

Protocol

Building a Global Research Network for Fair, Accountable, Interpretable, and Responsible AI in Emergency Care: Protocol for a FAIR-EC Study

Chuan Hong^{1*}, PhD; Jonathan Chong Kai Liew^{2*}, BEng; Jaeyong Yu³, PhD; Tomás Barry⁴, MSc, MPH, MBBCh, BAO, PhD; Audrey L Blewer^{5,6,7}, MPH, PhD; Daniel M Buckland⁸, MS, MD, PhD; Tianrun Cai^{9,10}, MD; Won Chul Cha¹¹, MD, PhD; Bibhas Chakraborty¹, PhD; Wei Chen¹², PhD; Jun Cheng¹³, PhD; Shu-Ling Chong¹⁴, MPH, MCI, MRCPCH (UK); Therese Djärv^{15,16}, MD, PhD; Arul Earnest¹⁷, BSocSc (Hons), MSc, PhD; Matthew Engelhard¹, MD, PhD; Xiuyi Fan^{18,19}, MSc, PhD; Mengling Feng^{20,21}, PhD; Jean Feng²², PhD; Huazhu Fu²³, ME, PhD; Wilson Wen Bin Goh^{24,25}, MSc, PhD; Benjamin A Goldstein¹, PhD; Jessica Gronsbell²⁶, PhD; Andrew Fu Wah Ho^{27,28,29}, MMed, MPH, PhD; Kendall Ho³⁰, MD; Taku Iwami³¹, MD, PhD; Anjni Joiner^{8,32}, DO, MPH; Aaron Kornblith³³, MD, MS; Siqi Li³⁴, PhD; Shir Lynn Lim^{20,27,35}, MBBS, MMED; Molei Liu³⁶, PhD; Zhenghong Liu²⁸, MBBS; Lei Lu³⁷, PhD; Yuan Luo³⁸, PhD; Yih Yng Ng¹⁸, MBBS, MPH, MBA; Yilin Ning³⁴, PhD; Yohei Okada^{27,31}, MD, PhD; Ju Ok Park^{39,40}, MD, MPH, PhD; Yu Rang Park⁴¹, PhD; Junaid Razzak^{42,43}, MBBS, PhD; Yuzeng Shen²⁸, MBBS, MMed; Fahad Javaid Siddiqui²⁷, MBBS, MSc (Epidemiology & Biostatistics), MSc (Public Health); Peter A D Steel⁴², MA, MBBS; Kenneth Boon Kiat Tan²⁸, MBBS; Salinelat Teixayavong²⁴, MSc; Bella Vakulenko-Lagun⁴⁴, PhD; Joao Ricardo Nickenig Vissoci³², PhD; Grzegorz Waligora^{45,46}, MD, PhD; Fei Wang⁴⁷, PhD; Haibo Wang⁴⁸, MBBS, MSc, MPH; Haoyuan Wang^{49,50}, MSc; An-Kwok Ian Wong^{1,32,51}, MD, PhD; Feng Xie⁵², PhD; Jie Yang^{53,54,55}, PhD; Yiye Zhang⁴², MS, PhD; Doudou Zhou⁵⁶, PhD; Li Zhou⁵⁷, MD, PhD; Tingting Zhu⁵⁸, MSc, DPhil; Robert Neumar⁵⁹, MD, PhD; David Page¹, PhD; Roger Vaughan⁶⁰, PhD; Marcus Eng Hock Ong^{27,28}, MBBS, MPH; Nan Liu^{1,34,61}, PhD

¹Department of Biostatistics and Bioinformatics, School of Medicine, Duke University, Durham, NC, United States

²Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

³Hallym University, Chuncheon, Republic of Korea

⁴School of Medicine, University College Dublin, Dublin, Ireland

⁵Department of Family Medicine and Community Health, School of Medicine, Duke University, Durham, NC, United States

⁶Department of Population Health Sciences, School of Medicine, Duke University, Durham, NC, United States

⁷Health Services Research and Population Health, Duke–National University of Singapore (NUS) Medical School, Singapore, Singapore

⁸Department of Emergency Medicine, University of Wisconsin - Madison, Madison, Wisconsin, United States

⁹VA Boston Healthcare System, Boston, MA, United States

¹⁰Brigham and Women's Hospital, Boston, MA, United States

¹¹Samsung Medical Center, Seoul, Republic of Korea

¹²School of Biomedical Engineering, The University of Sydney, Sydney, Australia

¹³Institute for Infocomm Research (I2R), Agency for Science, Technology and Research, Singapore, Singapore

¹⁴KK (Kandang Kerbau) Women's and Children's Hospital, Singapore, Singapore

¹⁵Karolinska University Hospital, Stockholm, Sweden

¹⁶Karolinska Institutet, Stockholm, Sweden

¹⁷School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

¹⁸Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

¹⁹College of Computing and Data Science, Nanyang Technological University, Singapore, Singapore

²⁰Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

²¹Institute of Data Science, National University of Singapore, Singapore, Singapore

²²Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

²³Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore, Singapore

²⁴Nanyang Technological University, Singapore, Singapore

²⁵Imperial College London, London, United Kingdom

²⁶Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada

²⁷Duke–National University of Singapore (NUS) Medical School, Singapore, Singapore

²⁸Singapore General Hospital, Singapore, Singapore

²⁹Centre for Population Health Research and Implementation, SingHealth, Singapore, Singapore

³⁰Department of Emergency Medicine, University of British Columbia, Vancouver, British Columbia, Canada

³¹Department of Preventive Services, School of Public Health, Kyoto University, Kyoto, Japan

- ³²Department of Emergency Medicine, Duke University, Durham, NC, United States
- ³³Department of Emergency Medicine, University of California, San Francisco, San Francisco, CA, United States
- ³⁴Centre for Biomedical Data Science, Duke-NUS Medical School, Singapore, Singapore
- ³⁵National University Heart Centre Singapore, Singapore, Singapore
- ³⁶Department of Biostatistic, Columbia University's Mailman School of Public Health, New York, NY, United States
- ³⁷School of Life Course & Population Sciences, King's College London, London, United Kingdom
- ³⁸Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States
- ³⁹College of Medicine, Hallym University, Chuncheon, Republic of Korea
- ⁴⁰Hallym University Dongtan Sacred Hospital, Hwaseong-si, Republic of Korea
- ⁴¹Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, South Korea
- ⁴²New York Presbyterian-Weill Cornell Medicine, New York, NY, United States
- ⁴³Center of Excellence for Trauma and Emergencies, Aga Khan University, Karachi, Pakistan
- ⁴⁴University of Haifa, Haifa, Israel
- ⁴⁵Gdańsk Medical University, Gdansk, Poland
- ⁴⁶Wroclaw Medical University, Wroclaw, Poland
- ⁴⁷Weill Cornell Medicine, New York, NY, United States
- ⁴⁸Research Centre of Big Data and Artificial Intelligence for Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China
- ⁴⁹Center for Health AI and Synthesis of Evidence (CHASE), Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
- ⁵⁰The Graduate Group in Applied Mathematics and Computational Science, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, United States
- ⁵¹Section of Pulmonary and Critical Care Medicine, Department of Medicine, Durham VA Medical Center, Durham, NC, United States
- ⁵²Department of Surgery, Division of Computational Health Sciences, University of Minnesota, Minneapolis, MN, United States
- ⁵³Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Harvard University, Boston, MA, United States
- ⁵⁴Broad Institute of MIT and Harvard, , Cambridge, MA, United States
- ⁵⁵Harvard Data Science Initiative, Harvard University, Cambridge, MA, United States
- ⁵⁶Department of Statistics and Data Science, National University of Singapore, Singapore, Singapore
- ⁵⁷Harvard Medical School, Boston, MA, United States
- ⁵⁸University of Oxford, Oxford, United Kingdom
- ⁵⁹Department of Emergency Medicine, Michigan Medicine, University of Michigan-Ann Arbor, Ann Arbor, MI, United States
- ⁶⁰University of Pennsylvania, Philadelphia, PA, United States
- ⁶¹NUS Artificial Intelligence Institute, National University of Singapore, Singapore, Singapore
- *these authors contributed equally

Corresponding Author:

Nan Liu, PhD
Centre for Biomedical Data Science
Duke-NUS Medical School
8 College Road
Singapore S169857
Singapore
Phone: 65 66016503
Email: liu.nan@duke-nus.edu.sg

Abstract

Background: The current landscape of emergency care (EC) is marked by high demand, leading to issues such as emergency department boarding, overcrowding, and subsequent delays that impact the quality and safety of patient care. Integrating data science into EC can enhance decision-making with predictive, preventative, personalized, and participatory approaches. However, gaps in adherence to fairness, accountability, interpretability, and responsibility are evident, particularly due to barriers to data-sharing, which often result in a lack of transparency and robust oversight in these applications.

Objective: The FAIR-EC (Fair, Accountable, Interpretable, and Responsible-Emergency Care) collaboration adapts the existing Fair, Accountable, Interpretable, and Responsible principles to address emerging challenges as data science integrates with EC. This initiative aims to transform EC by establishing ethical artificial intelligence standards specifically tailored for this integration. By bridging the gap between EC professionals, data scientists, and other stakeholders, the collaboration promotes international cooperation that leverages advanced data science techniques to enhance EC outcomes across different care settings.

Methods: We propose a federated research design to analyze extensive datasets from various global institutions without compromising patient privacy. This approach transforms epidemiological research with advanced data science techniques, emphasizing the harmonization of data for comprehensive analyses across different health care systems.

Results: The FAIR-EC initiative has facilitated the identification and harmonization of datasets from diverse geographical regions, enabling the examination of regional variations in EC practices. As of paper submission, participating sites have identified retrospective EC datasets totaling >2 million records (eg, Duke Health >400,000 and Singapore General Hospital >1.7 million records). Initial projects have demonstrated feasibility and operational readiness, including implementation of federated workflows and ongoing development of a federated scoring system, cross-site evaluation, and adaptation of association studies and predictive models across various regions. Cross-site harmonization and pilot analyses are underway (with local ethics approvals in progress), and first multisite results are expected to be submitted in mid-late 2026, with additional project-level publications anticipated in 2027. These efforts highlight the feasibility of leveraging advanced data science techniques to address the complexities of EC while preserving patient privacy without centralizing individual-level data. This project was funded from September 1, 2022, to August 31, 2023.

Conclusions: FAIR-EC integrates data science ethically and effectively into EC, addressing challenges such as fragmented data, real-time handoffs, and public health crises. Its federated design harmonizes diverse data streams while preserving privacy, and its emphasis on ethical artificial intelligence aligns with the dynamic nature of EC. Despite challenges in data variability and system complexity, FAIR-EC establishes a strong foundation for innovation in global EC.

International Registered Report Identifier (IRRID): DERR1-10.2196/74202

JMIR Res Protoc 2026;15:e74202; doi: [10.2196/74202](https://doi.org/10.2196/74202)

Keywords: emergency care; emergency medicine; pre-hospital Intensive Care Unit (ICU); federated learning; research network

Introduction

Current Status and Demands of Emergency Care

Emergency care (EC) involves the rapid diagnosis, evaluation, and treatment of patients with acute illness or injury and decompensated chronic illness, spanning prehospital care (eg, ambulances), emergency departments (EDs), intensive care units, and beyond (eg, operating rooms and procedural units). Its critical role is to ensure timely diagnosis and interventions that reduce morbidity and mortality across medical, surgical, and psychiatric emergencies [1].

Global demand for EC is surging, straining health care systems with ED overcrowding, delayed admissions, and prolonged wait times [2]. These challenges stem from factors such as aging populations, longer lifespans, and rising care expectations, creating a mismatch between demand and system capacity [3,4]. In the United States, ED overcrowding is primarily driven by inpatient boarding, where admitted patients remain in the ED due to inadequate inpatient capacity, rather than a significant increase in ED visits or hospitalizations over the past decade.

To address these issues, innovative solutions are needed to improve efficiency and decision-making across the entire EC continuum, from initial health system contact to postacute recovery. This includes optimizing prehospital and ED management, expediting inpatient admissions, and integrating data across hospital wards, skilled nursing facilities, rehabilitation centers, and outpatient follow-up clinics. Leveraging electronic health records (EHRs) and complementary data sources (such as secure messaging, wearable devices, and home monitoring) systemically enables

better coordinated, comprehensive, and longitudinal patient care.

Integrating Data Science and Artificial Intelligence in EC

Clinical decision-making in EC has traditionally relied on physicians' experience, supplemented by heuristic decision-support tools. However, in high-pressure environments such as overcrowded EDs, this approach can be inconsistent due to variations in adherence to evidence-based frameworks. Artificial intelligence (AI)-based clinical support algorithms offer a solution by synthesizing vast amounts of data, reducing cognitive biases, and ensuring more standardized, data-driven decisions. The widespread adoption of EHRs in EC presents opportunities to enhance decision-making through AI-driven analysis [5]. By leveraging structured data and real-time monitoring, AI models can predict conditions such as sepsis before clinical symptoms appear, optimizing workflows and improving patient outcomes [6] though their real-world implementation remains limited [7-9].

Gaps and Opportunities

Current EC methodologies lack full integration of ethical AI principles, leading to oversimplified decision-making that fails to address inherent biases, disproportionately affecting marginalized populations [10]. Additionally, disparities in access to AI tools hinder low-resource EDs from benefiting from association and predictive analytics, and real-time decision support, widening global health inequities [11]. Transparency and interpretability remain major challenges, limiting clinician trust and adoption [10,12,13]. Insufficient accountability and regulatory oversight further risk the ethical and responsible use of AI in EC [14]. Ethical concerns, such as patient autonomy and privacy, also require more rigorous

safeguards [15]. International, interdisciplinary collaboration presents a critical opportunity to bridge these gaps [16,17]. By integrating expertise from data scientists, EC professionals, ethicists, and patient advocates, one can develop scalable, transparent, and interpretable AI models to ensure equitable care and set new global standards for EC.

Ethical AI Standards in EC

Ethical AI integration in EC faces significant technical and systemic barriers, including data accessibility [18], integration challenges, privacy concerns, incomplete data, and real-time processing constraints [15,19,20]. To address these challenges, AI implementation must adhere to the core principles of the Fair, Accountable, Interpretable, and Responsible framework (Multimedia Appendix 1 [21-30]). “Fairness” ensures equitable outcomes, preventing algorithmic bias and promoting equal care across diverse patient populations [23]. Methods such as data rebalancing [31], threshold recalibration [32], and performance monitoring [33] can help mitigate biases. “Accountability” requires that AI developers, clinicians, and regulators ensure justifiable algorithmic decisions and establish clear protocols for addressing errors or biases, fostering transparency and trust [34-37]. “Interpretability” is crucial for clinical adoption, allowing health care providers to understand AI-driven recommendations and ensuring that AI supports rather than dictates clinical care [12, 24,25]. “Responsibility” entails the ethical use of AI, ensuring patient safety, privacy, and dignity while enhancing decision-making [26,38-41].

FAIR-EC Network

The FAIR-EC (Fair, Accountable, Interpretable and Responsible-Emergency Care) network is built on the principle that international collaboration is key to developing ethical, scalable data science techniques in EC. Given the variability in patient demographics, health care infrastructure, and resources across regions, evaluating data-driven solutions in diverse settings—including low-resource environments—is essential. The Pan-Asian Resuscitation Outcomes Study [42] underscores disparities in EC data availability, highlighting the need for adaptable techniques. By uniting emergency professionals and data scientists, FAIR-EC ensures solutions are both theoretically sound and effective in real-world applications.

Master Data Representation

A key aspect of this framework is the master data representation, which facilitates interoperability across diverse EC datasets. This is achieved through the adoption of standardized data models, including the Patient-Centered Outcomes Research Network common data model, Observational Medical Outcomes Partnership, and Fast Healthcare Interoperability Resources standards, which provide structured, interoperable representations of patient records, diagnoses, interventions, and outcomes. Additionally, ontology mapping and terminology harmonization using the Unified Medical Language System and SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms)

ensure that medical concepts are consistently interpreted across institutions.

To further unify data across institutions, FAIR-EC leverages natural language processing [43] and large language models (LLMs) [44] for cross-institutional data harmonization. BERT (Bidirectional Encoder Representations From Transformers)-based models are used for entity recognition and concept extraction from clinical notes, while graph convolutional networks (GCN) are used for knowledge graph-based patient stratification. Additionally, LLM-enhanced data augmentation improves data imputation strategies and supports missingness-aware predictions. FedIMPUTE, a federated-transfer imputation method, is specifically designed to address missing data across institutions, enhancing model robustness and consistency [45].

Federated Learning and Transfer Learning for Model Development

Federated learning [46] enables decentralized model training without sharing raw patient data, preserving privacy while allowing for cross-institutional collaboration. Key implementations include FedScore, a federated scoring system for association and predictive modeling, and FedFML, a federated meta-learning approach that tailors models to individual institutions while leveraging global insights.

The Score for Emergency Admission Prediction project evaluates the generalizability of Singapore’s inpatient admission prediction models across international sites, ensuring their applicability beyond their original development context. Meanwhile, TRACER (Dynamic Transfer Learning) is designed to continuously adapt models as new data becomes available, ensuring they remain relevant and effective in evolving clinical environments.

Methods

Study Design and Setting

FAIR-EC is based on a federated research design, a strategic approach that enables the collaborative analysis of data from diverse institutions without necessitating the sharing of individual patient-level data, while applying transfer learning techniques [33,47-49]. While this avoids the centralization of data across different countries or regions, it is important to note that within each participating site, researchers still collect and store private and confidential data on individuals. This approach ensures the utmost respect for security while leveraging the collective strength of data to uncover insights that can significantly improve EC practices. Additionally, the federated learning framework adopted by the FAIR-EC study does not introduce new security risks, aligning with current best practices in AI to safeguard data integrity and confidentiality. Our federated learning framework is communication-efficient and easy to implement [50,51]. Therefore, even in low-resource settings, they can still implement federated learning algorithms and participate in this collaboration without requiring advanced hardware.

The collaboration will be conducted across various participating institutions worldwide, showcasing a commitment to international collaboration. By adopting this federated framework, FAIR-EC aims to foster a comprehensive understanding of EC dynamics on a global scale, facilitating the development of scalable, equitable, and effective data science solutions within the realm of EC.

Core Study Definition and Scope

FAIR-EC functions both as a collaborative research network and a unified study protocol. At its core, this study aims to evaluate the feasibility, validity, and fairness of federated, interpretable machine learning models for key EC tasks such as admission prediction, triage, and discharge planning across diverse international sites. The primary outcomes include predictive performance (eg, area under the receiver operating characteristic curve [AUROC] and calibration), equity across subgroups using fairness metrics such as the group-conditional concordance index, and technical feasibility as measured by successful implementation of federated workflows, participation rates, and data completeness. Secondary outcomes focus on operational utility, user trust, and robustness under data shift or incompleteness. Specific analytic modules and modeling frameworks that address these outcomes are introduced in later sections of the protocol. In contrast, spin-off or locally initiated projects that focus on narrower clinical questions or extended data modalities will not be considered part of the core study, although they may build on FAIR-EC's infrastructure and adhere to its principles.

Participating Sites and Study Cohorts

FAIR-EC will use datasets originating from a broad geographical spectrum, including regions and countries from Asia, Australasia, North America, Europe, and the Middle East. These datasets are critical to understanding the global landscape of EC, providing insights into regional practices, outcomes, and challenges. Within each participating site, we target a wide-ranging population of patients receiving EC, with inclusion criteria designed to capture a broad spectrum of emergency presentations. Exclusion criteria will be carefully defined to ensure clarity and relevance of the data, focusing on removing cases that do not meet the EC

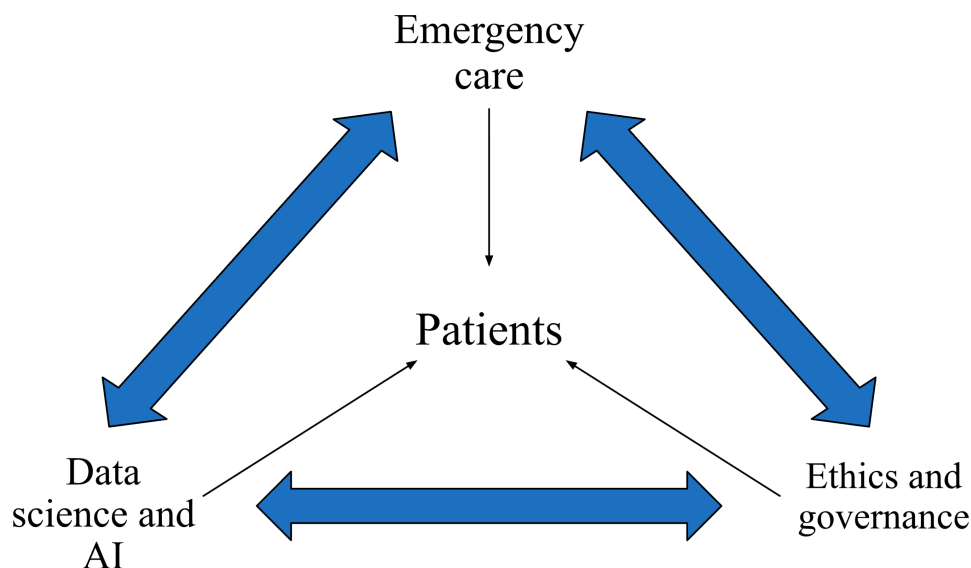
criteria or lack sufficient data for meaningful analysis. Data collection will be conducted in a site-specific manner, tailored to meet the unique needs and capabilities of each participating location to ensure the accuracy and applicability of the data obtained.

Given the broad scope of FAIR-EC, specific inclusion and exclusion criteria, variable sets, and analytic strategies vary across individual projects and participating sites. Each project within the FAIR-EC network is required to define its analytic cohort, outcomes, modeling strategy, and validation metrics in detail through locally approved institutional review board (IRB) protocols. While we do not prescribe a single patient-level definition or analytic pipeline across all studies, we ensure consistency through harmonization guidance, cross-site code templates, and shared evaluation principles. As an illustrative example, one active project evaluates hospital admission prediction across three Duke hospitals using 2019 ED visits. The cohort includes adult ED encounters with available demographics, vital signs, and comorbidity data (eg, pulse, blood pressure, oxygen saturation, acuity level, and *ICD* [*International Statistical Classification of Diseases*]-based comorbidity scores). The primary outcome is inpatient admission, and the project compares federated imputation approaches against local and pooled baselines, evaluating mean squared error and AUROC. Similar project-specific protocols are under development at other institutions and will follow comparable standards for cohort definition, model comparison, and validation. Details for each will be introduced in subsequent sections or appendices.

Interdisciplinary Collaboration Mechanisms

To facilitate interdisciplinary collaboration, a multifaceted infrastructure (Figure 1) will be established, encompassing emergency physicians and other EC providers, data scientists, social scientists, ethicists, patients, community supporters, and other relevant stakeholders. By establishing the mechanisms for interdisciplinary collaboration, communication, and iterative feedback, FAIR-EC aims to create a dynamic environment where innovative solutions are developed through the synergy of diverse expertise.

Figure 1. Interdisciplinary interaction mechanisms centered on the patient for collaboration in FAIR-EC, leveraging data across the entire episode of care—from initial health system contact to return to baseline or a new steady state. AI: artificial intelligence; FAIR-EC: Fair, Accountable, Interpretable, and Responsible-Emergency Care.

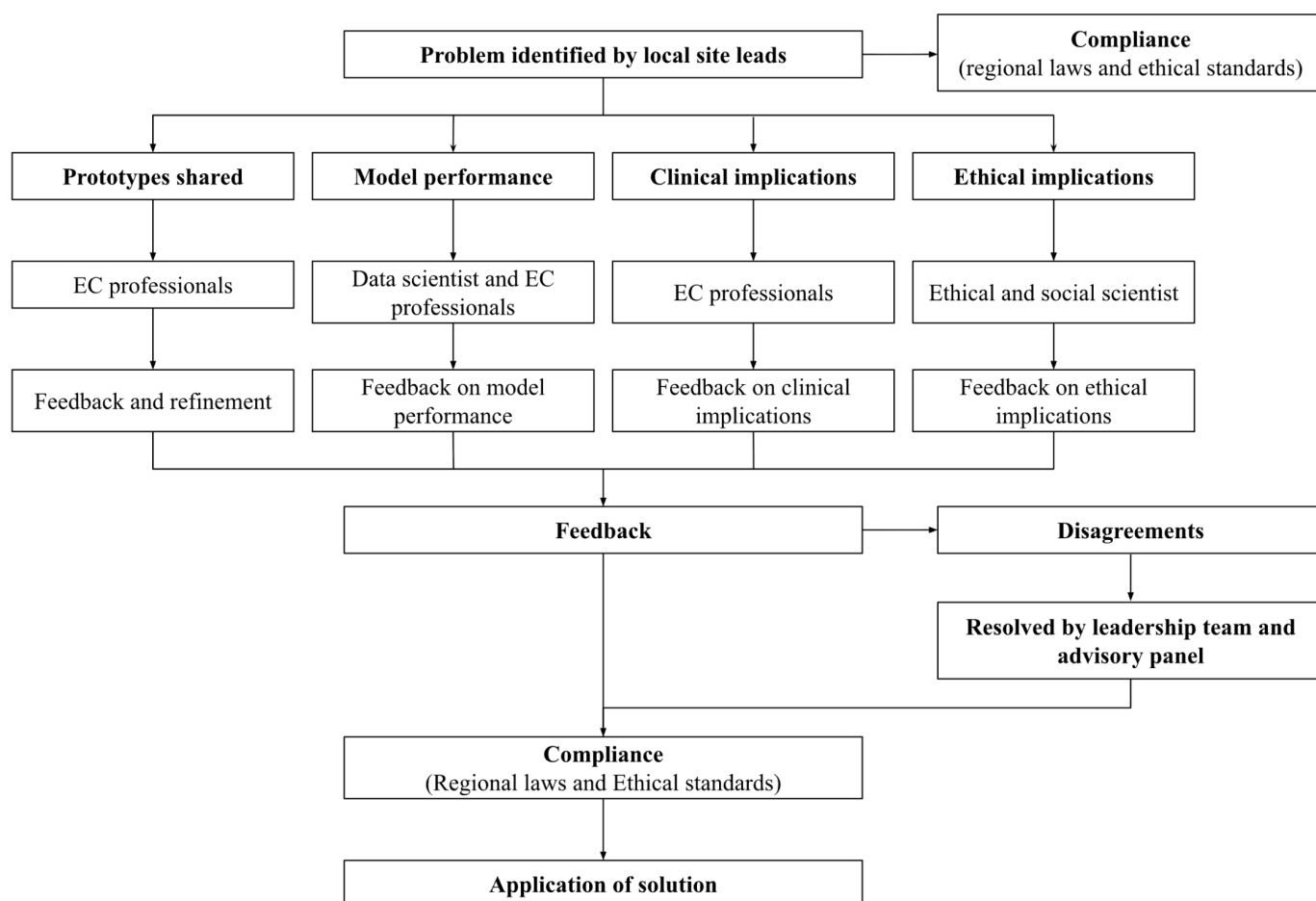


Collaborative Platforms

A digital collaboration platform (Figure 2) will be used to foster communication and project management among the interdisciplinary team members. This platform will support

document sharing, project tracking, and real-time discussions, ensuring that all participants, regardless of their geographic location, can contribute effectively to this project.

Figure 2. Iterative feedback mechanisms and stakeholder engagement. EC: emergency care.



Cross-Disciplinary Teams

This study will organize participants into cross-disciplinary teams, each tasked with addressing specific aspects of this project. These teams will blend the expertise of EC professionals, data scientists, epidemiologists, biostatisticians, and ethicists to ensure that every solution is developed with a holistic view of its clinical applicability, technical feasibility, and ethical implications. Examples include federated admission prediction to support real-time bed management and reduce ED boarding, deterioration models to streamline triage and accelerate clinician decision-making, and predictive discharge tools to assist case managers with patient flow coordination. By embedding implementation partners from the outset, FAIR-EC functions as both a data science network and an implementation platform, enabling the alignment of technical innovation with practical workflow improvements across diverse health care systems.

Implementation and Workflow Impact

A core objective of FAIR-EC, beyond developing predictive models, is to ensure that these models translate into meaningful efficiency gains for clinicians, administrators, and institutional leaders, who are the end users of EC systems. FAIR-EC brings together data scientists, emergency physicians, nursing staff, and hospital leadership in cross-disciplinary teams to co-design projects around pressing operational challenges.

Communication Channels and Workshops

Scheduled virtual and in-person meetings, when feasible, will ensure ongoing dialogue between all team members. These meetings will serve as forums for progress updates, brainstorming sessions, and resolving any challenges that arise during the collaboration. The FAIR-EC collaboration will organize workshops that bring together participants from various disciplines to share knowledge, discuss emerging trends in EC and data science, and explore the ethical considerations relevant to the project. These workshops will be crucial for maintaining the alignment of the project with the latest scientific and ethical standards.

Iterative Feedback Mechanisms and Stakeholder Engagement

Iterative feedback mechanisms and the incorporation and assessment of stakeholder input are central to refining our models and tools within the EC context (Figure 2). Prototypes developed will be tested by EC professionals in simulated or real-world settings, and their feedback will be crucial for ensuring that these solutions meet the practical needs of EC settings. Additionally, ethical review panels, including ethicists and social scientists, will periodically review the project's progress to ensure ethical considerations are thoroughly addressed. These panels will guide necessary adjustments based on the ethical implications of the data science techniques used. This proposed feedback loop will allow for continuous refinement of models based on clinical impacts observed in practice and insights from ethical

reviews, ensuring alignment with clinical needs, ethical standards, and societal expectations.

We will also use feedback from various stakeholder groups to inform and enhance the presentation of risk estimates. If different subgroups, such as health care administrative leadership and patients, highlight distinct priorities, we will consider tailoring the presentation to better meet these specific needs. This feedback will be assessed through a second phase of focus groups, which will evaluate the appropriateness and effectiveness of the presentation adjustments. This phase will include discussions and ranking processes to finalize communication options. We plan to conduct 3 focus groups, one for each type of stakeholder, drawing from the same participant pool as the initial phase.

Governance and Conflict Resolution

The FAIR-EC network operates under a multitiered governance structure designed to balance clinical, technical, and ethical perspectives. A central leadership team coordinates strategic direction and day-to-day activities, supported by an international advisory panel that provides independent guidance. Interdisciplinary working groups are responsible for reviewing project proposals, developing study protocols, and monitoring ongoing activities.

To address conflicts that may arise between data science objectives and clinical perspectives, FAIR-EC uses structured consensus-building methods such as moderated discussions and Delphi processes. If consensus cannot be reached at the working group level, issues are escalated to the leadership team for resolution in consultation with the advisory panel.

Given the network's global scope, we anticipate differences in regulations, cultural norms, and data governance requirements across sites. Local site leads are responsible for ensuring compliance with regional laws and ethical standards, while FAIR-EC provides overarching principles to align practices. Communication materials and protocols are designed to support multilingual participation, and data harmonization workflows account for both technical heterogeneity and cultural considerations in the interpretation of clinical data. This governance structure ensures that FAIR-EC remains inclusive, ethically grounded, and adaptable to diverse health care contexts.

Ethical Oversight Structure

Ethical oversight within FAIR-EC is embedded at multiple levels to balance collective accountability with local autonomy.

At a network level, steering and ethics working groups establish overarching ethical principles, provide strategic guidance, promote consistency across projects, and mediate conflicts when they arise, while preserving the decentralized ethos of federated learning.

On the project level, for each project undertaken within FAIR-EC, the project's authorship team serves as the coordinating body. These teams are responsible for applying FAIR-EC principles in study design and implementation,

integrating interdisciplinary expertise, and reviewing project protocols in collaboration with participating sites.

Finally, at the local site level, each participating institution maintains responsibility for compliance with its own IRBs, regulatory requirements, and cultural norms. Local investigators ensure that implementation respects regional laws and governance structures while contributing to the shared ethical standards of the network.

This layered structure ensures that ethical oversight is exercised consistently across the network while remaining sensitive to local contexts and governance requirements, as shown in [Figure 2](#).

Planned Recruiting Strategy

To strategically expand our network of participating sites, FAIR-EC plans to engage with global and regional EC organizations and conferences. We will present our research goals and preliminary findings at international health care symposiums to attract interest from potential site partners. Additionally, targeted outreach will be conducted to leading hospitals and research institutions known for their innovative approaches to EC. By offering these institutions the opportunity to contribute to and benefit from our federated research model, we aim to ensure a diverse and representative array of sites that can provide valuable data and insights, enhancing the overall impact and applicability of our findings in the global context of EC. Although the protocol primarily targets ED patients, the inclusion of patients is scalable, and we will not limit our scope solely to ED patients, allowing for the expansion of patient categories as this study progresses.

For additional institutions interested in joining the FAIR-EC network, we have established a clear and accessible pathway for collaboration. Interested parties can reach out through a dedicated section on the FAIR-EC project website, which details the collaboration process, including eligibility criteria, expected contributions, and benefits of participation. Institutions can also contact the FAIR-EC administrative team directly via email to express interest and receive guidance on the application process. Furthermore, we will offer informational webinars and question-and-answer sessions to provide prospective sites with a comprehensive understanding of the project's scope, goals, and collaborative framework. This open and structured approach ensures that any institution, regardless of its location or size, could become part of this transformative research endeavor.

Multimodal Data

We use data from 2 distinct modalities: structured codified data and unstructured data. Structured data encompasses quantifiable and easily searchable elements such as demographics, diagnosis codes, medications, laboratory tests and results, procedures, and genomic or genetic data (if available). Some social determinants of health are also included in this category, enabling the systematic analysis of trends in diagnoses and treatments (eg, Index of Relative Socio-Economic Disadvantage in Australia [52] and indexes of socioeconomic disadvantage or socioeconomic advantage indices in Singapore [53]). Such variables should be carefully

selected and organized to keep track of the socioeconomic status over time.

In contrast, unstructured data provides richer context and nuance. This includes clinical narratives (progress notes, discharge summaries, and radiology reports) that offer insights into patient symptoms and treatments not captured by structured fields. It also covers narrative accounts of social determinants of health, imaging data such as plain radiography and advanced diagnostic imaging (eg, computed tomography and magnetic resonance imaging) for rapid diagnosis, and waveform data (eg, electrocardiograms and electroencephalograms) that reveal real-time physiological responses. These unstructured modalities are indispensable for immediate, accurate medical decision-making in EC, supplementing the foundational insights obtained from structured data.

It is important to note that there may be variations in the availability, format, and definitions of each data modality across different participating sites in the FAIR-EC collaboration. Not all sites have access to the same range of modalities, such as imaging data or detailed genotype data, and even when similar data types are available, they may differ significantly in format and detail depending on regional practices and technological infrastructure. This variation necessitates comprehensive data harmonization efforts proposed in the next section, which are critical to ensure the integrity and comparability of the data across all sites. By allowing for the unique contributions of each site's available data and addressing discrepancies in data formats systematically, this study maximizes the breadth and depth of its insights while maintaining rigorous standards of data quality and consistency.

Planned Privacy-Preserving Federated Learning System

A key strategy in our approach is the use of federated learning, which will be adapted from our previous work [54,55]. The federated learning framework allows for the collaborative analysis of data from various sites without the need to centralize data, thus preserving patient privacy and data confidentiality. By distributing the computational tasks across multiple participating institutions, federated learning enables the FAIR-EC study to benefit from the rich diversity of global data while adhering to strict privacy standards. This approach is particularly effective in EC, where the variability of EC practices and outcomes across different regions can be vast.

To ensure broad participation across institutions with varying technical capacities, FAIR-EC adopts a tiered participation model. At one level, high-resource sites are able to run local model training directly on institutional servers, contributing model updates to the central coordinating layer while keeping raw data fully local. For sites with access to institutionally governed cloud environments, participation can be facilitated through the deployment of widely available federated learning toolkits (eg, Flower, TensorFlow Federated, and PySyft), which reduce local infrastructure

demands while maintaining institutional oversight of data. In settings without access to cloud resources, containerized deployments (eg, Docker or Kubernetes) provide a lightweight, reproducible environment that can operate on modest local hardware, thereby lowering barriers to entry. This approach preserves flexibility while ensuring comparability across sites, making it possible for institutions with different resource levels to contribute meaningfully to FAIR-EC. By combining centralized coordination of model aggregation with multiple pathways for site-level engagement, the framework enables scalable participation without compromising privacy or creating undue technical burden.

While federated learning avoids direct data sharing, we acknowledge that residual privacy risks remain, including potential gradient or model inversion attacks. To mitigate these risks, FAIR-EC incorporates secure aggregation, differential privacy mechanisms, and gradient clipping. For small sites or rare cohorts, where limited sample size may heighten reidentification risk, we adopt additional obfuscation strategies such as update padding, feature bucketing, and stronger noise injection, with minimum cohort thresholds required before updates are shared. Where thresholds are not met, sites may participate through pooled aggregation or contribute via federated analytics rather than direct model updates. To further reduce disparities in participation, our framework emphasizes communication-efficient algorithms, adaptive update protocols, and lightweight containerized deployments that minimize setup requirements for sites with limited infrastructure. These safeguards collectively ensure that privacy is preserved while enabling fair engagement across institutions with diverse capacities.

Planned Data Harmonization and Knowledge Network

Participating sites will use common data standards such as the Observational Medical Outcomes Partnership [56], Patient-Centered Outcomes Research Network [57], and Informatics for Integrating Biology and the Bedside [58]. We do not select a single model to ensure flexibility across diverse data environments and infrastructure levels among international sites. These models facilitate the standardization and interoperability of data across different health care systems, enabling a cohesive analysis despite the heterogeneity of the datasets. To address the challenges of data diversity and ensure a unified dataset for analysis, this current study proposes using the established BERT architecture for its proven effectiveness in understanding contextual relationships within text [59-61], and augmenting it with newer transformer models such as GPT-4, for enhanced data harmonization and metadata creation [62]. This dual approach allows us to leverage the robust foundational capabilities of BERT while benefiting from the advanced contextual processing powers of recent transformers. To integrate data from different common data models, these tools will be instrumental in translating disparate data elements into a standardized format, ensuring consistency and comparability across the datasets. Notably, in some countries, LLM workflows have been actively implemented to extract data

from electronic medical records for research purposes, though these technically feasible implementations still await regulatory approval.

To effectively manage the diverse multimodal and multisource datasets, we propose to implement sophisticated algorithms and tools designed to aggregate and process data from various modalities. Specifically, we are using our previously developed semantic learning approaches [63,64], state-of-the-art multimodal deep learning architectures [65], and contrastive learning [66] to derive meaningful representations from these large datasets without the need for phenotypic (labeled) data, which is often scarce in health care settings. This approach enables us to harness the full potential of available data, improving our ability to identify and analyze complex patterns associated with outcomes that are inherent in the multifaceted nature of health data. We focus on semantic integration, creating cohesive semantic spaces that allow for seamless interaction across different data modalities, such as structured data, unstructured clinical narratives, and genotype data.

As part of our broader strategy for data harmonization and representation learning, we will develop a GCN that incorporates a prototype graph [67]. A prototype graph will be developed to serve as a conceptual model to define and visualize the fundamental relationships and interactions among various dataset elements, such as multimodal covariates, risk factors, and health outcomes, which is crucial for mapping complex interdependencies. By using the prototype graph, we can precisely characterize and understand the intricate relationships within our data, informing the development and refinement of our association studies' predictive models tailored to EC settings. We plan to train the GCN using a combination of publicly available datasets from participating sites that have been annotated by a team of domain experts and data scientists. These datasets will encompass a diverse range of EC scenarios to ensure robust feature representation. Labeling will be overseen by clinical professionals to maintain accuracy and relevance. To validate the GCN's generalizability, we will conduct cross-validation studies across multiple participating sites, comparing performance metrics against baseline models to ensure it effectively models the relationships among data from various settings.

In addition, we recognize the role of vendor-facilitated platforms such as Epic Cosmos (Epic Systems Corporation), which demonstrate the potential for large-scale data aggregation within Epic-using institutions. However, reliance on such platforms alone would be restrictive, as not all participating sites use Epic and many already operate research infrastructures based on other data models. We therefore position Epic Cosmos as one option within a broader ecosystem, while ensuring that FAIR-EC remains flexible and inclusive across diverse EHR systems and international contexts.

For semantic alignment across sites, we supplement these mappings with ontology bridging (eg, Unified Medical Language System and SNOMED CT) and transformer-based methods for harmonizing unstructured text. This layered

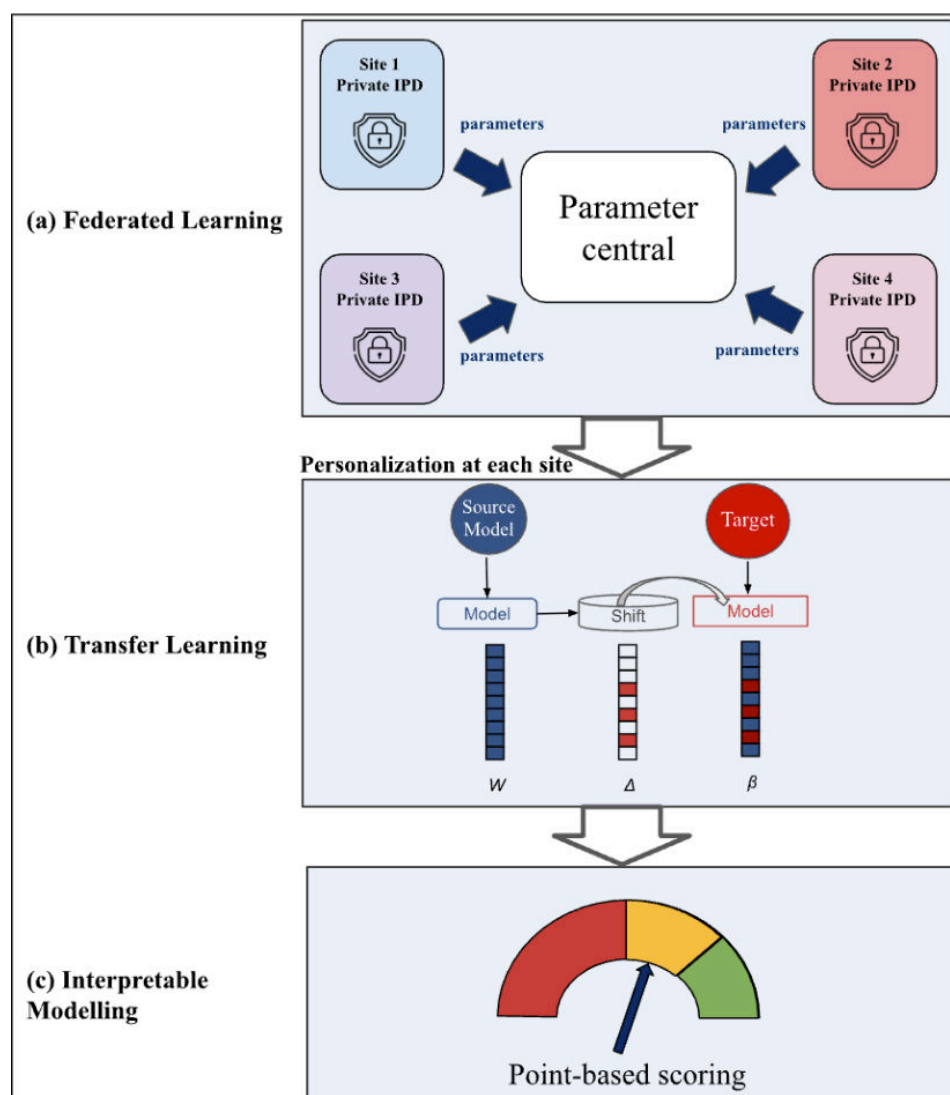
approach allows local flexibility while preserving a shared analytical layer, enabling federated analyses that are both robust and adaptable to the diverse data environments represented in FAIR-EC.

Planned Data Analytics and Model Development

Our approach to data analysis and model development in FAIR-EC is to be adaptive, allowing the application of

different and emerging techniques based on the specific needs of each project. Under the privacy-preserving federated learning system, our general approach is structured around two main modules: transfer learning for local adaptation, and point-based scoring systems to enhance interpretability, with a multimodal transformer as the backbone, adept at managing diverse data across various modalities, sample rates, and formats, even when unaligned (as illustrated in Figure 3).

Figure 3. Framework for model development under federated learning: transfer learning and interpretable modeling. IPD: individual patient data.



Federated Transfer Learning and Negative Transfer Mitigation

Transfer learning involves taking a preexisting model developed for one task or from one population and adapting it for another related task or population [68]. In the context of EC, this includes leveraging knowledge and patterns learned from one health care system or dataset to improve or tailor models for another, with particular attention to addressing disparities in underrepresented groups. Transfer learning can function in two distinct ways: (1) as a model calibration step following federated learning, where it fine-tunes global models to local contexts, and (2) as an independent approach,

particularly useful in low-resource settings. In such cases, researchers can directly leverage knowledge from published studies to adapt and enhance models for their new studies, even when they cannot establish a suitable collaboration in a short time. A notable example from our previous work illustrates how we address demographic disparity in the presence of class imbalance, by leveraging recent advances in imbalance learning, transfer learning, and federated learning [69].

Transfer learning can extend model utility to sites with limited data, but its effectiveness depends strongly on domain similarity. Applying models across heterogeneous health care

systems raises risks of negative transfer and bias amplification. To address these concerns, FAIR-EC adopts a federated transfer learning strategy that draws from multiple potential source sites rather than relying on a single external dataset. By evaluating similarity across candidate source domains, the framework preferentially borrows knowledge from institutions with population and practice patterns most aligned with the target site. When similarity is low, adaptation techniques are used to recalibrate models, and contributions from dissimilar sites are down-weighted or excluded to avoid negative transfer.

Equity Considerations, Ethical Implications, and Engagement Plan

Ensuring that EC models are equitable also necessitates a rigorous ethical framework guiding the collection, use, and interpretation of EHR data. This framework should prioritize transparency, accountability, and inclusiveness in model development and deployment, particularly regarding how data are used and how decisions based on model predictions impact patient care across different populations. The commitment to these ethical principles not only enhances the credibility and acceptability of association studies' predictive models but also ensures that these innovations contribute positively to public health goals, reinforcing trust among all stakeholders, including patients, health care providers, and health care administrative leadership. Active engagement of stakeholders is crucial for developing models that are technically sound, ethically grounded, and practically relevant. To this end, we will hold regular workshops and feedback sessions with clinicians, patients, and policymakers to integrate diverse perspectives directly into the development and implementation processes.

Planned Evaluation Strategy

We propose a comprehensive evaluation strategy designed to rigorously assess the performance and impact of the developed models, which not only assesses the technical performance of the developed models but also critically examines their ethical implications, fairness, and real-world utility. By rigorously benchmarking against traditional tools, comparing with black box algorithms, and focusing on fairness and equity, the FAIR-EC collaboration seeks to pave the way for the responsible and effective use of data science in EC.

Benchmarking Against Current State-of-the-Art Standards in EC

A crucial aspect of the evaluation strategy involves benchmarking the performance of the developed data science models against traditional EC tools currently in use. Specifically, we will compare our models to traditional clinical practice guidelines, which are driven by evidence and expert opinion, as well as to current decision support tools integrated within EHRs. This comparison will focus on various metrics such as accuracy, efficiency, and AUROC. The objective is to demonstrate tangible improvements and advantages of adopting advanced data science techniques in

EC settings, thereby justifying the transition from traditional practices to more innovative, data-driven approaches.

Comparison With Black Box AI Algorithms

This study will undertake a critical comparison between the interpretable models developed within the FAIR-EC framework and existing “black box” algorithms, which are known for their advanced and sophisticated architectures but lack transparency. While black box models often demonstrate superior predictive power, we aim to show that incorporating value-added principles such as transparency, fairness, and interpretability can achieve noninferior performance [70]. This comparison will highlight the ethical and practical advantages of interpretable models, emphasizing their role in clinical decision-making tools that clinicians can trust and understand. By demonstrating comparable performance alongside enhanced explainability, we aim to facilitate the broader adoption of these models in EC practices, ensuring they align with both ethical considerations and clinical needs.

Fairness and Equity Assessment

This study will use advanced methodologies specifically designed to assess algorithmic fairness across sociodemographic groups, ensuring that the models do not perpetuate existing biases or create new disparities in EC. This rigorous assessment will involve a detailed analysis of the models' performance and outcomes across different patient populations, with a focus on identifying any discrepancies or biases in predictions or recommendations. Methods such as disparity impact analysis, fairness metrics (eg, Canadian Index of Multiple Deprivation [71]) and equity rubric lists (eg, PROGRESS Plus) [72] will be used to systematically evaluate how algorithms perform across various groups defined by age, gender, ethnicity, socioeconomic status, and other relevant factors, depending on the data available at each site.

To further enhance the evaluation of fairness and equity, we will engage a diverse group of stakeholders—including patients, health care providers, health care administrators, population health scientists, and economists—both before and after model deployment. This comprehensive approach ensures that our models are developed and continuously improved with input from those most impacted by their implementation. We will also identify and use specific indices, such as access to care and health outcome disparities, to adjust and calibrate the model for greater equity throughout its lifecycle.

Potential Bias Mitigation Strategies

Recognizing the critical importance of addressing and mitigating bias, this study will explore and implement a range of strategies designed to ensure the fairness and equity of the developed models. These strategies may include rebalancing datasets, using algorithmic fairness techniques, and incorporating feedback loops that allow for the continuous monitoring and adjustment of models based on real-world performance data. By actively engaging in bias mitigation, the

FAIR-EC study aims to set a new standard for the development and deployment of equitable data science solutions in EC.

Ethical Considerations

As this paper describes a research protocol for a federated, multi-institutional collaboration, no individual patient-level data are shared centrally, and no centralized analyses of identifiable data have been conducted to date. As of the time of writing, ethics review board (IRB) applications are under preparation or review at participating institutions for their respective data contributions and analyses. Each site in the FAIR-EC network is responsible for obtaining approval from its own ethics board before engaging in data analysis or federated learning. The Duke University site operates under the oversight of the Duke Health IRB (Pro00120917), which covers federated infrastructure development and use of deidentified EHR data for modeling. On the other hand, at the coordinating site at Duke-NUS Medical School (Singapore), this study operates under the ethics policies of the NUS-IRB-2021-766. Under the National Healthcare Group Domain Specific Review Board, it adheres to SingHealth's Centralised Institutional Review Board Ref: 2021/2122.

This protocol paper does not report any direct analysis of patient-level data and is thus exempt from centralized ethics board review under institutional policy (Duke IRB SOP 1.2; National Healthcare Group Domain Specific Review Board guidelines on multisite research, version 12). Future analyses will be conducted only after appropriate IRB or ethics board approval is obtained at each participating site, in compliance with regional and institutional guidelines for human participants research.

Beyond technical safeguards and fairness metrics, the ethical foundation of FAIR-EC centers on the responsible stewardship of innovation in high-stakes, time-sensitive clinical environments. EC often involves rapid decisions under uncertainty; integrating AI must not exacerbate power imbalances, erode clinical autonomy, or marginalize vulnerable populations. FAIR-EC explicitly rejects approaches that prioritize performance over transparency or that rely on opaque, nongeneralizable models. Instead, we emphasize deliberative governance, participatory design, and contextual adaptability across geographies. Our ethical stance informs not only model development but also decisions about deployment, data sharing, and global collaboration, ensuring

that FAIR-EC remains accountable not just to regulators or researchers but to the patients and communities it aims to serve.

As this protocol involves secondary analysis of routinely collected, deidentified EHR data, individual informed consent is waived at participating sites in accordance with local IRB determinations. Importantly, no identifiable patient data will leave institutional boundaries; only aggregated model parameters are shared during federated learning, and all data are deidentified in accordance with applicable regulations (eg, Health Insurance Portability and Accountability Act, and Personal Data Protection Act) [73,74]. As ED populations inherently include vulnerable populations, the FAIR-EC protocol addresses these considerations through fairness-aware modeling that evaluates performance across demographic subgroups to prevent algorithmic bias. As this is a protocol paper, demographic characteristics will be reported in future publications describing model development and validation. No compensation is provided to patients, as this study involves retrospective analysis of existing records with no direct patient contact. All future analyses will adhere to the WMA (World Medical Association) Declaration of Helsinki.

Current Status and Planned Projects

The FAIR-EC study has made significant strides, starting with establishing a comprehensive network of participating institutions across the globe. These efforts are not presented as formal analyses under this protocol, but rather as examples of ongoing and planned work that align with the FAIR-EC framework. Specific modeling strategies, evaluation metrics, and validation procedures will be introduced in project-level protocols or subsequent publications.

Participating Institutions and Teams

Overview

The FAIR-EC collaborative framework has already facilitated a range of projects targeting association and predictive modeling, efficiency optimization, and fairness in EC delivery. The initiative's global reach, from North America and Asia to Europe, and now the Middle East, sets a strong foundation for impactful research aimed at improving emergency medicine practices and patient outcomes worldwide (Figure 4).

Figure 4. Global collaboration in the FAIR-EC Initiative, more than 20 contributions from multinational research networks across North America, Asia, Europe, and the Middle East. FAIR-EC: Fair, Accountable, Interpretable, and Responsible-Emergency Care.



North America

The initiative is bolstered by contributions from institutions in the United States, including Duke University, Columbia University, Harvard Medical School's Brigham and Women's Hospital, Cornell University, Northwestern University, University of Minnesota, and University of California San Francisco. These institutions have made available substantial patient records, with Duke University alone contributing data on more than 400,000 patients. In addition, the University of British Columbia and the University of Toronto represent our Canadian cohorts.

Asia

This study's Asian contingent features Singapore's Duke-National University of Singapore Medical School, Kangar Kerbau Women's and Children's Hospital, National University Health System, Lee Kong Chian School of Medicine, and Agency for Science, Technology and Research, with Singapore General Hospital providing access to over 1.7 million patient records. South Korea is represented by Samsung Medical Center, Yonsei University, and Hallym University Medical Center, which add additional patient records to this study's database. Japanese patients are also represented by Kyoto University. The First Affiliated Hospital of Sun Yat-Sen University extends this study's reach into China, enriching the diversity of data sources.

Europe, Australia, and the Middle East

In Europe, the University of Oxford and King's College London in the United Kingdom, University College Dublin in Ireland, Karolinska Institutet and Karolinska University Hospital in Sweden, and Provincial Specialist Hospital in Poland contribute their expertise, while Monash University

and the University of Sydney extend this study's reach to Australia. Notably, the University of Haifa in Israel represents the Middle East, adding to this study's geographical and cultural diversity. This inclusion ensures that the FAIR-EC study encompasses a wide range of health care systems and patient demographics, offering a comprehensive perspective on EC practices worldwide.

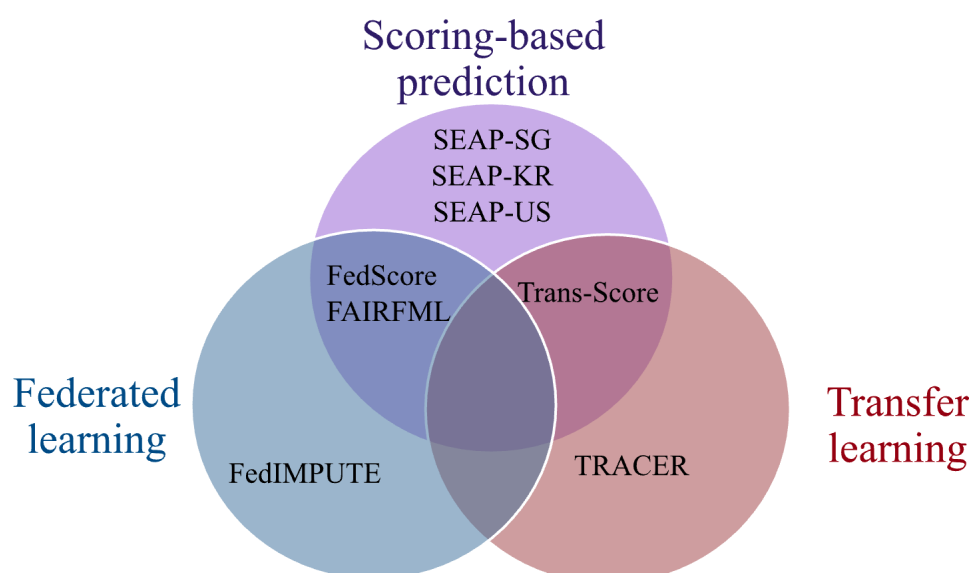
Results

All individual sites participating in this initiative (if their data are provided for collaboration) will obtain the necessary ethical approvals to ensure that the research complies with local and international ethical standards. Since its inception in September 2022 (funded under the Duke/Duke-NUS Collaboration Pilot Grant period September 1, 2022, to August 31, 2023), FAIR-EC has initiated a series of projects that exemplify the practical application of advanced data science techniques in EC, grounded in the principles of international collaboration (Figure 5). Among these, a federated scoring system is being developed to enable cross-institutional association and predictive modeling without compromising patient data privacy, illustrating the potential of federated learning in health care [33-35]. For this protocol submission (January 2026), participating sites had identified retrospective EC datasets totaling >2 million records, including US cohorts (eg, Duke Health ED encounters from calendar year 2019 across 3 hospitals for admission prediction and federated-imputation pilot work) and large regional datasets in Asia (eg, Singapore General Hospital with >1.7 million patient records), as well as additional international partners. As of January 2026, cross-site data harmonization and federated workflow implementation had been initiated, with local IRB or ethics review at participating

sites in progress. Concurrently, a project is underway to evaluate the generalizability of the inpatient admission prediction models originally developed in Singapore across multiple international sites, with initial cross-site benchmarking analyses expected to be submitted in mid-late 2026. This initiative assesses the accuracy of the Score for Emergency Admission Prediction-Singapore and its adaptations, the Score for Emergency Admission Prediction-Korea and the Score for Emergency Admission Prediction-United States, in diverse health care settings. Another notable effort involves the use of transfer learning to adapt these Singapore-developed models for use at Duke University Hospital, tailoring the models to better fit the specific patient demographics and care practices of the hospital environment (FedScore [48,49,75] and FairFML [33]). Additionally, the current study is exploring dynamic transfer learning to ensure that

these association and predictive models can be continuously updated over time, maintaining their relevance and accuracy in the face of evolving health care practices and patient populations (TRACER and Trans-Score, WO2025221209A1 [76]). Moreover, a federated-transfer imputation method has been proposed to handle missing data across different datasets, further enhancing the robustness and applicability of the federated models (FedIMPUTE). Collectively, these projects reflect the types of collaborative efforts enabled by the FAIR-EC infrastructure. While not part of a single analytic protocol, they demonstrate the feasibility of federated modeling, transfer learning, and harmonized data strategies across sites. Project-level papers are planned for submission in 2027, aligned with the completion of harmonized extracts, federated analyses, and site-specific ethics milestones.

Figure 5. Ongoing projects enabled by the FAIR-EC study. FAIR-EC: Fair, Accountable, Interpretable, and Responsible-Emergency Care; SEAP-KR: Score for Emergency Admission Prediction-Korea; SEAP-SG: Score for Emergency Admission Prediction-Singapore; SEAP-US: Score for Emergency Admission Prediction-United States.



Discussion

Principal Findings

FAIR-EC is dedicated to integrating ethical AI principles into the development and application of data science techniques within EC. By leveraging advanced methodologies such as federated learning and transfer learning, this study aims to enhance association and predictive modeling, improve patient stratification, and ensure equitable care across diverse health care settings. The significance of FAIR-EC lies in its pioneering approach to addressing complex challenges in EC through data science. By focusing on the ethical application of AI [14], this current study not only seeks to improve patient outcomes but also to set a new standard for the responsible use of technology in health care. The international and interdisciplinary collaboration at the heart of FAIR-EC underscores the global relevance of its mission, highlighting the universal need for improved EC solutions.

FAIR-EC boasts a broad geographical representation and a rich diversity of data, enabling a comprehensive analysis of EC practices and outcomes. The adoption of federated learning and other privacy-preserving techniques ensures the ethical handling of patient data, fostering trust and collaboration among participating institutions. Moreover, the collaboration's focus on transfer learning facilitates the adaptation of models across different settings, enhancing their applicability and impact.

The variability in data quality and completeness across sites poses challenges to model development and validation. Additionally, the complex nature of health care data and the evolving landscape of emergency medicine require ongoing adjustments to modeling approaches. At present, the network is primarily composed of well-resourced sites, limiting our capacity to engage lower-resource settings. To address these issues, FAIR-EC plans to establish localized data analytical teams at each site, enhancing our federated model with site-specific expertise and resources, and exploring mentorship mechanisms to support future partners in less-resourced

environments. Future efforts will also focus on developing more robust mechanisms for model evaluation and on exploring new areas where data science can contribute to EC, such as operational efficiency and patient experience.

In assessing the use of multimodal data in EC, it is crucial to acknowledge the inherent biases in data collection. Availability and depth of data—ranging from structured codified data to imaging and genotype data—vary significantly worldwide. Some regions have advanced systems capable of extensive data generation, while others might lack these resources. This disparity not only impacts data quality but also the applicability of advanced data-driven methodologies, potentially reinforcing health inequities.

Implementing challenges include technological integration, workflow understanding and integration, staff training, and the potential resistance from health care professionals accustomed to traditional methods of care delivery. Addressing these issues would require comprehensive education and demonstration of the model's efficacy and benefits.

The pioneering projects initiated by the FAIR-EC study facilitate further innovation in EC. By leveraging the feasibility and effectiveness of federated learning, transfer learning, and dynamic model updating in real-world health care settings, these projects lay the groundwork for future research endeavors. As these initial projects validate the use of advanced data science techniques across different geographies and health care systems, they are expected to inspire a broader adoption of these approaches, facilitating more personalized, efficient, and equitable EC solutions. Furthermore, the successful implementation and outcomes of these projects will likely attract participation from additional sites and foster stronger interdisciplinary collaborations, expanding the scope and impact of FAIR-EC. This momentum is anticipated to generate a self-reinforcing cycle of innovation, where each success story paves the way for new questions, hypotheses, and projects, ultimately driving continuous improvement in EC practices worldwide.

To ensure that these efforts translate into tangible scientific contributions, the FAIR-EC leadership has established a

coordinated publication pipeline, with writing teams, target journals, and timelines prespecified for each project. Lead authors are assigned at project kickoff, and each publication is expected to include investigators from multiple institutions to reflect FAIR-EC's collaborative spirit. Dedicated paper development workshops, internal peer review processes, and milestone tracking within the collaboration platform will support timely and high-quality dissemination. The first set of papers is expected to be submitted in early to mid-2026, with additional publications planned as subsequent analyses mature.

Conclusion

FAIR-EC is a pioneering initiative that applies data science ethically, equitably, and effectively in EC. By leveraging federated learning, transfer learning, and interpretable AI, FAIR-EC establishes a scalable and privacy-preserving framework for collaborative model development across diverse health care settings. While challenges related to data variability, health care system complexity, and cross-institutional harmonization persist, this initiative lays a strong foundation for ongoing innovation in global EC and medicine.

Moving forward, FAIR-EC commits to three key promises: (1) expanding the network by integrating more diverse institutions and health care systems to improve model generalizability and equity; (2) enhancing methodological rigor through continuous refinement of federated, transfer, and explainable learning approaches to improve clinical utility; and (3) ensuring real-world impact by translating research findings into actionable clinical tools that empower emergency physicians, policymakers, and health care providers.

With these commitments, FAIR-EC aspires to set a new global standard for fair, accountable, and interpretable AI applications in EC—bridging the gap between cutting-edge data science and real-world patient care.

Acknowledgments

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading and editing. The GenAI tool used was OpenAI. Responsibility for this final paper lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by CH. Additional note: ChatGPT was used only for light language editing (eg, spelling and grammar checks, and minor wording improvements) using short, generic prompts (eg, “fix spelling/grammar and propose rephrasing for language improvement”) applied to text primarily written by the authors. All scientific content, study design decisions, interpretation, and the final wording were determined by the authors. No protected or identifiable patient data were entered into the tool. Any citations and factual statements were verified by the authors using primary sources.

Funding

This research protocol was funded by the Duke/Duke–National University of Singapore (NUS) Collaboration Pilot Grant titled “Privacy-Preserving, Adaptive Federated and Swarm Learning: Emergency Department Triage Decision-Making across Two International Healthcare Systems” (Duke/Duke-NUS/RECA(Pilot)/2022/0064). This project was awarded US\$100,000 to the Duke budget and S\$100,000 (US \$79,180) to the Duke-NUS budget, with a project period starting September 1, 2022, and an initial end date of August 31, 2023. This study was supported under the auspices of the Duke Office of Duke-NUS

Affairs and the Duke-NUS Research Affairs Department. Funding covered activities related to infrastructure development, federated learning implementation, and preliminary modeling. The funders of this study had no role in this study's design, data collection, data analysis, data interpretation, or writing of this report.

Data Availability

Due to the federated and privacy-preserving design of the FAIR-EC (Fair, Accountable, Interpretable and Responsible-Emergency Care) initiative, individual-level patient data are not centralized or shared outside of the participating institutions. Each site retains full control over its own data in accordance with local ethical approvals, regulatory requirements, and institutional governance. However, summary statistics, model artifacts (eg, global weights and performance metrics), and analysis protocols developed through FAIR-EC will be shared through a secure collaboration platform, with appropriate safeguards to prevent reidentification. Participating institutions interested in replicating analyses or contributing new data may do so via the FAIR-EC onboarding process. In addition, code used for federated learning, data harmonization, and model evaluation—including modules for FedScore, TRACER, FedIMPUTE, and FairFML—will be made available via open-source repositories (eg, GitHub) upon publication of corresponding results, in accordance with FAIR-EC's commitment to transparency and reproducibility.

Conflicts of Interest

TB is a part of the Editorial Board of Resuscitation Plus, a member of the statutory board in the Pre-Hospital Emergency Care Council, Ireland, and receives funding from the Health Research Board, Ireland. ALB receives funding for research related to out-of-hospital cardiac arrest and cardiopulmonary resuscitation from the Laerdal Foundation and the National Institutes of Health's National Heart, Lung, and Blood Institute, and the National Institutes of Health's National Institute on Minority Health and Health Disparities. ALB receives in-kind support from World Point, a manikin manufacturer. ALB is a BIRCWH scholar (K12AR084231). TD is an Editor of Resuscitation Plus and receives research funding from the Laerdal Foundation. AK is a consultant and cofounder for UCSF (University of California, San Francisco) spinout, Capture Dx, and supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under (K23HD110716). This information or content and conclusions are those of the author and should neither be construed as the official position or policy of, nor should any endorsements be inferred by, HRSA (Health Resources and Services Administration), HHS (Department of Health & Human Services), or the US government. MEHO is the Scientific Advisor of TIIM Healthcare SG.

Multimedia Appendix 1

Motivations behind the FAIR-EC framework. FAIR-EC: Fair, Accountable, Interpretable, and Responsible-Emergency Care. [\[DOCX File \(Microsoft Word File\), 22 KB-Multimedia Appendix 1\]](#)

References

1. Razzak J, Beecroft B, Brown J, Hargarten S, Anand N. Emergency care research as a global health priority: key scientific opportunities and challenges. *BMJ Glob Health*. Jul 2019;4(Suppl 6):e001486. [doi: [10.1136/bmjgh-2019-001486](https://doi.org/10.1136/bmjgh-2019-001486)]
2. Andrew E, Nehme Z, Cameron P, Smith K. Drivers of increasing emergency ambulance demand. *Prehosp Emerg Care*. May 3, 2020;24(3):385-393. [doi: [10.1080/10903127.2019.1635670](https://doi.org/10.1080/10903127.2019.1635670)]
3. Darraj A, Hudays A, Hazazi A, Hobani A, Alghamdi A. The association between emergency department overcrowding and delay in treatment: a systematic review. *Healthcare (Basel)*. Jan 29, 2023;11(3):385. [doi: [10.3390/healthcare11030385](https://doi.org/10.3390/healthcare11030385)]
4. Butun A, Kafdag EE, Gunduz H, et al. Emergency department overcrowding: causes and solutions. *Emerg Crit Care Med*. 2023;3(4):171-176. [doi: [10.1097/EC9.0000000000000078](https://doi.org/10.1097/EC9.0000000000000078)]
5. Huang J, Pang WS, Mak FY, et al. Factors associated with adoption of Electronic Health Record Sharing System (eHRSS) among private physicians. *Health Policy Technol*. Mar 2025;14(2):100995. [doi: [10.1016/j.hlpt.2025.100995](https://doi.org/10.1016/j.hlpt.2025.100995)]
6. Islam KR, Prithula J, Kumar J, et al. Machine learning-based early prediction of sepsis using electronic health records: a systematic review. *JCM*. Aug 30, 2023;12(17):5658. [doi: [10.3390/jcm12175658](https://doi.org/10.3390/jcm12175658)]
7. Obermeyer Z, Lee TH. Lost in thought - the limits of the human mind and the future of medicine. *N Engl J Med*. Sep 28, 2017;377(13):1209-1211. [doi: [10.1056/NEJMp1705348](https://doi.org/10.1056/NEJMp1705348)] [Medline: [28953443](https://pubmed.ncbi.nlm.nih.gov/28953443/)]
8. Zhou D, Zhang Y, Sonabend-W A, Wang Z, Lu J, Cai T. Federated offline reinforcement learning. *J Am Stat Assoc*. Oct 2024;119(548):3152-3163. [doi: [10.1080/01621459.2024.2310287](https://doi.org/10.1080/01621459.2024.2310287)]
9. Deliu N, Williams JJ, Chakraborty B. Reinforcement learning in modern biostatistics: constructing optimal adaptive interventions. *Int Statistical Rev*. Dec 2025;93(3):385-424. [doi: [10.1111/insr.12583](https://doi.org/10.1111/insr.12583)]
10. Nazer LH, Zatarah R, Waldrip S, et al. Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digit Health*. Jun 2023;2(6):e0000278. [doi: [10.1371/journal.pdig.0000278](https://doi.org/10.1371/journal.pdig.0000278)] [Medline: [37347721](https://pubmed.ncbi.nlm.nih.gov/37347721/)]

11. Moresky RT, Razzak J, Reynolds T, et al. Advancing research on emergency care systems in low-income and middle-income countries: ensuring high-quality care delivery systems. *BMJ Glob Health*. 2019;4(Suppl 6):e001265. [doi: [10.1136/bmjgh-2018-001265](https://doi.org/10.1136/bmjgh-2018-001265)] [Medline: [31406599](https://pubmed.ncbi.nlm.nih.gov/31406599/)]
12. Okada Y, Ning Y, Ong MEH. Explainable artificial intelligence in emergency medicine: an overview. *Clin Exp Emerg Med*. Dec 2023;10(4):354-362. [doi: [10.15441/ceem.23.145](https://doi.org/10.15441/ceem.23.145)] [Medline: [38012816](https://pubmed.ncbi.nlm.nih.gov/38012816/)]
13. Ning Y, Teixayavong S, Shang Y, et al. Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. *Lancet Digit Health*. Nov 2024;6(11):e848-e856. [doi: [10.1016/S2589-7500\(24\)00143-2](https://doi.org/10.1016/S2589-7500(24)00143-2)] [Medline: [39294061](https://pubmed.ncbi.nlm.nih.gov/39294061/)]
14. Ahun E, Demir A, Yiğit Y, et al. Perceptions and concerns of emergency medicine practitioners about artificial intelligence in emergency triage management during the pandemic: a national survey-based study. *Front Public Health*. Oct 26, 2023;11:1285390. [doi: [10.3389/fpubh.2023.1285390](https://doi.org/10.3389/fpubh.2023.1285390)]
15. Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: a narrative review. *Heliyon*. Feb 29, 2024;10(4):e26297. [doi: [10.1016/j.heliyon.2024.e26297](https://doi.org/10.1016/j.heliyon.2024.e26297)] [Medline: [38384518](https://pubmed.ncbi.nlm.nih.gov/38384518/)]
16. Holliman JJ. International emergency medicine. In: *An Emergency Physician's Path*. Springer, Cham; 2023:109-114. [doi: [10.1007/978-3-031-47873-4_16](https://doi.org/10.1007/978-3-031-47873-4_16)]
17. Hsu SL, Tsai KT, Tan TH, et al. Interdisciplinary collaboration and computer-assisted home healthcare referral in the emergency department: a retrospective cohort study. *Aging Clin Exp Res*. Aug 2022;34(8):1939-1946. [doi: [10.1007/s40520-022-02109-9](https://doi.org/10.1007/s40520-022-02109-9)]
18. Wu J, Liu X, Li M, et al. Clinical text datasets for medical artificial intelligence and large language models — a systematic review. *NEJM AI*. May 23, 2024;1(6):AIra2400012. [doi: [10.1056/AIra2400012](https://doi.org/10.1056/AIra2400012)]
19. Puskarich MA, Callaway C, Silbergleit R, et al. Priorities to overcome barriers impacting data science application in emergency care research. *Acad Emerg Med*. Jan 2019;26(1):97-105. [doi: [10.1111/acem.13520](https://doi.org/10.1111/acem.13520)]
20. Harron K, Wade A, Gilbert R, Muller-Pebody B, Goldstein H. Evaluating bias due to data linkage error in electronic healthcare records. *BMC Med Res Methodol*. Mar 5, 2014;14(1):36. [doi: [10.1186/1471-2288-14-36](https://doi.org/10.1186/1471-2288-14-36)] [Medline: [24597489](https://pubmed.ncbi.nlm.nih.gov/24597489/)]
21. Queralt-Rosinach N, Kaliyaperumal R, Bernabé CH, et al. Applying the FAIR principles to data in a hospital: challenges and opportunities in a pandemic. *J Biomed Semantics*. Apr 25, 2022;13(1):12. [doi: [10.1186/s13326-022-00263-7](https://doi.org/10.1186/s13326-022-00263-7)] [Medline: [35468846](https://pubmed.ncbi.nlm.nih.gov/35468846/)]
22. Inau ET, Sack J, Waltemath D, Zeleke AA. Initiatives, concepts, and implementation practices of the findable, accessible, interoperable, and reusable data principles in health data stewardship: scoping review. *J Med Internet Res*. Aug 28, 2023;25:e45013. [doi: [10.2196/45013](https://doi.org/10.2196/45013)] [Medline: [37639292](https://pubmed.ncbi.nlm.nih.gov/37639292/)]
23. Lauridsen S. Emergency care, triage, and fairness. *Bioethics*. Jun 2020;34(5):450-458. [doi: [10.1111/bioe.12689](https://doi.org/10.1111/bioe.12689)] [Medline: [31721247](https://pubmed.ncbi.nlm.nih.gov/31721247/)]
24. Blackburn J, Ousey K, Goodwin E. Information and communication in the emergency department. *Int Emerg Nurs*. Jan 2019;42:30-35. [doi: [10.1016/j.ienj.2018.07.002](https://doi.org/10.1016/j.ienj.2018.07.002)] [Medline: [30122462](https://pubmed.ncbi.nlm.nih.gov/30122462/)]
25. Dhillon G, Zhang Z, Grewal H, Kashyap R. Editorial: clinical application of artificial intelligence in emergency and critical care medicine, volume IV. *Front Med*. 2023;10:1346070. [doi: [10.3389/fmed.2023.1346070](https://doi.org/10.3389/fmed.2023.1346070)]
26. Lavoie B, Bourque CJ, Côté AJ, et al. The responsibility to care: lessons learned from emergency department workers' perspectives during the first wave of the COVID-19 pandemic in Canada. *Can J Emerg Med*. Aug 2022;24(5):482-492. [doi: [10.1007/s43678-022-00306-z](https://doi.org/10.1007/s43678-022-00306-z)]
27. Dehon E, Weiss N, Jones J, Faulconer W, Hinton E, Sterling S. A systematic review of the impact of physician implicit racial bias on clinical decision making. *Acad Emerg Med*. Aug 2017;24(8):895-904. [doi: [10.1111/acem.13214](https://doi.org/10.1111/acem.13214)] [Medline: [28472533](https://pubmed.ncbi.nlm.nih.gov/28472533/)]
28. Rubio-Navarro A, Garcia-Capilla DJ, Torralba-Madrid MJ, Ruttu J. Ethical, legal and professional accountability in emergency nursing practice: an ethnographic observational study. *Int Emerg Nurs*. Sep 2019;46:100777. [doi: [10.1016/j.ienj.2019.05.003](https://doi.org/10.1016/j.ienj.2019.05.003)] [Medline: [31331840](https://pubmed.ncbi.nlm.nih.gov/31331840/)]
29. Schilling S, Armaou M, Morrison Z, Carding P, Bricknell M, Connelly V. Understanding teamwork in rapidly deployed interprofessional teams in intensive and acute care: A systematic review of reviews. *PLOS ONE*. 2022;17(8):e0272942. [doi: [10.1371/journal.pone.0272942](https://doi.org/10.1371/journal.pone.0272942)] [Medline: [35980893](https://pubmed.ncbi.nlm.nih.gov/35980893/)]
30. Gorick H. Factors that affect nurses' triage decisions in the emergency department: a literature review. *Emerg Nurse*. May 3, 2022;30(3):14-19. [doi: [10.7748/en.2022.e2123](https://doi.org/10.7748/en.2022.e2123)] [Medline: [35224917](https://pubmed.ncbi.nlm.nih.gov/35224917/)]
31. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. Oct 25, 2019;366(6464):447-453. [doi: [10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342)] [Medline: [31649194](https://pubmed.ncbi.nlm.nih.gov/31649194/)]
32. Nikolić M, Nikolić D, Stefanović M, Koprivica S, Stefanović D. Mitigating algorithmic bias through probability calibration: a case study on lead generation data. *Mathematics*. Jul 3, 2025;13(13):2183. [doi: [10.3390/math13132183](https://doi.org/10.3390/math13132183)]

33. Li S, Wu Q, Zhou D, et al. FairFML: fair federated machine learning with a case study on reducing gender disparities in cardiac arrest outcome prediction. *NPJ Health Syst.* Oct 7, 2024;2(1):29. [doi: [10.1038/s44401-025-00035-2](https://doi.org/10.1038/s44401-025-00035-2)]
34. Busuioc M. Accountable artificial intelligence: holding algorithms to account. *Public Adm Rev.* Sep 2021;81(5):825-836. [doi: [10.1111/puar.13293](https://doi.org/10.1111/puar.13293)]
35. Novelli C, Taddeo M, Floridi L. Accountability in artificial intelligence: what it is and how it works. *AI Soc.* Aug 2024;39(4):1871-1882. [doi: [10.1007/s00146-023-01635-y](https://doi.org/10.1007/s00146-023-01635-y)]
36. Barsotti C. Commentary: leveraging accountability for advocacy. *Ann Emerg Med.* Nov 2019;74(5):S69-S70. [doi: [10.1016/j.annemergmed.2019.08.462](https://doi.org/10.1016/j.annemergmed.2019.08.462)]
37. DeMarco JP, Jones GE, Daly BJ. *Ethical and Legal Issues in Nursing*. Broadview Press; 2019. ISBN: 9781554813964
38. Agarwal S, Mishra S. *Responsible AI*. Springer Nature; 2021. ISBN: 9783030768591
39. Kohane I. Policy in progress — the race to frame AI in health care. *NEJM AI.* Feb 22, 2024;1(3):Aie2400044. [doi: [10.1056/Aie2400044](https://doi.org/10.1056/Aie2400044)]
40. Hosseini MM, Hosseini STM, Qayumi K, Ahmady S, Koohestani HR. The aspects of running artificial intelligence in emergency care; a scoping review. *Arch Acad Emerg Med.* 2023;11(1):e38. [doi: [10.22037/aaem.v11i1.1974](https://doi.org/10.22037/aaem.v11i1.1974)] [Medline: [37215232](https://pubmed.ncbi.nlm.nih.gov/37215232/)]
41. Dubber MD, Pasquale F, Das S. *The Oxford Handbook of Ethics of AI*. Oxford University Press; 2020. ISBN: 9780190067397
42. Doctor NE, Ahmad NSB, Pek PP, Yap S, Ong MEH. The Pan-Asian Resuscitation Outcomes Study (PAROS) clinical research network: what, where, why and how. *Singapore Med J.* Jul 2017;58(7):456-458. [doi: [10.11622/smedj.2017057](https://doi.org/10.11622/smedj.2017057)] [Medline: [28741005](https://pubmed.ncbi.nlm.nih.gov/28741005/)]
43. Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl.* Jan 2023;82(3):3713-3744. [doi: [10.1007/s11042-022-13428-4](https://doi.org/10.1007/s11042-022-13428-4)]
44. Radford A, WuJ, ChildR, LuanD, AmodeiD, Sutskever I. Language models are unsupervised multitask learners. OpenAI; 2019. URL: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf [Accessed 2026-02-18]
45. Li S, Yan M, Yuan R, Liu M, Liu N, Hong C. FedIMPUTE: privacy-preserving missing value imputation for multi-site heterogeneous electronic health records. *J Biomed Inform.* May 2025;165:104780. [doi: [10.1016/j.jbi.2025.104780](https://doi.org/10.1016/j.jbi.2025.104780)] [Medline: [40054590](https://pubmed.ncbi.nlm.nih.gov/40054590/)]
46. Kairouz P, McMahan HB, Avenet B. Advances and open problems in federated learning. *Foundations and Trends in Machine Learning.* Jun 23, 2021;14(1-2):1-210. [doi: [10.1561/22000000083](https://doi.org/10.1561/22000000083)]
47. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng.* 2010;22(10):1345-1359. [doi: [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191)]
48. Li S, Ning Y, Ong MEH, et al. FedScore: a privacy-preserving framework for federated scoring system development. *J Biomed Inform.* Oct 2023;146:104485. [doi: [10.1016/j.jbi.2023.104485](https://doi.org/10.1016/j.jbi.2023.104485)]
49. Li S, Wang Z, Shang Y, et al. Developing federated time-to-event scores using heterogeneous real-world survival data. *Comput Biol Med.* Oct 2025;197(Pt B):111084. [doi: [10.1016/j.combiomed.2025.111084](https://doi.org/10.1016/j.combiomed.2025.111084)] [Medline: [40976210](https://pubmed.ncbi.nlm.nih.gov/40976210/)]
50. Li S, Miao D, Wu Q, et al. Federated learning in healthcare: a benchmark comparison of engineering and statistical approaches for structured data analysis. *Health Data Sci.* 2024;4:0196. [doi: [10.34133/hds.0196](https://doi.org/10.34133/hds.0196)] [Medline: [39635226](https://pubmed.ncbi.nlm.nih.gov/39635226/)]
51. Li S, Liu P, Nascimento GG, et al. Federated and distributed learning applications for electronic health records and structured medical data: a scoping review. *J Am Med Inform Assoc.* Nov 17, 2023;30(12):2041-2049. [doi: [10.1093/jamia/ocad170](https://doi.org/10.1093/jamia/ocad170)]
52. Januszewski AS, O'Connell RL, Li L, Sullivan DR, Jenkins AJ, Keech AC. Online tool for cross-sectional and longitudinal comparison of socio-economic status indices based on postcodes in Australia. *Intern Med J.* Aug 2025;55(8):1360-1366. [doi: [10.1111/imj.70117](https://doi.org/10.1111/imj.70117)] [Medline: [40540631](https://pubmed.ncbi.nlm.nih.gov/40540631/)]
53. Earnest A, Ong MEH, Shahidah N, Chan A, Wah W, Thumboo J. Derivation of indices of socioeconomic status for health services research in Asia. *Prev Med Rep.* 2015;2:326-332. [doi: [10.1016/j.pmedr.2015.04.018](https://doi.org/10.1016/j.pmedr.2015.04.018)]
54. Cai T, Liu M, Xia Y. Individual data protected integrative regression analysis of high-dimensional heterogeneous data. *J Am Stat Assoc.* Oct 2, 2022;117(540):2105-2119. [doi: [10.1080/01621459.2021.1904958](https://doi.org/10.1080/01621459.2021.1904958)]
55. Hong C, Wang Y, Cai T. A divide-and-conquer method for sparse risk prediction and evaluation. *Biostatistics.* Apr 13, 2022;23(2):397-411. [doi: [10.1093/biostatistics/kxaa031](https://doi.org/10.1093/biostatistics/kxaa031)]
56. Voss EA, Makadia R, Matcho A, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc.* May 1, 2015;22(3):553-564. [doi: [10.1093/jamia/ocu023](https://doi.org/10.1093/jamia/ocu023)]
57. 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.* Jul 1, 2014;21(4):578-582. [doi: [10.1136/amiajnl-2014-002747](https://doi.org/10.1136/amiajnl-2014-002747)]

58. Murphy SN, Mendis ME, Berkowitz DA, Kohane I, Chueh HC. Integration of clinical and genetic data in the i2b2 architecture. *AMIA Annu Symp Proc.* 2006;2006:1040. [Medline: [17238659](#)]
59. Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics.* Feb 15, 2020;36(4):1234-1240. [doi: [10.1093/bioinformatics/btz682](#)]
60. Zhou D, Gan Z, Shi X, et al. Multiview incomplete knowledge graph Integration with application to cross-institutional EHR data harmonization. *J Biomed Inform.* Sep 2022;133:104147. [doi: [10.1016/j.jbi.2022.104147](#)] [Medline: [35872266](#)]
61. Alsentzer E, Murphy J, Boag W, et al. Publicly available clinical BERT embeddings. In: *Proceedings of the 2nd Clinical Natural Language Processing Workshop.* 2019:72-78. [doi: [10.18653/v1/W19-1909](#)]
62. Choi I, Long Q, Getzen E. Filling the gaps: leveraging large language models for temporal harmonization of clinical text across multiple medical visits for clinical prediction. *medRxiv.* Preprint posted online on May 7, 2024. [doi: [10.1101/2024.05.06.24306959](#)] [Medline: [38765975](#)]
63. Yang D, Zhou D, Cai S, et al. Robust automated harmonization of heterogeneous data through ensemble machine learning: algorithm development and validation study. *JMIR Med Inform.* 2025;13:e54133-e54133. [doi: [10.2196/54133](#)]
64. Hong C, Rush E, Liu M, et al. Clinical Knowledge Extraction Via Sparse Embedding Regression (KESER) with multi-center large scale electronic health record data. *Health Inf.* Preprint posted online on 2021. [doi: [10.1101/2021.03.13.21253486](#)]
65. Lee SJ, Rho M. Multimodal deep learning applied to classify healthy and disease states of human microbiome. *Sci Rep.* Jan 17, 2022;12(1):824. [doi: [10.1038/s41598-022-04773-3](#)]
66. Wang WC, Ahn E, Feng D, Kim J. A review of predictive and contrastive self-supervised learning for medical images. *Mach Intell Res.* Aug 2023;20(4):483-513. [doi: [10.1007/s11633-022-1406-4](#)]
67. Gao C, Zhu J, Zhang F, Wang Z, Li X. A novel representation learning for dynamic graphs based on graph convolutional networks. *IEEE Trans Cybern.* Jun 2023;53(6):3599-3612. [doi: [10.1109/TCYB.2022.3159661](#)] [Medline: [35333735](#)]
68. Li S, Li X, Yu K, et al. Bridging data gaps in healthcare: a scoping review of transfer learning in structured data analysis. *Health Data Sci.* 2025;5:0321. [doi: [10.34133/hds.0321](#)] [Medline: [40917080](#)]
69. Hong C, Liu M, Wojdyla DM, Hickey J, Pencina M, Henao R. Trans-balance: reducing demographic disparity for prediction models in the presence of class imbalance. *J Biomed Inform.* Jan 2024;149:104532. [doi: [10.1016/j.jbi.2023.104532](#)] [Medline: [38070817](#)]
70. Chong A, Phua SX, Xiao Y, Ng WY, Li HY, Goh WWB. Establishing the foundations for a data-centric ai approach for virtual drug screening through a systematic assessment of the properties of chemical data. *Elife.* Dec 17, 2024. [doi: [10.7554/eLife.97821.1](#)]
71. Canadian Index of Multiple Deprivation 2021. Halifax Data, Mapping and Analytics Hub. URL: <https://data-hrm.hub.arcgis.com/datasets/de9edfea914645399104df70473f6c31/about> [Accessed 2026-02-18]
72. PROGRESS-plus. The Cochrane Collaboration. URL: <https://methods.cochrane.org/equity/projects/evidence-equity/progress-plus> [Accessed 2026-02-18]
73. Personal Data Protection Commission Singapore. URL: <https://www.pdpc.gov.sg/about/the-legislation/pdpa-overview> [Accessed 2026-04-28]
74. Health Insurance Portability and Accountability Act of 1996 (HIPAA). URL: <https://www.cdc.gov/phlp/php/resources/health-insurance-portability-and-accountability-act-of-1996-hipaa.html> [Accessed 2026-04-28]
75. Wu Q, Li S, Miao D, Shang Y, Li X, Liu N. Evaluating the efficacy of federated scoring systems with heterogeneous electronic health records. The Second Tiny Papers Track at ICLR; 2024. URL: <https://openreview.net/pdf?id=c4GVRbEx1g> [Accessed 2026-02-18]
76. Liu N, Li S, Ong EHM, Hong C, inventors; Patient risk level estimation. WO2025221209A1. Oct 23, 2025. URL: <https://patents.google.com/patent/WO2025221209A1/en?q=WO+Patent+2025221209A1> [Accessed 2026-02-18]

Abbreviations

AI: artificial intelligence
AUROC: area under the receiver operating characteristic curve
BERT: Bidirectional Encoder Representations From Transformers
EC: emergency care
ED: emergency department
EHR: electronic health record
FAIR-EC: Fair, Accountable, Interpretable, and Responsible-Emergency Care
GCN: graph convolutional network
ICD: *International Statistical Classification of Diseases*

IRB: institutional review board

LLM: large language model

PROGRESS: Place of Residence; Race, Ethnicity, Culture, and Language; Occupation; Gender or Sex; Religion; Education; Socioeconomic Status; Social Capital

SNOMED CT : Systematized Nomenclature of Medicine – Clinical Terms

WMA : World Medical Association

Edited by Javad Sarvestan; peer-reviewed by Enze Bai, Varada Vivek Khanna; submitted 22.Mar.2025; final revised version received 22.Jan.2026; accepted 23.Jan.2026; published 04.Aug.2026

Please cite as:

Hong C, Liew JCK, Yu J, Barry T, Blewer AL, Buckland DM, Cai T, Cha WC, Chakraborty B, Chen W, Cheng J, Chong SL, Djärv T, Earnest A, Engelhard M, Fan X, Feng M, Feng J, Fu H, Goh WWB, Goldstein BA, Gronsbell J, Ho AFW, Ho K, Iwami T, Joiner A, Kornblith A, Li S, Lim SL, Liu M, Liu Z, Lu L, Luo Y, Ng YY, Ning Y, Okada Y, Park JO, Park YR, Razzak J, Shen Y, Siddiqui FJ, Steel PAD, Tan KBK, Teixayavong S, Vakulenko-Lagun B, Vissoci JRN, Waligora G, Wang F, Wang H, Wang H, Wong AKI, Xie F, Yang J, Zhang Y, Zhou D, Zhou L, Zhu T, Neumar R, Page D, Vaughan R, Ong MEH, Liu N *Building a Global Research Network for Fair, Accountable, Interpretable, and Responsible AI in Emergency Care: Protocol for a FAIR-EC Study*

JMIR Res Protoc 2026;15:e74202

URL: <https://www.researchprotocols.org/2026/1/e74202>

doi: [10.2196/74202](https://doi.org/10.2196/74202)

© Chuan Hong, Jonathan Chong Kai Liew, Jaeyong Yu, Tomás Barry, Audrey L. Blewer, Daniel M. Buckland, Tianrun Cai, Won Chul Cha, Bibhas Chakraborty, Wei Chen, Jun Cheng, Shu-Ling Chong, Therese Djärv, Arul Earnest, Matthew Engelhard, Xiuyi Fan, Mengling Feng, Jean Feng, Huazhu Fu, Wilson Wen Bin Goh, Benjamin A. Goldstein, Jessica Gronsbell, Andrew Fu Wah Ho, Kendall Ho, Taku Iwami, Anjni Joiner, Aaron Kornblith, Siqi Li, Shir Lynn Lim, Molei Liu, Zhenghong Liu, Lei Lu, Yuan Luo, Yih Yng Ng, Yilin Ning, Yohei Okada, Ju Ok Park, Yu Rang Park, Junaid Razzak, Yuzeng Shen, Fahad Javaid Siddiqui, Peter A D Steel, Kenneth Boon Kiat Tan, Salinelat Teixayavong, Bella Vakulenko-Lagun, Joao Ricardo Nickenig Vissoci, Grzegorz Waligora, Fei Wang, Haibo Wang, Haoyuan Wang, An-Kwok Ian Wong, Feng Xie, Jie Yang, Yiye Zhang, Doudou Zhou, Li Zhou, Tingting Zhu, Robert Neumar, David Page, Michael Pencina, Roger Vaughan, Marcus Eng Hock Ong, Nan Liu. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 04.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.