DUHS Inpatient General Decompensation Prediction: literature review

Ziyuan Shen [*], Mengxuan Cui [†]

Last Revision: June 12, 2019

Abstract

Patients in general wards suffer decompensation for complex reasons and it tends to be harder to detect than in Intensive Care Unit (ICU). Inpatient general decompensation prediction is of great importance in terms of optimizing response before deterioration. The project goal is to define and predict a general decompensation problem in the university hospital. By early detection of decompensation, adverse events such as calling Rapid Response Team (RRT) or code blue events can be hopefully avoided. Health record datasets will be explored and transformed into a useable format. Ultimately, machine learning models will be designed and evaluated to achieve the best possible performance. By this way, we expect to set a standard to alarm deterioration and to put forward a series of actionable responses.

1 Purpose of Study

Patients in hospital sometimes decompensate for complex reasons and the general ward is usually a harder setting than Intensive Care Unit (ICU) to detect deterioration [1]. This is because patients in ICU have higher level of care and monitoring. Also, compared with patients admitted directly to ICU [2], patients who suffer unanticipated or delayed transfers to ICU have higher mortality [3, 4]. Based on the current reactive rather than proactive caregiving process, Rapid Response Team (RRT) is usually called when the patient is sick enough and may not manage to rescue the patient in time. Therefore, for patient safety, an early alerting system is required to be equipped in general wards. Advanced machine learning methodologies are worth deploying to predict early clinical deterioration, thus to better manage resources, hopefully to provide high quality care and prevent deterioration. Our study aims to well define a general decompensation detection problem by following Duke Hospital's clinical workflow, and create a machine learning model that predicts general decompensation in real time. Such work will also help to reduce deterioration [5] and to standardize Rapid Response Team (RRT) protocols.

2 Background and Significance

From a statistical point of view, studies show that unexpected ICU transfers [3], delayed transfers [4], late ICU admission [6] increase patients' in-hospital mortality and length of stay if the patient survived to discharge. Delayed recognition of deterioration rise from monitoring failures [7] to a great extent. Nurses can fail to detect decompensation due to various reasons such as too much workload [8] and limitation of an individual's capability. In addition, constantly observable information, such as heart rhythms and oxygen saturation that is displayed by monitors, is only a subset of available data and is insufficient for decision making [9]. The caregiving process is too dependent on frontline staff and an automatic monitoring system remains to be developed. Machine learning techniques will be a helpful tool for reducing the rate of preventable adverse events.

[*]ziyuan.shen@duke.edu, Department of Electrical and Computer Engineering, Duke University
[†]mengxuan.cui@duke.edu, Department of Biomedical Engineering, Duke University

Before the emergence of machine learning techniques, risk scores (Sepsis-related Organ Failure Assessment (SOFA) [10], Early Warning Score (EWS) [11], Modified Early Warning Score (MEWS) [12], etc) had been widely used for identifying and predicting patient deterioration (cardiac arrest [13], ICU admission [11], mortality [14] etc). Identified risk factors include abnormal respiratory rate, abnormal breathing indicator, abnormal pulse etc [13]. A large amount of work focus on detecting a specific type of deterioration, including cardiac arrest [13, 15, 16], sepsis [17–23], acute kidney injury (AKI) [22, 24] and acute lung injury [25, 26]. Proposed machine learning methods to solve relevant problems include logistic regression [22, 24, 27, 28], generalized additive model (GAM) [22], naive Bayes [22, 24], support vector machine [22, 24, 29], decision trees (random forest included) [24, 28], gradient boost ensemble method [20], artificial neural network (ANN) [16], time-to-event modeling based on deep network [30]. Recently, researchers also make an effort to develop models that capitalize on streaming data to aid real-time decision-making, for instance, Gaussian process predictor combined with recurrent neural network (RNN) [19], RNN combined with gated recurrent units (GRU) [31] and hierarchical analysis [21].

In spite of voluminous literature attempting to predict patients' deterioration, a lot of problems remain unsolved. Unlike specific types of deterioration, definition of non-specific deterioration lacks a consensus or criteria to follow. Subjectivity and variations of what forms clinical decompensation can lead to the barrier of identifying deteriorating patients, which necessitates the clarification of the concept [32]. Basically, there are two ways to define clinical deterioration [33]. One way is retrospective, and focuses on the end events such as cardiac arrest, severe sepsis and in-hospital mortality. The other way is prospective, which defines deterioration before it actually happens. Related variables contain abnormal vital signs and patient's general condition evolves in a negative direction. Existing systems for detecting patient general condition include National Early Warning Score (NEWS) [34] and Rothman Index (RI) [35], but measures of risk vary and depend on the research team. Furthermore, current electronic detection systems do not always benefit clinical outcomes [36]. EWS is found to have poor predictive capabilities [1]. Most early warning scores are not individual-based, which means they cannot fit specific patients. Numerous and sophisticated alert-generating devices can result in alert fatigue, potentially worsening caregivers' performance [37]. Although NEWS is already in use in Duke Hospital, the performance is far from optimal. Therefore, research efforts should lay emphasis on model quality rather than quantity.

3 Data Analysis and Statistical Considerations

3.1 Aim 1: Data exploration and preprocessing

Cohort Identification:
Identifying cohort is important before any preprocessing and model design, as models working well in one specific population may not be generalized to others. As a normal case, we consider solely adults for general decompensation.

General Decompensation Definition:
One barrier of early decompensation detection is the variation of medical data sources. Definition of decompensation here is also of great significance. As the project goal is to prevent RRT and code blue events, these two events can be directly used as the label from a machine learning perspective. However, information regarding the calling of RRT is recorded elsewhere instead of along with other data. Preprocessing is crucial for extracting and combining all the useful data. Another alternative way is to look into what triggers RRT. Mostly used conditions include acute respiratory failure, acute cardiac failure, acute changes in consciousness, hypotension, arrhythmias, pulmonary edema, and sepsis [38]. Ref. [28] predicts the combined outcome of cardiac arrest, intensive care unit transfer, or death. Similarly, common measures of risk such as ICU transfer, in-hospital mortality and 30-readmission will probably be integrated in our model. By following the clinical workflow, who will be exposed to the decompensation alert should also be taken into consideration. Determining which clinical staff and how they will provide treatment will help clarify the information in need.

Feature Extraction:

Features to be incorporated should mainly come from Electrical Health Record (EHR) and vary from demographics, disease diagnosis, medication record to vital signs and laboratory results [1, 39], dependent on the decompensation definition. Predictors proven to be useful also compose of clinical history, physical examination, presenting symptoms, etc [40]. SQL will be used to integrate all kinds of information, generating features and labels in a usable format for statistical models and ultimately applications in real clinical scenarios.

3.2 Aim 2: Model creation and evaluation

Once we have well transformed the raw data into right and usable format, we will test our own designed models or models from existing literature. NEWS and traditional classifiers will also be implemented as benchmarks. A Python-based platform will be used. Model performance should be compared with existing methods as well as clinicians' performance to validate whether early alerts have the potential to influence clinical outcomes [25]. Due to the dynamic nature of patients' condition, time series data should be utilized and the real time prediction should be implemented. Another issue will come up when evaluating the model. As part of ICU care, clinicians always take action to treat deteriorating patients. Such intervention prevents adverse events that would occur without practitioners' surveillance. Neglecting this fact will underestimate the model performance, thus strategies are worth developing to mitigate this effect.

3.3 Design Pipeline

Pipeline for our planned model design is illustrated in Figure 1, consistent with Section 3.1 and 3.2. Additionally, model performance and label definition are desired to be optimized iteratively through a feedback loop.

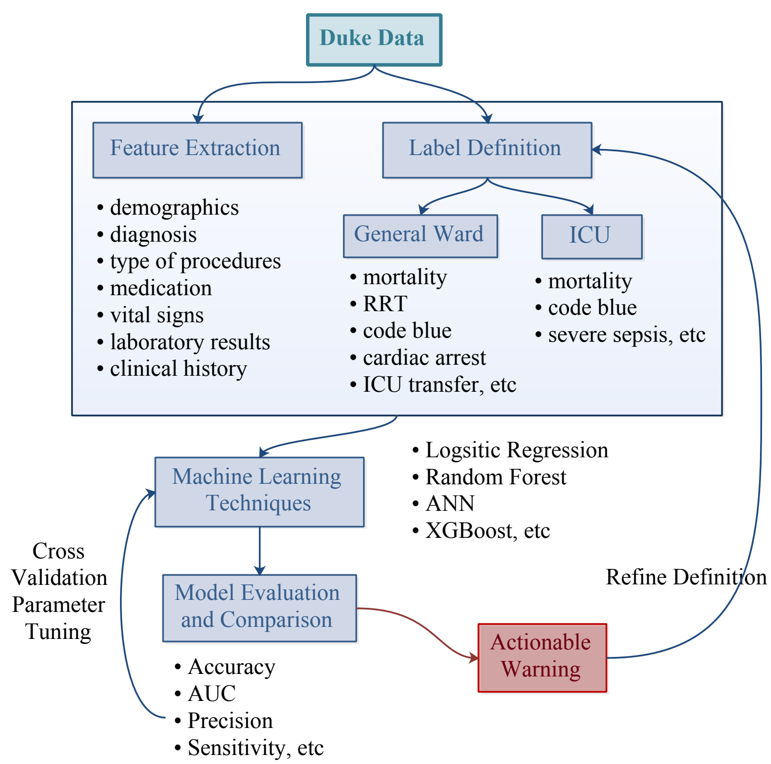

Figure 1: Design pipeline.

References

[1] C. Petit, R. Bezemer, and L. Atallah, “A review of recent advances in data analytics for post-operative patient deterioration detection,” *Journal of clinical monitoring and computing*, vol. 32, no. 3, pp. 391–402, 2018.

[2] G. J. Escobar, B. J. Turk, A. Ragins, J. Ha, B. Hoberman, S. M. LeVine, M. A. Ballesca, V. Liu, and P. Kipnis, “Piloting electronic medical record–based early detection of inpatient deterioration in community hospitals,” *Journal of hospital medicine*, vol. 11, pp. S18–S24, 2016.

[3] V. Liu, P. Kipnis, N. W. Rizk, and G. J. Escobar, “Adverse outcomes associated with delayed intensive care unit transfers in an integrated healthcare system,” *Journal of hospital medicine*, vol. 7, no. 3, pp. 224–230, 2012.

[4] M. M. Churpek, B. Wendlandt, F. J. Zadravecz, R. Adhikari, C. Winslow, and D. P. Edelson, “Association between intensive care unit transfer delay and hospital mortality: a multicenter investigation,” *Journal of hospital medicine*, vol. 11, no. 11, pp. 757–762, 2016.

[5] P. Kipnis, B. J. Turk, D. A. Wulf, J. C. LaGuardia, V. Liu, M. M. Churpek, S. Romero-Brufau, and G. J. Escobar, “Development and validation of an electronic medical record-based alert score for detection of inpatient deterioration outside the icu,” *Journal of biomedical informatics*, vol. 64, pp. 10–19, 2016.

[6] B. Renaud, C. Brun-Buisson, A. Santin, E. Coma, C. Noyez, M. J. Fine, D. M. Yealy, and J. Labarère, “Outcomes of early, late, and no admission to the intensive care unit for patients hospitalized with community-acquired pneumonia,” *Academic Emergency Medicine*, vol. 19, no. 3, pp. 294–303, 2012.

[7] L. S. van Galen, P. W. Struik, B. E. Driesen, H. Merten, J. Ludikhuize, J. I. van der Spoel, M. H. Kramer, and P. W. Nanayakkara, “Delayed recognition of deterioration of patients in general wards is mostly caused by human related monitoring failures: a root cause analysis of unplanned icu admissions,” *PloS one*, vol. 11, no. 8, p. e0161393, 2016.

[8] P. R. DeLucia, T. E. Ott, and P. A. Palmieri, “Performance in nursing,” *Reviews of human factors and ergonomics*, vol. 5, no. 1, pp. 1–40, 2009.

[9] M. McNett, M. Doheny, C. A. Sedlak, and R. Ludwick, “Judgments of critical care nurses about risk for secondary brain injury,” *American Journal of critical care*, vol. 19, no. 3, pp. 250–260, 2010.

[10] J.-L. Vincent, R. Moreno, J. Takala, S. Willatts, A. De Mendonça, H. Bruining, C. Reinhart, P. Suter, and L. Thijs, “The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure,” *Intensive care medicine*, vol. 22, no. 7, pp. 707–710, 1996.

[11] J. McGaughey, F. Alderdice, R. Fowler, A. Kapila, A. Mayhew, and M. Moutray, “Outreach and early warning systems (EWS) for the prevention of intensive care admission and death of critically ill adult patients on general hospital wards,” *Cochrane Database of Systematic Reviews*, no. 3, 2007.

[12] J. Gardner-Thorpe, N. Love, J. Wrightson, S. Walsh, and N. Keeling, “The value of modified early warning score (MEWS) in surgical in-patients: a prospective observational study,” *The Annals of The Royal College of Surgeons of England*, vol. 88, no. 6, pp. 571–575, 2006.

[13] T. J. Hodgetts, G. Kenward, I. G. Vlachonikolis, S. Payne, and N. Castle, “The identification of risk factors for cardiac arrest and formulation of activation criteria to alert a medical emergency team,” *Resuscitation*, vol. 54, no. 2, pp. 125–131, 2002.

[14] A. M. Mudge, C. Douglas, X. Sansome, M. Tresillian, S. Murray, S. Finnigan, and C. R. Blaber, “Risk of 12-month mortality among hospital inpatients using the surprise question and spict criteria: a prospective study,” *BMJ supportive & palliative care*, vol. 8, no. 2, pp. 213–220, 2018.

[15] M. M. Churpek, T. C. Yuen, S. Y. Park, D. O. Meltzer, J. B. Hall, and D. P. Edelson, “Derivation of a cardiac arrest prediction model using ward vital signs,” *Critical care medicine*, vol. 40, no. 7, p. 2102, 2012.

[16] D.-H. Jang, J. Kim, Y. H. Jo, J. H. Lee, J. E. Hwang, S. M. Park, D. K. Lee, I. Park, D. Kim, and H. Chang, "Developing neural network models for early detection of cardiac arrest in emergency department," *The American journal of emergency medicine*, 2019.

[17] V. Herasevich, M. S. Pieper, J. Pulido, and O. Gajic, "Enrollment into a time sensitive clinical study in the critical care setting: results from computerized septic shock sniffer implementation," *Journal of the American Medical Informatics Association*, vol. 18, no. 5, pp. 639–644, 2011.

[18] A. M. Harrison, C. Thongprayoon, R. Kashyap, C. G. Chute, O. Gajic, B. W. Pickering, and V. Herasevich, "Developing the surveillance algorithm for detection of failure to recognize and treat severe sepsis," in *Mayo Clinic Proceedings*, vol. 90, no. 2. Elsevier, 2015, pp. 166–175.

[19] J. Futoma, S. Hariharan, and K. Heller, "Learning to detect sepsis with a multitask gaussian process RNN classifier," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 2017, pp. 1174–1182.

[20] C. Barton, U. Chettipally, Y. Zhou, Z. Jiang, A. Lynn-Palevsky, S. Le, J. Calvert, and R. Das, "Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs," *Computers in biology and medicine*, 2019.

[21] F. van Wyk, A. Khojandi, and R. Kamaleswaran, "Improving prediction performance using hierarchical analysis of real-time data: A sepsis case study," *IEEE journal of biomedical and health informatics*, 2019.

[22] P. Thottakkara, T. Ozrazgat-Baslanti, B. B. Hupf, P. Rashidi, P. Pardalos, P. Momcilovic, and A. Bihorac, "Application of machine learning techniques to high-dimensional clinical data to forecast postoperative complications," *PloS one*, vol. 11, no. 5, p. e0155705, 2016.

[23] J. S. Calvert, D. A. Price, U. K. Chettipally, C. W. Barton, M. D. Feldman, J. L. Hoffman, M. Jay, and R. Das, "A computational approach to early sepsis detection," *Computers in biology and medicine*, vol. 74, pp. 69–73, 2016.

[24] R. J. Kate, R. M. Perez, D. Mazumdar, K. S. Pasupathy, and V. Nilakantan, "Prediction and detection models for acute kidney injury in hospitalized older adults," *BMC medical informatics and decision making*, vol. 16, no. 1, p. 39, 2016.

[25] V. Herasevich, M. Yilmaz, H. Khan, R. D. Hubmayr, and O. Gajic, "Validation of an electronic surveillance system for acute lung injury," *Intensive care medicine*, vol. 35, no. 6, pp. 1018–1023, 2009.

[26] N. W. Chbat, W. Chu, M. Ghosh, G. Li, M. Li, C. M. Chiofolo, S. Vairavan, V. Herasevich, and O. Gajic, "Clinical knowledge-based inference model for early detection of acute lung injury," *Annals of biomedical engineering*, vol. 40, no. 5, pp. 1131–1141, 2012.

[27] M. Gad Segal, M. Yoram Levo, and M. Rami Hershkovitz, ""general deterioration": a diagnosis that is a marker for risk of mortality upon re-admission," 2013.

[28] M. M. Churpek, T. C. Yuen, C. Winslow, D. O. Meltzer, M. W. Kattan, and D. P. Edelson, "Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards," *Critical care medicine*, vol. 44, no. 2, p. 368, 2016.

[29] L. Clifton, D. A. Clifton, P. J. Watkinson, and L. Tarassenko, "Identification of patient deterioration in vital-sign data using one-class support vector machines," in *2011 federated conference on computer science and information systems (FedCSIS)*. IEEE, 2011, pp. 125–131.

[30] P. Chapfuwa, C. Tao, C. Li, C. Page, B. Goldstein, L. Carin, and R. Henao, "Adversarial time-to-event modeling," *arXiv preprint arXiv:1804.03184*, 2018.

[31] B. Shickel, T. J. Loftus, L. Adhikari, T. Ozrazgat-Baslanti, A. Bihorac, and P. Rashidi, "Deepsofa: A continuous acuity score for critically ill patients using clinically interpretable deep learning," *Scientific reports*, vol. 9, no. 1, p. 1879, 2019.

[32] R. M. Padilla and A. M. Mayo, “Clinical deterioration: A concept analysis,” *Journal of clinical nursing*, vol. 27, no. 7-8, pp. 1360–1368, 2018.

[33] D. Jones, I. Mitchell, K. Hillman, and D. Story, “Defining clinical deterioration,” *Resuscitation*, vol. 84, no. 8, pp. 1029–1034, 2013.

[34] G. B. Smith, D. R. Prytherch, P. Meredith, P. E. Schmidt, and P. I. Featherstone, “The ability of the national early warning score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death,” *Resuscitation*, vol. 84, no. 4, pp. 465–470, 2013.

[35] M. J. Rothman, S. I. Rothman, and J. Beals IV, “Development and validation of a continuous measure of patient condition using the electronic medical record,” *Journal of biomedical informatics*, vol. 46, no. 5, pp. 837–848, 2013.

[36] M. H. Hooper, L. Weavind, A. P. Wheeler, J. B. Martin, S. S. Gowda, M. W. Semler, R. M. Hayes, D. W. Albert, N. B. Deane, H. Nian *et al.*, “Randomized trial of automated, electronic monitoring to facilitate early detection of sepsis in the intensive care unit,” *Critical care medicine*, vol. 40, no. 7, p. 2096, 2012.

[37] “Alert fatigue,” https://psnet.ahrq.gov/primers/primer/28/alert-fatigue, accessed: 2019-05-27.

[38] D. A. Jones, M. A. DeVita, and R. Bellomo, “Rapid-response teams,” *New England Journal of Medicine*, vol. 365, no. 2, pp. 139–146, 2011.

[39] L. A. Despins, “Automated deterioration detection using electronic medical record data in intensive care unit patients: a systematic review,” *CIN: Computers, Informatics, Nursing*, vol. 36, no. 7, pp. 323–330, 2018.

[40] E. E. Tripoliti, T. G. Papadopoulos, G. S. Karanasiou, K. K. Naka, and D. I. Fotiadis, “Heart failure: diagnosis, severity estimation and prediction of adverse events through machine learning techniques,” *Computational and structural biotechnology journal*, vol. 15, pp. 26–47, 2017.