“THE LAST MILE” FOR AI IN HEALTHCARE: SUSTAINABLY IMPROVING PATIENT HEALTH OUTCOMES – THE ESSENTIAL ROLE OF MULTIDISCIPLINARY COOPERATION AND COLLEGIALITY

Ryan P. Moore, Henry J. Domenico, Daniel W. Byrne

Vanderbilt Departments of Biostatistics and Biomedical Informatics

Abstract

Artificial intelligence (AI) and machine learning have made great strides in healthcare, however, these fields still fall behind others in medicine when it comes to using rigorous science and measurably improving care for patients. Here we describe the milestones to sustainably improve patient outcomes using AI and machine learning: project groundwork, model development and validation, model implementation, and rigorous evaluation. Additionally, we outline the necessary milestone components, along with challenges and common pitfalls. We present examples of past and current projects with lessons learned. The Vanderbilt Clinical Informatics Center has been key to the success of these projects.

Both AI and healthcare are infinitely complex and successfully integrating even a small aspect of these fields requires a deep and mature understanding of what AI can do and what healthcare needs. Equally important is understanding what AI cannot do and what healthcare does not need. Much of the slow progress can be attributed to misunderstandings. AI experts have misunderstandings about the problems in healthcare and healthcare professionals have misunderstandings about the solutions that AI tools can provide.

A pragmatic randomized controlled trial is the key to real progress of AI in healthcare. Implementation of AI must be compared with current usual medical care. These comparisons must assess real-world evaluations for effectiveness, harm, and unintended consequences. Using observational data and adjusting for confounding factors is inadequate for assessing the impact of AI on patients. In these studies, regression to the mean is the Achilles’ heel of AI research. Regression to the mean occurs when one selects outliers at baseline and then assesses these outliers in the future. These outliers nearly always move closer to the mean, or average, of the entire group, without imposing any intervention on them. AI researchers must learn to recognize this and avoid flawed conclusions by using stronger study designs.

The ultimate goal is to have clinicians adopt AI models as their complementary work partners. AI will augment a specific task, not automate an entire job in healthcare. Success requires leveraging the scalable, fast, and economical benefits of AI to predict and classify many clinical events. We must design a frictionless nondisruptive point of care AI user interface. The technology should run silently in the background. The AI model is not a “magic bullet”. Understanding how physicians make decisions in the clinical workflow and implementing AI tools seamlessly with an effector arm is the key to success. AI researchers should plan for an iterative approach. The first implementation with the AI tool is unlikely to improve patient outcomes; perhaps the second or third approach will. Much of the hard work involves having clinicians incorporate the AI tool in their workflow.

In summary, most healthcare organizations are struggling to execute AI. Overcoming this requires identifying and recognizing how previous AI implementation for medicine failed and solve these problems with modern reproducible research.

The Last Mile for Improving Health Outcomes with AI

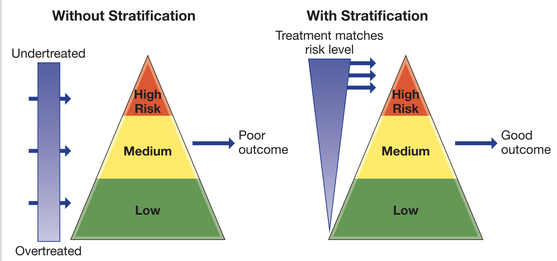
Researchers Often End Projects After Model Development

Project Groundwork	Model Development and Validation	Model Implementation	Rigorous Evaluation

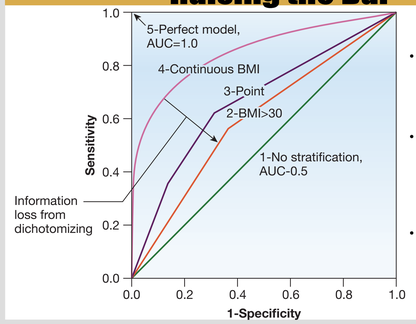
• Clinical team and leaders responsible for clinical outcome need to be involved from earliest stages of project.
• P.I. of project should be a clinician.
• Obtaining data in analyzable format currently takes significant time and resources and is a major bottleneck for investigators.
	• Previously published models often do not meet requirements for implementation:
	  • Applicable to general patient population.
	  • Use variables that are routinely collected during a patient’s stay.
	  • Use predictors that avoid reverse-causation.
	  • Do not dichotomize predictor variables.
	• External models need to be validated and calibrated within our patient population.
		• Display risk predictions in EHR or within a module in a way that is meaningful for users. Clinician input is crucial.
		• Risk predictions should be transparent and show the key drivers.
		• Clinical team develops approach for how to handle high-risk patients.
		• Display in a way that can be rigorously studied.
			• A pre/post evaluation or comparison of clinician-selected patients to non-selected patients will be fatally biased. Adjustment methods will not eliminate these biases.
			• Randomization allows for straightforward, unbiased estimation of model-guided intervention effect overall and within subgroups.
			• Randomization is the best way to test for harm and bias in potentially vulnerable populations and predict patient-level treatment effects.

AVAIL Goal
Sustainably improve patient outcomes with machine learning and AI

Risk Stratification is Crucial

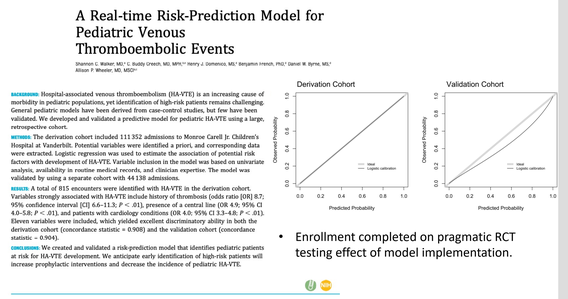


Raising the Bar



- Model performance should not be judged based solely on apparent performance
- Calibration and discrimination in external validation data are what are most important
- AI often has better AUC (discrimination), but worse Calibration

CLOT Trial


A Real-time Risk-Prediction Model for Pediatric Venous Thromboembolic Events

Shannon C. Walker, MD,* C. Buddy Creech, MD, MPH,** Henry J. Domenico, MS,* Benjamin Irvinch, PhD,* Daniel W. Byrne, MS,* Allison P. Wheeler, MD, MSCI**

Derivation Cohort

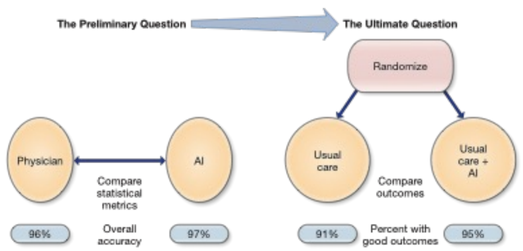
Validation Cohort

Observed Probability

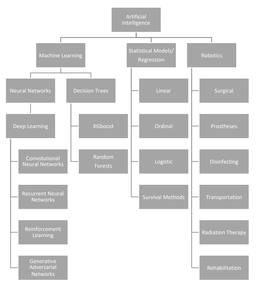
Predicted Probability

- Enrollment completed on pragmatic RCT testing effect of model implementation.

The Secret Sauce - *Randomization*



Choosing the Best Tool for the Job



- Artificial intelligences encompasses a broad range of techniques, algorithms and models – each with their own strengths and weaknesses
- Practice model agnosticism and choose the best tool for the job

Teamwork

- Advanced Vanderbilt Artificial Intelligence Laboratory (AVAIL)
- Quality, Safety, and Risk Prevention (QSRP)

VCLIC
VANDERBILT CLINICAL INFORMATICS CENTER

Conclusions/Take-home messages

- Multidisciplinary collaboration is vital for the success of AI integration into healthcare.
- Projects need team members with clinical, informatics, and biostatistical expertise to be successful.
- Use the best tool for the job – we should practice model agnosticism.
- Avoid methods that ignore information by using dichotomized predictors or scoring systems.
- AI as a field needs rigorous RCTs to test effectiveness, protect against harm and bias, and improve care for patients.
- Pragmatic trials of AI can be low cost and doable but very few have been carried out.
- AVAIL-supported projects are advancing healthcare and AI down the last mile.

Reference: Byrne DW. *Artificial Intelligence for Improved Patient Outcomes - Principles for Moving Forward with Rigorous Science.* Wolters Kluwer; 2022.

This work was supported by AVAIL and the Department of Quality, Safety and Risk Prevention

Please contact the corresponding author at henry.domenico@vumc.org

VANDERBILT UNIVERSITY®