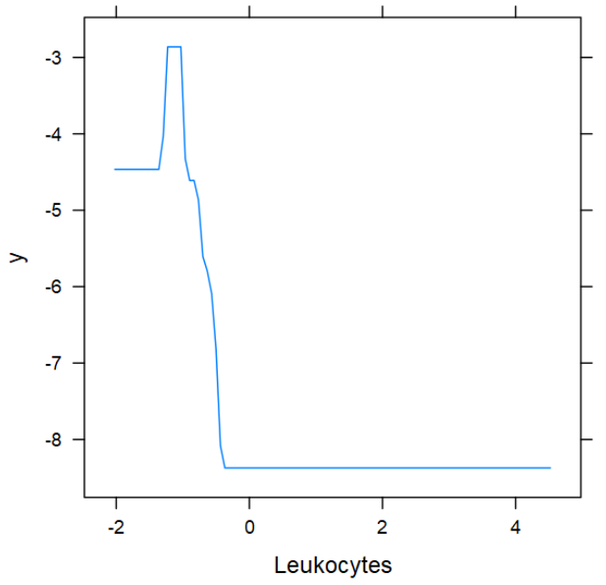
DRIVER 1



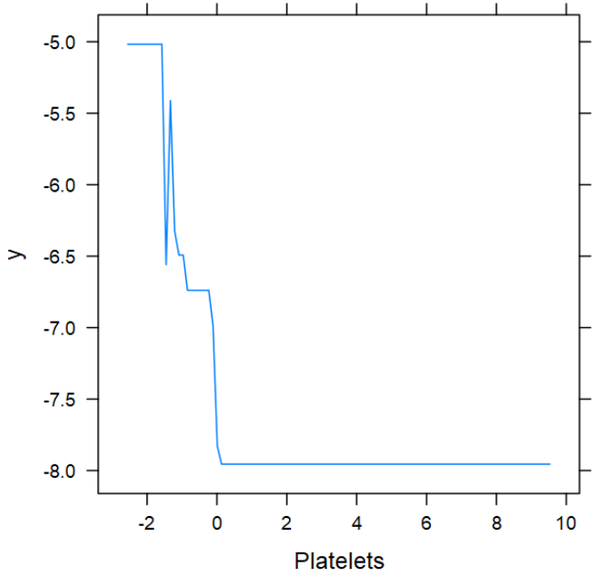
DRIVER 2



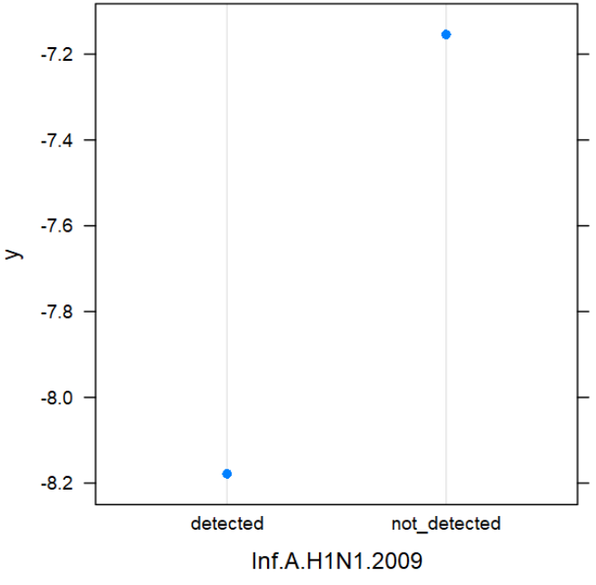
DRIVER 3



DRIVER 4



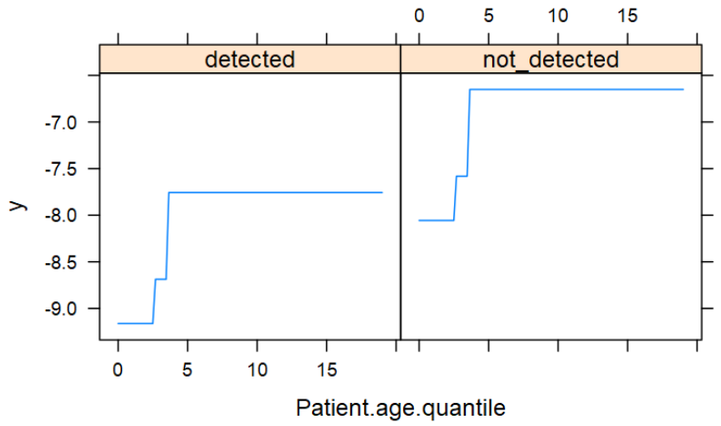
DRIVER 5



Patients admitted with COVID-19 symptoms who tested negative for Rhinovirus Enterovirus, Influenza B and Inf.A.H1N1.2009 and presented low levels of Leukocytes and Platelets were more likely to test positive for SARS-CoV-2.

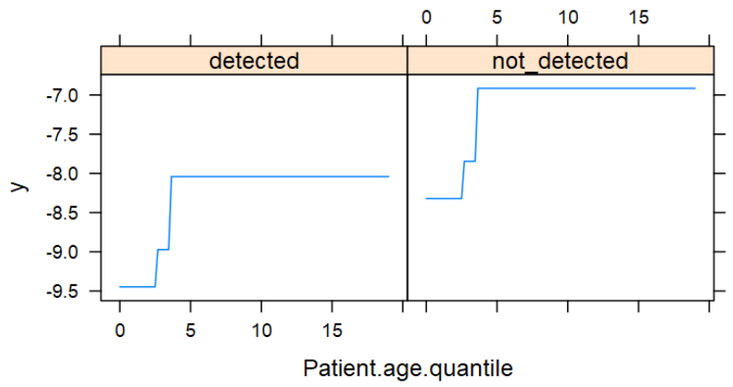
FACTOR 1

Rhinovirus.Enterovirus



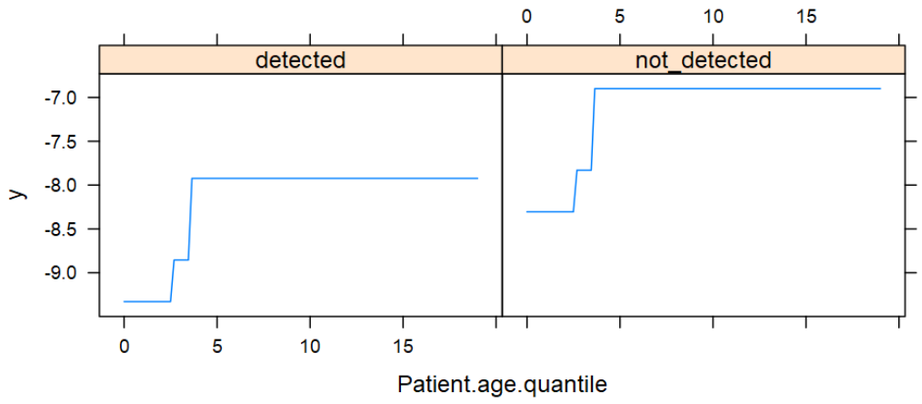
FACTOR 2

Influenza.B



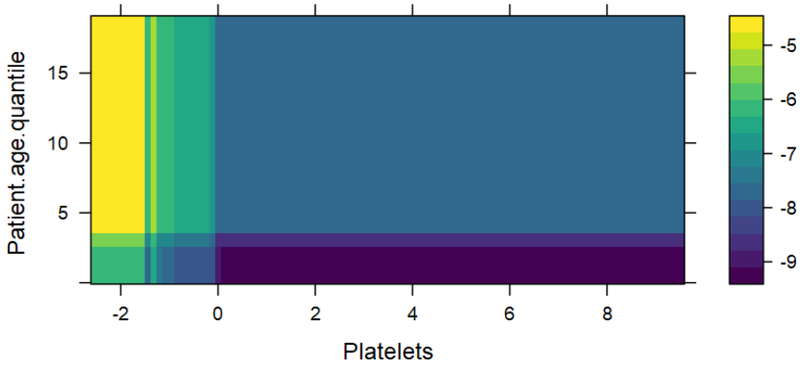
FACTOR 3

Inf.A.H1N1.2009



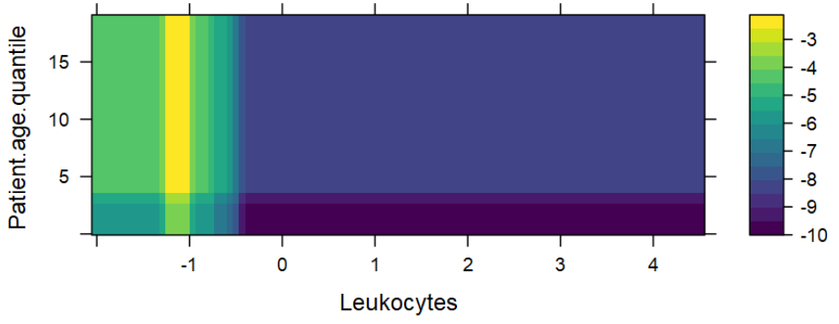
FACTOR 4

Platelets

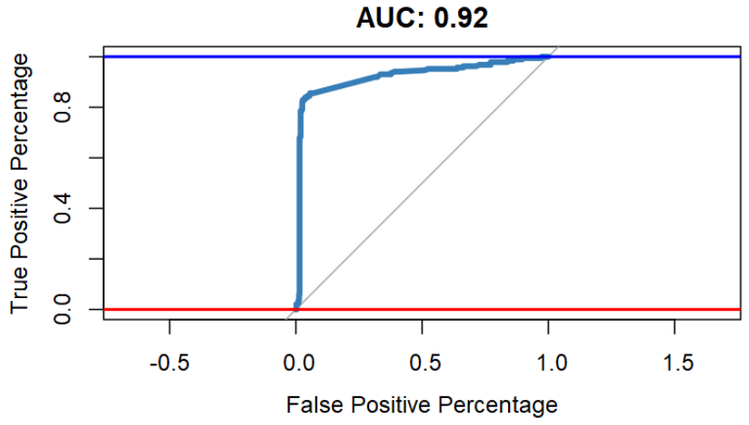


FACTOR 5

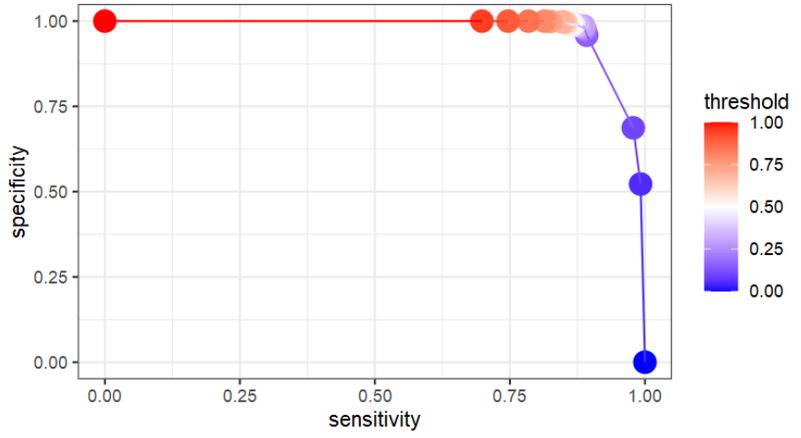
Leukocytes



ROC for 688 Test patients

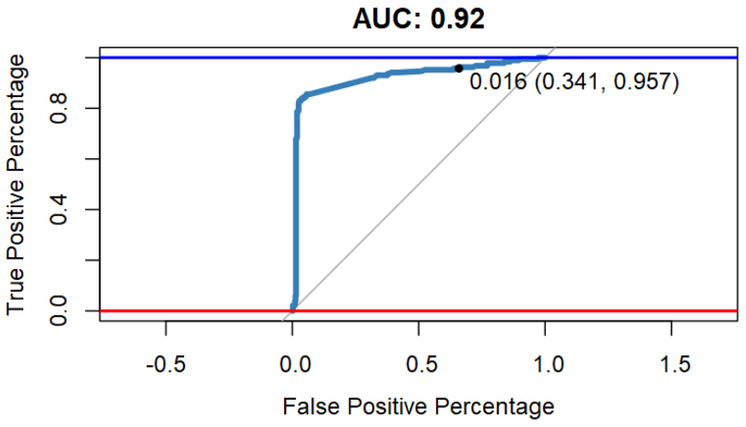


Test Threshold setting for optimal sensitivity



Ideal model is one that is well-balanced, i.e., one that has high sensitivity but it does not over-assign patients with positive labels

High Availability of Resources



Specificity levels above 98% can be achieved if the hospital were willing to consider a test that selects as many as 75% of the patients as likely to test positive

BALANCED SCENARIO

Ideal model is one that is well-balanced, i.e., one that has high sensitivity but it does not over-assign patients with positive labels