COVID-19 and its Clinical Spectrum Prediction using Bayesian Networks
(Summary)

Sanchit Kumar, Boroju Arthik (ID_29)
(May, 2021)

ABSTRACT

In the recent covid outbreak, physicians had to predict the severity of patients' deteriorating health conditions and allocate the appropriate clinical resources to each patient. This was a challenge as the clinical resources, test equipment for SARS-CoV-2 are limited and health care professionals knew very little about how the virus and its mutations would affect the body to correctly gauge the patient's condition. So In this paper we propose a probabilistic predictive model based on Bayesian Networks, using which health care professionals can predict how probable a person is SARS-CoV-2 positive and also the probability of his admission into Regular ward, Semi-intensive Care and Intensive Care.

METHOD

Bayesian Networks are useful for decision makers to make sense out of complex information i.e, reason under uncertainty using a probabilistic approach.Formally, Bayesian Networks is a Joint Probability Distribution over a set of random varaibles.It is represented by a Directed Acyclic Graph, where nodes correspond to random variables and arcs correspond to the probabilistic dependence between them.
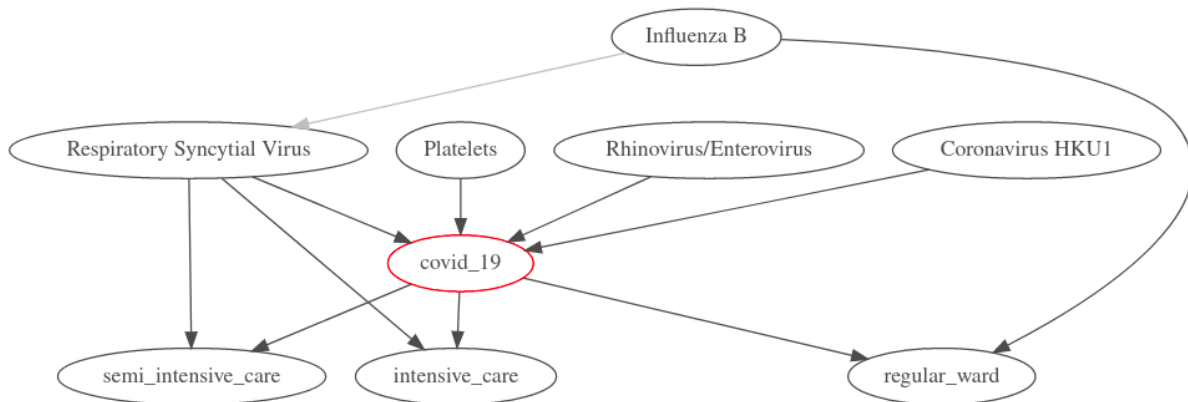

Fig 1 .Markov Blanket for Covid 19 (This image is only for Illustrative purposes*)***

In the above DAG the probability of a person having covid-19 given Influenza B, Respiratory Syncytial test result, Platelets test result, Rhinovirus test result, Coronavirus HKU1 can be calculated.
In the same manner, we can calculate the probability that a person would be allocated a

Regular Ward or semi-intensive ward or Intensive care ward based on the DAG obtained. See Fig 2. The conditional probability tables can be calculated using a combination of frequentist approach(learning from dataset) and explicitly encoding domain knowledge . In this paper, we have encoded compulsory arcs into the Bayesian Network like {CoronaVirus HKU1, covid_19}, {Respiratory Syncytial virus,covid_19}, etc. and we have avoided some direct arcs {Influenza B rapid test, covid_19}.

Fig 2. Purposed Bayesian Network

DATASET AND LIBRARY USED

The dataset is taken from a hospital in Israel, It has 5644 records each corresponding to a patient that was tested for SARS-CoV2. Along with the result of the covid-19 test we have the information of whether the person is admitted to Regular ward, semi-intensive or intensive care. It also has results of various other health tests like Bordetella pertussis, Adenovirus, CoronavirusNL63, etc. to name a few.

No two patients are the same, so doctors often need to prescribe different types of tests for each patient upon admission. Hence, our dataset contains a big amount of missing values, in some columns it went upto 99 %. As discussed earlier, We have dropped columns with more than 95 % missing values.

We used the pyAgrum library to do the analysis, it is a wrapper for c++ aGrUM library. The library is used for building applications using graphical models like Bayesian Network, Markov Decision processes.

EXPERIMENT RESULTS

From the proposed model, we can easily test the probability of a person likely to get admitted to ICU, SEMI-ICU and regular ward with particular comorbidities.
Although we can test on the basis of various comorbidities, here we are showing results for few comorbidities such as platelets, red blood cells , bordetella pertussis and covid 19 with set with respective categories for demonstration of the mode and how experiment results are promising.

1. **Regular Ward**
    This experiment is done with comorbidities i.e platelets, red blood cells ,bordetella pertussis, respiratory Syncytial Virus and covid 19 set as 3(category value), 3 (category value), detected, detected,detected respectively. The results show the particular person is less likely to get admitted to a regular ward.

		Regular ward	
COVID-19	**Influenza B**	**0**	**1**
0	**Not detected**	**0.9713**	**0.0287**
	detected	**0.9983**	**0.0017**
1	**Not detected**	**0.7495**	**0.2505**
	detected	**0.9286**	**0.0714**

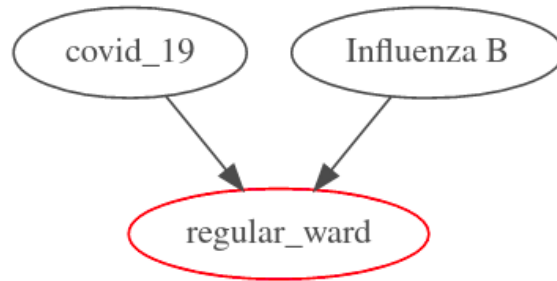**Table 1. Probability of getting admitted to regular ward (probability table)**

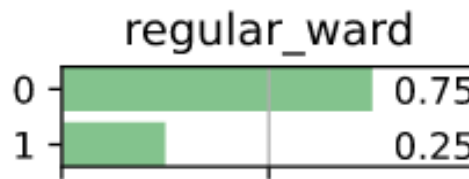**Fig 3. Markov blanket of regular ward**

Fig 4. Posterior probability of regular ward

2. Intensive Care Unit (ICU)

This experiment is done with comorbidities i.e platelets, red blood cells ,bordetella pertussis, respiratory Syncytial Virus and covid 19 set as 3(category value), 3 (category value), detected, detected,detected respectively. The results show the particular person is likely to get admitted to ICU when tested covid positive.

		Intensive Care Unit (ICU)	
COVID-19	**Respiratory Syncytial Virus**	**0**	**1**
0	Not detected	0.9851	0.0149
	detected	0.8396	0.1604
1	Not detected	0.9620	0.0380
	detected	0.5000	0.5000

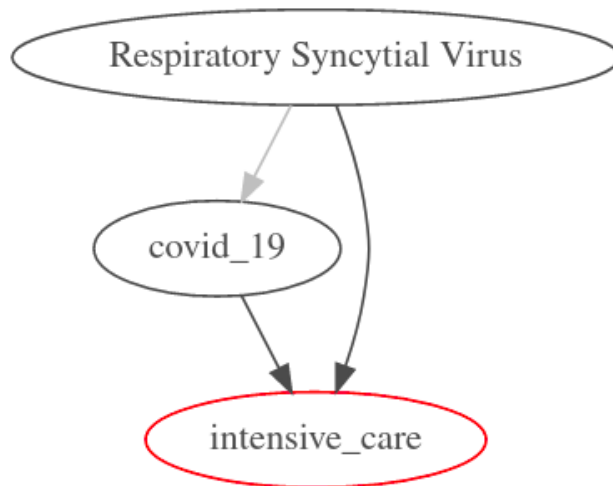**Table 2. Probability table of getting admitted to ICU**

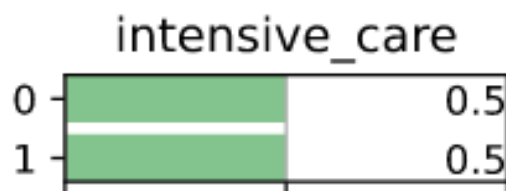**Fig 5. Markov blanket of intensive care unit**

**Fig 6. Posterior Probability of getting admitted to ICU
when person is covid19 positive**

The probability will get affected if any of the parameters will get changes. Let's see the posterior when tested covid19 negative.
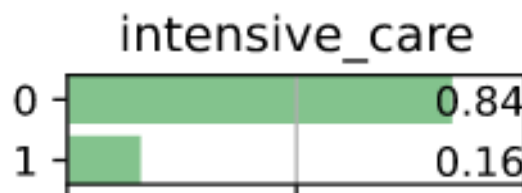

**Fig 7. Posterior Probability of getting admitted to ICU
when person is covid19 negative**

3. Semi-intensive Care Unit (ICU)

This experiment is done with comorbidities i.e platelets, red blood cells ,bordetella pertussis, respiratory Syncytial Virus and covid 19 set as 3(category value), 3 (category value), detected, detected,detected respectively. The results show the particular person

is likely to get admitted to semi-ICU.

COVID-19	Respiratory Syncytial Virus	Semi Intensive Care Unit (S-ICU)	
		0 (S-ICU not required)	1 (S-ICU not required)
0	Not detected	0.9783	0.0217
	detected	0.9620	0.0380
1	Not detected	0.8623	0.1377
	detected	0.5000	0.5000

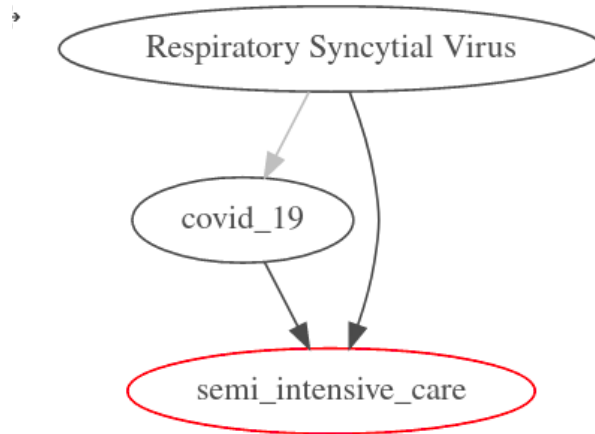**Table 3. Probability table of getting admitted to Semi-ICU**

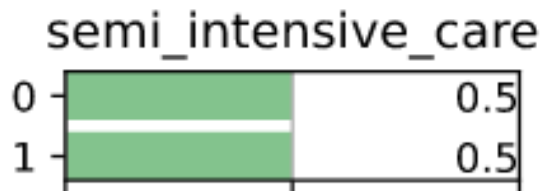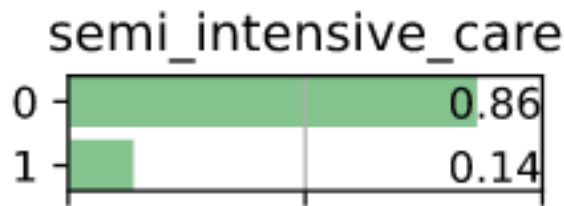**Fig 8. Markov blanket of semi intensive care unit**

Fig 9. Posterior Probability of getting admitted to Semi- ICU when person is covid19 positive

The probability will get affected if any of the parameters will get changes. Let's see the posterior when tested covid19 negative.

Fig 10. Posterior Probability of getting admitted toSemi- ICU when person is covid19 negative

The experimental results are much dependent on covid 19 dependency. The following figures show the probability calculation and dependence of covid 19 and markov blanket for that on the basis of the used dataset and proposed model.

Respiratory Syncytial Virus	Platelets	Rhinovirus/ Enterovirus	Coronavirus HK1	COVDI19	
				0 (negative)	1 (positive)
0 (not detected)	2	Not detected	Not detected	0.8880	0.1120
			detected	0.9997	0.0003
		detected	Not detected	0.9871	0.0129
			detected	0.9977	0.0023
	3	Not detected	Not detected	0.4375	0.5625
			detected	0.5000	0.5000
		detected	Not detected	0.9954	0.0046
			detected	0.5000	0.5000
	0	Not detected	Not detected	0.9999	0.0001
			detected	0.5000	0.5000
		detected	Not detected	0.9999	0.0046
			detected	0.9998	0.0002
	1	Not detected	Not detected	0.7494	0.2506

			detected	0.5000	0.5000
		detected	Not detected	0.9954	0.0046
			detected	0.5000	0.5000
1 (detected)	2	Not detected	Not detected	0.9998	0.0002
			detected	0.5000	0.5000
		detected	Not detected	0.9995	0.0005
			detected	0.5000	0.5000
	3	Not detected	Not detected	0.9977	0.0023
			detected	0.5000	0.5000
		detected	Not detected	0.9954	0.0046
			detected	0.5000	0.5000
	0	Not detected	Not detected	0.9977	0.0023
			detected	0.5000	0.5000
		detected	Not detected	0.5000	0.5000
			detected	0.5000	0.5000
	1	Not detected	Not detected	0.5000	0.5000
			detected	0.5000	0.5000
		detected	Not detected	0.5000	0.5000
			detected	0.5000	0.5000

Table 4. Probability table proposed covid 19 model

CONCLUSION

- From the proposed probability bayesian network, we can effectively predict the clinical spectrum of covid19 and effectively manage the resources and save the lives of the people during this pandemic.
- This model will allow us to predict the patient's need in the near future whether he/she will require the ICU, semi-ICU, general ward etc. facility on the basis of various

comorbidities.
- ● Patient can also use this model for their self analysis so they can become more careful for their actions on the basis of their risk category.

FUTURE WORK

Currently Bayesian networks work well with random variables which take discrete values. In the current model, we have discretized continuous random variables. In this process, Information is lost. We can improve this model for continuous random variables. We will look for a solution to handle the continuous variable in our future work.