# Grounding Multimodal AI Algorithms in Real-World Evidence for Critical Traumatic Brain Injury Outcome Prediction

Pranav Manjunath[1] • Timothy W. Dunn, Ph.D[1]
[1] Dept. of Biomedical Engineering, Duke University

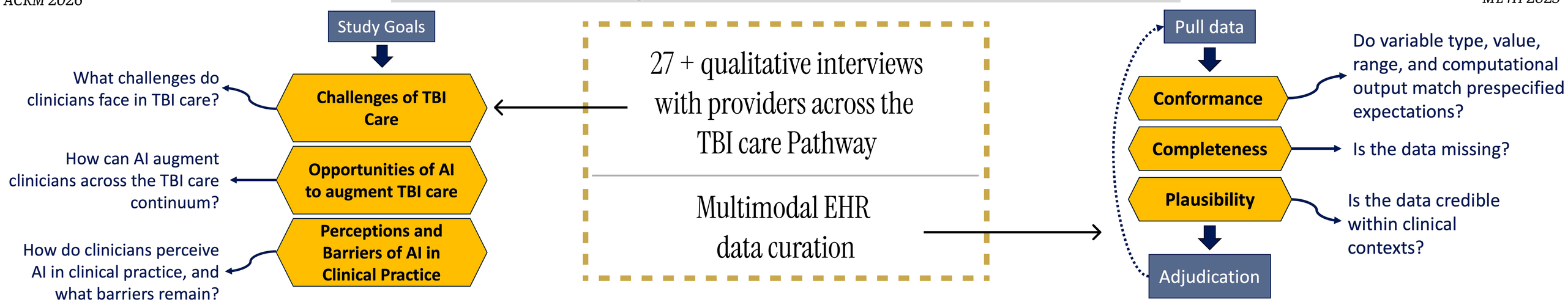
Duke BIOMEDICAL ENGINEERING

*Mixed-Methods Research to Design and Develop AI Algorithms for Augmenting Traumatic Brain Injury (TBI) Care [Health + AI + HCI]*

Portfolio

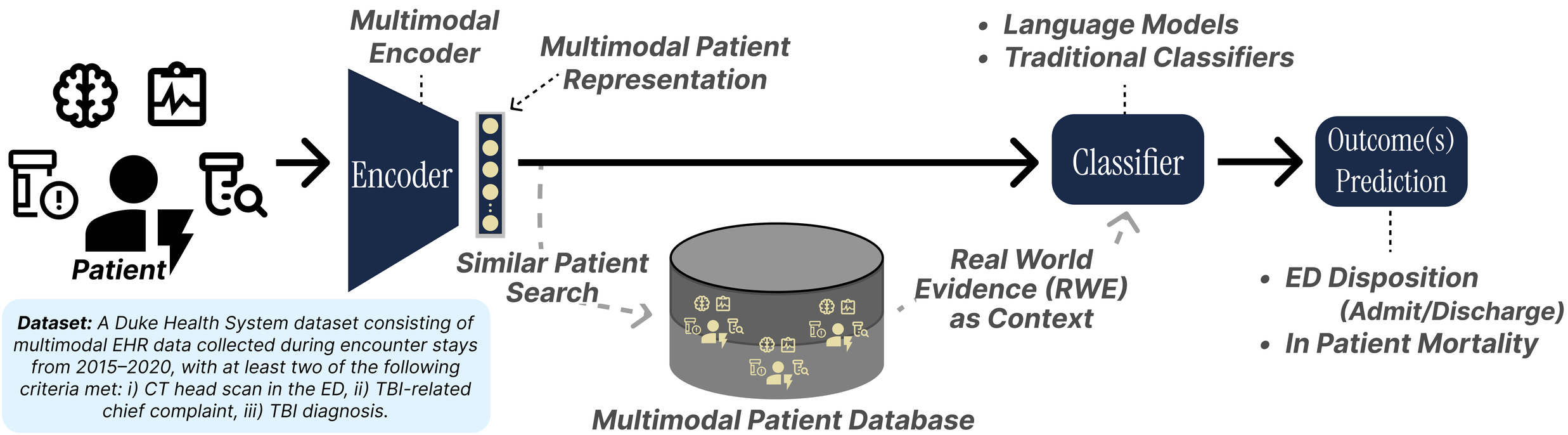
## Determining Critical Outcomes for TBI

*ACRM 2026**




*ML4H 2025*

## Multimodal Patient Representation for Real World Evidence




***Dataset:*** *A Duke Health System dataset consisting of multimodal EHR data collected during encounter stays from 2015–2020, with at least two of the following criteria met: i) CT head scan in the ED, ii) TBI-related chief complaint, iii) TBI diagnosis.*

**Q: Does personalized RWE as context improve outcome prediction?**

A: Yes! Furthermore, random patients as context harms the classification performance. Selecting the right context is important!

Outcome: ED Disposition

| | Phi | Qwen |
|---|---|---|
| | F1 | F1 |
| No RWE | 0.539 | 0.383 |
| RWE (1 Patient) | 0.555 | 0.522 |
| RWE (5 Patient) | 0.621 | **0.642** |
| RWE (7 Patient) | **0.629** | 0.627 |
| RWE (7 Random Patient) | 0.501 | 0.525 |

*JBHI 2025*

**Q: How do guidelines compare to RWE, as context, in predicting outcomes?**

A: RWE as context, represents more real world outcomes while guidelines tends to over admit patients to the hospital

Outcome: ED Disposition

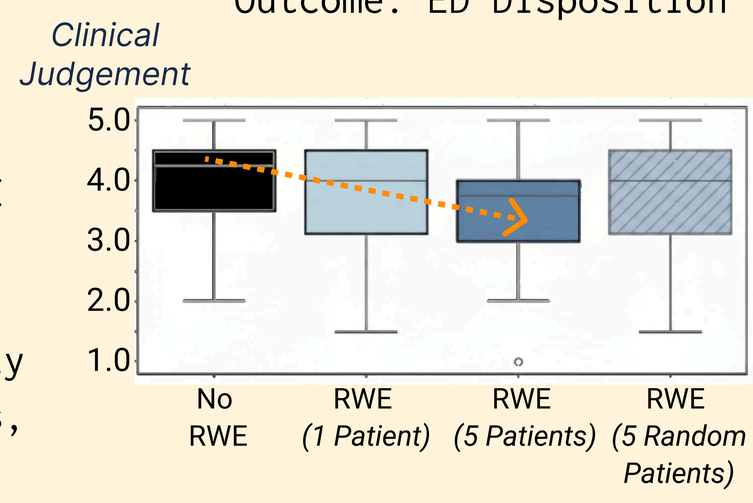
| | | No Context | Guidelines | RWE (7 Patients) | Guidelines and RWE |
|---|---|---|---|---|---|
| Phi | F1 | 0.586 | 0.512 | **0.634** | 0.626 |
| | SN | 0.728 | **0.948** | 0.884 | 0.698 |
| | SP | 0.656 | 0.202 | 0.589 | **0.758** |
| Qwen | F1 | 0.386 | 0.547 | **0.626** | 0.563 |
| | SN | 0.311 | **0.923** | 0.683 | 0.912 |
| | SP | **0.865** | 0.340 | 0.774 | 0.399 |

*BHI 2025*

**Q: How does RWE as context affect reasoning?**

A: Interestingly, RWE can actually hurt reasoning, language models may incorrectly link similar patients, producing hallucinations.

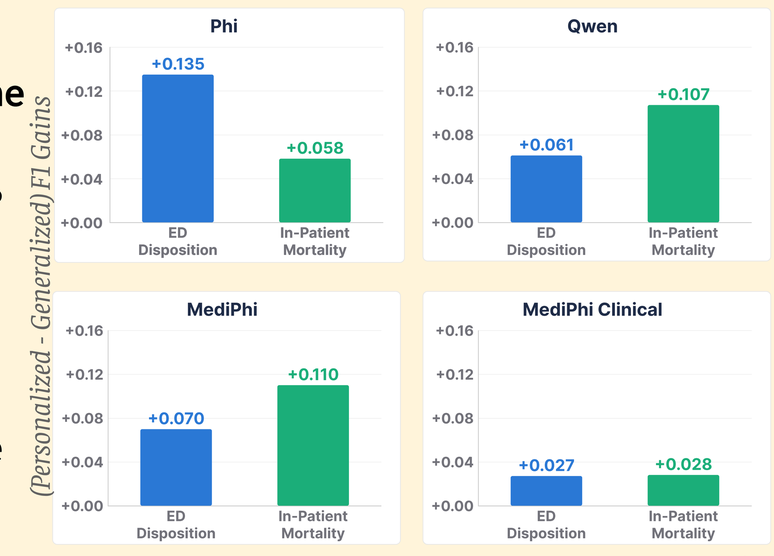
Outcome: ED Disposition




*BHI 2025*

**Q: Is personalized RWE dependent on the outcome and robust to language models?**

A: RWE tends to be outcome specific and is robust to a variety of language models




*AMIA 2026**

Fun Fact: RWE holds true temporally within an encounter, but the similar patients at each time point vary significantly*

**Q: How can RWE help support information needs of patients and caregivers when in the Emergency Department?***

Develop a taxonomy of information needs in the ED → Design a personalized RWE-based answer engine to support information needs → Evaluate answer engine with existing sources of information needs → Conduct User Studies to study its effectiveness in practice

*CHI 2026 Workshop*

*In Collaboration with Dr. Adam Perer, HCII at Carnegie Mellon University*

*This work is supported by the National Institutes of Health (NIH) under Grant R01NS123275 and Bass Connections Research Grant, Duke University*

**Indicates current on-going projects*