

# Integrating Predictive Models Into Care: Facilitating Informed Decision-Making and Communicating Equity Issues

Paige Nong, BA; Minakshi Raj, PhD, MPH; and Jodyn Platt, PhD, MPH

**P**redictive models continue to gain attention and investment in the US health care system.<sup>1-3</sup> By leveraging large quantities of data from electronic health records (EHRs), clinical research, and other sources, these models can be used to anticipate patient outcomes such as disease progression, predict health service utilization, or identify appropriate treatments.<sup>3-6</sup> In addition to commercial vendors, health systems are building predictive analytics tools and using them for clinical decision support and administrative purposes.<sup>7-10</sup> Although some promising applications of predictive models have been identified, such as sepsis prediction,<sup>11-14</sup> the evidence base remains fragmented and marked by both methodological and implementation issues.<sup>5,15-17</sup> Furthermore, regulation of predictive models is minimal, with only limited oversight of particular diagnostic models by the FDA.<sup>18</sup> In the context of little regulatory oversight and a highly varied body of evidence, the risk of exacerbating existing inequalities or creating new types of inequity is high.

Because predictive models are built on historic data from the health care system, they reflect and can intensify long-standing racism and inequities in health care.<sup>19</sup> Bias in these models has been documented, discussed, and conceptualized in a growing body of literature.<sup>20-23</sup> In a well-known example, an algorithm used to allocate care management resources systematically discriminated against Black patients. The model used health expenditures as a proxy for health conditions. At any given risk score used to indicate predicted need, Black patients were sicker than their White counterparts who received the same score, resulting in comparatively healthier White patients receiving additional resources.<sup>24</sup> There are many historical and contemporary racial inequities in health care expenditures, quality of care, and aggressive interventions that disadvantage Black patients.<sup>19,24</sup> Reflecting these inequities, the algorithm underestimated the health needs of Black patients. Similar patterns have been identified in predictions related to kidney disease, whereby some models have contributed to or worsened racial inequity by building barriers to treatment for Black patients.<sup>25,26</sup> These examples highlight cases in which harms have been identified post facto, in part because they did not account for inequity prior to model development. They also

## ABSTRACT

**OBJECTIVES:** As predictive analytics are increasingly used and developed by health care systems, recognition of the threat posed by bias has grown along with concerns about how providers can make informed decisions related to predictive models. To facilitate informed decision-making around the use of these models and limit the reification of bias, this study aimed to (1) identify user requirements for informed decision-making and utilization of predictive models and (2) anticipate and reflect equity concerns in the information provided about models.

**STUDY DESIGN:** Qualitative analysis of user-centered design (n = 46) and expert interviews (n = 10).

**METHODS:** We conducted a user-centered design study at an academic medical center with clinicians and stakeholders to identify informational elements required for decision-making related to predictive models with a product information label prototype. We also conducted equity-focused interviews with experts to extend the user design study and anticipate the ways in which models could interact with or reflect structural inequity.

**RESULTS:** Four key informational elements were reported as necessary for informed decision-making and confidence in the use of predictive models: information on (1) model developers and users, (2) methodology, (3) peer review and model updates, and (4) population validation. In subsequent expert interviews, equity-related concerns included the purpose or application of a model and its relationship to structural inequity.

**CONCLUSIONS:** Health systems should provide key information about predictive models to clinicians and other users to facilitate informed decision-making about the use of these models. Implementation efforts should also expand to routinely incorporate equity considerations from inception through the model development process.

*Am J Manag Care.* 2022;28(1):18-24. doi:10.37765/ajmc.2022.88812

demonstrate how high the stakes of predictive models in health care are.

Models used to manage many health conditions are similarly reflective of social context and inequity. For example, hepatitis C is curable but very costly to treat. Although predictive models can be helpful in anticipating disease progression and targeting financial resources, there are serious equity concerns related to access to treatment.<sup>27</sup> Hepatitis C disproportionately affects racial and ethnic minorities, and White patients are more likely to receive treatment.<sup>28</sup> Given this context, it is important to vet models and account for existing inequity prior to their use.

Model auditing and analysis of potentially unequal outcomes have been recommended as potential safeguards against these harms.<sup>23</sup> Although such steps could be highly useful, they generally take place and identify bias post implementation.<sup>24,29,30</sup> To anticipate bias and inequity before they cause harm, social context and equity need to inform predictive models from inception through utilization.<sup>31</sup> Traditional approaches to building and using predictive analytics are limited in this capacity. These efforts tend to fall into silos of expertise such that experts in equity or social sciences, who might anticipate the “unintended consequences” of predictive models, are not regularly included in their development.<sup>32</sup>

It is critically important that inequity is considered from inception and that clinicians are provided necessary information about predictive models to make informed decisions about their use.<sup>33</sup> Much of the work on this topic to date has been focused on concepts and commentary. Although this literature is foundational to our understanding of predictive models, there is a dearth of related empirical analysis in health care. This exploratory study aimed to contribute to developing the literature by understanding how stakeholders assess the value of these models, identifying ways to facilitate informed decision-making about their use, and building equity into basic communication about the models from inception. To accomplish this, we applied user-centered design to develop a product information label that would help clinicians and administrators evaluate predictive models. We conducted follow-up expert interviews to investigate the equity implications inherent in the development and use of predictive analytics and to offer an alternative to the current approach to model development.

## METHODS

### Phase 1: Eliciting User Needs Related to Predictive Model Use

In phase 1 of this project, our design team followed the double diamond method to identify the needs and priorities of predictive model users. This is a type of user-centered design method widely deployed for information technology (IT) development that is

## TAKEAWAY POINTS

Communicating effectively about predictive models and their equity implications is critical for (1) informed decision-making and (2) preventing potential harms to patients.

- ▶ In this article we identify provider and administrator informational needs for making informed decision-making about the use of predictive models. We provide a prototype product information label for predictive tools and incorporate equity concerns.
- ▶ As the health care system works toward evidence-based medicine, the evidence on bias and inequity must be integrated into decision-making around predictive analytics. When this evidence is treated as secondary and unrelated to model development, the negative consequences of predictive models will remain “unintended” but preventable. Designing policy around model transparency, communicating key informational elements, and analyzing equity implications are important steps that health systems can take to facilitate informed decision-making and to prevent harm to patients.

characterized by iterative development cycles that produce a solution or prototype that responds to user needs.<sup>34</sup> In this study, the double diamond process was used to (1) define problems faced by clinicians and administrators in the use of predictive models, (2) develop a prototype product information label for a predictive model, and (3) elicit reactions from users.<sup>34</sup>

The first step of the double diamond method was a needs assessment.<sup>35</sup> Interviews were conducted with 46 clinicians, administrators, model developers, IT staff, and project managers at a large academic medical institution with predictive analytics infrastructure. We interviewed clinicians from a variety of specialties, including emergency medicine, pediatrics, general medicine, oncology, and primary care. The semistructured interviews sought to understand participants’ needs, concerns, and preferences related to a hypothetical hepatitis C model prototype that predicted serious illness to allocate treatment resources. A prototype product information label was developed and updated iteratively through rapid prototyping as interviews were conducted.<sup>35</sup> Participant needs and concerns informed the prototype design, which was then used to validate and expand on the findings from earlier interviews.<sup>34</sup> Participant quotes were also used to construct analytic matrices for validation and comparison of interviewees’ priorities and concerns.

Equity and bias were raised in these interviews, although they were not central to the priorities of interviewed stakeholders. It remained unclear how communication about predictive models could better incorporate information related to equity. To better understand how equity could be incorporated and addressed in a product information label, we continued the double diamond method with a second phase of the research focused specifically on these topics.

### Phase 2: Expert Interviews on Equity Implications

Building on the findings in phase 1 and recognizing the pressing concern of inequity in health care, the second phase of the study included 10 interviews with investigators and clinicians specializing in predictive models, clinical care, and health equity at the same institution who did not participate in phase 1 interviews. Using the hypothetical hepatitis C case developed in phase 1, these semistructured interviews focused on predictive models,

**TABLE 1.** Informational Elements Required for Informed Decision-Making

Informational element	Description
A. Model developers and users	<ul style="list-style-type: none"> <li>Includes who developed the model and which institutions have used it/are using it currently</li> <li>Includes endorsement or use by professional societies, specific health care systems, and EHR vendors</li> </ul>
B. Methodology	Describes how the tool was created, the problem it aims to solve, and the evidence base used for its development
C. Peer review and updates	<ul style="list-style-type: none"> <li>Notifies the user of the verification and review process that the model has undergone</li> <li>Lists if and when a model has been updated since its initial creation</li> </ul>
D. Population validation	<ul style="list-style-type: none"> <li>Describes the population for which the model was developed</li> <li>Indicates in which populations the model has been validated</li> </ul>

EHR, electronic health record.

high-cost treatment, and equity to better understand the ways that predictive models can either exacerbate or reflect inequity. All interviews were recorded and transcribed professionally. The rigorous and accelerated data reduction (RADaR) technique was used to produce concise data tables and identify themes.<sup>36</sup> Using this approach, interview content is reduced through successive iterations of coding to identify and refine codes and produce core insights from interview content. One author (P.N.) conducted the RADaR process with regular review, input, discussion, and oversight from the senior author (J.P.). Text reduction was reviewed and any thematic coding issues were resolved during regular analysis meetings. Phase 2 interviews informed an updated product information label that communicated key equity concerns to potential users.

## RESULTS

Trust in predictive models and willingness to use them depend on characteristics of and communication about the models.

### Phase 1

Four key informational elements required for informed decision-making emerged from the phase 1 interviews with clinicians, administrators, and model developers (Table 1). These were (1) information about the developers and users of a given model, (2) the methodology with which it was developed, (3) the extent to which it has been peer reviewed and updated, and (4) the populations in which it has been validated. These types of information were necessary for evaluation and confidence in a given predictive model, especially among clinicians.

The informational elements required for informed engagement with a model, and potential use, are included in the product information label prototype (Figure 1).

When viewing the product information label prototype, interviewees searched for the developer and other users of a model. Specific trusted institutions, such as leading academic medical centers or professional societies, were critical for trust in a predictive model.

Utilization of the model by well-known institutions was seen as evidence that it was vetted and reliable. Relatedly, peer review was an important mechanism of verification. Because administrators and clinicians expressed that they might not have the time to evaluate specific statistical methods used in a model, they required evidence that trusted institutions, experts, or organizations had approved it.

Although nondevelopers expressed that they did not have the time or specialized knowledge necessary to evaluate specific methodological decisions like the use of random forest models, key methodological information was required for model evaluation. The data used to create the model, specific evidence underlying its function, and its original purpose were of particular interest. Clinicians were especially concerned that a

given model was validated on their patient population, or a very similar one, to ensure that the model was appropriate for their patients.

### Phase 2

Table 2 includes the equity implications of phase 1 key informational elements, as well as 2 additional informational elements that emerged from the interviews: (1) purpose and application and (2) relationship to structural inequity.

### Implications of Phase 1 Informational Elements

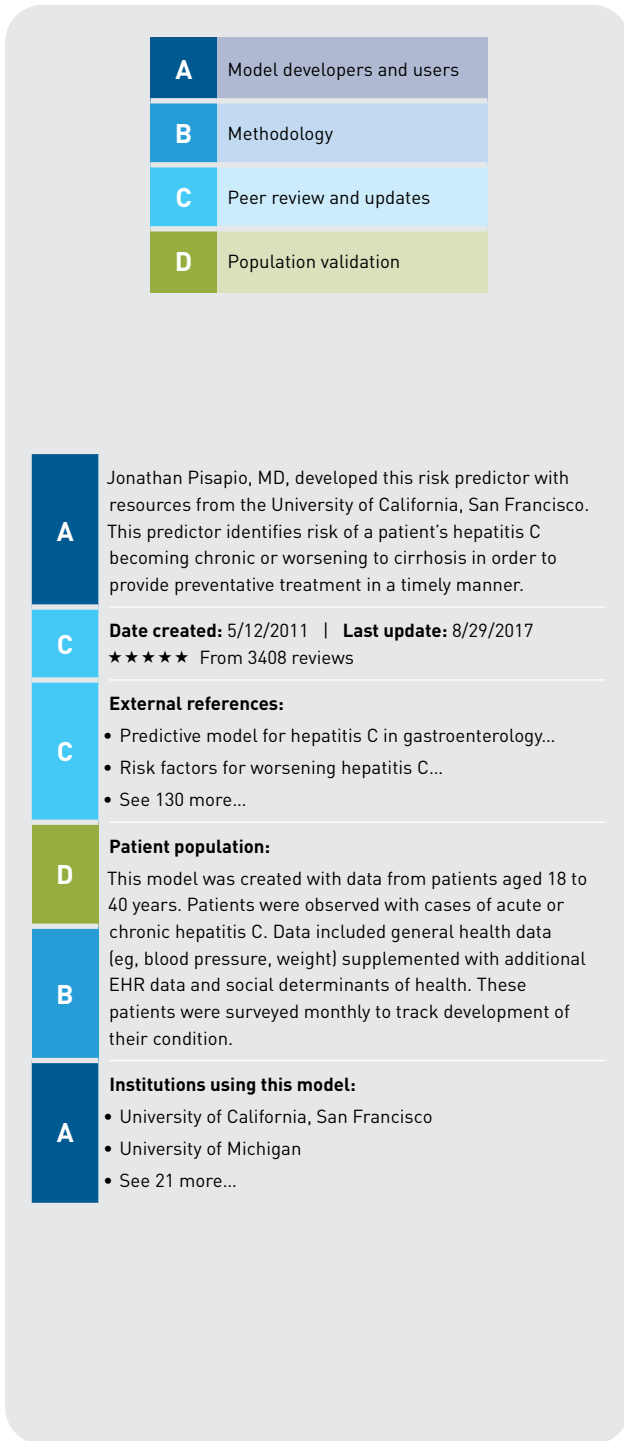
Experts reported that ethics and equity were often treated as separate from model development processes, which prioritized pragmatism and speed over ethical concerns. This separation of model development from the consideration of social context meant that equity was largely absent from predictive model development and implementation. Experts emphasized the risks to marginalized patients, including exacerbating existing inequities and building barriers to care or treatment. Accounting for and communicating these risks in model development and implementation were critically important, although there was considerable variation in opinions about where and with whom this responsibility should lie.

The lack of standards and safeguards against bias in predictive analytics was also a key threat to equity. Without clear requirements for updating and modifying a model, for example, bias cannot be systematically identified. Model development, methodology, and peer review/updates were especially affected by this lack of standards (Table 2). The implications include harms to marginalized patients, inability to systematically identify bias, and inadequate population validation on minority populations. Validation and calibration are not currently standard or required, meaning that the efficacy of a given model for subpopulations may be unclear, perpetuating inequity by advantaging patients who are already advantaged.

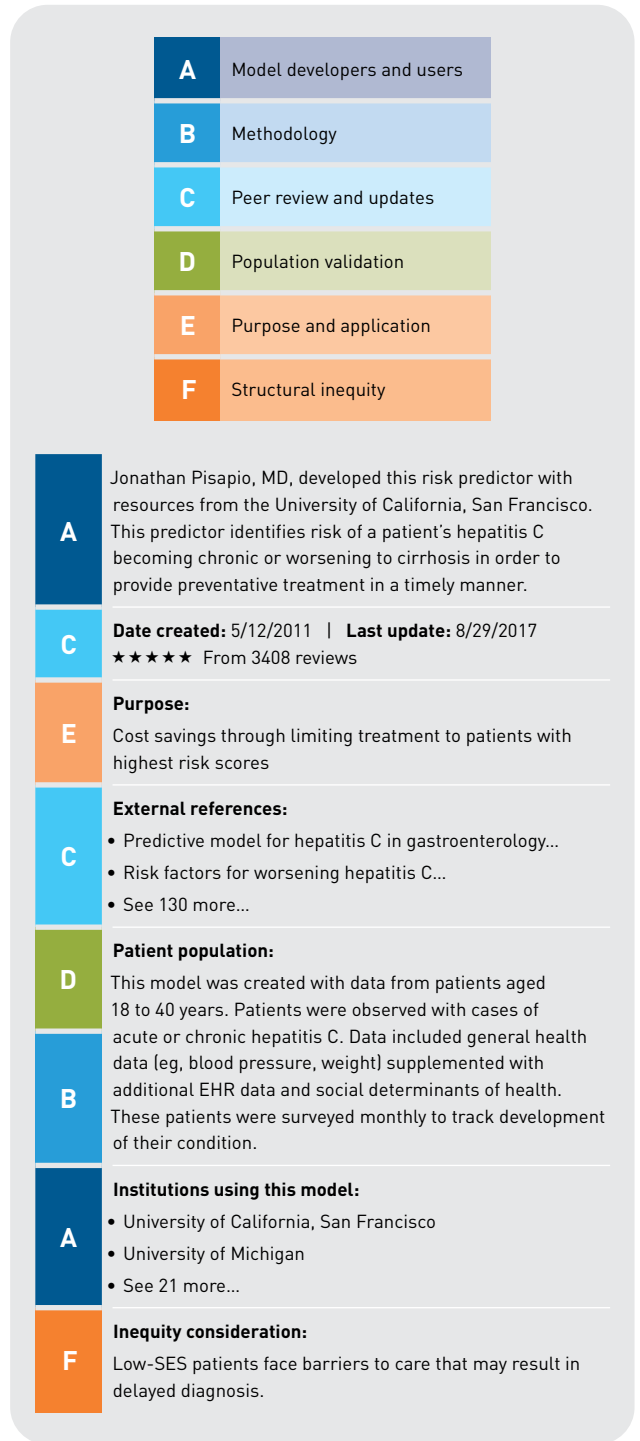
### Additional Informational Elements

Two additional key informational elements emerged from our interviews with experts. Questions about the *purpose and application*

**FIGURE 1.** Example Product Information Label Prototype<sup>a</sup>



**FIGURE 2.** Example Product Information Label Incorporating Equity Informational Elements<sup>a</sup>



EHR, electronic health record.

<sup>a</sup>This prototype product information label presents an example describing a hypothetical predictive model with the elements from Table 1, such as its creation date, date of most recent update, users, developer, references, and details about patient population. Alphabetical labels map to Table 1: (A) model developers and users, (B) methodology, (C) peer review and updates, and (D) population validation.

EHR, electronic health record; SES, socioeconomic status.

<sup>a</sup>Alphabetical labels map to Table 2: (A) model developers and users, (B) methodology, (C) peer review and updates, (D) population validation, (E) purpose and application, and (F) structural inequity.

**TABLE 2.** Equity Implications Related to Key Informational Elements of Predictive Analytics

Informational element	Current state	Equity implication
A. Model developers and users	Disciplinary silos (social dynamics and ethics are seen as separate)	Equity is not considered in development.
B. Methodology	Lack of transparency in audits and validation	Marginalized patients are at higher risk of harm when audits do not take place because bias is not identified.
C. Peer review and updates	Lack of consistent policy or standards for updates and peer review	Issues with models, such as bias, will not be identified systematically.
D. Population validation	Lack of population validation procedures and requirements	Models may be responsive to the needs of majority or dominant groups but harm smaller or underrepresented populations.
E. Purpose and application	Cost savings are a priority.	Models may focus on cost savings to the detriment of marginalized patients.
F. Relationship to structural inequity	Models are assumed to be neutral.	Predictions reflect existing inequities or create new ones.

of a given model were central to evaluation among experts. A model applied to anticipate high cost and limit access to care for patients, for example, would have fundamentally different implications than a model facilitating access to necessary treatment (see [eAppendix](#) [available at [ajmc.com](#)] for example quotes). Clarity about the issue to be addressed by a model was fundamentally important for experts considering a model's value.

A model's interaction with or reflection of *structural inequity* was also critically important. For example, multiple experts separately described biased algorithms related to kidney disease and cardiovascular function that exacerbated existing racial disparities. Because Black patients face structural barriers to health care, for example, diagnosis of kidney disease is often delayed. Black patients historically have been sicker by the time they receive a diagnosis, and a "race correction" reflecting this dynamic was built into an algorithm for future treatment. Knowledge of these social and historical dynamics and information about how a model engages with or addresses them was of primary concern. For the experts interviewed, clear communication about this aspect of a given model was necessary both for consideration in model development and for providers to understand the risks to their patients.

## DISCUSSION

In this study, we identified information necessary for informed decision-making related to predictive models and explored equity implications with experts. Our findings focus on 2 domains for health systems to consider in the development and use of predictive models: critical informational elements required for users and foundational equity concerns to address.

### Informational Elements

In phase 1, administrators and clinicians described some of the informational elements required to evaluate and potentially use predictive models. In addition to demonstrating an understanding of foundational methodological concerns relevant to clinical care

such as population validation, they also noted the challenge of evaluating every detail of the specific statistical methods. For some, it was necessary to rely on information about the use of models by trusted institutions, experts, or organizations that adopted a model. This was especially true when combined with prioritized methodological information affecting patient care. Our interviews also revealed the consistent desire on the part of clinicians to know if and how a model is validated on their patient population, or a very similar one, to ensure the appropriateness of the model. This informational element is important for health equity because homogenous training data sets lead to less accurate predictions for minority populations and marginalized groups.<sup>15,37</sup>

### Implications for Trust and Utilization

Our study results underscore ways in which informed decision-making and equity matter in the use of predictive models. The confidence and trust of clinicians are required for utilization of predictive models, but they must also be earned through validated, trustworthy models and communication of key informational elements, some of which we identify here.<sup>38</sup> For some, leveraging trust in scientific processes such as peer review and trust in organizations that are using specific models will inform whether a user will adopt a model, suggesting the importance of those organizations themselves being trusted and trustworthy agents.

Our findings also highlight the importance of communicating evidence that models are trustworthy and responsive to equity concerns. Health systems would be well served by establishing transparent policies related to population validation and communicating this critical informational element to clinicians and other users. Communication about predictive models also needs to include information and evidence on their relationships to structural inequity. Clarifying the purpose of a model in practice will facilitate critical engagement with potential negative consequences for patients. Individuals and systems considering use of these models should engage in discussions from early stages to identify any potential limitations of the models and could even



work with health systems and designers to ensure equity of the model design and its subsequent use. These critical reflections, at a system level, could be important for improving the validity of models.

### Alternative Approach to Development and Use of Predictive Models in Health Care

These study findings revealed some of the negative consequences of excluding expertise in equity from model development. Based on this outcome, we propose that perspectives of equity experts are incorporated throughout the development and implementation of predictive models and in the evaluation of and communication about predictive models in order to anticipate negative consequences of a given predictive tool. In the traditional approach, models are often developed and used in silos of expertise. After implementation, analysis of negative consequences may be conducted and bias can be identified. However, this means that patients have already been harmed. To better anticipate these consequences, an alternative approach (Figure 3) elicits and integrates the perspectives of a more diverse group of experts who are equipped to identify potential risks for marginalized patients. Future studies should expand on existing literature to examine specific outcomes and metrics that could be used to assess the equity implications of predictive models in health care. Patient perspectives should also be sought to understand perceived risks and benefits associated with predictive models.

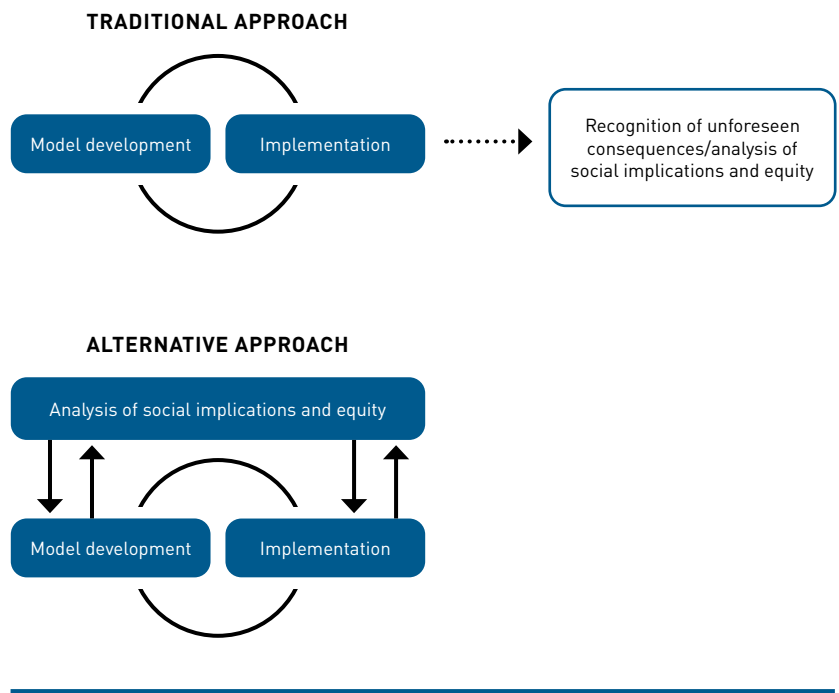
### Limitations

This is an exploratory study based in a single academic institution.<sup>39</sup> Our findings may not generalize, especially to smaller, nonacademic institutions that do not have similar model development capacity or familiarity with predictive models. Because this varies significantly and could be associated with specific EHR characteristics, future work should analyze the perspectives of various health systems.<sup>40</sup> Further validation, evaluation, and expansion in other settings will be important in future work. However, information labels, such as the one prototyped in this project, are increasingly in demand in consumer and health care domains, and our work provides an initial application that should be further explored.

## CONCLUSIONS

As the health care system continuously works toward evidence-based medicine, the evidence on bias and inequity must be included and integrated into predictive analytics. When this expertise or evidence

**FIGURE 3.** Model for Developing Equity-Informed Predictive Analytics



is treated as secondary and unrelated to model development, the negative consequences of predictive models will remain “unintended” but preventable. Designing policy around model transparency, communication of key informational elements, and analysis of equity implications are important steps that health systems can take to facilitate informed decision-making and to prevent harm to patients. If models are used caveat emptor, this leaves patients and clinicians vulnerable. More extensive evaluation and communication about model functions will be critical. Formal oversight mechanisms within and across organizations that incorporate expertise on equity in the key informational elements about predictive models will guide users in their assessments of model value and facilitate more broadly informed model utilization. ■

### Acknowledgments

The design team for phase 1 of this study included Jade Crump, Katherine Jones, Sarang Modi, and Sam Bertin. They designed the product information label prototypes and elicited responses from interviewees. The authors also want to acknowledge Tevah Platt and Drs Marie Grace Trinidad, Daniel Thiel, and Sharon Kardia for their input and contributions to the research project.

**Author Affiliations:** Department of Health Management and Policy (PN) and Department of Learning Health Sciences (JP), University of Michigan, Ann Arbor, MI; Department of Kinesiology and Community Health, University of Illinois at Urbana Champaign (MR), Champaign, IL.

**Source of Funding:** Knowledge to Treatment Optimization Program through the Michigan Department of Health and Human Services.

**Author Disclosures:** The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

**Authorship Information:** Concept and design (PN, JP); acquisition of data (PN, MR, JP); analysis and interpretation of data (PN, MR); drafting of the manuscript (PN); critical revision of the manuscript for important intellectual content (PN, MR, JP); obtaining funding (JP); administrative, technical, or logistic support (PN); and supervision (JP).

**Address Correspondence to:** Paige Nong, BA, Department of Health Management and Policy, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109. Email: ptassie@umich.edu.

## REFERENCES

- Wang F, Preinerger A. AI in health: state of the art, challenges, and future directions. *Yearb Med Inform.* 2019;28(1):16-26. doi:10.1055/s-0039-1677908
- Naylor CD. On the prospects for a (deep) learning health care system. *JAMA.* 2018;320(11):1099-1100. doi:10.1001/jama.2018.11103
- Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* 2019;20(5):e262-e273. doi:10.1016/S1470-2045(19)30149-4
- Alanazi HO, Abdullah AH, Qureshi KN. A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. *J Med Syst.* 2017;41(4):69. doi:10.1007/s10916-017-0715-6
- Futoma J, Morris J, Lucas J. A comparison of models for predicting early hospital readmissions. *J Biomed Inform.* 2015;56:229-238. doi:10.1016/j.jbi.2015.05.016
- Singh K, Valley TS, Tang S, et al. Validating a widely implemented deterioration index model among hospitalized COVID-19 patients. *medRxiv.* Preprint posted online April 29, 2020. doi:10.1101/2020.04.24.20079012
- Bhardwaj R, Nambiar AR, Dutta D. A study of machine learning in healthcare. In: *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*. Institute of Electrical and Electronics Engineers; 2017:236-241. doi:10.1109/COMPSAC.2017.164
- Tang C, Lorenzi N, Harle CA, Zhou X, Chen Y. Interactive systems for patient-centered care to enhance patient engagement. *J Am Med Inform Assoc.* 2016;23(1):2-4. doi:10.1093/jamia/ocv198
- Rosella LC, Kornas K, Yao Z, et al. Predicting high health care resource utilization in a single-payer public health care system: development and validation of the high resource user population risk tool. *Med Care.* 2018;56(10):e61-e69. doi:10.1097/MLR.0000000000000837
- Liu W, Stansbury C, Singh K, et al. Predicting 30-day hospital readmissions using artificial neural networks with medical code embedding. *PLoS One.* 2020;15(4):e0221606. doi:10.1371/journal.pone.0221606
- Singh K, Betensky RA, Wright A, Curhan GC, Bates DW, Waikar SS. A concept-wide association study of clinical notes to discover new predictors of kidney failure. *Clin J Am Soc Nephrol.* 2016;11(12):2150-2158. doi:10.2215/CJN.02420316
- Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. *J Am Med Inform Assoc.* 2020;27(11):1764-1773. doi:10.1093/jamia/ocaa143
- Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. *J Am Med Inform Assoc.* 2017;24(1):198-208. doi:10.1093/jamia/ocw042
- Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Crit Care Med.* 2018;46(4):547-553. doi:10.1097/CCM.0000000000002936
- Barda N, Yona G, Rothblum GN, et al. Addressing bias in prediction models by improving subpopulation calibration. *J Am Med Inform Assoc.* 2021;28(3):549-558. doi:10.1093/jamia/ocaa283
- Holmberg L, Vickers A. Evaluation of prediction models for decision-making: beyond calibration and discrimination. *PLoS Med.* 2013;10(7):e1001491. doi:10.1371/journal.pmed.1001491
- Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of Covid-19: systematic review and critical appraisal. *BMJ.* 2020;369:m1328. doi:10.1136/bmj.m1328
- Software as a medical device (SaMD). FDA. Updated December 4, 2018. Accessed March 28, 2021. <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>
- Benjamin R. Assessing risk, automating racism. *Science.* 2019;366(6464):421-422. doi:10.1126/science.aaz3873
- McCadden MD, Joshi S, Mazwi M, Anderson JA. Ethical limitations of algorithmic fairness solutions in health care machine learning. *Lancet Digit Health.* 2020;2(5):e221-e223. doi:10.1016/S2589-7500(20)30065-0
- McCadden MD, Joshi S, Anderson JA, Mazwi M, Goldenberg A, Zlotnick Shaul R. Patient safety and quality improvement: ethical principles for a regulatory approach to bias in healthcare machine learning. *J Am Med Inform Assoc.* 2020;27(12):2024-2027. doi:10.1093/jamia/ocaa085
- Ferryman K. Addressing health disparities in the Food and Drug Administration's artificial intelligence and machine learning regulatory framework. *J Am Med Inform Assoc.* 2020;27(12):2016-2019. doi:10.1093/jamia/ocaa133
- Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc.* 2018;25(8):1080-1088. doi:10.1093/jamia/ocy052
- 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. doi:10.1126/science.aax2342
- Braun L, Wentz A, Baker R, Richardson E, Tsai J. Racialized algorithms for kidney function: erasing social experience. *Soc Sci Med.* 2021;268:113548. doi:10.1016/j.socscimed.2020.113548
- Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight — reconsidering the use of race correction in clinical algorithms. *N Engl J Med.* 2020;383(9):874-882. doi:10.1056/NEJMs2004740
- Konerman MA, Beste LA, Van T, et al. Machine learning models to predict disease progression among veterans with hepatitis C virus. *PLoS One.* 2019;14(11):e0208141. doi:10.1371/journal.pone.0208141
- Vutien P, Hoang J, Brooks L Jr, Nguyen NH, Nguyen MH. Racial disparities in treatment rates for chronic hepatitis C. *Medicine (Baltimore).* 2016;95(22):e3719. doi:10.1097/MD.0000000000003719
- Harrison MI, Koppel R, Bar-Lev S. Unintended consequences of information technologies in health care—an interactive sociotechnical analysis. *J Am Med Inform Assoc.* 2007;14(5):542-549. doi:10.1197/jamia.M2384
- Murray SG, Wachter RM, Cucina RJ. Discrimination by artificial intelligence in a commercial electronic health record—a case study. *Health Affairs.* January 31, 2020. Accessed November 21, 2021. <https://www.healthaffairs.org/doi/10.1377/hlthaff.20200128.626576/full/>
- Rajkumar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Ann Intern Med.* 2018;169(12):866-872. doi:10.7326/M18-1990
- d'Aquin M, Troulinou P, O'Connor NE, Cullen A, Fattler G, Holden L. Towards an "ethics by design" methodology for AI research projects. In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery; 2018:54-59. doi:10.1145/3278721.3278765
- Sendak MP, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. *NPI Digit Med.* 2020;3:41. doi:10.1038/s41746-020-0253-3
- What is the framework for innovation? Design Council's evolved Double Diamond. Design Council. 2015. Accessed August 11, 2021. <https://www.designcouncil.org.uk/news-opinion/what-framework-innovation-design-councils-evolved-double-diamond>
- Kinzie MB, Cohn WF, Julian MF, Knaus WA. A user-centered model for web site design: needs assessment, user interface design, and rapid prototyping. *J Am Med Inform Assoc.* 2002;9(4):320-330. doi:10.1197/jamia.m0822
- Watkins DC. Rapid and rigorous qualitative data analysis: the "RADaR" technique for applied research. *Int J Qual Methods.* 2017;16:1-9. doi:10.1177/1609406917712131
- Kim MP, Ghorbani A, Zou J. Multiaccuracy: black-box post-processing for fairness in classification. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery; 2019:247-254. doi:10.1145/3306618.3314287
- Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. *J Med Internet Res.* 2020;22(6):e15154. doi:10.2196/15154
- Benda NC, Novak LL, Reale C, Ancker JS. Trust in AI: why we should be designing for APPROPRIATE reliance. *J Am Med Inform Assoc.* Published online November 2, 2021. doi:10.1093/jamia/ocab238
- Apathy NC, Holmgren AJ, Adler-Milstein J. A decade post-HITECH: critical access hospitals have electronic health records but struggle to keep up with other advanced functions. *J Am Med Inform Assoc.* 2021;28(9):1947-1954. doi:10.1093/jamia/ocab102

Visit [ajmc.com/link/88812](https://ajmc.com/link/88812) to download PDF and eAppendix

## eAppendix

### Example Phase 2 participant quotes

Participant ID	Quote	Code
5	“I think the question, ‘why are you developing the prediction model’, is really important... Let’s take hepatitis C, for example, and say, ‘Who is going to benefit most from a drug that doesn’t work really well and has a lot of side effects?’ That’s a very different model than, ‘Hey, you can use this drug for everybody, but you don’t have enough money. Who do you treat first?’ So, I think understanding the context of why you’re developing the predictive model is extremely important.”	Purpose/application
4	“There are certain populations of people who are going to score less well in these algorithms. It will exacerbate disparities by race and social class.”	Structural inequity
1	“What do we mean when we say we’re going to provide treatment,” you know, and “who gets it and who doesn’t?” These are questions that have very little to do with like medicine and probably more to do with... “What is justice in terms of getting this available treatment and at what level?”	Structural inequity
4	“So it turns out that African-Americans who are really, you know, have a much higher prevalence in terms of need for kidneys, fall worse on these algorithms because they get put on the need for a kidney, on the list, later than other patients”	Structural inequity
2	“(I would ask) are we really looking at parity for all patients who are infected? If you’re focusing on a certain at-risk population, I think it’s okay ... but if you’re using algorithms to- not dissuade populations, but putting some populations in front of other populations, then I don’t know that we’re doing any better.... I think it can highlight some populations that are at risk and that we should pay attention to, but I would be a little bit concerned... that it would leave other people at the wayside when the goal is really to eradicate (hepatitis c) and treat everybody.”	Purpose/application