

Research Letters

Cite this article: Zhu J, Pradhan D, Emeruwa IO and Walsh BC (2025). The Hidden Bias of Missing Data in Crisis Standards of Care Simulation Studies: Not So Random, Rethinking Missing Data in Crisis Standards of Care Simulation Studies. *Disaster Medicine and Public Health Preparedness*, 19, e301, 1–3 <https://doi.org/10.1017/dmp.2025.10239>

Received: 07 June 2025

Revised: 11 August 2025

Accepted: 04 October 2025

Keywords:

Standard of Care; Public Health; Vital Statistics

Corresponding author:

Jianan Zhu;

Email: jz4698@nyu.edu

The Hidden Bias of Missing Data in Crisis Standards of Care Simulation Studies: Not So Random, Rethinking Missing Data in Crisis Standards of Care Simulation Studies

Jianan Zhu MS¹, Deepak Pradhan MD, MHPE, FCCP, ATSF²,

I. Obi Emeruwa MD, PhD, MBA³ and B. Corbett Walsh MD, MBE³

¹Department of Biostatistics, New York University School of Global Public Health, New York, NY, USA; ²Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, New York University Grossman School of Medicine, New York, NY, USA and ³Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Dear Editor,

Introduction: The COVID-19 pandemic highlighted the critical need for robust crisis standards of care (CSC) protocols to handle extreme strain when scarce resources require rationing. Evaluating how such policies might perform in the real world remains paramount; however, to date, study of their performance has been limited to retrospective cohort designs using virtual simulations.^{1,2} The Sequential Organ Failure Assessment score (SOFA)—a composite 0–24 score of organ dysfunction incorporating neurologic, pulmonary, cardiovascular, hematologic, hepatobiliary, and renal subscores—remains ubiquitous in nearly all crisis standards of care protocols^{3,4,5} despite concerns regarding the utilization and potentially exacerbating existing racial inequities.⁵ Existing simulation studies have handled missing SOFA values by either imputing zero or assuming data are missing at random, followed by complex computational statistical modeling.^{6,7} This approach may introduce significant bias, with larger outcome implications than missing data in other forms of medical research, as these values directly affect decisions on who receives life-sustaining therapies. Our study aims to better understand the frequency, structure, and consequence of missing data in CSC simulation studies.

Methods: We conducted a retrospective simulation study of all mechanically ventilated patients across a single New York City academic hospital system between March 1, 2020, and June 30, 2020 (surge period). SOFA scores were collected for each day of mechanical ventilation as close to 10 a.m. as possible, with a 24-hour window permitted for data inclusion. We defined the “crisis period” as initiating once 95% of the health system’s pre-pandemic ventilatory supply was utilized and lasting 2 weeks (crisis cohort), consistent with previously published CSC simulations under the New York State Ventilator Allocation Guidelines (NY).⁸ The “surge cohort” included all ventilated patients during the surge period but excluded those whose ventilation only occurred during the crisis period. The primary outcome was the daily frequency of missing SOFA subscores. The secondary outcome was the cumulative number of missing subscores over time throughout the surge.

Results: In total, 1671 patients were ventilated during the surge period: 1091 (65.3%) male, mean age 64.2 years, 887 (53.1%) COVID-19 positive, mean duration of intubation 13.5 days (Q25 1, Q50 6, Q75 17.5 days), and hospitalized mortality 887 (50.0%). In the crisis period, 674 patients were ventilated: 465 (69.9%; $P = .09$) male, mean age 63.7 years ($P = .46$), 571 (84.7%; $P < .0001$) COVID-19 positive, mean duration of intubation 19.8 days (Q25 5, Q50 12, Q75 25 days; $P < .0001$), and hospital mortality 395 (59.4%; $P < .0001$). [Figure 1](#) depicts the frequency of six missing SOFA categories throughout the surge period (primary outcome). Patients were significantly more likely to have at least one missing SOFA subscore during the crisis period compared to the non-crisis surge period ($P < .0001$). The neurologic subscore was more frequently missing during the crisis period compared to the surge period ($P < .0001$). Conversely, the hepatic category was more likely to be present during the crisis period ($P < .0001$). [Figure 2](#) presents the distributions of missing SOFA categories within the patient cohort (secondary outcome). While patients missing one or two categories during the crisis period were common, missing more than two categories was rare.

Discussion: To our knowledge, this is the first study to specifically evaluate missing data patterns and their implications in CSC simulation studies. We found that the severity of missing SOFA subcategories correlates with periods of peak clinical strain during the surge period, i.e., the crisis period, and suggests that data are not missing at random but are instead associated with operational pressures during crisis periods.

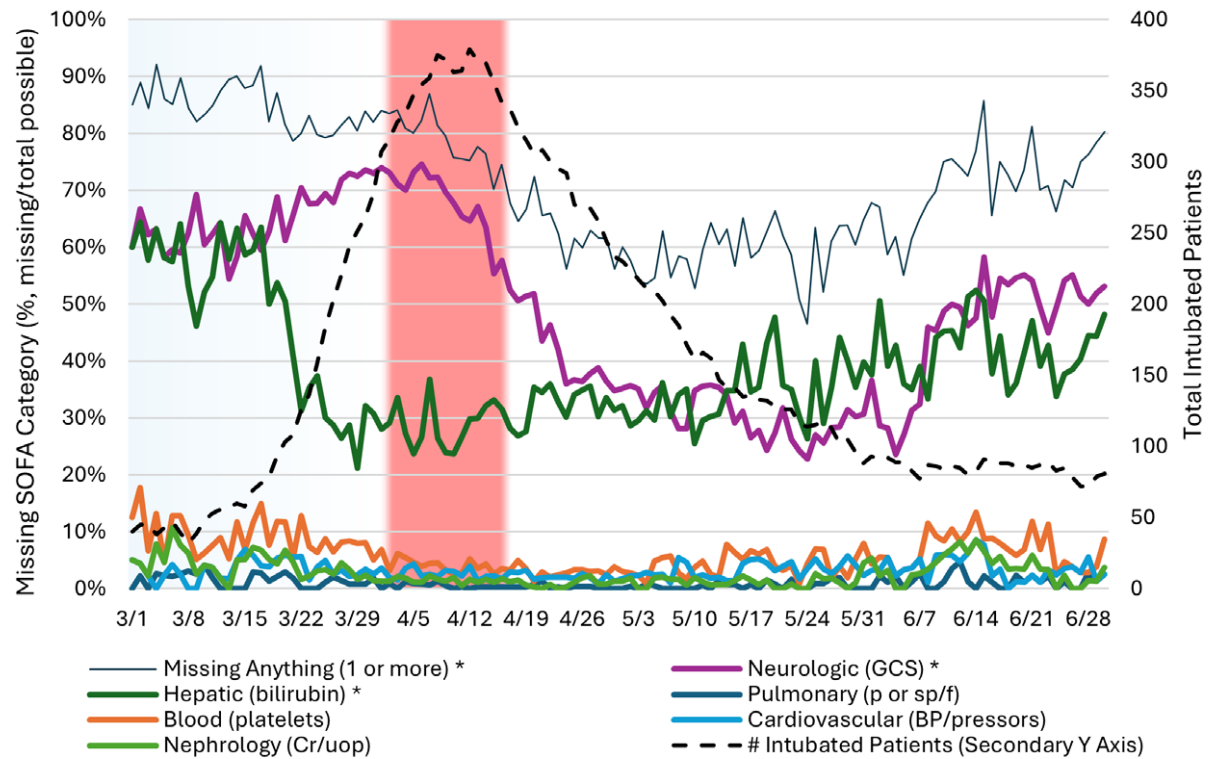


Figure 1. Frequency of missing SOFA categories throughout the Spring NYC 2020 COVID-19 surge.

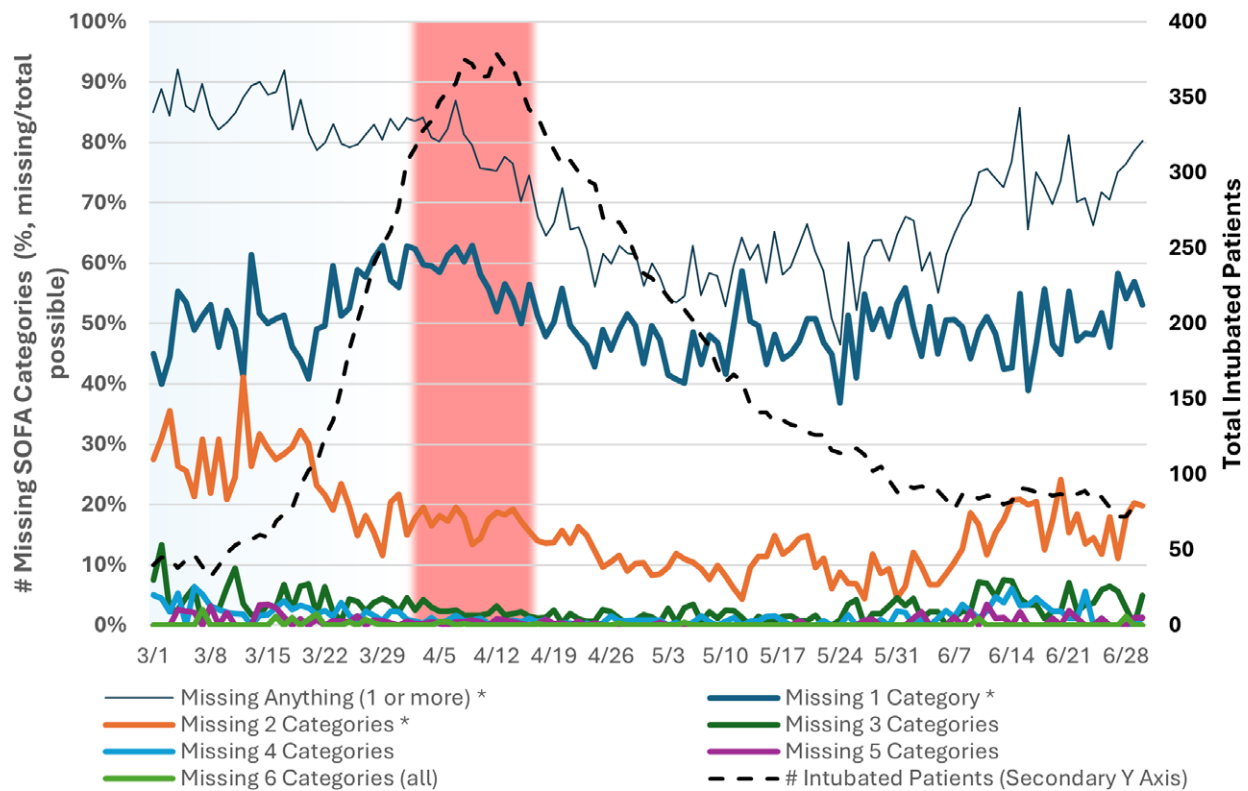


Figure 2. No. of missing categories throughout the Spring NYC 2020 COVID-19 surge.

Complex computational replacement methods, such as zero imputation, regression substitution, or multiple substitutions,^{6,7} often assume data are missing at random—an assumption contradicted by our data. Missing-at-random assumption is methodologically unsound and may introduce biases that distort CSC performance and undermine fair allocation.

Zero imputation, commonly used in prior CSC simulation studies, replaces missing values with 0, and thereby risks underestimating severity for patients whose missing values would otherwise raise their SOFA score.⁸ Zero imputation CSC protocols using a static SOFA score may introduce an optimism bias,⁹ as missing data on the day of evaluation might underestimate illness severity and increase a patient's triage priority. In CSC protocols relying on dynamic SOFA comparisons,¹⁰ where current and previous SOFA scores are compared to assess clinical trajectory, missing current-day data may falsely suggest improvement (optimism bias), while missing prior-day data may be falsely interpreted as suggesting clinical deterioration (pessimistic bias), and lead potentially to deprioritizing patients for limited resources. Both scenarios highlight how missing information can profoundly impact resource allocation decisions.

Importantly, our study suggests that missing data is likely to occur and will impact real CSC triage decisions. The neurologic subscore, most frequently missing during crisis periods, may be under-reported due to factors such as examination limitations under continuous neuromuscular blockade, efforts to limit provider exposure, or diagnostic time constraints when patients arrive in extremis.⁸ We were surprised that hepatobiliary scores were more common during the crisis period, hypothesizing that the relatively higher frequency may reflect automated lab ordering or protocolized care rather than clinical necessity.

Our pragmatic approach to missing data—imputing missing SOFA scores from the closest available time point, followed by zero imputation if none available—acknowledges that perfect data may not be feasible during crisis conditions. While this too could introduce potential bias, it better considers operational realities.

In summary, missing data in CSC simulations are common, non-random, and can be impactful on triage and reallocation decision-making of lifesaving therapies. A thorough understanding of their frequency and patterns is essential for accurately evaluating current CSC performance. Beyond addressing missingness, maintaining accurate, consistent, and timely data input is equally critical to ensuring the integrity and fairness of CSCs under real-world crisis conditions.¹¹ Together, these efforts form the foundation for

designing future CSCs that ensure equitable, accurate, and effective triage decisions.

Jianan Zhu, MS

Deepak Pradhan, MD, MHPE, FCCP, ATSF

I. Obi Emeruwa, MD, PhD, MBA

B. Corbett Walsh, MD, MBE

Competing interests. The authors have no conflicts of interest to disclose.

Reference

1. Bhavani SV, Luo Y, Miller WD, et al. Simulation of ventilator allocation in critically ill patients with COVID-19. *Am J Respir Crit Care Med*. 2021; **204**(10):1224–1227. doi:10.1164/rccm.202106-1453LE
2. Walsh BC, Pradhan D. The importance of incorporating patient throughput in crisis standards of care simulations. *Disaster Med Public Health Prep*. 2023; **17**:e390. doi:10.1017/dmp.2023.53
3. Raschke RA, Agarwal S, Rangan P, et al. Discriminant accuracy of the SOFA score for determining the probable mortality of patients with COVID-19 pneumonia requiring mechanical ventilation. *JAMA*. 2021; **325**(14):1469–1470. doi:10.1001/jama.2021.1545
4. Ashana DC, Anesi GL, Liu VX, et al. Equitably allocating resources during crises: racial differences in mortality prediction models. *Am J Respir Crit Care Med*. 2021; **204**(2):178–186. doi:10.1164/rccm.202012-4383OC
5. Miller WD, Han X, Peek ME, et al. Accuracy of the sequential organ failure assessment score for in-hospital mortality by race and relevance to crisis standards of care. *JAMA Netw Open*. 2021; **4**(6):e2113891. Published June 1, 2021. doi:10.1001/jamanetworkopen.2021.13891
6. Brinton DL, Ford DW, Martin RH, et al. Missing data methods for intensive care unit SOFA scores in electronic health records studies: results from a Monte Carlo simulation. *J Comp Eff Res*. 2022; **11**(1):47–56. doi:10.2217/ce-2021-0079
7. Molinnus D, Beulertz M, Bickenbach J, et al. Observational study of missing SOFA score data frequency in RCTs relative to ICU length of stay. *Sci Rep*. 2024; **14**(1):16160. Published July 12, 2024. doi:10.1038/s41598-024-67089-4
8. Walsh BC, Zhu J, Feng Y, et al. Simulation of New York City's Ventilator Allocation Guideline during the Spring 2020 COVID-19 surge. *JAMA Netw Open*. 2023; **6**(10):e2336736. Published October 2, 2023.
9. Valiani S, Terrett L, Gebhardt C, et al. Development of a framework for critical care resource allocation for the COVID-19 pandemic in Saskatchewan. *CMAJ*. 2020; **192**(37):E1067–E1073. doi:10.1503/cmaj.200756
10. New York State Task Force on Life and the Law. Ventilator allocation guidelines. November 2015. Accessed March 28, 2025. <https://int.nyt.com/data/documenthelper/6849-new-york-triage-guidelines/02cb4c58460e57ea9f05/optimized/full.pdf>
11. Hick JL, Hanfling D, Wynia M. Hospital planning for contingency and crisis conditions: crisis standards of care lessons from COVID-19. *Joint Commission Journal on Quality and Patient Safety*. 2022; **48**(6):354–361. doi: 10.1016/j.jcjq.2022.02.003