Predictive Adverse Event Detection and Proactive Clinical Intervention: Real-Time AI Architectures for Patient Safety Enhancement in Acute Care

Dr. Beatrice Kern, Professor of Information Systems, University of Applied Sciences Potsdam, Germany

1. Introduction

Although there has been a rapid increase in morbidity and mortality rates due to adverse events in healthcare settings, it has been difficult to properly monitor them because the number of sensors installed to cover various areas has increased, and monitoring has become more complicated. In this regard, there have been attempts to monitor healthcare settings using machine learning, which is a data-driven approach. When commercial companies monitor healthcare settings to prevent adverse events, all the systems they introduce must undergo performance tests at the time of installation. However, as hospitals change gradually, it is necessary to periodically validate and redevelop the existing systems.

Given new advancements in the fields of artificial intelligence, machine learning, and deep learning, healthcare has become one of the main beneficiaries of these techniques. In this regard, we explore the tools for real-time monitoring and provide support to prevent adverse events. In particular, we explore AI-powered tools that have the potential to enhance patient safety. One such technological solution is machine-learning imperceptible physiologic assessment and related techniques, which aims to obtain valuable patient physiologic data. We argue that AI has the potential to support data integration or the personalization of patient safety. This paper outlines the basic design and methodology of data collection and handling, describes the technical specification of the underlying algorithms, briefly presents their feasibility, and discusses data potential, implementation costs, and development possibilities. The paper is organized as follows. In the next section, we describe adverse events in healthcare and the scope of this research. In Section 3, we present the data employed in our work. We discuss the technical potential for real-time AI in personalized safety in Section 4. In Sections 5 and 6, we elaborate on technical specifications and feasibility. Costs, data, and implementation benefits are presented in Section 7. Finally, the paper ends with a discussion in Section 8.

1.1. Background and Significance

Patient safety has been a key concern in healthcare for the better part of the past century. In 1959, the Food and Drug Administration reported that 4,500 patients had been injured as a result of medical devices, pushing safety concerns to the forefront of public opinion. Not much has changed since then: between 1999 and 2013, the adverse event rate remained stable, despite significant advancements in technology and the advent of patient safety research. This is particularly troubling given the changing face of healthcare delivery: increasingly complex environments involving tightly coupled sociotechnical systems have been the backdrop to an increasing rate of adverse events and near misses. Adverse events are particularly prevalent in the operating room, where the probability of an adverse event occurring rises from 22% on dual-task days to 29% on control days.

The prevalence of adverse events poses significant implications for patient outcomes, accounting for an estimated 36,000 deaths and 1 million injuries each year in the United States alone. Additionally, adverse events cost $6,900 in additional healthcare costs per patient and an aggregate impact of $16.7 billion on medical negligence. Given the prevalence of adverse events and the associated high costs, healthcare practitioners are increasingly prioritizing technological solutions to enhance patient care delivery, as well as patient prognosis. The emphasis on patient safety within healthcare environments complements the interdisciplinary team approach adopted by healthcare practitioners. Patient safety, with a focus on systemic accidents that lead to harm, calls for relevant individuals, work systems and creators of these systems to be intertwined, each feeding into and dependent on the other. Each has the ability to make or break safety. It is the responsibility of the healthcare professional to identify and mitigate the risks. Unfortunately, healthcare is a high-pressure environment characterized by high stakes, uncertainty, complexity, constant time pressure, quantity of tasks, frequency of change and inadequate or changing information. Many practitioners are no longer able to be reflexive and adopt everyday normal practices that enhance patient safety.

Patient safety is an ongoing investigation: it is an emergent phenomenon, vulnerable to constant change from numerous interacting systems. Public health professionals are specifically calling for an end to the fragmented, siloed approach of patient safety studies, given the high connectivity and interdependency of healthcare systems that creates an increased potential for errors. AI-based surveillance is an emergent research field that has shown great potential in unveiling new and hitherto unheralded patient safety issues within clinical environments. Previous research in this space has shown similar concern for invisible risks in surgery. We contend that there are factors not accounted for in previous research and that only a narrow section of the people, task and system interfaces have been included that are fraught with impact on patient care delivery in an operating room environment.

1.2. Purpose of the Study

The main purpose of this study is to evaluate and review in what ways machine learning can be used to monitor and prevent adverse events in order to improve patient safety. The two most important questions we aim to address using our review are: 1) What machine learning strategies could be useful in patient safety and how well do they work? 2) What is the desired input data and output of these machine learning strategies according to healthcare professionals? The outcomes of this study are expected to be beneficial for healthcare professionals in supporting their targeted, 'in practice' applications for the future, particularly for systems aiming to integrate with a healthcare setting. The use of machine learning in real-time monitoring can demonstrate how the results of an algorithm using data can be used by healthcare professionals to facilitate a faster decision-making process. Adverse events (AEs) represent an important public health issue, resulting in a significant overload of patient care services and increasing the costs of healthcare systems. In the last decade, different AI modalities have been applied for monitoring AEs using an automatic algorithm for real-time surveillance. The purpose of this study is to review and rate the machine learning techniques for real-time monitoring and AE prevention, investigating what statistical and machine learning strategies are used and providing an evaluation of which are most successful. We also aim to present a synthesis of the needs of data scientists, healthcare professionals, patients, caregivers, and researchers for the system to function effectively in terms of the input and output requirements of these machine learning strategies and encompass this within a systematic review. The output of this systematic review will provide directions and policy, including insights for healthcare professionals in bridging the gap between implementation and healthcare provisions.

2. Current Challenges in Patient Safety

Patient safety continues to challenge existing efforts and ways of monitoring for patient harm effectively. The following existing systems are currently used in patient safety: costly manual chart validation, sentinel events, and reporting based on regulatory criteria, along with other sample surveillance methods to easily identify flaggable criteria for patient harm. This task is inefficient, expensive, haphazard, and limited in scope. It can lead to underestimation by rarely reporting certain adverse events. Research shows that most incidents have no adverse health effects, and some incidents with negative health effects are not linked to adverse event experiences. Trigger criteria, which have shown positive detection of patient safety signals in billing and medical records, include diagnoses, procedures, laboratory results, medications, blood product data, and radiology results. Trigger criteria have also proven to be valuable in identifying specific types of adverse events or patient harm but are less practical in identifying new patient safety monitors on a larger scale. Therefore, it is important to handle patient safety monitors or triggers. Triggers using administrative or billing data have also been developed, but they are often used after the secrecy has been fulfilled and are still less sensitive or specific. Guidelines with similar problems have not been established.

2.1. Types of Adverse Events

The definition of an adverse event is an injury that is caused by medical management rather than by the underlying disease or condition of the patient. It is both unintended and further sustains the relationship between the injury and services provided by healthcare professionals. Adverse events that occur in a hospital setting can span a wide range of injuries, including both an act of commission (an unplanned act or procedure that leads to an adverse event) and an act of omission (not carrying out a planned act or procedure that leads to an adverse event), or errors related to medications, nosocomial infections, or even falls.

Essentially, among all types of adverse events, critical incidents account for the overwhelming majority of avoidable patient morbidity and mortality, while an adverse event caused by negligence leads to a tip of the iceberg in terms of patient harm caused by healthcare management, often with an underlying cultural connection to safety. The occurrence of adverse events can be divided into three categories: injuries due to medical management represent the subset; injuries are preventable and need to be removed from the healthcare system in which they occur, whereas injuries can lead to various types of adverse events. Therefore, chronic diseases and trauma patient care are excludable from the event. The intentional perverse patient outcomes arise from a healthcare provider who seeks to maim or injure a patient. Such injuries could derive from activities, mistakes, adverse actions, or lack of errors.

2.2. Limitations of Traditional Monitoring Systems

The principal proposed safeguards are reporting systems, state licensing of hospitals and health professionals, and organizations for standard setting, certification, and buyers and sellers. Discussions of all of these protections against errors and negligence typically look to external regulation and monitoring. Reliance on state licensing and hospital accrediting organizations assumes that external regulators have sufficient expertise to ensure that health professionals are not incompetent, fraudulent, or corrupt, and that organizational structure and process requirements are consistent with health industry standards for safety and performance. While this assumption may hold for many hospital-related construction features, it is no panacea for identifying and addressing all latent safety hazards of clinical and health care processes.

The basic problem is that adverse outcomes often are not due to incompetence, fraud, or willful misconduct, and that those involved in the care delivery process may not know that an error has had an adverse effect on a patient. Furthermore, the impact of many health care decisions, particularly those of a diagnostic and treatment nature, is made in real-time in complex and dynamic environments where sufficient reliable information to ascertain the appropriateness and correctness of decisions is unavailable. Traditional, expert-based, forward-looking monitoring and control technologies identify only cumulative or subsumed events; they cannot take into account contextual factors that may lead to an adverse event. They are often costly and time-consuming and are often not able to take into full consideration information asymmetry in the patient-provider relationship. Post Analysis Act Review also is not a panacea, for discovering adverse events and their causes is difficult.

3. Role of AI and Machine Learning in Patient Safety

The use of AI and machine learning has been fundamental in revolutionizing patient safety initiatives, especially in driving and facilitating innovations that help prevent the incidence of adverse events and errors. Both AI and machine learning have their applications in clinical as well as non-clinical or hospital settings. Machine learning algorithms aim to generate models with the ability to improve automatically through experience. In healthcare, a machine learning model's experience comes as it continuously learns from new data and practices. There is no need to program the software or application specifically so that it can carry out or undergo a certain situation; it learns as the volume of information and data increases. Regarding patient safety, predictive analytics and analyzing risks or potential implications of the clinical setting are on the rise, and new software tools are under development. This is an area strongly supported and driven by AI and machine learning tools and software.

Given the vast quantity of data from EHRs now available, machine learning's key role is learning from the data what can improve patient immediate outcomes. In numerous instances, algorithmic output may lead to real-time, immediate actions, such as an alert for severe kidney failure or high temperature. From another perspective, incorporating AI insight in multi-agent systems might also improve intraoperative patient safety and patient outcomes. In addition, case managers and others could monitor the multi-agent systems' evolution in the population with the possibility of reacting or intervening further. Thus, because of the potential for improved adaptive solutions that AI might offer in patient care, the application and pursuits of AI in the patient setting of the future will soon develop from this point. A noticeably higher emphasis on proactive patient and staff safety features and solutions is robustly supported.

3.1. Overview of AI and Machine Learning Technologies

Current machine learning (ML) and AI-based tools contain multiple algorithms that can analyze and learn from data to manipulate it using different interconnected layers called neural networks. There are multiple types of artificial neural network (ANN) tools and methods scattered throughout the world for all industries to leverage, including healthcare. ML algorithms such as graphs, associative rules, neural networks, Bayesian networks, support vector machines, decision trees, and clustering are used to process data such as clustering, classification, regression, forecasting, anomaly detection, concept description, summarization, and visualization. There are different types of ML techniques that can be applied to healthcare data, such as supervised, unsupervised, semi-supervised, and reinforcement learning.

Reinforcement learning is used in a dynamic environment to make sequences of decisions. Generative adversarial networks are another emerging type of ML technology developed for unsupervised learning that can capture complex data distributions in a short period. Other subfields of ML, where these technologies can be combined and used to prevent harm to patients and staff in real-time, will be discussed in the section that follows. AI tools evolved gradually from expert systems, neural networks, natural language processing, and cognitive analytics to deep learning; their capacity in analyzing and re-engineering healthcare data has improved massively over the decades. The main driving force behind the takeoff in AI tools and solutions in healthcare was their ability to pick up vital signals and identify early warning signs for potential fatal clinical deterioration that can be cataloged and rationalized with great accuracy in emergency care, in acute medical and surgical wards, in critical care units, in operating theaters, before, during, and after treatment, in admission rooms, outpatient clinics, and in community care situations.

3.2. Benefits of Real-Time Monitoring

One of the main advantages of real-time monitoring and interventions relates to the preventability of adverse events. Many adverse events are caused by dry runs or people not taking actions that can potentially prevent harm. By continuously monitoring and analyzing data, the likelihood of these events occurring can be reduced. It also means that if a patient's condition deteriorates, timely alerts and notifications occur that give health care professionals the time to intervene to prevent further harm. Safe and effective care is also about processes that offer a timely response to any patient care issues that might arise, and real-time interventions assist with this. Also, through real-time monitoring, information about patient care can be promptly communicated within a healthcare team, so people are kept in the loop to take action and can also follow up with the patient. With patients and relatives also part of this communication loop, the patient can also be kept informed about their own care. Real-time monitoring also empowers patients to take action themselves with the trust that their impact will be felt at the healthcare personnel level. A few examples of this include the prevention of patient deterioration and cardiac arrest, on the back of real-time illness severity monitoring and high to normal patient throughput in the emergency department, to the implementation of early warning systems with news-like scores for early sepsis treatment. More recently, we see the increasing use of AI to predict deteriorations, such as in patients with COVID-19. All of these closely monitored opportunities for real-time intervention testify to the relationship between improved patient outcomes and timely health care intervention. In the end, due to their stable and consistent outcomes, these interventions result in preventing hospital adverse events and are thus, by extension, cost-effective.

4. Implementation and Integration in Healthcare Settings

the training and support necessary for staff to adapt to the new technology effectively. Continuous feedback loops and iterative improvements are essential to refine the integration process and ensure that the AI tools are meeting the needs of the users. Additionally, monitoring and evaluation mechanisms should be established to assess the impact of AI on clinical outcomes and workflow efficiency. Engaging stakeholders throughout the process, including patients, can provide valuable insights and foster a sense of ownership over the new technology. Ultimately, successful AI integration in healthcare hinges on a comprehensive understanding of the unique challenges and opportunities within the clinical environment, as well as a commitment to ongoing collaboration and adaptation.

4.1. Technical Considerations

Real-time AI in patient safety and monitoring for early adverse event detection requires robust and scalable data infrastructure with low-latency processing. The infrastructure should be able to manage large volumes of data generated in the real world while also being able to process this data and perform inference in real time. Software must interoperate with many information systems used in hospitals. These systems are highly diversified and include several health information exchanges, electronic health records, medication delivery systems, an array of wearable and implantable devices, digital imaging and communications in medicine, and picture archiving and communication systems. As a result of this and America's ever-growing network of independent hospitals, the algorithm must be flexible, work as is or with slight modification across different hospital systems with the need for minimal adaptation.

An important aspect of the software design is to ensure that data used for decision-making is accurate and up to date by implementing a reliable system. Other challenges include fast software development cycles, periodic cybersecurity assessments, vulnerability and compliance scans, continuous monitoring, and software maintenance. User-friendliness is of great importance, as the software is primarily a clinical decision-making aid. Early user experience interviews have indicated strong pushback against using the program if it requires significant additional manual data entry or system interfacing to operate cleanly in the hospital environment. Hospital systems also may not be up to date with the newest information protocol, which creates its own considerations when designing the input to the system. A comprehensive validation process needs to be set in place for new algorithm versions to test them under real-world conditions as close to their final deployment as possible to mitigate as much real-world risk as feasible. Additionally, the algorithms in use should be validated for fairness in classification and analyze the potential for algorithmic bias, which presents an additional consideration to a monitoring solution because it needs to assess whether the monitoring signal in use may be biased against a particular race, gender, etc. in the patient population seen in a hospital with likely confounding factors. Finally, secondary reviews for all adverse event predictions will be in place in order to safeguard against rapid, unanticipated drift in the AI's performance.

4.2. Ethical and Legal Implications

Currently, very few considerations are made in terms of ethical and legal matters associated with the integration of advanced monitoring systems relying on AI in healthcare. The overriding policy recommendation that arises from the discussion is that the consent of patients regarding the use of specific data and AI-driven patient monitoring and support solutions should be taken seriously. Patients or their representatives may therefore also need to be given the ability to consent (or refuse consent) to the use of AI-driven decision support systems on their behalf by default.

Given that AI-assisted technology within the healthcare context has developed only in recent years, several ethical issues have not received much attention in this literature. Key among these considerations is the potential for biased AI systems when applied to healthcare settings. In addition to biased algorithms, it is not always clear why an AI system makes a particular decision, which poses a problem for those who must defend their decision. Reliability of AI algorithms is also of concern in safety-critical settings. Important ethical questions arise about the roles and responsibilities of healthcare professionals who use AI tools as support. Ethical discussions on AI technology can provide valuable insights into the potential pitfalls of patient safety technologies as they are developed in the future. Law and theory must strike a balance between the flow of innovation and consider ethical considerations from day one. As there are no best practice guidelines for any of this in the AI domain, these are the forms of instruments we propose be developed moving forward.

Decisions made by an AI system must comply with the discretionary systems established by regulations such as the relevant regulations, the regulations for medical devices and intelligent systems, and the ethics and safety standards published by the government. Numerous conversations we had with these stakeholders discussed the hype and the need to ensure that it is done efficiently and effectively, as good use cases are necessary and the wealth of potential projects to research is endless. The risk, according to one commentator, is that society grows used to AI applications delivered with the "design it and see" approach and we miss all those people who refused to be beta testers and thus become users of last resort. Even if AI was affordable and effective, there needs to be a sound foundation and pathway for its uptake. Because it both results from and potentially results in outputs that impact big ideas and major technical trends, it is important to be informed about AI strategic directions and to seek input from these stakeholders on AI strategic directives. A challenge is to ensure that dialogue is ongoing and does not become static at a point in time when the future impact of such technologies is changing. Addressing the agility of the dialogue and ongoing engagement is as important as issuing a document. In many areas such as government and industrial consultation regarding AI, ethically aligned design, trust, data governance including privacy and security, and the responsible use of AI is particularly important.

5. Case Studies and Success Stories

Another track is to present case studies and success stories that demonstrate the success of AI in enhancing patient safety. Covering several case studies with different methods and results can provide valuable evidence for the practicability of such tools in patient safety. Some of the content of case studies should include the following: (1) Case study on using deep reinforcement learning and electronic health records for automatic adverse event prediction. The result of Study 1 is that the area under the curve (AUC) of the method on a real-world dataset is 0.92, in both in-sample and out-of-sample validation. (2) Case study on real-world adverse events: implementation of a clinically operational predictive model for hypoglycemia. This should report a study with good performance.

The results of the study may include AUC-ROC or AUC-PRC, including confidence intervals, sensitivity and specificity, or an F1 Score. Other practical examples where AI, predictive modeling, and machine learning tools have demonstrated success are also of interest. (3) A case study on using machine learning to monitor and prevent adverse events. This case study will discuss anonymizing the data and using machine learning, as well as data integration, scalability, and the success of AI. Collaboration uses real-world data for sepsis and vital signs, showing how the model can assess changes over time and the gradient of change. The presentation of lessons learned and barriers to further use of this model is pertinent. This workshop aims to bring together a variety of views, discuss recent advances and work in progress, and share experiences in developing and deploying AI and machine learning tools in healthcare. The event will focus on demonstrable success stories and scientific case studies for public scrutiny. It aims to foster open discussion both on the performance of the tools and the practical implementation in the hospital setting, and build consensus on the critical aspects of creating AI medical products for improved patient safety.

6. Future Direction

While the focus of this study is on preventing physical adverse events, the future of AI for patient safety will also lead to the prevention of medical errors related to diagnosis and incorrect treatment, as well as the emergence of a new approach to immunization safety monitoring for preventive purposes. Such monitoring improvements can be attributed to advances in natural language processing, pattern recognition, and computer vision. In addition to being able to function as a triage system, predictive analytics will assist care staff in generating hospital-based personalized treatment strategies for individual patients. This also involves embedding machine learning in technological solutions using big data uploaded from the patient's wearable health technology devices.

By introducing algorithms for precise adverse event detection and individual trajectory prediction, the current work will contribute to the drafting of initial guidelines for testing and implementing AI algorithms in healthcare settings. Real-time processing will detect problems and future predictions, and it will also support adaptive planning to provide for preventive measures formulated in existing clinical guidelines. In general, AI research can greatly benefit not only from in-depth interdisciplinary collaboration to understand and capture complementary knowledge, but also from methods of combining and shaping innovation between health professionals and technologists. It is pivotal and timely for patient safety in healthcare, as super digital intelligent systems evolve to be safe and reliable in real time, learn new data, and tailor their function to the changing demands. Frameworks should adapt to such systems and regularly update, manage, and analyze their risks using real data. It is essential to proactively engage closely with health professionals and patients. Their perspectives will inform how beneficial care and treatments are and how we might continue to reap such benefits.

7. Conclusion

In this policy piece, we argue that a significant potential for improving patient safety lies in the integration of machine learning technologies into the continuous monitoring of patient records for adverse events. By today, various real-time AI tools are at the development stage, each designed for the real-time identification of different types of adverse events that can cause substantial morbidity and mortality. These tools alert the clinician when a highly adverse event is likely to occur, thereby giving the clinician the possibility to intervene and prevent the event from happening in the first place. We briefly summarize the current state of research on these preliminary-stage real-time AI tools. We discuss some of the challenges in implementing AI-powered real-time monitoring tools for patient safety, including the current lack of educational material for healthcare professionals on how to develop and validate such models, and the absence of a clear ethical framework and legal regulations for integrating them into practice.

In conclusion, there is a need for further cross-disciplinary research on real-time AI monitoring of patient safety. We must conduct studies of these tools in a clinical setting to ensure that they effectively improve patient outcomes and do not create any new problems. The deployment of any such methods must be guided by clear ethical guidelines. We advocate a paradigm shift in which we move from predominantly learning from errors post hoc to enabling real-time systems to proactively support patient safety. A crucial part of this involves equipping healthcare professionals with the necessary skills to monitor patients in real time.