Published in final edited form as:

Crit Care Med. 2017 May ; 45(5): 912–913. doi:10.1097/CCM.0000000000002351.

Will Artificial Intelligence Contribute to Overuse in Healthcare?*

Matthieu Komorowski, MD and

Laboratory of Computational Physiology, Institute for Medical Engineering and Science,
Massachusetts Institute of Technology, Cambridge, MA
Department of Surgery and Cancer,
Faculty of Medicine, Imperial College London, London, United Kingdom

Leo Anthony Celi, MD

Laboratory of Computational Physiology, Institute for Medical Engineering and Science,
Massachusetts Institute of Technology, Cambridge, MA
Division of Pulmonary Critical Care and
Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, MA

Keywords

alarm; artificial intelligence; atrial fibrillation; overdiagnosis; overuse

Atrial fibrillation (AF) and atrial flutter (AFL) are relatively common in critically ill patients and have been previously associated with worse outcomes in the ICU (1). In the research presented by Moss et al (2), in this issue of *Critical Care Medicine*, the authors employed an automated rhythm classification methodology to analyze continuous electrocardiograms (ECGs) of critically ill patients to detect AF and AFL. The atrial arrhythmias were classified as either new-onset or recurrent/persistent depending on whether they had been previously diagnosed prior to the ICU admission. New-onset AF was further classified as clinical or subclinical depending on whether there was documentation of that rhythm in the electronic health record during the ICU admission. Using propensity matching, the authors then investigated the association between the different categories of the atrial arrhythmias and clinical outcomes. Of note, 7.5% of all ICU admissions had new-onset subclinical AF as identified by the automated rhythm classification.

The authors demonstrate how machine learning applied to routinely captured clinical data can generate new information and potentially new insights that are missed by the clinician who cannot continuously observe all ongoing ECG patterns. Only computers can process the volume of data collected in the process of care. This work is an illustration of how artificial intelligence (AI) has the potential to lead to the development of tools to assist clinicians and potentially improve patient outcomes. In a recent editorial, Beam and Kohane (3) discussed the opportunities presented in translating AI into clinical care. Despite several decades of research and hype, the AI field has failed to deliver on its promises of automated and improved disease detection, more effective monitoring, and efficiency boosts in workflow (4). Still, significant, if slow, progress has been achieved in recent years, especially in the

*See also p. 790.

area of computer vision, where algorithmic advances have started to trickle into areas such as medical image processing in fields like radiology and pathology (5).

During the course of clinical monitoring, there is significant opportunity to generate a variety of signals with little to no clinical relevance. Therefore, caution is warranted when interpreting the results of research like that of Moss et al (2), especially so before considering the implementation of such analytic tools into practice. Some findings of the research by Moss et al (2) are important to highlight. The magnitude of the impact of AF on outcomes was greatly reduced or disappeared after adjustment for severity of illness and the use of vasopressor agents. Of note, only new-onset “clinical” AF was associated with hospital mortality and longer length of stay in this study. Outcomes were not significantly worse among patients with new-onset subclinical AF (detected by the algorithm but missed, or at least never documented, by the clinicians) in propensity-adjusted regression analysis. In this latter situation, AF could simply represent a surrogate for patient severity, but one that does not specifically impact clinical outcomes. In multivariate logistic regression, the only rhythm abnormality that remained significantly associated with hospital mortality was new-onset, clinical AF, and even that with a borderline CI barely dodging the 1.0 threshold for statistical significance.

Even more important, the present study did not investigate what is perhaps the most important issue regarding whether detection and treatment of new-onset clinical AF in critically ill patients alter outcomes. Obviously, the article could not address the treatment of subclinical AF since, by definition, the clinicians were not aware of the dysrhythmia. The article suggests that new-onset AF not picked up by clinicians (i.e., subclinical AF) is not associated with increased mortality or length of stay, and therefore, its detection may not add any value. In fact, there is a possibility that better detection of new-onset AF not currently detected in the ICU might only lead to unnecessary treatment, added expense and potential for treatment-related adverse events.

If detecting and treating new-onset AF or AFL do not or very minimally improve outcomes, then this intervention could become yet another example of “overuse” of healthcare resources, where investigations lead to diagnoses and treatments that are either unnecessary (because they are ineffective), and unfortunately often detrimental, to both the patient and the healthcare segment of the economy (6). Overdiagnosis is a growing problem, one which may increase further as a result of technical improvements in the sensitivity of detection methods (6–8). An important distinction exists between overdiagnosis and a false positive test result. Although a false positive refers to a healthy patient being diagnosed with a disease (a false alarm), overdiagnosis is the detection of an actual disease (according to some recognized definition of illness) that, if it had remained undetected, would not have affected a person’s life (8). The downsides of overdiagnosis include patient anxiety, the harm from further testing and unnecessary treatment, and the opportunity cost of wasted time on the part of both patient and provider and healthcare resources that could be better used to treat or prevent genuine illness (7). Quantifying overdiagnosis is extremely challenging, but estimates suggest, for example, that up to a third of the cancers detected via screening may be overdiagnosed (7) and that 30% of people with a diagnosis of asthma may not have it (9). Medical overuse also includes overtesting. To illustrate, over 50% of

preoperative testing prior to cataract surgery is deemed unnecessary given the surgical and anesthetic risk (8).

Efforts are now targeting the issue of overdiagnosis. Conferences are held (such as the National Cancer Institute meeting on overdiagnosis in 2012 and the annual International Preventing Overdiagnosis Conference since 2013), and initiatives have been launched by medical journals, including the BMJ's "Too Much Medicine" project since 2002, JAMA Internal Medicine's "Less is More" Collection in 2010, and The Lancet's "Right Care" initiative in January 2017, which aims to improve global health by targeting overuse and underuse of medical treatment (10).

Such educational campaigns will hopefully convince providers to recognize that all medical specialties, including critical care, likely overtest, overdiagnose, and overtreat in some way or another. The use of AI may well contribute to this problem by discovering patterns undetected by the human mind that are not actually causing problems and never will. Clinically meaningful advances in this field will be an iterative process, where new algorithms are developed and systematically tested in real clinical settings for relevance against hard clinical endpoints. Only those that demonstrate value should be refined and improved before consideration for systematic bedside implementation.

Acknowledgments

Dr. Komorowski received support from the U.K. Engineering and Physical Sciences Research Council and is the recipient of an Imperial College President's PhD Scholarship. Dr. Celi received support from the National Institutes of Health.

References

1. Shaver CM, Chen W, Janz DR, et al. Atrial fibrillation is an independent predictor of mortality in critically ill patients. *Crit Care Med*. 2015; 43:2104–2111. [PubMed: 26154932]
2. Moss TJ, Calland JF, Enfield KB, et al. New-Onset Atrial Fibrillation in the Critically Ill. *Crit Care Med*. 2017; 45:790–797. [PubMed: 28296811]
3. Beam AL, Kohane IS. Translating artificial intelligence into clinical care. *JAMA*. 2016; 316:2368–2369. [PubMed: 27898974]
4. Schwartz WB. Medicine and the computer. The promise and problems of change. *N Engl J Med*. 1970; 283:1257–1264. [PubMed: 4920342]
5. Jha S, Topol EJ. Adapting to artificial intelligence: Radiologists and pathologists as information specialists. *JAMA*. 2016; 316:2353–2354. [PubMed: 27898975]
6. Marcus PM, Prorok PC, Miller AB, et al. Conceptualizing overdiagnosis in cancer screening. *J Natl Cancer Inst*. 2015; 107:djv014. [PubMed: 25663695]
7. Moynihan R, Doust J, Henry D. Preventing overdiagnosis: How to stop harming the healthy. *BMJ*. 2012; 344:e3502. [PubMed: 22645185]
8. Brownlee S, Chalkidou K, Doust J, et al. Evidence for overuse of medical services around the world. *Lancet*. 2017 Jan 6. Epub ahead of print.
9. Aaron SD, Vandemheen KL, Boulet LP, et al. Canadian Respiratory Clinical Research Consortium. Overdiagnosis of asthma in obese and nonobese adults. *CMAJ*. 2008; 179:1121–1131. [PubMed: 19015563]
10. Berwick DM. Avoiding overuse—the next quality frontier. *Lancet*. 2017 Jan 6. Epub ahead of print.