International Journal of
Environmental Research and Public Health

MDPI

Article

Using Machine Learning Algorithms to Develop a Clinical Decision-Making Tool for COVID-19 Inpatients

Abhinav Vepa [1], Amer Saleem [1], Kambiz Rakhshan [2], Alireza Daneshkhah [3,*], Tabassom Sedighi [4], Shamarina Shohaimi [5,*], Amr Omar [1], Nader Salari [6], Omid Chatrabgoun [7], Diana Dharmaraj [1], Junaid Sami [1], Shital Parekh [1], Mohamed Ibrahim [1], Mohammed Raza [1], Poonam Kapila [1] and Prithwiraj Chakrabarti [1]

1 Milton Keynes University Hospital, Standing Way, Eaglestone, Milton Keynes MK6 5LD, UK; Vepa.abhinav@nhs.net (A.V.); Amer.Saleem@mkuh.nhs.uk (A.S.); Amr.Omar@mkuh.nhs.uk (A.O.); Diana.Dharmaraj@mkuh.nhs.uk (D.D.); MianJunaid.sami@mkuh.nhs.uk (J.S.); Shital.Parekh@mkuh.nhs.uk (S.P.); MohamedElbadri.Ibrahim@mkuh.nhs.uk (M.I.); Mohammed.Raza@mkuh.nhs.uk (M.R.); Poonam.Kapila@mkuh.nhs.uk (P.K.); Prithwiraj.Chakrabarti@mkuh.nhs.uk (P.C.)
2 Leeds Sustainability Institute, Leeds Beckett University, Leeds LS1 3HE, UK; k.rakhshanbabanari@leedsbeckett.ac.uk
3 Research Centre for Computational Science and Mathematical Modelling, Coventry University, Coventry CV1 5FB, UK
4 Centre for Environment and Agricultural Informatics, Cranfield University, Bedfordshire MK43 0AL, UK; t.sedighi@cranfield.ac.uk
5 Department of Biology, Faculty of Science, University Putra Malaysia, Serdang, Selangor 43400, Malaysia
6 Department of Biostatistics, School of Health, Kermanshah University of Medical Sciences, Kermanshah 6715847141, Iran; n.salari@kums.ac.ir
7 Faculty of Mathematical Sciences & Statistics, Malayer University, Malayer 6571995863, Iran; o.chatrabgoun@malayeru.ac.ir
* Correspondence: ac5916@coventry.ac.uk (A.D.); shamarina@upm.edu.my (S.S.)

Abstract: Background: Within the UK, COVID-19 has contributed towards over 103,000 deaths. Although multiple risk factors for COVID-19 have been identified, using this data to improve clinical care has proven challenging. The main aim of this study is to develop a reliable, multivariable predictive model for COVID-19 in-patient outcomes, thus enabling risk-stratification and earlier clinical decision-making. Methods: Anonymised data consisting of 44 independent predictor variables from 355 adults diagnosed with COVID-19, at a UK hospital, was manually extracted from electronic patient records for retrospective, case–control analysis. Primary outcomes included inpatient mortality, required ventilatory support, and duration of inpatient treatment. Pulmonary embolism sequala was the only secondary outcome. After balancing data, key variables were feature selected for each outcome using random forests. Predictive models were then learned and constructed using Bayesian networks. Results: The proposed probabilistic models were able to predict, using feature selected risk factors, the probability of the mentioned outcomes. Overall, our findings demonstrate reliable, multivariable, quantitative predictive models for four outcomes, which utilise readily available clinical information for COVID-19 adult inpatients. Further research is required to externally validate our models and demonstrate their utility as risk stratification and clinical decision-making tools.

Keywords: Bayesian network; COVID-19; SARS CoV; random forest; risk stratification; synthetic minority oversampling technique (SMOTE)

check for updates

Citation: Vepa, A.; Saleem, A.; Rakhshan, K.; Daneshkhah, A.; Sedighi, T.; Shohaimi, S.; Omar, A.; Salari, N.; Chatrabgoun, O.; Dharmaraj, D.; et al. Using Machine Learning Algorithms to Develop a Clinical Decision-Making Tool for COVID-19 Inpatients. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6228. https://doi.org/10.3390/ijerph18126228

Academic Editors: Alireza Daneshkhah, Amin Hosseinian-Far and Samer A. Kharroubi

Received: 18 April 2021
Accepted: 1 June 2021
Published: 9 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

On Thursday, the 5th of March 2020, within the UK, COVID-19 claimed the life of its first victim and has since contributed towards over 103,000 deaths [1,2]. The Office of National Statistics (ONS) has since issued statements, based on population data, in conjunction with the National Health Service (NHS), indicating an increased risk of mortality from COVID-19 amongst poorer socioeconomic groups, Black and Minority Ethnics

(BAME), males and the elderly [2–5]. In addition to demographics, various biochemical parameters and co-morbidities, such as obesity, diabetes, hypertension, chronic obstructive pulmonary disease (COPD) and malignancy, have been identified as risk factors for poor COVID-19 outcomes [6–9]. However, using this vast data to improve clinical care has proven challenging. One particular challenge that remains is relatively quantifying the impact of various prognostic indicators upon COVID-19 outcomes, especially whilst using combinations of variables, in order to assist clinical decision-making and risk stratification.

Due to the limited nature of healthcare resources, such as hospital beds and ventilators, clinicians are often faced with difficult decisions where they must ration resources between patients, often having ethical implications [10–12]. Currently, clinicians are allocating healthcare resources to COVID-19 patients semi-quantitatively, and often as a response to clinical deterioration. Various risk stratification models have been described in the literature such as the 4C tool [13], but are currently not being used clinically due to criticism in recent systematic reviews [14,15]. Some of the key problems with existing risk-stratification tools are unclear methodologies, the exclusion of patients diagnosed with COVID-19 using Computed Tomography (CT) imaging but with negative Real Time-Polymerase Chain Reaction (RT-PCR) nasopharyngeal swabs, small sample sizes, many patients not reaching a study outcome, automated data extraction relying on clinical coding and many studies only exploring inpatient mortality as a primary outcome. In addition, many predictive models have been developed using patient data from other parts of the world, which may not be generalizable to the UK population due to patient factors, hospital factors and virus factors. Finally, only a small selection of risk-stratification tools analysed a wide host of independent variables including vital observations, biochemical markers, demographics and co-morbidities.

The multivariate predictive model showcased in this study uses Bayesian Networks (BNs), which have received increasing attention during the last two decades [16,17] for their efficacy in tackling challenging and complex problems whilst also aiding in making decisions under uncertainty [18]. The ever-increasing volumes of health data has created potential for developing new knowledge that could improve clinical practise and patient care. The BNs and other machine learning (ML) methods have been extensively utilised in a diverse range of health topics from genomics [19–21] to treatment selection, and outcome, prognosis and prediction [22]. A compelling advantage of BNs over other suitable data-driven methods is that they do not explicitly need massively large datasets. Furthermore, BNs can combine the elicited knowledge of experts in circumstances where data are limited, and still produce meaningful and accurate decision-support systems [23–26].

This paper seeks to develop a quantitative tool to aid risk-stratification, and earlier clinical decision-making for adult COVID-19 inpatients by benefiting from the properties of BNs, including making reliable predictions, and being robust in making decisions under various sources of uncertainties in data.

2. Materials and Methods

2.1. Study Design and Setting

This retrospective case–control study was conducted at Milton Keynes University Hospital (MKUH), which is a medium-sized, 550 bed, district general hospital in the United Kingdom. Data was collected during the routine clinical care of patients for auditing purposes, and upon receiving Health Research Authority (HRA) approval, the anonymous data was then also used for research purposes. The study aimed to follow the Transparent Reporting of a multivariable prediction model for the Individual Prediction of Diagnosis (TRIPOD) checklist [27] and was conducted according to a pre-defined study protocol.

2.2. Study Population

Adult patients diagnosed with positive RT-PCR nasopharyngeal swabs or CT scans with changes suggestive of COVID-19 [28], between 01/03/2020 (date of first COVID-19 patient diagnosis) and 22/04/2020 (date of initiating independent predictor variable data

collection) at MKUH, were included in this study. Sixty-nine patients were excluded, which is shown below in Table 1, to produce a final *n* number of 355. The sample size was determined by using the maximum number of COVID-19 patients diagnosed during the study period.

Table 1. Patient selection process.

Sample Population	
Patients diagnosed with COVID-19 between 01/03/2020 and 22/04/2020, at Milton Keynes University Hospital (*n* = 424)	
Inclusion Criteria	**Exclusion Criteria**
1. Patients diagnosed with at least 1 positive RT-PCR Nasopharyngeal swab	1. Patients diagnosed in the Outpatient setting
2. Patients diagnosed with CT scan changes consistent with COVID-19 [28]	2. Staff Members who were diagnosed via Occupational Health, and who did not receive a formal medical assessment
3. Age 18 years and above	
Final Study Participant Number (*n* = 355)	

2.3. Data Collection

The hospital Picture Archiving and Communication System (PACS) was searched to get the details of the CT scan reports of patients with suspected COVID-19 changes, from 01/03/2020 until 22/04/2020. Reports dictated by a consultant radiologist, and CT scan images where required, were screened for all patients who had changes suggestive of COVID-19 [28]. The radiologically positive cases were included in the study. A record of all the COVID-19 RT-PCR positive swabs was obtained from the microbiology department. After removing the duplicates, the CT positive and RT-PCR swab positive cases were populated to a Microsoft Excel spreadsheet. Further patient data from the hospital Electronic Patient Record System (EPR), was collected in accordance with data protection and Good Clinical Practise (GCP) guidelines, on a hospital computer, by a team of physicians. Specific instructions were issued to the team of physicians to use during the collection of data to ensure homogenous, standardised interpretation of data from EPR. Healthcare staff who had historically recorded patient information on the EPR during clinical assessment were, of course at the time, blind to the outcomes and hypotheses of this study. All data was checked for systematic error by at least 1 other physician. After data collection, data was fully anonymised.

2.4. Independent Predictor Variables

Independent predictor variables were selected for inclusion in this study a priori based on three criteria; (i) having a postulated role for influencing COVID-19 severity based on surrounding literature, (ii) values expected to be available for at least one-third of study participants and (iii) values being collected during the routine care of study participants. Forty-four independent predictor variables were used for analysis in this study, which are shown and defined in Supplementary Table S1. Patient characteristic information is shown in Supplementary Table S2.

2.5. Outcomes

All patients were either discharged or deceased and thus achieved all four outcomes; (i) inpatient mortality (IPD), (ii) duration of COVID-19 treatment (ADT), and (iii) maximum level of oxygen or ventilatory support during inpatient stay (MOoVS), which was divided into 4 categories: (A) requiring room air or non-high-flow oxygen, (B) requiring high-flow oxygen defined as using a venturi mask, (C) requiring non-invasive ventilation (continuous positive airway pressure (CPAP) or bi-level positive airway pressure (BiPAP)) and (D) requiring intubation. A new radiologically confirmed diagnosis of pulmonary embolism during inpatient stay (NCPE) was the fourth outcome. The outcomes were selected for

this study a priori based on 3 criteria: (i) all patients will be able to reach one of the pre-defined study outcomes, (ii) the pre-defined study outcomes are representative of COVID-19 severity based on the surrounding literature at the time of study design and (iii) the pre-defined study outcomes will involve data collected in the routine care of study participants. The follow-up period was defined as 2 months to give all patients ample time to achieve a study outcome prior to outcome data collection. During the above-mentioned data collection time-period, if a patient was admitted more than once, and both times were for COVID-19 related reasons within 5 days of each other, this was counted as a failed discharge and thus 1 admission. Days spent in hospital for social reasons or alternative diagnoses prior to developing COVID-19 were subtracted from the duration of inpatient stay to derive the duration of inpatient treatment outcome. The day of COVID-19 clinical presentation was retrospectively determined by physicians during data collection, after careful analysis of the patient notes, in order to determine the first day during the inpatient stay where COVID-19 was diagnosed clinically, based on the full repertoire of available clinical information.

2.6. Data Analysis

2.6.1. Missing Values

In this study, no outcome (or dependent) variable was missing, but there were several independent variables with high numbers of missing values. The biochemistry features, including max CRP levels at the different days were among the independent random variables with the highest percentages of missing values ranging from 40% to 70%. However, limited numbers of missing values could be observed on the rest of independent variables. Due to the high number of missing values in many of the independent random variables, the multiple imputation (MI) technique was selected as the most suitable technique to estimate the missing values in the dataset.

2.6.2. Balancing Outcomes

After converting the continuous and multi-scale discrete independent and dependent random variables into the categorical ones using the discretisation method [29], and estimating the missing values using the MI technique [30,31], it was observed that the resulting outcome variables suffer from a considerable imbalance. The predictions derived from fitting a suitable statistical model to the imbalanced datasets, whereby one class is dominant, would be inherently biased towards the dominant class, thus decreasing the reliability of the predictions made by the models [32–34].

In this study, the authors used the Synthetic Minority Oversampling Technique (SMOTE) to overcome the imbalance in the dataset [35]. The advantage of this technique over other oversampling methods is that it decreases the imbalance in a dataset by synthetically creating new examples of the minority class, and not duplicating them [35,36]. The authors applied the SMOTE on the entire dataset, in concordance with the surrounding literature [37–40].

2.6.3. Feature Selection

One essential stage in the development of predictive models using supervised Machine Learning (ML) techniques is feature selection, which includes identifying and choosing the best combination of independent variables in a dataset for efficient and optimum analysis of the problem at hand [32,33].

For its feature selection, this study adopts the recursive feature elimination (RFE) method, which is a backward variable selection wrapper technique [40]. For this purpose, the authors computed the RFE method in R (version 4.0.2) using the random forest (RF) function embedded in the Caret package [41,42]. In this study, the performance of the wrappers is assessed using k-fold Cross-Validation (k = 10), which repeats five times. The result of the feature selection using the RFE method for each of these responses is shown below in Table 2. In this table, the numbers against each variable for the corresponding

response indicates the predictive importance of that factor. The RFE method is a multi-step process. Firstly, the dataset is randomly split into 70% training and 30% testing using the validation set approach [35]. Subsequently, using the training data, a predictive model containing all features is developed based on the random forest method. This model then ranks the features based on a measure of importance. The RFE method then eliminates the least important feature, develops a new model based on a smaller number of independent variables, and re-ranks the remaining predictors [40]. RFE identifies two parameters: the number of subsets to evaluate and the number of predictors in each of the subsets. For each subset, the process of eliminating the least-important features continues until it reaches a determined subset size. Eventually, RFE compares the predictive performance of all subsets and determines the best subset size with the best accuracy [40].

Table 2. Feature selection results for four different outcomes; IPD, ADT, NCPE and MOoVS.

Predictor	RFE (NCPE)	RFE (MOoVS)	RFE (ADT)	RFE (IPD)
Age	2	1	7	1
Gender (Ge)	7	17	17	33
Ethnicity	17	5	24	4
Oxygen Saturations (OS)	16	2	9	2
Respiratory Rate (BPM)	19	12	26	9
Temperature	6	10	19	35
Obesity	11	8	11	28
Previous Venous Thromboembolism (PVTE)	Rejected	33	33	36
Chronic Obstructive Pulmonary Disease (COPD)	Rejected	37	37	37
Bronchiectasis	Rejected		41	31
Asthma	21	27	34	10
Interstitial Lung Disease (ILD)	Rejected	21	40	38
Lung Cancer (LC)	Rejected	41	38	39
Diabetes Mellitus (DM)	29	16	27	21
Hypertension (HTN)	14	26	8	25
Ischaemic Heart Disease (IHD)	31	28	28	16
Chronic Kidney Disease (CKD)	32	31	31	8
Non-steroidal anti-inflammatory drugs (ANNC)	33	38	25	15
Anticoagulant	23	35	29	17
Long-Term Antibiotic (LTA)	Rejected	34	36	30
Long Term Oral Steroid (LTO)	Rejected	39	42	42
Immunosuppressants (ISES)	Rejected	32	39	32
Oral NSAIDs (ONS)	Rejected	40	32	41
Angiotensin Converting Enzyme Inhibitors (ACEI)	28	36	30	40
Angiotensin Receptor Blockers (ARBB)	27	29	35	27
CT imaging severity of COVID-19 related changes (UoB)	1	4	2	23
COVID-19 related Chest X-ray changes (CCX)	30	7	10	11
Lactate (LDP)	12	25	21	20
Lymphocytes (LyDP)	4	23	16	18
Neutrophils (NDP)	5	18	15	26
Albumin (MADA)	3	6	1	6

Table 2. *Cont.*

Predictor	RFE (NCPE)	RFE (MOoVS)	RFE (ADT)	RFE (IPD)
Ferritin	24	20	23	24
D-Dimer (MDD)	8	11	6	7
C-Reactive Protein (CRP) Day 0	18	13	3	19
CRP Day 1–2 (MCRP1)	13	22	4	12
CRP Day 3–4 (MCRP3)	20	19	14	22
CRP Day 5–6 (MCRP5)	10	15	12	5
CRP Day 7–8 (MCRP7)	9	3	5	3
CRP Day 9–10 (MCRP9)	22	14	13	14
CRP Day 11–12 (MCRP11)	26	9	18	29
CRP Day 13–14 (MCRP13)	15	24	20	34
CRP Day 15–20 (MCRP15)	25	30	22	13

Figure 1 shows the performance of the RFE method based on the ranks of the variables. Table 2 and Figures 1–4 both show that for each outcome, there is a specific combination of independent variables that produce the highest predictive performance among all other possible combination of variables for the selected outcome.

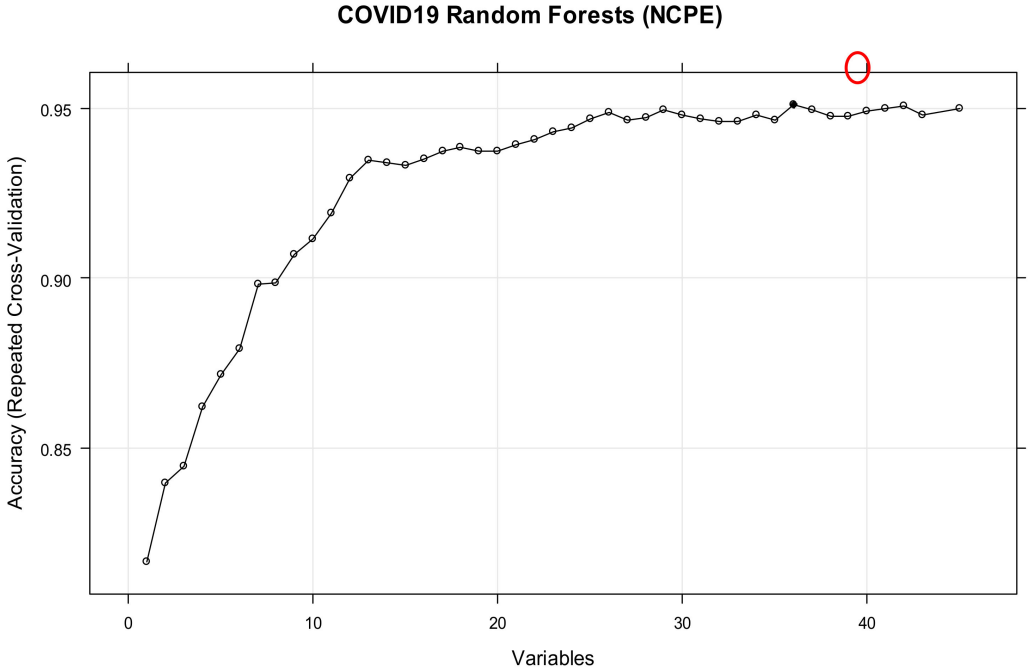

Figure 1. Performance of the RFE based on the ranks of the features of NCPE. The red circle shows the maximum achievable performance based on the best combination of variables.

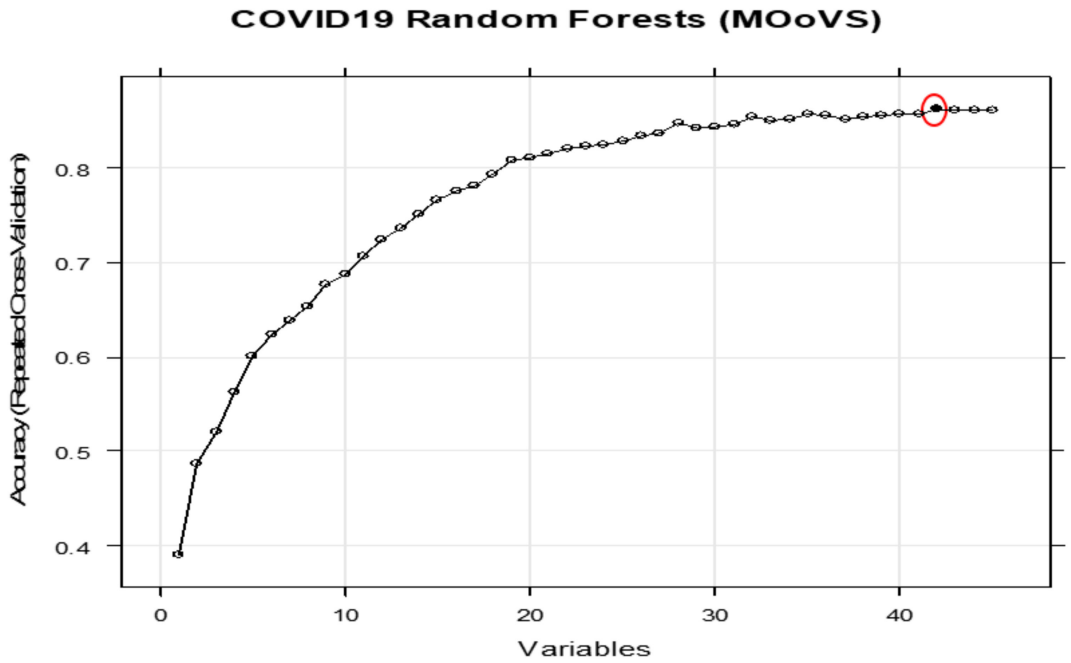

Figure 2. Performance of the RFE based on the ranks of the features of MOoVS. The red circle shows the maximum achievable performance based on the best combination of variables.

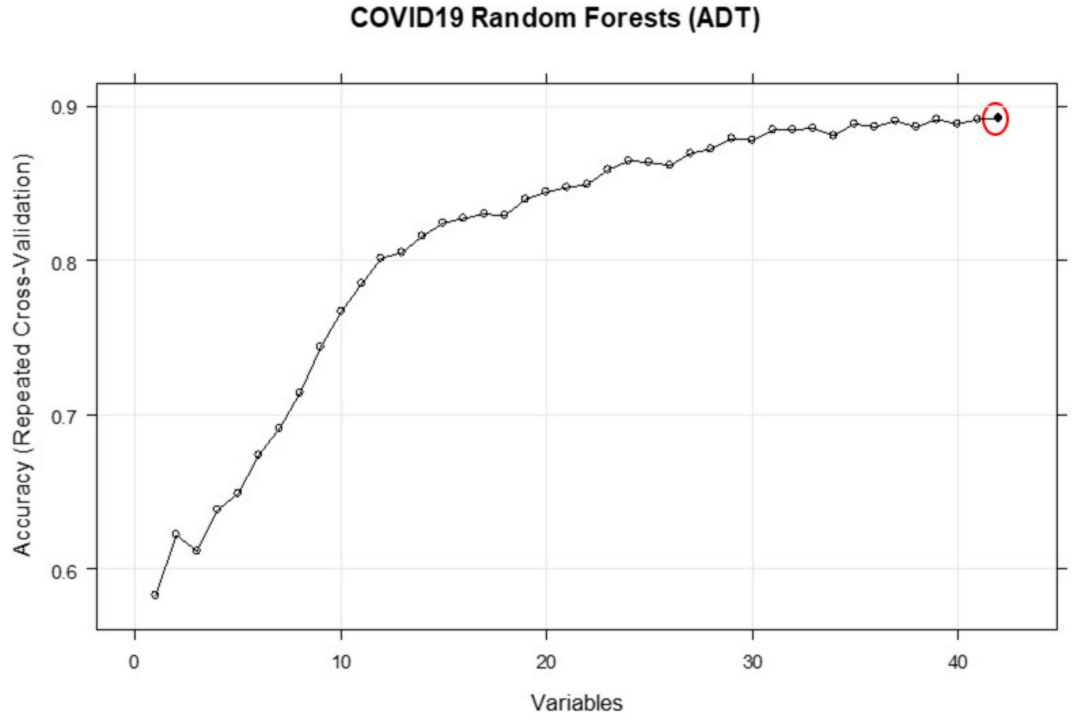

Figure 3. Performance of the RFE based on the ranks of the features of ADT. The red circle shows the maximum achievable performance based on the best combination of variables.

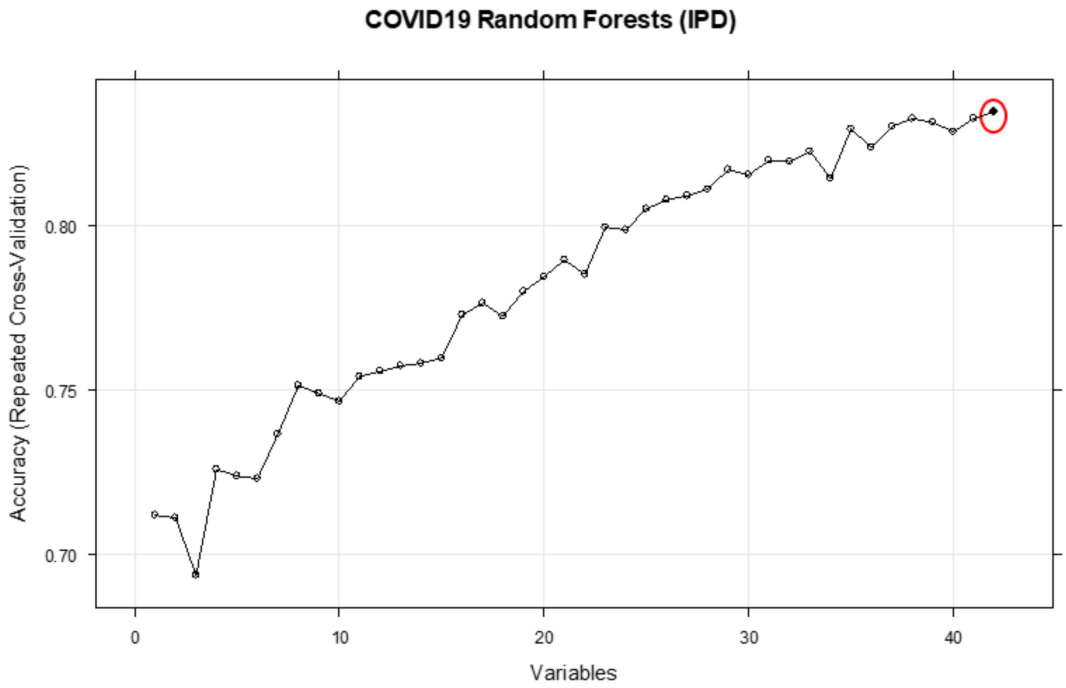

Figure 4. Performance of the RFE based on the ranks of the features of IPD. The red circle shows the maximum achievable performance based on the best combination of variables.

2.7. Bayesian Network Modelling

A Bayesian network (BN) is a probabilistic graphical model that is used to represent knowledge about an uncertain domain [43]. Applications of BN methods are found in a growing number of disciplines and policies [44]. BN learning consists of two general steps: (i) Finding Directed Acyclic Graphs (DAG), which illustrates the inter dependency between the variables/nodes and is denoted by G, and (ii) Finding Conditional Probability Tables (CPT) for each node given the values of its parents on the learned network structure G.

Finding the best DAG is the crucial step in BN design. Construction of a graph to describe a BN is commonly achieved based on probabilistic methods, which utilise databases of records [45], such as the search and score approach. In this approach, a search through the space of possible DAGs is performed to find the best DAG. The number of DAGs, f(p), as a function of the number of nodes, p, grows exponentially with p [46].

The BN structure learned from the data only for IPD based on the feature selected factors affecting IPD is shown in Figure 5. This network structure was learned from the completed data by evaluating the best model out of various score-based or constraints-based methods [47]. In particular, the BN shown in Figure 5 was selected by employing hill-climbing (or hc) algorithm and benchmarked with other suitable learning algorithms (e.g., Tabu Search or simply tabu) available in "bnlearn" library in R package. We then used the cross-validation, which is a standard way to obtain unbiased estimates of a model's goodness of fit to select the best models out of the learned networks using the learning algorithms mentioned above. The cross-validation method used in this paper is 5-fold cross-validation that can be simply computed using "bn.cv" function in "bnlearn" package. This function provides us with log-likelihood loss, its standard deviation and BIC. Both "tabu" and "hc" algorithms suggested the network structure illustrated in Figure 5, as the best networks learn from data, with the same BIC.

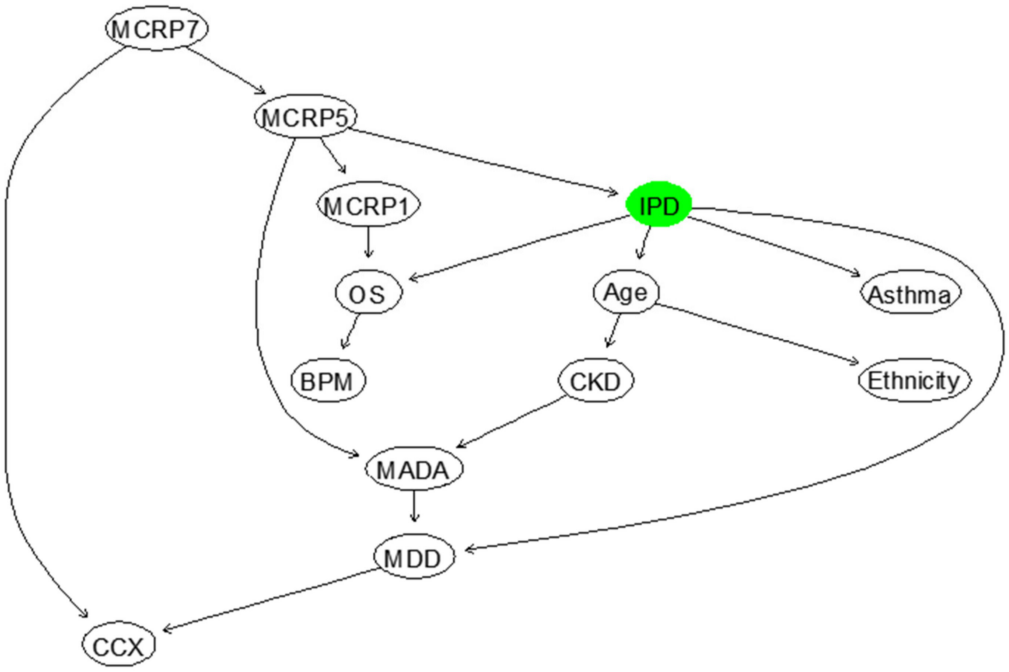

Figure 5. The BN that is fully learned from data to model "IPD" in terms of other relevant factors.

From this model shown in Figure 5, it is evident that the way that several independent variables affect IPD is incorrect. Therefore, it was important to discuss the resulting BN model, illustrated in Figure 5 with the domain medical experts, and consequently revise this BN by considering experts' opinions. The revised BN model learned based on the combination of data and expert opinions, whilst also validated using several model diagnostic algorithms, such as k-fold cross validation, is illustrated in Figure 6. The computed BIC metric for the network shown in Figure 6 (−4190.4, equivalent to 9.03 of log-likelihood loss value and 0.008 standard deviation of the loss) is smaller than BIC of the model shown in Figure 5, which was computed to be, −4000.3 (equivalent to 8.69 of log-likelihood loss value and 0.011 standard deviation of the loss).

In the BN model proposed for modelling IPD, the strength of the link, as well as the associated uncertainty, is captured using probabilities and statistical distributions, which are estimated or derived based on the observed data. Figure 7 shows the learned BN with the estimated marginal probabilities shown on each node. In this BN, three nodes (Age, Chronic Kidney Disease (CKD), and Ethnicity) are considered as root nodes, and their parameters are learned by estimating these probabilities using the maximum likelihood method or Bayes estimate. The estimated marginal and conditional probabilities for the variables can be updated in the light of new evidence or data using a statistical algorithm known as the Bayes rule [45]. Hence, the BN can compute the probability of surviving or dying due to COVID-19 based on the different combination of the parent nodes, including Age, the minimum Albumin level during admission (MADA), and the mean C-Reactive Protein (CRP) level during days 7–8 since clinical presentation of COVID-19 (MCRP7).

Using the same methods, BNs were constructed for the other 3 outcomes and are shown in Supplementary Figures S1–S6.

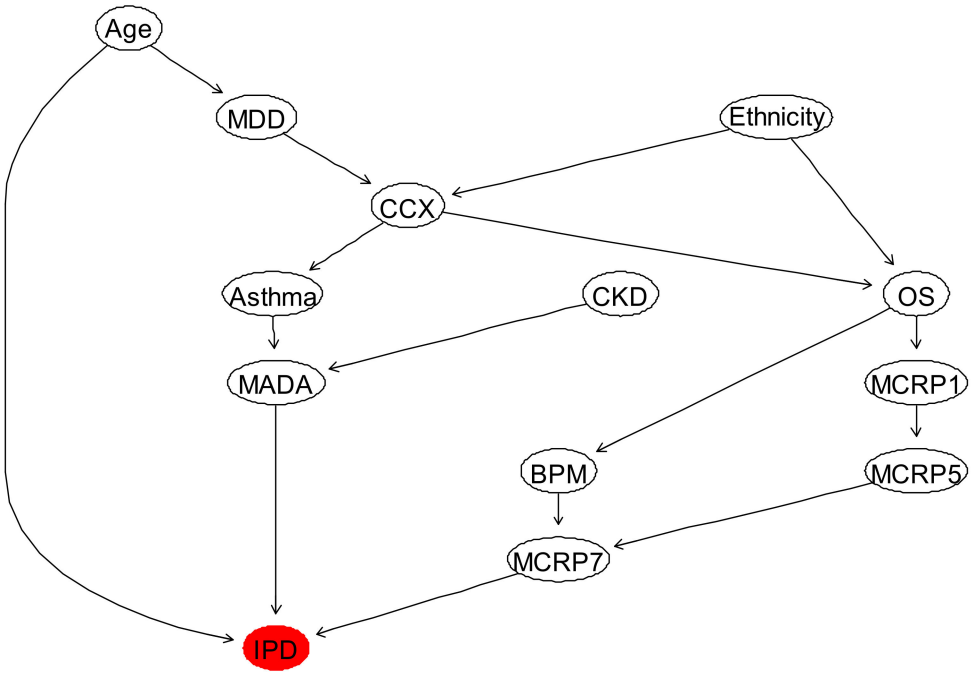

Figure 6. The BN learned by eliciting the domain expert combined with the (balanced and completed) data.

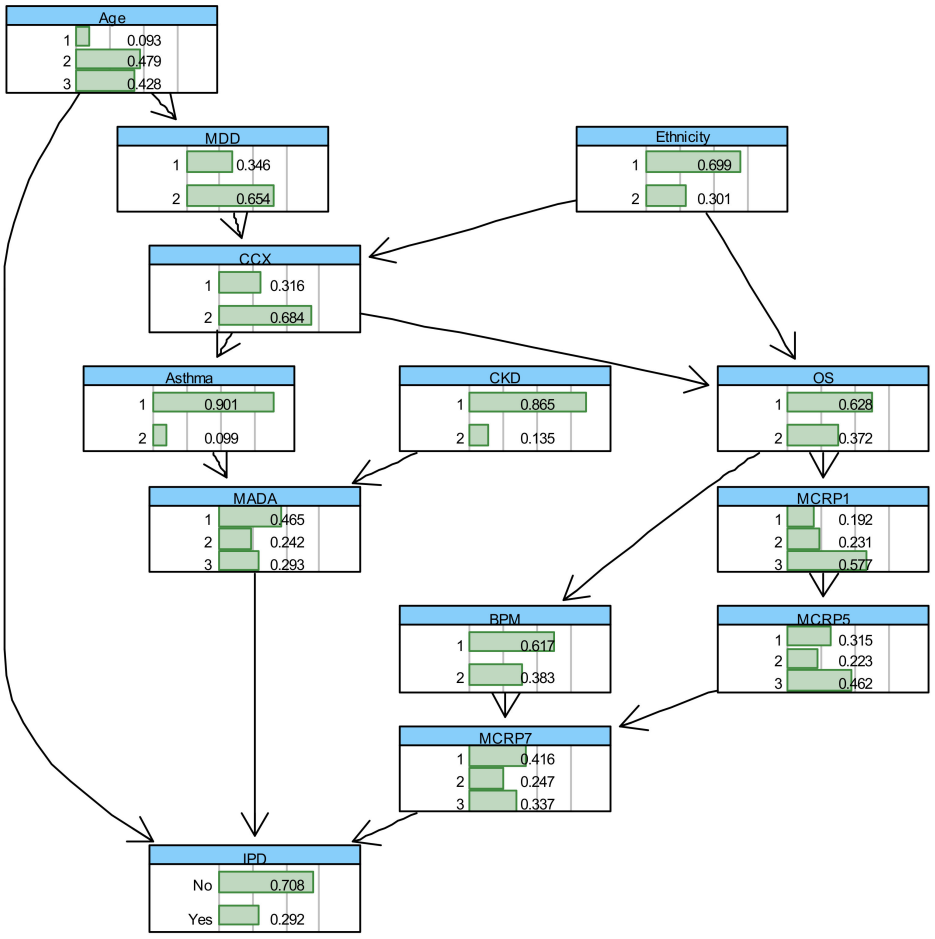

Figure 7. The BN with conditional probability tables (CPT) learned for "IPD" outcome based on the combined elicited domain expert opinions with the (balanced) data.

3. Results

3.1. Inpatient Mortality (IPD)

The results section describes the performance metrics of the predictive models constructed in this study, as well as the conditional probabilities of outcome occurrence given different exemplary combinations of independent factor variables.

Based on the BN model shown in Figure 7, the probabilities of IPD and survival of COVID-19 inpatients at different age groups can be computed, as shown in Supplementary Table S3. It highlights that the death rate of COVID-19 patients with ≥70 years is five times larger than patients' with ≤40 years. These probabilities can be updated by observing more evidence about the states of other influencing variables. Table 3 shows the conditional probabilities of IPD given MADA (1 = <30, 2 = 30–35, 3 = 35<), MCRP7 (1 = <50, 2 = 51–100, 3 = 100<), and Age in years (1 = <40, 2 = 40–70, 3 = 70<). It can be concluded that for the patients in the first age group (<40 years) if the MADA is above 35 g/L, they will survive COVID-19 regardless of MCRP7. For a patient in this age group and with the MADA level less than 30 g/L, they would more likely survive if their MCRP7 were less than 50 mg/L. This means that when the MADA level is less than 30 g/L, patients are at a particularly high risk, especially if their MCRP7 level is above 50 mg/L. Interestingly, similar patterns can be found for the patients aged > 70 years old, but the corresponding survival probabilities are considerably lower.

Table 3. The conditional probability of IPD given different configurations of the parent nodes.

Risk Factor (MADA, MCRP7, Age)	Probability of Inpatient Mortality
(3, 1, 1)	0
(2, 1, 2)	0
(3, 3, 1)	0
(2, 2, 2)	0.20
(1, 1, 1)	0.33
(1, 3, 1)	0.397
(3, 1, 3)	0.417
(1, 1, 3)	0.513
(1, 2, 2)	0.594
(3, 3, 3)	0.813
(1, 3, 3)	0.866

Overall, these results indicate how low albumin (reflective of malnourishment, as well as infection owing to its negative acute phase protein property), high CRP, and old age correlate with inpatient mortality in an additive manner.

Oxygen Saturations (OS) at the time of presentation with COVID-19 were also associated with mortality. As OS decreases from >92% to <92%, the risk of mortality increases by 2%, suggesting that there is a negative association between these two variables, as shown in Supplementary Table S4. We expected this association to be stronger, but found it difficult to measure the oxygen saturations during admission at any other time, or as an average, as it would be heavily affected by the level of oxygen therapy being received by the patient. Although these findings are useful, it is also fascinating to observe how OS could jointly, with other influencing independent variables, affect the risk of inpatient mortality, as shown below in Table 4, where OS in % (2 = <91, 1 = 92<), CCX (1 = 'No', 2 = 'Yes'), Ethnicity (1 = 'Caucasian', 2 = 'Non-Caucasian'), and Age in years (1 = <40, 3 = 70<). The illustrated results in this table suggest that the risk of inpatient mortality is elevated for patients with reduced oxygen saturations and older patients. Ethnicity seems to increase the risk of death in patients 70< years, which is concordant with surrounding

literature [3]; however, the results were not true for younger patients. This may have been due to a skewed population demographic, whereby older patients tended to be Caucasian and younger patients reflected a more multicultural demographic. Changes on chest X-ray (CCX) did not seem to significantly affect the risk of IPD, perhaps because the presence of changes is more likely an indicator of the time-point that an individual is along in their COVID-19 infection rather than an indicator of severity.

Table 4. The conditional probability of IPD given different configurations of OS, Ethnicity, CCX and Age.

Risk Factor (OS, CCX, Ethnicity, Age)	Probability of Inpatient Mortality
Age < 40	
(1, 2, 2, 1)	0.1375
(1, 2, 1, 1)	0.1376
(2, 1, 1, 1)	0.1387
(1, 1, 1, 1)	0.1391
(2, 2, 1, 1)	0.1395
(1, 1, 2, 1)	0.1401
(2, 1, 2, 1)	0.1407
(2, 2, 2, 1)	0.1417
Age < 70	
(1, 2, 2, 3)	0.6755
(1, 2, 1, 3)	0.6756
(1, 1, 1, 3)	0.6761
(1, 1, 2, 3)	0.6764
(2, 1, 2, 3)	0.6927
(2, 1, 1, 3)	0.6944
(2, 2, 1, 3)	0.6944
(2, 2, 2, 3)	0.6946

The next important research question is how the trend in CRP levels during the clinical course of a COVID-19 infection can be incorporated and evaluated using an appropriate model. This is because CRP levels can often correlate with infection severity, with a small associated lag time. Therefore, the trend in CRP is clinically useful for predicting what will happen to the patient. For example, if the gradient between the latest two CRP variables were positive, it would indicate that the infection is getting worse, whereas if the gradient were negative, it would indicate infection resolution. To account for the gradient between the CRPs, a dynamic version of BN needs to be developed, which is not possible due to the lack of training data. However, the BN model, illustrated in Figure 7, can be used to compute the risk of inpatient mortality for different levels of CRP at the different days during the clinical course of COVID-19 infection, as shown below in Table 5. This table shows the conditional probability of IPD given different configurations of Mean CRP between days 1–2 since clinical COVID-19 presentation (MCRP1) (1 = <30, 2 = 31–100, 3 = 100<) and Mean CRP between days 7–8 since clinical COVID-19 presentation (MCRP7) (1 = <50, 2 = 50–100, 3 = 100<), Age in years (1 = <40, 3 = 70<) and Minimum Albumin During Admission (MADA) (1 = <30, 2 = 30–35, 3 = 35<). As shown, if the level of MADA is 35<, increases or decreases in CRP levels during the clinical course of COVID-19 infection will not impose a mortality risk in patients <40 years. However, in patients aged 70< years, any increase in CRP levels (mg/L) from days 1–2 to days 7–8, would significantly increase mortality risk.

Table 5. The conditional probability of IPD given different configurations of MADA, Age, MCRP1 and MCRP7.

Risk Factor (MADA, Age, MCRP1, MCRP7)	Probability of Inpatient Mortality
MADA > 35 and Age < 40 years	
(3, 1, 2, 1)—Small CRP Decrease	0
(3, 1, 3, 1)—Large CRP Decrease	0
(3, 1, 1, 2)—Small CRP Increase	0
(3, 1, 1, 3)—Large CRP Increase	0
MADA > 35 and Age < 70 years	
(3, 3, 2, 1)—Small CRP Decrease	0.418
(3, 3, 3, 1)—Large CRP Decrease	0.416
(3, 3, 1, 2)—Small CRP Increase	0.496
(3, 3, 1, 3)—Large CRP Increase	0.812
MADA < 30 and Age < 70 years	
(1, 3, 2, 1)—Small CRP Decrease	0.515
(1, 3, 3, 1)—Large CRP Decrease	0.513
(1, 3, 1, 2)—Small CRP Increase	0.734
(1, 3, 1, 3)—Large CRP Increase	0.865

Table 6 below indicates the promising Positive Predictive Value (PPV), Negative Predictive Value (NPV), Sensitivity, Specificity, Overall accuracy and F-Score, which are used to evaluate the predictive performance of the BN suggested to model IPD in Figure 7, in terms of the feature selected risk factors. The definitions and details of how these metrics can be computed are described in [35]. F1 Score is the Harmonic Mean between precision and recall. The F1 score (83.7%) and accuracy (84.1) of our BN model developed for IPD is high, despite the small dataset used to train and test our BN model. Eighty-two percent (PPV) of adult patients predicted to die as inpatients during clinical COVID-19 infection, by our model, will die. However, only 67.86% (NPV) of adult COVID-19 patients predicted to survive the inpatient admission will indeed survive. This indicates that our model may fail to predict inpatient death of a sub-set of adult COVID-19 patients, but we expect this to improve with a larger dataset, which also incorporates more variables such as socioeconomic factors.

Table 6. Summary of the predictive performance results of the BN model developed to model IPD as Illustrated in Figure 7.

Predictive Performance Metric	PPV	NPV	Specificity	Sensitivity	Overall Accuracy	F1-Score
BN for IPD	82%	67.86%	82.6%	85.7%	84.1%	83.7%

3.2. Duration of Inpatient Treatment for COVID-19 (ADT)

Understanding the simultaneous impact of MADA, obesity, and MCRP1 on the duration of COVID-19 treatment in patients is increasingly important to manage the growing, unrelenting pressures on hospitals and the NHS. Table 7 shows the probabilities of several important queries computed from the learned BN. In this table, ADT categories are '1' (<1 day), '2' (>2 days but < 3 days) and '3' (>3 days). MCRP1 has been divided into 3 categories: '1' (<50), '2' (51–100) and '3' (>100), and MADA has been divided into 3 categories: '1' (<30), '2' (30–35) and '3' (>35). As is evident from this table, the duration of treatment of 71% of non-obese COVID-19 patients with normal MADA levels (>35 g/L) and low MCRP1

(<30 mg/L) is up to 1 day. In addition, the treatment duration of 95% of the patients with the above characteristics would be less than 3 days. In comparison to the obese patients, it can be observed that this probability (i.e., probability that the duration of treatment is up to one day) will be reduced to 54.4% (with the same characteristics). On the other hand, the predicted probabilities for duration of treatment of the obese and non-obese patients, with very low levels of MADA (<30 g/L), regardless of levels of MCRP1, are not significantly different from each other. Furthermore, the model and results reported in Table 7 suggest that the levels of MCRP1 alone would not be adequate to accurately predict the probabilities of treatment duration of more than 3 days. These probabilities must be updated by adding more evidence about the levels of MCRP at other days. It would be straightforward to revise and update the BN by augmenting the other outcomes, for example "IPD", to understand what proportion of patients with a treatment duration <3 days may not survive. Overall, it can be observed that the COVID-19 treatment duration is higher for obese patients with high CRP levels and low Albumin levels.

Table 7. The heat-mapped, conditional probabilities of ADT given different configurations of Obesity, MADA and MCRPI.

Probability of ADT Given Obesity, MADA and MCRP7	MADA (3) and MCRP1 (1)	MADA (3) and MCRP1 (3)	MADA (1) and MCRP1 (1)	MADA (1) and MCRP1 (3)
BMI < 30 (Non-Obese patients)				
<1 day	71.2%	68.7%	10.5%	10.5%
>2 days but <3 days	23.7%	25%	32.7%	30.3%
>3 days	5.1%	6.3%	56.8%	59.2%
BMI > 30 (Obese patients)				
<1 day	54.4%	49.4%	13.2%	11%
>2 days but <3 days	40.7%	45.2%	34.6%	32.5%
>3 days	4.8%	5.4%	52.2%	56.5%

Several metrics to assess the predictive performance of the proposed BN for the three different categories of ADT are shown in Table 8.

Table 8. Summary of the predictive performance results of BN model developed to model ADT.

Predictive Performance Metrics of ADT Category	Balanced Accuracy	Sensitivity (Recall)	Specificity	Precision	Overall Accuracy	F1-Score
<1 day	74.8%	66.2%	83.5%	73.7%	61.5%	69.8%
>2 days but <3 days	60.6%	41.4%	79.8%	53.5%	61.5%	46.7%
>3 days	71.3%	76.9%	65.7%	57.9%	61.5%	66.1%

The computed sensitivity measures, which is the metric to evaluate the learned BN ability to predict true positives of each available category of ADT, suggest that "ADT > 3" days has the highest rate (77%), and '2 $\leq$ ADT < 3^' days has the lowest sensitivity rate (41.4%). We also compute specificity, which is the metric to evaluate the fitted BN ability to predict true negatives of each ADT category. The results suggest "ADT < 1" days (83.5%) and "ADT > 3" days are the categories with the highest and lowest specificity rates, respectively. The next important metric is F1-score that can be interpreted as a weighted average of the precision and sensitivity values, where an F1 score reaches its best value at 1 and worst value at 0. Since, the F1-score takes both false positives and false negatives into account; it will be usually more useful than accuracy, especially if the original test dataset has an uneven class distribution. The computed F1-scores for the ADT categories suggest promising accuracy for 'ADT < 1' days and 'ADT > 3' days.

3.3. Max Oxygen or Ventilatory Support (MOoVS)

It is of great important to understand how the right level of oxygen therapy or ventilatory support should be selected to enhance the survival rate and recovery speed of COVID-19 patients. Table 9 shows the conditional probabilities (as heat-mapped) of Max Oxygen or Ventilatory Support (MOoVS) given the different configurations of MADA (1 = '<30', 2 = '30–35', 3 = '35<'), OS ('1' = >92 and '2' < 92) and MCRP11 (1 = <100 and 2 = >100).

Table 9. The heat-mapped, conditional probabilities of MOoVS given the different configurations of MADA, OS and MCRP11.

Probability of MOoVS Given Category of OS, MADA and MCRP11	OS (1), MADA (3) and MCRP11 (1)	OS (1), MADA (1) and MCRP11 (2)	OS (1), MADA (1) and MCRP11 (1)	OS (2), MADA (3) and MCRP11 (1)	OS (2), MADA (1) and MCRP11 (1)	OS (2), MADA (1) and MCRP11 (2)
NHF	72.80%	39.30%	34.80%	20.10%	9.80%	1.80%
HF	12.50%	10.10%	38.90%	45.90%	25.40%	12.40%
CPN	14.70%	26.90%	18.60%	34%	39.30%	36.90%
ITU	0%	23.70%	7.70%	0%	25.50%	48.90%

Patients either required no high-flow oxygen (NHF), high-flow oxygen (HF), CPAP/NIV (CPN) or ITU admission (ITU). The illustrated results in this table suggest that most of the patients with the better health characteristics such as high MADA (>35 g/L), high OS (≥92%), and low MCRP11 (≤100 mg/L) are likely to only require no-high flow O2 (73%) or high flow O2 (12.5%). These patients are thus suitable for ward-based care and will not need ITU admission. On the contrary, when MADA and OS levels, respectively, drop to below 30 g/L and 91%, and MCRP11 level increases to over 100 mg/L, the need is increased for these patients to receive CPAP/NIV (37%) or ITU admission (49%).

The probabilities given in Table 9 can be significantly altered in the light of new evidence, such as patient age. The updated probabilities are shown below in Table 10. Table 10 shows the conditional probabilities (as heat-mapped,) of MOoVS given the different configurations of MADA, OS, MCRP11, and Age (1 = '<40 years' and 3 = '>70 years'). As mentioned above, patients required either NHF, HF, CPN or ITU.

Table 10. The heat-mapped, conditional probabilities of MOoVS given the different configurations of Age, MADA, OS and MCRP11.

Probability of MOoVS Given OS, MADA, MCRP11 and Age	OS (1), MADA (3), MCRP11 (1) and Age (1)	OS (1), MADA (1), MCRP11 (1) and Age (1)	OS (1), MADA (3), MCRP11 (1) and Age (3)	OS (1), MADA (1), MCRP11 (1) and Age (3)	OS (2), MADA (1), MCRP11 (2) and Age (1)	OS (2), MADA (1), MCRP11 (2), and Age (3)
NHF	92.60%	79.80%	61%	39.90%	0%	2.10%
HF	7.40%	20.20%	33.60%	53.90%	0%	22%
CPN	0%	0%	5.40%	3.10%	36.50%	75.90%
ITU	0%	0%	0%	3.10%	63.50%	0%

From this table, it can be concluded that all young patients with OS ≥ 92, MADA ≤ 30 and MCRP11 ≤ 100 require either NHF (80%) or HF (20%) to recover. If MADA levels of young patients increase to over 35 g/L, with the same levels of OS (≥92) and MCRP11 (<30), they will more likely need to use only NHF (93%) to recover. However, if the health of the patients starts to deteriorate, as MADA and OS levels, respectively, drop to <30 g/L and <92%; and MCRP11 level increases to over 100 mg/L, their need for HF, CPAP/NIV (CPN) or ITU will significantly increase depending to the age of patient. For the young patients, CPAP/NIV (36.5%) or ITU (63.5%) would be recommended. However, for patients > 70 years, either HF (22%) or CPAP/NIV (76%), and paradoxically not ITU (0%), would be recommended, usually because they are deemed unsuitable for ITU admission due to

the futility of ITU-based treatment relative to a younger patient. This reflects the rationing of healthcare resources that occurs in hospitals after difficult medico-ethical decisions.

Table 11 illustrates several metrics to assess predictive performance of the proposed BN for the different states of "MOoVS". The overall classification accuracy suggests that over 60% of cases have been correctly classified. Despite the small sample size of data and high percentages of missing values of the raw data, the computed overall accuracy (60.25%) is quite promising. The ability to predict true positives of each available category of MOoVS is measured by Recall, or Sensitivity, which suggests ITU admission has the highest rate (89%), and "CPN" has the lowest sensitivity rate (36%). We also compute specificity, whereby "CPN" (91%) and ITU admission (73%) are the categories with the highest and lowest rates, respectively. Since the original raw data has an uneven class distribution, F1-score as a weighted average of the precision and sensitivity values were calculated.

Table 11. Summary of the predictive performance results of BN developed to model MOoVS.

Predictive Performance Metrics	Balanced Accuracy	Recall (Sensitivity)	Specificity	Precision	Overall Accuracy	F1-Score
NHF	72.3%	56%	88.7%	70.8%	60.25 %	62.5%
HF	68.5%	61.2%	75.8%	51.3%	60.25 %	55.8%
CPN	63.6%	36.2%	91.1%	64.8%	60.25%	66.4%
ITU	80.7%	88.84%	72.7%	59.9%	60.25 %	71.5%

3.4. New Confirmed Pulmonary Embolism during Admission (NCPE)

Understanding the influence of various independent variables upon the conditional probability of having an NCPE is important for clinicians in guiding their use of thromboprophylaxis within the context of COVID-19. Our results confirmed that having bilateral COVID-19 changes on CT scan can increase the risk of NCPE from 13.9% to 72.4%. This suggests that more extensive ground-glass or consolidative changes on CT scan secondary to COVID-19 may be associated with NCPE secondary to COVID-19. Another well-known key predictor variable is the maximum D-dimer during admission (MDD), which is often an indicator of thrombosis. Our results indicated that having an MDD of 400<, as opposed to <400, increases the risk of NCPE from 34.6% to 54.5%. Interestingly, NCPE appeared to be more significantly influenced by the presence of MADA < 30 or bilateral ground-glass or consolidative CT scan changes secondary to COVID-19, rather than levels of MDD, as shown below in Table 12. This may be explained by the fact that many other conditions can also increase D-Dimer, such as disseminated intravascular coagulation, deep vein thrombosis and infection, which could thus be leading to high rates of PE false positives amongst COVID-19 patients. Furthermore, not all patients with raised D-dimers would have had CT-scans to investigate for PE, especially if it was deemed futile and the patient was palliative.

Table 12. The heat-mapped, conditional probabilities of NCPE given the different states of MDD, MADA and UoB. The results suggest that the presence of NCPE is more significantly influenced by the presence of bilateral ground-glass or consolidative CT scan changes.

Probability of NCPE Given UoB and MDD	Bilateral CT Changes and MDD < 400	Bilateral CT Changes and MDD > 400	Unilateral CT Changes and MDD < 400	Unilateral CT Changes and MDD > 400
No NCPE	27.20%	27.80%	86.50%	85.90%
NCPE	72.80%	72.20%	13.50%	14.10%
Probability of NCPE Given Categories of UoB and MADA	**Bilateral CT Changes and MADA < 30**	**Bilateral CT Changes and MADA > 35**	**Unilateral CT Changes and MADA < 30**	**Unilateral CT Changes and MADA > 35**
No NCPE	27%	51%	85.30%	98.60%
NCP	73%	49%	14.70%	1.40%

We finally investigated the impact of obesity alongside MADA and MCRP7 levels on NCPE in Table 13. It can be concluded that there is a strong association (70%) between the presence of NCPE and MADA < 30, as well as MCRP7 > 100, for the non-obese patients. If the patient is obese, MADA < 30 seems to be more influential in contributing towards the risk of NCPE, as opposed to MCRP7 > 100.

Table 13. The heat-mapped, conditional probabilities of NCPE given the different states of MADA, MCRP7 and Obesity.

Probability of NCPE Given Categories of MADA, MCRP7 and Obesity	MADA (1) and MCRP7 (1)	MADA (1) and MCRP7 (3)	MADA (3) and MCRP7 (1)	MADA (3) and MCRP7 (3)
BMI < 30 (Non-Obese Patients)				
No NCPE	44.20%	30.90%	63%	44.70%
NCPE	55.80%	69.10%	37%	55.30%
BMI > 30 (Obese Patients)				
No NCPE	68.20%	51.50%	94.30%	87.80%
NCPE	31.80%	48.50%	5.70%	12.20%

Various metrics, including PPV, NPV, sensitivity, specificity, overall accuracy, and F-Score, have been used to evaluate the predictive performance of the BN created to model NCPE, as shown in Table 14. The computed F1-score of almost 86% shows the classification prediction of the learned BN for NCPE is precise and robust. In addition, the PPV (83.7%), sensitivity (88%) and NPV (80.9%), which collectively represent the BN model's ability to predict inpatient NCPE, was high.

Table 14. Summary of predictive performance results of the BN learned for "NCPE". The computed F1-score of almost 86% shows the classification prediction of the learned BN for NCPE is precise and robust.

Predictive Performance Metric	PPV	NPV	Specificity	Sensitivity	Overall Accuracy	F1-Score
BN for IPD	83.7%	80.9%	75%	87.9%	82.7%	85.8%

4. Discussion

In this study, in addition to quantifying the significance of feature-selected risk factors, we showcase the use of Bayesian Networks to accurately predict four different COVID-19 inpatient outcomes, using different combinations of readily available clinical data, which serve as the independent predictor variables, whilst also accounting for interdependency between these variables.

Various COVID-19 prognostic indicators have been described in the literature, such as neutrophil:lymphocyte ratio, CRP, age, gender, ethnicity, oxygen saturation on admission, diabetes mellitus, hypertension, malignancy, obesity and COPD [48]. However, drawing insights from this information is impeded by the lack of clarity as to the relative influence each of these indicators has on mortality. In the clinical setting, patients often present with different combinations of these risk factors and biomarkers. Consequently, weighing them all up to allocate scarce healthcare resources can be challenging. This highlights a role for a predictive, quantitative risk-stratification tool. Our model has been constructed so that it can utilise data at first clinical presentation, for example, in the emergency department, but also after 3 and 7 days of inpatient treatment. This can allow clinicians to risk stratify at different time-points during an inpatient stay. In addition, our model can also predict the duration of inpatient COVID-19 treatment and maximum level of oxygen requirement that a patient may need during their inpatient stay. This may aid emergency physicians with the decision as to whether to admit a patient to hospital and avoid failed discharges, but also medical physicians with the decision as to whether to

refer to ITU prior to a patient's clinical deterioration. The predicted duration of inpatient COVID-19 treatment can be especially useful for bed managers in orchestrating patient flow, which is essential to prevent the growing problem of hospital-acquired COVID-19 secondary cross-contamination [49]. Furthermore, within the context of hospitals, which have reached their maximum ITU capacity, by using our predictive model to identify high-risk patients earlier on in their disease course, clinicians can transfer these high-risk patients to neighbouring hospitals prior to their clinical deterioration.

One of the major strengths of our study is that we predict four outcomes. This increases the amount of clinical utility that can be offered to guide clinical decision making. Secondly, this data set has incorporated 44 different variables from 355 patients who all received the primary study outcomes, reducing the risk of confounding errors and selection bias. Another strength of this study is that data extraction was conducted and checked manually by trained medical physicians rather than by using coding. It has been well documented that hospital clinical coding is still not entirely accurate within the UK [50,51], and therefore insights drawn from national databases may be prone to significant information bias and thus systematic error. Furthermore, it has now been estimated that the sensitivity of the COVID-19 RT-PCR nasopharyngeal swab is likely to be 50–75% [52–56]. This has created a huge global problem in diagnosing and identifying COVID-19 patients, especially because not all patients exhibit symptoms [57–61]. To overcome this issue, the cohort of patients with COVID-19 who had negative RT-PCR swabs but positive CT scan imaging (n = 55) were also included in this study.

The biggest limitation to our study is that since the study was conducted in a single centre, not only is the n number limited, but also the data only reflects the demographics of the surrounding population. Within the UK, geographical location and socio-economic factors are heavily influencing death rates [3] and, thus, our data may have limited generalizability to the wider UK population by not accounting for these factors. The level of generalizability could have also been better assessed by using external validation methods to test the performance of the model, but this would have required an additional independent dataset. Secondly, data was not always available, or accurate, for all patients. This was sometimes due to a lack of documentation, usually if the attending physicians at the time did not deem the information relevant, or if the information was not available, especially in patients who were cognitively impaired without any next of kin to provide collateral histories. Moreover, not all investigations, such as CT scans, were required for every patient and may have not been done due to the limited resources available in the NHS. Subsequently, only patients deemed to have abnormal results would have been the patients to receive the investigation. Additionally, some patients had different treatment goals to others. For example, patients with severe COVID-19 who could have had ventilatory support may have not had it because their treatment goal was palliation instead. All these factors together introduce information bias secondary to data measurement. Finally, although most parameters were objectively documented, some data, such as ethnicity, was self-reported by patients, thus also introducing a modest element of recall bias.

As a future work, it would be appealing to further investigate the impact of SMOTH technique on the validity of the BN classifiers. As discussed above, it was observed that the response variables in this study are significantly imbalance (e.g., the minority class percentages of NCPE or MOoVS were less than 5%). In order to overcome the imbalance in the dataset, the SMOTE technique was applied on the entire dataset, as recommended in [62–65]. However, the alternative approach, which is more plausible, would be to apply the SMOTH on training datasets only, and re-evaluate the performance measures using the test data.

5. Conclusions

In this study, we were facing several numerical challenges whilst constructing a robust, reliable and computationally efficient probabilistic data-driven model, including a very small sample size in comparison to the dimension of input variables, and significant rates

of missing values for several variables. These challenges were all efficiently resolved by selecting and employing a range of ML methods. The output of this study was a quantitative tool, which can aid in both risk-stratification, and earlier clinical decision making, for COVID-19 inpatients.

Overall, our findings demonstrate reliable, ML-based predictive models for four outcomes that utilise readily available clinical information for COVID-19 adult inpatients. Our model not only computes the probability distributions of children nodes given the values of their parent nodes, but also the distributions of the parents given the values of their children. In other words, they can proceed not only from causes to consequences, but also deduce the probabilities of different causes given the consequences. All these probabilities can be instantly computed using the codes developed in R. The codes are available on request from the corresponding author (A. D).

The developed models in this study, if provided with more training data, have the potential to be refined even further. Future research is required to externally validate our models and demonstrate their utility as clinical decision-making tools.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijerph18126228/s1, Table S1: List and definitions of independent predictor variables, Table S2: Patient Characteristics and descriptive statistics, Table S3: The heat-mapped probabilities of inpatient death of patients at different age groups. As shown, the death rate of COVID-19 patients' $\geq$70 years is 5 times larger than patients' $\leq$40 years, Table S4: The heat-mapped probabilities of death and survival of patients based on OS, Figure S1: The BN that is learned from the combination of data and expert opinions, to model "ADT" in terms of the first 11 most relevant factors, Figure S2: The BN with conditional probability tables learned for "ADT" outcome based on the combined elicited domain expert opinions with the (balanced) data, Figure S3: The BN that is learned from the combined data and non-domain expert opinions to model "MOoVS" in terms of the most influencing factors, Figure S4: The BN with conditional probability tables learned for "MOoVS" outcome based on the combined elicited domain expert opinions with the (balanced) data, Figure S5: The BN that is learned from the combined data and domain expert opinions to model "NCPE" in terms of its most influencing factors, Figure S6: The BN for "NCPE", as shown in Supplementary Figure S5, augmented with the computed conditional probability tables learned from data. The marginal probability table of each variable, after the whole dataset was balanced with respect to the NCPE, can be observed on each node.

Author Contributions: A.V. and A.S. contributed to the study design. A.V., A.S., A.O., D.D., J.S., M.I., P.C., P.K., M.R. and S.P. contributed to data collection. K.R., T.S., O.C., N.S., S.S. and A.D. contributed to data analysis. A.V., A.S., K.R., S.S. and A.D. contributed to original draft preparation, review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Health Research Authority (HRA) provided research approval and waived the need for Research Ethics Committee (REC) approval.

Informed Consent Statement: N/A. All data extracted from electronic patient records was collected during standard hospital care and was subsequently anonymised.

Data Availability Statement: Restrictions apply to the availability of the anonymised data used in this paper. Data was obtained from the Milton Keynes hospital electronic patient records, in accordance with good clinical practice guidelines, after obtaining health research authority (HRA) approval. The datasets are available from the authors with the permission of the Milton Keynes hospital authorities.

Acknowledgments: The authors would like to thank Coventry University and Milton Keynes University Hospital to support us in conducting this research.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. The Independent Coronavirus: A Timeline of How Britain Went from 'Low Risk' to an Unprecedented National Shutdown. 2020. Available online: https://www.independent.co.uk/news/uk/home-news/coronavirus-uk-timeline-deaths-cases-covid-19-nhs-social-distancing-a9416331.html (accessed on 23 January 2021).
2. NHS COVID-19 Daily Deaths, England. Available online: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/ (accessed on 23 January 2021).
3. Razaq, A.; Harrison, D.; Karunanithi, S.; Barr, B.; Asaria, M.; Routen, A.; Khunti, K. BAME COVID-19 DEATHS—What Do We Know? Rapid Data & Evidence Review. Available online: https://www.cebm.net/covid-19/bame-covid-19-deaths-what-do-we-know-rapid-data-evidence-review/ (accessed on 23 January 2021).
4. Office for National Statistics Coronavirus (COVID-19) Related Deaths by Ethnic Group, England and Wales: 2 March 2020 to 10 April 2020. Available online: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/coronavirusrelateddeathsbyethnicgroupenglandandwales/2march2020to10april2020#ethnic-breakdown-of-deaths-by-age-and-sex (accessed on 23 January 2021).
5. Office for National Statistics Deaths Involving COVID-19 by Local Area and Socioeconomic Deprivation: Deaths Occurring between 1 March and 17 April 2020. Available online: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19bylocalareasanddeprivation/deathsoccurringbetween1marchand17april (accessed on 23 January 2021).
6. Zhou, F.; Yu, T.; Du, R.; Fan, G.; Liu, Y.; Liu, Z.; Xiang, J.; Wang, Y.; Song, B.; Gu, X.; et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet* **2020**, *395*, 1054–1062. [CrossRef]
7. Guan, W.-J.; Liang, W.-H.; Zhao, Y.; Liang, H.-R.; Chen, Z.-S.; Li, Y.-M.; Liu, X.-Q.; Chen, R.-C.; Tang, C.-L.; Wang, T.; et al. Comorbidity and its impact on 1,590 patients with Covid-19 in China: A nationwide analysis. *Eur. Respir. J.* **2020**, *55*. [CrossRef]
8. Richardson, S.; Hirsch, J.S.; Narasimhan, M.; Crawford, J.M.; McGinn, T.; Davidson, K.W.; Barnaby, D.P.; Becker, L.B.; Chelico, J.D.; Cohen, S.L.; et al. Presenting Characteristics, Comorbidities, and Outcomes among 5700 Patients Hospitalized with COVID-19 in the New York City Area. *JAMA J. Am. Med. Assoc.* **2020**, *323*, 2052–2059. [CrossRef] [PubMed]
9. Williamson, E.; Walker, A.J.; Bhaskaran, K.; Bacon, S.; Bates, C.; Morton, C.E.; Curtis, H.J.; Mehrkar, A.; Evans, D.; Inglesby, P.; et al. OpenSAFELY: Factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. *medRxiv* **2020**. [CrossRef]
10. Vincent, J.-L.; Creteur, J. Ethical aspects of the COVID-19 crisis: How to deal with an overwhelming shortage of acute beds. *Eur. Heart J. Acute Cardiovasc. Care* **2020**, *9*, 248–252. [CrossRef]
11. Salari, N.; Khazaie, H.; Hosseinian-Far, A.; Ghasemi, H.; Mohammadi, M.; Shohaimi, S.; Daneshkhah, A.; Khaledi-Paveh, B.; Hosseinian-Far, M. The prevalence of sleep disturbances among physicians and nurses facing the COVID-19 patients: A systematic review and meta-analysis. *Glob. Health* **2020**, *16*, 1–14. [CrossRef]
12. Salari, N.; Khazaie, H.; Hosseinian-Far, A.; Khaledi-Paveh, B.; Kazeminia, M.; Mohammadi, M.; Shohaimi, S.; Daneshkhah, A.; Eskandari, S. The prevalence of stress, anxiety and depression within front-line healthcare workers caring for COVID-19 patients: A systematic review and meta-regression. *Hum. Resour. Health* **2020**, *18*, 1–14. [CrossRef] [PubMed]
13. Knight, S.R.; Ho, A.; Pius, R.; Buchan, I.; Carson, G.; Drake, T.M.; Dunning, J.; Fairfield, C.J.; Gamble, C.; Green, C.A.; et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Development and validation of the 4C Mortality Score. *BMJ* **2020**, *370*. [CrossRef]
14. Wynants, L.; Van Calster, B.; Collins, G.S.; Riley, R.D.; Heinze, G.; Schuit, E.; Bonten, M.M.J.; Damen, J.A.A.; Debray, T.P.A.; De Vos, M.; et al. Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ* **2020**, *369*. [CrossRef] [PubMed]
15. Gupta, R.K.; Marks, M.; Samuels, T.H.A.; Luintel, A.; Rampling, T.; Chowdhury, H.; Quartagno, M.; Nair, A.; Lipman, M.; Abubakar, I.; et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: An observational cohort study. *Eur. Respir. J.* **2020**, *56*. [CrossRef]
16. McLachlan, S.; Dube, K.; Hitman, G.A.; Fenton, N.E.; Kyrimi, E. Bayesian networks in healthcare: Distribution by medical condition. *Artif. Intell. Med.* **2020**, *107*, 101912. [CrossRef] [PubMed]
17. Kyrimi, E.; McLachlan, S.; Dube, K.; Fenton, N. Bayesian Networks in Healthcare: The chasm between research enthusiasm and clinical adoption. *medRxiv* **2020**. [CrossRef]
18. Constantinou, A.C.; Fenton, N. Things to know about Bayesian networks: Decisions under uncertainty, part 2. *Significance* **2018**, *15*, 19–23. [CrossRef]
19. Libbrecht, M.W.; Noble, W.S. Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* **2015**, *16*, 321–332. [CrossRef]
20. Chatrabgoun, O.; Hosseinian-Far, A.; Chang, V.; Stocks, N.G.; Daneshkhah, A. Approximating non-Gaussian Bayesian networks using minimum information vine model with applications in financial modelling. *J. Comput. Sci.* **2018**, *24*, 266–276. [CrossRef]
21. Chatrabgoun, O.; Hosseinian-Far, A.; Daneshkhah, A. Constructing gene regulatory networks from microarray data using non-Gaussian pair-copula Bayesian networks. *J. Bioinform. Comput. Biol.* **2020**, *18*, 2050023. [CrossRef]
22. Parmar, C.; Grossmann, P.; Rietveld, D.; Rietbergen, M.M.; Lambin, P.; Aerts, H.J.W.L. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. *Front. Oncol.* **2015**, *5*, 272. [CrossRef] [PubMed]

23. O'Hagan, A.; Buck, C.E.; Daneshkhah, A.; Eiser, J.R.; Garthwaite, P.H.; Jenkinson, D.J.; Oakley, J.E.; Rakow, T. *Uncertain Judgements: Eliciting Experts' Probabilities*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
24. Daneshkhah, A.; Oakley, J.E. Eliciting Multivariate Probability Distributions. *Rethink. Risk Meas. Rep.* **2010**, *1*, 23.
25. Daneshkhah, A.; Hosseinian-Far, A.; Sedighi, T.; Farsi, M. Prior elicitation and evaluation of imprecise judgements for Bayesian analysis of system reliability. In *Strategic Engineering for Cloud Computing and Big Data Analytics*; Springer: Cham, Switzerland, 2017; pp. 63–79.
26. Smith, J.Q.; Daneshkhah, A. On the robustness of Bayesian networks to learning from non-conjugate sampling. *Int. J. Approx. Reason.* **2010**, *51*, 558–572. [CrossRef]
27. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G.M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *Eur. Urol.* **2015**, *67*, 1142–1151. [CrossRef]
28. British Society of Thoracic Imaging **UPDATED**Version 2 BSTI COVID-19 Guidance for the Reporting Radiologist. Available online: https://www.bsti.org.uk/standards-clinical-guidelines/clinical-guidelines/bsti-covid-19-guidance-for-the-reporting-radiologist/ (accessed on 23 January 2021).
29. Chen, Y.-C.; Wheeler, T.A.; Kochenderfer, M.J. Learning discrete Bayesian networks from continuous data. *J. Artif. Intell. Res.* **2017**, *59*, 103–132. [CrossRef]
30. Allison, P.D. *Missing Data*, 1st ed.; Sage Publications, Inc.: New York, NY, USA, 2001; ISBN 9780761916727.
31. Kang, H. The prevention and handling of the missing data. *Korean J. Anesthesiol.* **2013**, *64*, 402–406. [CrossRef] [PubMed]
32. Torgo, L. *Data Mining with R: Learning with Case Studies*, 2nd ed.; Kumar, V., Ed.; Taylor & Francis: Boca Raton, FL, USA; Chapman & Hall/CRC: Boca Raton, FL, USA, 2016.
33. Ding, C.; Peng, H. Minimum Redundancy Feature Selection from Microarray Gene Expression Data. *J. Bioinform. Comput. Biol.* **2005**, *3*, 185–205. [CrossRef]
34. Urbanowicz, R.J.; Meeker, M.; La Cava, W.; Olson, R.S.; Moore, J.H. Relief-based feature selection: Introduction and review. *J. Biomed. Inform.* **2018**, *85*, 189–203. [CrossRef] [PubMed]
35. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer: Berlin/Heidelberg, Germany, 2017.
36. Guyon, I.; Gunn, S.; Nikravesh, M.; Zadeh, L.A. (Eds.) *Feature Extraction Foundations and Applications*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2006; ISBN 3-540-35487-5.
37. Saeys, Y.; Inza, I.; Larrañaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **2007**, *23*, 2507–2517. [CrossRef] [PubMed]
38. McGill, W.J. Multivariate information transmission. *Psychometrika* **1954**, *19*, 97–116. [CrossRef]
39. Iguyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
40. Kuhn, M.; Johnson, K. *Feature Engineering and Selection: A Practical Approach for Predictive Models*, 1st ed.; Chapman and Hall/CRC: New York, NY, USA, 2020; ISBN 9781315108230.
41. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. Available online: https://www.r-project.org/ (accessed on 6 June 2021).
42. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
43. Jensen, K.L.; Toftum, J.; Friis-Hansen, P. A Bayesian Network approach to the evaluation of building design and its consequences for employee performance and operational costs. *Build. Environ.* **2009**, *44*, 456–462. [CrossRef]
44. Sedighi, T.; Varga, L. *CECAN Evaluation and Policy Practice Note (EPPN) for Policy Analysts and Evaluators—A Bayesian Network for Policy Evaluation*; The Centre for the Evaluation of Complexity across the Nexus (CECAN): Guildford, UK, 2019.
45. Zhou, Y.; Fenton, N.; Neil, M. Bayesian network approach to multinomial parameter learning using data and expert judgments. *Int. J. Approx. Reason.* **2014**, *55*, 1252–1268. [CrossRef]
46. Lee, C.; Van Beek, P. Metaheuristics for score-and-search Bayesian network structure learning. In *Canadian Conference on Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 129–141.
47. Fenton, N.; Neil, M. *Risk Assessment and Decision Analysis with Bayesian Networks*, 2nd ed.; CRC Press: Boca Raton, FL, USA; Taylor & Francis Group: Boca Raton, FL, USA, 2019; ISBN 9781315269405.
48. Han, J.; Shi, L.-X.; Xie, Y.; Zhang, Y.-J.; Huang, S.-P.; Li, J.-G.; Wang, H.-R.; Shao, S.-F. Analysis of factors affecting the prognosis of COVID-19 patients and viral shedding duration. *Epidemiol. Infect.* **2020**, *148*. [CrossRef] [PubMed]
49. Heneghan, C.; Howdon, D.; Oke, J.; Jefferson, T. The Ongoing Problem of UK Hospital Acquired Infections. Available online: https://www.cebm.net/covid-19/the-ongoing-problem-of-hospital-acquired-infections-across-the-uk/ (accessed on 23 January 2021).
50. Mahbubani, K.; Georgiades, F.; Goh, E.L.; Chidambaram, S.; Sivakumaran, P.; Rawson, T.; Ray, S.; Hudovsky, A.; Gill, D. Clinician-directed improvement in the accuracy of hospital clinical coding. *Future Healthc. J.* **2018**, *5*, 47–51. [CrossRef] [PubMed]
51. Saunders, C.L.; Abel, G.A.; El Turabi, A.; Ahmed, F.; Lyratzopoulos, G. Accuracy of routinely recorded ethnic group information compared with self-reported ethnicity: Evidence from the English Cancer Patient Experience survey. *BMJ Open* **2013**, *3*. [CrossRef]
52. Feng, H.; Liu, Y.; Lv, M.; Zhong, J. A case report of COVID-19 with false negative RT-PCR test: Necessity of chest CT. *Jpn. J. Radiol.* **2020**, *38*, 409–410. [CrossRef]

53. He, J.-L.; Luo, L.; Luo, Z.-D.; Lyu, J.-X.; Ng, M.-Y.; Shen, X.-P.; Wen, Z. Diagnostic performance between CT and initial real-time RT-PCR for clinically suspected 2019 coronavirus disease (COVID-19) patients outside Wuhan, China. *Respir. Med.* **2020**, *168*, 105980. [CrossRef]
54. Mawaddah, A.; Gendeh, H.S.; Lum, S.G.; Marina, M.B. Upper respiratory tract sampling in COVID-19. *Malays. J. Pathol.* **2020**, *42*, 23–35.
55. Vinh, D.B.; Zhao, X.; Kiong, K.L.; Guo, T.; Jozaghi, Y.; Yao, C.; Kelley, J.M.; Hanna, E.Y. Overview of COVID-19 testing and implications for otolaryngologists. *Head Neck* **2020**, *42*, 1629–1633. [CrossRef] [PubMed]
56. Zitek, T. The appropriate use of testing for Covid-19. *West. J. Emerg. Med.* **2020**, *21*, 470–472. [CrossRef]
57. Al-Tawfiq, J.A. Asymptomatic coronavirus infection: MERS-CoV and SARS-CoV-2 (COVID-19). *Travel Med. Infect. Dis.* **2020**, *35*, 101608. [CrossRef]
58. Day, M. Covid-19: Four fifths of cases are asymptomatic, China figures indicate. *BMJ* **2020**, *369*, m1375. [CrossRef]
59. Hu, Z.; Song, C.; Xu, C.; Jin, G.; Chen, Y.; Xu, X.; Ma, H.; Chen, W.; Lin, Y.; Zheng, Y.; et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. *Sci. China Life Sci.* **2020**, *63*, 706–711. [CrossRef]
60. Nishiura, H.; Kobayashi, T.; Miyama, T.; Suzuki, A.; Jung, S.-M.; Hayashi, K.; Kinoshita, R.; Yang, Y.; Yuan, B.; Akhmetzhanov, A.R.; et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int. J. Infect. Dis.* **2020**, *94*, 154–155. [CrossRef] [PubMed]
61. Yu, X.; Yang, R. COVID-19 transmission through asymptomatic carriers is a challenge to containment. *Influenza Other Respi. Viruses* **2020**, *14*, 474–475. [CrossRef] [PubMed]
62. Agrawal, K.; Baweja, Y.; Dwivedi, D.; Saha, R.; Prasad, P.; Agrawal, S.; Kapoor, S.; Chaturvedi, P.; Mali, N.; Kala, V.U.; et al. A Comparison of Class Imbalance Techniques for Real-World Landslide Predictions. In Proceedings of the 2017 International Conference on Machine Learning and Data Science (MLDS), Noida, India, 14–15 December 2017.
63. Naseriparsa, M.; Kashani, M.M.R. Combination of PCA with SMOTE resampling to boost the prediction rate in lung cancer dataset. *Int. J. Comput. Appl.* **2013**, *77*, 33–38. [CrossRef]
64. Taft, L.M.; Evans, R.S.; Shyu, C.R.; Egger, M.J.; Chawla, N.; Mitchell, J.A.; Thornton, S.N.; Bray, B.; Varner, M. Countering imbalanced datasets to improve adverse drug event predictive models in labor and delivery. *J. Biomed. Inform.* **2009**, *42*, 356–364. [CrossRef]
65. Al-Bahrani, R.; Agrawal, A.; Choudhary, A. Colon Cancer Survival Prediction Using Ensemble Data Mining on SEER data. In Proceedings of the 2013 IEEE International Conference on Big Data, Silicon Valley, CA, USA, 6–9 October 2013.

https://dspace.org/

Using machine learning algorithms to develop a clinical decision-making tool for COVID-19 inpatients

Vepa, Abhinav; Saleem, Amer; Rakhshan, Kambiz; Daneshkhah, Alireza; Sedighi, Tabassom ...

2021-06-09
MDPI
https://doi.org/10.3390/ijerph18126228
Vepa A, Saleem A, Rakhshan K, et al., (2021) Using machine learning algorithms to develop a clinical decision-making tool for COVID-19 inpatients. International Journal of Environmental Research and Public Health, Volume 18, Issue 12, June 2021, Article number 6228

Downloaded from a DSpace repository. https://dspace.org/