Anticipatory intervention, rather than a reactionary response to the deterioration of critically ill children, is ideal. Even for experienced clinicians, however, it is challenging to differentiate who among high-risk patients are at highest probability of decompensation, especially among patients with single-ventricle physiology. Making this prediction with certainty and foresight that afford time for preventive intervention has not been achievable thus far. In this issue of the Journal, Ruiz and colleagues\(^1\) approach this dilemma by creating prediction models that are based on clinician knowledge and machine learning to attempt to anticipate critical events as early as 8 hours in advance. The study suggests that routinely collected data can be leveraged to build an early warning system, and that models centered on a combination of machine learning and expert knowledge are superior to models reliant exclusively on expert knowledge.

Although the findings of Ruiz and colleagues\(^1\) are of interest, various aspects of the study design and interpretation merit thoughtful reflection. Although a single-ventricle population was targeted, the study cohort may not be representative because of the number of patients housed outside a cardiac intensive care unit, the distribution of congenital heart disease lesions, and the timeline and frequency of adverse events. Specifically, the majority of the events occurred greater than 7 days after surgical palliation, which is incongruent with literature, suggesting that the highest risk of decompensation occurs during the immediate postoperative period.\(^2,3\) The event frequency also appears lower than expected. Furthermore, a critical event was defined solely through medical records; however, many procedures may have been elective rather than truly emergency in nature. For example, many extracorporeal membrane oxygenation cannulations were not preceded by either cardiopulmonary resuscitation or intubation,\(^4\) and only 5% of intubations were followed by further deterioration. In terms of variables, the data were not longitudinal, may have been outdated, and are not all typically collected at other institutions. Finally, the expert knowledge models relied on a single clinician, rather than a combination or variety of experts.

Because of these limitations, the study directs attention to the potential impact of early warning systems in clinical medicine. Predictive modeling will likely gain increasing clinical traction, and the critical issues of adopting this technology are the necessity, potential, model scheme, data set, implementation, and patient impact.

Predictive modeling should not be viewed as an attempt to replace the physician, but rather as an enhancement of clinical effectiveness achieved by affording providers the opportunity to focus their cognitive energy on higher-order reasoning. Physicians are often unaware, fatigued, overwhelmed, or inexperienced, and predictive models promote the mental transition from the data gathering and information phase to the knowledge and wisdom stages of clinical management.\(^5,6,7\)

Model scheme is essential to unlocking the full potential of predictive modeling and the ability to adapt to different populations, institutions, and clinical settings. Ruiz and colleagues\(^1\) begin the transition from physician-based decision making to models reliant on expert knowledge to guide machine learning. A structured approach is intuitive, but it is a form of brute force logic that has limitations. Earnest consideration should be given to unstructured,
deep learning. The unstructured approach affords experiential learning, self-directed algorithm development, and, ultimately, a more powerful predictive system. In a 1997 chess match, supercomputer Deep Blue beat Gary Kasparov, arguably the premier grandmaster of all time, by means of structured algorithms. Today, after learning chess by simply receiving the rules of the game and driven by deep learning, AlphaZero did not lose a single match against the preeminent brute force computer.9 AlphaZero taught itself by playing millions of games against itself and now is regarded as unbeatable, often playing with intuition and decision making not previously seen in artificial intelligence.9,10 The same opportunity lies ahead for predictive modeling in medicine. An unstructured, deep learning scheme is unencumbered by traditional human logic giving it the potential to unlock relationships and associations previously unknown to physicians.11,12 Furthermore, expert-driven models require institution-specific tailoring, which is labor and cost intensive. Deep learning models can be trained on a variety of variables, thereby increasing the ease of adaptability.

Similar to model scheme, the data set used to generate these predictions is equally important. Although routinely collected data are convenient, their use is fraught with signal noise, is prone to errors in data entry and collection, and may be inadequate to generate a robust early prediction system. An industry similarly entrenched in tradition and learning gleaned from years of apprenticeship underwent a revolution in 2003. With the publication of Moneyball,13 athletics were thrust into the golden age of analytics.14 Previously, baseball was dependent on the traditional box score, which captured a limited number of statistics; offensive metrics were scant, while defensive value was neglected entirely. The ability to determine player effectiveness was consequently limited. The box score was outdated; a new data set was needed. New databases came in the form of multivariate video systems that captured player, ball, and referee movements with 0.04-second granularity.15 These new data allowed predictive algorithms to be developed to determine a player’s true value, whether on offense or defense. Early adoption of these new data sets equated to an increased winning percentage, and often to championships. Basketball, football, and other sports experienced a similar evolution.16 In pediatric cardiology, we are potentially facing a similar conundrum. The Pediatric Cardiac Critical Care Consortium has made great strides in ensuring data integrity and limiting signal noise to a degree; however, Pediatric Cardiac Critical Care Consortium and other current databases still rely on traditionally collected data. New metrics are necessary. These may come in the form of video monitoring, emerging technology such as wearable devices, or novel technology still to be developed. In athletics, traditionalists claimed that databases could not be generated and that even if they were, the hurdles to analyzing large volumes of data were insurmountable. They have unequivocally been proved wrong. In medicine, the same advances should be anticipated.

The integration of early warning systems into clinical care deserves attention to maximize clinical utility and patient impact. This study’s predictive outputs were area under the curve, sensitivity, and specificity, which do not translate easily to the bedside. Any system, such as light versus numeric scales, has intrinsic advantages and limitations, which should be scrutinized. Furthermore, it is insufficient for a model to predict only risk of decomposition in advance. Identification of likely etiologies of decomposition is essential. A model that only informs about a patient’s risk profile is limited unless it provides a means of physician intervention. This is the next phase of unstructured, deep learning, and it is up to the medical field to not to throw away its shot.

References