

ARTIFICIAL INTELLIGENCE FOR IMPROVED PATIENT OUTCOMES – THE PRAGMATIC RANDOMIZED CONTROLLED TRIAL IS THE SECRET SAUCE

Daniel W. Byrne, Henry J. Domenico, Shannon C. Walker, Ryan P. Moore

From AVAIL (the Advanced Vanderbilt Artificial Intelligence Laboratory) and the Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN.

ABSTRACT

Artificial Intelligence (AI) has exploded in the media for both its astonishing power and disturbing weaknesses. Of the potential applications of AI that are most likely to benefit society, most thought leaders point to medicine. Yet, to date, we have almost no rigorous evidence that AI improves patient health outcomes. Why is there a dearth of evidence? What needs to change?

In one word: "Randomization". The excuses for why randomization is not an option are plentiful and roll off the tongue, yet in our experience the excuses frequently turn out to be unjustified. The reason that we have safe and effective prescription drugs is that the US Food and Drug Administration (FDA) demands evidence from rigorous randomized controlled trials (RCT). The FDA would never consider the excuses used for not randomizing that are tolerated in the AI field.

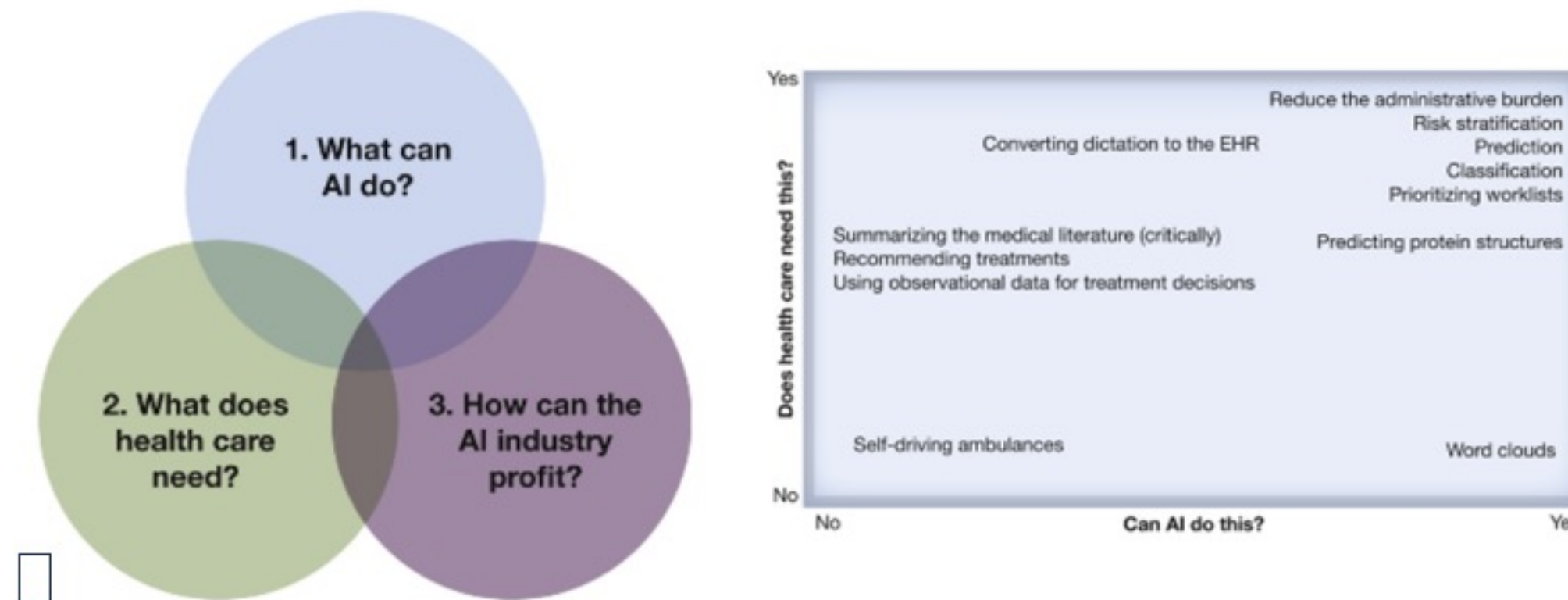
Using AI algorithms to improve patient health outcomes can be viewed as progress on the "last mile" between a model creation and improved outcomes. Building an AI model that works is only the first step. The application must be accepted and seamlessly integrated into the healthcare workflow in a frictionless manner. This work requires a collaborative multidisciplinary team of AI experts, physicians, nurses, informaticians, biostatisticians, and others. AI tools should be evaluated in the healthcare workflow using pragmatic patient-level randomized controlled trials. Because successful implementation will likely require several iterations, an adaptive platform trial is optimal; an adaptive platform trial enables multiple interventions to be tested in parallel, dropping study arms for futility and adding new arms in a perpetual manner without long delays between studies. For example, various forms of escalation can be assessed regarding changing clinician behavior.

By using pragmatic RCTs to know when AI has moved the needle regarding outcomes, healthcare can move more rapidly into a modern era of AI that benefits both patients and clinicians. Rather than comparing the performance of AI models to physician decisions, future trials should measure the performance of physician alone versus physician assisted by AI models.

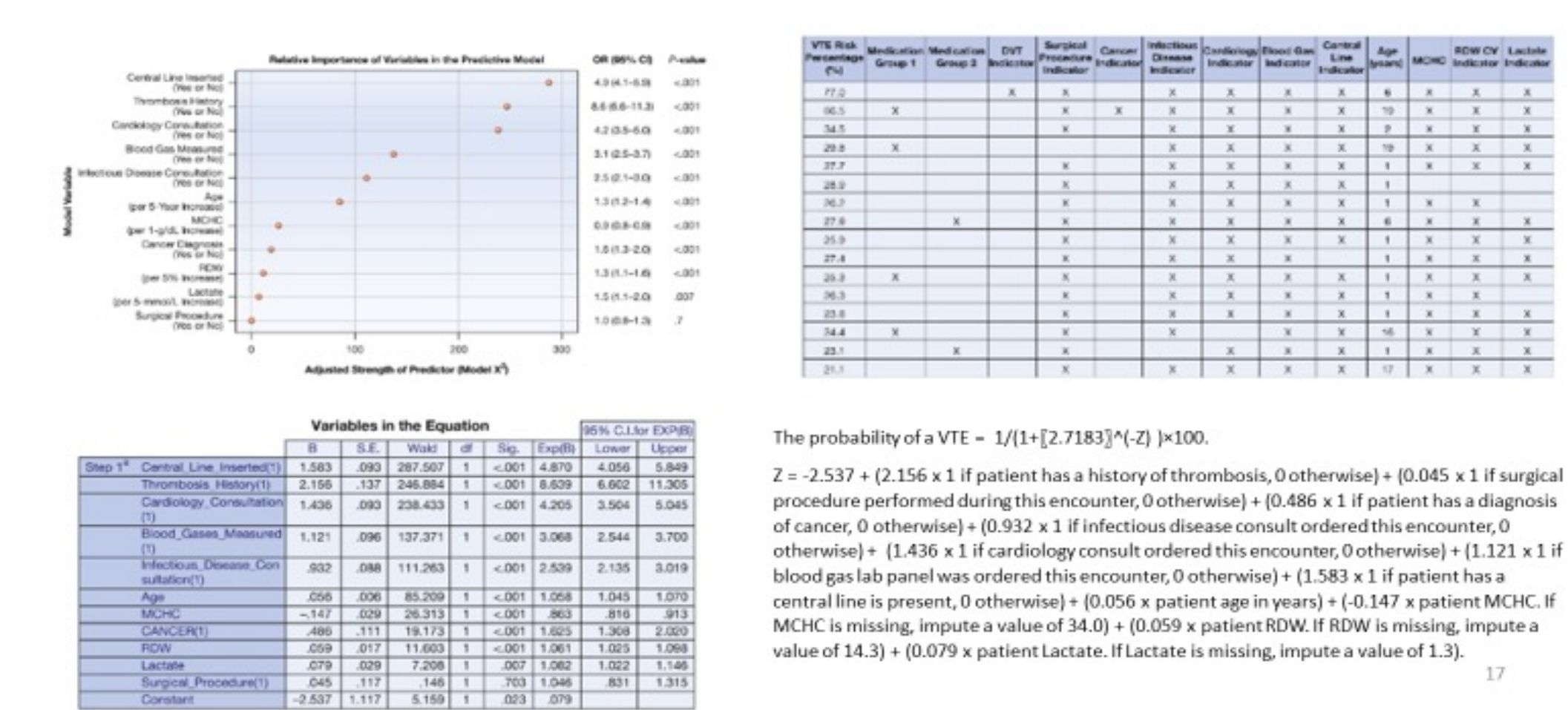
In 2016, the "godfather of AI," Geoffrey Hinton, said "People should stop training radiologists now. It is just completely obvious that within 5 years deep learning is going to do better than radiologists." Perhaps this should be updated to "People should start training physicians to create, use and evaluate AI tools in a modern and rigorous manner." This training needs to include AI evaluation skills, such as: regression to the mean, reverse causation, the issues with overall accuracy, residual confounding, pragmatic patient-level randomized controlled trials, and adaptive platform trials.

Once we embrace reproducible research in the form of pragmatic randomized controlled trials with prespecified study designs and end points, Artificial Intelligence will begin to improve patient outcomes – overall and within important subgroups. AI also holds great potential to reduce the workload and burnout among physicians. Despite the concerns, AI will not replace physicians - but physicians who use AI effectively will replace physicians who do not. The pragmatic randomized controlled trial is the secret sauce - but only if scientists with modern training lead these projects.

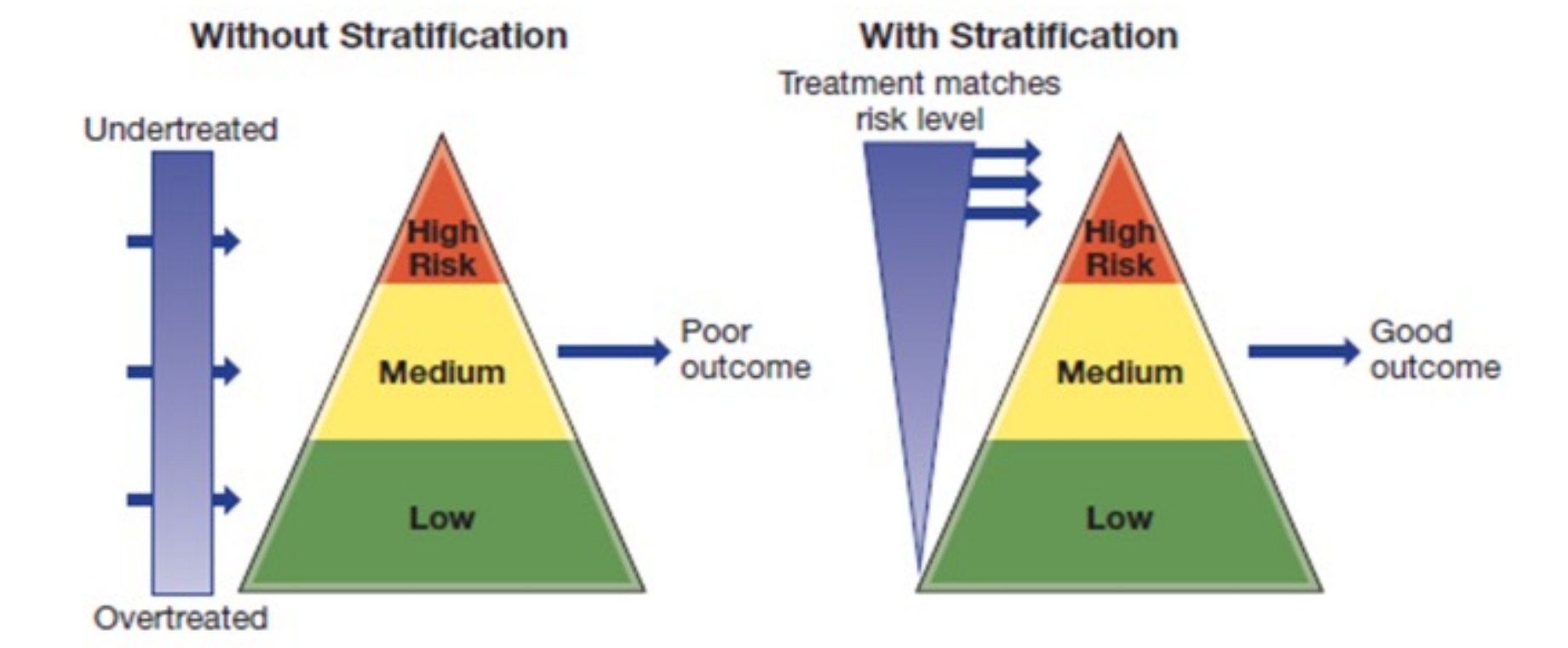
The Big Question:
Why is there a dearth of evidence that AI results in improved health for patients?



CLOT – Children’s Likelihood Of Thromboembolism



Risk stratification is essential for focused prevention and treatment.

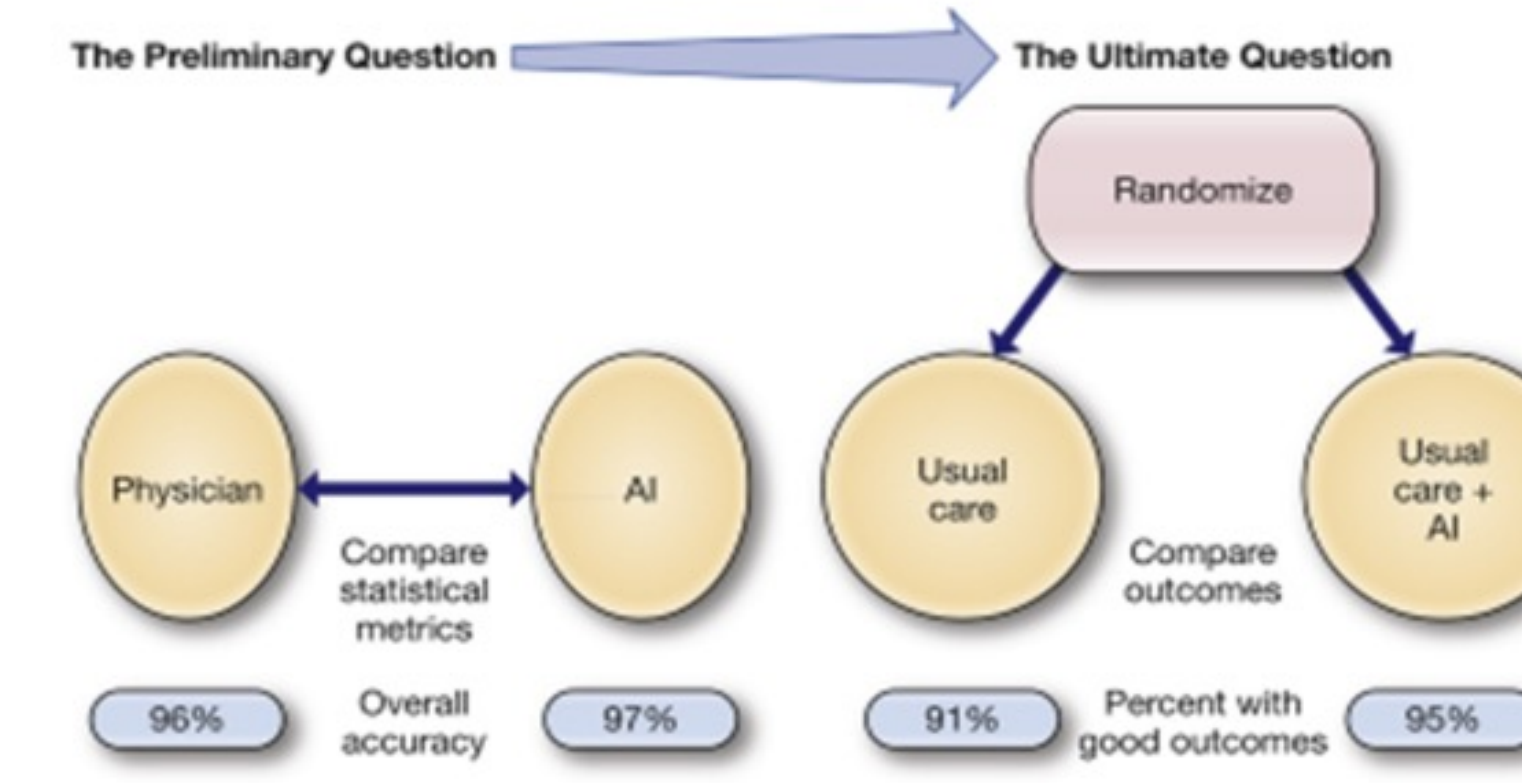


Identify an accuracy scam, which is one of the most common forms of pseudo-innovation in hyping AI.

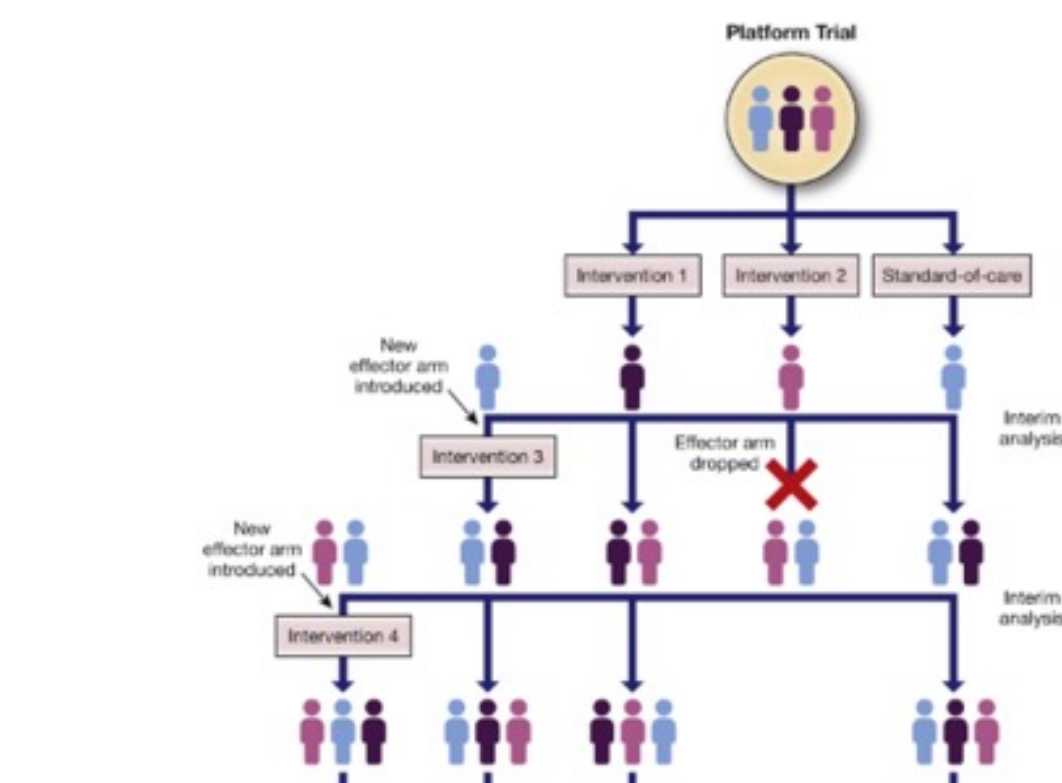
	No Hemochromatosis	Hemochromatosis	Total
Predicted No	996	4	1000
Predicted Yes	0	0	0
Total	996	4	1000

Overall accuracy = (996 + 0)/1000 = 99.6%

Focus on the Big Question.



Use Adaptive Platform Trials (APT) to assess specific forms of implementation.



Vaporware of what a primary care provider will see in the EHR at an annual checkup.

Health Condition	Current	Suboptimal	1 Year	5 Years	10 Years
Aspirin	95%	5%	95%	95%	95%
Cholesterol	85%	15%	85%	85%	85%
Diabetes	75%	25%	75%	75%	75%
Hypertension	65%	35%	65%	65%	65%
Smoking	55%	45%	55%	55%	55%
Weight	45%	55%	45%	45%	45%
Depression	35%	65%	35%	35%	35%
Alcohol	25%	75%	25%	25%	25%
Medication	15%	85%	15%	15%	15%
Immunizations	5%	95%	5%	5%	5%
Other	1%	99%	1%	1%	1%

Use AI tools to prevent complications before they happen.

Complication	Risk on admission	Current risk
Pressure injury	15%	30%
Readmission	14%	15%
Acute kidney injury	7%	8%
Blood clot (venous thromboembolism)	2%	7%
Hospital acquired infection	2%	4%
Adverse drug event	1%	2%
Delirium	1%	2%
Fall in the hospital	1%	1%
ECU readmission	1%	1%
Pneumonia	1%	1%
Torso traumatic stress disorder	1%	1%
Unintended death	1%	1%
Clostridiales difficile infection	0%	1%
Methicillin-resistant Staphylococcus aureus (MRSA) infection	0%	1%
Significant complication after surgery	2%	0%
Ventilator-associated pneumonia (VAP)	1%	0%
Catheter-associated urinary tract infection (CAUTI)	0%	0%
Central line bloodstream infection (CLABSI)	0%	0%
Surgical site infection	0%	0%
Need for hospice	0%	0%
Non-ICU cardiac arrest	0%	0%

Aligning with Leadership: Enhancing Care & Efficiency

- Enhancing Patient Care:**
- Improves patient outcomes.
 - Champions a culture of safety.
 - Advocates for personalized medicine through rigorous science.
 - Measures progress consistently towards VUMC Quality Pillar Goals.
 - Minimizes bias; ensures effectiveness for all.
- Optimizing Resources & Time:**
- Streamlines tasks, eliminating low-priority items.
 - Reduces clinician burnout, aiding in recruitment and retention.
 - Harnesses EHR data effectively.
 - Uses clinical and prevention resources efficiently; reduces wasted effort.
- Promoting Teamwork:**
- Fosters collaboration across departments.
 - Shares ownership of projects between clinical, informatics, analytics, and research teams.
 - Positions VUMC as a top-tier name in AI and safety research.

REFERENCES

Byrne DW. Artificial Intelligence for Improved Patient Outcomes: Principles for Moving Forward with Rigorous Science. Wolters Kluwer. 2023.
 Walker SC, Creech CB, Domenico HJ, French B, Byrne DW, Wheeler AP. A Real-time Risk-Prediction Model for Pediatric Venous Thromboembolic Events. Pediatrics. 2021.
 Walker SC, French B, Moore RP, Domenico HJ, Wanderer JP, Mixon AS, Creech CB, Byrne DW, Wheeler AP. Model-Guided Decision-Making for Thromboprophylaxis and Hospital-Acquired Thromboembolic Events Among Hospitalized Children and Adolescents - The CLOT Randomized Clinical Trial. JAMA Network Open. 2023;6(10).

