Implementing Integrated Diagnostics Into Precision Oncology Care

'Academic' : How to IMPLEMENT ADVANCED ALGORITHMS (Precision Oncology) with COMPLEX DATA (needed for diagnosis) into workflow (so it is Integrated)

David Dorr

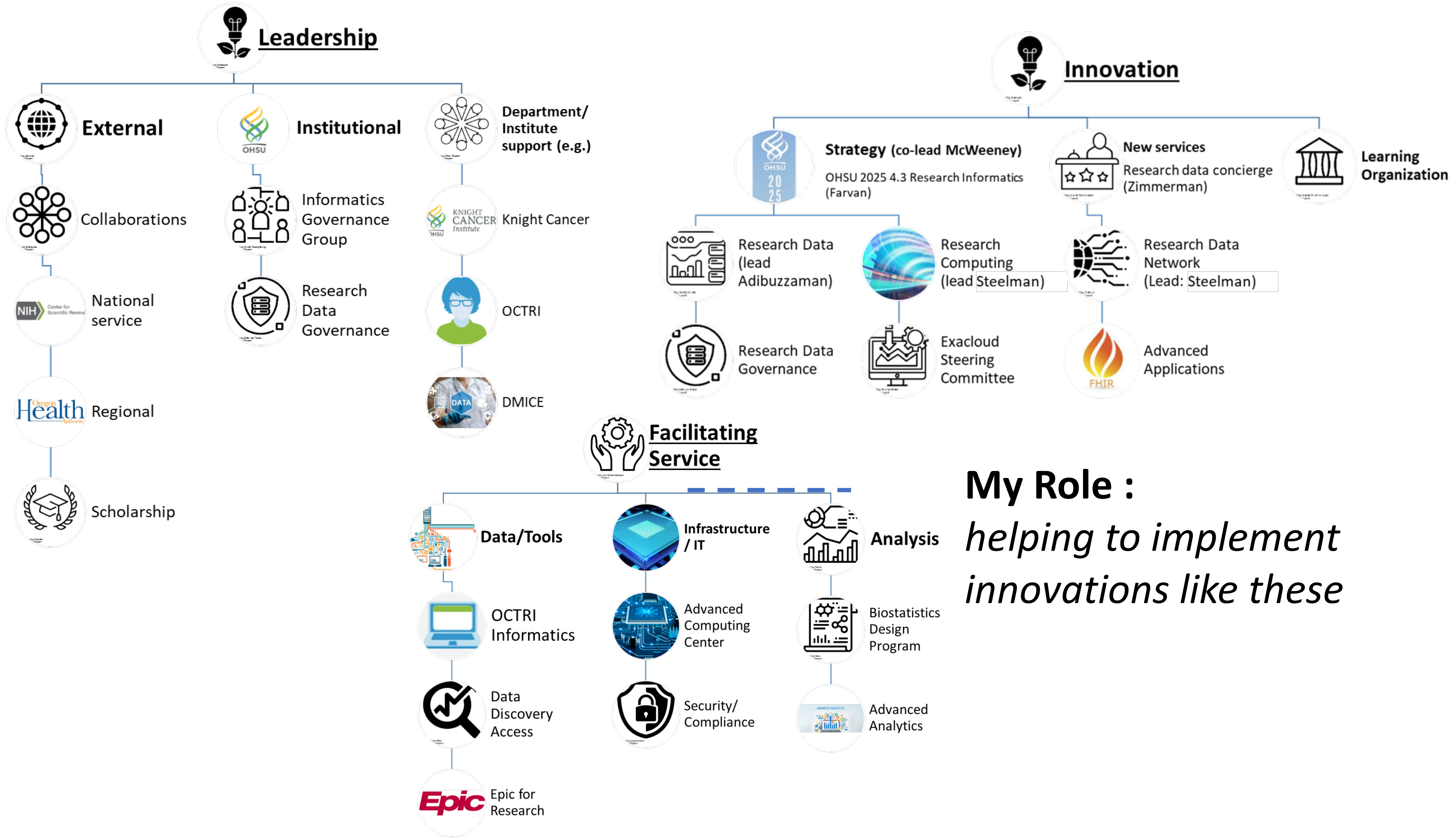
Chief Research Informatics Officer

Professor and Vice Chair

Oregon Health & Science University

What is health informatics?

- The science of the use of data, information and knowledge to improve health; interdisciplinary by nature
- Thus:
  - We strive to understand data biases and quality assessment – with higher risk for more complex data
  - We study cognitive biases and sociotechnical issues – for instance ***trust***, but also heuristics, ux and their impact on decision making and use of algorithms
  - ***Implementation*** Science – challenging for advanced algorithms
    - How does it integrate into workflow?
  - We work to mitigate these issues – one example: OHSU's Statement on Artificial Intelligence and an upcoming Code of Conduct effort



My Role :

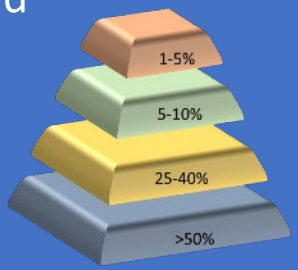
helping to implement innovations like these

The Mission of Care Management Plus

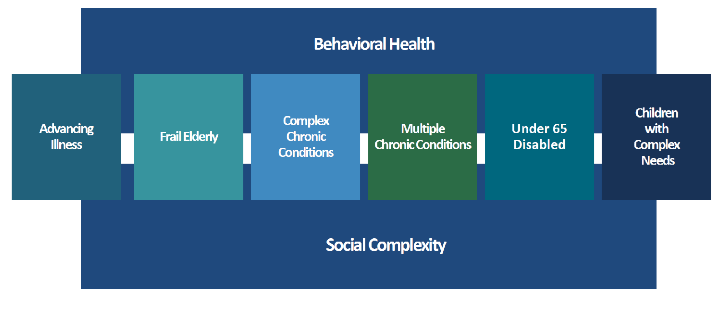
is to improve systems and outcomes for vulnerable populations through research, technology, and collaboration.

Identifying vulnerable people

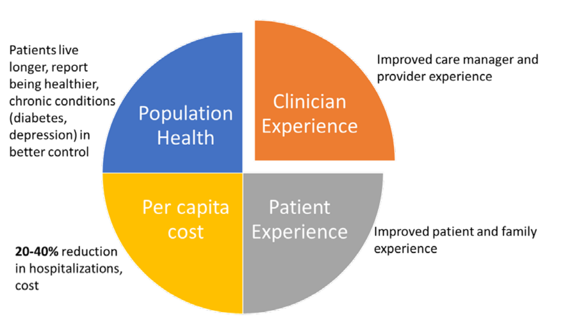
Risk stratification and segmentation



Tailoring care to these needs



Improving outcomes

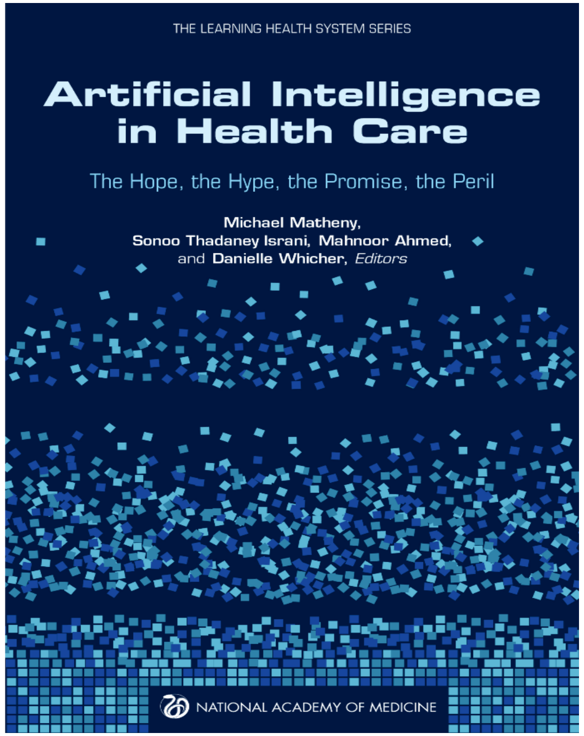


Funded by the John A. Hartford Foundation, AHRQ, NIH, Commonwealth Fund, Gordon and Betty Moore Foundation, CHCF, CMS, OHSU Foundation, San Francisco Health Plan, and the Oregon Health Authority

We are eager to implement advanced algorithms … but also cautious

Table 7
Methods, pros, cons, challenges and future directions in medical deep learning in 2019.

Publication	Methods	Pros	Cons	Challenges	Future Directions
Medical Imaging Biswas et al. [50]	AE, (fully) CNN, DBN, DRN, FCN, SVM	-DBM has easy inference -Automated feature extraction -Learning of complicated and composite relationships in data -DL methods surpass in robustness and performance	-Unknown generalization capabilities (DBN) -Vanishing gradient problems during training (AE)	-Improvement needed before techniques could be integrated into clinical workflows -Only trained on small datasets	-Widespread use in research and clinical routine -Develop real-time applications
Brain cancer classification Tandel et al. [51]	ANN, CNN, EM, KNN, NB, RF, SVM	-Automatically produce features that are stable to deformation and translation invariant -DL outperforms other ML methods	-Computationally more expensive	-Not discussed	-Provide the fast, non-invasive diagnosis tool that the field needs
Electroencephalogram Craik et al. [52]	AE, CNN, DBN, LSTM, MLP, RBM, RNN, SAE, SVM	-Successfully applied to motor imagery, seizure detection, mental workload, sleep stage scoring, event related potential, and emotion recognition	-Not discussed	-Formulation of the input data (PSD, wavelet decomposition, etc.)	-Combine convolutions and recurrent or RBM architectures -Use de-noised EEG data
Neuro-oncology Shaver et al. [54]	ANN, CNN, CRNN, LSTM, SVM	-Do not require human-constructed features -CNN architectures provide high accuracies on segmentation, characterization, grading and survival prediction tasks	-Requires large quantities of annotated data, necessitating medical expert knowledge and significant amounts of time -Overfitting	-Lack of large amounts of annotated data	-Undisruptive integration into workflows -Work with regulatory bodies who currently restrict the use of ML/DL in clinical practice
Diabetic retinopathy Asiri et al. [55]	AE, CNN, DBN, RNN	-Automatic discovery of relevant features -Ability to train and deliver solutions in an end-to-end manner -Successfully applied to vessel and optic disk segmentation, lesion detection and classification, diabetic retinopathy diagnosis	-Require large amounts of labeled data -Tendency to overfit -Convergence of DL methods not always guaranteed -Lack of interpretability -Class imbalance of datasets	-Lack of large-scale annotated uniform training data -Generalization of DL methods	-More standardization in data, labels, and test metrics -Research GANs



Egger J, Gsaxner C, Pepe A, et al. Medical deep learning-A systematic meta-review. *Comput Methods Programs Biomed*. 2022;221:106874. doi:10.1016/j.cmpb.2022.106874

Cognition : to act on the integrated diagnostics, several issues must be addressed

Frameworks for taking action on recommendations (AIM-ACT)

- **A**ffect
- **I**ntegration
- **M**otivation

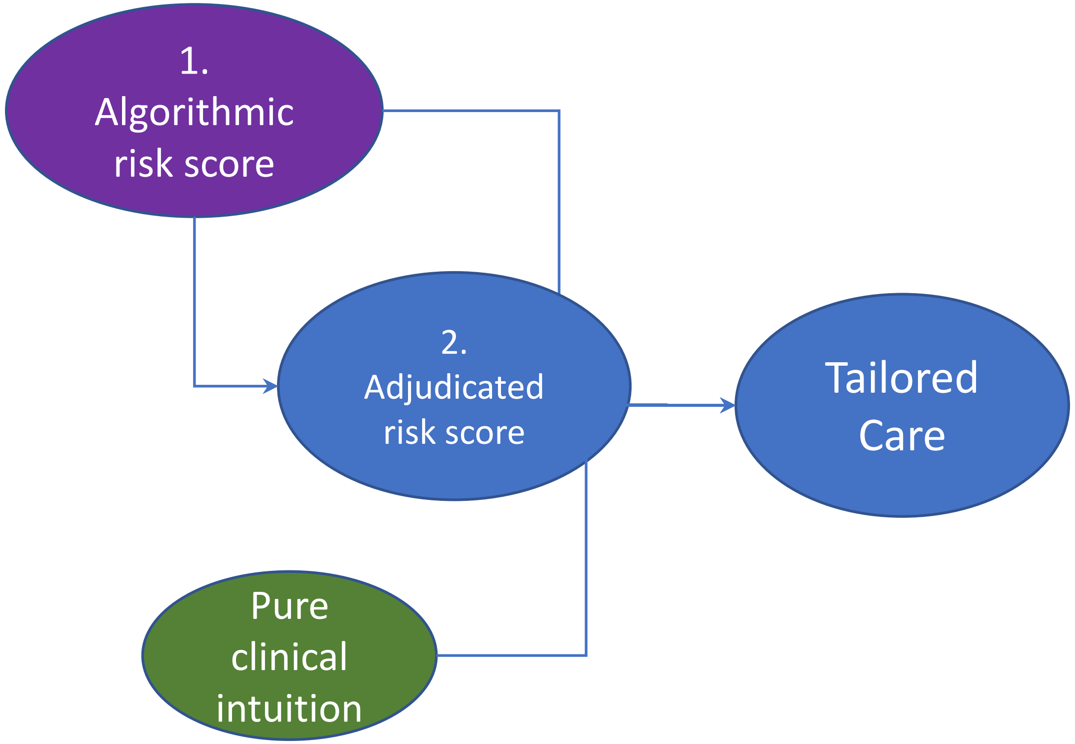
- **A**ttention
- **C**ontext
- **T**ranslation of Motivation into Action

Key issues

- Trust
- Understanding
- Timing, Options

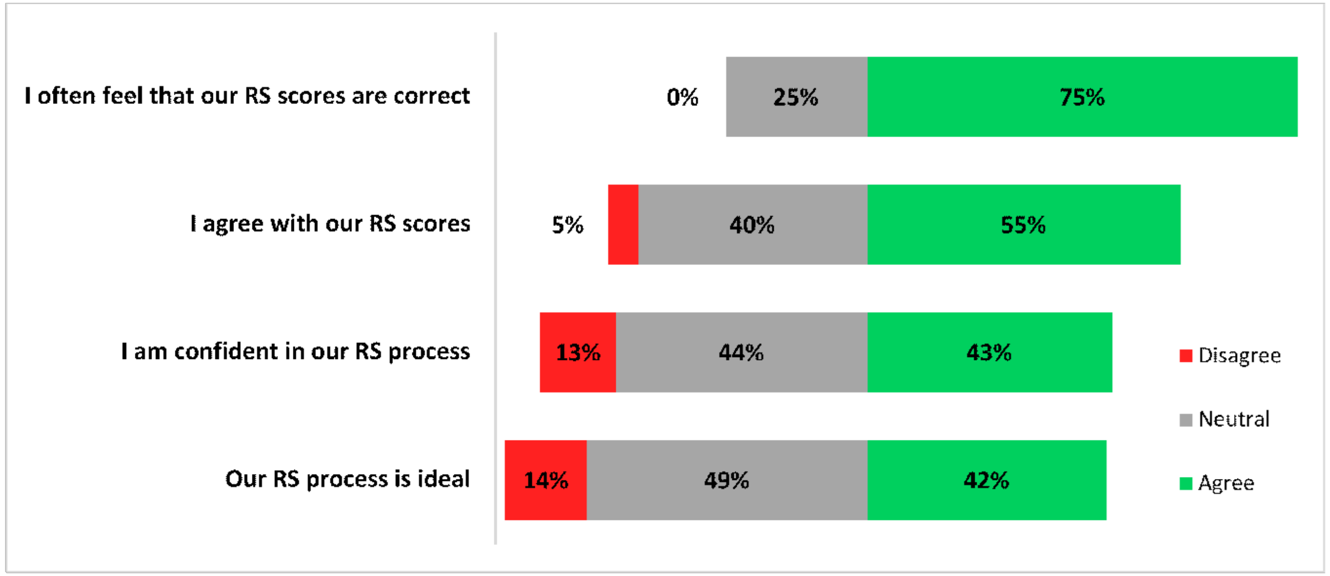
- Prioritization
- Adjudication / Annotation
- Actionability

Addressing cognition in integrating algorithms



“Of your patients with MCC, who would you not be surprised if they ended up in the hospital / died in the next year?”

Figure 1a. Individual perceptions of risk stratification outcomes.



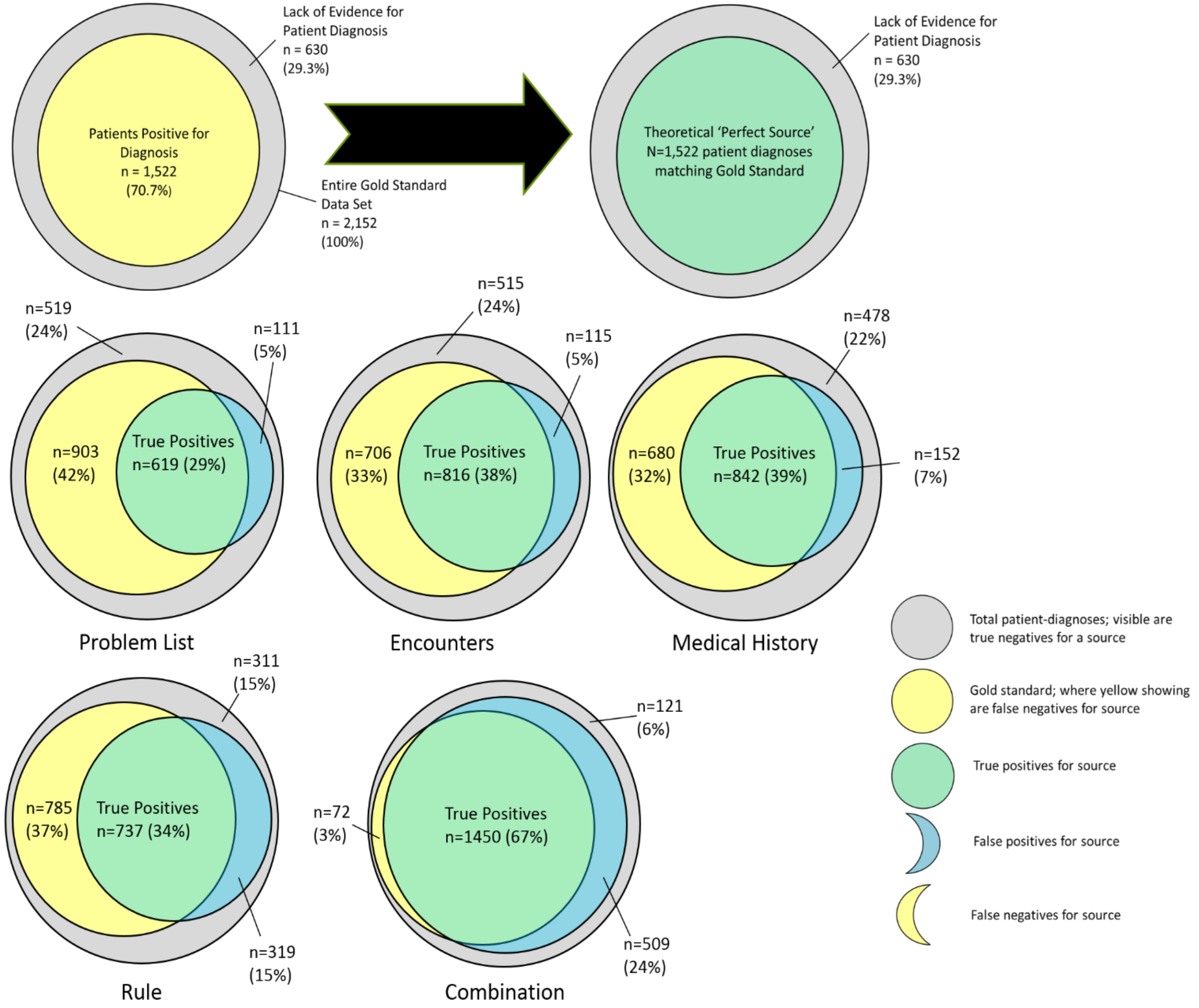
Ross et al/ BMC Med Inform Decis Mak.
2021 Mar 18;21(1):104. doi: 10.1186/s12911-021-01455-4.

Data is messy

Most algorithms require integrating 'real-world' data, such as from EHRs

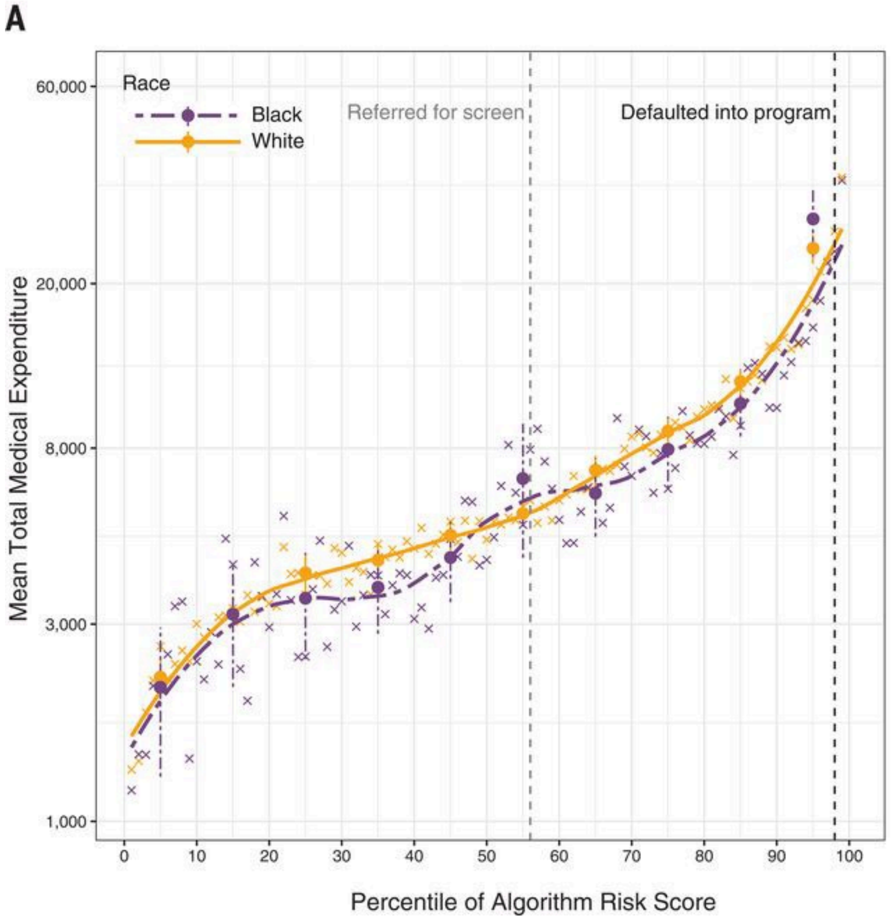
Data from EHRs is not collected for precise identification

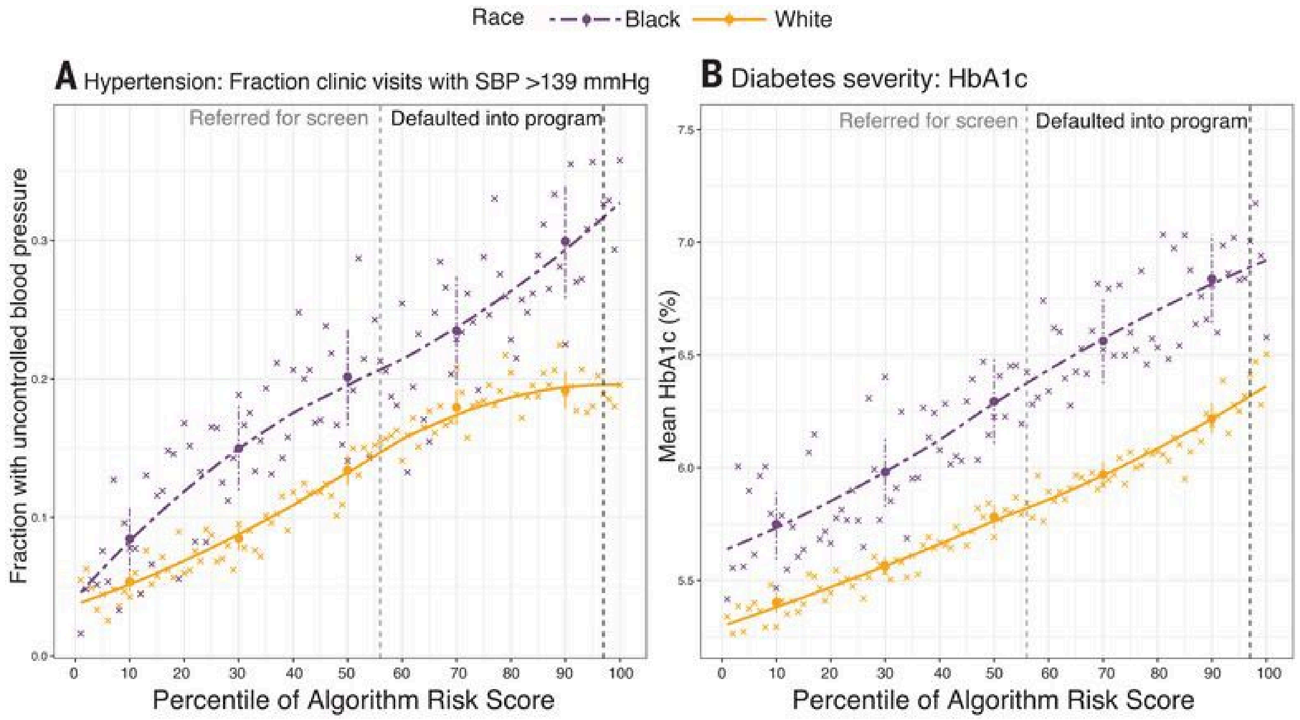
Thus, error rates for any condition as high as ~30%



Martin, ACI, 2017;
AHRQ grant number 1R21HS023091-01

Data contains biases (and complex data is worse)





These biases reflect our societal biases

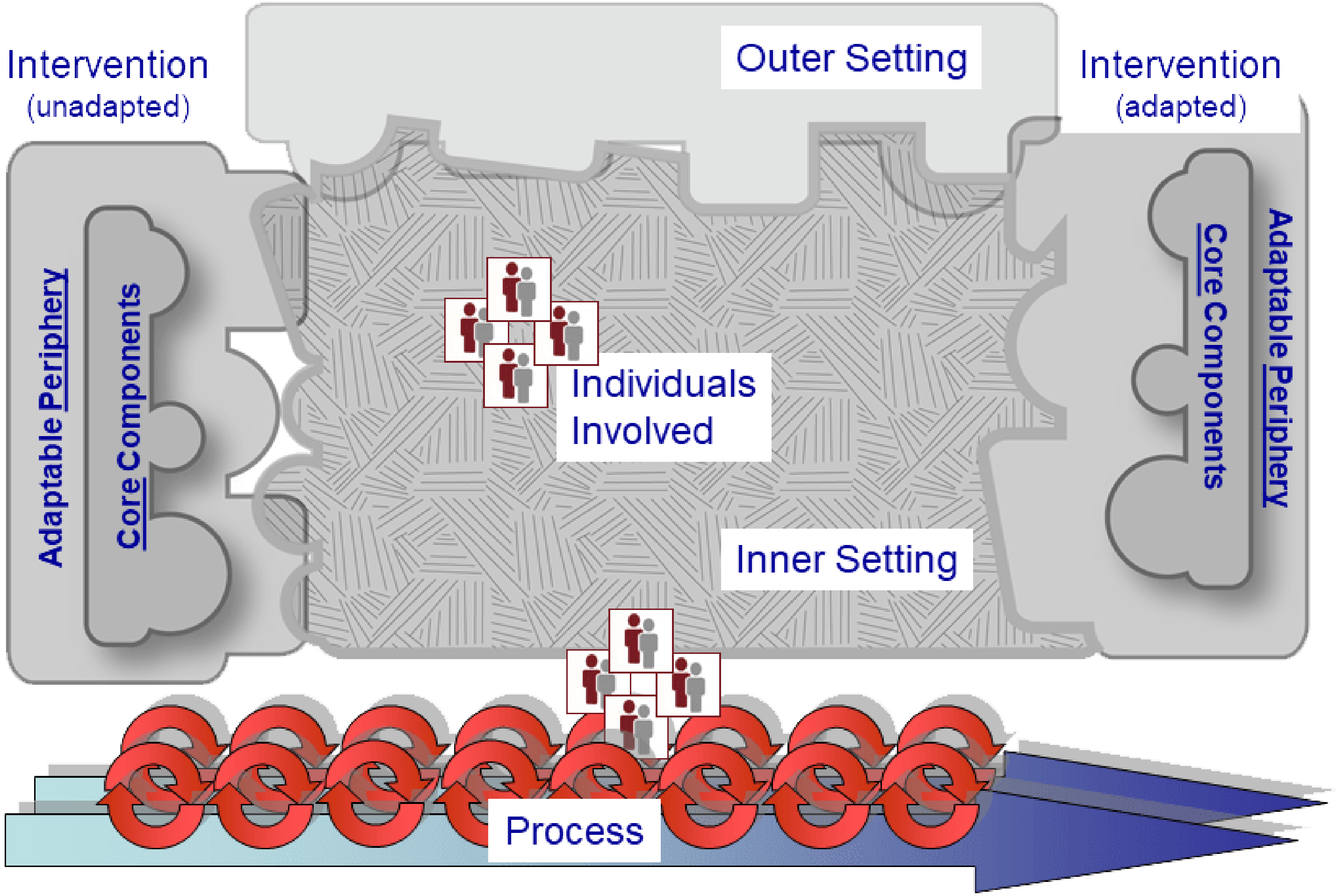
and the algorithms may carry biases depending on how they are developed

Selection biases, information biases, even 'colliders'

Ziad Obermeyer et al. Science 2019;366:447-453
Colliders and bias: Weiskopf et al. JAMIA 2023

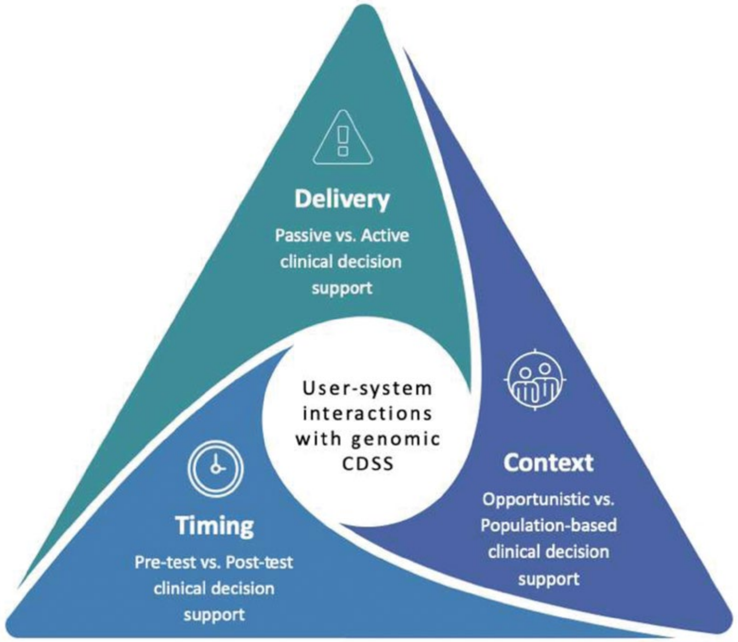
Science
AAAS

Integrating these issues through Implementation Science : a consolidated framework for implementation research (CFIR)



Damschroder et al BMC 2022 https://implementationscience.biomedcentral.com/articles/10.1186/s13012-022-01245-0

Examples of challenges in implementation and unintended consequences



- 12 eMERGE sites tried to implement genomic CDSS in care
- Issues across the 5 rights of CDS (the right information, to the right person, in the right intervention format, through the right channel, at the right time in workflow)
- Data sharing requirements grant patients the right to these data, but how to provide context and help with understanding
- Care teams need support to understand and exercise good judgment

Overby Taylor C et al. ACI Open 2021https://www.thieme-connect.com/products/ejournals/html/10.1055/s-0041-1729981

Thus, we need a code of conduct: Here's ours at OHSU

- For implementation, AI must meet these criteria:
  - Proven utility
  - A clear and evidence-based risk-benefit calculation.
  - Consideration of the ethics of the model, especially when vulnerable patient populations are involved.
  - Local validation and evaluation.
  - An implementation plan that includes monitoring for harms as well as benefit.
  - A training and support plan to help all persons at OHSU, including patients, understand and use the tools effectively and safely.

Van Calster *et al. BMC Medicine* (2023) 21:70
https://doi.org/10.1186/s12916-023-02779-w

BMC Medicine

OPINION Open Access

There is no such thing as a validated prediction model

Ben Van Calster[1,2,3], Ewout W. Steyerberg[1], Laure Wynants[1,2,4] and Maarten van Smeden[5*]

Abstract

Background Clinical prediction models should be validated before implementation in clinical practice. But is favorable performance at internal validation or one external validation sufficient to claim that a prediction model works well in the intended clinical context?

Main body We argue to the contrary because (1) patient populations vary, (2) measurement procedures vary, and (3) populations and measurements change over time. Hence, we have to expect heterogeneity in model performance between locations and settings, and across time. It follows that prediction models are never truly validated. This does not imply that validation is not important. Rather, the current focus on developing new models should shift to a focus on more extensive, well-conducted, and well-reported validation studies of promising models.

Conclusion Principled validation strategies are needed to understand and quantify heterogeneity, monitor performance over time, and update prediction models when appropriate. Such strategies will help to ensure that prediction models stay up-to-date and safe to support clinical decision-making.

Keywords Risk prediction models, Predictive analytics, Internal validation, External validation, Heterogeneity, Model performance, Calibration, Discrimination

CODE OF CONDUCT FOR ALGORITHMS IN HEALTH CARE

Proposed project of the NAM Leadership Consortium for a Value & Science-driven Health System

Compelling Aim: Ensure algorithms in health, health care, and biomedical science perform accurately, safely, reliably, ethically, and in a manner that advances knowledge leading to improved effectiveness, efficiency, and equity in the health and well-being of individuals and populations.

Issue: Algorithms that provide augmented or artificial intelligence and direct automated functions hold the promise to guide individual and population health decisions and reveal insights that can improve the effectiveness, reliability, person-centeredness, equity, and efficiency of care. However, the development and use of algorithms are subject to the fragility of the formula designs and data to which they are applied. The nature and quality of the data and the embedded analytic assumptions will inherently shape their accuracy and generalizability. For example, biases related to misplaced assumptions about data samples, representativeness of content, and analytics applied through algorithms can lead to erroneous conclusions and actions with the potential to not only miss health improvement opportunities, but also lead to unintended consequences, including the potential for worsening of health and societal inequities.

Given the increasing ubiquity and reliance of algorithms in health and health care research and practice, careful review and oversight is very important to the successful performance of advanced algorithms. The challenges are complex and stakeholders include federal level agencies, such as AHRQ, CDC, CMS, DOD, FDA, HRSA, IHS, ONC, and the VA, as well as virtually all public and private national, state, and local organizations and professional societies that develop and draw on data sets to guide decision-making. The impact of the evolving oversight on algorithm structure and function is unclear, but there is a pressing need for anchor principles to guide and assess, the development and use of health algorithms, drawing upon the relevant ethical frameworks, oaths, and standards germane to health care and biomedicine.

Approach: Under the direction of a multi-stakeholder, inter-disciplinary steering committee consisting of, but not limited to, health care providers, public and private payers, academia, public health departments, federal agencies, patients and community leaders, technology companies (including HIT and AI developers), ethics and equity experts, the National Academy of Medicine (NAM) will organize a set of activities to explore the components, characteristics, and implementation of a Code of Conduct for use in the development and application of algorithms in health, health care, and

Thank you!

dorrd@ohsu.edu