Apply Four Laboratory Characteristics to Classify Critical Patients With COVID-19 After Admission

anping guo  
THE first affiliated hospital of USTC

Zejian Kuang  
Second Affiliated Hospital of Shantou University Medical College

Ying Wang  
the first affiliated hospital of USTC

Dongxue Luo  
the first affiliated hospital of shantou university medical college

Xiaoying Zheng  
Shantou Central Hospital

Benjun Yang  
the first affiliated hospital of USTC

Tian Yang  
the first affiliated hospital of USTC

Jianwei Dai  
Second Affiliated Hospital of Shantou University Medical College

Canhong Wen  
China University of Science and Technology

Haizhu Tan  
linnanqia@126.com  
https://orcid.org/0000-0001-9410-2961

Research article

Keywords: COVID-19, laboratory characteristics, critical patients, classification

DOI: https://doi.org/10.21203/rs.3.rs-36760/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background

The sudden outbreaking of COVID-19 worldwide has brought into sharp increased burden of economic and treatment in worldwide. All confirmed patients with different severity not only share the limited healthcare systems simultaneously but increase the risk of cross-infection among patients and health care workers. Hence, effective separation of critical COVID-19 patients from the common COVID-19 will be the key to success for ensuring critical patients to obtain treatment priorities and avoiding cross-infections in the hospital.

Methods:

A total of 105 patients with complete medical records were collected, including 84 blood samples of patients who confirmed in the First Affiliated Hospital of the University of Science and Technology at Anhui and 25 blood samples of patients in two hospitals at Shantou. Series of machine learning tools were introduced to explore and validate the most significant laboratory characteristics. Meanwhile, we compared it to three current popular assessment systems for pneumonia by using three methods, including the AUC index, NRI index and the net benefit.

Results:

We identified four significant potential laboratory characteristics for the classification of critical patients, including C-reactive protein, albumin, globulin, and sodium levels. The results also suggested the accurate and prediction efficacy of these selected indicators are the highest.

Conclusions

In conclusion, four easily available and low-cost laboratory characteristics appear to be import predictors of classification in critical patients after hospital admission. They guide therapeutic options and help clinicians make clinical decisions. Hence, we believe that such classification is essential for a more rational allocation scarce medical resource.

Introduction

Coronavirus disease (COVID-19), a coronavirus pneumonia, is a highly infectious disease and is an ongoing outbreak in the world. Symptoms of patients with COVID-19 always include fever, cough, fatigue and respiratory complications. All confirmed patients with different severity not only share the scare healthcare systems simultaneously but increase the risk of cross-infection among patients and health care workers. Hence, effective separation of critical COVID-19 patients from the common COVID-19 will be the key to success for ensuring critical patients to obtain treatment priorities and avoiding cross-infections [1].

Little attention has been paid to the classification of critical patients with COVID-19 who require immediate medical attention after hospital admission. No simple operable classification system has been specially designed for separating patients with COVID-19. Hence, clinicians have to apply three main pneumonia severity scoring systems to classify patients with COVID-19 in clinical trial, including the Clinical Pulmonary Infection Score (CPIS), Confusion-Urea-Respiratory Rate-Blood pressure-65 (CURB-65) and the pneumonia severity index (PSI). For example, Hankunyuan et al. reported a higher proportion of older than young and middle-aged COVID-19 patients with PSI grade IV and V [2]. Wu suggested that the PSI can be used to stratify patients with COVID-19 after hospitalization [3]. Liu indicated that the increase in CURB-65 score occurred concomitantly with the aggravation of acute respiratory distress syndrome in patients with COVID-19 [4]. In a multi-center study in Zhejiang province, patients with COVID-19 were classified by PSI and CURB-65 together, treated as a supplementary classification system for clinical assessment after admission[5].

In this study, we aim to identify some robust and interpretable laboratory biomarkers to separate critical patients from the common patients on three retrospective cohort studies from 3 hospitals in 2 provinces in China. We believe that they can provide effective medical resources allocation and treatment for patients and avoid cross-infections in the hospital.

Materials And Methods

Study design and patients
This was a retrospective study of three cohorts with COVID-19 initial diagnosed the “Diagnosis and treatment protocol for novel coronavirus pneumonia (Trial version 6)” published by the National Health Commission of China [6] before admission. 84 patients admitted from January 20 to February 20, 2020 to the First Affiliated Hospital of the University of Science and Technology of China were collected. 13 patients from the First Affiliated Hospital of Shantou University Medical College and 12 patients from Shantou Central Hospital admitted from January 19 to February 20, 2020 were collected. All these patients were confirmed to have SARS-CoV-2 infection by RT-PCR of samples from the respiratory tract by the Centers for Disease Control and Prevention. All patients with COVID-19 were hospitalized and admitted to the same ward without making distinction between common patients and critical patients[6].

**Data collection**

A total of 105 patients with complete medical records were collected. We reviewed all clinical data, laboratory characteristics and chest CT scans (see Table 1). The clinical data included demographic information, underlying comorbidities, symptoms and signs. Laboratory characteristics included routine blood tests, biomarkers for monitoring functions of multiple organs, and infection-related biomarkers. All data were collected within 24 hr after admission. According to the guideline for patients with confirmed COVID-19 from the National Health Commission in China, patients with mild clinical presentations (no pneumonia) may not initially require hospitalization. Hence, we removed data for 3 patients with clinical presentations because of possible bias. Data for 81 patients with 35 variables were retained.

**Statistical analysis**

In three datasets, we found that some laboratory characteristics had missing data. After deleting 2 variables with high missing rate (> 25%), we imputed the remaining data by using multiple imputation [7, 8]. We also handled the collinearity and filtering with mis-measured outliers by considering the results of variance inflation factor and correlation analysis together [9, 10]. Subsequently, a classification model including 35 candidate predictors was fitted by using cforest implementation with the random forest (RF) classification model [11]. During this analysis, the importance of various conditioning factors can be measured quantitatively, and we found several negative importance variables. We kept running a loop function to remove negative values. The importance of selected variables was weighted by using the weight of evidence (woe) method to improve the classification accuracy [12]. Finally, four significant selected biomarkers were selected by using a generalized linear model (glm) with the stepwise Bayesian information criterion method. The prediction model was depicted by the nomogram.

Internal validation was conducted 100 times by spitting 80% of data into a training set with \( n_{\text{train}} = 62 \) samples and 20% into a test set with \( n_{\text{test}} = 15 \) samples. Then we counted the total times each predictive variable was present in each model. Moreover, external validation involved using data for 25 patients with COVID-19 from 2 hospitals in Shantou.

We compared the new classification to the three current popular assessment systems for pneumonia by using three methods, including the area under the receiver operating characteristic (ROC) curve (AUC) index, the net reclassification (NRI) index and the net benefit. AUC was used to describe the diagnostic ability of a binary classifier system [13]. The net reclassification index (NRI) was used to evaluate the improvement in risk prediction by adding a marker to a set of baseline predictors [14, 15]. Decision curve analysis (DCA) was used to evaluate and compare prediction models that incorporate clinical consequence [16]. In this study, we used DCA to graphically describe the clinical usefulness of each classifier based on a potential threshold for misclassification (x axis) and the net benefit of using the model to risk-stratify patients (y axis) relative to assuming that no patient will be misclassified.

Statistical analyses were performed with R v3.6.3 and \( p < 0.05 \) was considered statistically significant.

**Results**

Table 1 describes participant characteristics. Three cohorts show few major differences existed.

After filtering collinearity and outliers, 77 patients from Hefei were retained. The full model was approximated by a small model including the 14 most predictive variables by using RF. Only 13 predictive variables were retained after deleting variables with strength of evidence less than “very strong”. Figure 1 shows the weight of evidence of importance of variables. Finally, four significant biomarkers were selected, including: CRP (\( P = 0.001 \)), ALB (\( P = 0.014 \)), GLB (\( P = 0.013 \)) and sodium (\( P = 0.006 \)) (see Table 2). The nomogram was depicted in Fig. 2 and the final classification model was described in formula 1:

\[
\text{Logit}(p) = 76.579 + 0.064 \times \text{CRP} - 0.259 \times \text{ALB} + 0.287 \times \text{GLB} - 0.567 \times \text{sodium}. \quad \ldots \ldots (1)
\]

In internal validation, the random-splitting was repeated 100 times and results are described in Table 3. CRP and sodium level appeared 100 times, ALB level 72 times and GLB level 85 times and so were selected to build the model in the training dataset. Table 2 showed the results
of GLM analysis. Both them demonstrated that four selected laboratory characteristics can be regarded as potential biomarkers for identifying the critical patients.

The AUC for the CPIS score was highest (AUC_{CPIS} = 0.988) and that for four biomarkers was lower, 0.881 (Fig. 2). However, the ability to discriminate patients with critical and common disease was better by using four biomarkers than the CPIS, mainly because the CPIS overestimates the variance when the AUC is close to 1 and it is not realistic in clinical trials [17]. DCA demonstrated that the prediction model built by four biomarkers improved the accuracy of classification against the threshold probabilities of three popular classifiers. Table 4 suggested that the new classification model performed the best because the values of three NRIIs were larger than 0. The new classification model was always superior to other 3 models across a wide range of threshold probabilities (Fig. 3). For example, the highest difference between the new prediction model and CPIS was at a threshold probability around 0.41. At that threshold, the net benefit for the new prediction model was about 0.29 and 0.1 for CPIS. At that threshold, using the new prediction model over the CPIS to classify patients and make clinical decisions, the probability of more profitable treatment was 28% (95% CI 0.29 – 0.1).

**Discussion**

According to the results of our study, four easily available and low-cost laboratory characteristics appear to be import predictors of classification in critical patients after hospital admission. The new classification model based on four laboratory characteristics was demonstrated to had better discriminative ability than other 3 current popular systems. The results of AUC, NRI and DCA analysis also demonstrated that it was the best classifier. The discriminative ability of it was also externally validated.

CRP level were positively correlated with the severity of COVID-19. It was consistent with some previous studies. CRP can activate the complement system to enhance the regulation of lymphocytes and promote the phagocytosis of macrophages to eliminate the invading pathogens [18, 19]. Some studies of COVID-19 showed that CRP level was significantly increased specifically in patients with severe disease [20, 21]. The reason might be some inflammatory factors such as interleukin 6, interleukin 1, tumor necrosis a could promote the synthesis of CRP by hepatocytes [18]. Ko et al. found that CRP ≥ 2 mg/dl was one of the predictive factors for pneumonia development of Middle East respiratory syndrome (MERS), while CRP ≥ 4 mg/dl, low albumin level, male, hypertension, thrombocytopenia, lymphopenia were regarded as the predictive factors for respiratory failure [22]. A recent retrospective study also showed that CRP levels of patients with COVID-19 were also significantly higher in the death group on admission [23]. Liu et al. reported that IL-6 and CRP could be used as independent factors to predict the severity of COVID-19, and those patients were more likely to have severe complications while their CRP level larger 41.8 mg/L [24]. Wang also suggested that CRP level can be regarded as an important biomarker in the early stage of COVID-19 because CRP could reflect lung lesions and disease severity [25]. Albumin was the second potential biomarker found in our study. It could be detected in the blood and was a protein made in the liver. Albumin could prevent leakage of the fluids from the blood into other organs [26]. Increasing number of studies showed that low albumin levels were associated with poorer outcomes of patients with COVID-19 [27]. Albumin concentration was suggested as an independent risk factor for mortality in patients with pneumonia and also found associated with COVID-19 [28, 29]. A systematic reviewed and meta-analysis showed that hypoalbuminemia status increased risk of severe COVID-19 [30]. Our study also described that lower sodium was a risk factor for severe COVID-19 infection. Sodium was considered a predictor in several scoring systems for assessing pneumonia, including the PSI and Acute Physiology and Chronic Health Evaluation II. Hyponatremia was the most common electrolyte disorder in clinical practice and severe hyponatremia was associated with increased mortality [31]. Berni et al. found that sodium was inversely correlated with IL-6 in COVID-19 patients, directly correlated with PaO₂/FiO₂ ratio [32]. Stephan J.L Bakker gave a hypothesis about that low sodium balance may augment cellular damage at a certain virus load and increase the risk of developing severe and fatal COVID-19 infection by their experimental and epidemiological data [33]. Finally, globulin was suggested to be positively relative with the severity of COVID-19. Yafei Zhang demonstrated that the globulin level in severe COVID-19 patients is significantly increased while comparing to the mild patients because the promoted immunoglobulin synthesis [27].

In addition, the CPIS, a diagnostic algorithm, is mainly applied for ventilator-associated pneumonia and community-acquired pneumonia. Most studies indicated that CPIS had inaccurate sensitivity and specificity [34–37]. The CPIS was suggested to have high inter-observer variability and is not available for multiple centers study [37, 38]. The CURB-65 score consists of 5 separate elements: confusion, uremia, respiratory rate, blood pressure, and age ≥ 65. The CURB-65 is relatively simple to use. The PSI involves 20 clinical variables defining 5 classes of increasing risk of mortality. It has been extensively validated. However, the inappropriate weights of age or inappropriate threshold values for both the PSI and CURB-65 result in a potential underestimation of severe pneumonia, especially in young people [39, 40].

A major limitation of the current study is the insufficient sample size. As more raw data will be collected in the future, we would have the ability to optimize our new model. Another limitation of our study was that we had to combine patients with critical presentation to severe presentation because there were only 6 patients with a critical clinical presentation in our study.
In conclusion, four easily available and low-cost laboratory characteristics appear to be important predictors of classification in critical patients after hospital admission. They guide therapeutic options and help clinicians make clinical decisions. Hence, we believe that such classification is essential for a more rational allocation of scarce medical resources.

**Declarations**

**Acknowledge**

Thanks to three hospital staff members for their efforts in collecting the data that was used in this study, and all patients who consented to donate their data for our analysis and the medical staff members who insist on working on the front line of caring for patients.

**Data availability**

Mrs. AP Guo Tan had full access to all of the data in the study. After publication, the data will be made available to others on reasonable requests after approval from the author (Mrs. AP Guo guoanpingah@126.com).

**Conflict of interest statement:**

The authors have declared that no conflict of interest exists.

**Study approval**

This study was approved by the Medical Ethical Committees of the First Affiliated Hospital of the University of Science and Technology of China (Hefei, China), the Medical Ethical Committees of the First Affiliated Hospital of Shantou University Medical College (Shantou, China) and the Medical Ethical Committees of the Shantou Central Hospital (Shantou, China) respectively (approval letter 2020-P-018, 2020-046, 2020-019). Written informed consent was waived.

**Contributors:**

ZK and AG contributed to the concept and design of the study, acquisition of data, and interpretation of data, drafting the article and final approval of the version to be published, they contributed equally to this work and are joint first authors. DL, XZ, BY and JD were involved in data collection, verification and final approval of the version to be published. TY and YW were involved in the interpretation of the data, drafting the article. HZ and CW contributed to the obtaining funding, design of the study, acquisition of data, statistical analysis and interpretation of data, drafting the article, and approval of the final version to be published. AG is the data guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Funding source**

This work was supported by grants from the National Natural Science Foundation of China (11801540); the Natural Science Foundation of Guangdong (2017A030310572); the Fundamental Research Funds for the Central Universities (WK2040170015, WK2040000016); Shantou Science and Technology Project ([2018] 131 and Shantou Science and Technology Project ([2018] 121).

**References**

1. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. 2003;58(5):377–82.
2. Liu K, Chen Y, Lin R. K. Han.Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. J Infect. 2020;80(6):e14–8.
3. Tian S, Chang Z, Wang Y, Wu M, Zhang W, Zhou G, et al. Clinical characteristics and reasons of different duration from onset to release from quarantine for patients with COVID-19 Outside Hubei province, China. MedRxiv 2020.
4. Liu Y, Sun W, Li J, Chen L, Wang Y, Zhang L, et al. Clinical features and progression of acute respiratory distress syndrome in coronavirus disease 2019. MedRxiv 2020.
5. Yang W, Cao Q, Qin L, Wang X, Cheng Z, Pan A, et al. Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19): A multi-center study in Wenzhou city. Zhejiang China J Infect. 2020;80(4):388–93.
6. Diagnosis. and treatment guidelines for novel coronavirus pneumonia (draft version 6), http://www.nhc.gov.cn/yzwj/s7653p/202002/8334a8326dd94d329df351d7da8ae6c2/files/b218cfeb1bc54639af227f922bf6b817.pdf (accessed Feb 19 2020).
34. Post A, Dullaart RPF, Bakker SJL. Is low sodium intake a risk factor for severe and fatal COVID-19 infection? Eur J Intern Med. 2020;75:109.

35. Harde Y, Rao SM, Sahoo JN, Bharuka A, Betham S, Pulla S. Detection of ventilator associated pneumonia, using clinical pulmonary infection score (CPIS) in critically ill neurological patients. Journal of Anesthesiology and Clinical Science. 2013;2. https://doi.org/10.7243/2049-9752-2-20.

36. Fartoukh M, Maitre B, Honore S, Cerf C, Zahar JR. C. Brun-Buisson. Diagnosing pneumonia during mechanical ventilation: the clinical pulmonary infection score revisited. Am J Respir Crit Care Med. 2003;168:173–9.

37. Cook DJ, Walter SD, Cook RJ, Griffith LE, Guyatt GH, Leasa D, et al. Incidence of and risk factors for ventilator-associated pneumonia in critically ill patients. Ann Intern Med. 1998;129:433–40.

38. Zilberberg MD, Shorr AF. Ventilator-associated pneumonia: the clinical pulmonary infection score as a surrogate for diagnostics and outcome. Clin Infect Dis. 2010;51(Suppl 1N):131–5.

39. Schurink CAM, Nieuwenhoven CAV, Jacobs JA, Rozenberg-Araska M, Joore HCA, Buskens E, et al. Clinical pulmonary infection score for ventilator-associated pneumonia: accuracy and inter-observer variability. Intensive Care Med. 2004;30:217–24.

40. Chen JH, Chang SS, Liu JJ, Chan RC, Wu JY, Wang WC, et al. Comparison of clinical characteristics and performance of pneumonia severity score and CURB-65 among younger adults, elderly and very old subjects. Thorax. 2010;65:971–7.

41. Shah BA, Ahmed W, Dhobi GN, Shah NN, Khursheed SQ. I. Haq. Validity of pneumonia severity index and CURB-65 severity scoring systems in community acquired pneumonia in an Indian setting. Indian J Chest Dis Allied Sci. 2010;52:9–17.

Tables
Table 1 Demographics, Clinical Indicators and Laboratory Characteristics in Two Cohorts

| Group          | Hefei Derivation Cohort (n=84) | Shantou Validation Cohort (n=25) |
|----------------|--------------------------------|----------------------------------|
|                | Mild                           | 3(3.57%)                         | 9(36.00%)                        |
|                | Common                         | 56(66.67%)                       | 10(40.00%)                       |
|                | Severity/Critical              | 27(32.14%)                       | 6(24.00%)                        |
| gender         | male                           | 50(59.52%)                       | 10(40.00%)                       |
|                | female                         | 34(40.48%)                       | 15(60.00%)                       |
| age            |                                | 47(5-91)                         | 38(12-75)                        |
| Respiratory    |                                | 20(15-40)                        | 20(18-24)                        |
| Temperature    |                                | 36.85(36-39.5)                   | 36.8(36.1-39.2)                  |
| SBP            |                                | 121(100-175)                     | 124(105-166)                     |
| DBP            |                                | 76(60-108)                       | 81(66-113)                       |
| HR             |                                | 85(60-112)                       | 90(55-128)                       |
| CRP            |                                | 17.7(0.5-150)                    | 4.375(0.499-93.3)                |
| ESR            |                                | 38.5(3-145)                      | 8(2-30)                          |
| BNP            |                                | 1116.89(20-5601)                 | 60.7(18-1433)                    |
| cTnl           |                                | 0.08(0.02-3.10)                  | 0.01(0-0.1)                      |
| ALB            |                                | 43.05(30-54)                     | 40.86(34.12-50.8)                |
| GLB            |                                | 28(15.6-52.8)                    | 31.5(24.26-39.52)                |
| TG             |                                | 2.64(0.59-9.51)                  | 1.08(0.51-10.92)                 |
| ALT            |                                | 22.5(8-236)                      | 23(7.22-126.79)                  |
| AST            |                                | 27(14-182)                       | 24.23(12-65.69)                  |
| CK             |                                | 87.25(23-1062.6)                 | 71.47(15.77-827)                 |
| CK-MB          |                                | 10.95(4.4-57.6)                  | 13(5.77-33)                      |
| Cr             |                                | 65(4-1561)                       | 82.26(46-119.5)                  |
| BUN            |                                | 4.155(1.57-269)                  | 4.13(2.48-6.57)                  |
| Cys C          |                                | 1.08(0.43-13.79)                 | 0.93(0.68-10.3)                  |
| LDH            |                                | 243.5(130-936)                   | 202(128-354)                     |
| Na             |                                | 137(126-146)                     | 139(135.6-143.2)                 |
| Glu            |                                | 6.24(4.05-21.08)                 | 5.26(2.23-17)                    |
| PCT            |                                | 0.15(0.01-1.87)                  | 0.05(0.02-1.39)                  |
| WBC            |                                | 5.34(2.08-21.77)                 | 4.9(2.2-21.6)                    |
| NE             |                                | 3.57(1.37-20.89)                 | 2.485(0.89-9.9)                  |
| LY             |                                | 1.075(0.22-5.61)                 | 1.4(0.7-3.2)                     |
| HCT            |                                | 0.411(0.257-0.506)               | 0.431(0.32-44.8)                 |
| PLT            |                                | 163.5(72-363)                    | 220(86-367)                      |
| PT             |                                | 14.5(10.7-19.9)                  | 10.65(9.7-12.8)                  |
| APTT           |                                | 38.65(25.4-52.3)                 | 26.2((21.1-33.3)                 |
|                  | Estimate 1   | Estimate 2   |
|------------------|-------------|-------------|
| Fib-Fibrinogen   | 3.23(1.04-6.45) | 3.05(1.53-6.78) |
| D-dimer D-Dimer  | 0.25(1.5-38.2)  | 180(60-850)  |
| I-Bil Indirect bilirubin | 8.7(2.0-29.0)       | 10.735(6.1-30.05) |
| D-Bil Direct bilirubin   | 5.5(1.5-38.2)       | 2.71(1.2-5.75)   |

### Table 2: Results of generalized linear model

|                  | Estimate | P-value |
|------------------|----------|---------|
| **new-model**    |          |         |
| Intercept        | 76.573   | 0.005   |
| CRP              | 0.064    | 0.001   |
| ALB              | -0.259   | 0.014   |
| GLB              | 0.282    | 0.013   |
| sodium           | -0.567   | 0.006   |
| **CPIS**         |          |         |
| Intercept        | -5.212   | 1.63E-05|
| CPIS             | 1.836    | 3.61E-05|
| **CURB-65**      |          |         |
| Intercept        | -1.027   | 2.77E-04|
| CURB.65          | 1.455    | 0.017   |
| **PSI**          |          |         |
| Intercept        | -4.578   | 8.78E-06|
| PSI              | 0.078    | 5.09E-05|

### Table 3: Occurrences for significant variables by 100 repeats

|                  | Times | Times | Times | Times |
|------------------|-------|-------|-------|-------|
| CRP              | 94    | 72    | 85    | 94    |
| ALB              |       |       |       |       |
| GLB              |       |       |       |       |
| sodium           |       |       |       |       |

### Table 4: Rustls of NRI index

|                  | CPIS    | CURB.65 | PSI    |
|------------------|---------|---------|--------|
| **new-model**    | 0.533   | 0.513   | 0.659  |

**Figures**
Figure 1

Weight of evidence of importance of variables related to severity of COVID-19. CRP, C-reactive protein; ALB, albumin; GLB, globulin; BUN, blood urea nitrogen; PCT, procalcitonin; LDH, lactate dehydrogenase; ALT, alanine aminotransferase; AST, aspartate transaminase; BP, blood pressure; Na, sodium; cTN.I, cardiac troponin I; D.Bil, direct bilirubin; PT, prothrombin time.
Figure 2

The plot of nomogram.
Figure 3

Area under the receiver operating characteristic curve (AUC) for the new model. CPIS, Clinical Pulmonary Infection Score; CURB-65, Confusion-Urea-Respiratory Rate-Blood pressure-65; PSI, pneumonia severity index.
Figure 4

Threshold probabilities of models.