Computer-aided prediction of in-hospital COVID-19 mortality utilizing neurologic symptoms

brajwa@purdue.edu jrochet@purdue.edu

Bartek Rajwa[1,2], Md Mobasshir Arshed Naved[3], Mohammad Adibuzzaman[4], Ananth Y. Grama[3], Babar Khan[5], M. Murat Dundar[6], Jean-Christophe Rochet[2]

[1]Bindley Bioscience Center, [2]Purdue Institute for Integrative Neuroscience, [3]Dept. of Computer Science, Purdue University, West Lafayette, IN, [4]Oregon Clinical & Translational Research Institute, Oregon Health and Science University, Portland, OR, [5]Regenstrief Institute, Indianapolis, IN, [6]Dept. of Computer & Information Sciences, IUPUI, Indianapolis, IN

GILEAD Creating Possible

The authors gratefully acknowledge financial support from the COMMITTM (COvid-19 unMet MedIcal needs and associated research exTension) Program of the Gilead Foundation.

ABSTRACT

As the world emerges from the pandemic caused by SARS-CoV-2, there is a need to understand factors that determine the effects of COVID-19, as well as the diagnostic features that may be used to predict the occurrence of severe cases and mortality. Patients with severe COVID-19 are often afflicted with neurologic symptoms, and individuals with a pre-existing neurodegenerative disease have an increased risk of severe COVID-19 [1–4]. We conducted a study to determine the relationship between the lethality of COVID-19 and CNS-related symptoms. The electronic health records of 471 patients with severe COVID were analyzed to automatically identify the clinical characteristics predictive of COVID-19 mortality. The feature discovery was conducted by a regularized logistic regression classifier with an embedded feature selection capability [5]. The initial selection followed by SHAP analysis revealed that a small ensemble of readily observable clinical features, notably including characteristics associated with cognitive impairment, could predict in-hospital mortality with an accuracy greater than 0.85 (expressed as the area under the ROC curve of the classifier). These findings have important implications for the prioritization of clinical measures used to identify patients with COVID-19 (and, potentially, other forms of acute respiratory distress syndrome) having an elevated risk of mortality.

DATA INPUT

- The data represent 471 patients admitted to the ICU at IU Health Methodist Hospital and Sydney & Lois Eskenazi Hospital (Indianapolis, IN) with severe SARS-CoV-2 infection.
- 399 patients were eventually discharged, and 72 died (see Table below).
- 196 patients identified as *white*, and 246 as *Black or African American*.
- 245 of the patients were females, and 226 were males.
- No significant difference in age between the African-American and white patients.
- *Hispanic/Latino* patients were significantly younger than other patients.

	Females					Males				
Race	**Alive**	**Alive Per.**	**Died**	**Died %**	**Total**	**Alive**	**Alive %**	**Died**	**Died %.**	**Total**
Asian	3	100.0%	0	0.0%	**3**	5	83.3%	1	16.7%	**6**
Black or Afr. Amer.	123	87.9%	17	12.1%	**140**	85	80.2%	21	19.8%	**106**
Refused to identify	1	33.3%	2	66.7%	**3**	2	100.0%	0	0.0%	**2**
Unknown	10	100.0%	0	0.0%	**10**	5	100.0%	0	0.0%	**5**
White	77	86.5%	12	13.5%	**89**	88	82.2%	19	17.8%	**107**
Total	**214**	**87.3%**	**31**	**12.70%**	**245**	**185**	**81.9%**	**41**	**18.1%**	**226**

MACHINE-LEARNING MODEL

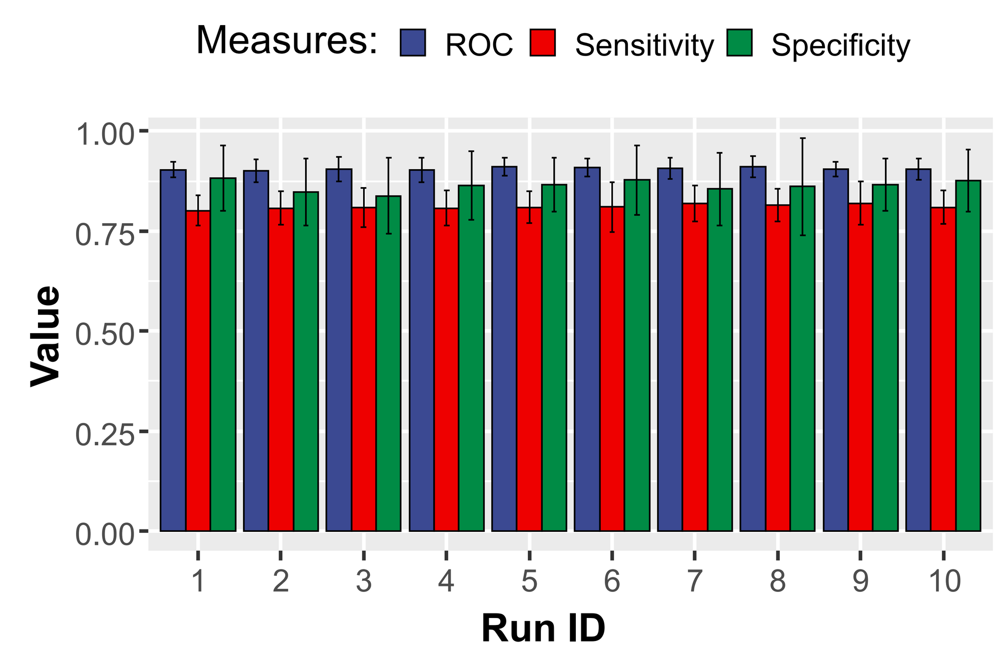
- Set of relevant diagnostic features established by implementing an ante-hoc explainable, predictable statistical model with embedded feature selection capability.
- We utilized a logistic regression model regularized with a ridge (ℓ2), LASSO (ℓ1), or a combination of both penalties (elastic net) [5–7].
- This approach allowed: (1) creation of a simple model capturing all significant sources of variability, incorporating all diverse clinical descriptors/features; and (2) performance of simultaneous feature selection and feature ranking, enabling identification of the major drivers of correct prediction [8].

$$\left(\hat{\beta}_0, \hat{\beta}\right) = \underset{(\beta_0,\beta)\in\mathbb{R}^{p+1}}{\operatorname{argmin}} \left[\frac{1}{2n}\sum_{i=1}^{n}\left(y_i - \beta_0 - x_i^T\beta\right)^2 + \lambda\left(\frac{1-\alpha}{2}\|\beta\|_2^2 + \alpha\|\beta\|_1\right)\right]$$

where λ is the tuning hyperparameter controlling the overall strength of the LASSO and ridge penalties and α controls the balance between them.

FEATURE SELECTION

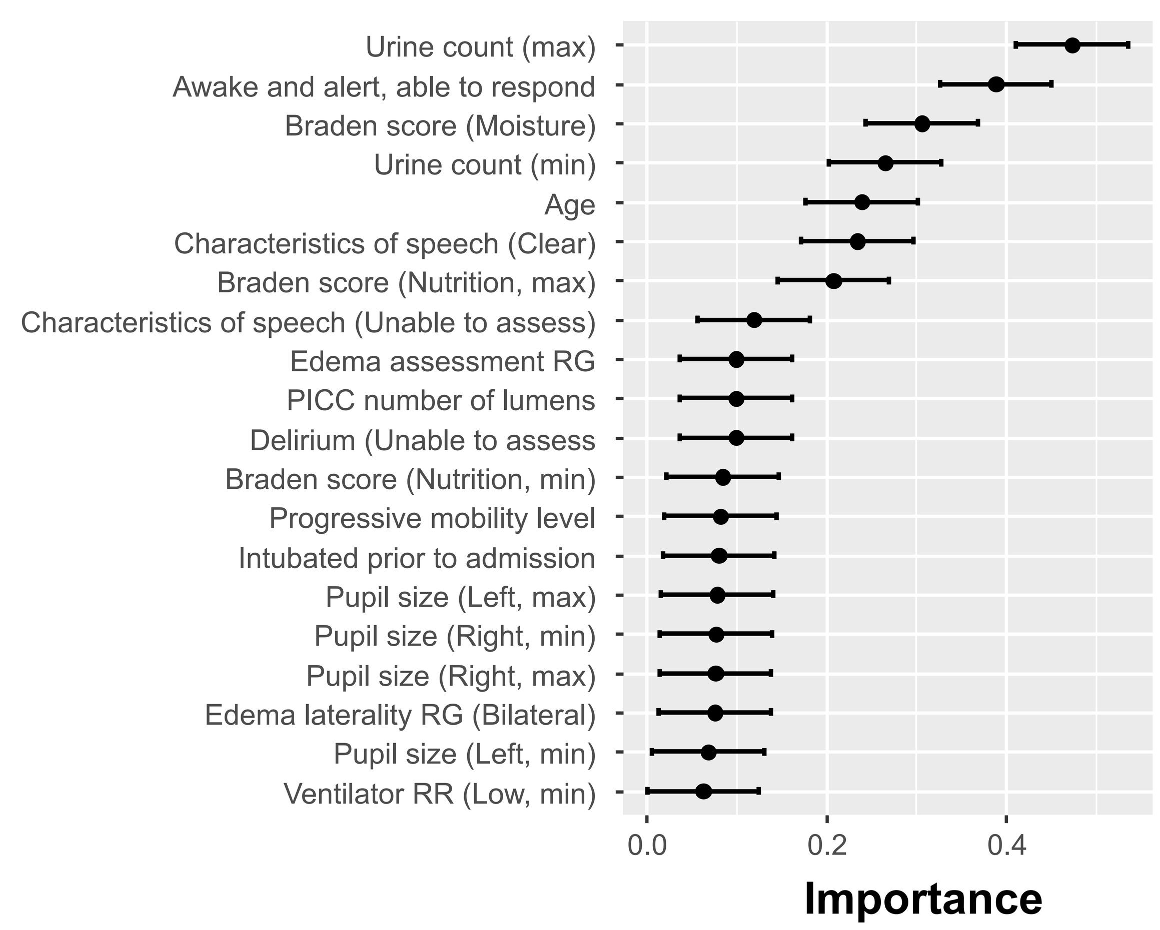
- Minor adjustments in the model's random initialization or train-test split led to variances in the selected feature set for the embedded feature selection methods due to the known problem of selection instability.
- Investigating an ensemble of independent models, each of which was initiated with a different random seed, resolved this issue.



FEATURE IMPORTANCE

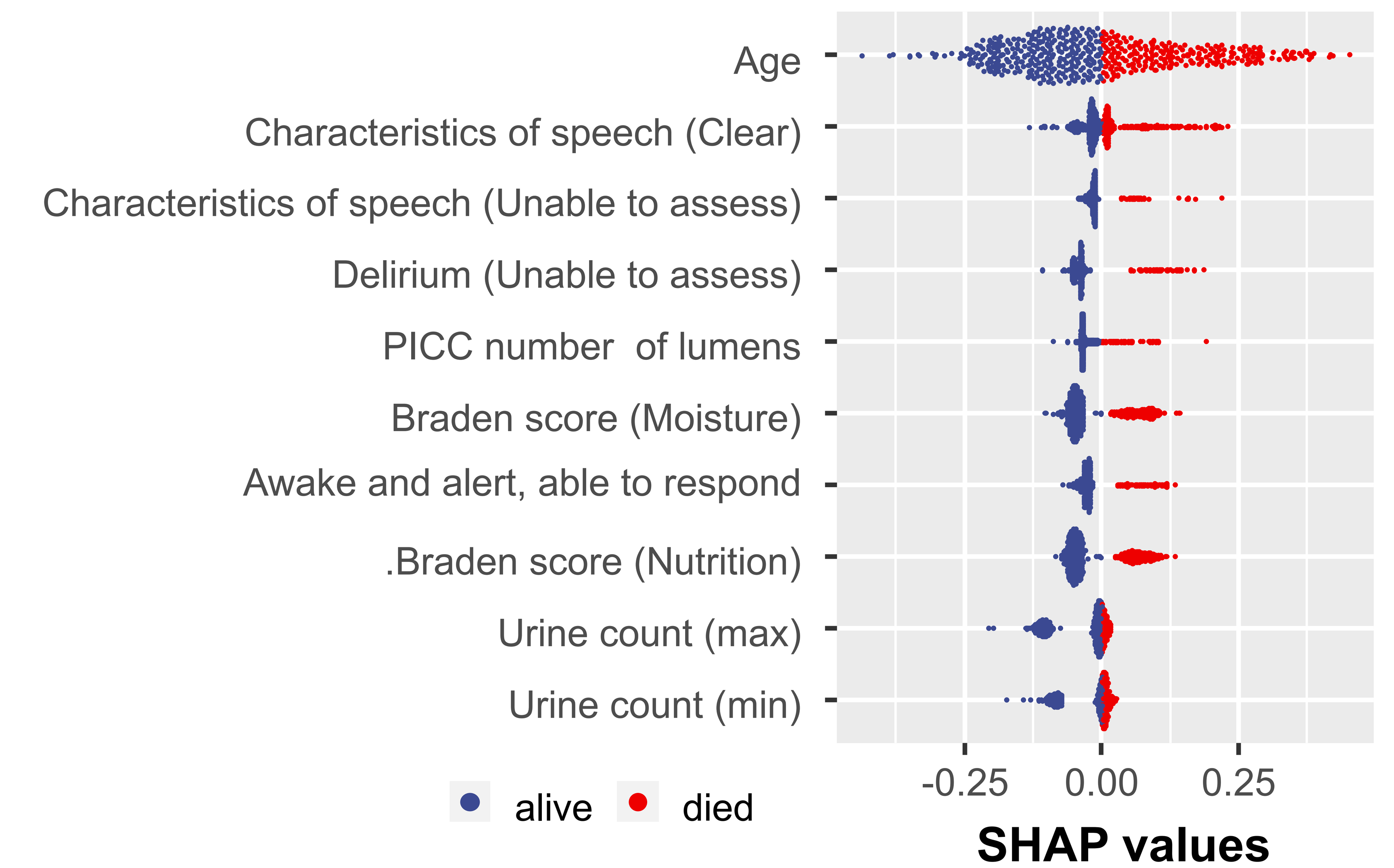
Global feature importance

- ENET regularization penalizes the size of the coefficients, sets some irrelevant values to 0, and minimizes the impact of irrelevant features.
- Feature importance can be expressed by the absolute values of the non-zero coefficients of the covariates.



Local feature importance

- We analyzed the local importance (i.e., judged for each patient) of each feature.
- SHapley Additive exPlanation (SHAP) values can determine the importance of a feature and its directionality influence by comparing what a model predicts with and without that feature for each observation in the training data by calculating the marginal contribution [9].
- Based on game theory, the SHAP values illustrate how vital each player (i.e., classification feature) is to the overall cooperation and what payoff (i.e., classification accuracy) the player can expect from participation in the game [10].



RESULTS

Besides age, the feature selection pipeline identified several clinical descriptors that can be divided into two categories: CNS-related and frailty-related features:

CNS-related neuropsychiatric features include

- *Inability to assess the patient's speech,* possibly caused by sedation, loss of consciousness, or delirium. Delirium has been described as one of the most common neuropsychiatric manifestations of severe COVID-19 [3,4,11–13].
- *Unclear/slurred speech.* Nonsensical speech, confusion, and disorientation are described as initial neurological symptoms of severe COVID-19 [2,14,15].
- *Inability to assess delirium* feature communicates that a patient was unable to be subjected to the Confusion Evaluation Method [16].
- *Awake and able to respond* feature demonstrates that a patient is awake, responding appropriately, and aware of self, place, and time.

Patient well-being and frailty features

- *The urine voiding count* feature is connected with the lower urinary tract symptoms known to be associated with COVID-19 [17,18].
- The majority of COVID-19 patients may experience urinary incontinence, increased urination frequency, nocturia, and urgency during the infection, perhaps caused indirectly by COVID-19-related general dysfunction in the autonomic nervous system [16,19].
- *Braden score,* created to identify early pressure sore-prone patients, contains six subscales measuring sensory perception, skin wetness, activity, mobility, friction and shear, and nutrition [20]. It describes the overall condition of the patient.

REFERENCES

1. Almeria, M., Cejudo, J. C., Sotoca, J., Deus, J. & Krupinski, J. Cognitive profile following COVID-19 infection: Clinical predictors leading to neuropsychological impairment. *Brain, Behavior, & Immunity - Health* **9**, 100163 (2020).
2. Paterson, R. W. *et al.* The emerging spectrum of COVID-19 neurology: clinical, radiological and laboratory findings. *Brain* **143**, 3104–3120 (2020).
3. Ahmad, I. & Rathore, F. A. Neurological manifestations and complications of COVID-19: A literature review. *Journal of Clinical Neuroscience* **77**, 8–12 (2020).
4. Beach, S. R. *et al.* Delirium in COVID-19: A case series and exploration of potential mechanisms for central nervous system involvement. *General Hospital Psychiatry* **65**, 47–53 (2020).
5. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J Royal Statistical Soc B* **67**, 301–320 (2005).
6. Tibshirani, R. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B (Methodological)* **58**, 267–288 (1996).
7. Hoerl, A. E. & Kennard, R. W. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* **12**, 55–67 (1970).
8. Zou, H. Classification with high dimensional features. *WIREs Computational Statistics* **11**, e1453 (2019).
9. Song, E., Nelson, B. L. & Staum, J. Shapley effects for global sensitivity analysis: theory and computation. *SIAM/ASA J. Uncertainty Quantification* **4**, 1060–1083 (2016).
10. Shapley, L. S. A value for n-person games. in *The Shapley Value: Essays in Honor of Lloyd S. Shapley* (ed. Roth, A. E.) 31–40 (Cambridge University Press, 1988). doi:10.1017/CBO9780511528446.003.
11. Forget, M.-F. *et al.* Delirium and inflammation in older adults hospitalized for COVID-19: a cohort study. *CIA* **16**, 1223–1230 (2021).
12. Kroon, B. *et al.* Delirium in older COVID-19 patients: Evaluating risk factors and outcomes. *International Journal of Geriatric Psychiatry* **37**, (2022).
13. Pranata, R. *et al.* Delirium and mortality in coronavirus disease 2019 (COVID-19) – a systematic review and meta-analysis. *Archives of Gerontology and Geriatrics* **95**, 104388 (2021).
14. Ramage, A. E. Potential for cognitive communication impairment in COVID-19 survivors: a call to action for speech-language pathologists. *American Journal of Speech-Language Pathology* **29**, 1821–1832 (2020).
15. Saeed, S. A., Pastis, I. S. & Santos, M. G. COVID-19 and its impact on the brain and Mind- A conceptual model and supporting evidence. *Psychiatr Q* **93**, 271–284 (2022).
16. Inouye, S. K. *et al.* Clarifying confusion: the confusion assessment method. *Ann Intern Med* **113**, 941–948 (1990).
17. Mumm, J.-N. *et al.* Urinary frequency as a possibly overlooked symptom in COVID-19 patients: does SARS-CoV-2 cause viral cystitis? *European Urology* **78**, 624–628 (2020).
18. Swatesutipun, V. & Tangpaitoon, T. Lower urinary tract symptoms (LUTS) related to COVID-19: review article. *Journal of the Medical Association of Thailand* **104**, 1045–9 (2021).
19. Inouye, S. K. *et al.* A chart-based method for identification of delirium: validation compared with interviewer ratings using the confusion assessment method. *Journal of the American Geriatrics Society* **53**, 312–318 (2005).
20. Bergstrom, N., Braden, B. J., Laguzza, A. & Holman, V. The Braden scale for predicting pressure sore risk. *Nurs Res* **36**, 205–210 (1987).