JAMIA Open, 2023, **6(4)**, ooad084
https://doi.org/10.1093/jamiaopen/ooad084
Research and Applications

Research and Applications

Medical Informatics Operating Room Vitals and Events Repository (MOVER): a public-access operating room database

Muntaha Samad, PhD[1,2], Mirana Angel[1,2], Joseph Rinehart, MD[3], Yuzo Kanomata[1,2], Pierre Baldi, PhD[1,2,*], Maxime Cannesson, MD, PhD[4,*]

[1]Department of Computer Science, University of California, Irvine, Irvine, CA 92697, United States, [2]Institute for Genomics and Bioinformatics, University of California, Irvine, Irvine, CA 92697, United States, [3]Department of Anesthesiology & Perioperative Care, University of California, Irvine, Irvine, CA 92697, United States, [4]Department of Anesthesiology & Perioperative Medicine, University of California, Los Angeles, Los Angeles, CA 90095, United States

*Corresponding author: Pierre Baldi, PhD, Department of Computer Science, University of California, Irvine, Irvine, CA 92697 (pfbaldi@ics.uci.edu); Maxime Cannesson, MD PhD, Department of Anesthesiology & Perioperative Medicine, UCLA David Geffen School of Medicine, 757 Westwood Plaza, Suite 2331-L, Los Angeles, CA 90095 (mcannesson@mednet.ucla.edu)

Author contributions: Muntaha Samad and Mirana Angel contributed equally and are considered co-first authors of this work.

Abstract

Objectives: Artificial intelligence (AI) holds great promise for transforming the healthcare industry. However, despite its potential, AI is yet to see widespread deployment in clinical settings in significant part due to the lack of publicly available clinical data and the lack of transparency in the published AI algorithms. There are few clinical data repositories publicly accessible to researchers to train and test AI algorithms, and even fewer that contain specialized data from the perioperative setting. To address this gap, we present and release the Medical Informatics Operating Room Vitals and Events Repository (MOVER).

Materials and Methods: This first release of MOVER includes adult patients who underwent surgery at the University of California, Irvine Medical Center from 2015 to 2022. Data for patients who underwent surgery were captured from 2 different sources: High-fidelity physiological waveforms from all of the operating rooms were captured in real time and matched with electronic medical record data.

Results: MOVER includes data from 58 799 unique patients and 83 468 surgeries. MOVER is available for download at https://doi.org/10.24432/C5VS5G, it can be downloaded by anyone who signs a data usage agreement (DUA), to restrict traffic to legitimate researchers.

Discussion: To the best of our knowledge MOVER is the only freely available public data repository that contains electronic health record and high-fidelity physiological waveforms data for patients undergoing surgery.

Conclusion: MOVER is freely available to all researchers who sign a DUA, and we hope that it will accelerate the integration of AI into healthcare settings, ultimately leading to improved patient outcomes.

Lay Summary

Despite many publications showing artificial intelligence algorithms to be successful in retrospective healthcare studies, there is a very limited amount of freely and publicly available medical data for researchers to work with, to develop and benchmark predictive and other methods in a reproducible manner. This is even more significant in the perioperative setting for patients undergoing surgery and anesthesia. In this article, we present and release a new repository we have constructed called MOVER: Medical Informatics Operating Room Vitals and Events Repository. This repository contains data (electronic medical record data and high-fidelity physiological waveforms data obtained from the bedside physiological monitors) associated with hospital visits for patients undergoing surgery and anesthesia. MOVER is freely available for download for all researchers who sign a data usage agreement: https://doi.org/10.24432/C5VS5G. MOVER is intended to advance a wide variety of healthcare research and serve as a resource to evaluate new clinical decision support and monitoring algorithms.

Key words: anesthesiology; surgery; electronic medical record; physiology; artificial intelligence.

Background and significance

In 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act was enacted to promote the adoption of healthcare information technology in hospitals. This act includes incentives for using electronic health record (EHR) systems.[1] The passage of the HITECH Act has resulted in widespread hospital EHR adoption with 80.5% of hospitals in the United States using an EHR system, as of 2015.[2] The increased adoption of EHR systems and subsequent rise in digitally available healthcare data has resulted in a newfound ability to perform predictive modeling on healthcare data using artificial intelligence (AI), primarily in the form of machine learning. Applications of AI in a healthcare setting include providing more accurate diagnoses, recommending treatment plans, predicting patient outcomes, tracking patient engagement and adherence, reducing the burden

Received: June 8, 2023; **Revised:** August 18, 2023; **Editorial Decision:** August 28, 2023; **Accepted:** October 6, 2023
Published by Oxford University Press on behalf of the American Medical Informatics Association 2023.

of administrative tasks, among others.[3–16] Despite many publications showing AI algorithms to be very successful in retrospective healthcare studies, there is a very limited amount of freely and publicly available medical data for researchers to work with, to develop and benchmark predictive and other methods in a reproducible manner.[8]

In addition, it is important to note that the benefits of open-source data extend beyond AI alone. Various research methods, including statistical analyses, epidemiological investigations, and other data-driven approaches, can greatly benefit from access to freely and publicly available medical data.

Therefore, we present and release a new repository we have constructed over the years called MOVER. This repository contains data associated with hospital visits for patients undergoing surgery at the University of California, Irvine (UCI) Medical Center. The data included in MOVER was collected over 7 years and contains comprehensive EHR and high-fidelity physiological waveforms for patients who underwent surgery at UCI.[8] These records include general information about each patient and their medical history, and specific information regarding the surgical procedure being performed including medicines used, lines or drains used, and postoperative complications. The repository includes 58 799 unique patients with data from 83 468 surgeries. MOVER is freely available for download for all researchers who sign a data usage agreement (DUA) and is intended to advance a wide variety of healthcare research and serve as a resource to evaluate new clinical decision support and monitoring algorithms.

Materials and methods

This study was approved by the Institutional Review Boards at the UCI Medical Center and on the main UCI campus. Requirement for individual patient consent was waived because we removed or deidentified all protected health information (PHI) in a Health Insurance Portability and Accountability Act (HIPAA) compliant manner.

Patient population

This first release of MOVER includes adult patients who underwent surgery at the UCI Medical Center from 2015 to 2022. The UCI Medical Center is the only level I trauma center in Orange County, California, a burn treatment center, and a National Cancer Institute-designated comprehensive cancer center. In addition, the UCI Douglas Hospital has some of the most technologically advanced surgical suites including state-of-the-art endovascular hybrid suites and intraoperative computed tomography and magnetic resonance imaging suites.

Data acquisition and electronic health records

The data acquisition process did not interfere with the clinical care of patients or methods of monitoring. Data for patients who underwent surgery were captured from 2 different sources. First, high-fidelity waveforms (EKG, pulse oximetry, and arterial line, if present) from all of the operating rooms (ORs) were captured in real-time using Bernoulli Health's hardware and software platform (Bernoulli, Cardiopulmonary Corp., New Haven, CT, USA). All of the waveforms were saved to a server on the medical center's network organized by source location (OR) and date/time. Subsequently, the medical center's informatics team delivered a data extract from the hospital EHR system from 2015 to 2022. For the years 2015-2017, the EHR system used was the Surgical Information Systems (SIS, Alpharetta, GASIS) and from 2017 to 2022 the EPIC EHR system was used (EPIC, Verona, WI, USA). For this reason, MOVER includes 2 datasets: the first contains 2 years of data from the SIS EHR system (SIS dataset) and the second contains 5 years of data from the EPIC EHR system (EPIC dataset).

Without question, the most challenging part of building a repository like this is getting the source waveforms out of the system. For the most part, EHR data are available. Each anesthesia workstation manufacturer has their own interface for the acquisition of waveforms that must be custom integrated into a data capture apparatus.

Repository development and curation

Data processing

Developing MOVER involved significant data postprocessing and organization. Following delivery of the data extract, the start and end time of each case were used to extract the appropriate waveform data for that case (based on location and date/time) and link it to the case. The EHR data for both datasets was then restructured and organized into logical tables (comma separated value [csv] files) for simplicity and to help facilitate data analyses. The SIS dataset has a single identifier representing each surgery, surgeries are not linked to patients and therefore it is not possible to track patients temporally across surgeries using this dataset. The raw EHR data for the EPIC dataset contained several redundant identifiers for patients and patient visits that differed between tables. To simplify this, the number of patient/event identifiers in the data were reduced to just 2: the patient identifier and the patient visit identifier. This identification system allows patients to be tracked over time if they have multiple surgeries.

Deidentification and HIPAA compliance

Under HIPAA Privacy Rule, all patient identifiers were removed or deidentified. For deidentification, all patient identifiers and patient visit identifiers were encoded via 1-way hash functions. Additionally, PHI was removed from free text using regular expressions and manually reviewed to ensure that all PHI was removed. Patient ages were capped at 90, so any patient with a recorded age of more than 90 years old was set to 90 years old. Ages were capped to protect patient anonymity because extreme ages are considered identifying. Finally, dates were shifted by a random number of days. The number of days by which to shift the data is linked to each patient to ensure that the data for a single patient is internally consistent. For example, if a patient had 2 surgeries 2 months apart in the raw data, then the deidentified data would also reflect the surgeries as being 2 months apart. It is important to note that this temporal consistency can only be observed for a single patient and not across distinct patients. For example, in the deidentified data 2 patients who are listed as having surgery on the same day in reality did not necessarily have surgery on the same day or even in the same year.

Repository description

The MOVER repository contains 2 datasets: the first contains 2 years of data from the SIS EHR system (SIS dataset) collected from 2015 to 2017 and the second contains 5 years of data from the EPIC EHR system (EPIC dataset) collected from 2017 to 2022. The datasets contain comprehensive

EHR and high-fidelity physiological waveforms of patients who underwent surgeries (Figure 1).

Table names and field names in the repository have been kept consistent wherever possible with the source names from the respective source EHRs. The hope is that this will make integration with other databases from the same EHRs easier in the future, reducing the need for manual field mapping across the datasets. Additionally, while merging the 2 databases into a single unified structure for MOVER would facilitate the use of the data, it would also entail *loss* of some information to unify the elements that differ between the sets. Keeping the data close to the original source is the only "lossless" way to present the data. This means that users will need to do their own work to unify the data if they wish to use both sets, but the benefits are that users will have complete control over how they perform that merge and can tailor the process to their own goals.

SIS dataset

The SIS dataset includes 19 114 patients and is separated into 9 tables: patient information, patient I/O (intake and output), patient vitals, patient observations, patient medications, patient laboratory measurements, patient procedure events, patient ventilator, and patient arterial line (Table 1).

These tables contain patient demographics, information regarding the surgical procedure and anesthesia, laboratory data, and administered medications. This dataset is unique because, in addition to waveforms, it contains high temporal resolution vital signs including cardiac output, blood pressure, and stroke volume variation.

EPIC dataset

The EPIC dataset is the larger of the 2 datasets, containing 39 685 patients, and is separated into 10 tables: patient information, patient history, patient visit, patient medications, patient LDA (lines, drains, and airway devices), patient laboratory measurements, patient measurements, patient postoperative complications, patient procedure events, and patient coding (Table 2).

Similar to the SIS dataset, the EPIC dataset includes patient demographics and specific information regarding the surgical procedure being performed including medicines used, surgical events, and laboratory data. Although similar, a major difference between these 2 datasets is that the EPIC dataset contains outcome information including postoperative complications, mortality, and if the patient was admitted to the ICU and the SIS dataset does not. Additionally, the EPIC dataset includes information about a patient's medical history prior to surgery and their American Society of Anesthesiologists (ASA) physical status, while the SIS dataset does not. The final major difference is that the EPIC dataset includes billing codes.

Repository distribution and documentation

MOVER is available for download at: https://doi.org/10.24432/C5VS5G. It can be downloaded by anyone who signs a DUA, to restrict traffic to legitimate researchers. The website landing page has 3 buttons that lead to corresponding pages: documentation, data download, and article. The documentation page outlines the content of each downloadable table including the meaning of each column, an explanation of the possible values of each column (where applicable), and the unit of measurement for each column (where applicable). On the data download page, users can sign the DUA and download the SIS and EPIC datasets. The paper page displays this publication and the citation to use when citing MOVER.

Results

MOVER includes 58 799 unique patients with data from 83 468 surgeries.

SIS dataset

The SIS dataset is the smaller of the 2 datasets with 19 114 patients and surgeries. Table S1 shows summary statistics and patient demographics for all surgeries in the SIS dataset. The SIS dataset does not contain outcome information however, it does include high temporal resolution vital signs which would

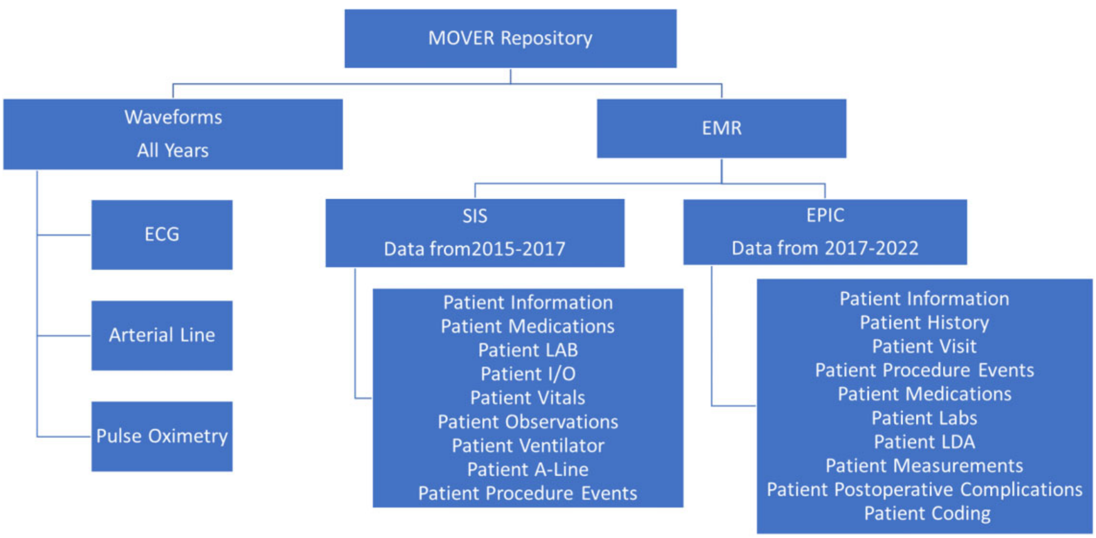

Figure 1. Structured overview of the MOVER data repository. ECG, electrocardiogram; EMR, electronic medical record; SIS, surgical information service; I/O, input/output.

Table 1. Description of the 9 tables included in the SIS dataset.

SIS dataset table	Description
Patient information	Patient demographic information including age, sex, height, and weight; information regarding the surgery being performed including: the type of surgery, the start and end time of the surgery
Patient medications	Medications which were prescribed to patients including: the dose of the medication, the unit of measurement for the dose, and the time the medication was administered or prescribed
Patient labs	Labs ordered for a patient, the corresponding observed measurements, and the time at which each measurement was taken
Patient I/O	Patient fluid input and output information including type of fluid, volume of fluid, and time of fluid input and or output
Patient vitals	Vital signs including HR from EKG, HR from pulse oximetry, noninvasive SBP, noninvasive MAP, noninvasive DBP, and SPO_2
Patient observations	High temporal resolution vital signs including: CO, measurements for the femoral arterial line, measurements from the radial arterial line, SVV, cerebral oximetry, intracranial pressure mean, train of 4, and several others
Patient ventilator	Ventilator information including anesthetic agent, anesthetic agent inspired fraction concentration, respiratory tidal volume, respiratory rate, airway positive inspiratory pressure, inspired fraction of nitrous oxide, and end-tidal fraction of nitrous oxide
Patient arterial line	Patient arterial line placement time, location, and laterality
Patient procedure events	Preoperative, perioperative, and postoperative procedural events and the corresponding time of the event

I/O, input/output; HR, heart rate; SBP, systolic blood pressure; MAP, mean arterial pressure; DBP, diastolic blood pressure; CO, cardiac output; SVV, stroke volume variation.

Table 2. Description of the 10 tables included in the EPIC dataset.

EPIC dataset table	Description
Patient information	Patient demographic information including age, sex, height, and weight; information regarding the surgery being performed including: the type of surgery, the start and end time of both the surgery and anesthesia, ASA status, and discharge disposition
Patient history	Patient's health history including all patient diagnoses available in the EHR
Patient visit	Diagnosis and diagnosis code for a particular visit
Patient medications	Medications which were prescribed to patients including: the dose of the medication, the time the medication was administered or prescribed, and the medication route
Patient LDA	Description of lines, drains, and airway devices used on the patient, the time of placement, the time of removal, and location of placement
Patient labs	Labs ordered for a patient, the corresponding observed measurements, and the reference measurements for each lab
Patient measurements	All preoperative and postoperative patient measurements including: vitals, intake and output, pain levels, complications, and disposition
Patient postoperative complications	The type of complication and a free text note field outlining more specific details
Patient procedure events	Preoperative, perioperative, and postoperative procedural events and the corresponding time of the event
Patient coding	Patient billing codes

ASA, American Society of Anesthesiologists; EHR, electronic health record; LDA, lines, drains and airway devices.

be invaluable for making real-time predictions to assist anesthesiologists.

EPIC dataset

The EPIC dataset makes up the majority of the repository, with 39 685 patients and 64 354 surgeries. Table 3 shows summary statistics and patient demographics for all surgeries in the EPIC dataset.

Of the 64 354 surgeries we can see, for instance, that the average patient age was 55 and that the average length of stay was 7 days. Additionally, looking at the ASA scores, we see a diverse distribution of scores with a mode of 3. The ASA score is a system used to represent a patient's preanesthesia medical comorbidities, with a higher score representing a patient in worse health. Having a diverse distribution of ASA scores in this dataset shows that patients undergoing surgery are in varying degrees of health, which makes this dataset more generalizable than datasets exclusively containing patients in critical condition.

Table 4 characterizes the outcomes available in the EPIC dataset in MOVER. This characterization is useful to investigators to get an idea of what predictions they can make using MOVER.

In the EPIC dataset, 45.3% of patients are transferred to the ICU after surgery and there is a 1.6% mortality rate. Table 4 also shows the percentages of the 11 classes of postoperative complications. Each postoperative complication is assigned to a class and more specific details surrounding the complication can be found in the associated free text. Investigators would be able to use these outcomes individually for specific outcome prediction or use them in combination to understand what factors contribute to a bad outcome of any kind.

Physiological waveforms

Both the SIS and EPIC datasets contain high-fidelity physiological waveform data for certain patients during surgery. Typically, the waveforms contain the electrocardiogram, the arterial waveform, and the pulse oximetry waveform. These

Table 3. Characterization of the EPIC dataset reported as number of records (%) or mean ± SD.

Characteristic	EPIC dataset
Gender	
Female	30 139 (46.8%)
Male	34 214 (53.2%)
Age (years)	55 ± 17
ASA rating	
1	2960 (4.6%)
2	18 068 (28.1%)
3	29 449 (45.8%)
4	6370 (9.9%)
5	657 (1.0%)
6	41 (0.06%)
Length of stay in the hospital (days)	7 ± 14
10 most performed procedures	
Catheterization, heart, left, with intervention if indicated	1521 (2.3%)
Cholecystectomy, laparoscopic	1189 (1.8%)
Laparoscopy, diagnostic	1040 (1.6%)
Laparotomy, exploratory	985 (1.5%)
Dilation and evacuation, uterus	887 (1.4%)
Debridement, with split-thickness skin graft application	741 (1.2%)
Arthroplasty, knee	680 (1.1%)
Irrigation and debridement, lower extremity	623 (1.0%)
ORIF, fracture, femur	608 (0.9%)
AV fistulogram, with angioplasty if indicated	573 (0.9%)

ASA, American Society of Anesthesiologists.

Table 4. Characterization of the EPIC dataset outcomes reported as number of records (%).

Outcome	EPIC dataset
Transfer to the intensive care unit	29 131 (45.3%)
Death	1023 (1.6%)
Postoperative complications	
Other	1093 (1.7%)
Cardiovascular	861 (1.3%)
Respiratory	735 (1.1%)
Airway	373 (0.6%)
Metabolic	154 (0.2%)
Neurological	147 (0.2%)
Administrative	118 (0.2%)
Injury/infection	117 (0.2%)
Medication	94 (0.1%)
Regional	60 (0.1%)
Chronic pain	32 (0.05%)

waveforms provide real-time information about the patients, presenting an invaluable resource for a multitude of retrospective studies.

For instance, utilizing the arterial waveform, it is possible to compute the mean arterial pressure (MAP) for the entire duration of the surgical procedure. Plotting these data points facilitates the visual interpretation of the patient's hemodynamic changes during surgery (Figure 2).

In addition, the arterial waveforms and electrocardiograms can be graphically represented as a snapshot at any given moment during the surgery. This provides a crucial tool for probing the precise physiological changes in the patient at that specific point (Figure 3). Such instantaneous data are not only vital for retrospective examination but also serve as an indispensable tool for making real-time predictions, thereby significantly aiding anesthesiologists in their clinical decision-making processes.

Comparison with other publicly available clinical repositories

To the best of our knowledge, MOVER stands out as the only freely available public data repository that consists of both EHRs and high-fidelity physiological waveform data for patients undergoing surgery. While there are other medical datasets that have been made publicly available, such as MIMIC-IV, some datasets in the UCI Machine Learning Repository, and n2c2: National NLP Clinical Challenges.[14,15]

The MIMIC-IV comprises data from 180 733 unique patients who stayed in the hospital and 50 920 unique patients admitted to the ICU (Table 5). While MIMIC-IV contains a larger patient population than MOVER, its primary focus is on patients in critical condition. On the other hand, MOVER's emphasis lies in data recorded during surgical procedures, encompassing comprehensive details about the procedure itself, patients' physiological changes during the procedure, fluids input-output, as well as anesthesia and ventilator information.

MIMIC-IV focuses on the aspects of hospital and ICU stays, providing rich information about inpatient care, while MOVER concentrates on surgical data, offering valuable insights into procedures and related factors. These 2 datasets serve different purposes, and their synergy is evident as they collectively cover various aspects of medical care across different scenarios.

The UCI Machine Learning Repository has a very limited number of small clinical datasets focusing on very specific health issues, such as diabetes, and does not include complete EHR data. The n2c2 exclusively contains unstructured text and therefore can only be used for natural language processing applications. We believe that the publication of MOVER will help address some of these limitations and complement these other publicly available datasets.

Discussion

While US hospitals have adopted EHR documentation of patient care, interoperability of these systems remains an open issue, leading to challenges in data integration. In the OR setting, besides EHR clinical data, physiological waveforms represent a large source of information.[17–20] Monitors analyze physiological waveforms to extract and display information used by clinicians to make decisions but modern science must transform these monitors into early warning systems and forecasting devices, without increasing alarm pollution.[19–21] Although anesthesiologists can use available monitoring solutions to rescue unstable patients—decreasing the incidence of cardiac arrests and mortality[22–24]—the more proactive approach would be to enable them to recognize impending instability during surgery before it happens. This would allow an even more decisive impact on patients' outcomes. However, to achieve this goal, close collaboration between industry, academia, and healthcare is required.[25] Especially, in order to accelerate the discovery of new knowledge and the creation of new, impactful monitoring devices, EHR and monitoring data need to be less fragmented and more accessible to the research community.[25,26] In the ICU setting, such

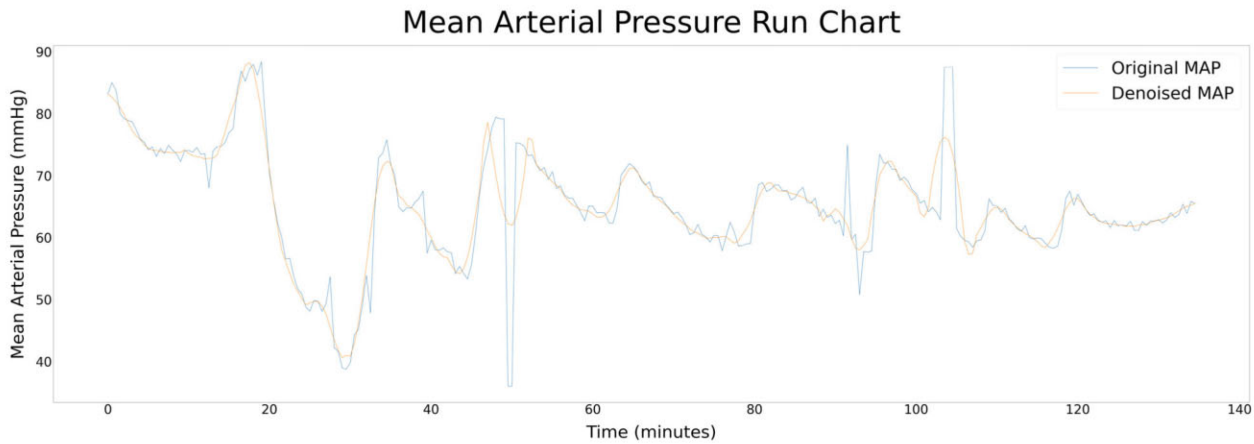

Figure 2. The raw and denoised MAP run chart during surgery. MAP, mean arterial pressure

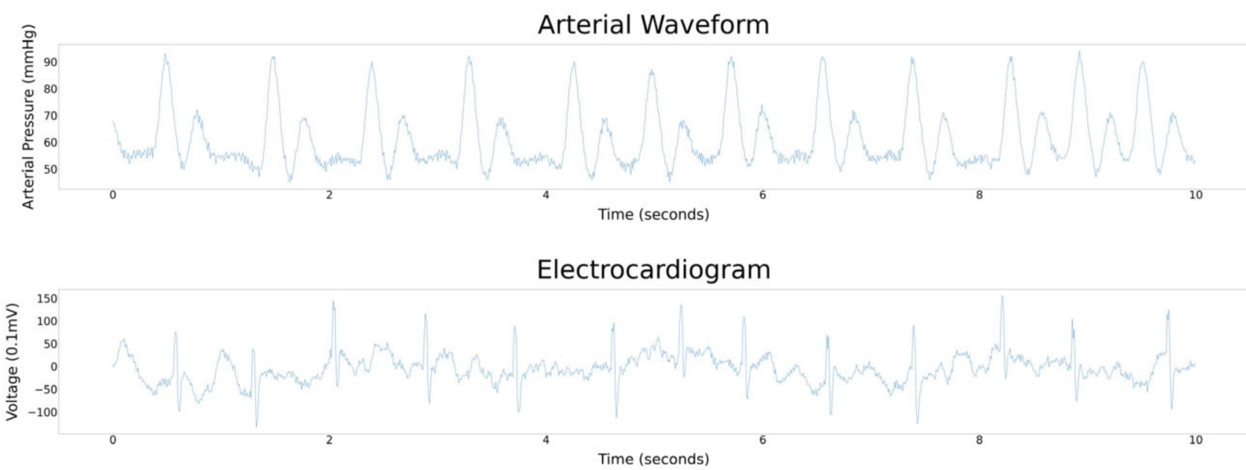

Figure 3. A 10-s snapshot of the arterial waveform and ECG during surgery.

Table 5. The comparison of patient demographic information in MOVER and MIMIC-IV.

Data repository	MIMIC-IV	MIMIC-IV	MOVER
Patient type	Hospital admissions	ICU admissions	Surgeries
Number of unique patients	180 733	50 920	58 799
Patient age, mean (SD)	58.8 (19.2)	64.7 (16.9)	55.0 (17.0)
Female administrative gender, %	52.2	44.2	46.8
Length of hospital stay days, mean (SD)	4.5 (6.6)	11.0 (13.3)	7.0 (14.0)

an initiative has been developed and has led to the development of the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) database.[17,18,27] This database is public and has been used for the development and validation of several new monitoring devices, predictive analytics solutions, and improvement in our understanding of physiology in the ICU.[28–32] Similarly, the Physionet project offers access to large collections of physiologic signals and related open-source softwares[15,16,33,34] but none is specifically based in the surgical setting. For the perioperative setting, the Multicenter Perioperative Outcome Group (MPOG)[35] and the National Surgical Quality Improvement Program (N-SQIP)[36,37] provide high quality perioperative process and outcome measures but do not provide fused EHR and high-fidelity physiological waveform data that can be used to developed novel predictive physiological tools. This is a significant gap because while 5.7 million Americans are admitted to an ICU each year, more than 50 million undergo surgery annually and the incidence of postoperative complications remains high and represents a burden to our society.[38–42]

The creation and release of the MOVER database is innovative in many ways. First, while the MPOG[35] and the N-SQIP[36,37] provide high-quality perioperative process and outcome measures, they are not publicly available and they do not provide fused EHR and high-fidelity physiological waveform data that can be used to develop novel predictive physiological tools. The MOVER database is the first public database to offer this capability. Second, while publicly accessible databases including EHR and physiological waveforms exist in the critical care setting (MIMIC database and Physionet), the MOVER database is the first to offer the same access for a large census of surgical patients. We have developed catalogs and syllabus of perioperative EHR and physiological waveform data and created the architecture of a scalable and searchable perioperative database for future sister database integration. Third, the surgical/OR environment is unique in the sense that knowledge of the baseline/presurgical stress status of essentially all patients before surgery allows normalization, calibration, and markedly enhances development of predictive tools. In addition, the continuous and immediate presence of dense skilled acute care practitioners in the OR allows implementation of complex treatment algorithms much faster than in the ICU. Finally, defined stages (anesthesia induction, tracheal intubation, skin incision, anesthesia emergence and others), procedures, and stressors (intra-abdominal air insufflation, prone positioning for back surgery, and other surgery-specific interventions) allow machine learning approaches to build large common relational database registries.

Conclusion

AI holds great promise for transforming healthcare, but its widespread deployment in clinical settings has been limited due to the scarcity of publicly available clinical data and transparent AI algorithms. Few accessible clinical data repositories exist for training and testing AI algorithms, especially in the perioperative setting. To bridge this gap, we introduce the freely available MOVER. Researchers can access MOVER after signing a DUA. Our aim is to accelerate the integration of AI into healthcare, ultimately improving patient outcomes.

Acknowledgments

We also wish to express our gratitude towards undergraduate researcher Kaiwen Yu for his assistance in the data mining process.

Author contributions

M.S. conducted the study, drafted the manuscript, and final approval of the manuscript, J.R. designed the study, conducted the study, drafted the manuscript, and final approval of the manuscript, M.A. conducted the study, drafted the manuscript, and final approval of the manuscript, Y.K. conducted the study, drafted the manuscript, and final approval of the manuscript, P.B. designed the study, conducted the study, drafted the manuscript, and final approval of the manuscript, and M.C. designed the study, conducted the study, drafted the manuscript, and final approval of the manuscript.

Supplementary material

Supplementary material is available at *JAMIA Open* online.

Funding

This work was supported by National Institute of Health (NIBIB R01EB029751) and by Edwards Lifesciences (PA-1411-001).

Conflict interest

Dr. Cannesson is a consultant for Edwards Lifesciences and Masimo Corp, and has funded research from Edwards Lifesciences and Masimo Corp. He is also the founder of Sironis and Perceptive Medical and he owns patents and receives royalties for closed loop hemodynamic management technologies that have been licensed to Edwards Lifesciences. Dr. Rinehart is a consultant for Edwards Lifesciences. He is also the founder of Sironis and Perceptive Medical and he owns patents and receives royalties for closed loop hemodynamic management technologies that have been licensed to Edwards Lifesciences.

Data availability

The data presented in this manuscript is available in its entirety at: https://doi.org/10.24432/C5VS5G.

References

1. Everson J, Rubin JC, Friedman CP. Reconsidering hospital EHR adoption at the dawn of HITECH: implications of the reported 9% adoption of a "basic" EHR. *J Am Med Inform Assoc*. 2020;27(8):1198-1205. https://doi.org/10.1093/jamia/ocaa090
2. Adler-Milstein J, Holmgren AJ, Kralovec P, *et al*. Electronic health record adoption in US hospitals: the emergence of a digital "advanced use" divide. *J Am Med Inform Assoc*. 2017;24(6):1142-1148. https://doi.org/10.1093/jamia/ocx080
3. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94-98. https://doi.org/10.7861/futurehosp.6-2-94
4. Lee CK, Hofer I, Gabel E, Baldi P, Cannesson M. Development and validation of a deep neural network model for prediction of postoperative in-hospital mortality. *Anesthesiology*. 2018;129(4):649-662. https://doi.org/10.1097/ALN.0000000000002186
5. Lee CK, Samad M, Hofer I, Cannesson M, Baldi P. Development and validation of an interpretable neural network for prediction of postoperative in-hospital mortality. *NPJ Digit Med*. 2021;4(1):8. https://doi.org/10.1038/s41746-020-00377-1
6. Lilot M, Ehrenfeld JM, Lee C, *et al*. Variability in practice and factors predictive of total crystalloid administration during abdominal surgery: retrospective two-Centre analysis. *Br J Anaesth*. 2015;114(5):767-776. https://doi.org/10.1093/bja/aeu452
7. Hill BL, Brown R, Gabel E, *et al*. An automated machine learning-based model predicts postoperative mortality using readily-extractable preoperative electronic health record data. *Br J Anaesth*. 2019;123(6):877-886. https://doi.org/10.1016/j.bja.2019.07.030
8. Cannesson M, Hofer I, Rinehart J, *et al*. Machine learning of physiological waveforms and electronic health record data to predict, diagnose and treat haemodynamic instability in surgical patients: protocol for a retrospective study. *BMJ Open*. 2019;9(12):e031988. https://doi.org/10.1136/bmjopen-2019-031988
9. Kopitar L, Kocbek P, Cilar L, Sheikh A, Stiglic G. Early detection of type 2 diabetes mellitus using machine learning-based prediction models. *Sci Rep*. 2020;10(1):11981. https://doi.org/10.1038/s41598-020-68771-z
10. Datta A, Flynn NR, Barnette DA, *et al*. Machine learning liver-injuring drug interactions with non-steroidal anti-inflammatory drugs (NSAIDs) from a retrospective electronic health record

(EHR) cohort. *PLoS Comput Biol*. 2021;17(7):e1009053. https://doi.org/10.1371/journal.pcbi.1009053
11. Lauritsen SM, Kristensen M, Olsen MV, *et al*. Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nat Commun*. 2020;11(1):3852. https://doi.org/10.1038/s41467-020-17431-x
12. Baldi P., ed. *Deep Learning in Science*. Cambridge University Press; 2021.
13. Meng C, Trinh L, Xu N, Enouen J, Liu Y. Interpretability and fairness evaluation of deep learning models on MIMIC-IV dataset. *Sci Rep*. 2022;12(1):7166. https://doi.org/10.1038/s41598-022-11012-2
14. Johnson AEW, Bulgarelli L, Shen L, *et al*. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10(1):1. https://doi.org/10.1038/s41597-022-01899-x
15. Goldberger AL, Amaral LA, Glass L, *et al*. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. 2000;101(23):E215-220.
16. Moody GB, Mark RG, Goldberger AL. PhysioNet: a research resource for studies of complex physiologic and biomedical signals. *Comput Cardiol*. 2000;27:179-182.
17. Lee J, *et al*. Open-Access MIMIC-II database for intensive care research. *Annu Int Conf IEEE Eng Med Biol Soc*. 2011;2011:8315-8318. https://doi.org/10.1109/IEMBS.2011.6092050
18. Saeed M, Villarroel M, Reisner AT, *et al*. Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): a public-access ICU database. *Crit Care Med*. 2011;39(5):952-960. https://doi.org/10.1097/CCM.0b013e31820a92c6
19. Shelley K, Cannesson M. Off-label use of clinical monitors: what happens when new physiologic understanding meets state-of-the-art technology. *Anesth Analg*. 2014;119(6):1241-1242. https://doi.org/10.1213/ANE.0000000000000479
20. Bickler PE, Cannesson M, Shelley KH. Trends and challenges in clinical monitoring: Papers from the 2015 IAMPOV symposium. *Anesth Analg*. 2017;124(1):2-3. https://doi.org/10.1213/ANE.0000000000001495
21. Pinsky MR, Dubrawski A. Gleaning knowledge from data in the intensive care unit. *Am J Respir Crit Care Med*. 2014;190(6):606-610. https://doi.org/10.1164/rccm.201404-0716CP
22. Winters BD, Pham JC, Hunt EA, *et al*. Rapid response systems: a systematic review. *Crit Care Med*. 2007;35(5):1238-1243. https://doi.org/10.1097/01.CCM.0000262388.85669.68
23. Chan PS, Khalid A, Longmore LS, *et al*. Hospital-wide code rates and mortality before and after implementation of a rapid response team. *JAMA*. 2008;300(21):2506-2513. https://doi.org/10.1001/jama.2008.715
24. Goodrich C. Endpoints of resuscitation: what should we be monitoring? *AACN Adv Crit Care*. 2006;17(3):306-316. https://doi.org/10.4037/15597768-2006-3008
25. Pronovost PJ, Powers J, Jin W. Technology development in health care is broken. *Am J Med Qual*. 2017;32(2):215-217. https://doi.org/10.1177/1062860616666165
26. Celi LA, Marshall JD, Lai Y, Stone DJ. Disrupting electronic health records systems: the next generation. *JMIR Med Inform*. 2015;3(4):e34. https://doi.org/10.2196/medinform.4192
27. Scott DJ, Lee J, Silva I, *et al*. Accessing the public MIMIC-II intensive care relational database for clinical research. *BMC Med Inform Decis Mak*. 2013;13:9. https://doi.org/10.1186/1472-6947-13-9
28. Ghassemi MM, Richter SE, Eche IM, *et al*. A data-driven approach to optimized medication dosing: a focus on heparin. *Intensive Care Med*. 2014;40(9):1332-1339. https://doi.org/10.1007/s00134-014-3406-5
29. Badawi O, Brennan T, Celi LA, *et al*.; MIT Critical Data Conference 2014 Organizing Committee. Making big data useful for health care: a summary of the inaugural MIT critical data conference. *JMIR Med Inform*. 2014;2(2):e22. https://doi.org/10.2196/medinform.3447
30. Ghassemi M, Marshall J, Singh N, Stone DJ, Celi LA. Leveraging a critical care database: selective serotonin reuptake inhibitor use prior to ICU admission is associated with increased hospital mortality. *Chest*. 2014;145(4):745-752. https://doi.org/10.1378/chest.13-1722
31. Pirracchio R, Petersen ML, Carone M, *et al*. Mortality prediction in intensive care units with the super ICU learner algorithm (SICULA): a population-based study. *Lancet Respir Med*. 2015;3(1):42-52. https://doi.org/10.1016/S2213-2600(14)70239-5
32. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. *Sci Transl Med*. 2015;7(299):299ra122. https://doi.org/10.1126/scitranslmed.aab3719
33. Costa M, Moody GB, Henry I, Goldberger AL. PhysioNet: an NIH research resource for complex signals. *J Electrocardiol*. 2003;36(Suppl):139-144. https://doi.org/10.1016/j.jelectrocard.2003.09.038
34. Moody GB, Mark RG, Goldberger AL. PhysioNet: a web-based resource for the study of physiologic signals. *IEEE Eng Med Biol Mag*. 2001;20(3):70-75. https://doi.org/10.1109/51.932728
35. Kheterpal S. Clinical research using an information system: the multicenter perioperative outcomes group. *Anesthesiol Clin*. 2011;29(3):377-388. https://doi.org/10.1016/j.anclin.2011.06.002
36. Khuri SF, Daley J, Henderson W, *et al*. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. *Ann Surg*. 1998;228(4):491-507. https://doi.org/10.1097/00000658-199810000-00006
37. Berian JR, Zhou L, Hornor MA, *et al*. Optimizing surgical quality datasets to care for older adults: Lessons from the American college of surgeons NSQIP geriatric surgery pilot. *J Am Coll Surg*. 2017;225(6):702-712.e701. https://doi.org/10.1016/j.jamcollsurg.2017.08.012
38. Birkmeyer JD, Gust C, Dimick JB, Birkmeyer NJ, Skinner JS. Hospital quality and the cost of inpatient surgery in the United States. *Ann Surg*. 2012;255(1):1-5. https://doi.org/10.1097/SLA.0b013e3182402c17
39. Khan NA, Quan H, Bugar JM, *et al*. Association of postoperative complications with hospital costs and length of stay in a tertiary care center. *J Gen Intern Med*. 2006;21(2):177-180. https://doi.org/10.1111/j.1525-1497.2006.00319.x
40. Pradarelli JC, Healy MA, Osborne NH, *et al*. Variation in Medicare expenditures for treating perioperative complications: the cost of rescue. *JAMA Surg*. 2016;151(12):e163340. https://doi.org/10.1001/jamasurg.2016.3340
41. International Surgical Outcomes Study Group. Global patient outcomes after elective surgery: prospective cohort study in 27 low-, Middle- and high-income countries. *Br J Anaesth*. 2016;117:601-609. https://doi.org/10.1093/bja/aew316
42. Whitlock EL, Feiner JR, Chen LL. Perioperative mortality, 2010 to 2014: a retrospective cohort study using the national anesthesia clinical outcomes registry. *Anesthesiology*. 2015;123(6):1312-1321.