

---

## Research and Applications

# The MIMIC Code Repository: enabling reproducibility in critical care research

Alistair EW Johnson,<sup>1,\*</sup> David J Stone,<sup>2</sup> Leo A Celi,<sup>1,3</sup> and Tom J Pollard<sup>1</sup>

<sup>1</sup>Massachusetts Institute of Technology, Cambridge, MA, USA, <sup>2</sup>University of Virginia School of Medicine, Charlottesville, VA, USA, and <sup>3</sup>Beth Israel Deaconess Medical Center, Boston, MA, USA

\*Corresponding Author: Alistair EW Johnson, Massachusetts Institute of Technology, E25-505, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. E-mail: aewj@mit.edu.

Received 25 May 2017; Revised 11 July 2017; Editorial Decision 21 July 2017; Accepted 27 July 2017

### ABSTRACT

**Objective:** Lack of reproducibility in medical studies is a barrier to the generation of a robust knowledge base to support clinical decision-making. In this paper we outline the Medical Information Mart for Intensive Care (MIMIC) Code Repository, a centralized code base for generating reproducible studies on an openly available critical care dataset.

**Materials and Methods:** Code is provided to load the data into a relational structure, create extractions of the data, and reproduce entire analysis plans including research studies.

**Results:** Concepts extracted include severity of illness scores, comorbid status, administrative definitions of sepsis, physiologic criteria for sepsis, organ failure scores, treatment administration, and more. Executable documents are used for tutorials and reproduce published studies end-to-end, providing a template for future researchers to replicate. The repository's issue tracker enables community discussion about the data and concepts, allowing users to collaboratively improve the resource.

**Discussion:** The centralized repository provides a platform for users of the data to interact directly with the data generators, facilitating greater understanding of the data. It also provides a location for the community to collaborate on necessary concepts for research progress and share them with a larger audience. Consistent application of the same code for underlying concepts is a key step in ensuring that research studies on the MIMIC database are comparable and reproducible.

**Conclusion:** By providing open source code alongside the freely accessible MIMIC-III database, we enable end-to-end reproducible analysis of electronic health records.

**Key words:** critical care, reproducibility, mimic-iii, data mining, intensive care, electronic health record

---

### INTRODUCTION

Concerns about the reproducibility of results in science are becoming increasingly prominent in both scientific and mainstream literature.<sup>1</sup> Some commentators have gone so far as to call the current state a crisis, citing causes such as pressure to publish positive results, the cost of replicating studies such as double-blind randomized controlled clinical trials, and the lack of emphasis on reproducibility as a requirement for sound science.

In parallel, health care has been undergoing a digital revolution in recent years. The Health Information Technology for Economic and Clinical Health Act has catalyzed the transition of hospitals and care institutions from paper-based to electronic-based systems.<sup>2</sup> Vast quantities of digital data are now routinely collected by modern hospital monitoring systems, even more so in intensive care units (ICUs), where patients require close observation. There is optimism that increasing the availability of large-scale clinical databases will

offer opportunities to overcome many of the challenges associated with the lack of evidence in medical practice.

The Medical Information Mart for Intensive Care (MIMIC-III) database is an example of such a data repository.<sup>3,4</sup> The database comprises detailed clinical information regarding >60 000 stays in ICUs at the Beth Israel Deaconess Medical Center in Boston, Massachusetts, collected as part of routine clinical care. The MIMIC-III dataset is freely available to researchers around the world and has been widely used in the development of predictive models, epidemiological studies, and educational courses.

Perhaps the most important insight since the database was made open access is how challenging research using electronic health records (EHRs) can be, requiring close collaboration between domain experts and data scientists. As MIMIC-III is a deidentified version of raw data stored during routine clinical care, a nontrivial body of work is required to transform the data into a usable form for research. This derivation of clinical concepts on an EHR database is a resource-intensive task, however, and is a significant barrier to those unfamiliar with the clinical environment or the database structure. Moreover, if concepts are not defined collaboratively with those who are familiar with the workflows, including how the data are captured, the validity of the findings may be suspect.

In this paper, we describe the MIMIC Code Repository, a centralized location for derived concepts that are relevant to critical care research. Detailed descriptions on how the concepts are defined and extracted from the database are provided, including the assumptions that are made and the conditions for which codes or queries are valid. Additional tools are provided to educate researchers on best practices for conducting a fully reproducible study using the database. The code is open source, follows good documentation practices, and is contributed to by members of the research community using MIMIC-III.

The repository provides a framework for collaboration around research. While the case for open data has been already been strongly made elsewhere, we believe open code is equally important. We would make the argument that the use of an openly available code repository will improve secondary analysis of health data by accelerating the understanding of datasets by researchers, and improving the consistency and validity of future studies.

## THE MIMIC CODE REPOSITORY

The MIMIC Code Repository is available online<sup>5</sup> and is open source. Code is available as standardized scripts in languages including Structured Query Language (SQL), Python, and R. Scripts are modified to allow an individual who has been granted access to the MIMIC-III database to generate a number of “views” of the data, with each view being an extraction from the raw data. Each script is associated with an automatically generated unique commit hash that acts as an identifier for the code. Publications that use the code repository can further cite the commit hash, allowing other researchers to download a copy of the code used regardless of any modifications since. All code follows the principles of good scientific programming as outlined by Wilson et al.,<sup>6</sup> including incremental development with a distributed version control system, unit tests, and a public issue tracker. The repository was tested on MIMIC-III v1.4 at the time of this publication.

There are 3 components to the repository that facilitate navigation of the data for research purposes. These components are:

- *Concepts*: Code to extract important concepts from the health records. For example, a module on acute kidney injury uses the criteria as specified by Kidney Disease: Improving Global Outcomes and provides the code to identify patients with acute kidney injury in MIMIC.
- *Executable documents*: Notebooks that allow text and analytical code to be seamlessly combined into a single executable document, allowing studies and tutorials to be reproduced.
- *Community*: Public discussions to facilitate contributions from members of the MIMIC research community.

## Concepts

Code to extract concepts that are broadly applicable to research questions in critical care are provided in the repository. For example, severity of illness scores are frequently required to adjust for confounding factors in a study, but are complex to derive, and so scripts are provided for reuse. These and other concepts are coded in a modular fashion to reduce redundancy in code and allow for extension. The following sections describe various concepts currently available in the repository.

### Severity of illness scores

Severity of illness scores have been developed over recent decades to provide an assessment of the patient’s acuity, particularly but not exclusively at the time of admission to the ICU.<sup>7</sup> The principal aim of these scores is to risk-adjust patient populations for benchmarking and research purposes, such as comparing cohorts in clinical trials and observational studies. In the context of performing research using MIMIC-III, the use of severity of illness scores for risk adjustment is almost always required to address confounding.

While severity of illness scores are integral to risk adjustment, their calculation, if done retrospectively, presents challenges. Most severity scores were developed with well-curated datasets, put together through prospective data collection or manual data abstraction by dedicated trained personnel. As a result, the data tend to be cleaner and often have, perhaps more importantly, a distribution that is markedly different from routinely collected data, such as that present in an electronic health record.

Secondly, routinely collected data often lack data elements required to compute the score. For example, the comorbidity “biopsy proven cirrhosis” is required for the Acute Physiology Score and Chronic Health Evaluation system, but this concept is not documented in a structured manner during routine care. Finally, the data definitions for the same concept can vary between the original dataset used to define the severity score and the EHRs being analyzed. To illustrate this potential disparity, the Glasgow Coma Scale (GCS), a common marker of neurological dysfunction that ranges from 3 (worst) to 15 (best), is usually assumed to be 15 for patients who are unable to be assessed due to sedation or ventilation but otherwise appear to be neurologically intact. In an EHR, however, this definition is not strictly adhered to, as there is no defined protocol, and as a result, sedated patients may be assigned a score of 15 by some care providers and a score of 3 by others.

Working with local nurses and doctors has helped us to address the kinds of issues that potentially impact the code, helping to ensure that the derived scores accurately reflect the true severity of patient illness. There are 5 severity of illness scores currently implemented in the MIMIC Code Repository: acute physiology score (APS)-III,<sup>8</sup> simplified acute physiology score (SAPS),<sup>9</sup> SAPS-II,<sup>10</sup> and the Oxford acute severity of illness score (OASIS).<sup>11</sup>

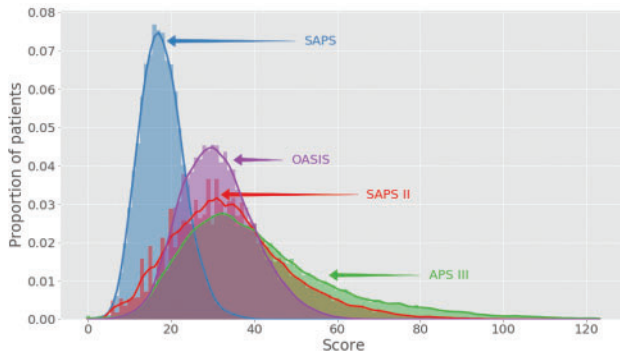


Figure 1. Comparison of severity of illness score distributions.

A more detailed comparison of the severity scores is provided in the [supplementary material](#), along with a discussion of the assumptions made in calculating the scores. Organ dysfunction scores are also available and detailed later.

Each score comprises at least 10 independent components. The APS III, SAPS II, Sequential Organ Failure Assessment (SOFA), Logistic Organ Dysfunction system (LODS), and OASIS scores are generally calculated using data from the first 24 h of the patient's stay. Systemic inflammatory response syndrome score and qSOFA are screening tools with scores calculated on admission to the ICU, which is concretely defined as up to 2 h after the admission time. The distribution of these scores is shown in in [Figure 1](#).

### Organ dysfunction scores

Organ failure is a hallmark of acute illness and is quantified in numerous scores. Some scores assess multiple organ systems: the SOFA score<sup>12</sup> and LODS<sup>13</sup> both assess 6 organ systems for failure. Others are organ-specific. Examples include the Model for End-stage Liver Disease,<sup>14</sup> the Risk/Injury/Failure/Loss/End-stage renal disease criteria,<sup>15</sup> the Acute Kidney Injury Network classification,<sup>16</sup> and the Kidney Disease: Improving Global Outcomes criteria.<sup>17</sup> The latter 3 scores assess the degree of acute kidney injury in patients. A variety of lab, diagnostic, and therapeutic data are needed to calculate these scores.

To highlight the discrepancies that can arise from the way a concept is defined, we contrast 2 versions of the SOFA score, 1 derived by prior researchers and 1 available in the MIMIC Code Repository. [Figure 2](#) shows the area under the receiver operator characteristic curve for hospital mortality of patients admitted in the MIMIC-III database between 2001 and 2008 using 2 versions of SOFA, grouped by year of admission.

The disagreement between the 2 modalities is multifactorial, but a major contributing factor relates to an important variable: the GCS. In the original paper describing the SOFA score, clinicians were instructed to set GCS to its maximum value, 15, if they were unable to assess patients fully (eg, if patients were sedated to facilitate mechanical ventilation). In contrast, the documentation of GCS for these patients in the MIMIC-III database is usually a value of 3, the minimum value, with a note that they are unable to assess the patients. Naive use of GCS values results in a dramatic difference in the capability of the score to discriminate severely ill patients and highlights the need to understand variables and how they are captured or derived. In the MIMIC Code Repository, special extraction steps are used to detect a GCS value of 3 due to sedation, and these values are corrected to 15 in the calculation of scores.

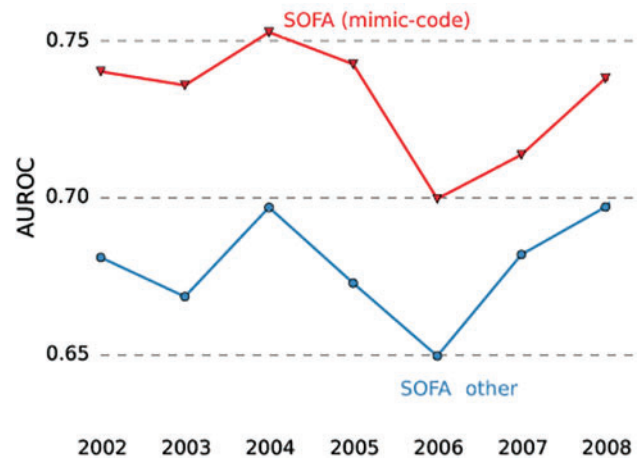


Figure 2. Comparison of areas under the receiver operating curve for SOFA scores calculated from MIMIC code and a prior research report.

### Timing of treatment

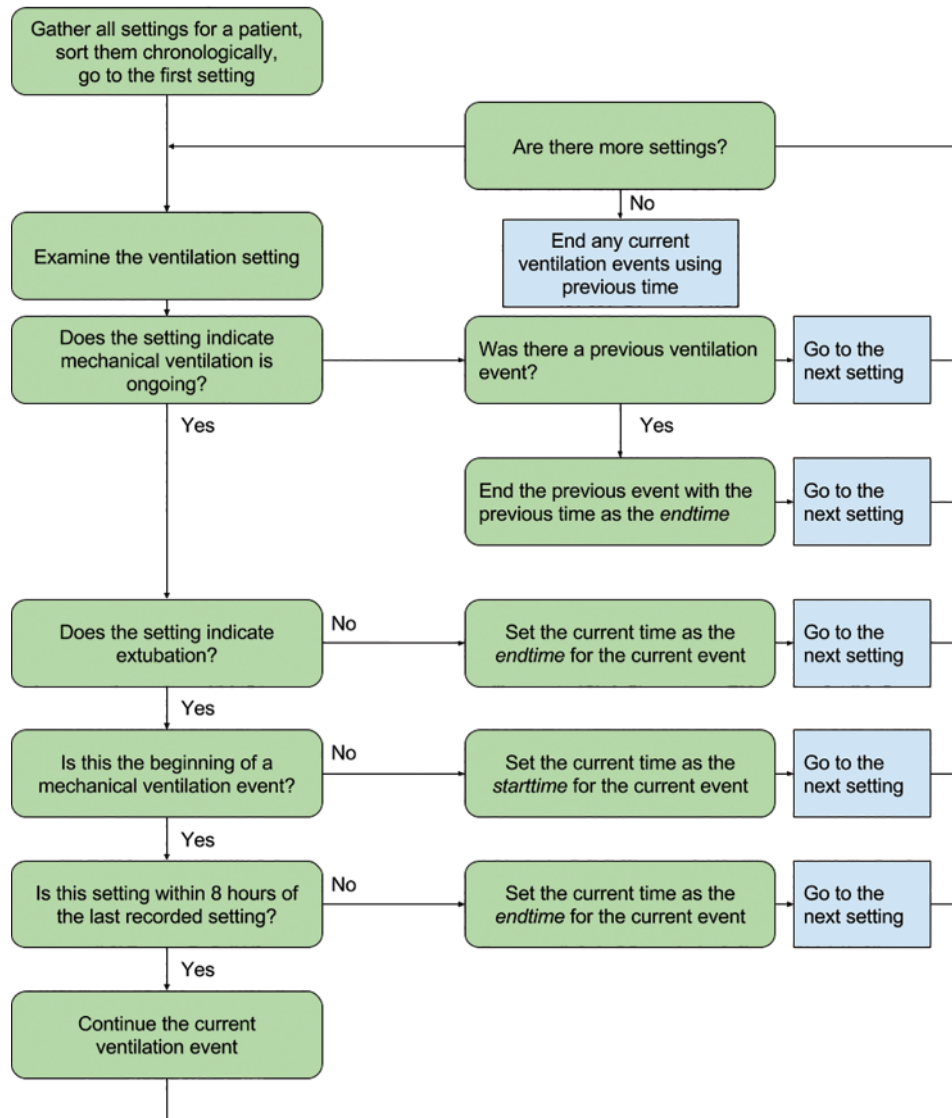
The timing and duration of treatment are important concepts for researchers seeking to understand issues that relate to the intensity of an administered intervention. Duration can serve as an indirect metric of severity and has been used in the development of decision support tools.<sup>18</sup>

As a result of data-capture limitations in the hospital, the exact timing and duration of many medications and treatments are not explicitly available and so must be derived. Derivation can involve identifying surrogate data known to be carried out with a high level of compliance, documented by clinical staff contemporaneous to the treatment. [Figure 3](#) shows a schema for the derivation of the start and stop times of mechanical ventilation. Similar rules are used to define the timing of vasopressor administration and continuous renal replacement therapy (CRRT) available in the repository. Clinical expertise is invaluable in developing these rules and interpreting the fine points of the medical chart that determine them.

An example of a patient undergoing mechanical ventilation and receiving vasopressor agents is provided in [Figure 4](#).

### Sepsis

Sepsis is a major and costly disease in the ICU, costing over \$20 billion in the United States in 2011 (5.2% of all US hospital costs),<sup>19</sup> and growing to over \$23 billion in 2013 (6.2% of all US hospital costs).<sup>20</sup> Sepsis has traditionally been defined as the concurrent presence of systemic inflammation and infection, but a recent reexamination of the problem has suggested redefining the disease as life-threatening organ dysfunction caused by a dysregulated host response to infection.<sup>21</sup> The precise onset of sepsis is not typically documented in the EHR, and is, in fact, a difficult item to capture clinically. In their quantitative evaluation of septic patients, Seymour et al.<sup>22</sup> first identified patients suspected of having infection by cross-referencing antibiotic use with requests for a microbiology assessment. We implemented a similar approach, defining suspected infection as the acquisition of a microbiology culture followed by or shortly after ICU admission. Using this definition, and following the Sepsis-3 guidelines, we define sepsis as suspicion of infection associated with organ failure as quantified by an increase in  $SOFA \geq 2$ . This definition is admittedly a proxy for the actual onset of sepsis, but in the absence of more precise markers, it serves as an approximation of onset time and could be used for the development of deci-



**Figure 3.** Logic behind the query for converting aperiodically recorded ventilator settings into durations of mechanical ventilation.

sion support tools. Scripts for these concepts are available and a notebook describing the derivation is also available.

Identification of sepsis has also been done retrospectively using administrative data, and in particular billing codes acquired on hospital discharge. Two billing codes explicitly denote sepsis (International Classification of Diseases, Ninth Revision [ICD-9] codes 785.52 and 995.92). Angus et al.<sup>23</sup> and Martin et al.<sup>24</sup> describe algorithms for defining sepsis using a set of diagnostic and procedural ICD-9 codes. The criteria as proposed by Angus et al.<sup>23</sup> were validated in a later study by Iwashyna et al.<sup>25</sup> These 3 criteria – explicit coding, those proposed by Angus et al.,<sup>23</sup> and those proposed by Martin et al.<sup>24</sup> – are available in the repository.

### Comorbidities

Many ICU patients have chronic conditions prior to their acute presentation that affect the probability of their surviving critical illness. Elixhauser et al.<sup>26</sup> codified these comorbidities into 29 categories using administrative data, specifically ICD-9 codes. The American Health and Research Quality group continues to maintain these ad-

ministrative codes via the Healthcare Cost and Utilization Project, adapting them accordingly as changes are made to diagnosis and treatment coding.<sup>27</sup> Finally, Quan et al.<sup>28</sup> proposed an enhanced ICD-9 coding methodology based on examining inconsistencies among previous definitions. Diagnosis-related groups, which are used to bill for the principal diagnosis for a patient hospitalization, are used to filter out those conditions that are not present prior to hospitalization. A comparison of these 3 methods is provided in Figure 5. These representations of comorbidities are provided in the repository, both with and without diagnosis-related group filtering.

Van Walraven et al.<sup>29</sup> later aggregated comorbidities codified by Elixhauser et al.<sup>26</sup> into a single point score for in-hospital mortality prediction, which is also available in the repository.

### Concept road map

Table 1 provides a list of currently available concepts in the repository, as well as concepts that are planned for future development (italicized). As code is demand-driven, the planned concepts are not exhaustive.

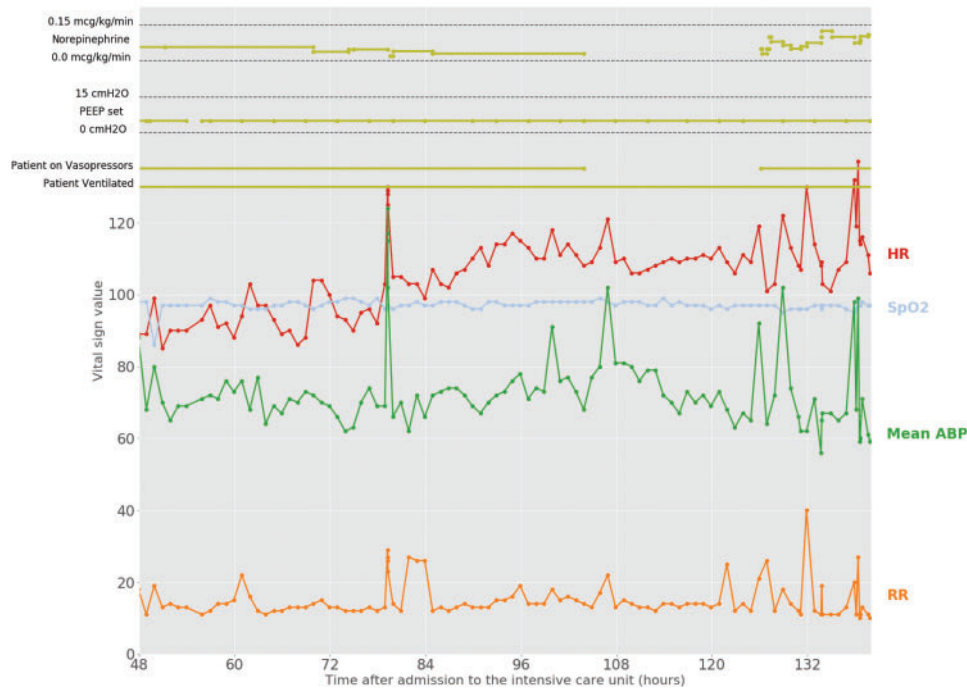


Figure 4. Example of a patient who was both mechanically ventilated and receiving vasopressors for cardiovascular support.

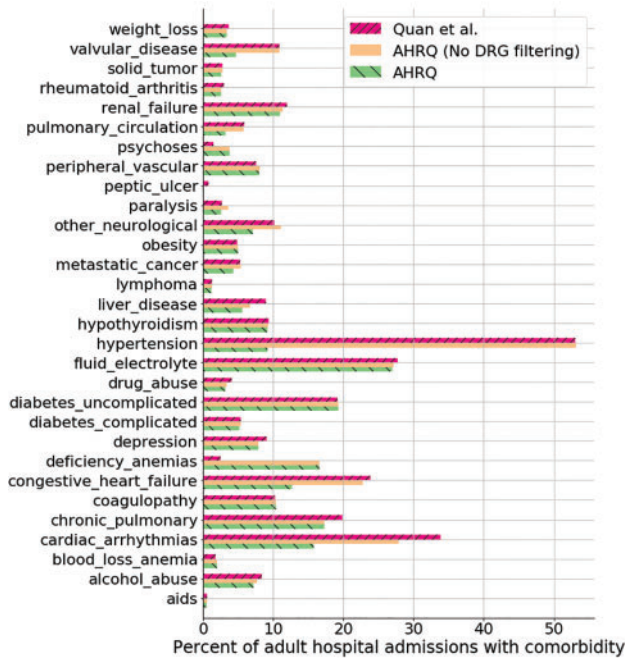


Figure 5. Comparison of 3 methods for calculating presence of a comorbidity for a patient using billing data: an updated coding from the AHRQ which uses diagnosis-related group (DRG) codes to mask non-comorbid conditions, the same coding without the DRG masking, and finally an alternative coding that does not use DRG masking, proposed by Quan et al.<sup>28</sup>

Executable documents

When both data and code are freely available to researchers, as is now the case for MIMIC-III, this provides a framework that allows a study to be entirely reproduced. This is especially powerful when toolkits such as R Markdown and Jupyter Notebook are employed,

Table 1. Concepts available in the repository

Category	Concepts
Severity of illness scores	APS III, SAPS, SAPS II, OASIS
Organ dysfunction scores	SOFA, qSOFA, LODS, SIRS, MELD, KDIGO, AKIN
Treatments	Continuous renal replacement therapy, intermittent hemodialysis, vasopressors, mechanical ventilation
Sepsis	Suspicion of infection, Angus et al. criteria, Martin et al. criteria, explicit ICD-9 coding of sepsis, <i>CMS sepsis criteria</i> , <i>CDC sepsis criteria</i>
Comorbid burden	Elixhauser et al. (AHRQ), Quan et al., <i>Charlson et al.</i>
First 24 h aggregates	Vital signs, laboratory values, blood gas values, urine output
Diagnosis groups	Certified Coding Specialist groups
Demographics	Weight, height, age, gender, <i>service type</i>
<i>Hourly data</i>	<i>Vasopressor doses, vital signs, laboratory values, blood gas values</i>
<i>Fluid balance</i>	<i>Total fluid intake, total fluid output</i>

Concepts that are italicized are planned for future release. MELD: Model for End-stage Liver Disease; SIRS: systemic inflammatory response syndrome; KDIGO: Kidney Disease: Improving Global Outcomes; AKIN: Acute Kidney Injury Network; CMS: Centers for Medicare and Medicaid Services; CDC: Centers for Disease Control and Prevention; AHRQ: Agency for Healthcare Research and Quality.

allowing documentation and code to be seamlessly combined to create executable documents. Figure 6 shows an example of a Jupyter Notebook that extracts patient demographics and displays the results for the user to view. Jupyter Notebooks are language agnostic, supporting code written in Python, R, MATLAB, SAS, and others.<sup>30,31</sup>

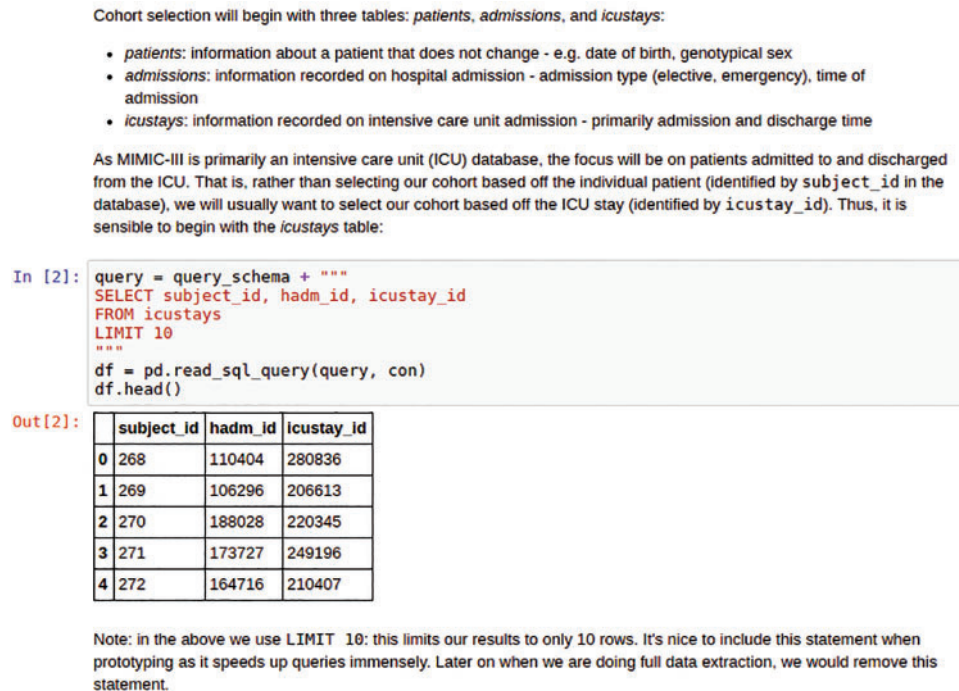


Figure 6. Example of a notebook providing a tutorial with MIMIC-III data.

We have found executable documents particularly valuable for research in cross-disciplinary fields such as health care, because they facilitate collaboration between data analysts and domain experts. Notebooks primarily serve 3 purposes: (1) they allow documentation of the logic behind the code in an organized and easy-to-read manner; (2) they aid rapid writing of the code, particularly during group discussions; and (3) they provide a means of sharing details of a published study that captures the learning that takes place during the evolution of a research project. To encourage sharing of research code, we have reproduced a previously published study on indwelling arterial catheters and their association with in-hospital mortality for hemodynamically stable patients with respiratory failure.<sup>32</sup> This study was initially performed in MIMIC-II, which has since been superseded by MIMIC-III. As the structure of the databases differs, the study was reimplemented based on the manuscript. The executable documents perform data extraction, preprocessing of the data, and construction of a propensity score, and provide an interpretation of the results. Specifically, a Jupyter Notebook, `aline.ipynb`, extracts the study population and necessary data, outputting data to a plain-text file. An R Markdown file, `aline.Rmd`, subsequently loads data from the plain-text file and tests the study hypothesis after matching cohorts with a propensity score. The executable documents provide a template for creating completely reproducible studies using the MIMIC-III database.

Executable documents are also a platform well suited to tutorials. Harmonization of text and code allows for explanations of the subject matter, while the interactive nature of the document allows for experimentation and facilitates learning. A number of tutorials have been made available to explain key concepts important for working with MIMIC. For example, the transformation of recorded clinical parameters, such as hemofiltration settings, into desired clinical concepts, such as length of CRRT, is nontrivial and requires both domain and database expertise. An executable document is provided, which overviews the process of exploring MIMIC-III, assessing the data stored within and creating the definition of CRRT provided in the

database. In addition to explaining the logic behind the definition of CRRT, the tutorial also acts as a template for defining other concepts in the MIMIC database and potentially other similar ICU EHRs. Other tutorials include an introduction to SQL, a step-by-step guide to selecting a study cohort, and an outline of the data-capture process for commonly recorded parameters in the database.

## Community

The MIMIC Code Repository provides many benefits regarding distribution of the source code and enhancing reproducibility, as previously mentioned. An additional advantage is the communication channel opened between the maintainers and distributors of the MIMIC database and users. Longo et al.<sup>33</sup> argue, in a well-publicized editorial on data sharing, that researchers not involved in the collection of data may lack an understanding of its underlying details. Our framework connects researchers who reuse the MIMIC-III dataset with the laboratory and clinical staff who collect and produce the data, helping to provide context for downstream data analysis. Researchers can post issues inquiring about aspects of the data collection and best practices for analyzing the data, and experienced users, some of whom are involved in collecting the data, can provide insight and advice. This correspondence facilitates appropriate and meaningful use of the data, and as all discussions are publicly available, the result is an organically growing set of documentation that spans both narrow and broad topics. Researchers are encouraged to contribute to the MIMIC Code Repository, progressively improving the code base and helping to accelerate research in critical care. Source code control allows for transparency both in the authorship of the code and in the nature of any changes.<sup>34</sup>

## CONCLUSION

Transparent research processes can help to improve the quality of evidence that underpins health care, and the case for open data has been

quite well described.<sup>35,36</sup> To achieve full transparency, researchers must be able to provide both the data used for analysis *and* the code used to process it. The MIMIC-III database is exceptional due in no small part to its publicly accessible nature: all researchers who undergo human subjects research training and who sign a data use agreement can freely access the data. By supplementing the MIMIC-III database with the MIMIC Code Repository, we provide a foundation for completely reproducible research in critical care. Examples of reproducible code and even reproducible studies are available and provide a framework for future work with the database. Future work will build new concepts as well as provide executable documents for other applications, such as predictive modeling.

There are some limitations of this approach. First, use of the repository requires familiarity with technical tools such as git and SQL. Second, the SQL code strongly conforms to the current American National Standards Institute–SQL standard (6th revision) and may need to be adapted for noncompliant database systems. Third, the repository is tailored to the MIMIC-III database, although we anticipate much of the code to be broadly applicable as common data models for critical care development. Finally, the content is not exhaustive and continues to be developed over time.

While cultural barriers exist that may discourage some researchers from sharing code, it is clear that the barriers in the case of MIMIC-III are not technical. The unique combination of open code with publicly accessible data allows for the creation of fully executable studies with diligent audit trails, and it would behoove researchers to adopt these approaches.

## FUNDING

This work has been supported by grants NIH-R01-EB017205 and NIH-R01-EB001659 from the National Institutes of Health.

## AUTHOR CONTRIBUTIONS

AEWJ and TJP collaborated to build the MIMIC Code Repository. All authors contributed to the paper.

## COMPETING INTERESTS

The authors have no competing interests to declare.

## ACKNOWLEDGMENTS

The authors would like to thank Professor Roger G Mark, the Massachusetts Institute of Technology Laboratory for Computational Physiology, Philips Healthcare, and the Beth Israel Deaconess Medical Center for the creation of the MIMIC-III database.

## REFERENCES

- Baker M. 1,500 scientists lift the lid on reproducibility. *Nature*. 2016;533(1):452–54.
- Gruber WH, Powell AC, Torous JB. Healthcare The power of capturing and using information at the point of care. *Healthcare [Internet]* 2016;(January):0–1. <http://dx.doi.org/10.1016/j.hjdsi.2016.09.004>. Accessed May 25, 2017.
- Johnson AEW, Pollard TJ, Shen L, *et al*. MIMIC-III, a freely accessible critical care database. *Scientific Data*. 2016;3:1–9.
- Pollard TJ, Johnson AEW. *The MIMIC-III Clinical Database*. <http://dx.doi.org/10.13026/C2XW26> 2016. Accessed May 25, 2017.
- Pollard TJ, Johnson AEW, Blundell J, *et al*. MIT-LCP/mimic-code: MIMIC-III v1.4. 2017. doi:10.5281/zenodo.821872. Accessed May 25, 2017.
- Wilson G, Aruliah DA, Brown CT, *et al*. Best practices for scientific computing. *PLOS Biol [Internet]*. 2014;12(1):e1001745. <http://dx.doi.org/10.1371/journal.pbio.1001745>. Accessed May 25, 2017.
- Knaus WA. APACHE 1978–2001: the development of a quality assurance system based on prognosis: milestones and personal reflections. *Arch Surg*. 2002;137(1):37–41.
- Knaus WA, Wagner DP, Draper EA, *et al*. The APACHE III prognostic system: Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*. 1991;100(6):1619–36.
- Le Gall J, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993;270(24):2957–63.
- Le Gall JR, Klar J, Lemeshow S, *et al*. The Logistic Organ Dysfunction system. A new way to assess organ dysfunction in the intensive care unit. ICU Scoring Group. *JAMA [Internet]*. 1996;276(10):802–10. <http://www.ncbi.nlm.nih.gov/pubmed/17166357>. Accessed May 25, 2017.
- Johnson AEW, Kramer AA, Clifford GD. A new severity of illness scale using a subset of acute physiology and chronic health evaluation data elements shows comparable predictive accuracy. *Critical Care Med*. 2013;41(7):1711–18.
- Vincent J-L, Moreno R, Takala J, *et al*. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22:707–10.
- Le Gall JR, Klar J, Lemeshow S, *et al*. The Logistic Organ Dysfunction system. A new way to assess organ dysfunction in the intensive care unit. ICU Scoring Group. *JAMA [Internet]*. 1996;276(10):802–10. <http://www.ncbi.nlm.nih.gov/pubmed/17166357>. Accessed May 25, 2017.
- Wiesner R, Edwards E, Freeman R, *et al*. Model for end-stage liver disease (mELD) and allocation of donor livers. *Gastroenterology*. 2003;124(1):91–96.
- Kellum JA, Mehta RL, Angus DC, Palevsky P, Ronco C. The first international consensus conference on continuous renal replacement therapy. *Kidney Int*. 2002;62(5):1855–63.
- Mehta RL, Kellum JA, Shah SV, *et al*. Acute kidney injury network: Report of an initiative to improve outcomes in acute kidney injury. *Critical Care*. 2007;11(2):R31.
- KDIGO CKD-MBD Work Group. KDIGO clinical practice guideline for the diagnosis, evaluation, prevention, and treatment of chronic kidney disease–mineral and bone disorder (CKD-MBD). *Kidney Int Suppl*. 2009;113:S1.
- Ghassemi M, Wu M, Hughes MC, Szolovits P, Doshi-Velez F. Predicting intervention onset in the ICU with switching state space models. In: *Proceedings of the 2017 AMIA Summit on Clinical Research Informatics*. 2017;82–91.
- Torio CM, Andrews RM. National inpatient hospital costs: The most expensive conditions by payer, 2011. hCUP statistical brief #160. Agency for Healthcare Research and Quality, Rockville, MD [Internet]. 2013. <http://www.hcup-us.ahrq.gov/reports/statbriefs/sb160.pdf>. Accessed May 25, 2017.
- Torio CM, Moore BJ. National inpatient hospital costs: The most expensive conditions by payer, 2013. hCUP statistical brief #204. Agency for Healthcare Research and Quality, Rockville, MD [Internet]. 2016. <http://www.hcup-us.ahrq.gov/reports/statbriefs/sb204-Most-Expensive-Hospital-Conditions.pdf>. Accessed May 25, 2017.
- Singer M, Deutschman CS, Seymour CW, *et al*. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA*. 2016;315(8):801–10.
- Seymour CW, Liu VX, Iwashyna TJ, *et al*. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA*. 2016;315(8):762–74.
- Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Critical Care Med*. 2001;29(7):1303–10.
- Martin GS, Mannino DM, Eaton S, Moss M. The epidemiology of sepsis in the United States from 1979 through 2000. *New Engl J Med*. 2003;348(16):1546–54.

25. Iwashyna TJ, Odden A, Rohde J, *et al.* Identifying patients with severe sepsis using administrative claims: patient-level validation of the Angus implementation of the International Consensus Conference definition of severe sepsis. *Med Care.* 2014;52(6):e39.
26. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care.* 1998;36(1):8–27.
27. Steiner C, Elixhauser A, Schnaier J. The healthcare cost and utilization project: an overview. *Effect Clin Pract.* 2001;5(3):143–51.
28. Quan H, Sundararajan V, Halfon P, *et al.* Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care.* 2005;43(11):1130–39.
29. Walraven C van, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Medical Care.* 2009;47(6):626–33.
30. Pérez F, Granger BE. iPython: A system for interactive scientific computing. *Comput Sci Engineering [Internet]* 2007;9(3):21–9. <http://ipython.org>. Accessed May 25, 2017.
31. Kluyver T, Ragan-Kelley B, Pérez F, *et al.* Jupyter Notebooks—a Publishing Format for Reproducible Computational Workflows. *Positioning and Power in Academic Publishing: Players, Agents and Agendas;* 2016:87.
32. Hsu DJ, Feng M, Kothari R, Zhou H, Chen KP, Celi LA. The association between indwelling arterial catheters and mortality in hemodynamically stable patients with respiratory failure: a propensity score analysis. *CHEST J.* 2015;148(6):1470–76.
33. Longo DL, Drazen JM. Data sharing. *N Engl J Med.* 2016;376:276–7.
34. Perez-Riverol Y, Gatto L, Wang R, *et al.* Ten simple rules for taking advantage of git and gitHub. *PLoS Comput Biol [Internet]* 2016;12(7):1–11. <https://doi.org/10.1371/journal.pcbi.1004947>. Accessed May 25, 2017.
35. Ross JS, Krumholz HM. Ushering in a new era of open science through data sharing: the wall must come down. *JAMA.* 2013;309(13):1355–56.
36. Bierer BE, Crosas M, Pierce HH. Data authorship as an incentive to data sharing. *N Engl J Med.* 2017;376:1684–87.