

Viewpoint

Applying Machine Learning Techniques to Implementation Science

Nathalie Huguet^{1,2*}, PhD; Jinying Chen^{3,4,5*}, PhD; Ravi B Parikh^{6*}, MPP, MD; Miguel Marino^{1,2*}, PhD; Susan A Flocke^{1,2*}, PhD; Sonja Likumahuwa-Ackman^{1,2*}, MID, MPH; Justin Bekelman^{6,7*}, MD; Jennifer E DeVoe^{1,2*}, MD, DPhil

¹Department of Family Medicine, Oregon Health & Science University, Portland, OR, United States

²BRIDGE-C2 Implementation Science Center for Cancer Control, Oregon Health & Science University, Portland, OR, United States

³Section of Preventive Medicine and Epidemiology, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

⁴Data Science Core, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

⁵iDAPT Implementation Science Center for Cancer Control, Wake Forest School of Medicine, Winston-Salem, NC, United States

⁶Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

⁷Penn Center for Cancer Care Innovation, Abramson Cancer Center, Penn Medicine, Philadelphia, PA, United States

* all authors contributed equally

Corresponding Author:

Nathalie Huguet, PhD

Department of Family Medicine

Oregon Health & Science University

3181 SW Sam Jackson Park Road

Portland, OR, 97239

United States

Phone: 1 503 494 4404

Email: huguetn@ohsu.edu

Abstract

Machine learning (ML) approaches could expand the usefulness and application of implementation science methods in clinical medicine and public health settings. The aim of this viewpoint is to introduce a roadmap for applying ML techniques to address implementation science questions, such as predicting what will work best, for whom, under what circumstances, and with what predicted level of support, and what and when adaptation or deimplementation are needed. We describe how ML approaches could be used and discuss challenges that implementation scientists and methodologists will need to consider when using ML throughout the stages of implementation.

(*Online J Public Health Inform* 2024;16:e50201) doi: [10.2196/50201](https://doi.org/10.2196/50201)

KEYWORDS

implementation science; machine learning; implementation strategies; techniques; implementation; prediction; adaptation; acceptance; challenges; scientist

Introduction

Implementation science is a research field developing and testing methods and strategies that can improve the uptake of evidence-based interventions (EBIs) and practices into routine use in targeted settings [1]. It has important applications in both clinical and public health settings, such as health care facilities, public health departments, schools, and workplaces [2-4]. For example, the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) framework, which was proposed by implementation scientists to guide the planning and evaluation of programs, has been used for health care- and

community-based programs promoting chronic disease prevention and management, healthy aging, mental health, and health behavior change [5]. In addition, implementation science methods have been applied in clinical settings (eg, clinic-initiated cancer screening, tobacco cessation, and mental health programs) to scale up effective interventions to improve population health [4].

Implementation strategies are the methods, actions, and activities that aim to enhance the adoption, implementation, and sustainability of EBIs in clinical and public health practice. Implementation strategies can target multiple levels (eg, communities, hospitals, health care clinics, public health

departments, clinical and public health practitioners, and individual patients and community members) and may involve multiple components (eg, information technology tools, workflow changes, and policies mandating services) and activities (eg, training and incentives) [6,7]. Numerous factors, such as target populations and targeted behavior change, varied uptake of strategies across settings, the actors that deliver the implementation strategies, and the timing of the EBI implementation, can influence the implementation processes and outcomes [6-8]. Further, there is often a need to tailor or adapt implementation strategies and the associated activities to the local, dynamic context to increase implementation success. Given the multifactorial drivers and their complex relationships, implementation science could benefit from advanced data analytics frameworks and methods for artificial intelligence and machine learning (ML).

As a subfield of artificial intelligence, ML [9,10] develops automated methods and algorithms that learn from data. With this learning, it can then perform tasks such as prediction and pattern discovery. To date, ML applications in health care settings have been focused on supplementing clinical work, predicting health-related outcomes (eg, disease severity and prognosis) [11-14], and supporting clinical decisions (eg, tailoring medications and other treatments) [15-17]. Applications of ML in public health include population health surveillance and outbreak mitigation, evaluating the effectiveness of public health strategies and campaign and disaster and emergency

alleviation [18-22]. Existing literature on the application of ML in the field of implementation science is sparse [23]. However, ML has great potential to be applied in areas such as tailoring strategies and support activities, supporting decision-making on the selection of actors or settings, and predicting and understanding the impact of implementation strategies on the adoption of EBIs across different settings and target populations. The aim of this viewpoint is to introduce a roadmap for applying ML techniques to address implementation science questions, describe a few limited real-world applications of ML related to implementation science, and discuss challenges that implementation scientists and methodologists may face along the way when using ML as a strategy to monitor EBI adoption or to inform the need for interventions.

A Roadmap for Applying ML in Implementation Science

ML approaches can be applied across the continuum of EBI implementation. Here, we use the strategic implementation framework (SIF) [24] as a roadmap to illustrate the potential application of ML at different stages of implementation, as summarized in Table 1. The SIF depicts 3 stages of implementation (ie, setting the stage; active implementation; and monitor, support, and sustain) and the distinct types of strategies needed for practice change in each stage to ensure that improvements are supported and sustained.

Table 1. Roadmap for implementation scientists and methodologists to use ML^a.

	Strategic implementation framework stages		
	Setting the stage	Active implementation	Monitor, support, and sustain
Implementation goals and activities	<ul style="list-style-type: none"> Understand the local context to select implementation strategies and prepare for implementation 	<ul style="list-style-type: none"> Implement strategies and support activities to improve EBI^b 	<ul style="list-style-type: none"> Monitor the sustained adoption of strategy and EBI improvements
Implementation challenges	<ul style="list-style-type: none"> Limited data available or used The context is not static End users with differing priorities 	<ul style="list-style-type: none"> The context is not static End users with differing priorities Need to adapt to setting and targeted population 	<ul style="list-style-type: none"> The context is not static (new guidelines, policy, and care delivery)
Opportunities for ML application	<ul style="list-style-type: none"> Predict who will adopt an EBI Determine the level of support needed Identify the need for change 	<ul style="list-style-type: none"> ML as a strategy 	<ul style="list-style-type: none"> Monitor progress Inform need for deimplementation Inform need for support
Considerations when using ML across the 3 stages	<ul style="list-style-type: none"> Setting characteristics Time period Data completeness Multilevel strategy 	<ul style="list-style-type: none"> N/A^c 	<ul style="list-style-type: none"> Risk prediction bias Recalibration of ML Adaptation of ML Deimplementation of ML

^aML: machine learning.

^bEBI: evidence-based intervention.

^cN/A: not applicable.

Setting the Stage

Setting the stage refers to preimplementation activities such as assessing readiness to change, identifying barriers and facilitators to implementing EBIs, selecting or developing

strategies to support implementation, and identifying and acquiring resources. Implementation scientists often find that an effective strategy in one setting may not work well in other settings and that some may need more or different types of support (eg, hours of training, intensity of coaching support, or

remote vs in-person training). As such, one of the biggest implementation science challenges is to identify what works, for whom, under what circumstances, and with what level of support.

Typical approaches for selecting and tailoring implementation strategies to fit the local context (eg, process mapping, intervention mapping, and coincidence analysis) address this challenge at the organization or population levels [25]. Often, data to inform the selection of an implementation strategy are limited to surveys, qualitative interviews, and organization-level data. However, clinical or public health data (eg, electronic health records [EHRs], administrative data, claims data, patient or disease registries, immunization registries, and health surveys), data linkages (eg, EHR data linked across practice sites, water quality, and air quality), and data related to implementation processes (eg, responses of patients, community members, and practitioners to a specific implementation science strategy from prior studies) are increasingly available. Implementation scientists could use ML to analyze large-scale, individual-level data to identify or predict who (individuals or subpopulations) is most likely (or least likely) to engage or respond to the intervention [26,27]. Specifically, the application of ML in the preimplementation stage could assist with the selection of the settings or actors, refinement of implementation strategies, and decisions about support activities. ML techniques could predict which sites, practitioners, or target populations will most likely respond well to certain implementation strategies (such as a training session or a health information technology tool), are most likely to need extra support, or might respond better to different strategies. These analyses could be based on prior engagement with strategies that led to increased adoption of EBIs or known characteristics of community (eg, census and environmental health), health systems (eg, geographic location), providers (eg, years of practice), patients (eg, race and ethnicity), and other targeted users.

There are currently no studies using ML approaches to tailor implementation strategies or support needs in the preimplementation stage. A few studies have used unsupervised statistical learning methods, such as latent class analysis and latent profile analysis [28], to identify subgroups of health care providers [27] and patients [26] responding differently to implementation strategies that promote provider-patient communication on critical illness or patients' physical activities for weight reduction. For example, one study identified 3 groups (or phenotypes) of oncologists based on demographics, practice patterns, and patient panel information [27]. These phenotypes showed different responses to an EHR-based intervention (EHR nudges) aimed at improving advance care planning (ACP) discussion. Oncologists with the lowest volume of patients and a higher rate of baseline ACP discussion showed the greatest improvement compared to those with higher volume or lowest baseline ACP and intermediate volume or baseline ACP. One study used a supervised learning model to identify areas where the implementation of HIV prevention programs should be prioritized. Using state surveillance data on substance use, sexually transmitted diseases, and community characteristics (eg, percent living in poverty), ML modeling identified high-priority areas, of which 79% did not have implemented

syringe services programs [29]. Similar modeling approaches could be used to better identify who will adopt what implementation strategies with what supports and tailor resource allocation before an implementation program is launched to improve the adoption and sustainability of EBIs.

Further, ML applications during the setting the stage could also facilitate monitoring when interventions are needed. For instance, using continuously collected clinical or public health data and ML-based phenotyping methods [27], it is possible to prioritize target populations who need the EBIs most at different time points or stages of the implementation of an intervention. Modeling could also trigger notifications to local clinics and public health departments about changes in quality metrics that require improvement, the resources needed to make an improvement (eg, additional staff), or changes in an environmental context (eg, climate change) [30] that could impact disease incidences and health care needs.

Active Implementation

During the active implementation stage, strategies and support activities are implemented to promote the adoption of an EBI (eg, disease surveillance, prescribing shingles vaccination, and lung cancer screening). During this stage, ML techniques could be incorporated as an implementation strategy. ML-based algorithms relating to the active implementation stage are currently being used to support making accurate diagnoses, disease risk estimation and surveillance, public health campaigns, and clinical decision-making. One example is the use of an ML model to identify foodborne illness in real time (FINDER). This model was developed, implemented, and tested in 2 US cities. FINDER would provide a daily list of restaurants identified as unsafe (likely to have health code violation). Health departments would then conduct an inspection in the restaurants identified by FINDER. The model identified accurately more unsafe restaurants than the previous system or reported complaints [31]. Examples in palliative care include a deep learning model that incorporates patients' EHR data to predict mortality (those patients most likely to die within 3-12 months). The model-generated estimates were used to inform providers' care recommendations and decisions about referring patients to palliative care [32,33]. In the context of cancer screening, ML models based on reinforcement learning or ensemble learning are being developed to more accurately identify patients with high risk of cancer [34,35]. These models could be used for cancer screening to balance the benefits of early detection and the costs of overscreening.

Further, in clinical care, clinical decision support (CDS) tools [36,37], including EHR alerts, are common implementation strategies used to promote guideline-concordant practice. ML can be used to develop "smarter" CDS tools to reduce alert fatigue. For example, an ML model was developed to predict whether a provider would respond to shingles vaccination alerts based on the provider's characteristics (eg, demographics and clinical roles), patient's demographics, and history of the provider's interaction with the alerts [38]. The ML model was shown to reduce over 45% of shingles vaccination alerts without reducing weekly shingles vaccination orders [38].

Monitor, Support, and Sustain

This stage focuses on activities that ensure the sustainability of an intervention. During the monitor, support, and sustain stage, ML can inform changes needed to ensure the adoption and sustainability of practice changes. ML-based methods can leverage vast amounts of data to inform more flexible and adaptive implementation strategies. ML can also facilitate the evaluation and adaptation of strategies and inform where deimplementation is needed. For instance, ML could be used to identify when public health campaigns have reached saturation, need to be refocused, or are missing the target population. For example, during the COVID-19 pandemic, studies use ML models to identify people at greatest risk for COVID-19 death and who should be prioritized for vaccination. Different studies using different populations showed variations in who should be prioritized in informing local public health efforts [39-42]. For example, in clinical practice, implementation scientists leveraged both EHR audit logs and innovative ML-based approaches to monitor the impact of implementing a tobacco control CDS tool in the EHR system [43-45]. According to the Health Information Portability and Accountability Act (HIPAA) [46] and the 2014 release of the Meaningful Use regulations [47], all the EHRs in the United States are required to implement audit logs to unobtrusively track users' EHR use. In a recent study, a latent-variable statistical ML model was developed to infer EHR-use activities from EHR audit log data [44]. Specifically, the ML model identified topics from EHR log data, where each topic was represented by a probability distribution of microlevel EHR actions such as loading a flow sheet, viewing a problem list, and using a favorite phrase predefined in EHR. Domain experts (3 physicians and 1 EHR specialist) reviewed these topics (eg, the top-ranked microlevel EHR actions belonging to each topic and example EHR sessions representative for each topic) and assigned an EHR-use activity (eg, visit documentation with record review and address CDS alerts) to each topic. This domain expert-informed model was then applied to EHR logs for 3703 encounters (before CDS implementation: n=2633 and after CDS implementation: n=1070) in 4 cancer clinics to monitor changes in providers' EHR-use between 2019 and 2020 [45]. This study found that clinicians spent more time addressing CDS (more than 32-35 seconds) during a patient visit after CDS implementation (vs before CDS implementation), with compensatory unintended reductions in time spent reviewing patient vital data (less than 61 seconds) and modifying EHR (less than 7-24 seconds) [45]. These findings pointed to potential adaptations of the CDS to improve efficacy and reduce burden [43]. These data-driven findings can inform qualitative studies that aim to understand the causes of the unintended consequences and further inform the decision on refining or deimplementing certain features of the CDS tool.

In summary, despite very few real-world applications of ML in implementation science, there are many options and opportunities to use ML at different stages of implementation; however, some factors are important to take into consideration.

What Are the Factors to Consider in Using ML for Implementation Science?

As illustrated earlier, ML applications can potentially benefit implementation science across each of the SIF stages. However, many factors can impact the use or validity of these ML-based applications in real-world settings, including achieving equitable outcomes across multiple settings or subpopulations [48].

There are various techniques used in ML [49]. Supervised learning methods can be used to build predictive models (eg, prediction of patients' risks in illness or poor prognosis and responses of community members, patients, or providers to EBIs and implementation science strategies). Unsupervised learning methods can be used to mine data to identify patterns (eg, identify subgroups of population, patients, and health systems who have different responses to EBIs and implementation strategies). A common practice to develop and validate supervised ML models includes two stages: (1) using a data set to develop and validate (ie, internal validation) the model and (2) using a separate data set (obtained from other similar settings or from a withheld sample) to validate (ie, external validation) the developed model [50,51]. In the first stage, the model can be trained or validated through cross-validation or using a random split of the data set (eg, training or development or validation sets). The model's parameters and hyperparameters are tuned or set using the training and development sets. In the second stage, the model's performance is further assessed on the external validation set. Different from supervised learning, there is no ground truth (eg, labels for clusters or subgroups identified by unsupervised learning) to validate results from unsupervised learning in a real-world setting. Consequently, the evaluation process for unsupervised learning is less standard than supervised learning, and the choice of evaluation measures often depends on the unsupervised learning algorithms that are used [52,53]. In general, the quality of clustering results can be measured in 2 aspects when no external references (ie, ground truth) are available: coherence (ie, the similarity of objects falling into the same cluster) and separation (ie, the separation between clusters). Manual chart review is also useful or even necessary for qualitatively validating the clustering results in clinical settings [54]. Both supervised and unsupervised models developed on a specific sample or data set may not be readily applicable to other samples or data sets—the issue with generalizing ML models to different settings [55,56]. This issue has important implications on the use of ML in implementation science and requires paying special attention to model design, development, and validation.

The first factor to consider is that implementation strategies can be implemented at multiple levels (eg, state, county, community, population, health systems, clinicians, and patients), which would determine at which level the ML models would be based. Models developed and validated using data from one level (eg, clinic or community) need further validation and adaptation before being used for predicting outcomes at another level (eg, patient) or an intervention implemented at multiple levels [57]. For example, within the setting the stage phase, a model could

be developed using clinician and clinic characteristics (eg, specialty, provider type, and clinic geographic location) to predict which clinicians or clinics will be most likely to adopt a CDS tool. This model, however, is unlikely to be sufficient or valid in predicting the adoption of a multilevel intervention that targets both clinicians and patients (using provider nudges via EHR and patient nudges via SMS text messages). Similarly, public health programs (eg, a tobacco control or vaccination program) often use strategies targeting various levels within a public health jurisdiction (eg, individual, city, county, and state). An ML model predicting the adoption or success of such programs needs to take into account multilevel factors.

Second, the setting (eg, type of clinic and social culture of a specific community), its geographic location, and the time period used in validating the ML model are important factors to consider. These contextual factors are important in implementation science as they impact which strategy or combination of strategies are selected to scale up or modify to ensure the adoption and sustainability of EBI. Models that predict the adoption or sustainability of an implementation strategy developed in primary care clinics are unlikely to have an adequate prediction in specialty clinics in the setting stage phase. Similarly, an ML-based strategy to improve an EBI in a rural community setting will likely need adaptation to be valid in an urban community setting. Additionally, the time period in which the model was developed needs to be taken into account. For instance, ML-based CDS developed prior to the COVID-19 pandemic may be obsolete or invalid after the pandemic in view of the widespread adoption of telehealth.

Third, when using ML models as an implementation strategy for risk prediction, they should be designed to predict the actual targeted outcome rather than the outcome that is easiest to obtain. For example, consider a risk prediction model being used to direct palliative care interventions. It is easier to train an ML-based tool to predict mortality, as a surrogate for palliative care needs, because mortality is less susceptible to measurement error and is available in palliative care medical records [58,59]. However, training an algorithm on mortality may not identify the individuals with high symptomatic or psychosocial needs who would benefit from palliative care the most. Targeting the risk prediction to the outcome that is most likely to matter for the EBI being implemented is imperative.

Finally, it is critically important to develop and validate models with equity in mind. Many of the algorithms developed in medicine are based on trials with nonrepresentative samples [60]. A recent publication examining various race-biased algorithms used for medical risk predictions demonstrated the potentially harmful consequences of biased algorithms [61]. Within implementation science, as noted earlier, strategies may not work for all. ML models validated in a specific population (eg, pediatric patients) within a specific setting (eg, hospital) could be misused and inequitable if used in a different population (eg, Latino pediatric patients receiving care in a community health center). The learning here is that ML-based implementation strategies need to be tested, validated, and adapted to fit the context of the targeted population to ensure health equity.

What Are the Challenges?

Overview

Despite the large amount of clinical data and data from pragmatic implementation trials, there are many challenges associated with data access and data quality. Further, the tools and resources needed to extract and preprocess these data for developing ML may not be easily accessible. For example, extracting and harmonizing patient-level data from the EHRs from multiple health systems to develop a preimplementation ML model could be particularly difficult and time-consuming if these health systems have different EHR vendors. Furthermore, the application of ML in implementation science may result in unintended consequences, and issues related to the sustainability and scalability of the model need to be addressed.

Data: Quality, Availability, and Type

Public health data and information systems vary with regard to data quality, completeness, collection methods by systems, sampling bias, and underreporting [62-64]. In addition, the collection and generation of public health data are often time-consuming, resulting in delays in data reporting. Similarly, clinical-related data, such as EHR or health insurance claims data, are not designed for research and as such may not be collected and recorded in a systematic standard way. For example, comprehensiveness, completeness, and availability of patient demographic information (eg, race or ethnicity), health insurance data, and clinician data vary greatly by health systems and EHR vendors [65-71]. Additionally, some information that can be critical in the accuracy of ML prediction may reside in unstructured data (eg, a scanned PDF and free text of an encounter note) and, therefore, would require additional preprocessing steps, such as natural language processing [72]. Missing clinical-related data are unlikely to be random [70]. Specifically, EHR data come from a combination of clinician notes, test orders and results, documentation of diagnoses, and patient-reported information. The accuracy and completeness of these data are dependent on the source of the information. For example, the history of a cancer diagnosis can be derived from clinician diagnosis, clinical exchange systems, and patient self-reported history. A study linked EHR data with cancer registry to assess the accuracy of cancer diagnosis in the EHR [66]. Authors found that approximately 45% of cases recorded in the registry did not have a cancer history in their EHR. This information may have been in unstructured data such as in encounter notes. Data used for training an ML model may underrepresent certain patient subgroups [71]. For example, the use of insurance claims data excludes patients without health insurance, and these patients are often socioeconomically disadvantaged individuals. Variation in data documentation and completeness impact not only predictor variables used in the ML models but also the outcome variables. For example, predictive models of emergency department admissions using claims data would miss patients who are uninsured and are more likely to rely on the emergency department for care [73]. Moreover, ML models designed to develop an intervention targeting health system, school system, or community-based

organization change may require data on staffing, supplies, or organizational capacity, which could be challenging to obtain.

Potential for Unintended Consequences

ML models, whether designed for predicting disease risk or for supporting clinical care management and decision-making, are susceptible to bias. Bias can be introduced at multiple points in the development and application process of ML [61,74,75]. As noted earlier, data sources and data representativeness (eg, the population, inclusion or exclusion of diseases, comorbidities, and health risk factors) can greatly influence the ML model and consequently the actions based on the ML model. Further, because ML models can generate data for other ML models, bias can be amplified and can lead to unintended consequences [76]. Char et al [77] proposed a framework for examining ethical considerations of ML models in health care settings, which poses questions about the values and ethics at multiple steps of the model development and implementation. This framework can guide decision-making to minimize bias and can promote accountability and transparency in model development.

Sustainability and Scalability of the Model

Public health interventions and campaigns are moving targets. For instance, climate change is leading public health departments to adapt or develop new initiatives for disaster preparedness efforts, disease surveillance, and carbon footprint reduction [78-80]. For instance, there is growing evidence of the mental health toll of climate-related events [81], yet strategies to monitor and intervene climate-related mental health burden are scarce [78]. Analogously, health care systems are ever-changing [82] as they must adapt to new clinical care guidelines, changes in reimbursement policies, care delivery modality (ie, telemedicine), quality improvement efforts, and local, state, or federal law amendments. For example, in April 2020, the American Society of Colposcopy and Cervical Pathology released new guidelines to provide recommendations on cervical cancer screening frequency and follow-up tests for abnormal cervical cancer results [83]. These guidelines significantly differ from the previous 2012 version [84]. Any implementation strategies designed to facilitate the adoption of the 2012 guidelines became obsolete and needed to be revised. For

another example, EHR-based patient portals are efficient systems for communication between patients and health care providers and platforms for health information exchange. These portals can be a platform for patient-centered implementation strategies to improve the uptake of evidence-based practice. Patient portal tools have been used to improve the uptake of ACP or lung cancer screening [85]. Patient portal adoption before the COVID-19 pandemic, however, remained relatively low and varied widely across patient subgroups (eg, by age and socioeconomic status), diminishing the effectiveness of strategies implemented within the portal [86,87]. The need for social distancing and the uptake of telemedicine during the COVID-19 pandemic led to a rise in patient portal use, which could improve the reach of such strategies [88]. The uptake in patient portal during the pandemic was also associated with a rise in “e-visits,” which were communications between patients and clinicians between in-person visits [89,90]. This led to health care systems to bill for these messages following existing federal rules [90,91], which in turn may limit the use of patient portals and impact their effectiveness as an implementation strategy. This example illustrates how the changes in the health care system can impact a specific implementation strategy. Consequently, the reach, adoption, and sustainability of the EBI it aimed to improve are also impacted. These ever-changing systems pose a significant complication when using ML models [92,93]. How frequently should an ML model be adapted or recalibrated to ensure that it has accurate predictions and is unbiased and ethical? This is a critical factor impacting the use of ML in implementation science and across the 3 stages of implementation and remains to be answered by future studies.

Conclusions

ML can assist with predicting what will work best, for whom, under what circumstances, and with what level of support, or what and when adaptation and deimplementation are needed. However, there are many remaining challenges with integrating ML into various stages of implementation, which require further research and investigation. Tackling these challenges has the potential to render ML as an innovative and useful tool in implementation science in years to come.

Acknowledgments

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health (awards P50CA244289, P50CA244690, and P50CA244693). This program is supported by funding provided through the Cancer Moonshot. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Generative artificial intelligence was not used for any portion of the paper writing.

Authors' Contributions

All authors substantially contributed to the conceptualization and edits of this viewpoint and approved the final version.

Conflicts of Interest

JB reports personal fees from Reimagine Care, Astrazeneca, and Healthcare Foundry and grants from Lilly Loxo and Gilead. The other authors have no conflicts of interest to declare.

References

1. Armstrong R, Sales A. Welcome to implementation science communications. *Implement Sci Commun*. 2020;1:1. [FREE Full text] [doi: [10.1186/s43058-020-00018-4](https://doi.org/10.1186/s43058-020-00018-4)] [Medline: [32885176](https://pubmed.ncbi.nlm.nih.gov/32885176/)]

2. Lobb R, Colditz GA. Implementation science and its application to population health. *Annu Rev Public Health*. 2013;34:235-251. [[FREE Full text](#)] [doi: [10.1146/annurev-publhealth-031912-114444](https://doi.org/10.1146/annurev-publhealth-031912-114444)] [Medline: [23297655](#)]
3. Estabrooks PA, Brownson RC, Pronk NP. Dissemination and implementation science for public health professionals: an overview and call to action. *Prev Chronic Dis*. 2018;15:E162. [[FREE Full text](#)] [doi: [10.5888/pcd15.180525](https://doi.org/10.5888/pcd15.180525)] [Medline: [30576272](#)]
4. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? *Psychiatry Res*. 2020;283:112376. [[FREE Full text](#)] [doi: [10.1016/j.psychres.2019.04.025](https://doi.org/10.1016/j.psychres.2019.04.025)] [Medline: [31036287](#)]
5. Kwan BM, McGinnes HL, Ory MG, Estabrooks PA, Waxmonsky JA, Glasgow RE. RE-AIM in the real world: use of the RE-AIM framework for program planning and evaluation in clinical and community settings. *Front Public Health*. 2019;7:345. [[FREE Full text](#)] [doi: [10.3389/fpubh.2019.00345](https://doi.org/10.3389/fpubh.2019.00345)] [Medline: [31824911](#)]
6. Leeman J, Birken SA, Powell BJ, Rohweder C, Shea CM. Beyond "implementation strategies": classifying the full range of strategies used in implementation science and practice. *Implement Sci*. 2017;12(1):125. [[FREE Full text](#)] [doi: [10.1186/s13012-017-0657-x](https://doi.org/10.1186/s13012-017-0657-x)] [Medline: [29100551](#)]
7. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci*. 2015;10:21. [[FREE Full text](#)] [doi: [10.1186/s13012-015-0209-1](https://doi.org/10.1186/s13012-015-0209-1)] [Medline: [25889199](#)]
8. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. *Implement Sci*. 2013;8:139. [[FREE Full text](#)] [doi: [10.1186/1748-5908-8-139](https://doi.org/10.1186/1748-5908-8-139)] [Medline: [24289295](#)]
9. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol*. 2019;28(2):73-81. [doi: [10.1080/13645706.2019.1575882](https://doi.org/10.1080/13645706.2019.1575882)] [Medline: [30810430](#)]
10. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920-1930. [[FREE Full text](#)] [doi: [10.1161/CIRCULATIONAHA.115.001593](https://doi.org/10.1161/CIRCULATIONAHA.115.001593)] [Medline: [26572668](#)]
11. Ting WC, Lu YCA, Ho WC, Cheewakriangkrai C, Chang HR, Lin CL. Machine learning in prediction of second primary cancer and recurrence in colorectal cancer. *Int J Med Sci*. 2020;17(3):280-291. [[FREE Full text](#)] [doi: [10.7150/ijms.37134](https://doi.org/10.7150/ijms.37134)] [Medline: [32132862](#)]
12. Alafif T, Tehame AM, Bajaba S, Barnawi A, Zia S. Machine and deep learning towards COVID-19 diagnosis and treatment: survey, challenges, and future directions. *Int J Environ Res Public Health*. 2021;18(3):1117. [[FREE Full text](#)] [doi: [10.3390/ijerph18031117](https://doi.org/10.3390/ijerph18031117)] [Medline: [33513984](#)]
13. Selya A, Anshutz D, Griese E, Weber TL, Hsu B, Ward C. Predicting unplanned medical visits among patients with diabetes: translation from machine learning to clinical implementation. *BMC Med Inform Decis Mak*. 2021;21(1):111. [[FREE Full text](#)] [doi: [10.1186/s12911-021-01474-1](https://doi.org/10.1186/s12911-021-01474-1)] [Medline: [33789660](#)]
14. Martinez O, Martinez C, Parra CA, Rugeles S, Suarez DR. Machine learning for surgical time prediction. *Comput Methods Programs Biomed*. 2021;208:106220. [doi: [10.1016/j.cmpb.2021.106220](https://doi.org/10.1016/j.cmpb.2021.106220)] [Medline: [34161848](#)]
15. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA*. 2018;320(21):2199-2200. [doi: [10.1001/jama.2018.17163](https://doi.org/10.1001/jama.2018.17163)] [Medline: [30398550](#)]
16. Howard J. Artificial intelligence: implications for the future of work. *Am J Ind Med*. 2019;62(11):917-926. [doi: [10.1002/ajim.23037](https://doi.org/10.1002/ajim.23037)] [Medline: [31436850](#)]
17. Jayatilake SMDAC, Ganegoda GU. Involvement of machine learning tools in healthcare decision making. *J Healthc Eng*. 2021;2021:6679512. [[FREE Full text](#)] [doi: [10.1155/2021/6679512](https://doi.org/10.1155/2021/6679512)] [Medline: [33575021](#)]
18. Payedimarri AB, Concina D, Portinale L, Canonico M, Seys D, Vanhaecht K, et al. Prediction models for public health containment measures on COVID-19 using artificial intelligence and machine learning: a systematic review. *Int J Environ Res Public Health*. 2021;18(9):4499. [[FREE Full text](#)] [doi: [10.3390/ijerph18094499](https://doi.org/10.3390/ijerph18094499)] [Medline: [33922693](#)]
19. Lu S, Christie GA, Nguyen TT, Freeman JD, Hsu EB. Applications of artificial intelligence and machine learning in disasters and public health emergencies. *Disaster Med Public Health Prep*. 2022;16(4):1674-1681. [doi: [10.1017/dmp.2021.125](https://doi.org/10.1017/dmp.2021.125)] [Medline: [34134815](#)]
20. English N, Anesetti-Rothermel A, Zhao C, Latterner A, Benson AF, Herman P, et al. Image processing for public health surveillance of tobacco point-of-sale advertising: machine learning-based methodology. *J Med Internet Res*. 2021;23(8):e24408. [[FREE Full text](#)] [doi: [10.2196/24408](https://doi.org/10.2196/24408)] [Medline: [34448700](#)]
21. Fisher S, Rosella LC. Priorities for successful use of artificial intelligence by public health organizations: a literature review. *BMC Public Health*. 2022;22(1):2146. [[FREE Full text](#)] [doi: [10.1186/s12889-022-14422-z](https://doi.org/10.1186/s12889-022-14422-z)] [Medline: [36419010](#)]
22. Rodrigues PM, Madeiro JP, Marques JAL. Enhancing health and public health through machine learning: decision support for smarter choices. *Bioengineering (Basel)*. 2023;10(7):792. [[FREE Full text](#)] [doi: [10.3390/bioengineering10070792](https://doi.org/10.3390/bioengineering10070792)] [Medline: [37508819](#)]
23. McFadden BR, Reynolds M, Inglis TJJ. Developing machine learning systems worthy of trust for infection science: a requirement for future implementation into clinical practice. *Front Digit Health*. 2023;5:1260602. [[FREE Full text](#)] [doi: [10.3389/fdgh.2023.1260602](https://doi.org/10.3389/fdgh.2023.1260602)] [Medline: [37829595](#)]
24. Mitchell SA, Chambers DA. Leveraging implementation science to improve cancer care delivery and patient outcomes. *J Oncol Pract*. 2017;13(8):523-529. [[FREE Full text](#)] [doi: [10.1200/JOP.2017.024729](https://doi.org/10.1200/JOP.2017.024729)] [Medline: [28692331](#)]

25. Powell BJ, Beidas RS, Lewis CC, Aarons GA, McMillen JC, Proctor EK, et al. Methods to improve the selection and tailoring of implementation strategies. *J Behav Health Serv Res.* 2017;44(2):177-194. [FREE Full text] [doi: [10.1007/s11414-015-9475-6](https://doi.org/10.1007/s11414-015-9475-6)] [Medline: [26289563](https://pubmed.ncbi.nlm.nih.gov/26289563/)]
26. Lienert J, Patel M. Patient phenotypes help explain variation in response to a social gamification weight loss intervention. *Am J Health Promot.* 2020;34(3):277-284. [doi: [10.1177/0890117119892776](https://doi.org/10.1177/0890117119892776)] [Medline: [31876175](https://pubmed.ncbi.nlm.nih.gov/31876175/)]
27. Li E, Manz C, Liu M, Chen J, Chivers C, Braun J, et al. Oncologist phenotypes and associations with response to a machine learning-based intervention to increase advance care planning: secondary analysis of a randomized clinical trial. *PLoS One.* 2022;17(5):e0267012. [FREE Full text] [doi: [10.1371/journal.pone.0267012](https://doi.org/10.1371/journal.pone.0267012)] [Medline: [35622812](https://pubmed.ncbi.nlm.nih.gov/35622812/)]
28. Oberski D. Mixture models: latent profile and latent class analysis. In: Robertson J, Kaptein M, editors. *Modern Statistical Methods for HCI.* Cham. Springer International Publishing; 2016;275-287.
29. Bartholomew TS, Tookes HE, Spencer EC, Feaster DJ. Application of machine learning algorithms for localized syringe services program policy implementation—Florida, 2017. *Ann Med.* 2022;54(1):2137-2150. [FREE Full text] [doi: [10.1080/07853890.2022.2105391](https://doi.org/10.1080/07853890.2022.2105391)] [Medline: [35900201](https://pubmed.ncbi.nlm.nih.gov/35900201/)]
30. DeVoe JE, Huguet N, Likumuwa-Ackman S, Bazemore A, Gold R, Werner L. Precision ecologic medicine: tailoring care to mitigate impacts of climate change. *J Prim Care Community Health.* 2023;14:1-6. [FREE Full text] [doi: [10.1177/21501319231170585](https://doi.org/10.1177/21501319231170585)] [Medline: [37086151](https://pubmed.ncbi.nlm.nih.gov/37086151/)]
31. Sadilek A, Caty S, DiPrete L, Mansour R, Schenk T, Bergtholdt M, et al. Machine-learned epidemiology: real-time detection of foodborne illness at scale. *NPJ Digit Med.* 2018;1:36. [FREE Full text] [doi: [10.1038/s41746-018-0045-1](https://doi.org/10.1038/s41746-018-0045-1)] [Medline: [31304318](https://pubmed.ncbi.nlm.nih.gov/31304318/)]
32. Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. *BMC Med Inform Decis Mak.* 2018;18(Suppl 4):122. [FREE Full text] [doi: [10.1186/s12911-018-0677-8](https://doi.org/10.1186/s12911-018-0677-8)] [Medline: [30537977](https://pubmed.ncbi.nlm.nih.gov/30537977/)]
33. Manz CR, Parikh RB, Small DS, Evans CN, Chivers C, Regli SH, et al. Effect of integrating machine learning mortality estimates with behavioral nudges to clinicians on serious illness conversations among patients with cancer: a stepped-wedge cluster randomized clinical trial. *JAMA Oncol.* 2020;6(12):e204759. [FREE Full text] [doi: [10.1001/jamaoncol.2020.4759](https://doi.org/10.1001/jamaoncol.2020.4759)] [Medline: [33057696](https://pubmed.ncbi.nlm.nih.gov/33057696/)]
34. Yala A, Mikhael PG, Lehman C, Lin G, Strand F, Wan YL, et al. Optimizing risk-based breast cancer screening policies with reinforcement learning. *Nat Med.* 2022;28(1):136-143. [doi: [10.1038/s41591-021-01599-w](https://doi.org/10.1038/s41591-021-01599-w)] [Medline: [35027757](https://pubmed.ncbi.nlm.nih.gov/35027757/)]
35. Sun L, Yang L, Liu X, Tang L, Zeng Q, Gao Y, et al. Optimization of cervical cancer screening: a stacking-integrated machine learning algorithm based on demographic, behavioral, and clinical factors. *Front Oncol.* 2022;12:821453. [FREE Full text] [doi: [10.3389/fonc.2022.821453](https://doi.org/10.3389/fonc.2022.821453)] [Medline: [35242711](https://pubmed.ncbi.nlm.nih.gov/35242711/)]
36. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ.* 2005;330(7494):765. [FREE Full text] [doi: [10.1136/bmj.38398.500764.8F](https://doi.org/10.1136/bmj.38398.500764.8F)] [Medline: [15767266](https://pubmed.ncbi.nlm.nih.gov/15767266/)]
37. McCoy AB, Thomas EJ, Krousel-Wood M, Sittig DF. Clinical decision support alert appropriateness: a review and proposal for improvement. *Ochsner J.* 2014;14(2):195-202. [FREE Full text] [Medline: [24940129](https://pubmed.ncbi.nlm.nih.gov/24940129/)]
38. Chen J, Chokshi S, Hegde R, Gonzalez J, Iturrate E, Aphinyanaphongs Y, et al. Development, implementation, and evaluation of a personalized machine learning algorithm for clinical decision support: case study with shingles vaccination. *J Med Internet Res.* 2020;22(4):e16848. [FREE Full text] [doi: [10.2196/16848](https://doi.org/10.2196/16848)] [Medline: [32347813](https://pubmed.ncbi.nlm.nih.gov/32347813/)]
39. Tiwari A, Dadhania AV, Rangunathrao VAB, Oliveira ERA. Using machine learning to develop a novel COVID-19 Vulnerability Index (C19VI). *Sci Total Environ.* 2021;773:145650. [FREE Full text] [doi: [10.1016/j.scitotenv.2021.145650](https://doi.org/10.1016/j.scitotenv.2021.145650)] [Medline: [33940747](https://pubmed.ncbi.nlm.nih.gov/33940747/)]
40. Jamshidi E, Asgary A, Tavakoli N, Zali A, Dastan F, Daaee A, et al. Symptom prediction and mortality risk calculation for COVID-19 using machine learning. *Front Artif Intell.* 2021;4:673527. [FREE Full text] [doi: [10.3389/frai.2021.673527](https://doi.org/10.3389/frai.2021.673527)] [Medline: [34250465](https://pubmed.ncbi.nlm.nih.gov/34250465/)]
41. Cheong Q, Au-Yeung M, Quon S, Concepcion K, Kong JD. Predictive modeling of vaccination uptake in US counties: a machine learning-based approach. *J Med Internet Res.* 2021;23(11):e33231. [FREE Full text] [doi: [10.2196/33231](https://doi.org/10.2196/33231)] [Medline: [34751650](https://pubmed.ncbi.nlm.nih.gov/34751650/)]
42. Couto RC, Pedrosa TMG, Seara LM, Couto CS, Couto VS, Giacomini K, et al. COVID-19 vaccination priorities defined on machine learning. *Rev Saude Publica.* 2022;56:11. [FREE Full text] [doi: [10.11606/s1518-8787.2022056004045](https://doi.org/10.11606/s1518-8787.2022056004045)] [Medline: [35319671](https://pubmed.ncbi.nlm.nih.gov/35319671/)]
43. Chen J, Cutrona SL, Dharod A, Bunch SC, Foley KL, Ostasiewski B, et al. Monitoring the implementation of tobacco cessation support tools: using novel electronic health record activity metrics. *JMIR Med Inform.* 2023;11:e43097. [FREE Full text] [doi: [10.2196/43097](https://doi.org/10.2196/43097)] [Medline: [36862466](https://pubmed.ncbi.nlm.nih.gov/36862466/)]
44. Chen J, Cutrona SL, Dharod A, Bridges A, Moses A, Ostasiewski B, et al. Characterizing clinical activity patterns by topics inferred from electronic health record audit logs. 2022. Presented at: AMIA 2022 Annual Symposium; 2022; Washington, DC. URL: https://s4.goeshow.com/amia/annual/2022/schedule_at_a_glance.cfm
45. Chen J, Cutrona SL, Dharod A, Moses A, Bridges A, Ostasiewski B, et al. Monitoring for unintended consequences of EHR-based implementation strategies: a novel approach using EHR audit logs and machine learning. 2022. Presented at: 15th Annual Conference on the Science of Dissemination and Implementation in Health; 2022; Washington, DC.

46. United States. Health Insurance Portability and Accountability Act of 1996. Public Law 104-191. US Statut Large. 1996;110:1936-2103. [Medline: [16477734](#)]
47. 45 CFR § 170.210—Standards for health information technology to protect electronic health information created, maintained, and exchanged. Cornell Law School. URL: <https://www.law.cornell.edu/cfr/text/45/170.210> [accessed 2023-11-06]
48. Bates DW, Auerbach A, Schulam P, Wright A, Saria S. Reporting and implementing interventions involving machine learning and artificial intelligence. *Ann Intern Med*. 2020;172(Suppl 11):S137-S144. [FREE Full text] [doi: [10.7326/M19-0872](#)] [Medline: [32479180](#)]
49. Alpaydin E. Introduction to Machine Learning. 4th Edition. Cambridge, MA. MIT Press; 2020.
50. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. *Behav Ther*. 2020;51(5):675-687. [FREE Full text] [doi: [10.1016/j.beth.2020.05.002](#)] [Medline: [32800297](#)]
51. Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? *Clin Kidney J*. 2021;14(1):49-58. [FREE Full text] [doi: [10.1093/ckj/sfaa188](#)] [Medline: [33564405](#)]
52. Gan G, Ma C, Wu J. Data Clustering: Theory, Algorithms, and Applications. 2nd Edition. Philadelphia, PA. Society for Industrial and Applied Mathematics; 2020.
53. Tan PN, Steinback M, Kumar V. Introduction to Data Mining. Chennai, Tamil Nadu. Pearson India; 2016.
54. Gao CX, Dwyer D, Zhu Y, Smith CL, Du L, Folia KM, et al. An overview of clustering methods with guidelines for application in mental health research. *Psychiatry Res*. 2023;327:115265. [FREE Full text] [doi: [10.1016/j.psychres.2023.115265](#)] [Medline: [37348404](#)]
55. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. 2019;17(1):195. [FREE Full text] [doi: [10.1186/s12916-019-1426-2](#)] [Medline: [31665002](#)]
56. Maleki F, Ovens K, Gupta R, Reinhold C, Spatz A, Forghani R. Generalizability of machine learning models: quantitative evaluation of three methodological pitfalls. *Radiol Artif Intell*. 2023;5(1):e220028. [FREE Full text] [doi: [10.1148/ryai.220028](#)] [Medline: [36721408](#)]
57. Oikonomidi T, Norman G, McGarrigle L, Stokes J, van der Veer SN, Dowding D. Predictive model-based interventions to reduce outpatient no-shows: a rapid systematic review. *J Am Med Inform Assoc*. 2023;30(3):559-569. [FREE Full text] [doi: [10.1093/jamia/ocac242](#)] [Medline: [36508503](#)]
58. Courtright KR, Chivers C, Becker M, Regli SH, Pepper LC, Draugelis ME, et al. Electronic health record mortality prediction model for targeted palliative care among hospitalized medical patients: a pilot quasi-experimental study. *J Gen Intern Med*. 2019;34(9):1841-1847. [FREE Full text] [doi: [10.1007/s11606-019-05169-2](#)] [Medline: [31313110](#)]
59. Vu E, Steinmann N, Schröder C, Förster R, Aebbersold DM, Eychmüller S, et al. Applications of machine learning in palliative care: a systematic review. *Cancers (Basel)*. 2023;15(5):1596. [FREE Full text] [doi: [10.3390/cancers15051596](#)] [Medline: [36900387](#)]
60. Stipelman CH, Kukhareva PV, Trepman E, Nguyen QT, Valdez L, Kenost C, et al. Electronic health record-integrated clinical decision support for clinicians serving populations facing health care disparities: literature review. *Yearb Med Inform*. 2022;31(1):184-198. [FREE Full text] [doi: [10.1055/s-0042-1742518](#)] [Medline: [36463877](#)]
61. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. [FREE Full text] [doi: [10.1126/science.aax2342](#)] [Medline: [31649194](#)]
62. Geneviève LD, Martani A, Wangmo T, Elger BS. Precision public health and structural racism in the United States: promoting health equity in the COVID-19 pandemic response. *JMIR Public Health Surveill*. 2022;8(3):e33277. [FREE Full text] [doi: [10.2196/33277](#)] [Medline: [35089868](#)]
63. Martin LT, Nelson C, Yeung D, Acosta JD, Qureshi N, Blagg T, et al. The issues of interoperability and data connectedness for public health. *Big Data*. 2022;10(S1):S19-S24. [FREE Full text] [doi: [10.1089/big.2022.0207](#)] [Medline: [36070509](#)]
64. Acharya JC, Staes C, Allen KS, Hartsell J, Cullen TA, Lenert L, et al. Strengths, weaknesses, opportunities, and threats for the nation's public health information systems infrastructure: synthesis of discussions from the 2022 ACMI Symposium. *J Am Med Inform Assoc*. 2023;30(6):1011-1021. [FREE Full text] [doi: [10.1093/jamia/ocad059](#)] [Medline: [37146228](#)]
65. Marino M, Angier H, Valenzuela S, Hoopes M, Killerby M, Blackburn B, et al. Medicaid coverage accuracy in electronic health records. *Prev Med Rep*. 2018;11:297-304. [FREE Full text] [doi: [10.1016/j.pmedr.2018.07.009](#)] [Medline: [30116701](#)]
66. Hoopes M, Voss R, Angier H, Marino M, Schmidt T, DeVoe JE, et al. Assessing cancer history accuracy in primary care electronic health records through cancer registry linkage. *J Natl Cancer Inst*. 2021;113(7):924-932. [FREE Full text] [doi: [10.1093/jnci/djaa210](#)] [Medline: [33377908](#)]
67. Haneuse S, Arterburn D, Daniels MJ. Assessing missing data assumptions in EHR-based studies: a complex and underappreciated task. *JAMA Netw Open*. 2021;4(2):e210184. [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.0184](#)] [Medline: [33635321](#)]
68. Gutman CK, Lion KC, Waidner L, Bryan L, Sizemore A, Holland C, et al. Gaps in the identification of child race and ethnicity in a pediatric emergency department. *West J Emerg Med*. 2023;24(3):547-551. [FREE Full text] [doi: [10.5811/westjem.59145](#)] [Medline: [37278792](#)]
69. Anand P, Zhang Y, Merola D, Jin Y, Wang SV, Lii J, et al. Comparison of EHR data-completeness in patients with different types of medical insurance coverage in the United States. *Clin Pharmacol Ther*. 2023;114(5):1116-1125. [doi: [10.1002/cpt.3027](#)] [Medline: [37597260](#)]

70. Beaulieu-Jones BK, Lavage DR, Snyder JW, Moore JH, Pendergrass SA, Bauer CR. Characterizing and managing missing structured data in electronic health records: data analysis. *JMIR Med Inform.* 2018;6(1):e11. [[FREE Full text](#)] [doi: [10.2196/medinform.8960](https://doi.org/10.2196/medinform.8960)] [Medline: [29475824](https://pubmed.ncbi.nlm.nih.gov/29475824/)]
71. Cook L, Espinoza J, Weiskopf NG, Mathews N, Dorr DA, Gonzales KL, et al. Issues with variability in electronic health record data about race and ethnicity: descriptive analysis of the National COVID Cohort Collaborative Data Enclave. *JMIR Med Inform.* 2022;10(9):e39235. [[FREE Full text](#)] [doi: [10.2196/39235](https://doi.org/10.2196/39235)] [Medline: [35917481](https://pubmed.ncbi.nlm.nih.gov/35917481/)]
72. Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J Biomed Inform.* 2017;73:14-29. [[FREE Full text](#)] [doi: [10.1016/j.jbi.2017.07.012](https://doi.org/10.1016/j.jbi.2017.07.012)] [Medline: [28729030](https://pubmed.ncbi.nlm.nih.gov/28729030/)]
73. Zhou RA, Baicker K, Taubman S, Finkelstein AN. The uninsured do not use the emergency department more—they use other care less. *Health Aff (Millwood).* 2017;36(12):2115-2122. [[FREE Full text](#)] [doi: [10.1377/hlthaff.2017.0218](https://doi.org/10.1377/hlthaff.2017.0218)] [Medline: [29200330](https://pubmed.ncbi.nlm.nih.gov/29200330/)]
74. Maguolo G, Nanni L. A critic evaluation of methods for COVID-19 automatic detection from X-ray images. *Inf Fusion.* 2021;76:1-7. [[FREE Full text](#)] [doi: [10.1016/j.inffus.2021.04.008](https://doi.org/10.1016/j.inffus.2021.04.008)] [Medline: [33967656](https://pubmed.ncbi.nlm.nih.gov/33967656/)]
75. Bektaş M, Tuynman JB, Pereira JC, Burchell GL, van der Peet DL. Machine learning algorithms for predicting surgical outcomes after colorectal surgery: a systematic review. *World J Surg.* 2022;46(12):3100-3110. [[FREE Full text](#)] [doi: [10.1007/s00268-022-06728-1](https://doi.org/10.1007/s00268-022-06728-1)] [Medline: [36109367](https://pubmed.ncbi.nlm.nih.gov/36109367/)]
76. 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):2247-2249. [[FREE Full text](#)] [doi: [10.1038/s41591-022-01993-y](https://doi.org/10.1038/s41591-022-01993-y)] [Medline: [36163298](https://pubmed.ncbi.nlm.nih.gov/36163298/)]
77. Char DS, Abràmoff MD, Feudtner C. Identifying ethical considerations for machine learning healthcare applications. *Am J Bioeth.* 2020;20(11):7-17. [[FREE Full text](#)] [doi: [10.1080/15265161.2020.1819469](https://doi.org/10.1080/15265161.2020.1819469)] [Medline: [33103967](https://pubmed.ncbi.nlm.nih.gov/33103967/)]
78. Hajat S, Gasparrini A. The excess winter deaths measure: why its use is misleading for public health understanding of cold-related health impacts. *Epidemiology.* 2016;27(4):486-491. [[FREE Full text](#)] [doi: [10.1097/EDE.0000000000000479](https://doi.org/10.1097/EDE.0000000000000479)] [Medline: [26986872](https://pubmed.ncbi.nlm.nih.gov/26986872/)]
79. Watts N, Amann M, Ayeb-Karlsson S, Belesova K, Bouley T, Boykoff M, et al. The Lancet countdown on health and climate change: from 25 years of inaction to a global transformation for public health. *Lancet.* 2018;391(10120):581-630. [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(17\)32464-9](https://doi.org/10.1016/S0140-6736(17)32464-9)] [Medline: [29096948](https://pubmed.ncbi.nlm.nih.gov/29096948/)]
80. Zeuli K, Nijhuis A, Macfarlane R, Ridsdale T. The impact of climate change on the food system in Toronto. *Int J Environ Res Public Health.* 2018;15(11):2344. [[FREE Full text](#)] [doi: [10.3390/ijerph15112344](https://doi.org/10.3390/ijerph15112344)] [Medline: [30355969](https://pubmed.ncbi.nlm.nih.gov/30355969/)]
81. Clayton S. Climate change and mental health. *Curr Environ Health Rep.* 2021;8(1):1-6. [doi: [10.1007/s40572-020-00303-3](https://doi.org/10.1007/s40572-020-00303-3)] [Medline: [33389625](https://pubmed.ncbi.nlm.nih.gov/33389625/)]
82. Nilsen P, Seing I, Ericsson C, Birken SA, Schildmeijer K. Characteristics of successful changes in health care organizations: an interview study with physicians, registered nurses and assistant nurses. *BMC Health Serv Res.* 2020;20(1):147. [[FREE Full text](#)] [doi: [10.1186/s12913-020-4999-8](https://doi.org/10.1186/s12913-020-4999-8)] [Medline: [32106847](https://pubmed.ncbi.nlm.nih.gov/32106847/)]
83. Guidelines overview. American Society of Colposcopy and Cervical Pathology. 2023. URL: <https://www.asccp.org/guidelines> [accessed 2023-11-06]
84. Massad LS, Einstein MH, Huh WK, Katki HA, Kinney WK, Schiffman M, et al. 2012 updated consensus guidelines for the management of abnormal cervical cancer screening tests and cancer precursors. *J Low Genit Tract Dis.* 2013;17(5 Suppl 1):S1-S27. [doi: [10.1097/LGT.0b013e318287d329](https://doi.org/10.1097/LGT.0b013e318287d329)] [Medline: [23519301](https://pubmed.ncbi.nlm.nih.gov/23519301/)]
85. Lum HD, Brungardt A, Jordan SR, Phimphasone-Brady P, Schilling LM, Lin CT, et al. Design and implementation of patient portal-based advance care planning tools. *J Pain Symptom Manage.* 2019;57(1):112-117.e2. [[FREE Full text](#)] [doi: [10.1016/j.jpainsymman.2018.10.500](https://doi.org/10.1016/j.jpainsymman.2018.10.500)] [Medline: [30595147](https://pubmed.ncbi.nlm.nih.gov/30595147/)]
86. Ramsey A, Lanzo E, Huston-Paterson H, Tomaszewski K, Trent M. Increasing patient portal usage: preliminary outcomes from the MyChart Genius Project. *J Adolesc Health.* 2018;62(1):29-35. [[FREE Full text](#)] [doi: [10.1016/j.jadohealth.2017.08.029](https://doi.org/10.1016/j.jadohealth.2017.08.029)] [Medline: [29169768](https://pubmed.ncbi.nlm.nih.gov/29169768/)]
87. Wallace LS, Angier H, Huguet N, Gaudino JA, Krist A, Dearing M, et al. Patterns of electronic portal use among vulnerable patients in a nationwide practice-based research network: from the OCHIN Practice-Based Research Network (PBRN). *J Am Board Fam Med.* 2016;29(5):592-603. [[FREE Full text](#)] [doi: [10.3122/jabfm.2016.05.160046](https://doi.org/10.3122/jabfm.2016.05.160046)] [Medline: [27613792](https://pubmed.ncbi.nlm.nih.gov/27613792/)]
88. Pullyblank K, Krupa N, Scribani M, Chapman A, Kern M, Brunner W. Trends in telehealth use among a cohort of rural patients during the COVID-19 pandemic. *Digit Health.* 2023;9:20552076231203803. [[FREE Full text](#)] [doi: [10.1177/20552076231203803](https://doi.org/10.1177/20552076231203803)] [Medline: [37799503](https://pubmed.ncbi.nlm.nih.gov/37799503/)]
89. Nouri S, Lyles CR, Sherwin EB, Kuznia M, Rubinsky AD, Kemper KE, et al. Visit and between-visit interaction frequency before and after COVID-19 telehealth implementation. *JAMA Netw Open.* 2023;6(9):e2333944. [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2023.33944](https://doi.org/10.1001/jamanetworkopen.2023.33944)] [Medline: [37713198](https://pubmed.ncbi.nlm.nih.gov/37713198/)]
90. Holmgren AJ, Byron ME, Grouse CK, Adler-Milstein J. Association between billing patient portal messages as e-visits and patient messaging volume. *JAMA.* 2023;329(4):339-342. [[FREE Full text](#)] [doi: [10.1001/jama.2022.24710](https://doi.org/10.1001/jama.2022.24710)] [Medline: [36607621](https://pubmed.ncbi.nlm.nih.gov/36607621/)]

91. Medicare telemedicine health care provider fact sheet. Centers for Medicare & Medicaid Services. 2020. URL: <https://www.cms.gov/newsroom/fact-sheets/medicare-telemedicine-health-care-provider-fact-sheet> [accessed 2023-11-06]
92. Davis SE, Lasko TA, Chen G, Siew ED, Matheny ME. Calibration drift in regression and machine learning models for acute kidney injury. *J Am Med Inform Assoc*. 2017;24(6):1052-1061. [FREE Full text] [doi: [10.1093/jamia/ocx030](https://doi.org/10.1093/jamia/ocx030)] [Medline: [28379439](https://pubmed.ncbi.nlm.nih.gov/28379439/)]
93. Zhang A, Xing L, Zou J, Wu JC. Shifting machine learning for healthcare from development to deployment and from models to data. *Nat Biomed Eng*. 2022;6(12):1330-1345. [FREE Full text] [doi: [10.1038/s41551-022-00898-y](https://doi.org/10.1038/s41551-022-00898-y)] [Medline: [35788685](https://pubmed.ncbi.nlm.nih.gov/35788685/)]

Abbreviations

ACP: advance care planning

CDS: clinical decision support

EBI: evidence-based intervention

EHR: electronic health record

FINDER: foodborne illness in real time

HIPAA: Health Information Portability and Accountability Act

ML: machine learning

RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance

SIF: strategic implementation framework

Edited by E Mensah; submitted 22.06.23; peer-reviewed by X Ruan, L Weinert, C Pollack, S McRoy; comments to author 25.09.23; revised version received 15.11.23; accepted 14.03.24; published 22.04.24

Please cite as:

Huguet N, Chen J, Parikh RB, Marino M, Flocke SA, Likumahuwa-Ackman S, Bekelman J, DeVoe JE

Applying Machine Learning Techniques to Implementation Science

Online J Public Health Inform 2024;16:e50201

URL: <https://ojphi.jmir.org/2024/1/e50201>

doi: [10.2196/50201](https://doi.org/10.2196/50201)

PMID:

©Nathalie Huguet, Jinying Chen, Ravi B Parikh, Miguel Marino, Susan A Flocke, Sonja Likumahuwa-Ackman, Justin Bekelman, Jennifer E DeVoe. Originally published in the Online Journal of Public Health Informatics (<https://ojphi.jmir.org/>), 22.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Online Journal of Public Health Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://ojphi.jmir.org/>, as well as this copyright and license information must be included.