Breakout Session 4: Track A

Development of a Method for Identifying Ethical Considerations Arising from Healthcare AI Deployments

Dr. Danton Char (Moderator)
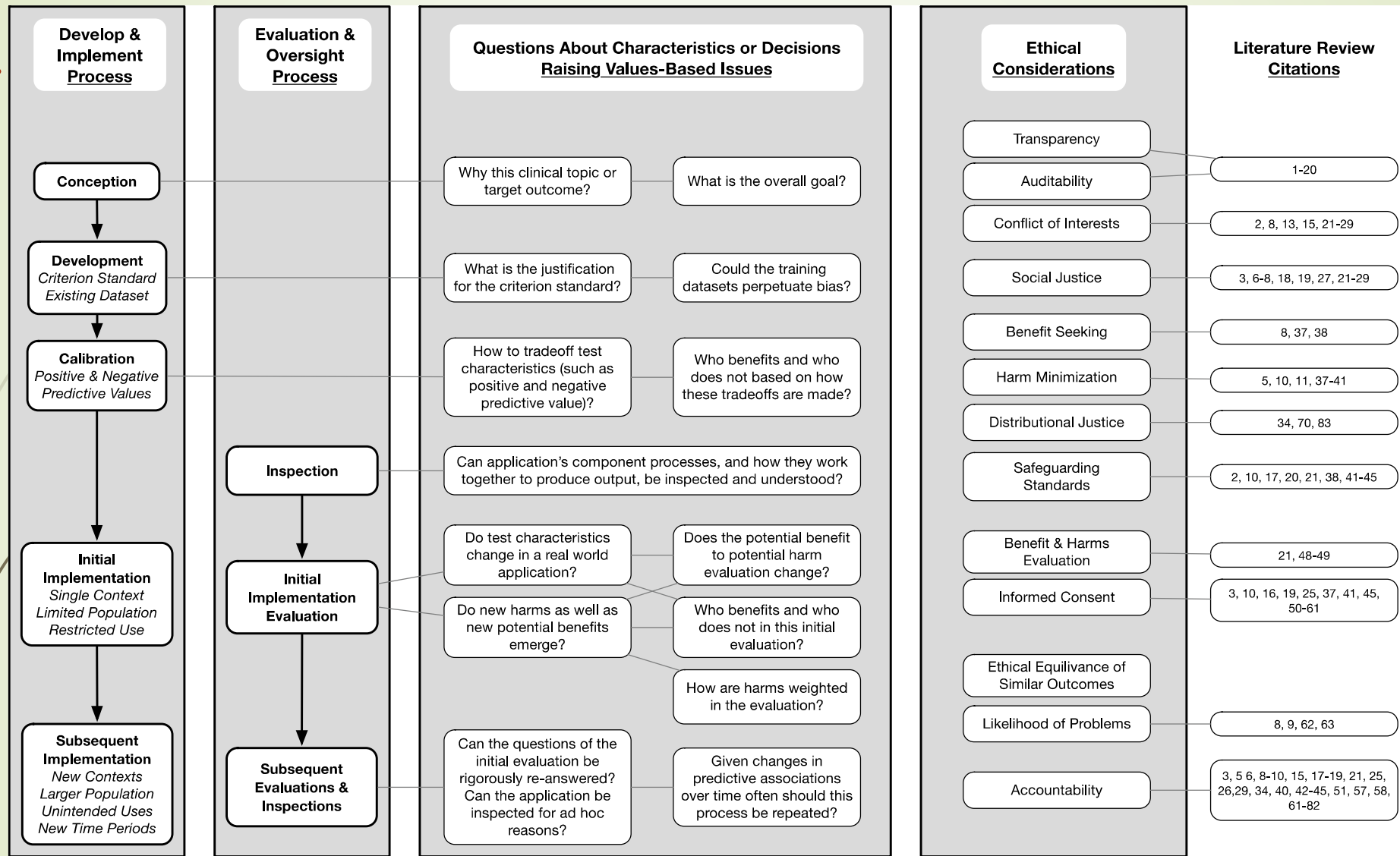*Associate Professor, Stanford Medicine*

Development of a Method for Identifying Ethical Considerations Arising from Healthcare AI Deployments

Danton Char, Supplement Investigator

Nigam Shah, Curt Langlotz, Co-PIs on Parent Award (R01HL155410)

- **Identifying ethical concerns with AI-HC before problems result** has become a stated goal of design oversight groups and regulatory agencies, such as the U.S. F.D.A and Department of Health and Human Services.

- However, **the lack of an accepted, scalable methodology for ethical analysis of AI-HC is a critical obstacle to achieving this goal**.

- Other promising medical technologies, like gene therapy, have led to direct patient harm in clinical research because of failures to identify and address ethical concerns early on.

- Although many frameworks and sets of principles for ethical evaluation of AI have been developed, concrete approaches for applying these principles are lacking. Our work fills this gap.

Develop & Implement Process

Conception

Development
Criterion Standard
Existing Dataset

Calibration
Positive & Negative
Predictive Values

Initial Implementation
Single Context
Limited Population
Restricted Use

Subsequent Implementation
New Contexts
Larger Population
Unintended Uses
New Time Periods

Evaluation & Oversight Process

Inspection

Initial Implementation Evaluation

Subsequent Evaluations & Inspections

Questions About Characteristics or Decisions Raising Values-Based Issues

Why this clinical topic or target outcome?

What is the overall goal?

What is the justification for the criterion standard?

Could the training datasets perpetuate bias?

How to tradeoff test characteristics (such as positive and negative predictive value)?

Who benefits and who does not based on how these tradeoffs are made?

Can application's component processes, and how they work together to produce output, be inspected and understood?

Do test characteristics change in a real world application?

Does the potential benefit to potential harm evaluation change?

Do new harms as well as new potential benefits emerge?

Who benefits and who does not in this initial evaluation?

How are harms weighted in the evaluation?

Can the questions of the initial evaluation be rigorously re-answered? Can the application be inspected for ad hoc reasons?

Given changes in predictive associations over time often should this process be repeated?

Ethical Considerations

Transparency

Auditability

Conflict of Interests

Social Justice

Benefit Seeking

Harm Minimization

Distributional Justice

Safeguarding Standards

Benefit & Harms Evaluation

Informed Consent

Ethical Equilivance of Similar Outcomes

Likelihood of Problems

Accountability

Literature Review Citations

1-20

2, 8, 13, 15, 21-29

3, 6-8, 18, 19, 27, 21-29

8, 37, 38

5, 10, 11, 37-41

34, 70, 83

2, 10, 17, 20, 21, 38, 41-45

21, 48-49

3, 10, 16, 19, 25, 37, 41, 45, 50-61

8, 9, 62, 63

3, 5 6, 8-10, 15, 17-19, 21, 25, 26,29, 34, 40, 42-45, 51, 57, 58, 61-82

Char D, Shah N, Magnus D. *N Eng J Med* 2018; 378: 981
Char D, Abramoff M, Feudtner C. *Am J Bioeth*; 2020 Nov

Values Collisions Anticipate Ethical Concerns

- Multiple stakeholders are impacted by any AI-HC. Stakeholders can be identified by examining the design/deployment contexts

- Stakeholders have different values, and explicit or implicit goals for the AI-HC, that should and can be ascertained

- Process of design and development of an AI-HC involves making a series of decisions

- How a stakeholder makes these decisions, or would want these decisions to be made, reflects their underlying values

- ***Where stakeholder groups disagree or their values are at odds about resolving these decisions—where values collide—are where ethical problems are most likely to emerge***

- Some value collisions may mark novel ethical concerns. Many can be resolved by drawing on prior scholarship on similar or related problems.

Challenge: No clear use case for parent award PE AI

Stanford Health Care does not have a pulmonary embolism response team (PERT), so there is no clinical service line "owner" for PE specific care management or institutional performance targets

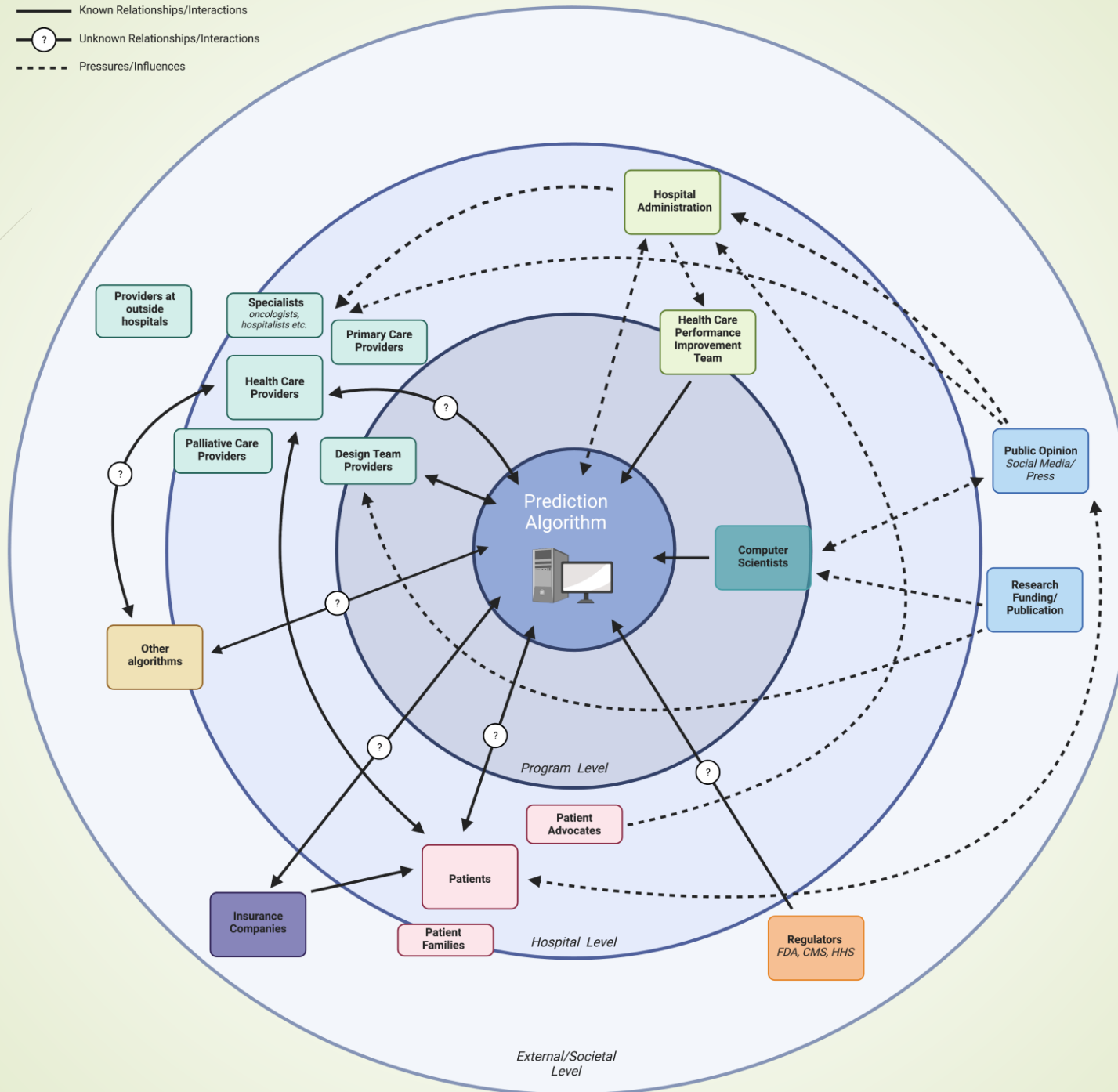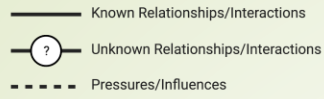Different providers face different challenges

Our interviews suggest that there is not a clear use case or immediate value for AI in PE management

Alison Callahan, Keith Morse

Case study: ML-mortality prediction to guide advance care planning

Our model is an 18-layer Deep Neural Network that inputs the EHR data of a patient, and outputs the probability of death in the next 3-12 months.

We train the model on the historic data from the Stanford Hospital EHR data base, which contains data of over 2 million patients. The model is trained to predict probability of patient mortality in the next 3-12months. Training uses patient's EHR data from the past 12 months, specifically the diagnostic codes, procedure codes, medication codes, and encounter details. All this data is converted into a feature vector for 13,654 dimensions. The trained model achieves an AUROC score of 0.93 and an Average Precision score of 0.69 on cross validation.
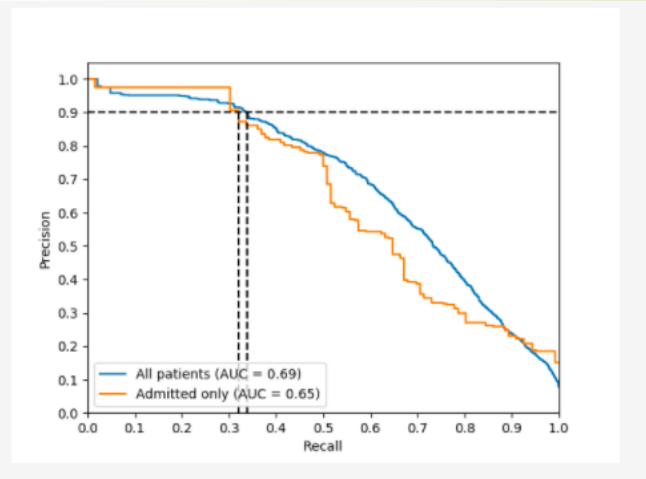
Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. BMC Med Inform Decis Mak. 2018 12 12;18(Suppl 4):122

Legend:
— Known Relationships/Interactions
—○?— Unknown Relationships/Interactions
- - - - Pressures/Influences

Cagliero D, Deuitch N, Shah N, Feudtner C, Char D. A framework to identify ethical concerns with ML-guided care workflows: a case study of mortality prediction to guide advance care planning. J Am Med Inform Assoc. 2023 Apr 19;30(5):819-827.

"*At one point they were asking me can you guys predict if they've [patients] got 24 hours or less? Because if they've got 24 hours or less, we're going to put them in Obs and not admit them, and Obs means they're not officially admitted, and if they die in Obs, they don't count as a death*. **And I was like, I feel like I'm going to vomit into my mouth right now because you're telling me you want to know they're going to die in 24 hours because you wouldn't put them in a normal inpatient acute care bed, you'd put them in Obs!?!**"

Other Collisions Raising Ethical Considerations

- Who gets the mortality prediction? (patient-clinician-designer collision)

- Unintended uses/uses other than ACP referral (clinician-designer collision)

- What to do when prediction is right but ACP is inappropriate? (patient-designer collision)

Cagliero D, Deuitch N, Shah N, Char D. Evaluating Ethical Concerns with Machine Learning to Guide Advance Care Planning. Journal of Investigative Medicine, 2021;69:152
Cagliero D, Deuitch N, Shah N, Feudtner C, Char D. A framework to identify ethical concerns with ML-guided care workflows: a case study of mortality prediction to guide advance care planning. J Am Med Inform Assoc. 2023 Apr 19;30(5):819-827.

Broader Vision: Ways to Strengthen Further

- Outcome Studies: Do recommendations result in actions to mitigate ethical concerns identified? Are recommendations helpful to design/development teams? Are recommendations helpful to healthcare systems? To patients?

- Balancing Speed With Thoroughness

- Generalizability of our process to differently resourced healthcare systems

- Case Categorization

- How to evaluate generative AI

- Ongoing additional ethical evaluations predictive tools and now generative models (LLMs)

Viewpoint | AI in Medicine

ONLINE FIRST   FREE

November 30, 2023

President Biden's Executive Order on Artificial Intelligence—Implications for Health Care Organizations

Michelle M. Mello, JD, PhD, MPhil[1,2]; Nigam H. Shah, MBBS, PhD[3]; Danton S. Char, MD[4,5]

» Author Affiliations | Article Information

Thank you

- ODSS Administrative Supplement to R01HL155410
- Nigam Shah
- Curt Langlotz
- Michelle Mello
- Alaa Youssef
- Elisabeth Grosvenor
- Alison Callahan
- Keith Morse
- TDS@SHC