Medical Information Mart for Intensive Care (MIMIC-III): A Real-World Data Foundation for Reproducible Artificial Intelligence Machine Learning Research

Paul B Rogers, Dong Wang and Zhiyuan Lu, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, FDA

Introduction
The Medical Information Mart for Intensive Care (MIMIC) is a database of de-identified electronic health records (EHR) associated with patients who stayed in intensive care units (ICU) at the Beth Israel Deaconess Medical Center in Boston, MA. All four versions of MIMIC (Table 1) are accessible and available to the public, supporting the concept of reproducibility within ICU research.

Conduit for Collaboration
Publicly available critical care datasets like MIMIC provide a collaborative “sandbox” that brings together clinicians, academia, medical technology companies, and the pharmaceutical industry. Currently, MIMIC is studied in over 30 countries, with 4,000 plus users in industry and academia.

MIMIC-III Data Flow, Structure and Versions

TABLE 1 | Versions of MIMIC.
<table>
<thead>
<tr>
<th>Version</th>
<th>Years Spanned</th>
<th>Release Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIMIC-II</td>
<td>2001–2007</td>
<td>February 2010</td>
</tr>
<tr>
<td>MIMIC-IV</td>
<td>2008–2019</td>
<td>AUGUST 2020</td>
</tr>
</tbody>
</table>

*MIMIC-IV was released as an extension to MIMIC-III.*

Background & Summary
In recent years there has been a concerted move towards the adoption of digital health record systems in hospitals. In the US, for example, the number of non-federal acute care hospitals with basic digital systems increased from 9.4 to 75.5% over the 7-year period between 2008 and 2014. In parallel, the scientific research community is increasingly coming under criticism for the lack of reproducibility of studies. MIMIC-III integrates deidentified, comprehensive clinical data of patients and makes it widely accessible to researchers internationally under a data use agreement. The open nature of the data allows clinical studies to be reproduced and improved in ways that would not otherwise be possible.

Literature cited

Acknowledgments
Dong Wang
Julian Lu
Zhichao Liu
Weida Tong
Stephen Turner

Further information
Please see https://physionet.org/for more information. I’m at Paul.Rogers@fda.hhs.gov if you have a question or comment.