



Applying machine learning to continuously monitored physiological data

Barret Rush¹ · Leo Anthony Celi^{2,3} · David J. Stone⁴

Received: 12 March 2018 / Accepted: 8 November 2018 / Published online: 11 November 2018
© Springer Nature B.V. 2018

Abstract

The use of machine learning (ML) in healthcare has enormous potential for improving disease detection, clinical decision support, and workflow efficiencies. In this commentary, we review published and potential applications for the use of ML for monitoring within the hospital environment. We present use cases as well as several questions regarding the application of ML to the analysis of the vast amount of complex data that clinicians must interpret in the realm of continuous physiological monitoring. ML, especially employed in bidirectional conjunction with electronic health record data, has the potential to extract much more useful information out of this currently under-analyzed data source from a population level. As a data driven entity, ML is dependent on copious, high quality input data so that error can be introduced by low quality data sources. At present, while ML is being studied in hybrid formulations along with static expert systems for monitoring applications, it is not yet actively incorporated in the formal artificial learning sense of an algorithm constantly learning and updating its rules without external intervention. Finally, innovations in monitoring, including those supported by ML, will pose regulatory and medico-legal challenges, as well as questions regarding precisely how to incorporate these features into clinical care and medical education. Rigorous evaluation of ML techniques compared to traditional methods or other AI methods will be required to validate the algorithms developed with consideration of database limitations and potential learning errors. Demonstration of value on processes and outcomes will be necessary to support the use of ML as a feature in monitoring system development: Future research is needed to evaluate all AI based programs before clinical implementation in non-research settings.

Keywords Machine learning · Intensive care · Monitoring · Physiological · Patient monitoring · Artificial intelligence

✉ Barret Rush
bar890@mail.harvard.edu
Leo Anthony Celi
lceli@bidmc.harvard.edu
David J. Stone
djs4v@virginia.edu

- ¹ Division of Critical Care Medicine, St. Paul's Hospital, University of British Columbia, 1081 Burrard Street, Vancouver, BC V6Z 1Y6, Canada
- ² Laboratory for Computational Physiology, MIT Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
- ³ Division of Pulmonary Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA
- ⁴ Departments of Anesthesiology and Neurosurgery, University of Virginia School of Medicine, Charlottesville, VA 22904, USA

1 Introduction

Machine Learning (ML) is a term likely spoken of more than understood. ML is most simply defined as the use of various statistical techniques that can be employed to make predictions and decisions based on similarities in what is being analyzed to what has previously been observed. Ideally, artificial intelligence (AI), represents an application that has the capacity to learn, dynamically updating and changing its rules based on evaluation of its performance. However, the ML currently incorporated in monitoring systems does not have this learning capability. In an ML based monitoring system, the goals of the application remain the same (i.e. the detection of clinical anomalies), but the precise nature of the approach to reaching those goals changes. ML applications are created with a variety of techniques; these will not be addressed in this paper in detail. While ML lacks many of the nuances of human intelligence, it also

possesses advantages over human intelligence: a potential for very powerful memory and computational power as well as the ability to examine large volumes of data continuously while identifying subtle and dynamic patterns that would likely remain obscure to humans.

Clearly, the current application of expert static rules to alarm thresholds in conventional monitoring does not involve ML. A fundamental definitional issue arises when ML is utilized to formulate algorithms that are subsequently incorporated in an expert system in a static fashion. We define this kind of monitoring system as one possessing a hybrid format, incorporating elements derived from ML approaches that are subsequently applied in static fashion to newly observed data. In monitoring applications, the hybrid format—application of ML as a “fixed” software (one that does not learn from and change on the basis of its inputs and outputs, and employs an unchanged or ‘fixed’ algorithm) appears to be the dominant, if not exclusive, form so far reported in these very early stages of the use of ML in clinical monitoring.

Many challenges face the implementation of ML tools in the healthcare setting. Regulatory constraints from governmental agencies, insurance providers and hospitals themselves may stifle burgeoning advancements in this area. The reliance on potentially unreliable datasets for algorithm function can be the source of so-called machine bias. Thus, errors and gaps in data input and structure can lead to erroneous outputs from otherwise well-designed algorithms. Additionally, where does clinician decision and input fall in a ML centered process of care? Will these algorithms only serve to augment the data synthesis skills of the physician or will the clinician depend on output from these algorithms to make a therapeutic decision? Will ML algorithms detect subclinical events leading to over-diagnosis and unnecessary or potentially harmful treatments? How will clinicians respond to the incorporation of ML into the clinical process of monitoring? Very few, if any, ‘true’ self-learning ML applications have been conceived, and it appears that none are in actual clinical practice outside of the research sphere at the time of this paper’s publication. This review attempts to examine some of the nascent work implementing early ML techniques into actual clinical monitoring systems in adult patients. This is by no means an exhaustive survey and evaluation of all the different approaches to physiologic monitoring and clinical alert development. Additionally, as the field of ML and AI is rapidly advancing, during the necessary delay in publication of this work further publications in this field may have been released which were not captured.

2 Why apply ML to monitoring?

Why do we monitor patients at all? We want to know if problems are present or developing, further identify the nature of these problems, and track responses to our interventions. Can continuous physiological monitoring move beyond the current practice in which alarms (often false positive) are generated when rigid, static value thresholds are crossed? Are we able to make engineering design advances to support the interpretation of basic vital sign values as well as to capture and analyze complex dynamic patterns that may represent important clinical entities [1]? Should we now proceed further in the practice of monitoring to track and more thoroughly analyze the trajectory of disease and host response at deeper, more data driven levels of engineering expertise? We would suggest that the analytic tools of ML have the potential both to support the workflow inherent in interpreting vital signs, as well as providing insights into data patterns and complexities that are beyond the perceptive capabilities of the normal human clinical observer [2, 3]. While ML is demanding in terms of input data volume requirements, that is not a problem in this case, especially if the wealth of continuous monitor data is supplemented by pertinent data from the EHR (electronic health record) or the increasingly ubiquitous wireless wearable devices that measure physiologic signals.

3 Other non-ML data driven approaches

The potential use of ML techniques to improve the delivery of hospital care by utilizing complex data is but one method of possible data driven solutions. Other applications of data science and artificial intelligence may be equally as useful for improving patient care. Methods modelling expert based opinions as well as mathematical physiological models have been and are also employed to inform clinical decision-making. Large static administrative databases can also be leveraged to improve clinical processes. ML is one method available to researchers attempting to utilize the vast amount of clinical data created in the modern health care setting.

Rigorous evaluation and experimentation will be needed to select the best AI techniques to handle the vast amounts of data created in clinical medicine. Each methodology has strengths and weaknesses which will impact the value of the technology but a full comparative discussion of these methods is beyond the scope of this review. ML itself has significant inherent limitations that must be taken into account. For example, as a data driven entity,

errors can introduce bias into algorithms which may lead to biased information. No matter the method of AI, human clinical knowledge and expertise will always be necessary in the care of patients. It is the goal of ML or any AI assisted technology to augment and enhance rather than replace the clinical acumen and abilities of the health care team.

4 Current state of ML and monitors

In this section, we review a number of monitor-related ML applications for adult patients that have been published or are available publicly online as working papers. We acknowledge that these entries are not inclusive of everything that has been published in this area. The general theme of what has been done is that ML has been used to create algorithms for monitoring that are trained and validated on retrospective clinical data, and then applied as static expert systems. Furthermore, most of these hybrid creations have not been clinically implemented, or at least, the results of such use in pilot version has not yet been reported in the literature. We present the examples of ML in order of their anticipated clinical usefulness in the next few years.

4.1 Sepsis identification and clinical deterioration

Identification of new episodes of sepsis in hospitalized, especially ICU patients, is an important but challenging problem. Laboratory and electronic health record (EHR) based identification schemes can be beneficial, but are limited due to the time delay between sepsis onset and laboratory value changes and availability [4–6]. Addition of real-time vital sign monitoring to these algorithms can improve the prognostic accuracy of these models [7]. Recently, machine learning algorithms incorporating high resolution vital signs and EHR data have demonstrated in multiple retrospective cohorts the ability to accurately predict the onset of sepsis 4–12 h before clinical identification [8].

A small randomized controlled trial of 142 patients demonstrated that addition of a ML algorithm to existing EHR based sepsis identification systems was associated with decreased hospital length of stay and mortality [9]. The use of ML for this purpose is likely to increase as this problem is complex, dynamic (and time critical), and of high acuity.

4.2 Reducing false alarms

False alarms continue to be a problem whenever patients are monitored [10]. The interpretation of monitor data has traditionally centered on responding to data points that lie outside the bounds of what are considered the normal values for that particular variable (e.g. HR (heart rate), BP (blood

pressure), etc.). ML has the potential ability to make the interpretation of designated ‘abnormal’ values smarter (i.e. improve what we now called alarms) as well as give interpretive meaning to patterns within developing/dynamic and combinatorial/complex physiological values. This will aid clinicians in understanding what is going on in a diagnostic sense, what may be developing, and in conjunction with EHR/outcome data, start to help in optimizing responses. Drew et al. noted that an important cause of false alarms was the fact that alarms were not sufficiently precisely tailored to individual patients- ML provides at least the potential for more personalized medicine in this regard. Rigorous validation studies will be required to ensure that ML is accurately detecting alarms and abnormal values in order for the models to perform correctly [10].

Technical issues such as artifact, electromagnetic interference and poor recording quality present another issue in interpreting alarms. Borges and Brusamarello employed ML in conjunction with other methods to fuse physiological data to address the first two problems [11]. They found that the ML/neural networks approach to the issue produced the best results i.e. the most robust heart rate readings. Chen and co-workers employed machine learning to successfully classify monitor alerts as real vs artifactual [12]. In another publication, the research team used similar methods on the problem and achieved similar results, presumably providing increased evidence of the reliability of their approach [13]. Ansari et al. used machine learning to reduce the incidence of false arrhythmia alarms [14]. The application of ML to the area of alarms is probably the most extensively studied to date. Surprisingly, ML technology for this purpose has not yet been incorporated into clinical monitors, at least based on what has been published and what we have observed in available devices for clinical purposes.

4.3 Delirium identification

Delirium is a common entity in the ICU which has been linked to a myriad of adverse outcomes including prolonged ICU stay, poor cognitive recovery, increased hospital costs, as well as mortality [15]. While significant advances in delirium screening and management have been implemented over recent years, early accurate identification of delirium remains a significant problem in the ICU [16].

Real-time heart rate variability assisted by ML techniques have been investigated in preliminary studies as a method of delirium identification [17]. Excess sympathetic activity associated with delirium can be detected in real-time by continuous monitoring of vital signs. The authors demonstrated that their preliminary ML algorithm was equivalent to current delirium screening tools. Future studies may be able to incorporate more variables and eventually may be able to detect delirium before it is detected clinically. The main

question to clinicians is whether this information is actionable. Measures to prevent delirium should be undertaken for all ICU patients regardless of risk.

4.4 Waveform analysis-detecting ventilator dyssynchrony

Dyssynchrony occurs when a patient receiving mechanical ventilation breathes in a manner that is uncoordinated with the ventilator to a degree that has potentially deleterious implications. Sottile et al. employed machine learning to examine the details of the waveforms produced by various types of dyssynchrony [18]. The algorithm produced was employed to examine the data retrospectively to detect and classify these events. Problems with patient recruiting for this study raised the issue of selection bias in creating the data set, an important consideration for the development of ML [19]. It is essential to avoid machine bias to avoid creating decision support algorithms that could pose potential harm to patient subsets. Algorithms should therefore be evaluated with regard to their performance across different cohorts.

4.5 Sedation management in the ICU

While the harms of continuous sedation in the ICU have been well established, the use of sedative medications remains a cornerstone of the management of critically ill patients. Currently, nurse driven scoring systems such as the Richmond Agitation Sedation Scale (RASS) form the basis of quantifying levels of sedation. These types of scales are subjective and are often only calculated once or twice per day. It has also been shown that these nurse driven scores perform poorly in very unresponsive to comatose patients [20]. Utilizing continuous streams of data from ICU patients may allow for real-time analysis of patient sedation levels and may help to optimize the use of sedative medications.

Preliminary hypothesis generating studies have been conducted utilizing real-time heart rate variability and ML as a method to classify levels of sedation in ICU patients [21, 22]. Based on heart rate variability alone, the algorithms were able to classify patients into deep sedation versus low sedation levels with 75% accuracy. Electroencephalogram (EEG) monitoring has also been studied as a method for quantifying sedation levels [23]. Use of continuous EEG monitoring as well as bispectral index (BIS) monitoring have been investigated as adjuncts for sedation monitoring [24]. Future studies incorporating multiple data input variables may allow further refinement in these algorithms. Accurate, automatic, continuous sedation interpretation may allow for staff to monitor patients more closely and avoid complications related to over or under sedation.

4.6 Detecting overlooked arrhythmias-undocumented atrial fibrillation

Moss and co-workers used an automatic rhythm classification method to detect atrial fibrillation and atrial flutter in critically ill patients [25]. The rhythms were classified as new if not extant prior to intensive care unit (ICU) admission, and as subclinical if undocumented in the chart. This example of a machine learning generated algorithm was able to detect arrhythmias that were not detected (i.e. in the sense that if not documented, they had not been detected) by the clinicians at the bedside. This ‘failure’ by the clinicians to capture transient events was due to the volume of and need for continuous capture of the data. Notably, while outcomes were affected by those new onset anomalies that were documented (and therefore had been detected by clinicians), outcomes were not affected by the occurrence of the subclinical arrhythmias. An accompanying editorial pointed out that the ML-enabled identification of such subclinical events might lead to over-diagnosis and unnecessary treatments, potentially creating further expense and potential adverse events based on these treatments [26]. While the arrhythmia detection modality was created by machine learning, it was itself not ML in the sense that it represents a static expert system that is not constantly calibrated. Any benefit analysis of ML assisted techniques must include such rigorous scientific study to ensure that clinical outcomes are improved in a demonstrable and reproducible way before being widely adopted.

5 Potential ML monitoring applications

The remainder of this piece is explicitly speculative, focusing on how ML applied to monitor data could improve clinical workflows and decision making. This includes creation of an integrated system where monitor and EHR data is seamlessly employed for this purpose. ML has the potential to discover significant patterns in monitor data that human clinicians cannot detect- this kind of deeper analysis should benefit from the contextual support of EHR data. EHR data is also particularly important to provide the kind of outcomes feedback that will be required for real learning by clinical ML.

5.1 Supporting the workflow of monitor data interpretation

A great deal of time and effort is spent in the clinical interpretation of vital signs in real time by nurses and physicians caring for monitored patients. Accomplishing this interpretation more efficiently and effectively in each individual case should therefore, at the very least, improve clinical

workflows. Tele-ICU provides an extreme example of the challenge of copious and unrelenting monitor data input that requires interpretation [27, 28]. In this setting, dozens of monitors display electronic patient records continuously, presenting varying degrees of patient vital sign anomalies, each one requiring careful attention and analysis. As a tele-ICU real time population management clinician, one is constantly trying to decide issues such as why (remote) patients x, y, and z have fast heart rates (tachycardia), and why patients a, b, and c have low blood pressures (hypotension), for example. While ML is not going to completely close the loop on addressing these challenges, it has the potential to significantly reduce the detailed work required for each individual analysis by analyzing the appropriate supporting data, as well as potentially providing suggestions as to the underlying pathophysiological events that have produced these anomalies. It is ultimately up to the clinician to determine diagnosis and treatment; however, ML has the potential to reduce the required effort, time expenditure, and repetitive busy work of gathering information. This decreased workload could enable clinicians to spend more time actually thinking rather than rote assembling the pieces of the puzzle required. ML should also be able to provide a first approximation of the clinical situation based on this assembly and analysis process, which is then amenable to human review before clinical actions are taken (or not).

Integrating monitors in a bidirectional fashion with the EHR, including treatment and outcomes data, would be required to leverage the full value of health data. ML likely will be excellent, fast, and reliable at examining patterns such as a combination of increasing heart rate and decreasing blood pressure (or any combination of data elements) over time, particularly while incorporating other relevant information such as serum troponin levels, cardiac medications, infusion pump information, and other pertinent available data elements in real time. The ultimate goal would be to compare observed patterns with historical (short and long term, as well as population) outcomes to drive optimally informed interventions. This goal also applies to the next section which addresses more complex monitor data anomalies.

5.2 Waveforms, patterns, dynamics, and complexity

Another use case for ML could be for detecting, mining, and analyzing complex changes over long, or relatively short but high data density, time intervals (e.g. for HR variability patterns), that we humans are simply not capable of doing. We realize that it may be obvious to point out that computers are better than humans at rapidly capturing, storing, and analyzing large amounts of data, but can these obvious qualities be applied specifically and intelligently to clinical monitoring? It is incumbent on us to establish whether there are vital sign

constellations and dynamics that represent pathophysiological footprints that predictably precede events like cardiac arrest, sepsis, hemorrhagic shock, or respiratory failure: Even when presented in the format of relatively reviewable time series information, it is difficult or impossible to perform this analysis by routine clinical viewing and reviewing of screen data. With the aid of computers, we can learn what these data footprints consist of, and what they mean, with consideration of useful levels of probability and uncertainty [1]. These deeper and subtler patterns could then contribute to a more proactive, efficient, and informative diagnostic process.

5.3 Incorporating the electronic health record

In order to receive the full benefits of ML supported monitoring systems, a continuously analyzed monitor database could be dynamically created in order to compare the patterns of monitor data with the clinical and outcomes data available from the EHR. This kind of data-rich, population-level approach could shed much light on the nature of physiological data, its behavior, its relationships to disease trajectory, and optimal responses to what the monitors are telling us. While we have long known about certain broad stroke relationships between particular vital signs and clinical scenarios, we still lack a precise handle on the magnitude and speed of these changes, as well as their bounds of confidence, in terms of disease evolution. If a particular dynamic constellation of monitor findings occurs, being able to subsequently assign probabilities for actionable conditions based on these constellations would represent a significant clinical advance.

In addition, as ML systems are populated more fully with pertinent data, they may develop the ability to suggest further data acquisitions that should be carried out to narrow down the confidence interval around any single diagnosis. As described above, clinicians have spent and continue to spend an inordinate amount of time and effort ‘reinventing the wheel’ of associating vital sign changes with clinical events. This particular clinical process can clearly be supported and improved with artificial intelligence methodologies such as machine learning, computer vision and natural language processing. We are already dependent on monitors to tell us what the heart rate is without putting a hand on the patient- machines can also augment clinicians’ critical thinking and reduce cognitive bias without posing a threat to clinicians’ decision-making autonomy. The application of ML to monitor data should be further informed by available context, continuously learning from prior and incoming EHR data. Currently, we are severely underleveraging the data that is laboriously entered into EHRs and which could be gainfully employed to enhance the meaning and value of monitored data. We understand that full interoperability of

data among EHRs and monitors is not currently available, but this is a regulatory issue rather than a true technical or conceptual roadblock.

5.4 Caveats and next goals

In an age of increasingly ubiquitous ML, newer engineering modalities can be leveraged to improve the interpretation of monitored data in a way that supports clinical workflow and decision-making. The implementation of these advances must not add to over-diagnosis and overuse of healthcare resources. Indeed, by improving the sensitivity and information value of detection methods, ML could increase the risk of detecting conditions that, if they had remained undiagnosed, would not have affected the patient's outcomes or quality of life [19]. Care must therefore be taken to avoid the introduction of the kinds of unhelpful information that can lead to patient anxiety, unnecessary treatment, added expense, and potential for treatment-related adverse events. Clinical medicine has only recently begun to systematically address these issues for traditional practices; overdiagnosis and overtreatment will be brought even further to the forefront by modalities that leverage large amounts of granular data (some unreliable or erroneous, some of uncertain value, and some of no import) such as ML. It is also essential to avoid bias so far as possible in the creation of ML [26].

One major barrier to the widespread utilization of ML models in healthcare lies in the direct operationalization issues faced by many facilities. Outside of research settings, most hospitals are unable to currently gather or interpret real-time physiologic data. Large capital investments will need to be made in order to have the technological capabilities to implement ML models across diverse healthcare settings. Additionally, hiring clinical and technical staff that are trained in the implementation of these new models represents a further barrier to rapid implementation of this technology. The introduction of a conceptually new approach to monitoring interpretation will also face appropriate push-back from the conservative culture of medicine, requiring robust evidence to support its implementation and maintenance in the clinical sphere.

Another major concern about ML applied to health care functions can be described as the 'black box' factor—it may be very difficult to impossible to understand how and why ML came to a particular conclusion [29]. The rebuttals to this consist of claims that ML is no less transparent than the way traditional doctors think—that doctors themselves don't entirely know how they came to their own conclusions, and that ML, unlike most clinical decision making currently, is thoroughly and rigorously based on data. One approach to this problem is to attempt to build some degree of explicit explainability into the ML so that the user has a better and clearer idea of

how the ML came to its conclusion. ML should also only be applied to problems where the value it adds clearly exceeds this risk, understanding that this may be easier said than done. Additionally, when data scientists and engineers are designing ML systems, clinicians should be incorporated in this process to build trust and familiarity with the final models. Finally, ML requires expert human oversight so that the outputs can be screened for absurd or dangerous suggestions, all the while keeping in mind that the ML suggestions, while sometimes not immediately obvious, may be deeper, more robust, and more profound than the human evaluator first believes. This may require an entirely new kind of clinician who is an expert at the domain issues of medicine as well as with the issues that ML brings to the table when applied to clinical issues. Like an airline pilot who often uses sometime imperfect automatic control modalities, future clinicians will need to recognize when and how they need to step in to the situation.

Currently, while the conception of ML monitoring applications is dynamic in terms of information updating the rules, their execution is static. The introduction of true, learning ML into the actual care process will also raise regulatory issues—are machines making decisions that should be made by doctors? This will inevitably raise medico-legal issues as potentially adverse outcomes result from machine-human hybrid decisions, or even purely from machine decisions. But it is important not to let such concerns override our potential and ability to innovate to improve clinical processes and outcomes with intelligent digital partners.

It is possible that processing information more intelligently will still not provide us with the ability to prevent or mitigate adverse clinical events, or to make smarter, faster diagnoses, but we won't know until we try. Only future research and rigorous evaluation of ML and other AI techniques will allow for safe incorporation of these tools into clinical practice. Our aim is to recruit more clinicians and machine learning experts to engage in this brainstorming to integrate ML into an ideal digital care system that leverages every bit of data to add value to patient monitoring and to assist clinicians in prioritizing tasks and ultimately benefiting patients [29]. Our sometimes painful experience with the digitization of medicine should also force us to be cognizant that the introduction of ML into clinical care requires better educational and clinical training, as well as the production of more user-friendly clinical tools, than has heretofore been the case for the design and introduction of EHRs [30].

Author contributions All authors contributed equally to the manuscript, and read and approved the final manuscript.

Funding Leo Anthony Celi is funded by the National Institutes of Health through National Institute of Biomedical Imaging and Bioengineering Grant R01 EB017205-01A1 and by Philips Healthcare.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

References

- Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375:1216–9.
- Obermeyer Z, Lee TH. Lost in thought—the limits of the human mind and the future of medicine. *N Engl J Med*. 2017;377:1209–11.
- Chen JH, Asch SM. Machine learning and prediction in medicine - beyond the peak of inflated expectations. *N Engl J Med*. 2017;376:2507–9.
- Wang S-L, Wu F, Wang B-H. Prediction of severe sepsis using SVM model. *Adv Exp Med Biol*. 2010;680:75–81.
- Desautels T, Calvert J, Hoffman J, Jay M, Kerem Y, Shieh L, et al. Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach. *JMIR Med Informatics*. 2016;4:e28.
- McCoy A, Das R. Reducing patient mortality, length of stay and readmissions through machine learning-based sepsis prediction in the emergency department, intensive care unit and hospital floor units. *BMJ Open Qual*. 2017;6:e000158.
- Mayaud L, Lai PS, Clifford GD, Tarassenko L, Celi LA, Annane D. Dynamic data during hypotensive episode improves mortality predictions among patients with sepsis and hypotension. *Crit Care Med*. 2013;41:954–62.
- Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Crit Care Med*. 2017;46:547–53.
- Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. *BMJ Open Respir Res*. 2017;4:e000234.
- Drew BJ, Harris P, Zègre-Hemsey JK, Mammone T, Schindler D, Salas-Boni R, et al. Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients. *PLoS ONE*. 2014;9:e110274.
- Borges G, Brusamarello V. Sensor fusion methods for reducing false alarms in heart rate monitoring. *J Clin Monit Comput*. 2016;30:859–67.
- Chen L, Dubrawski A, Wang D, Fiterau M, Guillaume-Bert M, Bose E, et al. Using supervised machine learning to classify real alerts and artifact in online multisignal vital sign monitoring data. *Crit Care Med*. 2016;44:e456–63.
- Hravnak M, Chen L, Dubrawski A, Bose E, Clermont G, Pinsky MR. Real alerts and artifact classification in archived multi-signal vital sign monitoring data: implications for mining big data. *J Clin Monit Comput*. 2016;30:875–88.
- Ansari S, Belle A, Ghanbari H, Salamango M, Najarian K. Suppression of false arrhythmia alarms in the ICU: a machine learning approach. *Physiol Meas*. 2016;37:1186–203.
- Maldonado JR. Acute Brain Failure: Pathophysiology, diagnosis, management, and sequelae of delirium. *Crit Care Clin*. 2017;33:461–519.
- Pandharipande PP, Ely EW, Arora RC, Balas MC, Boustani MA, La Calle GH, et al. The intensive care delirium research agenda: a multinational, interprofessional perspective. *Intensive Care Med*. 2017;43:1329–39.
- Oh J, Cho D, Park J, Na SH, Kim J, Heo J, et al. Prediction and early detection of delirium in the intensive care unit by using heart rate variability and machine learning. *Physiol Meas*. 2018;39:035004.
- 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;46:e151–7.
- Rush B, Stone DJ, Celi LA. From big data to artificial intelligence: harnessing data routinely collected in the process of care. *Crit Care Med*. 2018;46:345–6.
- Barbato M, Barclay G, Potter J, Yeo W, Chung J. Correlation between observational scales of sedation and comfort and bispectral index scores. *J Pain Symptom Manage*. 2017;54:186–93.
- Nagaraj SB, Biswal S, Boyle EJ, Zhou DW, McClain LM, Bajwa EK, et al. Patient-specific classification of ICU sedation levels from heart rate variability. *Crit Care Med*. 2017;45:e683–90.
- Nagaraj SB, McClain LM, Zhou DW, Biswal S, Rosenthal ES, Purdon PL, et al. Automatic classification of sedation levels in ICU patients using heart rate variability. *Crit Care Med*. 2016;44:e782–9.
- West N, McBeth PB, Brodie SM, van Heusden K, Sunderland S, Dumont GA, et al. Feasibility of continuous sedation monitoring in critically ill intensive care unit patients using the NeuroSENSE WAVCNSindex. *J Clin Monit Comput*. 2018;32:1081–91.
- Wang Z-H, Chen H, Yang Y-L, Shi Z-H, Guo Q-H, Li Y-W, et al. Bispectral index can reliably detect deep sedation in mechanically ventilated patients: a prospective multicenter validation study. *Anesth Analg*. 2017;125:176–83.
- Moss TJ, Calland JF, Enfield KB, Gomez-Manjarres DC, Ruminiski C, DiMarco JP, et al. New-onset atrial fibrillation in the critically ill. *Crit Care Med*. 2017;45:790–7.
- Komorowski M, Celi LA. Will artificial intelligence contribute to overuse in healthcare? *Crit Care Med*. 2017;45:912–3.
- Breslow MJ, Rosenfeld BA, Doerfler M, Burke G, Yates G, Stone DJ, et al. Effect of a multiple-site intensive care unit telemedicine program on clinical and economic outcomes: an alternative paradigm for intensivists staffing. *Crit Care Med*. 2004;32:31–8.
- Lilly CM, Zubrow MT, Kempner KM, Reynolds HN, Subramanian S, Eriksson EA, et al. Critical care telemedicine: evolution and state of the art. *Crit Care Med*. 2014;42:2429–36.
- Knight W. The dark secret at the heart of AI. *MIT Technol Rev*. 2017;120:54–63.
- Moskowitz A, McSparron J, Stone DJ, Celi LA. Preparing a new generation of clinicians for the era of big data. *Harvard Med student Rev*. 2015;2:24–7.