Identifying and Anticipating Ethical Challenges with Machine Learning for Genomics

Danton Char, M.D., M.A.S.

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

Facebook Says Cambridge Analytica Harvested Data of Up to 87 Million Users

The Facebook chief executive, Mark Zuckerberg, is expected to appear before multiple congressional committees. Steven Senne/Associated Press

nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress

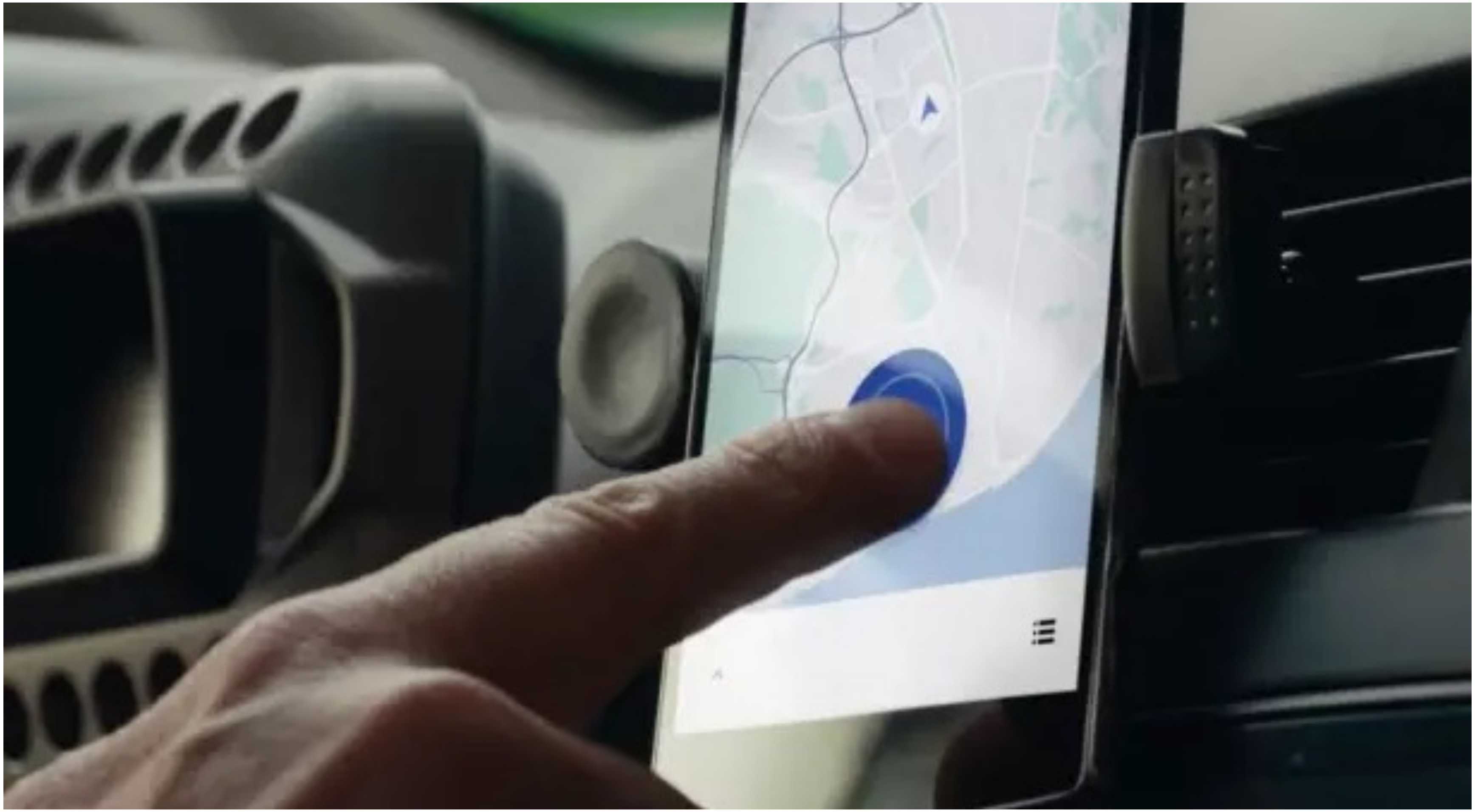


Image from https://www.uber.com/

http://www.huffingtonpost.com/entry/paris-climate-talks-artist-protests-corporations_us_565c5769e4b072e9d1c25108

https://sites.duke.edu/lit290s-1_02_s2017/2017/03/27/vw-scandal-explained-in-pictures/

Photo: Doug Mills, NYT, June 25, 2019

https://www.dni.gov/files/NCSC/documents/SafeguardingOurFuture/NCSC_China_Genomics_Fact_Sheet_2021.pdf

“So for us, ***one of the more immediate benefits of genomic sequencing is we could have that discussion with the parents and change our goals of care to comfort as opposed to prolongation with futile intensive care****.” (Neonatologist)*

Char DS, Lee SS, Magnus D, Cho M. Genet Med. 2018 11;20(11):1455-1461.

Genomic sequencing results	Likelihood of recommending: Surgical palliation	ECMO	Transplant candidate
Breast cancer	.00	.08	.04
	(.07)	(.10)	(.10)
Childhood onset cancer syndrome	-.14*	-.09	-.24**
	(.05)	(.07)	(.09)
Intellectual disability	-.27***	-.18*	-.35***
	(.07)	(.09)	(.10)
Autism	-.17*	-.07	-.14
	(.06)	(.09)	(.09)
Number of observations	198	197	197
R^2	.14**	.07	.17

Table notes: * p<.05, ** p<.01, ***p<.001. Entries are unstandardized regression coefficients, standard errors in parentheses. Comparison group is physicians who made recommendations without genomic sequencing results.

Char D, Deuitch N, Berent M, Chung W, Krosnick J, Magnus D. in submission

Medical option	Percent of physicians who would recommend the medical option by genomic sequencing condition: No genomic sequencing results (n=82)	Genomic sequencing indicates schizophrenia (N=100)	Difference
ECMO	75.2%	67.4%	7.8%
Transplant candidate	70.5%	55.1%	15.5%$^{+}$

Char D, Deuitch N, Berent M, Chung W, Krosnick J, Magnus D. in submission

"They've spent probably over four million dollars on him just in two years of the transplant, pre-transplant, post-everything. And if we had known about his [genetic mutation], his genes being bad, if we had known about it in advance, I always think what if they had declined to treat him." (F1)

"If there's preexisting conditions or the potential for conditions to come up in the future, how much does a medical institution invest in helping somebody that potentially is going to die?" (F22)

Gal DB, Deuitch N, Lee SSJ, Simon RT, Char DS. Pediatr Crit Care Med. 2021 Feb 12

“We see distrust across the board in all of our institutions, you see it with the measles outbreak and the anti-vaxxers, there’s distrust of pharmaceutical companies, there’s distrust of the mega-industry of healthcare. That will get worse and more intense with genome testing.” (F27)

Gal DB, Deuitch N, Lee SSJ, Simon RT, Char DS. Pediatr Crit Care Med. 2021 Feb 12

LEADERSHIP

Effort to restore trust in science must begin now

DEC 1, 2020

Susan R. Bailey, MD
President

PRINT PAGE

Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

Search

EMERGING INFECTIOUS DISEASES®

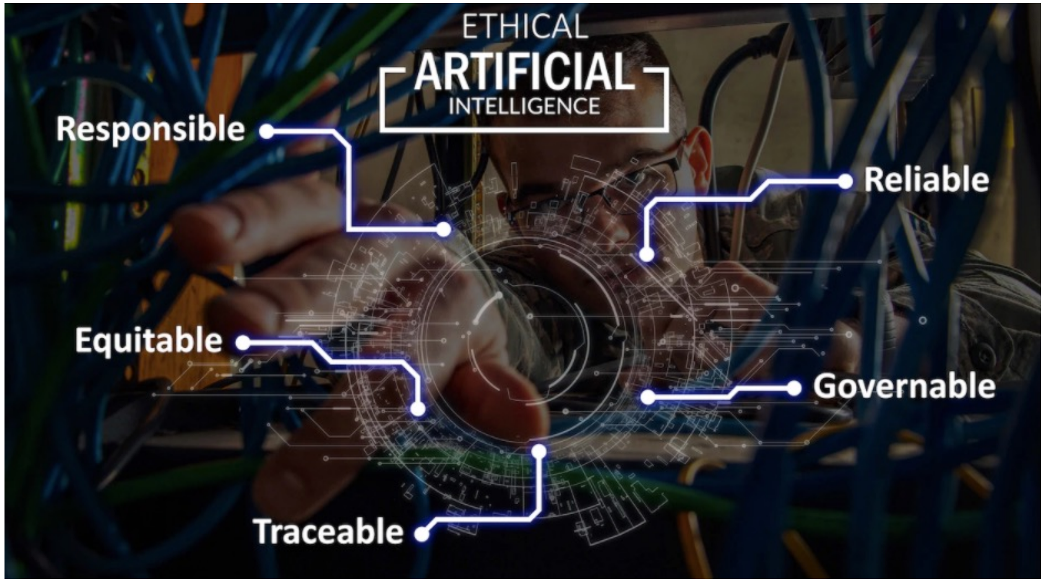
Volume 27, Number 2—February 2021

Research

Addressing COVID-19 Misinformation on Social Media Preemptively and Responsively

Emily K. Vraga✉ and Leticia Bode

DoD AI Ethical Principles



- Responsible: Humans should exercise judgment & remain responsible for use & outcomes
- Equitable: Avoid unintended bias & inadvertent harm
- Traceable: Transparent & Auditable methodologies, data sources, design procedures
- Reliable: Explicit domain of use; safety tested across entire life cycle of use in that domain
- Governable: Possess the ability to detect/avoid unintended harm & for human disengagement or deactivation

Defense Innovation Board, AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense, accessed online at media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF (Alka Patel)

IDx-DR

The first ever autonomous AI system cleared by the FDA to provide a diagnostic decision

Additional Ethical Principles for Healthcare Applications

- **Non-Maleficence:** Do no harm; patient benefit; improved clinical outcomes
- **Autonomy:** Patient still in control of their healthcare; liability for AI system malfunction related to degree of autonomy; ownership of data
- **Equity:** Absence of bias, fairness in distribution, access and benefits of groups

U.S. Food & Drug Administration (FDA) Digital Health Center of Excellence C, ,. Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. 2021. https://www.fda.gov/media/145022/download

Stanford University. Collaborative Community on Ophthalmic Imaging (CCOI). 2020:https://www.cc-oi.org/.

Abramoff MD, Tobey D, Char DS. Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process. *Am J Ophthalmol*. 2020;214(1):134-142. doi:10.1016/j.ajo.2020.02.022

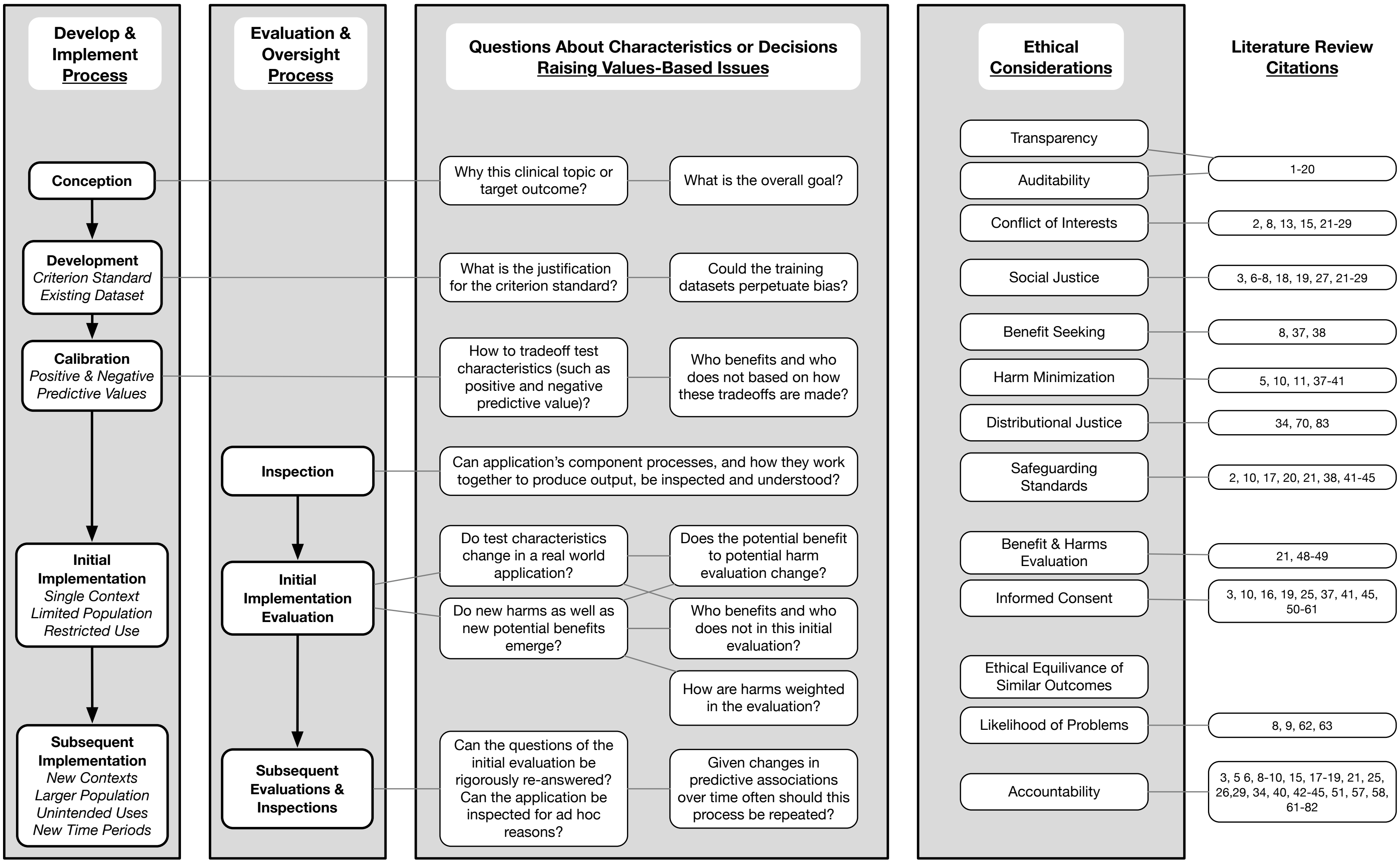
National Defense Authorization Act, January 1, 2021:
White House Interagency Coordination of AI including Ethical Issues

Image: Reuters, https://hai.stanford.edu/blog/congress-gets-serious-about-artificial-intelligence

"There is an old saying that a problem well put is half solved. This much is obvious. What is not so obvious, however, is how to put a problem well."

-Churchman, Ackoff, Arnoff

Introduction to Operations Research, 1957, page 67.



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

6 Premises

- Multiple stakeholders impacted by any ML-HCA. These stakeholders can be identified by examining the design/deployment contexts
- Stakeholder groups have different values, and explicit or implicit goals for the ML-HCA, that should and can be ascertained
- Process of design and development of an ML-HCA 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.

EXCLUSIVE AUTONOMOUS VEHICLES UBER/LYFT

Uber Finds Deadly Accident Likely Caused By Software Set to Ignore Objects On Road

By Amir Efrati May 7, 2018 9:48 AM PDT • Comments by Noah David, Michael D. Geer and 4 others

Subscribe now

Uber has determined that the likely cause of a fatal collision involving one of its prototype self-driving cars in Arizona in March was a problem with the software that decides how the car should react to objects it detects, according to two people briefed about the matter.

The car's sensors detected the pedestrian, who was crossing the street with a bicycle,

3 Interacting Data Elements

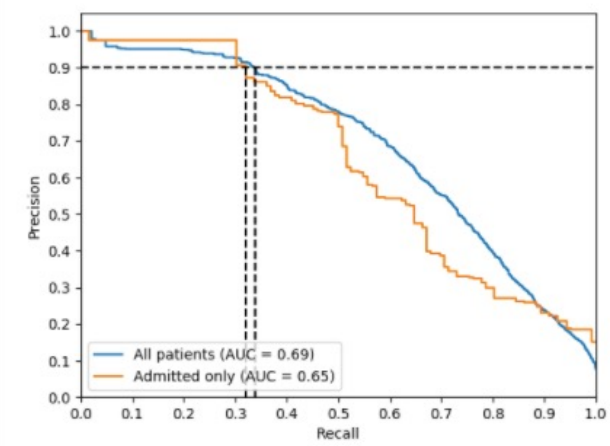
- 1) The model and the output it provides
- 2) The workflow into which the model output is introduced, and policy for allocating an intervention at a certain output recommendation
- 3) The benefit-harm trade-off of the intervention itself.
- Value mismatches can arise in any of these three elements.

Shah, N.

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.



Stakeholders

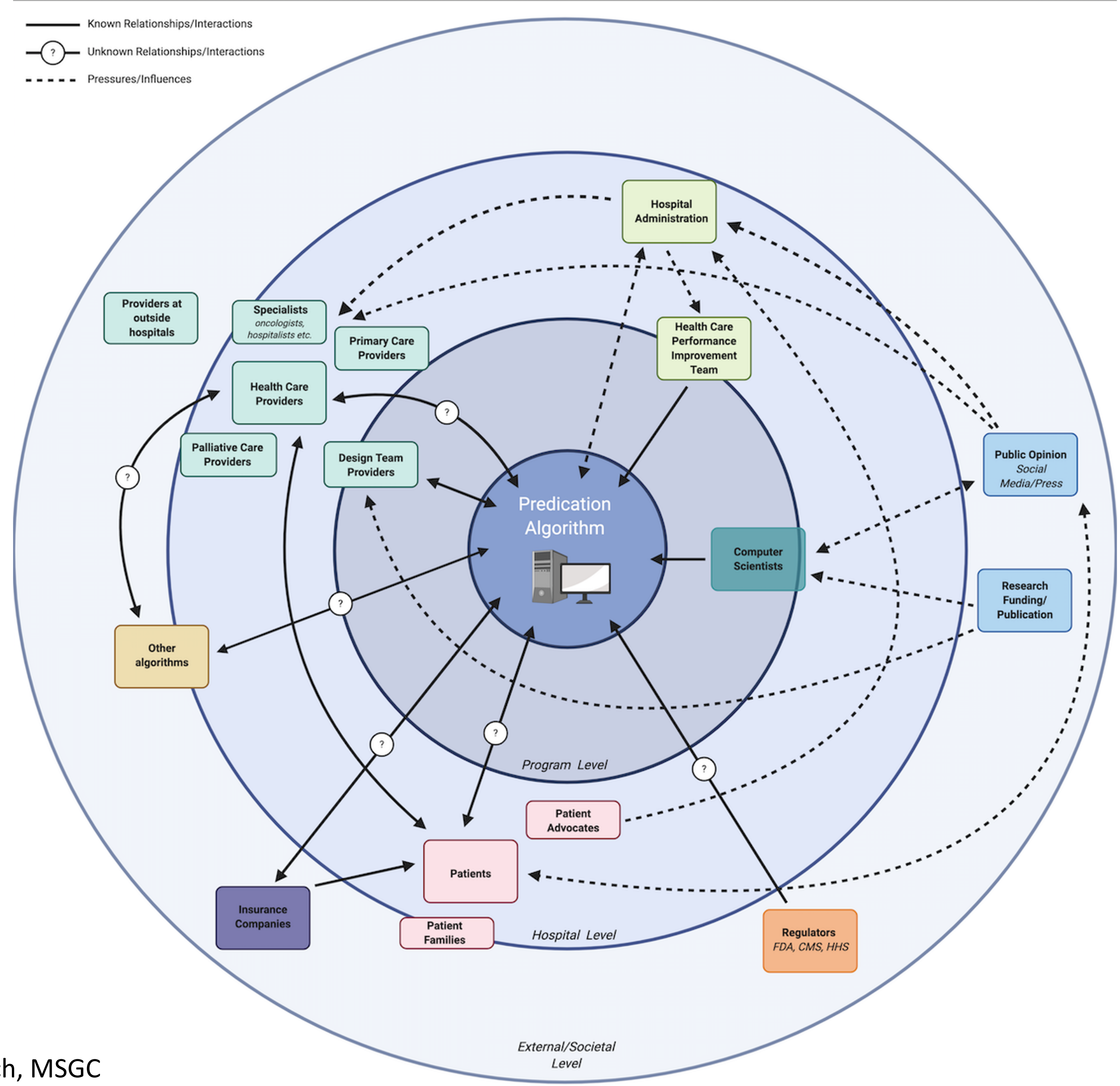


Figure: N. Deuitch, MSGC

Value Collisions

Ethical Concern	Patient Value	Clinician Value	Designer Value
Perspectives on death and end-of-life care	Want mortality prediction to inform ACP decisions	ML prediction gives numeric legitimacy to prognosis/prognostication	Concern that patients and clinicians won't know what to do with mortality prediction information
Implementation of algorithm in health care setting	Important to get this information from a trusted clinician such as a PCP	Concern around algorithm further burdening the Palliative Care team or being used in unintended ways	The algorithm has low pretest probability and the outcome is not harmful -- ideal ML "test case"
Patient involvement	Would like knowledge of mortality prediction	Agree with patient knowledge as long as accompanied with conversation	May not be an accurate predictor of mortality, so should not be shown to patients -- issue of misinterpretation
Transparency	Details not important but would like overall idea of how prediction works	Important to know about how algorithm works, emphasis on use of pre-specified trial endpoints	More important to demonstrate algorithm validation than methodology
External Pressures & Study Integrity		Concerned about media	Concern about PR blowback if misinterpreted

Design of the model:
-Perspectives on end-of-life care

Workflow:
-Who should receive the mortality prediction (i.e. Should patients have access to the mortality predictions? Should all clinicians? Should only palliative care clinicians?)
-Unintended uses of mortality prediction

Benefit-harm trade-off of the intervention:
-How and if to protect ML mortality prediction research from external pressures, like social media scrutiny before research is completed

"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!?!"***

Design team was able to prioritize needed efforts focused on:

- examining alternative implementation strategies to delivery of mortality predictions into the workflow (i.e. directly to patients or to hospitalist clinicians)
- explicitly clarifying to clinicians, administrators, and patients that the mortality prediction was only evaluated to predict need for ACP not other mortality-related needs, and renaming the prediction as “ACP needs probability” rather than “mortality prediction”;
- shielding their ongoing research into mortality prediction from social media scrutiny until endpoint driven studies were completed (i.e. enacting protections similar to blinded clinical trials)

- When should future ethical analyses should be conducted as this (or any ML-HCA) is revised and deployed more broadly?
- How to better streamline the ethical analysis process (whether questions can be delivered via survey, which questions are of the highest yield, and the optimal number of stakeholder assessments needed)?

Acknowledgments

- NHGRI K01HG008498
- Human Centered Artificial Intelligence (HAI) Seed Grant
- Natalie Deuitch, MSGC
- Dana Gal, MD
- Nigam Shah, MBBS, PhD
- Sandra Lee, PhD
- David Magnus, PhD
- Chris Feudtner, MD, PhD
- Michael Abramoff, MD, PhD
- Wendy Chung, MD, PhD
- Matt Berent, PhD
- Jon Krosnick, PhD