



Published in final edited form as:

Crit Care Med. 2018 February ; 46(2): 345–346. doi:10.1097/CCM.0000000000002892.

From Big Data to Artificial Intelligence: Harnessing data routinely collected in the process of care

Barret Rush, FRCPC MD MPH^{1,2}, David J. Stone, MD³, and Leo Anthony Celi, MS MD MPH^{4,5}

¹Division of Critical Care Medicine, St. Paul's Hospital, University of British Columbia, 1081 Burrard Street, Vancouver, BC, V6Z 1Y6, Canada

²Centre for Heart Lung Innovation (HLI), University of British Columbia, Vancouver, Canada

³Departments of Anesthesiology and Neurosurgery, University of Virginia School of Medicine, Charlottesville, VA

⁴Division of Pulmonary, Critical Care, and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA

⁵Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA

Keywords

ARDS; Big Data; Artificial Intelligence; Machine Bias

Thomas Frieden provides us with the most compelling reason to leverage data that is routinely collected in the process of care: “For much, and perhaps most of medical practice, RCT-based data are lacking and no RCT is being planned or is likely to be completed to provide evidence for action. ... [It] leaves practitioners with large information gaps for most conditions and increases reliance on clinical lore.” (1) With over 90% of care providers in the United States now using an electronic health record (EHR) system, health data is being collected at a scale (exabytes), resolution (up to 500Hz), and levels of heterogeneity, which are historically unprecedented. (2) The sheer magnitude of such data can leverage population data and facilitate the application of advanced algorithmic techniques which were previously not feasible due to small sample sizes (e.g. for deep learning). Indeed, recent investigations have reported impressive performances using algorithms to automate the diagnosis of skin cancers (3) and diabetic retinopathy. (4) Critically ill patients are an ideal population for clinical database investigations because while the data from ICUs is extensive, the value of many treatments and interventions remains largely unproven, and high-quality studies supporting or discouraging specific practices are relatively sparse. (5) The data-rich ICU

Send Correspondence to: Dr. Barret Rush MD, Division of Critical Care Medicine, St. Paul's Hospital, 1081 Burrard Street, Vancouver, BC, V6Z 1Y6, Canada, bar890@mail.harvard.edu.

Conflict of Interest: No authors disclose any conflicts of interest

Dr. Stone has disclosed that he does not have any potential conflicts of interest.

environment provides a potential area for uses of artificial intelligence (AI), a highly data dependent entity.

In this issue of Critical Care Medicine, Sottile et al. expand the scope of big data and machine learning (a current application of AI) to the realm of ventilator dyssynchrony (VD) in ventilated patients with and at risk for Acute Respiratory Distress Syndrome, or ARDS (7). In this small, single center prospective analysis, the authors explored the association between VD, delivered tidal volumes, and level of sedation in 62 patients meeting inclusion criteria. After analysis of 4.26 million breaths, the majority of which occurred in a proprietary mode delivered by a specific ventilator, the authors observed that compared with synchronous breaths, high tidal volume breaths (defined as $>10/\text{kg}$) were significantly more likely to be delivered with double triggered or flow limited breaths. Non synchronous breaths were observed in 34% of observed breaths. The rate of non-synchronous breaths was decreased with deep sedation; however, only neuromuscular blockade led to complete elimination.

The results described by Sottile et al. (7) must be interpreted with caution and analyzed in the context of the study design. This preliminary hypothesis generating study utilized a novel algorithm to detect VD; however, the generalizability and clinical implications of these findings are unknown. The clinical impact of VD as well as the impact of infrequent high tidal volume breaths in the era of low-tidal volume ventilation is not well known. Some authors have suggested that infrequent high-volume breaths may even be protective by promoting sustained alveolar recruitment. (8,9) Furthermore, the impact of ventilator mode, sedation level, and rate and type of VD needs to be examined. A large majority of the breaths analyzed were in APVCMV (adaptive pressure ventilation/controlled mandatory ventilation)- one particular and proprietary mode of ventilation so that extrapolation of these findings to other modes and models of ventilator are unknown. Other studies have demonstrated that altering ventilator mode and settings is a superior option to increased sedation in the reduction of VD. (10,11)

The authors noted (in the supplementary materials) that a non-uniform patient recruitment process occurred during the study because of unavoidable logistical issues. A careful examination of how this could have affected the make-up of the patient cohort was performed in order to address concerns that bias might have been introduced by feeding the model with data from a select group of patients for training. This is not just a theoretical concern in this case: Machine bias, a feared consequence of AI, arises from bias in creating the dataset with which an algorithm is developed. An investigative report was published last year on software that calculates risk assessment scores to inform decisions about who can be set free at every stage of the criminal justice system, from assigning bond amounts to granting parole. (12) The formula was particularly likely to falsely flag black defendants as future criminals, wrongly labeling them at almost twice the rate as white defendants. White defendants were mislabeled as low risk more often than black defendants. It is easy to see how this kind of fundamental machine bias in decision support algorithms could pose potential harm to patients. Imagine a tool that predicts response to treatment with variable accuracy based on a patient's ethnicity leading to withholding of the treatment to those who would benefit from it. This highlights the importance of being able to detect flaws in the

model and biases in the data as we debate whether interpretability should be required of AI tools, especially in healthcare. Shining the light on the black box to understand how deep learning algorithms classify, predict or optimize is a rapidly growing field in AI with methodologies such as layer-wise relevance propagation and sensitivity analysis. (13)

One currently unfulfilled but extremely exciting promise of data driven, precision medicine is that it will provide an important basis for the creation of clinical decision support tools that appropriately employ AI. In addition to careful analysis of such features regarding costs, efficiencies, risks, and a variety of clinical outcomes, the introduction of such modalities must be done with more careful thought and consideration for user and workflow issues than has heretofore been provided in current EHRs. The introduction of such AI based functions applied to routinely collected data will require appropriate responses by the healthcare system in terms of integrating these features into medical care in as seamless, smart, and painless way as possible. They should be created and utilized where there is a need rather than having to drum up a need for already created features that were interesting or fun to develop. Hospital cultures, along with medical undergraduate and graduate education systems, will need to adapt to a decision support environment that will be quite different in many ways. (14,15) People will learn differently, be trained differently, and practice differently in an environment of data driven AI, and we are just beginning to learn how to do this in the early stages of the use of digital tools in medicine. Meanwhile, those at the forefront of the health data revolution must earn and maintain clinicians' and society's trust, and demonstrate that careful and complete data collection, as well as sharing and reuse, are necessary steps to improve patient care.

Acknowledgments

Drs. Rush and Celi received support for article research from the National Institutes of Health.

References

1. Frieden T. Evidence for Health Decision Making — Beyond Randomized, Controlled Trials. *N Engl J Med*. 2017; 377:465–475. [PubMed: 28767357]
2. Cowie MR, Blomster JI, Curtis LH, Duclaux S, Ford I, Fritz F, Goldman S, Janmohamed S, Kreuzer J, Leenay M, Michel A, Ong S, Pell JP, Southworth MR, Stough WG, Thoenes M, Zannad F, Zaleski A. Electronic health records to facilitate clinical research. *Clin Res Cardiol*. 2017; 106(1): 1–9.
3. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017; 542(7639):115–118. [PubMed: 28117445]
4. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*. 2016; 316(22):2402–2410. [PubMed: 27898976]
5. Celi LA, Mark RG, Stone DJ, et al. “Big data” in the intensive care unit. Closing the data loop. *Am J Respir Crit Care Med*. 2013; 187:1157–60. [PubMed: 23725609]
6. Ghassemi M, Celi LA, Stone DJ. State of the Art Review: The data revolution in critical care. *Crit Care*. 2015; 19:118. [PubMed: 25886756]
7. Sottile P, Albers D, Higgins C, Mckeehan JMM. The Association between Ventilator Dyssynchrony, Delivered Tidal Volume, and Sedation using a Novel Automated Ventilator Dyssynchrony Detection Algorithm. *Crit Care Med*. 2017 IN PRESS.

8. Mauri T, Eronia N, Abbruzzese C, et al. Effects of Sigh on Regional Lung Strain and Ventilation Heterogeneity in Acute Respiratory Failure Patients Undergoing Assisted Mechanical Ventilation. *Crit Care Med.* 2015; 43:1823–31. [PubMed: 25985386]
9. Spieth PM, Carvalho AR, Pelosi P, et al. Variable tidal volumes improve lung protective ventilation strategies in experimental lung injury. *Am J Respir Crit Care Med.* 2009; 179:684–93. [PubMed: 19151194]
10. Vaschetto R, Cammarota G, Colombo D, et al. Effects of propofol on patient-ventilator synchrony and interaction during pressure support ventilation and neurally adjusted ventilatory assist. *Crit Care Med.* 2014; 42:74–82. [PubMed: 23982026]
11. Chanques G, Kress JP, Pohlman A, et al. Impact of ventilator adjustment and sedation-analgesia practices on severe asynchrony in patients ventilated in assist-control mode. *Crit Care Med.* 2013; 41:2177–87. [PubMed: 23782972]
12. Angwin, J., Larson, J., Mattu, S., Kirchner, L. Machine Bias. Available at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Accessed Oct 20 2017
13. Ribeiro, MT., Singh, S., Guestrin, C. “Why should I trust you?”: Explaining the predictions of any classifier; Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; San Francisco, CA, USA. August 13–17, 2016; 2016. p. 1135-1144.
14. Moskowitz A, McSparron J, Stone DJ, Celi LA. Preparing a new generation of clinicians for the era of big data. *Harvard Medical Student Review.* 2015; 2(1):24–7. [PubMed: 25688383]
15. Obermeyer Z, Lee TH. Lost in thought- the limits of the human mind and the future of medicine. *NEJM.* 2017; 377:1209–11. [PubMed: 28953443]