

The Role of Pragmatic Implementation Science Methods in Achieving Equitable and Effective Use of Artificial Intelligence in Healthcare



Anna M. Maw, MD, MS^{1,2} , Katy E. Trinkley, PharmD, PhD^{1,3}, and Russell E. Glasgow, PhD^{2,3}

¹Adult and Child Center for Outcomes Research and Delivery Science Center, Aurora, CO, USA; ²Division of Hospital Medicine, Department of Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, 12605 E 16th Ave, Aurora, CO 80045, USA; ³Department of Family Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

J Gen Intern Med 39(7):1242–4

DOI: 10.1007/s11606-023-08580-y

© The Author(s), under exclusive licence to Society of General Internal Medicine 2023

There is growing consensus that evaluation of the effects of artificial intelligence (AI) interventions within health systems is needed to ensure its safe, equitable, and patient-centered use. Given the speed with which AI is developing, the phrase “building the plane while flying it” could not be more apt. For AI’s full potential to be realized, methods that ensure its effectiveness is replicated across diverse clinical environments, benefits are distributed equitably, and adverse consequences are minimized must be quickly developed and implemented.

Pragmatic implementation science methods¹ that assess and enhance the impact of complex interventions in real-world environments have great potential to help AI achieve its goals while minimizing adverse unintended consequences including patient harm, system inefficiency, and disparities in care delivery in ways that are replicable. We use the example of predictive AI sepsis alerts which are already commonly used, and the Practical Robust Implementation Sustainability Model (PRISM)² (Fig. 1), a frequently used implementation science framework, to illustrate these methods.

AI INTERVENTIONS ARE COMPLEX AND CONTEXT-DEPENDENT

When considering what methods are needed to assess and enhance replicability across clinical settings, it is important to recognize AI innovations satisfy the definition of a complex intervention.³ Beyond the predictive AI model itself, additional intervention components are necessary to guide actions taken in response to a given AI prediction (Fig. 2). These “decision” components are environment- or context-dependent. They must be tailored to unique aspects of the clinical setting, including organizational culture, workflow, and infrastructure for the AI intervention to produce the desired outcome.

AI model performance also varies with context and data available. The “brittleness” or inability of AI models to maintain their predictive performance when applied to data sets other than those they were trained on is one of the most important challenges to AI improving clinical care. Thus, like other complex interventions, the benefits of AI are unlikely to be realized when implemented in new contexts without iterative tailoring, or adaptations, using carefully selected implementation strategies.

For instance, an AI model used to predict sepsis developed at one hospital will likely have worse prediction accuracy when initially deployed in another, requiring retraining on local data. Further, predictive AI models can only offer a probability. The intervention designers must decide on the probability above which a clinician will be notified. The clinician must then prioritize their response to the alert given multiple competing demands which are influenced by unique contextual factors such as the size and acuity of their patient census. This example illustrates why AI interventions are not likely to maintain the same magnitude of effectiveness when initially deployed in a new context.

Implementation science frameworks, like PRISM (Fig. 1), are useful in planning, implementing, and maintaining complex health interventions because they provide a scaffold by which to measure multiple contextual factors, process, and clinical outcomes across settings and subgroups. PRISM facilitates monitoring and iterative data-driven adjustments until the desired outcomes are achieved (Fig. 2). For instance, in addition to mortality, the RE-AIM constructs of PRISM support measuring process outcomes, like time to antibiotic administration, that are critical to understanding the effectiveness results of a sepsis alert. The contextual domains of PRISM facilitate understanding aspects of the environment that influence outcomes. Qualitative methods are often used to capture contextual drivers of outcomes that are otherwise difficult to capture, like distrust of an AI model’s prediction accuracy.

HEALTH EQUITY OUTCOMES MUST BE MONITORED TO PREVENT BIAS

Bias can be introduced at every phase of the AI “lifecycle,” from data creation to model deployment.⁴ Given how easily AI models can incorporate and conceal bias, proactive

Received September 11, 2023

Accepted December 11, 2023

Published online January 3, 2024

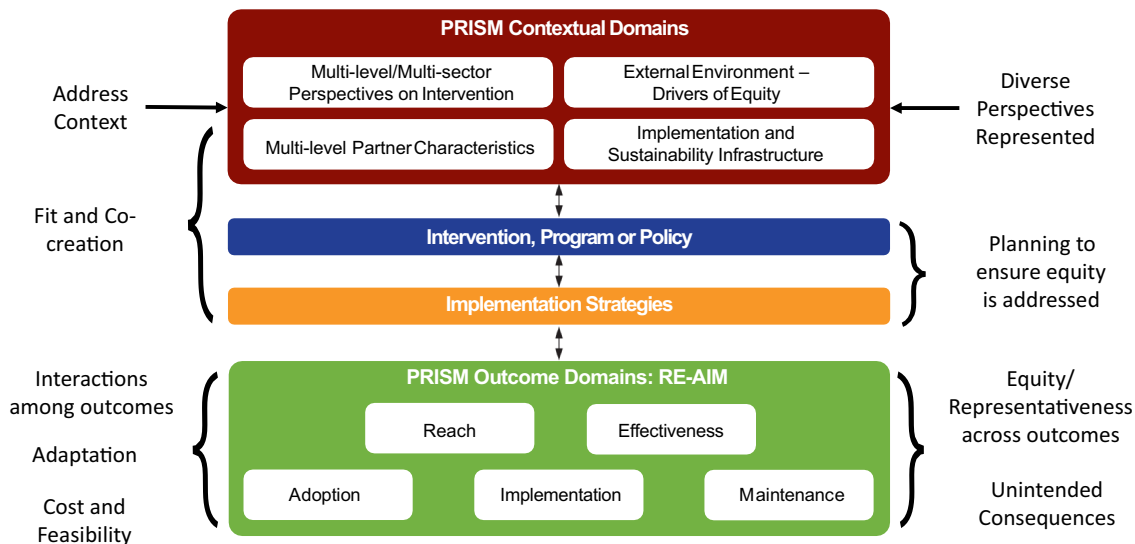


Figure 1 How the Practical Robust Implementation Sustainability Model (PRISM) facilitates equitable implementation of health interventions.

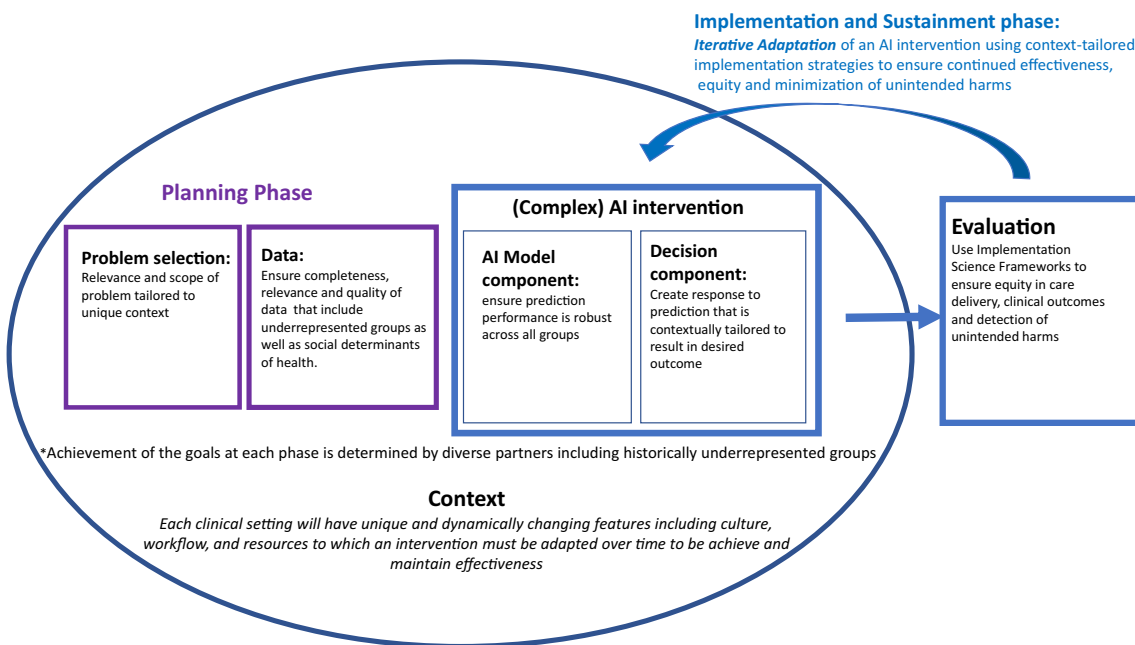


Figure 2 Using pragmatic implementation science methods to plan, implement, and sustain effective, equitable AI interventions.

and iterative monitoring of both care delivery and clinical outcomes that can rapidly identify and address disparities is needed. While the “black box” aspect of many AI models has been cited as an important barrier to trust and detection of bias, close monitoring by health systems to assess for implementation and outcome disparities can mitigate these drawbacks. This approach can also help address the lack of representativeness in existing data by promoting a higher level of scrutiny and transparency with regard to the completeness, relevance, and quality of data, ensuring appropriate inclusion of historically underrepresented populations and behavioral, environmental, and social determinants of

health (SDoH) measures. It is also important to consider whether adequate quantity of data is available to meaningfully apply AI in equitable ways. To avoid these potential pitfalls, they should be considered from the beginning when the problem and AI model are specified and iteratively revisited throughout the AI “lifecycle.”

For example, RE-AIM implementation outcomes, with their emphasis on representativeness, can measure whether an AI sepsis alert is being delivered at the same rate, in the same way, and resulting in the same outcomes across demographics like race/ethnicity and other SDoH measures. If, for instance, worse outcomes are found in non English-speaking

patients, qualitative data combined with quantitative process outcomes can be used to understand drivers of inequity, identify strategies to address them, and reevaluate outcomes once targeted implementation strategies have been deployed to assess if the disparity has resolved.

IMPLEMENTATION SCIENCE METHODS PAIRED WITH A LEARNING HEALTH SYSTEM INFRASTRUCTURE WILL MAKE TAILORING, RIGOROUS EVALUATION, AND MONITORING OF AI INTERVENTIONS FEASIBLE

The integration of pragmatic implementation science methods with the evolving informatics-driven learning health system (LHS)⁵ can help facilitate both the equity and replicability of AI intervention effectiveness in diverse contexts. As LHS infrastructures advance, the speed, feasibility, and robustness of implementation science–guided evaluations will also grow. Contextual, process, and effectiveness data that previously took days to months to collect can now be queried and displayed to implementers in real time allowing for more rapid, iterative adaptations to optimize the fit of the AI intervention with its context and correct any unanticipated negative outcomes. This LHS informatics infrastructure also makes pragmatic randomized trials⁶ and interrupted time series designs more feasible allowing for more accurate estimates of AI on the quintuple aim: clinical effectiveness, health equity, cost, patient and clinician experience.

For example, an operational, automated dashboard displaying RE-AIM outcomes⁷ of a sepsis alert populated with data extracted from the EHR allows for close monitoring of intervention delivery, effectiveness, and unintended harms in a manner that requires minimal health system resources to perform. Rapid qualitative assessments can be performed in response to these quantitative interval evaluations to understand drivers of desired outcomes.

The RE-AIM outcome of effectiveness can display not only the relative mortality rate associated with the sepsis alert but also balancing measures such as rates of *Clostridioides difficile* infection. Iterative, qualitative methods can be deployed to identify possible unintended consequences. For instance, if nurses express concerns that the sepsis alerts delay them from performing other duties, implementers can monitor patient falls and pressure wounds in response. Because healthcare environments are both complex and

dynamic, the iterative evaluations and adaptations should be continued even after an intervention has demonstrated effectiveness. This will ensure its *continued* effectiveness and equity *over time* in an ever-changing context.

In conclusion, the integration of pragmatic implementation science methods and the LHS can provide for the informed design, feasible monitoring, and iterative tailoring of AI interventions essential for effective and equitable use. Application of these approaches can offer a path for realization of AI's great potential to propel healthcare toward achievement of the quintuple aim.

Corresponding Author: Anna M. Maw, MD, MS, , Adult and Child Center for Outcomes Research and Delivery Science Center, Aurora, CO, USA (e-mail: anna.maw@cuanschutz.edu).

Funding REG was partially supported by the National Cancer Institute's Implementation Science Center grant P50CA244688. KET was partially supported by NHLBI K23 grant HL161352.

Declarations:

Conflict of Interest: AMM is a consultant for UltraSight.

REFERENCES

1. **Brownson RC, Colditz GA, Proctor EK.** Dissemination and Implementation Research in Health: Translating Science to Practice. Oxford University Press; 2023.
2. **Fort MP, Manson SM, Glasgow RE.** Applying an equity lens to assess context and implementation in public health and health services research and practice using the PRISM framework. *Front Health Serv Manage* 2023;3:1139788.
3. **Skivington K, Matthews L, Simpson SA, et al.** A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 2021;374:n2061.
4. **Ng MY, Kapur S, Blizinsky KD, Hernandez-Boussard T.** The AI life cycle: a holistic approach to creating ethical AI for health decisions. *Nat Med* 2022;28(11):2247–9.
5. **Trinkley KE, Ho PM, Glasgow RE, Huebschmann AG.** How dissemination and implementation science can contribute to the advancement of learning health systems. *Acad Med* 2022;97(10):1447–58.
6. **Weinfurt KP, Hernandez AF, Coronado GD, et al.** Pragmatic clinical trials embedded in healthcare systems: generalizable lessons from the NIH Collaboratory. *BMC Med Res Methodol* 2017;17(1):144.
7. **Maw AM, Morris MA, Glasgow RE, et al.** Using Iterative RE-AIM to enhance hospitalist adoption of lung ultrasound in the management of patients with COVID-19: an implementation pilot study. *Implement Sci Commun* 2022;3(1):89.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.