# FAIR-EC: A Global Research Network for Fair, Accountable, Interpretable, and Responsible AI in Emergency Care

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, Yong Chen, Jun Cheng, Shu-Ling Chong, Therese Djärv, Arul Earnest, Matthew Engelhard, Xiuyi Fan, Jean Feng, Mengling 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 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

Submitted to: JMIR Research Protocols
on: March 22, 2025

# *Table of Contents*

# FAIR-EC: A Global Research Network for Fair, Accountable, Interpretable, and Responsible AI in Emergency Care

Chuan Hong[1] PhD; Jonathan Chong Kai Liew[2, 3]; Jaeyong Yu[4]; Tomás Barry[5]; Audrey L. Blewer[6, 7, 8]; Daniel M. Buckland[9]; Tianrun Cai[10]; Won Chul Cha[11]; Bibhas Chakraborty[2]; Wei Chen[12]; Yong Chen[13]; Jun Cheng[14]; Shu-Ling Chong[15]; Therese Djärv[16, 17]; Arul Earnest[18]; Matthew Engelhard[1]; Xiuyi Fan[19, 20]; Jean Feng[21]; Mengling Feng[22]; Huazhu Fu[23]; Wilson Wen Bin Goh[19, 24, 25]; Benjamin A. Goldstein[1]; Jessica Gronsbell[26]; Andrew Fu Wah Ho[2, 27, 28]; Kendall Ho[29, 30, 31]; Taku Iwami[32]; Anjni Joiner[9, 33]; Aaron Kornblith[21]; Siqi Li[2]; Shir Lynn Lim[2, 22, 34]; Molei Liu[35]; Zhenghong Liu[27]; Lei Lu[36]; Yuan Luo[37]; Yih Yng Ng[19]; Yilin Ning[2]; Yohei Okada[2, 32]; Ju Ok Park[38, 39]; Yu Rang Park[40]; Junaid Razzak[41, 42]; Yuzeng Shen[27]; Fahad Javaid Siddiqui[2]; Peter Steel[42]; Kenneth Boon Kiat Tan[27]; Salinelat Teixayavong[43]; Bella Vakulenko-Lagun[44]; Joao Ricardo Nickenig Vissoci[9]; Grzegorz Waligora[45, 46]; Fei Wang[47]; Haibo Wang[48]; Haoyuan Wang[1]; An-Kwok Ian Wong[9]; Feng Xie[49]; Jie Yang[10]; Yiye Zhang[42]; Doudou Zhou[22]; Li Zhou[10]; Tingting Zhu[50]; Robert Neumar[51]; David Page[9]; Michael Pencina[9]; Roger Vaughan[52]; Marcus Eng Hock Ong[2, 27]; Nan Liu[2]

[1] Department of Biostatistics and Bioinformatics Duke University School of Medicine Durham US
[2] Duke-NUS Medical School Singapore SG
[3] School of Chemistry, Chemical Engineering and Biotechnology Nanyang Technological University Singapore SG
[4] Hallym University Chuncheon KR
[5] University College Dublin Dublin IE
[6] Department of Population Health Sciences, School of Medicine Duke University Durham US
[7] Department of Family Medicine and Community Health, School of Medicine Duke University Durham US
[8] Health Services and Systems Research SingHealth Duke-NUS Academic Medical Centre Singapore SG
[9] Duke University Durham US
[10] Harvard Medical School & Brigham and Women's Hospital Harvard University Boston US
[11] Samsung Medical Center Seoul KR
[12] The University of Sydney Sydney AU
[13] University of Pennsylvania Philadelphia US
[14] Agency for Science, Technology and Research Singapore SG
[15] KK Women's and Children's Hospital Singapore SG
[16] Karolinska University Hospital Stockholm SE
[17] Karolinska Institutet Stockholm SE
[18] Monash University Melbourne AU
[19] Lee Kong Chian School of Medicine Nanyang Technological University Singapore SG
[20] College of Computing and Data Science Nanyang Technological University Singapore SG
[21] University of California, San Francisco San Francisco US
[22] National University of Singapore Singapore SG
[23] Institute of High Performance Computing Agency for Science, Technology and Research Singapore SG
[24] Center of AI in Medicine Nanyang Technological University Singapore SG
[25] Division of Neurology, Department of Brain Sciences Faculty of Medicine Imperial College London London GB
[26] University of Toronto Toronto CA
[27] Singapore General Hospital Singapore SG
[28] Centre for Population Health Research and Implementation SingHealth Singapore SG
[29] University of British Columbia Vancouver CA
[30] Digital Emergency Medicine Unit University of British Columbia Vancouver CA
[31] BC Ministry of Health Vancouver CA
[32] Kyoto University Kyoto JP
[33] Health Services and Systems Research Duke-NUS Medical School Singapore SG
[34] National University Heart Centre Singapore Singapore SG
[35] Columbia University New York City US
[36] King's College London London GB
[37] Northwestern University Evanston US
[38] College of Medicine Hallym University Chuncheon KR
[39] Hallym University Dongtan Sacred Heart Hospital Chuncheon KR
[40] Yonsei University Seoul KR

[41] Aga Khan University Karachi PK
[42] Cornell University Ithaca US
[43] Nanyang Technological University Singapore SG
[44] University of Haifa Haifa IL
[45] Gda?sk Medical University Gda?sk PL
[46] Wroclaw Medical University Wroc?aw PL
[47] Weill Cornell Medicine New York City US
[48] Research Centre of Big Data and Artificial Intelligence for Medicine The First Affiliated Hospital, Sun Yat-sen University Guangzhou CN
[49] University of Minnesota Minneapolis US
[50] University of Oxford Oxford GB
[51]Department of Emergency Medicine Michigan Medicine University of Michigan–Ann Arbor Ann Arbor US
[52] Penn Medicine University of Pennsylvania Philadelphia US

**Corresponding Author:**
Nan Liu

Duke-NUS Medical School
8 College Road Singapore
Singapore
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 in data-sharing, which often result in a lack of transparency and robust oversight in these applications.

**Objective:** The Fair, Accountable, Interpretable and Responsible (FAIR)-EC collaboration adapts the existing FAIR principles to address emerging challenges as data science integrates with EC. This initiative aims to transform EC by establishing ethical artificial intelligence (AI) 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 that enables analyses of 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 healthcare systems.

**Results:** The FAIR-EC initiative has facilitated the collection and analysis of datasets from diverse geographical regions, enabling the examination of regional variations in EC practices. Initial projects have demonstrated promising outcomes, including the successful development of a federated scoring system and the adaptation of association studies and predictive models across various regions. These efforts highlight the feasibility of leveraging advanced data science techniques to address the complexities of EC while preserving patient privacy.

**Conclusions:** FAIR-EC integrates data science ethically and effectively into EC, addressing challenges like fragmented data, real-time handoffs, and public health crises. Its federated design harmonizes diverse data streams while preserving privacy, and its emphasis on ethical AI 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.

(JMIR Preprints 22/03/2025:74202)
DOI: https://doi.org/10.2196/preprints.74202

## Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
- ✓ **Please make my preprint PDF available to anyone at any time (recommended).**
- Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
- Only make the preprint title and abstract visible.
- No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

No. Please do not make my accepted manuscript PDF available to anyone.

# Original Manuscript

# FAIR-EC: A Global Research Network for Fair, Accountable, Interpretable, and Responsible AI in Emergency Care

## 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 in data-sharing, which often result in a lack of transparency and robust oversight in these applications.
**Objective:** The Fair, Accountable, Interpretable and Responsible (FAIR)-EC collaboration adapts the existing FAIR principles to address emerging challenges as data science integrates with EC. This initiative aims to transform EC by establishing ethical artificial intelligence (AI) 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 that enables analyses of 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 healthcare systems.
**Results:** The FAIR-EC initiative has facilitated the collection and analysis of datasets from diverse geographical regions, enabling the examination of regional variations in EC practices. Initial projects have demonstrated promising outcomes, including the successful development of a federated scoring system and the adaptation of association studies and predictive models across various regions. These efforts highlight the feasibility of leveraging advanced data science techniques to address the complexities of EC while preserving patient privacy.
**Conclusions:** FAIR-EC integrates data science ethically and effectively into EC, addressing challenges like fragmented data, real-time handoffs, and public health crises. Its federated design harmonizes diverse data streams while preserving privacy, and its emphasis on ethical AI 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.
**Trial Registration:**
**Keywords:** Emergency Care; Emergency Medicine; Pre-hospital ICU; Federated Learning

## 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 pre-hospital care (e.g., ambulances), emergency departments (EDs), intensive care units and beyond (e.g., 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 healthcare 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 U.S., 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 post-acute recovery. This includes optimizing pre-hospital and ED management, expediting inpatient admissions, and integrating data across hospital wards, skilled nursing facilities, rehabilitation centers, and outpatient follow-up clinics. By leveraging electronic health records (EHRs) and complementary data sources —such as secure messaging, wearable devices, and home monitoring—EC can be better coordinated to provide a comprehensive, longitudinal view of patient outcomes.

## Integrating Data Science and Artificial Intelligence (AI) in EC

Clinical decision-making in EC has traditionally relied on physicians' experience, supplemented by heuristic decision-support tools. However, in high-pressure environments like overcrowded EDs, this approach can be inconsistent due to variations in adherence to evidence-based frameworks. AI-based clinical support algorithms offer a solution by synthesizing vast amounts of data, reducing cognitive biases, and ensuring more standardized, data-driven decisions. Unlike human decision-making, AI models are not affected by fatigue or distractions, enabling more accurate and reliable clinical assessments, particularly in time-sensitive scenarios. The widespread adoption of EHRs in EC presents opportunities to enhance decision-making through AI-driven analysis.[5,6] By leveraging structured data and real-time monitoring, AI models can predict conditions like sepsis before clinical symptoms appear, optimizing workflows and improving patient outcomes,[7] though their real-world implementation remains limited.[8–10]

## 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.[11] 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.[12] Transparency and interpretability remain major challenges, limiting clinician trust and adoption.[11,13,14] Insufficient accountability and regulatory oversight further risk the ethical and responsible use of AI in EC.[15] Ethical concerns, such as patient autonomy and privacy, also require more rigorous safeguards.[16] International, interdisciplinary collaboration presents a critical opportunity to bridge these gaps.[17–20] 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,[29] integration challenges, privacy concerns, incomplete data, and real-time processing constraints.[30–33] To address these challenges, AI implementation must adhere to the core principles of fairness, accountability, interpretability, and responsibility (FAIR).[34,35] *Fairness* ensures equitable outcomes, preventing algorithmic bias and promoting equal care across diverse patient populations.[36] Methods such as data rebalancing,[37] threshold recalibration,[38] and performance monitoring[39] can help mitigate biases.[36] *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.[40,41,42,43] *Interpretability* is crucial for clinical adoption, allowing healthcare providers to understand AI-driven recommendations and ensuring that AI supports rather than dictates clinical care.[13,44,45] *Responsibility* entails the ethical use of AI, ensuring patient safety, privacy, and dignity while enhancing decision-making.[46–50] For further details, please refer to the supplementary materials. Integrating these principles into AI-driven EC systems is critical for ensuring ethical, effective, and trustworthy healthcare.

## FAIR-EC Network

The Fair, Accountable, Interpretable and Responsible-Emergency Care (FAIR-EC) 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, healthcare infrastructure, and resources across regions, evaluating data-driven solutions in diverse settings—including low-resource environments—is essential. The Pan-Asian Resuscitation Outcomes Study (PAROS) [21] 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.

For AI models to be optimally effective and minimally biased, they ideally should be trained and validated on data from the patient population in which they will be deployed. However, when such data is unavailable or insufficient, leveraging data from other systems can provide valuable insights—provided that models are carefully adapted to account for differences in population characteristics and healthcare practices. FAIR-EC facilitates this by enabling collaborative learning strategies that prioritize both population-specific adaptability and cross-institutional knowledge transfer.

Unlike traditional global studies that require data sharing, FAIR-EC is designed as a flexible, collaborative research network that enables institutions to participate in ways that suit their capabilities. Institutions can contribute through centralized data sharing, federated learning, or methodological and clinical expertise. Even if only a few sites share data, the broader network remains engaged, fostering bidirectional collaboration where data scientists refine analytical methods while physicians provide real-world challenges that guide research priorities. Data harmonization is voluntary, with guidance and support available for those interested.

## 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 (PCORNET) Common Data Model (CDM), Observational Medical Outcomes Partnership (OMOP), and Fast Healthcare Interoperability Resources (FHIR) standards, which provide structured, interoperable representations of patient records, diagnoses, interventions, and outcomes. Additionally, ontology mapping and terminology harmonization using Unified Medical Language Systems (UMLS) and SNOMED CT ensure that medical concepts are consistently interpreted across institutions.

To further unify data across institutions, FAIR-EC leverages Natural Language Processing[22] and Large Language Models (LLMs)[23] for cross-institutional data harmonization. Bidirectional Encoder Representations from Transformers (BERT) -based models are employed 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.

## Federated Learning and Transfer Learning for Model Development

The FAIR-EC initiative integrates federated learning and transfer learning to ensure robust and scalable predictive modeling across diverse healthcare settings. Federated learning[24] 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.

To enhance generalizability, FAIR-EC applies transfer learning techniques.[25,26–28] The SEAP 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 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.[51,52] 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.

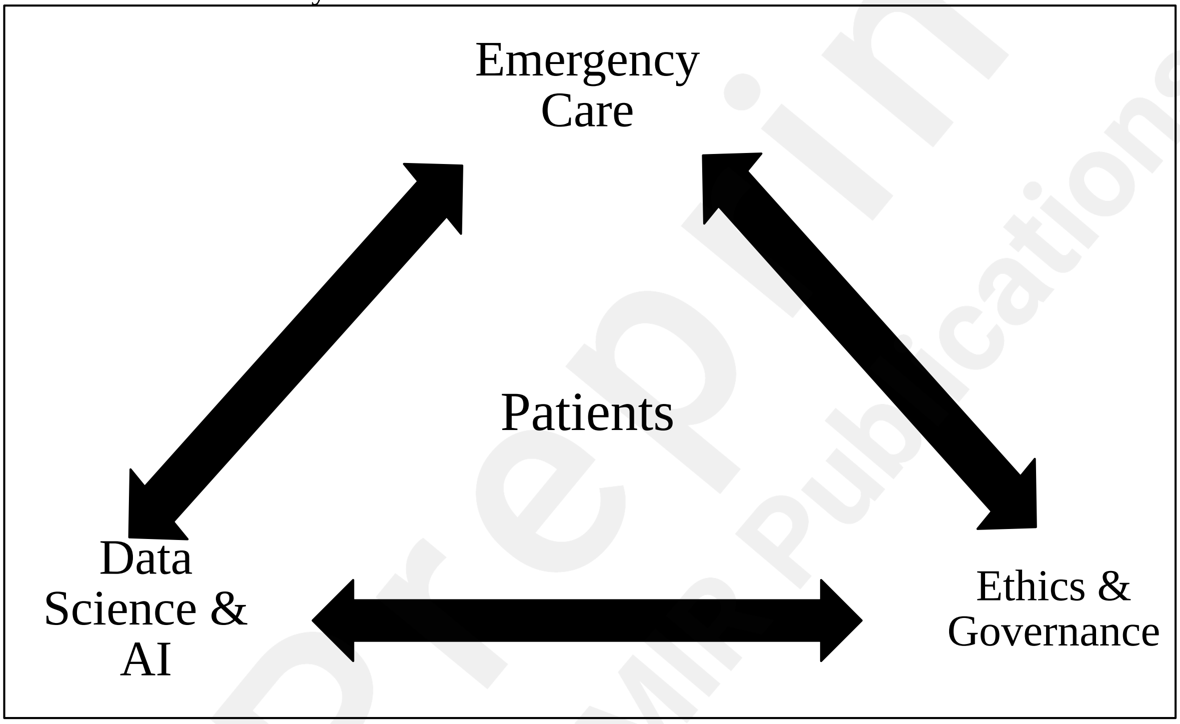
### *Participating Sites and Study Cohorts*

FAIR-EC will utilize 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.

### 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.




### Collaborative Platforms

A digital collaboration platform (Figure 2) will be utilized 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 the project.

### Cross-Disciplinary Teams

The study will organize participants into cross-disciplinary teams, each tasked with addressing specific aspects of the 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.

### Communication Channels and Workshops

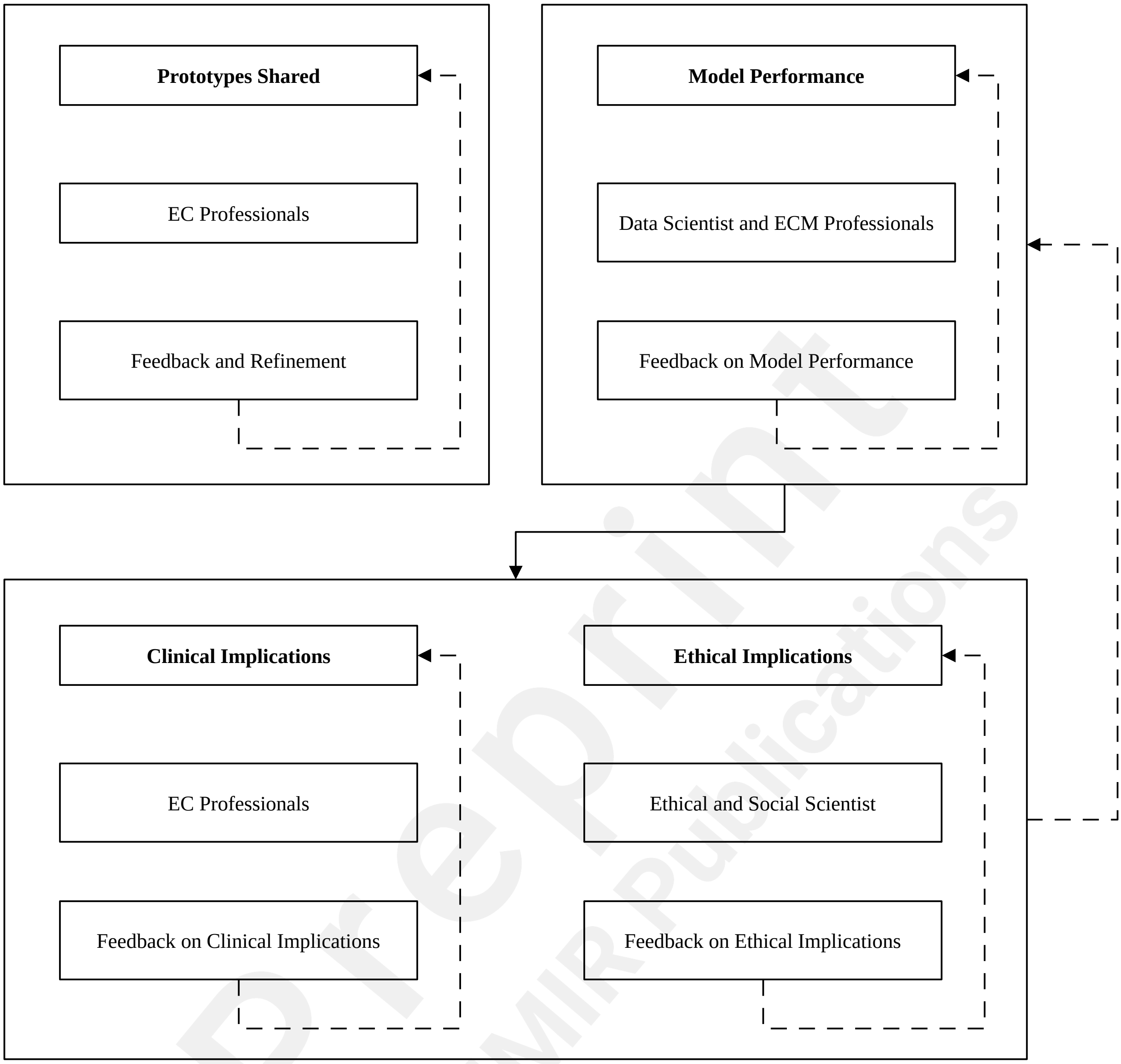
Scheduled virtual and in-person meetings when feasible will ensure ongoing dialogue between all team members. These meetings will serve as forums for progress updates, brainstorming sessions, and resolving any challenges that arise during the collaboration. The FAIR-EC collaboration will organize workshops 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 employed. 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 utilize feedback from various stakeholder groups to inform and enhance the presentation of risk estimates. If different subgroups, such as healthcare 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 three focus groups, one for each type of stakeholder, drawing from the same participant pool as the initial phase.

Figure 2. Iterative feedback mechanisms and stakeholder engagement.




## 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 healthcare 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 the 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 Q&A 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.

## Model Development

### Multimodal Data

We utilize data from two distinct modalities: structured codified data and unstructured data. Structured data encompasses quantifiable and easily searchable elements like demographics, diagnosis codes, medications, laboratory tests and results, procedures, and genomic/genetic data (if available). Some social determinants of health (SDOH), are also included in this category, enabling the systematic analysis of trends in diagnoses and treatments (e.g., IRSD index in Australia[53] and SEDI/SAD indices in Singapore[54]). Such variables should be carefully selected and organized to keep track of socio-economic status over time.

In contrast, unstructured data provides richer context and nuance. This includes clinical narratives (progress notes, discharge summaries, radiology reports) that offer insights into patient symptoms and treatments not captured by structured fields. It also covers narrative accounts of SDOH, imaging data like plain radiography and advanced diagnostic imaging (e.g. computed tomography and magnetic resonance imaging) for rapid diagnosis, and waveform data (e.g., electrocardiograms, 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, the 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.[55,56] 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.

### Planned Data Harmonization and Knowledge Network

Participating sites will utilize common data standards such as the OMOP,[57] PCORNET,[58] and Informatics for Integrating Biology & the Bedside (i2b2).[59] 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 healthcare 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, the study proposes utilizing the established BERT architecture for its proven effectiveness in understanding contextual relationships within text,[60–62] and augment it with newer transformer models such as GPT-4, for enhanced data harmonization and metadata creation.[63] 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 CDMs, 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 (EMRs) for research purposes, though these technically feasible implementations still await regulatory approval.

To effectively manage the diverse multi-modal and multi-source datasets, we propose to implement sophisticated algorithms and tools designed to aggregate and process from various modalities. Specifically, we are employing our previously developed semantic learning approaches,[64,65] state-of-the-art multimodal deep learning architectures[66] and contrastive learning[67] to derive meaningful representations from these large datasets without the need for phenotypic (labeled) data, which is often scarce in healthcare 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[68] that incorporates a prototype graph. 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 employing 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.

### 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).

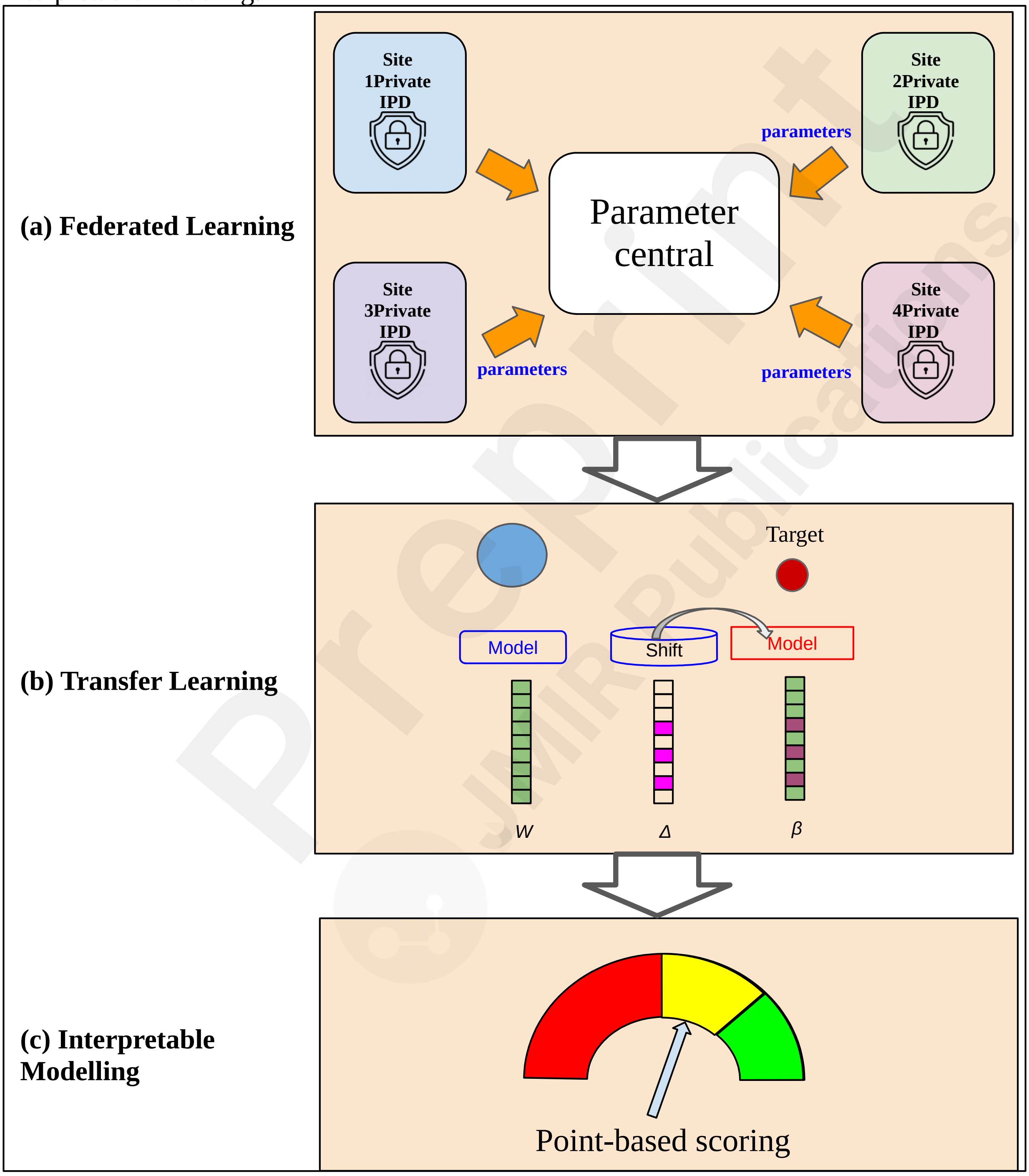
### Transfer Learning

Transfer learning involves taking a pre-existing model developed for one task or from one population and adapting it for another, related task or population.[69] In the context of EC, this includes leveraging knowledge and patterns learned from one healthcare 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.[70]

### Interpretable Models

Ensuring the interpretability of models is a cornerstone of the FAIR-EC study. Interpretable models are essential for gaining the trust of healthcare practitioners and patients alike, as they provide transparent insights into how decisions or predictions are made.[71] This transparency is crucial for clinical adoption, as it allows EC professionals to understand and validate the recommendations provided by the models, ensuring that these recommendations can be integrated seamlessly into clinical workflows.[72] Moreover, interpretable models facilitate accountability and responsibility by making it possible to trace and understand the outcome of each decision, thereby supporting ethical decision-making in EC settings. In our previous work, we have developed point-based scoring systems that quantify risk factors and treatment outcomes in a clear and accessible manner.[73,74]

Figure 3. Framework for model development under federated learning: transfer learning, and interpretable modeling.




## *Equity Considerations, Ethical Implications and Engagement Plan*

Integrating equity considerations into EC models is essential to ensure equitable health outcomes and to uphold ethical healthcare practices. This integration involves developing models specifically designed to handle data incompleteness and biases across diverse populations. It also includes implementing methodologies that actively adjust for disparities in data availability and quality. By doing so, we can mitigate the risk of perpetuating or worsening health disparities through technological applications.

Furthermore, 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 is 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, healthcare providers, and healthcare 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 electronic health records. 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*

The 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 non-inferior performance.[75] 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*

At the heart of the FAIR-EC's evaluation strategy is a deep-seated commitment to fairness and equity. The study will employ 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 (e.g., Canadian Index of Multiple Deprivation[76]) and equity rubric lists (e.g., PROGRESS Plus)[77] will be utilized to systematically evaluate how algorithms perform across various groups defined by age, gender, ethnicity, socioeconomic status (SES), 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, healthcare providers, healthcare 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 utilize 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, the study will explore and implement a range of strategies designed to ensure the fairness and equity of the developed models. These strategies may include re-balancing datasets, employing 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.

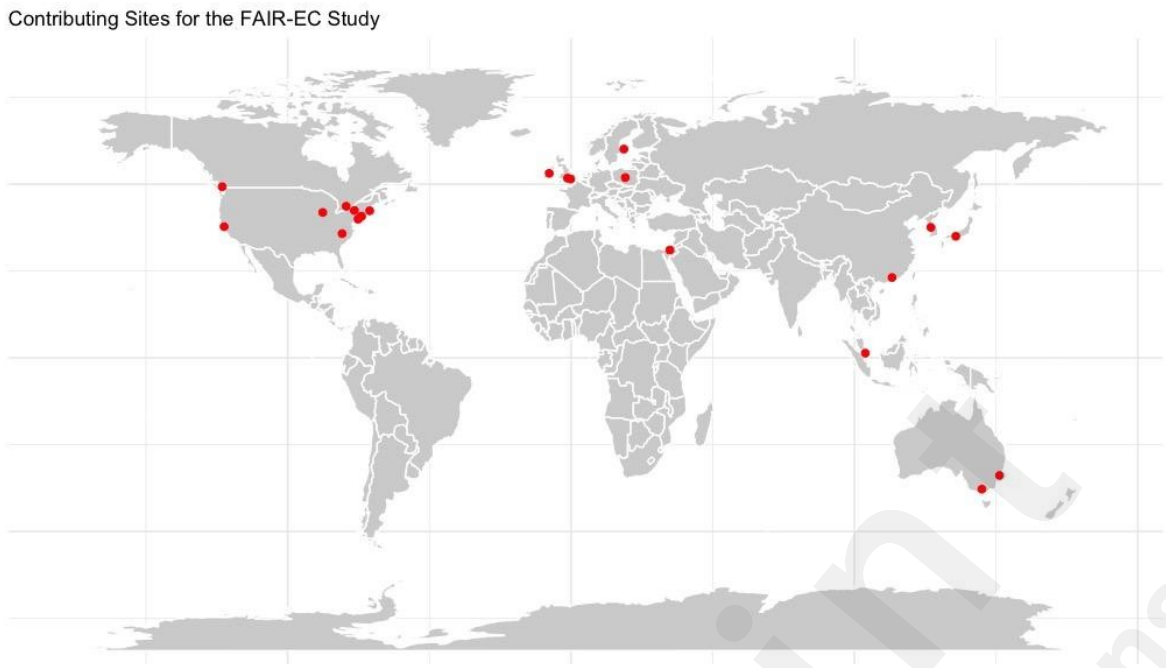
## Results

The FAIR-EC study has made significant strides, starting with establishing a comprehensive network of participating institutions across the globe.

## Participating Institutions and Teams

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.




### North America

The initiative is bolstered by contributions from institutions in the United States, including Duke University, Columbia University, Harvard Medical School & Brigham and Women's Hospital, Cornell University, Northwestern University, University of Pennsylvania, 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

The study's Asian contingent features Singapore's Duke-NUS Medical School, KK Women's and Children's Hospital, National University Health System, Lee Kong Chian School of Medicine, and Agency for Science, Technology and Research (A*STAR), with Singapore General Hospital (SGH) 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 the study's database. Japanese patients are also represented by Kyoto University. The First Affiliated Hospital of Sun Yat-Sen University extends the study's reach into China, enriching the diversity of data sources.

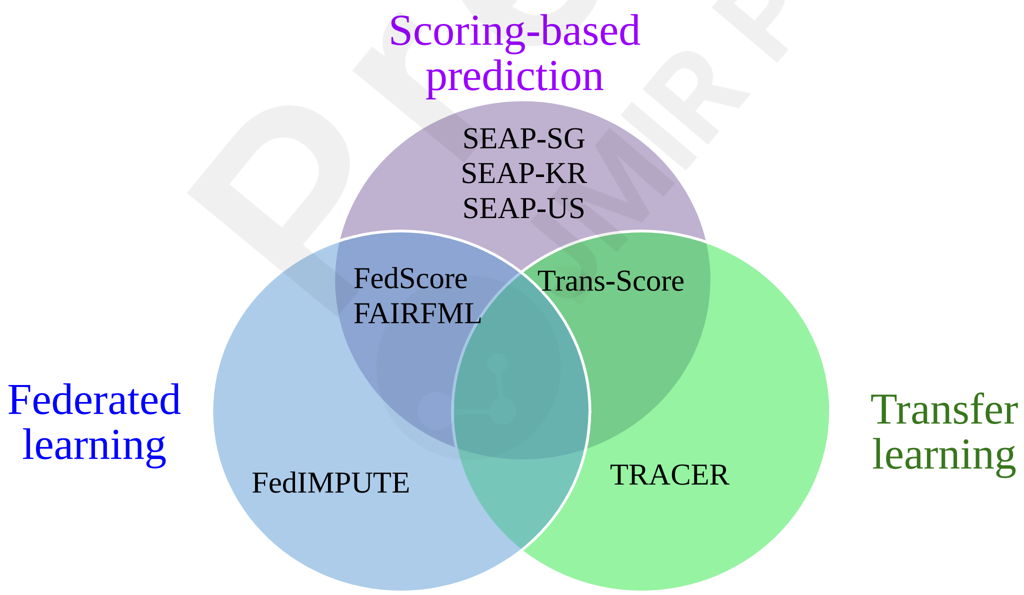
### Europe, Australia, and the Middle East

In Europe, the University of Oxford and King's College London in the UK, 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 the study’s reach to Australia. Notably, the University of Haifa in Israel represents the Middle East, adding to the study's geographical and cultural diversity. This inclusion ensures that the FAIR-EC study encompasses a wide range of healthcare systems and patient demographics, offering a comprehensive perspective on EC practices globally.

## Ongoing Projects Enabled by FAIR-EC

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. 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 healthcare.[26–28] Concurrently, a project is underway to evaluate the generalizability of the inpatient admission prediction models originally developed in Singapore across multiple international sites. This initiative assesses the accuracy of the Score for Emergency Admission Prediction-Singapore (SEAP-SG) and its adaptations, the SEAP-Korea (SEAP-KR) and the SEAP-United States (SEAP-US), in diverse healthcare 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 & FAIRFML). Additionally, the 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 healthcare practices and patient populations (TRACER). 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 underscore FAIR-EC's commitment to leveraging global insights and cutting-edge methodologies to enhance EC delivery worldwide.

Figure 5. Ongoing projects enabled by the FAIR-EC study.




## Discussion

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, the study aims to enhance association and predictive modeling, improve patient stratification, and ensure equitable care across diverse healthcare 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 AI14, the study not only seeks to improve patient outcomes but also to set a new standard for the responsible use of technology in healthcare. 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.

Despite its strengths, FAIR-EC faces several challenges. One challenge we anticipate is that not all equity issues translate globally. For example, race and ethnicity are significant in the U.S. but less so elsewhere. Therefore, we propose initially focusing on universally relevant indicators like SES or social deprivation to more consistently address disparities across countries. The variability in data quality and completeness across sites poses challenges to model development and validation. Additionally, the complex nature of healthcare data and the evolving landscape of emergency medicine require ongoing adjustments to modeling approaches. 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. 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. Looking ahead, FAIR-EC plans to expand its network of collaborating institutions, further diversify its data sources, and refine its methodologies to address these challenges. 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's 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 globally. 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 the proposed models within healthcare systems presents several challenges that warrant thorough consideration. The integration of complex data-driven frameworks into existing healthcare infrastructures involves not only technological adaptations but also significant changes in workflow and staff training. Healthcare providers must be equipped to understand and interact with the system effectively, ensuring that the model's outputs are used appropriately to enhance patient care. Additionally, the potential resistance from healthcare professionals accustomed to traditional methods of care delivery must be considered and addressed through comprehensive education and demonstration of the model's efficacy and benefits.

The pioneering projects initiated by the FAIR-EC study are set to facilitate further innovation in EC. By leveraging the feasibility and effectiveness of federated learning, transfer learning, and dynamic model updating in real-world healthcare 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 healthcare 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 globally.

## 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 healthcare settings. While challenges related to data variability, healthcare 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 healthcare 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 healthcare 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.

## Acknowledgements

This work was supported by the Duke/Duke-NUS Collaboration grant. The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

## 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.

AB 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 Institue on Minority Health and Health Disparities. AB receives in kind support from World Point, a manikin manufacturer. AB is a BIRCWH scholar K12AR084231.

TD is an Editor of Resuscitation Plus and receives research funding from Laerdal Foundation.

AK is a consultant and co-founder for Capture Dx and supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under award number K23HD110716 (AK). This information or content and conclusions are those of the author and should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS or the U.S. Government.

YO is a recipient of the KPFA, Fukuda, IMRF and JSPS Scholarships and the Zoll Foundation Research Grant.

MEHO is the Scientific Advisor of TIIM Healthcare SG.

## Abbreviations

AI: artificial intelligence
BERT: Bidirectional Encoder Representations from Transformers
CDM: Common Data Model
EHR: electronic health records
EC: emergency care
ED: emergency department
FAIR: fair, accountable, interpretable and responsible
FAIR-EC: Fair, Accountable, Interpretable and Responsible-Emergency Care
GCN: Graph Convolutional Networks
LLM: Large Language Model
OMOP: Observational Medical Outcomes Partnership
PCORNET: Patient-Centered Outcomes Research Network
SES: socioeconomic status
SDOH: social determinants of health

## References

1. Sakr, M., and Wardrope, J. (2000). Casualty, accident and emergency, or emergency medicine, the evolution. Emerg. Med. J. *17*, 314–319.

2. Andrew, E., Nehme, Z., Cameron, P., and Smith, K. (2020). Drivers of Increasing Emergency Ambulance Demand. Prehosp. Emerg. Care *24*. https://doi.org/10.1080/10903127.2019.1635670.

3. Bandiera, G., Gaunt, K., Sinclair, D., and Trafford, A. (2014). Emergency department overcrowding and long wait times: taking a corporate approach to improving patient flow. Healthc. Q. *17*. https://doi.org/10.12927/hcq.2015.24115.

4. Darraj, A., Hudays, A., Hazazi, A., Hobani, A., and Alghamdi, A. (2023). The Association between Emergency Department Overcrowding and Delay in Treatment: A Systematic Review. Healthc. Pap. *11*, 385.

5. Kruse, C.S., Kothman, K., Anerobi, K., and Abanaka, L. (2016). Adoption Factors of the Electronic Health Record: A Systematic Review. JMIR medical informatics *4*. https://doi.org/10.2196/medinform.5525.

6. EHR adopters vs. non-adopters: Impacts of, barriers to, and federal initiatives for EHR adoption (2014). Healthc. Pap. *2*, 33–39.

7. Islam, K.R., Prithula, J., Kumar, J., Tan, T.L., Reaz, M.B.I., Sumon, M.S.I., and Chowdhury, M.E.H. (2023). Machine Learning-Based Early Prediction of Sepsis Using Electronic Health Records: A Systematic Review. Journal of Clinical Medicine *12*, 5658.

8. Obermeyer, Z., and Lee, T.H. (2017). Lost in Thought — The Limits of the Human Mind and the Future of Medicine. https://doi.org/10.1056/NEJMp1705348.

9. Zhou, D., Zhang, Y., Sonabend-W, A., Wang, Z., Lu, J., and Cai, T. (2024). Federated Offline Reinforcement Learning. Journal of the American Statistical Association. https://doi.org/10.1080/01621459.2024.2310287.

10. Deliu, N., Williams, J.J., and Chakraborty, B. (2024). Reinforcement Learning in Modern Biostatistics: Constructing Optimal Adaptive Interventions. International Statistical Review. https://doi.org/10.1111/insr.12583.

11. Nazer, L.H., Zatarah, R., Waldrip, S., Ke, J.X.C., Moukheiber, M., Khanna, A.K., Hicklen, R.S., Moukheiber, L., Moukheiber, D., Ma, H., et al. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS Digital Health *2*, e0000278.

12. Moresky, R.T., Razzak, J., Reynolds, T., Wallis, L.A., Wachira, B.W., Nyirenda, M., Carlo, W.A., Lin, J., Patel, S., Bhoi, S., et al. (2019). Advancing research on emergency care systems in low-income and middle-income countries: ensuring high-quality care delivery systems. BMJ Global Health *4*. https://doi.org/10.1136/bmjgh-2018-001265.

13. Okada, Y., Ning, Y., and Ong, M.E.H. (2023). Explainable artificial intelligence in emergency medicine: an overview. Clinical and experimental emergency medicine *10*. https://doi.org/10.15441/ceem.23.145.

14. Ning, Y., Teixayavong, S., Shang, Y., Savulescu, J., Nagaraj, V., Miao, D., Mertens, M., Ting, D.S.W., Ong, J.C.L., Liu, M., et al. (2024). Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. Lancet Digit. Health *6*, e848–e856.

15. Ahun, E., Demir, A., Yiğit, Y., Tulgar, Y.K., Doğan, M., Thomas, D.T., and Tulgar, S. (2023). 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 *11*, 1285390.

16. Naess, A.-C., Foerde, R., and Steen, P.A. (2001). Patient autonomy in emergency medicine. Med. Health Care Philos. *4*, 71–77.

17. Arnold, J.L. (1999). International emergency medicine and the recent development of emergency medicine worldwide. Ann. Emerg. Med. *33*. https://doi.org/10.1016/s0196-0644(99)70424-5.

18. Al-Shaqsi, S. (2010). Models of International Emergency Medical Service (EMS) Systems. Oman Med. J. *25*. https://doi.org/10.5001/omj.2010.92.

19. James (Jim) Holliman, C. (2023). International Emergency Medicine. An Emergency Physician’s Path, 109–114.

20. Hsu, S.L., Tsai, K.T., Tan, T.H., Ho, C.H., Yang, P.C., Hsu, C.C., Lin, H.J., Hung, S.P., and Huang, C.C. (2022). Interdisciplinary collaboration and computer-assisted home healthcare referral in the emergency department: a retrospective cohort study. Aging Clin. Exp. Res. *34*. https://doi.org/10.1007/s40520-022-02109-9.

21. Doctor, N.E., Ahmad, N.S., Pek, P.P., Yap, S., and Ong, M.E. (2017). The Pan-Asian Resuscitation Outcomes Study (PAROS) clinical research network: what, where, why and how. Singapore medical journal *58*. https://doi.org/10.11622/smedj.2017057.

22. Khurana, D., Koli, A., Khatter, K., and Singh, S. (2022). Natural language processing: state of the art, current trends and challenges. Multimed. Tools Appl. *82*, 3713–3744.

23. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners.

24. Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. (2019). Advances and Open Problems in Federated Learning.

25. Pan, S.J., and Yang, Q. A Survey on Transfer Learning. https://ieeexplore.ieee.org/document/5288526.

26. FedScore: A privacy-preserving framework for federated scoring system development (2023). J. Biomed. Inform. *146*, 104485.

27. Li, S., Shang, Y., Wang, Z., Wu, Q., Hong, C., Ning, Y., Miao, D., Ong, M.E.H., Chakraborty, B., and Liu, N. (2024). Developing Federated Time-to-Event Scores Using Heterogeneous Real-World Survival Data.

28. Li, S., Wu, Q., Li, X., Miao, D., Hong, C., Gu, W., Shang, Y., Okada, Y., Chen, M.H., Yan, M., et al. (2024). FairFML: Fair Federated Machine Learning with a case study on reducing gender disparities in cardiac arrest outcome prediction. arXiv [cs.CY].

29. Wu, J., Liu, X., Li, M., Li, W., Su, Z., Lin, S., Garay, L., Zhang, Z., Zhang, Y., Zeng, Q., et al. (2024). Clinical Text Datasets for Medical Artificial Intelligence and Large Language Models — A Systematic Review. NEJM AI. https://doi.org/10.1056/AIra2400012.

30. Puskarich, M.A., Callaway, C., Silbergleit, R., Pines, J.M., Obermeyer, Z., Wright, D.W., Hsia, R.Y., Shah, M.N., Monte, A.A., Limkakeng, A.T., et al. (2019). Priorities to Overcome Barriers Impacting Data Science Application in Emergency Care Research. Acad. Emerg. Med. *26*, 97–105.

31. Weber, G.M., Adams, W.G., Bernstam, E.V., Bickel, J.P., Fox, K.P., Marsolo, K., Raghavan, V.A., Turchin, A., Zhou, X., Murphy, S.N., et al. (2017). Biases introduced by filtering electronic health records for patients with “complete data.” J. Am. Med. Inform. Assoc. *24*, 1134–1141.

32. Harron, K., Wade, A., Gilbert, R., Muller-Pebody, B., and Goldstein, H. (2014). Evaluating bias

due to data linkage error in electronic healthcare records. BMC Med. Res. Methodol. *14*, 1–10.

33. Kum, H.C., and Ahalt, S. (2013). Privacy-by-Design: Understanding Data Access Models for Secondary Data. AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science *2013*.

34. Queralt-Rosinach, N., Kaliyaperumal, R., Bernabé, C.H., Long, Q., Joosten, S.A., van der Wijk, H.J., Flikkenschild, E.L.A., Burger, K., Jacobsen, A., Mons, B., et al. (2022). Applying the FAIR principles to data in a hospital: challenges and opportunities in a pandemic. J. Biomed. Semantics *13*, 1–19.

35. Inau, E.T., Sack, J., Waltemath, D., and Zeleke, A.A. (2023). Initiatives, Concepts, and Implementation Practices of the Findable, Accessible, Interoperable, and Reusable Data Principles in Health Data Stewardship: Scoping Review. J. Med. Internet Res. *25*, e45013.

36. Lauridsen, S. (2020). Emergency care, triage, and fairness. Bioethics *34*. https://doi.org/10.1111/bioe.12689.

37. Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science. https://doi.org/10.1126/science.aax2342.

38. Kleinberg, J., Mullainathan, S., and Raghavan, M. (2016). Inherent Trade-Offs in the Fair Determination of Risk Scores.

39. Fairness and machine learning https://fairmlbook.org/.

40. Busuioc, M. (2021). Accountable Artificial Intelligence: Holding Algorithms to Account. Public Administration Review *81*, 825–836.

41. Novelli, C., Taddeo, M., and Floridi, L. (2023). Accountability in artificial intelligence: what it is and how it works. AI & SOCIETY *39*, 1871–1882.

42. Barsotti, C. (2019). Commentary: Leveraging Accountability for Advocacy. Ann. Emerg. Med. *74*. https://doi.org/10.1016/j.annemergmed.2019.08.462.

43. DeMarco, J.P., Jones, G.E., and Daly, B.J. (2019). Ethical and Legal Issues in Nursing (Broadview Press).

44. Zhang, Z., Kashyap, R., Su, L., Liu, N., and Meng, Q. (2024). Clinical Application of Artificial Intelligence in Emergency and Critical Care Medicine, Volume IV (Frontiers Media SA).

45. Blackburn, J., Ousey, K., and Goodwin, E. (2019). Information and communication in the emergency department. Int. Emerg. Nurs. *42*, 30–35.

46. Hosseini, M.M., Hosseini, S.T.M., Qayumi, K., Ahmady, S., and Koohestani, H.R. (2023). The Aspects of Running Artificial Intelligence in Emergency Care; a Scoping Review. Arch Acad Emerg Med *11*, e38–e38.

47. Kohane, I. (2024). Policy in Progress — The Race to Frame AI in Health Care. NEJM AI. https://doi.org/10.1056/AIe2400044.

48. Agarwal, S., and Mishra, S. Responsible AI (Springer International Publishing).

49. Oxford Handbook of Ethics of AI Google Books. https://books.google.com/books/about/Oxford_Handbook_of_Ethics_of_AI.html?id=9PQTEAAAQBAJ.

50. Lavoie, B., Bourque, C.J., Côté, A.-J., Rajagopal, M., Clerc, P., Bourdeau, V., Ali, S., Doyon-Trottier, E., Castonguay, V., Fontaine-Pagé, É., et al. (2022). The responsibility to care: lessons learned from emergency department workers' perspectives during the first wave of the COVID-19 pandemic in Canada. CJEM *24*, 482–492.

51. Li, S., Miao, D., Wu, Q., Hong, C., D'Agostino, D., Li, X., Ning, Y., Shang, Y., Wang, Z., Liu, M., et al. (2024). Federated Learning in Healthcare: A Benchmark Comparison of Engineering and Statistical Approaches for Structured Data Analysis. Health Data Science. https://doi.org/10.34133/hds.0196.

52. Website https://academic.oup.com/jamia/article-abstract/30/12/2041/7252875.

53. Walker, R., and Hiller, J.E. (2005). The index of relative socio-economic disadvantage: general population views on indicators used to determine area-based disadvantage. Aust. N. Z. J. Public Health *29*, 442–447.

54. Earnest, A., Ong, M.E.H., Shahidah, N., Chan, A., Wah, W., and Thumboo, J. (2015). Derivation of indices of socioeconomic status for health services research in Asia. Prev Med Rep *2*, 326–332.

55. Hong, C., Wang, Y., and Cai, T. (2020). A divide-and-conquer method for sparse risk prediction and evaluation. Biostatistics *23*, 397–411.

56. Cai, T., Liu, M., and Xia, Y. (2019). Individual Data Protected Integrative Regression Analysis of High-dimensional Heterogeneous Data.

57. Voss, E.A., Makadia, R., Matcho, A., Ma, Q., Knoll, C., Schuemie, M., DeFalco, F.J., Londhe, A., Zhu, V., and Ryan, P.B. (2015). Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. J. Am. Med. Inform. Assoc. *22*. https://doi.org/10.1093/jamia/ocu023.

58. Fleurence, R.L., Curtis, L.H., Califf, R.M., Platt, R., Selby, J.V., and Brown, J.S. (2014). Launching PCORnet, a national patient-centered clinical research network. J. Am. Med. Inform. Assoc. *21*, 578–582.

59. Murphy, S.N., Mendis, M.E., Berkowitz, D.A., Kohane, I., and Chueh, H.C. (2006). Integration of clinical and genetic data in the i2b2 architecture. AMIA Annu. Symp. Proc. *2006*.

60. Alsentzer, E., Murphy, J.R., Boag, W., Weng, W.-H., Jin, D., Naumann, T., and McDermott, M.B.A. (2019). Publicly Available Clinical BERT Embeddings.

61. Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C.H., and Kang, J. (2019). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. https://doi.org/10.1093/bioinformatics/btz682.

62. Zhou, D., Gan, Z., Shi, X., Patwari, A., Rush, E., Bonzel, C.-L., Panickan, V.A., Hong, C., Ho,

Y.-L., Cai, T., et al. (2022). Multiview Incomplete Knowledge Graph Integration with application to cross-institutional EHR data harmonization. J. Biomed. Inform. *133*, 104147.

63. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. (2017). Attention Is All You Need.

64. Hong, C., Rush, E., Liu, M., Zhou, D., Sun, J., Sonabend, A., Castro, V.M., Schubert, P., Panickan, V.A., Cai, T., et al. (2021). Clinical knowledge extraction via sparse embedding regression (KESER) with multi-center large scale electronic health record data. npj Digital Medicine *4*, 1–11.

65. SONAR: Enabling Robust Automated Harmonization of Heterogeneous Data through Ensemble Machine Learning JMIR Preprints. https://preprints.jmir.org/preprint/54133.

66. Lee, S.J., and Rho, M. (2022). Multimodal deep learning applied to classify healthy and disease states of human microbiome. Sci. Rep. *12*, 1–11.

67. A simple framework for contrastive learning of visual representations https://doi.org/10.5555/3524938.3525087.

68. Kipf, T.N., and Welling, M. (2022). Semi-Supervised Classification with Graph Convolutional Networks. In International Conference on Learning Representations.

69. Li, S., Li, X., Yu, K., Miao, D., Zhu, M., Yan, M., Ke, Y., D'Agostino, D., Ning, Y., Wu, Q., et al. (2024). Bridging Data Gaps in Healthcare: A Scoping Review of Transfer Learning in Biomedical Data Analysis.

70. Trans-Balance: Reducing demographic disparity for prediction models in the presence of class imbalance (2024). J. Biomed. Inform. *149*, 104532.

71. Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., and Zhong, C. (2021). Interpretable Machine Learning: Fundamental Principles and 10 Grand Challenges.

72. Abgrall, G., Holder, A.L., Dagdia, Z.C., Zeitouni, K., and Monnet, X. (2024). Should AI models be explainable to clinicians? Critical Care *28*, 301.

73. Xie, F., Chakraborty, B., Ong, M.E.H., Goldstein, B.A., and Liu, N. (2020). AutoScore: A Machine Learning-Based Automatic Clinical Score Generator and Its Application to Mortality Prediction Using Electronic Health Records. JMIR medical informatics *8*. https://doi.org/10.2196/21798.

74. Xie, F., Ong, M.E.H., Liew, J.N.M., Tan, K.B.K., Ho, A.F.W., Nadarajan, G.D., Low, L.L., Kwan, Y.H., Goldstein, B.A., Matchar, D.B., et al. (2021). Development and Assessment of an Interpretable Machine Learning Triage Tool for Estimating Mortality After Emergency Admissions. JAMA Netw Open *4*, e2118467–e2118467.

75. Chong, A., Phua, S.-X., Xiao, Y., Ng, W.Y., Li, H.Y., and Goh, W.W.B. (2024). Establishing the foundations for a data-centric AI approach for virtual drug screening through a systematic assessment of the properties of chemical data. eLife *13*. https://doi.org/10.7554/eLife.97821.2.

76. Halifax Data Mapping and Analytics Hub https://data-hrm.hub.arcgis.com/datasets/de9edfea914645399104df70473f6c31/about.

77. PROGRESS-Plus https://methods.cochrane.org/equity/projects/evidence-equity/progress-plus.

# Supplementary Files

# Multimedia Appendixes

Motivations behind the FAIREC framework.
URL: http://asset.jmir.pub/assets/3cb07c5c18ce0b3038d1b0ab8a89c5ff.docx