NEJM AI

NEJM AI 2026;3(6)
DOI: 10.1056/AIra2501253

REVIEW ARTICLE

MEDS — An Emerging Data Standard and Ecosystem for Health AI Research

Matthew B. A. McDermott, Ph.D.,[1] Ethan Steinberg, Ph.D.,[2,3] Jason A. Fries, Ph.D.,[2,4] Robin P. van de Water, M.S.,[5] Chao Pang, Ph.D.,[1] Patrick Rockenschaub, Ph.D.,[6] Pawel Renc, Ph.D.,[7,8] Jungwoo Oh, Ph.D.,[9] Kamilė Stankevičiūtė, M.Sc., M.A.,[10] Justin Xu,[11] Tom J. Pollard, Ph.D.,[12] Nassim Oufattole, S.M.,[13] Michael Wornow, Ph.D.,[14] Teya S. Bergamaschi, M.S.E.,[13] Hyewon Jeong, M.D., M.S.,[13] Simon A. Lee, M.S.,[15] Vincent Jeanselme, Ph.D.,[1] Kiril V. Klein, Ph.D.,[16] Mikkel Odgaard, Ph.D.,[16] Maria E. Montgomery, M.S.,[16] Arkadiusz Sitek, Ph.D.,[8] Mads Nielsen, Ph.D.,[16] Jeffrey N. Chiang, Ph.D.,[15,17] Noa Dagan, M.D., Ph.D., M.P.H.,[18,19] Isaac Kohane, M.D., Ph.D.,[20] Shalmali Joshi, Ph.D.,[1] Edward Choi, Ph.D.,[8] and Nigam H. Shah, Ph.D., M.B.B.S.[2]

Received: November 13, 2025; Revised: April 6, 2026; Accepted: April 9, 2026; Published: May 28, 2026

Abstract

While data standards have been well adopted and highly impactful for observational health informatics, the emerging application of artificial intelligence (AI) to electronic health record (EHR) data — known broadly as health AI — still lacks broadly adopted data standards. This gap limits reproducibility and the development of open-source ecosystems for health AI, as well as limiting the ability to collaboratively develop and characterize emerging capabilities such as the use of foundation models. The Medical Event Data Standard (MEDS) is an open-source data standard and ecosystem designed and promulgated by us (and others) to address this gap. MEDS was first presented at a workshop in May 2024 and has since been described in online materials and preprints. Here, we describe MEDS in detail and review its use in the community, highlighting its strengths and weaknesses, and the extent of its adoption to date.

We designed MEDS to emphasize simplicity, algorithm transportability, and support for workflows used in training foundation models, and to offer complementary strengths compared with existing health data standards. As of March 2026, we find usage of MEDS across 21 institutions, in at least 27 academic papers and preprints, and to support work with 17 datasets and 12 AI algorithms. Tools in the MEDS ecosystem have reported computational speedups that range from 1.9 to around 40,000 times faster than prior tools or common individual workflows. In novel case study comparisons, we find that codebases leveraging MEDS show reductions in necessary lines of functional Python code by up to 70%. Further, the MEDS standard has supported the development of key frontier models for EHR data. (Supported by Canadian Institutes of Health Research and others.)

Introduction

A data standard is a set of rules that governs how data are formatted, described, and exchanged to ensure consistency and interoperability across users and systems. When broadly adopted, successful data standards are key drivers of reproducible science, open-source ecosystem development, cost-sharing of intensive engineering efforts,

The author affiliations are listed at the end of the article.

Matthew B. A. McDermott can be contacted at mattmcdermott8@gmail.com.

and the "frictionless reproducibility" of findings.[1-4] In observational health informatics, for example, the adoption of data standards such as the Observational Health Data Sciences and Informatics (OHDSI) Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM),[5,6] the PCORnet CDM,[7] and the i2b2 CDM[8] has empowered observational studies in diverse settings.[9,10] In contrast, the emerging field of applying high-parameter artificial intelligence (AI) models to structured electronic health record (EHR) data — known broadly as health AI — still does not have widely adopted data standards. This lack of standardization is a significant barrier to reproducibility, the rigorous evaluation of progress, and the development of shared tools in health AI.[1,2]

We argue that this discrepancy in data standard adoption is in part due to fundamental differences between the key objectives and workflows of studies in observational health informatics and those in health AI. To highlight these differences, consider the following prototypical examples across these two fields.

First, consider the task of real-world evidence generation from observational data — a common, high-impact focus of observational health informatics.[10-12] Such studies aim to produce statistical evidence over targeted patient populations that generalizes across health systems; accordingly, they rely extensively on expert-curated cohorts and features and use interpretable statistical analyses to maximize robustness.

In contrast, for health AI, consider the recent methodologic advances in autoregressive foundation models, such as the Enhanced Transformer for Health Outcome Simulation (ETHOS),[13] ETHOS Adaptive Risk Estimation System,[14] Curiosity,[15] Delphi-2M,[16] and CEHR-GPT,[17] or in embedding models such as CLMBR[18] and MOTOR.[19] These models use the longitudinal nature of structured EHRs to train neural networks with hundreds of millions to billions of parameters to predict, given a patient's historical medical record, when and what the patient's next interaction with the source health system will be. Models are typically trained and evaluated on care patterns within — and often specific to — a single data source; trained over a general-purpose representation of the entire EHR, rather than disease-specific cohorts or manually engineered features; and rarely shareable publicly owing to concerns around the leakage of private information through their parameters.

As argued more extensively in the Supplementary Appendix, these two examples are indicative of key differences between the objectives and workflows of these two fields. In particular, classical observational health informatics studies are often motivated first and foremost by producing causal statistical models that are maximally generalizable across health systems and clinical settings. In contrast, predictive health AI studies are designed to maximize predictive power within a source health system (especially given privacy-related sharing restrictions) and, correspondingly, may be less concerned with cross-site generalizability.[20] As a result, the research community often focuses first on whether the algorithms used to produce these models can be reliably used on new data sources, rather than on whether the pretrained model itself generalizes across sites. This same difference in objectives motivates other workflow differences: Classical observational health informatics relies on deep feature harmonization and careful cohort curation, whereas health AI prioritizes the efficient use of all features across all patients, often with little to no modification or curation.

Ultimately, we argue that these key workflow differences motivate a set of design needs for data standards in health AI — needs that were not prioritized by existing informatics standards. In this work, we review adoption of the Medical Event Data Standard (MEDS), a recently introduced open-source data standard for health AI research that we developed.[21,22] Specifically, we designed MEDS to (1) enable transportability of model training algorithms first and foremost, rather than requiring full cross-site generalizability of trained models; (2) prioritize simplicity in representation by respecting EHR data's fundamental structure — a longitudinal timeline of events — without mandating deep harmonization of clinical vocabularies used; (3) empower development of an open-source Python ecosystem for health AI that offers some of the same advantages that widely used tools and platforms such as Hugging Face[23] or TorchVision[24] provide in other domains of AI; and (4) enable the computational performance necessary to work with near-total, raw clinical datasets at sufficient scale to train foundation models.[25] We further designed MEDS to be complementary to existing data standards such as the OMOP CDM or the Fast Healthcare Interoperability Resources (FHIR)[26] standards via public tools to transform data from these standards into MEDS.[27-29]

In this Review we examine, as of March 2026, the literature that cites any of the existing nonarchival publications describing MEDS, as well as the tools developed within the MEDS ecosystem. We also characterize EHR datasets and health AI models known to us to be MEDS-compliant. Approximately 40% of the reviewed works include our work, reflecting the collaborative, open-source nature of MEDS.

The Medical Event Data Standard

For full technical details, see the Supplementary Appendix and the MEDS technical documentation at https://medical-event-data-standard.github.io/.

KEY DESIGN PRINCIPLES OF MEDS

In this section, we expand upon the initial design principles presented in the workshop paper that first introduced MEDS[22] and highlight how they relate to the unique differences in field objective and workflow described previously.

Transportability of Model Training Algorithms

The primary objective of our work with MEDS is to reduce barriers to the transportability of algorithms used in model training, enabling frictionless reproduction of training recipes across institutions. Emphasizing portability allows us to make MEDS a streamlined data standard that does not assume or require prior data harmonization, aiding accessibility across datasets. This is a major, intentional departure from existing data standards. For example, all existing major observational health informatics standards assume a fundamental goal of transportability of trained statistical models, rather than focusing separately on transportability of the recipe to train a model on a dataset. The former goal requires a significantly more complex data standard to facilitate deep feature harmonization and cohort control. For example, the OHDSI–OMOP CDM specifies approximately 400 fields across 40 distinct tables; in contrast, MEDS requires only three fields with several optional fields over a single core data schema. Further, we note that this focus does not preclude pretrained model transfer; in cases where datasets are externally harmonized (e.g., when MEDS datasets are produced from existing OMOP datasets), this focus on algorithm transportability is intended to enable easy evaluation of pretrained models across institutions.

Adherence to the Fundamental Structure of Health Data

The fundamental shared structure of EHR data is that of a longitudinal, continuous time sequence of complex events (Fig. 1A). Therefore, we posited that these data should be represented in as simple a manner as possible while remaining true to that underlying structure.[30] Doing so allowed us to ensure that (1) MEDS reflects the core structural element of EHR data that is shared across health sites; (2) MEDS captures the aspects of data that differentiate it mathematically from other types of data on which model training is well explored, such as ordinal sequences, regularly sampled high-frequency time series, or fixed-size tabular data. While some existing standards conform in part to this principle (e.g., the i2b2 standard leverages an event stream–like view[8]), they also often enforce additional structural constraints to facilitate deep feature control and prioritize population- and cohort-level analyses (in accordance with typical informatics workflows), which often motivate structuring data via relational databases. In contrast, the focus of high-capacity AI on individual-patient prediction tasks means a health AI–focused standard can conform to this fundamental structure in a simpler, nonrelational manner.

Empower Development of an Open-Source Python Ecosystem for Health AI

As noted previously, we designed MEDS to empower the development of an accessible open-source Python ecosystem (Fig. 1C) for health AI research, spanning the triad of tools, models, and datasets that has been shown to be extremely impactful in other areas of AI use, such as sharing via Hugging Face[23] and TorchVision.[24] In particular, MEDS is a "data as interface" standard, meaning that MEDS constrains the way data are encoded on disk but does not mandate the use of a particular package or closed framework. This is designed to ensure that the open-source community can benefit from MEDS-compliant datasets across numerous platforms and packages, in the hopes that this flexibility will reduce barriers to communal ecosystem growth. Further, whereas many existing data standards prioritize usage with R, the development of core tools in Python and in line with Python development best practices helps facilitate adoption within the core computer science and machine learning and AI communities.

Support Computational Performance for Foundation Models

We designed MEDS to be in alignment with the computing and engineering needs of foundation models, which typically have hundreds of millions to billions of parameters and are usually trained on specialized hardware such as graphics processing units or tensor processing units.[31] Specifically, whereas many existing standards leverage relational database formats as their underlying storage backends, in line with our patient-centric view and prior design principles for MEDS, we can leverage file storage formats focused on parallelism-based scalability (e.g., MapReduce[32]), rather than requiring maximally efficient support for relational database workflows that are harder to parallelize.

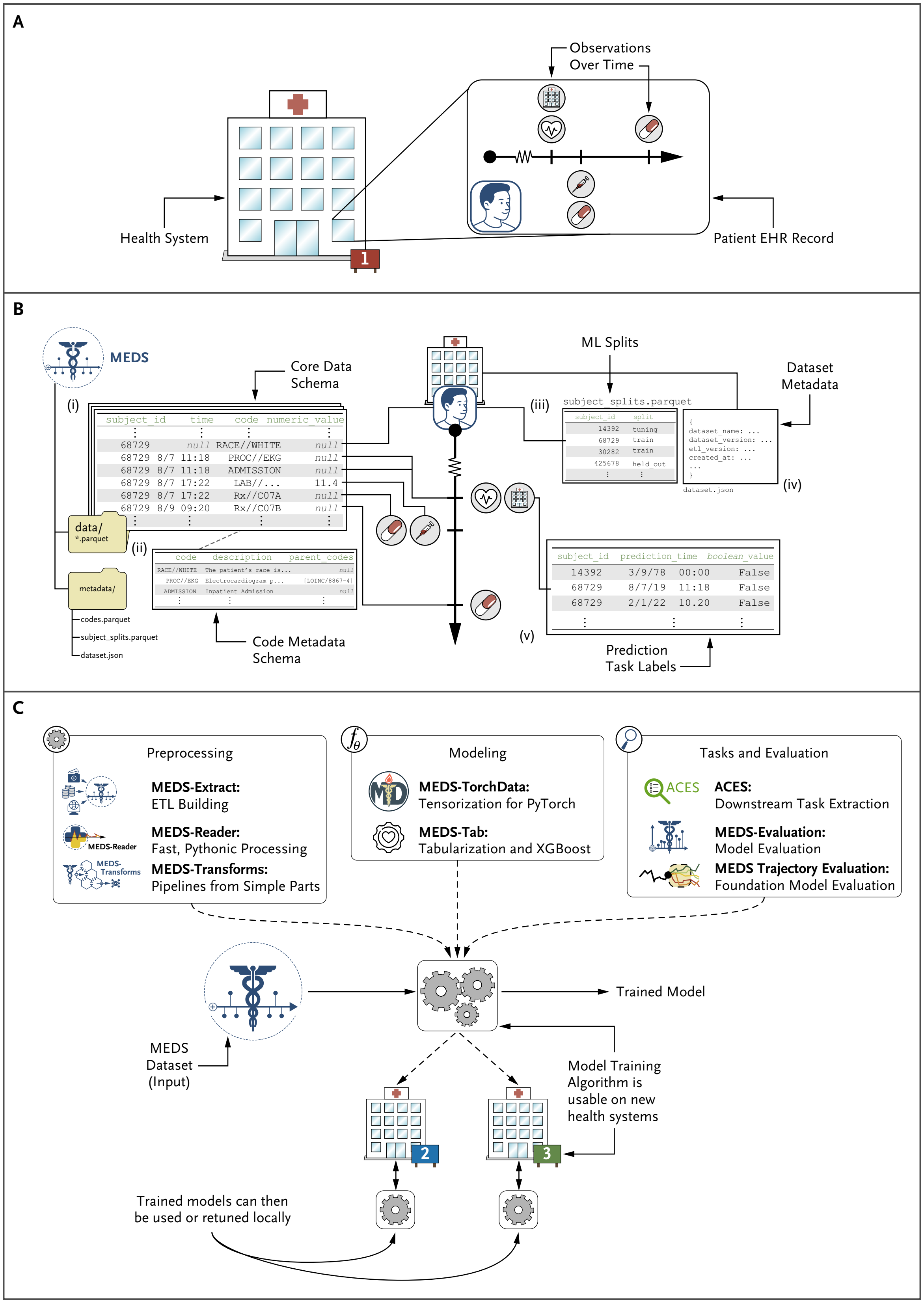

Figure 1. (*Continued*)

MEDS SCHEMA AND FILE STRUCTURES

Any data standard can be seen as a set of rules governing how data are formatted, described, and exchanged. In the case of MEDS, these rules are organized in the form of specific file schema and organization patterns for EHR data, as shown in Figure 1. We describe each of these rules here, with further details provided in the Supplementary Appendix.

The MEDS Data Schema

The MEDS data schema (Fig. 1Bi) has three required columns: a **subject_id** identifying to which patient the row corresponds, a **time** entry identifying when the measurement in the row took place, and a **code** column providing a categorical descriptor (from an unconstrained vocabulary) of the measurement contained in the row. Note that some clinical observations can be associated with multiple time stamps (e.g., a blood test is ordered at one point in time but the actual result of the test is only collected and entered into the patient record later). In general, to best reflect the longitudinal structure of the data, we recommend that such measurements be split into separate observations, each captured at their respective time and level of information (e.g., one row capturing the blood test order at a point in time followed by a separate row capturing the addition of the test result in the patient record, at the time of data storage and accompanied by the test result itself). If this is not possible, observations should generally default to the latest associated time stamp to minimize any data leakage, though this can risk other biases downstream.

This schema permits two additional optional columns — **numeric_value** and **text_value** — for numeric or textual values associated with the measurement. To maximize flexibility, other dataset-specific columns can be added, though most downstream tools or models will not leverage them. To ensure use at very large data scales, MEDS requires that compliant data files be organized on disk as a subject-sharded set of columnar parquet files, encouraging MapReduce-style parallel processing.[32] Note that this schema has fewer fields than all alternative health informatics schemas, most of which feature up to hundreds of distinct fields across many tables (e.g., the OMOP schema has ∼400 fields across ∼40 tables).[6-8,26]

The MEDS Code Metadata Schema

MEDS datasets accept metadata about the (unconstrained) vocabulary of **code** elements through a **metadata/codes.parquet** file (Fig. 1Bii). This file links **code** strings to free-text, human-readable descriptions and to identifiers from external ontologies. Doing so enables MEDS datasets sourced from fully or partially harmonized source datasets to benefit from the prior harmonization. Further, these descriptions can also empower large language model (LLM) applications over MEDS datasets, giving a natural textual interface for generative models over EHR data when site-specific code descriptions are available.

The MEDS Task Splits, Dataset Metadata, and Labels Schemas

MEDS datasets can also present a default subject-level split for the data before any model training occurs through **metadata/subject_splits.parquet** (Fig. 1Biii) and a **metadata/dataset.json** schema (Fig. 1Biv). Doing so is necessary to ensure data provenance for the end user and to prevent leakage of information that can inflate performance estimates of classifiers and predictors trained using EHR data. Within MEDS, we also define a formal schema by which downstream task labels and cohort membership can be specified (Fig. 1Bv). This ensures that task definitions can be created independently of model training pipelines and that task selection and extraction from source data can be automated within the MEDS ecosystem via dedicated tools.[33]

Figure 1. A Visual Overview of the Design Principles and Schema of MEDS.

Medical Event Data Standard (MEDS)–compliant datasets represent electronic health record data as a timeline consisting of a longitudinal sequence of events (Panel A). Observations on those timelines can be directly mapped to elements of the MEDS core data schema, which records subject identifiers (IDs), time stamps, and categorical codes in an unconstrained vocabulary (Panel B, i). Codes are further described via a metadata file that links to external ontologies (Panel B, ii). MEDS-compliant datasets also include (Panel B, iii) subject-level splits for artificial intelligence model development and (Panel B, iv) dataset metadata in JSON form for provenance tracking. MEDS-compliant datasets can be accompanied by label files defined by a subject ID, prediction time, and label that define a task to permit easy development of predictive models that can be run over arbitrary MEDS-compliant datasets (Panel B, v). Algorithms (i.e., workflows for training models) for MEDS-compliant datasets can benefit from the open-source ecosystem of MEDS-compatible tools and can be applied to other MEDS-compliant datasets (Panel C). ACES denotes Automatic Cohort Extraction System; EHR, electronic health record; EKG, electrocardiogram; ETL, extract, transform, load; ID, identifier; JSON, JavaScript Object Notation; MEDS, Medical Event Data Standard; and ML, machine learning.

Review and Case Study Results

METHODOLOGY

Analysis of Published Literature, Preprints, and the MEDS Ecosystem

To assess the usage of MEDS by the broader community, both within and beyond those usages associated with the original nonarchival MEDS report, we undertook a manual literature search through various sources, including articles that had cited either the initial nonarchival MEDS workshop description,[22] several core MEDS tools,[34-36] or the MEDS KDD 2025 Tutorial[37]; articles built atop code that was known to use MEDS internally; or works reported to us from the broader MEDS community. This literature search may not be exhaustive (e.g., there may be even more uses of MEDS in the health AI research space), but reported uses of MEDS were manually validated, so this set is a valid lower bound for the current adoption of MEDS in the community. All statistics were collected no later than March 2026.

Case Study Methodology

To examine how the use of MEDS impacts researcher code efficiency, we conducted a curated set of case studies, comparing public project repositories for projects produced using MEDS with those produced without using MEDS. We examined eight repositories (four using MEDS, four not), selected manually by us for these comparisons (the full set is summarized in Table S1), spanning tasks including data extraction from source, benchmarking multiple AI models, and autoregressive foundation models. Repositories were extracted at a particular, frozen commit identifier for reproducibility, and lines of code were counted using the tool **cloc**.[38] The number of lines of functional code (excluding documentation or tests) for repositories using MEDS and not using MEDS was then compared to assess possible efficiency gains due to the use of MEDS. While any comparison of this sort performed in a post hoc manner on repositories aiming to perform nonidentical tasks will always have many sources of bias, we feel these comparisons are still meaningful and we provide further commentary in Section C.3 in the Supplementary Appendix.

Potential Conflict of Interest Disclosure

As this review is authored by the creators of MEDS, who are, naturally, among the most active members of the MEDS community, a subset (~40%) of the works we characterize here have overlapping author lists with the authors of this review, risking a possible conflict of interest. When considering this risk, it is important first to note that, in fact, many of the authors of this work became core members of the MEDS community after or upon beginning to contribute to the MEDS ecosystem. Testament to this fact is that, of the individually highlighted authors on this work, approximately 70% were not authors of the workshop paper that initially highlighted MEDS,[22] demonstrating that the core MEDS community has seen significant growth from its initial development. Nevertheless, to address this potential bias, we highlight in our results below, and in our detailed methodology section (Section C in the Supplementary Appendix), which of the works referenced have overlapping author pools with this review.

REVIEW FINDINGS

Adoption

As of March 2026, MEDS has been used by researchers spanning at least 21 institutions across North America, Europe, and Asia. It has been referenced or used in 27 academic articles or preprints[14,33-36,39-60] (of which ~59% were produced by researchers outside the group of core MEDS developers) and supports 17 existing datasets, including 7 publicly available EHR intensive care unit (ICU) datasets,[61-67] datasets in OMOP CDM and FHIR formats,[27-29] and various private datasets used internally by members of the MEDS community. Twelve health AI algorithms have been developed and released atop MEDS, including tabular models,[35,68] supervised predictive neural network models,[69] classical pretraining and fine-tuning representation learners,[18,19,48,70,71] and autoregressive foundation models.[14,17,39,72] Finally, at least 14 tools have been developed to support health AI needs in the MEDS ecosystem, including tools that leverage LLMs to process EHR data through linearization strategies,[43] a model context protocol server implementation,[73] and visualization tools for MEDS datasets.[74] Some tools in this ecosystem highlight the ability to replace common steps with externally configurable prebuilt tools, such as replacing custom task-extraction code with Automatic Cohort Extraction System configuration files[33] or tensorization pipelines via MEDS TorchData.[75] See Figure 2 for a summary of MEDS adoption, ecosystem scale, and performance improvements.

Previously Reported Computational Speedups

Works we review strongly suggest that tools in the MEDS ecosystem can enable significant speedups in comparison

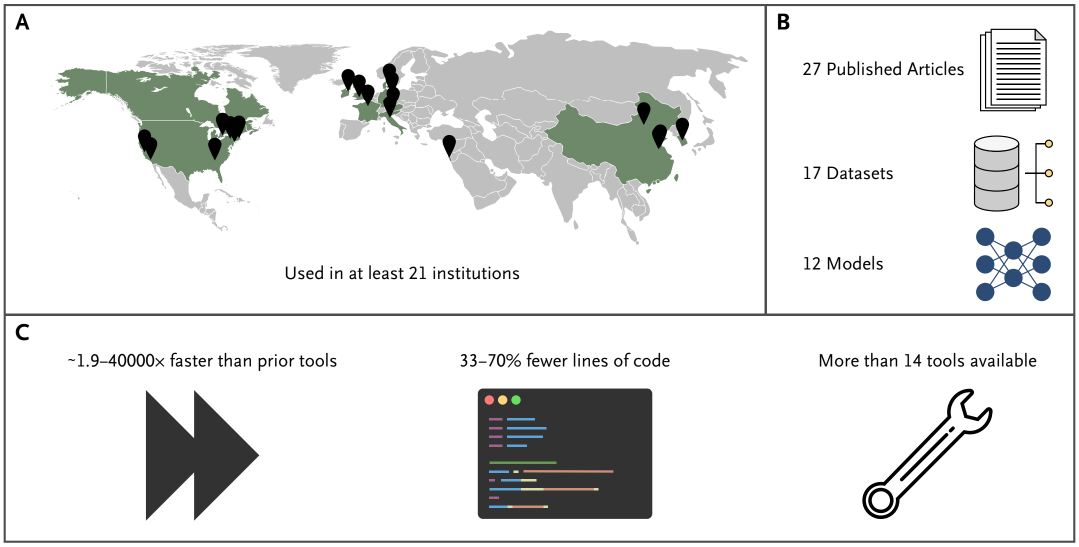

Figure 2. MEDS Usage.

The Medical Event Data Standard (MEDS) has been used by researchers in at least 21 institutions across 12 countries in North America, Europe, and Asia (Panel A). MEDS has been used in projects spanning 27 academic papers or preprints and supports 17 existing datasets or dataset formats, and at least 12 published model training recipes exist over the MEDS framework (Panel B). The more than 14 tools publicly available in the MEDS ecosystem enable workflows to be written with (in cases) up to 70% fewer lines of functional code and run orders of magnitude faster than prior research code (Panel C). Vide infra and refer to the Supplementary Appendix for details.

with existing tools. Steinberg et al. report that **meds_reader** offers reductions in computational speed ranging from 1.9× up to around 40,000× in comparison with other tools for general preprocessing applications,[42] and Oufattole et al. report that the tool MEDS-Tab reduces central processing unit memory and runtime use by factors ranging from 7× to 300× in comparison with competing tabularization tools.[35]

Transportability

Our review also highlights qualitative differences in the accessibility and styles of benchmarking over EHR data in the use of MEDS. For example, the FoMoH work provides broad comparisons of high-capacity models through the MEDS framework, and the MEDS Decentralized, Extensible Validation (MEDS-DEV) system presents an infrastructure for continuous model benchmarking through MEDS.[34,41] Works in the emerging area of generative autoregressive foundation models over EHR data have also made consistent use of MEDS.[14,39,76]

CASE STUDY FINDINGS ON CODE EFFICIENCY

Our analyses on the effect of using MEDS on researcher code efficiency suggest that use of MEDS reduces the required lines of code necessary to perform specific steps in training health AI models. When examining existing repositories corresponding to similar workflows with or without MEDS, we see reductions in necessary lines of code ranging from 33% to 70% (median 54.9%; see the Supplementary Appendix for full details).

Discussion

MEDS LIMITATIONS AND TARGET USAGE PATTERNS

We designed MEDS to be a computational representation of medical events at the individual-patient level, enabling efficient use of patient-level EHR data for AI development and research within or across multiple health systems. Under appropriate evaluation practices and within the bounds of the training distribution, such AI systems could be eventually used in deployment for clinical decision support or hospital operational workflows. It is important to emphasize that we designed MEDS for high-capacity predictive AI workflows — rather than traditional epidemiologic or other causal statistical analyses, which are often better supported by other standards such as the OHDSI–OMOP CDM. Without additional work, predictive algorithms over MEDS should be assumed to be associative in nature. In addition, given its focus on patient-level modeling, MEDS-compliant datasets may not be structured appropriately for use in regulatory evaluation of drugs and devices.

MEDS ROAD MAP

MEDS has several high-priority areas of future development, including (1) expanding support for additional data modalities, (2) strengthening integration with existing health standards or EHR software, (3) enhancing the ability to distinguish between providers or sites of care within MEDS datasets, and (4) improving support for enriching MEDS dataset metadata through external ontologies. Addressing these limitations will enable broader use of MEDS in AI workflows, especially as compute and multimodal capabilities in AI continue to accelerate. In parallel, a key priority is to use the reproducibility of MEDS to empower decentralized benchmarking studies within health AI across diverse tasks and datasets. This effort is already under way through the MEDS-DEV platform.[34]

ADDITIONAL RELATED WORK

Other Data Standards for Health

MEDS complements a rich landscape of many existing standards, such as the OHDSI–OMOP CDM,[6] the i2b2 standard, the PCORnet CDM, and the Health Level 7 FHIR standard.[26] These standards each optimize for different workflows and community needs in comparison with MEDS. The FHIR standard, for example, is designed for health data exchange rather than data analysis or informatics, and thus is not a true alternative to analytics. The remaining standards are more natural alternatives in the context of data analysis. In particular, the OMOP, i2b2, and PCORnet CDMs are all designed to facilitate classical observational studies, querying, cohort building, and evidence generation over health datasets. Given their focus, greater emphasis is placed on reliably harmonizing data elements to external ontologies. MEDS remains agnostic to data harmonization, instead prioritizing the minimal amount of standardization necessary to enable code transportability, which allows it to provide a simpler interface at the expense of reduced support for classical workflows that are adequately supported by other standards. MEDS integrates naturally with existing standards, and dedicated tools exist to transform several of these standards (OMOP and FHIR, with i2b2 support in development) into MEDS,[27-29] allowing users in those ecosystems to benefit from the advantages of MEDS in the context of health AI research without sacrificing the utility offered by these standards in other areas.

Existing Health AI Frameworks

The space of health AI has a number of published and unpublished computational frameworks for working with data and building models under different paradigms.[77-82] In our opinion, MEDS has several key advantages over these frameworks. First, many of these frameworks are specific to the Medical Information Mart for Intensive Care (MIMIC) datasets, or to a small collection of public ICU datasets; in contrast, MEDS is a general data standard designed to be usable across any EHR source, public or private. We believe that this makes MEDS much more translatable for general research use. In particular, while MIMIC has been essential for the development of health AI, many major advancements have happened primarily over private, internal, non-ICU datasets, and accordingly data standards specific to MIMIC will not support research on effective models for non-ICU EHR data. Ensuring MEDS can support general EHR data will, therefore, help ensure that developments pursued across private or non-ICU sources can still be assessed and used by the community on other datasets. Second, these frameworks are, in general, closed ecosystems designed to be leveraged through a Python package–based interface. In contrast, we think MEDS's decoupled data-as-interface design and growing open-source ecosystem make it much more extensible to researchers' needs and responsive to the changing AI landscape.

No-Schema Generative AI Approaches

In addition to existing health AI frameworks, the rise of LLMs and generative AI tools has led to the development of new, fully text-driven data interactivity tools. These typically use LLMs to explore the structure of a dataset in light of a user-provided query to enable dynamic data extraction or query response without ever requiring input data standardization or harmonization.[83,84] Preliminary work in this direction, such as text-vectorization packages and model context protocol server code for MEDS datasets, is in development by the community.[73,85] While these tools enable data exploration, they are not a complete solution in their own right for training new models, sharing training pipelines, or benchmarking performance of models.

CONCLUSION

In this review, we examine work leveraging MEDS, an emerging open-source data standard and ecosystem for high-capacity AI over EHR data. We designed MEDS to prioritize simplicity and utility for supporting such health AI workflows and to complement existing standards to meet the needs of this growing use case. Our review finds usage of MEDS across 21 institutions across North America, Europe, and Asia, and that more than 12 model training algorithms have been developed and released atop the MEDS framework, and the emerging MEDS ecosystem features at least 14 open-source tools. Examining the reported improvements offered by these tools suggests that they may offer improvements in computational efficiency and memory usage of up to two orders of magnitude over prior tools or bespoke solutions, and novel case studies conducted in this work show that algorithms developed in MEDS can use up to 70% less code than equivalent approaches in other data formats. Notably, a number of reports exploring new autoregressive foundation models for EHR data have leveraged MEDS as their source data format.[14,39,40] Ultimately, while these initial results are promising, future adoption of MEDS remains uncertain, though we believe that continued work and engagement by the health AI community in efforts toward standardization and reproducibility, such as MEDS, will be of significant long-term benefit to our field.

Disclosures

Author disclosures and other supplementary materials are available at ai.nejm.org.

Artificial intelligence–assisted technologies were used with author oversight in the production of this article, in particular for language editing and limited figure drafting and generation.

Various authors report sources of support that enabled them to pursue this work: Dr. McDermott gratefully acknowledges support from a Berkowitz Postdoctoral Fellowship, of the Ivan and Francesca Berkowitz Family Living Laboratory Collaboration at Harvard Medical School and Clalit Research Institute. Mr. Xu acknowledges support from the Canadian Institutes of Health Research and Oxford University Press. Ms. Stankevičiūtė acknowledges support from AstraZeneca. Mr. Lee acknowledges support from the Warren Alpert Foundation. Dr. Fries and Dr. Shah acknowledge support for this effort from the Debra and Mark Leslie Fund for Supporting AI in Healthcare Research. Dr. Pollard is supported by the National Institutes of Health (NIH) (grant nos. NIH OT2OD032701, NIH U24EB037545, and NIH R01EB030362). Dr. Choi is supported by the Institute for Information and Technology Planning and Evaluation grant (grant no. RS-2019-II190075) and the National Research Foundation of Korea grant (grant no. NRF-2020H1D3A2A03100945) funded by the Korean government. Dr. van de Water is funded by the European Commission in the Horizon 2020 project INTERVENE (grant agreement ID 101016775). Dr. Klein is supported in part by the Pharmacovigilance by AI Real-Time Analysis project supported by Innovation Fund Denmark. Dr. Nielsen acknowledges support from the Pioneer Centre for Artificial Intelligence, Danish National Research Foundation Grant P1.

The Medical Event Data Standard (MEDS) effort benefited tremendously from the contributions of all members of the MEDS community who offered time, insight, and development work. All members of the core MEDS community, beyond those highlighted as individual contributors, should be seen as authors of this article, as stated in the author line. A full list of such contributors, with information on the manner of their contribution, is provided in the Supplementary Appendix.

Author Affiliations

[1] Department of Biomedical Informatics, Columbia University, New York, NY, USA

[2] Center for Biomedical Informatics Research, Stanford University, Stanford, CA, USA

[3] Prealize Health, Palo Alto, CA, USA

[4] Department of Biomedical Data Science, Stanford University, Stanford, CA, USA

[5] Hasso Plattner Institute, University of Potsdam, Potsdam, Germany

[6] Institute of Clinical Epidemiology, Public Health, Health Economics, Medical Statistics and Informatics, Medical University of Innsbruck, Innsbruck, Austria

[7] AGH University of Kraków, Kraków, Poland

[8] Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston

[9] Kim Jaechul Graduate School of AI, Korea Advanced Institute of Science and Technology, Seoul

[10] Department of Computer Science and Technology, University of Cambridge, Cambridge, UK

[11] Big Data Institute, University of Oxford, Oxford, UK

[12] Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, USA

[13] Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, USA

[14] Department of Computer Science, Stanford University, Stanford, CA, USA

[15] Department of Computational Medicine, University of California, Los Angeles, Los Angeles, USA

[16] Department of Computer Science, University of Copenhagen, Copenhagen, Denmark

[17] Department of Neurosurgery, University of California, Los Angeles, Los Angeles, USA

[18] Clalit Research Institute, Clalit Health Services, Faculty of Computer and Information Science, Ramat-Gan, Israel

[19] Ben-Gurion University of the Negev, Beersheba, Israel

[20] Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA

References

1. McDermott M. The (lack of?) science of health AI. In: Proceedings of the 4th Machine Learning for Health Symposium. Vol 259. Vancouver, BC: PMLR, 2025:19-29 (https://proceedings.mlr.press/v259/mcdermott25a.html).

2. Donoho D. Data science at the singularity. Harv Data Sci Rev 2024;6. DOI: 10.1162/99608f92.b91339ef.

3. Gliklich RE, Leavy MB. Data standards. In: Gliklich RE, Leavy MB, Dreyer NA, eds. Tools and technologies for registry interoperability, registries for evaluating patient outcomes: a user's guide, 3rd edition, addendum 2. Rockville, MD: Agency for Healthcare

Research and Quality, 2019 (https://www.ncbi.nlm.nih.gov/books/NBK551886/).

4. Institute of Medicine (U.S.) Committee on Data Standards for Patient Safety. Health care data standards. In: Aspden P, Corrigan JM, Wolcott J, Erickson SM, eds. Patient safety: achieving a New standard for care. Washington, DC: National Academies Press, 2004 (https://www.ncbi.nlm.nih.gov/books/NBK216088/).

5. Hripcsak G, Duke JD, Shah NH, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. Stud Health Technol Inform 2015;216:574-578. DOI: 10.3233/978-1-61499-564-7-574.

6. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. J Am Med Inform Assoc 2012;19:54-60. DOI: 10.1136/amiajnl-2011-000376.

7. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc 2014;21:578-582. DOI: 10.1136/amiajnl-2014-002747.

8. Partners Healthcare System. Informatics for integrating biology to the bedside. i2b2, 2026 (https://www.i2b2.org/index.html).

9. Callahan A, Gombar S, Cahan EM, et al. Using aggregate patient data at the bedside via an on-demand consultation service. NEJM Catal Innov Care Deliv 2021;2(10). DOI: 10.1056/cat.21.0224.

10. Suchard MA, Schuemie MJ, Krumholz HM, et al. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. Lancet 2019;394:1816-1826. DOI: 10.1016/S0140-6736(19)32317-7.

11. Vashisht R, Jung K, Schuler A, et al. Association of hemoglobin A1c levels with use of sulfonylureas, dipeptidyl peptidase 4 inhibitors, and thiazolidinediones in patients with type 2 diabetes treated with metformin: analysis from the Observational Health Data Sciences and Informatics initiative. JAMA Netw Open 2018;1:e181755. DOI: 10.1001/jamanetworkopen.2018.1755.

12. Khera R, Dhingra LS, Aminorroaya A, et al. Multinational patterns of second line antihyperglycaemic drug initiation across cardiovascular risk groups: federated pharmacoepidemiological evaluation in LEGEND-T2DM. BMJ Med 2023;2:e000651. DOI: 10.1136/bmjmed-2023-000651.

13. Renc P, Jia Y, Samir AE, et al. Zero shot health trajectory prediction using transformer. NPJ Digit Med 2024;7:256. DOI: 10.1038/s41746-024-01235-0.

14. Renc P, Grzeszczyk MK, Oufattole N, et al. Foundation model of electronic medical records for adaptive risk estimation. Gigascience 2025;14:giaf107. DOI: 10.1093/gigascience/giaf107.

15. Waxler S, Blazek P, White D, et al. Generative medical event models improve with scale. November 7, 2025 (http://arxiv.org/abs/2508.12104). Preprint.

16. Shmatko A, Jung AW, Gaurav K, et al. Learning the natural history of human disease with generative transformers. Nature 2025;647:248-256. DOI: 10.1038/s41586-025-09529-3.

17. Pang C, Jiang X, Pavinkurve NP, et al. CEHR-GPT: generating electronic health records with chronological patient timelines. May 6, 2024 (http://arxiv.org/abs/2402.04400). Preprint.

18. Steinberg E, Jung K, Fries JA, Corbin CK, Pfohl SR, Shah NH. Language models are an effective representation learning technique for electronic health record data. J Biomed Inform 2021;113:103637. DOI: 10.1016/j.jbi.2020.103637.

19. Steinberg E, Fries JA, Xu Y, Shah N. MOTOR: a time-to-event foundation model for structured medical records. In: The Twelfth International Conference on Learning Representations. Vienna, Austria: ICLR, 2024 (https://openreview.net/forum?id=NialiwI2V6).

20. Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. Lancet Digit Health 2020;2:e489-e492. DOI: 10.1016/S2589-7500(20)30186-2.

21. McDermott MBA, Xu J, Bergamaschi TS, et al. MEDS: building models and tools in a reproducible health AI ecosystem. In: Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V2. New York, NY: ACM, 2025:6243-6244 (https://dl.acm.org/doi/10.1145/3711896.3737608).

22. MEDS Working Group, Arnrich B, Choi E, et al. Medical Event Data Standard (MEDS): facilitating machine learning for health. ICLR 2024 Workshop TS4H. Vienna, Austria: ICLR, 2024 (https://openreview.net/forum?id=IsHy2ebjIG).

23. Wolf T, Debut L, Sanh V, et al. HuggingFace's transformers: state-of-the-art natural language processing. July 14, 2020 (http://arxiv.org/abs/1910.03771). Preprint.

24. Marcel S, Rodriguez Y. Torchvision the machine-vision package of torch. In: Proceedings of the International Conference on Multimedia — MM '10. New York, NY: ACM Press, 2010:1485-1488. DOI: 10.1145/1873951.1874254.

25. Bommasani R, Hudson DA, Adeli E, et al. On the opportunities and risks of foundation models. July 12, 2022 (http://arxiv.org/abs/2108.07258). Preprint.

26. HL7 International. Fast healthcare interoperability resources (FHIR). Release 5. March 26, 2023 (http://hl7.org/fhir/).

27. van de Water R. OMOP_MEDS: an ETL pipeline for transforming OMOP datasets into the MEDS format using the MEDS-Transforms library. Github, April 3, 2026 (https://github.com/rvandewater/OMOP_MEDS).

28. Github. meds_etl: a collection of ETLs from common data formats to Medical Event Data Standard. August 4, 2025 (https://github.com/Medical-Event-Data-Standard/meds_etl).

29. van de Water R. FHIR2MEDS ETL (MEDS on FHIR). Github. July 21, 2025. DOI: 10.5281/zenodo.16269018.

30. Matheny M, Israni ST, Ahmed M, Whicher D, eds. Artificial intelligence in health care: the hope, the hype, the promise, the peril. Washington, DC: National Academy of Medicine, 2019. DOI: 10.17226/27111.

31. Silvano C, Ielmini D, Ferrandi F, et al. A survey on deep learning hardware accelerators for heterogeneous HPC platforms. ACM Comput Surv 2025;57:1-39. DOI: 10.1145/3729215.

32. Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters. Commun ACM 2008;51:107-113. DOI: 10.1145/1327452.1327492.

33. Xu J, Gallifant J, Johnson AEW, McDermott MBA. ACES: automatic cohort extraction system for event-stream datasets. In: Proceedings of the International Conference on Learning Representations. Singapore: ICLR, 2025 (https://openreview.net/forum?id=P4XmKjXTrM).

34. Kolo A, Pang C, Choi E, et al. MEDS decentralized, extensible validation (MEDS-DEV) benchmark: establishing reproducibility and comparability in ML for health. ML4H 2024 Demo Track, 2024 (https://openreview.net/pdf?id=DExp3tRRel).

35. Oufattole N, Bergamaschi T, Kolo A, et al. MEDS-Tab: automated tabularization and baseline methods for MEDS datasets. October 31, 2024 (http://arxiv.org/abs/2411.00200). Preprint.

36. Oufattole N, Bergamaschi T, Renc P, Kolo A, McDermott MBA, Stultz C. MEDS-torch: an ML pipeline for inductive experiments for EHR medical foundation models. In: NeurIPS Workshop on Time Series in the Age of Large Models. Vancouver, BC: NeurIPS, 2024 (https://openreview.net/forum?id=uwnK2LZtjX).

37. McDermott MBA, Xu J, Bergamaschi TS, Jeong H, Lee SA, Oufattole N, et al. MEDS: Building models and tools in a reproducible health AI ecosystem. In: Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V2. New York, NY, USA: ACM; 2025. pp. 6243-6244. DOI: 10.1145/3711896.3737608.

38. AlDanial, Snel S, Jolkdarr, et al. AlDanial/cloc: 1.92. Zenodo. December 6, 2021. DOI: 10.5281/ZENODO.5760077

39. Pang C, Park J, Jiang X, et al. CEHR-XGPT: a scalable multi-task foundation model for electronic health records. September 5, 2025 (http://arxiv.org/abs/2509.03643). Preprint.

40. Fallahpour A, Alinoori M, Ye W, Cao X, Afkanpour A, Krishnan A. EHRMamba: towards generalizable and scalable foundation models for electronic health records. In: Hegselmann S, Zhou H, Healey E, Chang T, Ellington C, Mhasawade V, Tonekaboni S, Argaw P, Zhang H, eds. Proceedings of the 4th Machine Learning for Health Symposium. Vol 259. Singapore: PMLR, 2025:291-307 (https://proceedings.mlr.press/v259/fallahpour25a.html).

41. Pang C, Jeanselme V, Choi YS, et al. FoMoH: a clinically meaningful foundation model evaluation for structured electronic health records. June 16, 2025 (http://arxiv.org/abs/2505.16941). Preprint.

42. Steinberg E, Wornow M, Bedi S, Fries J, McDermott MBA, Shah NH. Meds_reader: a fast and efficient EHR processing library. In: Machine Learning for Health Symposium (Findings Track). Vancouver, BC: ML4H, 2024 (https://arxiv.org/abs/2409.09095).

43. Hegselmann S, von Arnim G, Rheude T, et al. Large language models are powerful electronic health record encoders. April 13, 2025 (http://arxiv.org/abs/2502.17403). Preprint.

44. Zhang Y, Li S. ChronoFormer: time-aware transformer architectures for structured clinical event modeling. April 10, 2025 (http://arxiv.org/abs/2504.07373). Preprint.

45. Pham KM, Mai TT, Dao TTP, Tran M-T, Crane M, Bezbradica M. Zero-shot detection of nosocomial infections using clinical foundation models on structured electronic patient records. June 18, 2025 (https://doi.org/10.36227/techrxiv.175021755.55836808/v1). Preprint.

46. Lee SA, Jain S, Chen A, et al. Clinical decision support using pseudo-notes from multiple streams of EHR data. NPJ Digit Med 2025;8:394. DOI: 10.1038/s41746-025-01777-x.

47. Montgomery ME, Klein K, Odgaard M, Lorenzen SS, Sobhaninia Z. BONSAI: a framework for processing and analysing electronic health records (EHR) data using transformer-based models. J Open Source Softw 2025;10:8869. DOI: 10.21105/joss.08869.

48. Zhu M, Luo Z, Liu Y, Zhu T. MedTPE: compressing long EHR sequence for LLM-based clinical prediction with token-pair encoding. Presented at The 11th Mining and Learning from Time Series Workshop. Toronto, ON: MILETS, August 3-7, 2025 (https://kdd-milets.github.io/milets2025/papers/MILETS_2025_paper_16.pdf).

49. Fani R, Attrach RA, Restrepo D, Jia Y, Celi LA, Schüffler P. Coefficient of variation masking: a volatility-aware strategy for EHR foundation models. In: Proceedings of the 5th Machine Learning for Health Symposium. Singapore: PMLR, December 4, 2025 (https://arxiv.org/abs/2512.05216).

50. Marfoglia A, Jhee JH, Coulet A. Clinical data goes MEDS? Let's OWL make sense of it. February 2, 2026 (https://doi.org/10.48550/arXiv.2601.04164). Preprint.

51. Zhu J, Feng X, Shen Y, Ning G, Fang L, Pan Q. PMTG: personalized multivariate temporal graph learning with multi-granular encoding for clinical prediction on EHRs. In: 2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Wuhan, China: IEEE, 2025:5589-5596 (https://www.semanticscholar.org/paper/PMTG%3A-Personalized-Multivariate-Temporal-Graph-with-Zhu-Feng/f05fb2ecc5650326da63a3a41f0b82865ba793dd).

52. Zhang Y, Zhou H, Feng L, Hong Y, Li S. Learning longitudinal health representations from EHR and wearable data. January 18, 2026 (https://doi.org/10.48550/arXiv.2601.12227). Preprint.

53. Cho E, Kim J, Lee M, Park S, Choi E. Generating multi-table time series EHR from latent space with minimal preprocessing. The Thirty-Ninth Annual Conference on Neural Information Processing Systems. San Diego, CA: NeurIPS, April 21, 2026 (https://openreview.net/forum?id=k5TbfYPYuc).

54. Hu Y, Chuan C, Zhang W, et al. Icom: an intervention based foundation model for healthcare. January 9, 2026 (https://doi.org/10.36227/techrxiv.176800116.62500980/v1). Preprint.

55. Alonso E, Mendez N, Garcia-Navarro T, et al. EHRs enable robust lung cancer risk stratification with transformer-based models: a retrospective multi-center validation study. December 2, 2025 (https://doi.org/10.21203/rs.3.rs-8200847/v1).

56. Lowe S, Park G, Lee L, Smith P. Latent physiology as language: a state-space foundation model for multimodal ICU and EHR representation learning. 2026. DOI: 10.13140/RG.2.2.21597.27364.

57. Yakdan S, Warner B, Ghogawala Z, et al. Clinically-guided models or foundation models? Predicting cervical spondylotic myelopathy from electronic health records. NPJ Digit Med 2026;9:153. DOI: 10.1038/s41746-026-02337-7.

58. Qian L. On AI frameworks for healthcare time-series analysis: missingness, dynamics and representation. Kings College London, August 2025 (https://kclpure.kcl.ac.uk/ws/portalfiles/portal/351130102/2025_Qian_Linglong_1920412_ethesis.pdf).

59. Lawrence S, Zhang Y, Wang M. Markovian representation learning for longitudinal electronic health records. Research Square, February 24, 2026 (https://doi.org/10.21203/rs.3.rs-8935515/v1).

60. Ferrel W, Chang T, Lawrence S. Sparse foundation models for continuous-time EHRs. February 26, 2026 (https://doi.org/10.21203/rs.3.rs-8819405/v1).

61. McDermott M, Xu J. MIMIC-IV MEDS: an ETL pipeline to extract MIMIC-IV data into the MEDS format. Github. November 5, 2025. DOI: 10.5281/zenodo.17535580.

62. Rockenschaub P. AUMCdb_MEDS: the AUMCdb MEDS ETL. Github. November 5, 2025 (https://github.com/prockenschaub/AUMCdb_MEDS).

63. van de Water RP. NWICU_MEDS ETL. Github. March 14, 2025. DOI: 10.5281/zenodo.14892134.

64. van de Water RP. SICdb_MEDS ETL. Github. February 19, 2025. DOI: 10.5281/zenodo.14893939.

65. van de Water RP. INSPIRE_MEDS ETL. Github. February 19, 2025. DOI: 10.5281/zenodo.14891940.

66. McDermott M, Zhu M, van de Water RP. eICU_MEDS: eICU MEDS Extraction ETL. Github. November 5, 2025. DOI: 10.5281/zenodo.17535692.

67. Github. van de Water RP. HIRID MEDS ETL. October 16, 2025. DOI: 10.5281/zenodo.17370179.

68. Github. femr: FEMR (Framework for Electronic Medical Records) provides tooling for large-scale, self-supervised learning using electronic health records. October 16, 2025 (https://github.com/som-shahlab/femr).

69. Hur K, Oh J, Kim J, et al. GenHPF: general healthcare predictive framework for multi-task multi-source learning. IEEE J Biomed Health Inform 2024;28:502-513. DOI: 10.1109/JBHI.2023.3327951.

70. Pang C, Jiang X, Kalluri KS, et al. CEHR-BERT: incorporating temporal information from structured EHR data to improve prediction tasks. In: Roy S, Pfohl S, Rocheteau E, Tadesse GA, Oala L, Falck F, et al. eds. Proceedings of Machine Learning for Health. Vol 158. PMLR, 2021:239-260 (https://proceedings.mlr.press/v158/pang21a.html).

71. Odgaard M, Klein K, Thysen SM, Jimenez-Solem E, Sillesen M, Nielsen M. CORE-BEHRT: a carefully optimized and rigorously evaluated BEHRT. In: Deshpande K, Fiterau M, Joshi S, Lipton Z, Ranganath R, Urteaga I, eds. Machine Learning for Healthcare Conference. Vienna, Austria: PMLR, October 11, 2024. DOI: 10.48550/arXiv.2404.15201.

72. McDermott M. MEDS "everything-is-code" autoregressive model. Github, November 5, 2025. DOI: 10.5281/zenodo.17535559.

73. Fries J. meds-mcp: MEDS Model Context Protocol (MCP) server and client. Github, April 29, 2026 (https://github.com/VISTA-Stanford/meds-mcp).

74. van de Water R. MEDS-Inspect. Github. October 27, 2025. DOI: 10.5281/zenodo.14892253.

75. McDermott M, Oufattole N, Bergamaschi T, Chandak P, King R, Kolo A. MEDS TorchData. Github. April 21, 2026. DOI: 10.5281/zenodo.17535772.

76. Wornow M, Bedi S, Hernandez MAF, et al. Context clues: evaluating long context models for clinical prediction tasks on EHR data. In: The Thirteenth International Conference on Learning Representations. Singapore: ICLR February 13, 2025 (https://openreview.net/forum?id=zg3ec1TdAP).

77. Yang C, Wu Z, Jiang P, et al. Pyhealth: a deep learning toolkit for healthcare applications. In: Proceedings of the 29th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) 2023. New York, NY: Association of Computing Machinery, 2023:5788-5789 (https://dl.acm.org/doi/10.1145/3580305.3599178).

78. McDermott MBA, Nestor B, Argaw P, Kohane I. Event stream GPT: a data pre-processing and modeling library for generative, pre-trained transformers over continuous-time sequences of complex events. In: Proceedings of the Thirty-Seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track. New Orleans, LA: NeurIPS, 2023 (https://papers.neurips.cc/paper_files/paper/2023/file/4c8f197b24e9b05d22028c2de16a45d2-Paper-Datasets_and_Benchmarks.pdf).

79. van de Water R, Schmidt H, Elbers P, Thoral P, Arnrich B, Rockenschaub P. Yet another ICU benchmark: a flexible multi-center framework for clinical ML. In: The Twelfth International Conference on Learning Representations. Vienna, Austria: ICLR, March 15, 2024 (https://openreview.net/forum?id=ox2ATRM90I).

80. Jarrett D, Yoon J, Bica I, Qian Z, Ercole A, van der Schaar M. Clairvoyance: a pipeline toolkit for medical time series. In: The Ninth International Conference on Learning Representations,

Editor. Austria: ICLR, 2021 (https://openreview.net/pdf?id=xnC8YwKUE3k).

81. Saveliev ES, van der Schaar M. TemporAI: facilitating machine learning innovation in time domain tasks for medicine. January 28, 2023 (http://arxiv.org/abs/2301.12260). Preprint.

82. Wang S, McDermott MBA, Chauhan G, Ghassemi M, Hughes MC, Naumann T. MIMIC-Extract: a data extraction, preprocessing, and representation pipeline for MIMIC-III. Proceedings of the ACM Conference on Health, Inference, and Learning. New York, NY, USA: Association for Computing Machinery, 2020:222-235. DOI: 10.1145/3368555.3384469.

83. Shi W, Xu R, Zhuang Y, et al. EHRAgent: code empowers large language models for few-shot complex tabular reasoning on electronic health records. Proc Conf Empir Methods Nat Lang Process 2024;2024:22315-22339. DOI: 10.18653/v1/2024.emnlp-main.1245.

84. Attrach RA, Moreira P, Fani R, Umeton R, Celi LA. Conversational LLMs simplify secure clinical data access, understanding, and analysis. November 23, 2025 (http://arxiv.org/abs/2507.01053). Preprint.

85. Fries J. meds2text: transforms MEDS-formatted data to textual representations. Github. December 22, 2025 (https://github.com/VISTA-Stanford/meds2text).