PhD Research Diary

Mayank Keoliya

Started: March 28, 2025
Last Updated: April 9, 2025

Contents

Chapter 1

Research Overview

1.1 Research Objectives

- **Primary Research Question:** How can AI improve diagnostic accuracy in healthcare?
- **Key Research Domains:**
  1. Medical image analysis using deep learning
  2. Predictive risk assessment models
  3. Ethical AI in healthcare applications
- **Anticipated Contributions:**
  - Develop novel AI algorithms for medical diagnostics
  - Improve interpretability of medical AI systems
  - Address potential bias in healthcare AI

Chapter 2

Literature Review

The field of Artificial Intelligence in healthcare is rapidly evolving, with significant advances in: (i) predictive diagnostic models [1], (ii) Multimodal data integration, and (iii) Explainable AI techniques.

2.1 AI in Healthcare

2.1.1 Surveys

Here are some survey & position papers. Rajpurkar et al. [2] advocate for *generalist medical AI*, i.e. models which are trained on large, unlabelled, diverse datasets with self-supervision, can flexibly ingest different modalities (e.g. imaging, EHR, genomics) and produce expressive outputs (e.g. free-text explanations, spoken recommendations). They argue that such models will be capable of carrying out a diverse set of tasks using very little or no task-specific labelled data.

Add surveys on sepsis, etc.

2.1.2 Healthcare Applications

Sepsis

Sepsis is a life-threatening organ dysfunction caused by a dysregulated host response to infection [3]. It is a leading cause of morbidity and mortality in hospitals, with an estimated 11 million deaths annually [4]. Organ dysfunction can be identified as an acute change in total SOFA score ≥ 2 points consequent to the infection, Fig 2.1a. Prior sepsis definitions used two of 4 SIRS criteria, which are not specific to sepsis and can lead to overdiagnosis(patients who has SIRS but not sepsis).

The SOFA score is a clinical score used to track a patient's status during their stay in an ICU. It is used to determine the extent of a person's organ function or rate of failure. The SOFA score is based on the following six organ systems: respiratory, coagulation, liver, cardiovascular, renal, and neurological. Each system is assigned a score from 0 to 4, with higher scores indicating more severe dysfunction. The total SOFA score is the sum of the individual scores for each organ system.

ML-Driven Sepsis Detection

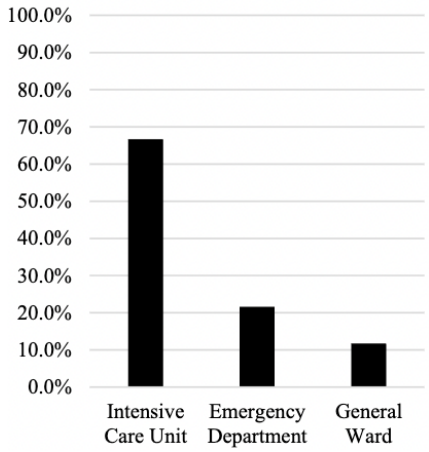
Surveys: Islam et al. [5] does a meta-study, showing that a majority of studies used ICU data (Fig 2.1b), as well as the MIMIC-III dataset. Most of the studies were carried out

Sequential (Sepsis-Related) Organ Function Assessment (SOFA) Score.[a]

System / Score	0	1	2	3	4
Respiration: PaO_2/FiO_2, mmHg (kPa)	≥400 (53.3)	<400 (53.3)	<300 (40)	<200 (26.7) with respiratory support	<100 (13.3) with respiratory support
Coagulation: Platelets x $10^3/\mu L$	≥150	<150	<100	<50	<20
Liver: Bilirubin, mg/dL (μmol/L)	<1.2 (20)	<1.2-1.9 (20-32)	2.0-5.9 (33-101)	6.0-11.9 (102-204)	>12.0 (204)
Cardiovascular	MAP ≥70 mm Hg	MAP <70 mm Hg	Dopamine <5 or dobutamine (any dose)[b]	Dopamine 5.1-15 or epinephrine ≤0.1, or norepinephrine ≤0.1[b]	Dopamine >15 or epinephrine, >0.1, or norepinephrine >0.1[b]
Central nervous system: Glasgow coma scale score[c]	15	13-14	10-12	6-9	<6
Renal: Creatinine mg/dL (μmol/L); Urine output, mL/day	<1.2 (110)	1.2-1.9 (110-170)	2.0-3.4 (171-299)	3.5-4.9 (300-440); <500	>5.0 (440); <200

Figure 2. Abbreviations: PaO_2/FiO_2, partial pressure of oxygen/fraction of inspired oxygen. a) Adapted from Vincent et al[3]; b) Catecholamine doss in μg/kg/min, >1 hour; c) Glasgow Coma Scale scores range from 3-15 (3 minimum,15 normal).

(a) SOFA score criteria for organ dysfunction.



(b) Sepsis data sources.

Figure 2.1: (a) SOFA score criteria for organ dysfunction and (b) Sepsis data sources.

using vital signs (73.1%) and laboratory data (65.4%).The median number of features was 22. The most common top features were heart rate, temperature, WBC count, systolic BP, age, dystolic BP and respiratory rate.

Moor et al. [6] show that .

2025: [7] 2025: FDA Device [8]

HCI 2020: **sandhu2022integrating** 2022: [9]

Papers **2015-2017**: ?? **2017-2020**: TREWS [10] **2020-2022**: ?? **2023**: ?? **2024**: [1], [11] shows that sepsis is uniquely high-stakes, uncertain, time-sensitive. Proposes an iterative loop where the model suggests labs to be drawn based on uncertainty.

Cardiac Arrest

Cardiac Arrest is a critical condition that requires immediate medical intervention. It occurs when the heart stops beating effectively, leading to a lack of blood flow to vital organs. There are two types of cardiac arrest: out-of-hospital cardiac arrest (OHCA) and in-hospital cardiac arrest (IHCA). OHCA survival rate to discharge is 10-12%, while IHCA survival rate is 20-25% [12]. 80% of presenting rhythms are non-shockable, meaning that defibrillation is not an option, i.e. that early detection is the best fix. The incidence is 9-10 per 1000 admissions [12].

Breast Cancer affects 1 in 8 women. They use recurrence gene assays (Breast Cancer Index, JCO).

2.2 Explainable AI

read the LIME, SHApe papers, and add a summary here.

2.3 Multimodal AI

read the papers on multimodal AI, and add a summary here.

RQ: how do we compress exomic analysis? Videos belong in low-dimensional space, so do sequences also?

Dimension-reduction with phenotypes!

HeLM HAIM

2.4 Time-Series Forecasting

Given $\mathbf{X}$, a vector of time-series data, and a binary labelling function $f(\mathbf{X}, t)$ which labels the current state as 0/1, our task is to given $\mathbf{X}[i]$, predict $f(\mathbf{X}, j)$ for some $j > i$.

Approaches:

1. Encode X[0...i] as text, ask an LLM
2. Encode X[0...i] as text, fine-tune an LLM classification head
3. Neurosymbolic: Somehow encode X[0...i], get distribution of $f(\hat{X}_{i+1})$, $f(\hat{X}_{i+2})...f(\hat{X}_j)$, train a model with binary cross-entropy loss
4. Neurosymbolic with Pretraining: Somehow encode X[0...i], First, pretrain a model M to do next-event prediction, i.e. given $\mathbf{X}[i]$, predict $\mathbf{X}_{i+1}, \mathbf{X}_{i+2}, ..., \mathbf{X}_j$, then fine-tune M using the neurosymbolic approach above.

Chapter 3

Research Meetings and Collaborations

3.1 Advisor Meetings

3.1.1 Meeting Log Template

Date	Add specific meeting dates
Attendees	[Advisor Name(s), Collaborators]
Key Discussions	• Research direction refinement • Methodology challenges • Publication strategy
Action Items	• Literature review update • Experiment design • Manuscript preparation

3.2 Collaboration Network

- **Clinical Collaborators:**
  - Penn Medicine
  - Clinicians specializing in sepsis research
- **Interdisciplinary Connections:**
  - Machine Learning Researchers
  - Clinical Informaticists

3.3 Research Alignment

Key research objectives discussed with advisors:

1. Develop trustworthy AI models for critical healthcare applications
2. Enable multimodal reasoning across clinical data sources
3. Ensure model explainability for high-stakes decisions

Update with specific meeting details and outcomes

Chapter 4

Research Experiments

4.1 Experimental Framework

4.1.1 Research Focus Areas

- Early Sepsis Detection
- Multimodal AI Explanations
- Treatment Recommendation Systems

4.2 Experiment Tracking Template

Experiment ID	Assign unique identifier
Hypothesis	Developing explainable AI models improves clinical decision-making
Methodology	• Multimodal data integration • Transformer-based architectures • Concept-based explanations
Data Sources	• Electronic Health Records • Time-series clinical data • Multi-modal patient information
Key Metrics	• Model accuracy • Explanation faithfulness • Clinical utility

4.3 Preliminary Experimental Directions

1. Develop novel explanation frameworks
2. Create multimodal reasoning mechanisms
3. Validate model performance across diverse clinical contexts

Detailed experiment protocols to be developed

Chapter 5

Career Development

5.1 Research Grants and Funding

- **Current Grant:**
  - Developing Trustworthy AI for Early Sepsis Detection
  - Lead PhD Student on Institutional Research Grant
- **Potential Funding Opportunities:**
  - NIH Research Grants
  - NSF Computing Innovations Fellowships
  - Institutional Research Support

5.2 Professional Development

1. Technical Skills Enhancement
   - Advanced Machine Learning Techniques
   - Clinical Informatics
   - Ethical AI Development
2. Soft Skills Development
   - Scientific Communication
   - Interdisciplinary Collaboration
   - Research Ethics

5.3 Career Trajectory

- **Short-term Goals:**
  - Complete PhD with impactful research
  - Publish in top-tier conferences/journals

  - Develop industry and academic network

- **Long-term Aspirations:**
  - Lead AI research in healthcare
  - Bridge machine learning and clinical practice
  - Contribute to ethical AI development

Regularly update career development plan

5.4 Logistics

[Logistics](Research/AI for Healthcare/Logistics.md)

5.5 Skills

Must read:

- https://arxiv.org/pdf/2409.10580
- http://proceedings.mlr.press/v119/rieger20a.html
- https://adelaidehsu.github.io/
- https://www.nature.com/articles/s42256-019-0048-x

Conferences:

- ML4H
- AAAI Symposium
- JAMA AI

5.6 People

5.6.1 Academia

Name	Institution	Field
Emily Alsentzer		
Bin Yu	UCB	LITERALLY WHAT I DO
Cynthia Rudin		
Peter Solovitz	MIT	Clinical Decision Making (CDM)
John Guttag	MIT	Adverse-Event Prediction, Treatment Suggestion
David Sontag	MIT	
Pranav Rajpurkar	Harvard	Foundation Models, Generalist MAI, GMAI
Hima Lakkaraju	Harvard	XAI
Zak Kohane	Harvard	CDM, genomic rare diseases

Nigam Shah	Stanford	GreenButton, Atropos Health
Saurabh Gombar	Stanford	
Purvesh Khatri	Stanford	
Dokyoon Kim	Penn	Multiomics data
Suchi Saria	JHU	Bayesian startup
Su-in Lee	UWash	everything!
Anshul Kundaje	Stanford	Immunology
Sanmi Koyejo	UIUC	Google
Irene Y Chen	UCB	Equitable AI
Matthew Abraham	Princeton	MedARC
Zhi Huang	Penn	
Ahmed Alaa	UCB	

5.6.2 Industry

Institution	**Name**	**Contacted?**
Microsoft Health	Matthew Lungren	
	Chandan Singh	
	Hoifung Poon	Y
	Tristan Naumann	Y
Google Health	Stephen Pfohl	
	Mayank Daswani	
	Chirag Nagpal	
	Priya Gupta	
	Khaled Saab	
	Wei-Hung Weng	
	Ryutaro Tanno	
a16z	Julie Yoo	
a16z	Vijay Pande	
Layer Health	David Sontag	
Apple Health		
Genesis Therapeutics		
PictureHealthAI		
JoriAI		

5.7 Conferences

SAIL (https://sail.health/)

e AI for Healthcare (https://sites.google.com/view/imlh2023/home?authuser=1)

ML4H (https://ml4h.cc)

5.8 Papers

5.8.1 Survey

Scoping Evaluation (https://www.medrxiv.org/content/10.1101/2023.09.12.23295381v1)

5.8.2 XAI

AI for Chest X-rays (https://www.nature.com/articles/s42256-022-00536-x)

currence prediction (https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(23)00377-2/fulltext)

el for Heart Failure (https://arxiv.org/abs/2310.15472)

5.8.3 Foundation Models

Models for Medicine (https://www.nature.com/articles/s41586-023-05881-4)

medical literature (https://www.medrxiv.org/cgi/content/short/2023.06.07.23291119v1)

NEJM Catalyst (https://catalyst.nejm.org/doi/full/10.1056/CAT.21.0224)

Bibliography

[1] E. Researcher, "Optimizing artificial intelligence in sepsis management: Opportunities in the present and looking closely to the future," *Artificial Intelligence in Medicine*, 2024.

[2] M. Moor, O. Banerjee, Z. S. H. Abad, *et al.*, "Foundation models for generalist medical artificial intelligence," *Nature*, vol. 616, pp. 259–265, 2023. DOI: 10.1038/s41586-023-05881-4.

[3] M. Singer *et al.*, "The third international consensus definitions for sepsis and septic shock (sepsis-3)," *JAMA*, vol. 315, no. 8, pp. 801–810, 2016.

[4] K. E. Rudd, S. C. Johnson, K. M. Agesa, *et al.*, "Global, regional, and national sepsis incidence and mortality, 1990–2017: Analysis for the global burden of disease study," *The Lancet*, vol. 395, no. 10219, pp. 200–211, 2020.

[5] K. R. Islam, J. Prithula, J. Kumar, *et al.*, "Machine learning-based early prediction of sepsis using electronic health records: A systematic review," *Journal of Clinical Medicine*, vol. 12, no. 17, p. 5658, 2023. DOI: 10.3390/jcm12175658.

[6] M. Moor, B. Rieck, M. Horn, C. R. Jutzeler, and K. Borgwardt, "Early prediction of sepsis in the icu using machine learning: A systematic review," *Frontiers in Medicine*, vol. 8, p. 607 952, 2021. DOI: 10.3389/fmed.2021.607952.

[7] E. G. Bignami, G. Arrigoni, S. Gramegna, G. Grasselli, and G. Foti, "Artificial intelligence in sepsis management: An overview for clinicians," *Journal of Clinical Medicine*, vol. 14, no. 1, p. 286, Jan. 2025. DOI: 10.3390/jcm14010286. [Online]. Available: https://doi.org/10.3390/jcm14010286.

[8] A. Bhargava, C. López-Espina, L. Schmalz, *et al.*, "Fda-authorized ai/ml tool for sepsis prediction: Development and validation," *NEJM AI*, vol. 1, no. 12, AIoa2400867, 2024. DOI: 10.1056/AIoa2400867. eprint: https://ai.nejm.org/doi/pdf/10.1056/AIoa2400867. [Online]. Available: https://ai.nejm.org/doi/full/10.1056/AIoa2400867.

[9] J. A. Silvestri, T. E. Kmiec, N. S. Bishop, S. H. Regli, and G. E. Weissman, "Desired characteristics of a clinical decision support system for early sepsis recognition: Interview study among hospital-based clinicians," *JMIR Human Factors*, vol. 9, no. 4, e36976, 2022. DOI: 10.2196/36976. [Online]. Available: https://doi.org/10.2196/36976.

[10] R. Adams, K. Henry, A. Sridharan, *et al.*, "Prospective, multi-site study of patient outcomes after implementation of the trews machine learning-based early warning system for sepsis," *Nature Medicine*, vol. 28, no. 7, pp. 1455–1460, 2022. DOI: 10.1038/s41591-022-01894-0.

[11] S. Zhang, J. Yu, X. Xu, *et al.*, "Rethinking human-ai collaboration in complex medical decision making: A case study in sepsis diagnosis," ser. CHI '24, Honolulu, HI, USA: Association for Computing Machinery, 2024, ISBN: 9798400703300. DOI: 10.1145/3613904.3642343. [Online]. Available: https://doi.org/10.1145/3613904.3642343.

[12] L. W. Andersen, M. J. Holmberg, K. M. Berg, M. W. Donnino, and A. Granfeldt, "In-hospital cardiac arrest: A review," *JAMA*, vol. 321, no. 12, pp. 1200–1210, 2019. DOI: 10.1001/jama.2019.1696.