CARE DELIVERY AND PRACTICE MANAGEMENT

Using Big Data and Predictive Analytics to Determine Patient Risk in Oncology

Ravi B. Parikh, MD, MPP[1,2]; Andrew Gdowski, DO, PhD[3]; Debra A. Patt, MD, MPH, MBA[3,4]; Andrew Hertler, MD, FACP[5]; Craig Mermel, MD, PhD[6]; and Justin E. Bekelman, MD[1,2]

overview

Big data and predictive analytics have immense potential to improve risk stratification, particularly in data-rich fields like oncology. This article reviews the literature published on use cases and challenges in applying predictive analytics to improve risk stratification in oncology. We characterized evidence-based use cases of predictive analytics in oncology into three distinct fields: (1) population health management, (2) radiomics, and (3) pathology. We then highlight promising future use cases of predictive analytics in clinical decision support and genomic risk stratification. We conclude by describing challenges in the future applications of big data in oncology, namely (1) difficulties in acquisition of comprehensive data and endpoints, (2) the lack of prospective validation of predictive tools, and (3) the risk of automating bias in observational datasets. If such challenges can be overcome, computational techniques for clinical risk stratification will in short order improve clinical risk stratification for patients with cancer.

INTRODUCTION

Big data are moving from hype to reality in medicine. Recent advances in computational capacity and machine learning have led to well-publicized breakthroughs in clinical practice: artificial intelligence algorithms can now detect pneumonia from chest x-rays[1] and diabetic retinopathy from fundoscopic images,[2,3] with performance augmenting and sometimes exceeding clinician diagnostic abilities. The field of predictive analytics has been particularly well positioned to make sense of the terabytes of data being produced by electronic health records (EHRs). Published predictive analytic algorithms have been shown to predict and sometimes prevent important events, such as readmissions from heart failure,[4] chronic obstructive pulmonary disease,[5] and neonatal sepsis.[6]

A data-rich field like oncology seems like a natural landing ground for predictive analytics. However, applications of predictive analytics are sparse in oncology despite the need for better predictions of life expectancy, acute care use, adverse effects, and genomic and molecular risk. We argue that current predictive analytic interventions could address sizable gaps in risk stratification strategies in oncology. Burgeoning applications of predictive analytics in pathology interpretation, drug development, and population health management provide a way forward for future tools to move into clinical practice. However, to achieve this potential, clinicians, developers, and policymakers must address the research, technical, and regulatory barriers that hamper applications of analytics in oncology.

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on May 17, 2019 and published at ascopubs.org on May 17, 2019: DOI https://doi.org/10.1200/EDBK_238891

GAPS IN RISK STRATIFICATION STRATEGIES IN ONCOLOGY

Risk stratification in oncology suffers from a lack of access to relevant prognostic information, a need for time-consuming manual input of data, a lack of access to comprehensive data, and, in some cases, over-reliance on clinician intuition. Consider the case of prognostication of a patient with a metastatic, incurable cancer. Prospective data suggest that clinicians are poor at estimating prognosis, even among patients with advanced solid malignancies.[7] Deficiencies in identifying patients at high risk of death can lead to overly aggressive end-of-life care or unnecessary acute care use among patients with cancer.[8] Although prognostic aids exist in oncology for some specific cancers, they are rarely used because they do not apply to most cancers,[9,10] may not incorporate genetic information, do not identify most patients who will die within 1 year,[11] and require time-consuming manual data input.[12] Furthermore, assessment of prognostic variables like performance status is subject to inter-clinician variability and bias.[13] Even less published data are available for determining risk of other important outcomes, such as hospitalization or adverse effects in patients with cancer.

A driving impetus for improving risk stratification models in oncology is the push toward more patient-centric care. Moreover, changing reimbursement models, namely alternative payment models such as bundled payments, will promote the right care for the right patient, no longer exclusively incentivizing the

PRACTICAL APPLICATIONS

- Although recent advances in computational capacity and machine learning have led to well-publicized breakthroughs in clinical risk stratification, these advances are noticeably absent in oncology.
- A driving impetus for improving risk stratification models in oncology is the push toward more patient-centric care.
- Several current use cases of predictive analytics in population health management, radiomics, and pathology interpretation are potentially generalizable in oncology practice.
- Predictive analytic tools are poised to make inroads in routine clinical decision support and genomic risk stratification within the next 5 years.
- Clinicians and policymakers must address difficulties in comprehensive data acquisition, the lack of prospective validation of predictive tools, and the risk of perpetuating bias in observational datasets prior to widespread implementation of predictive analytic tools.

volume of services rendered.[14] Oncologists are increasingly expected to tailor care based on a patient's formal risk for certain outcomes. This requires capturing and interpreting data about populations, episodes of care, and specific clinical conditions. The Centers for Medicare and Medicaid Services, as part of the Oncology Care Model, has assembled comprehensive datasets of Medicare beneficiaries and has been working with EHR vendors to improve data collection and data needs.[15-17] However, although increasingly rich data integrating clinical and use factors have become available, robust predictive tools are needed to determine future risk of acute use or other poor outcomes.

Decision aids based on predictive analytics have shown to improve value-based clinical decision-making in areas such as readmission risk prevention in the general inpatient setting.[18,19] There is a critical need to use similar tools to improve clinician decision-making and population health strategies in oncology.

CURRENT USE CASES OF PREDICTIVE ANALYTICS IN ONCOLOGY

Predictive analytic tools allow automated forecasting of future health outcomes for individuals or populations based on algorithms derived from historical patient data.[20] As the amount of EHR, radiology, genomic, and other data have increased in oncology, there have been several use cases that are potentially generalizable.

Population Health Management

A key area of population health management is directing interventions to high-risk patients to avoid poor outcomes. Predictive algorithms can identify patients at risk for mortality or acute care use among patients receiving chemotherapy.[20-22] This prediction could be used to influence clinician behavior along the cancer spectrum, such as after chemotherapy,[22] after colorectal cancer surgery,[23,24] or in discharge planning.[25] Intervening with these high-risk patients may curb overuse of resources. Indeed, institutions like Penn Medicine and New Century Health use predictive algorithms to identify patients with cancer who are at high risk of imminent hospitalization or emergency department visit to target care management solutions like proactive phone calls or visits.[26,27] Although EHR data can be notoriously difficult to use, individuals at Google[28] have reported on using the Fast Healthcare Interoperability Resources format to expedite the laborious process of extracting data from EHRs. The group developed deep learning methods from the Fast Healthcare Interoperability Resources format from over 46 billion data points to accurately predict multiple medical events, including in-hospital mortality, readmissions, prolonged length of stay, and discharge diagnoses.

Radiomics

The expanding field of radiomics is one example of how predictive analytics models are beginning to be used in oncology. Radiomics is a field of texture analysis that uses quantitative data from scans to study tumor characteristics.[29] These characteristics can be used to assist with detecting, characterizing, and monitoring solid tumors.[30] Computer-aided detection has applications in detecting cancerous lung nodules on CT[31] and prostate lesions on MRI[32] and may have applications in automated staging of tumors.[33] Perhaps most interestingly, artificial intelligence–based algorithms applied to lung cancer CTs can predict important outcomes, such as mutational status and risk of distant metastases.[34,35] Because radiologic data systems can inform decisions about care delivery, dynamic MRI can be used to detect early responses to treatment and inform clinicians of tumor response before standard predictors of response would be used.[36]

Pathology

Pathology is another field critical to oncology practice that is poised to gain benefits from predictive analytics. There is considerable heterogeneity among pathologists in areas such as non–small cell lung cancer detection from bronchoscopic biopsies and Gleason Score determination from prostate biopsies.[37,38] Inaccurate biopsy reads can lead to unnecessary or inappropriate treatment strategies. Artificial intelligence algorithms can detect metastatic breast cancer from images of sentinel lymph node biopsies with high discrimination (area under the receiving operator

characteristic curve, 0.99) comparable to pathologists' interpretations.[39] These models allow for improved ability to scan large tissue sections to identify cancer cells and may help improve the workflow of pathologists by allowing them to dedicate more time to other tasks.

Models built for tumor pathologic characteristics to prognosticate outcomes and predict response to therapy are widespread for some diseases. Examples of those widespread in clinical practice are the 21-gene recurrence score and the 70-gene recurrence score used in breast cancer to facilitate the determination of utility of chemotherapy in patients with early-stage breast cancer.

FUTURE USE CASES

Clinician Decision Support

As predictive analytics tools reach a threshold of performance, oncology clinicians will increasingly use such tools to influence routine aspects of patient care. Predictive algorithms have been shown in prospective settings to decrease the time necessary to respond to patients with sepsis and to ensure timelier treatment of patients with stroke.[40,41] Although advanced predictive models have not been routinely implemented in clinical practice, predicting adverse events from chemotherapy, likely duration of response from chemotherapy, recurrence risk, and overall life expectancy at the point of care could be potential applications of analytics used to improve clinicians' decision-making.[42] As a proof of concept, several real-time EHR-based algorithms have been developed to estimate oncology patients' risk of short-term mortality prior to chemotherapy initiation.[20,21,43] These algorithms, based on structured and unstructured EHR data, are theoretically applicable to any EHR. Although the prospective applications of these algorithms are unclear, accurate mortality predictions could be extremely useful to oncologists at the point of care.

Genomic Risk Stratification

Because germline testing and next-generation tumor sequencing increase among patients with cancer, it is necessary to develop robust algorithms that can predict risk based on thousands of genes sequenced. Predictive tools based on patient history and clinical characteristics can be used to target genetic testing to certain individuals because tools like next-generation sequencing are too expensive as a blindly administered screening approach for an entire population. Machine learning models applied to targeted next-generation sequencing panels have been shown to accurately stratify actual variants from artifacts; this is a potentially useful predictive tool because variants of unknown significance can cause considerable confusion in interpretation among physicians and patients.[44,45] Additionally, genomic risk stratification can predict which patients will benefit from breast cancer screening. A group in the United Kingdom found that offering breast cancer mammography to women with a high genetic risk of breast cancer decreased overdiagnosis and improved cost-effectiveness compared with the current breast cancer screening paradigm based on age.[46] Because the field of genomics has been progressing rapidly, predicting risk will continue to evolve and will likely be most valuable in a multimodal context with incorporation of other clinical data points.

CHALLENGES IN THE APPLICATION OF ANALYTICS IN ONCOLOGY

Data Acquisition

Developing robust risk stratification models based on the experience of large numbers of patients likely improves costs and outcomes, but the major limitation is a lack of quality data. The largest hurdle facing risk-based models, particularly in oncology, is that certain aspects of patient data are limited. In claims-based data sets, emergency room visits and hospitalizations are often not captured and aggregated into accessible big data sets in a timely fashion. In particular, accurate date of death often requires querying multiple data sources, making prediction of mortality difficult. Additionally, virtually no data are collected on patients at home, which is where patients spend the vast majority of their time. Novel ways to frequently collect real-time data on oncology patients may prove helpful in preventing unnecessary hospitalizations by exposing patterns that exist in patients at the early signs of illness.

Real-world data sources may increasingly allow for real-time collection of EHR-based data. Predictive algorithms based on real-world data could be immediately actionable and may be more applicable than predictive algorithms based on clinical trials, which often exclude relevant segments of the population.[47,48] Real-world datasets, such as those from Flatiron Health and ASCO CancerLinQ, could serve this purpose but also have substantial limitations due to their dependence on manual curation and limitations due to the variability of the user interface with the medical health record.[49,50]

Prospective Validation of Algorithms

Many recent U.S. Food and Administration (FDA) clearances of predictive algorithms have been primarily based on improvement in statistical endpoints, such as area under the receiving operator characteristic curve or positive predictive value.[51] However, few algorithms have rigorously studied the impact of predictive algorithms on meaningful clinical endpoints, such as overall survival or process metrics such as time to diagnosis, particularly in oncology.[52] The FDA's Digital Health Innovation Action Plan has provided a precertification program as a pipeline for streamlined prospective evaluations of potential analytic tools for

purposes of clinical use that may be a venue for future oncology predictive analytic devices to obtain regulatory approval.[53] Other standards for prospective evaluation and regulatory approval of advanced predictive algorithms have been proposed that could serve as a standardization tool for validating predictive algorithms in oncology.[51,54]

Ensuring Representativeness and Mitigating Bias

One risk of using historical retrospective data to train predictive analytic models is that predictions may reinforce existing biases in clinical care. Algorithms that are based on subjective clinical data or access to health care could systematically bias against certain groups of patients.[55] Consider the example of a predictive algorithm based on tumor genomic data for a particular cancer. Data sets used to train the algorithm may contain low numbers of patients of certain ethnic minorities. This may result in incorrect classification of tumor genetic variants for minority populations.[56] Conversely, a lack of data from unrepresented populations could preclude the ability to identify predictive genetic variants from under-represented populations, compromising the generalizability of the predictive model.[57] When generating predictive models for risk stratification, it will be important to ensure representativeness of all populations of interest in a training set and to ensure audit mechanisms after the predictive tool is devised to ensure that under-represented groups do not encounter systematic bias in predictive output.[51]

PREDICTIVE ANALYTICS: THE NEXT BREAKTHROUGH OF PRECISION ONCOLOGY

Just as discoveries in genetic and molecular classification of tumors have dramatically improved biologic risk stratification, oncology is poised to benefit from advances in computational techniques for clinical risk stratification of patients with cancer. Advanced algorithms predicting risk of use, costs, and clinical outcomes will likely play an increasing role in shaping the clinical care of patients with oncology. Merging insights from clinical-, genetic-, and molecular-based prediction may ensure a new era of comprehensive risk stratification in oncology with high accuracy, enabling true precision oncology.

ACKNOWLEDGMENT

R. B. Parikh and A. Gdowski served as co-first authors of this article.

AFFILIATIONS

[1]Penn Center for Cancer Care Innovation at the Abramson Cancer Center, University of Pennsylvania, Philadelphia, PA
[2]Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA
[3]Dell Medical School at The University of Texas at Austin, Austin, TX
[4]Texas Oncology, Dallas, TX
[5]Evolent Health, Arlington, VA
[6]Google, Mountain View, CA

CORRESPONDING AUTHOR

Ravi B. Parikh, MD, MPP, 3400 Civic Center Blvd., 12th Floor, Abramson Cancer Center, Philadelphia, PA 19104; email: ravi.parikh@uphs.upenn.edu.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTERESTAND DATA AVAILABILITY STATEMENT

Disclosures provided by the authors and data availability statement (if applicable) are available with this article at DOI https://doi.org/10.1200/EDBK_238891.

REFERENCES

1. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med. 2018;15:e1002686.
2. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316:2402-2410.
3. Ting DSW, Cheung CY-L, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA. 2017;318:2211-2223.
4. Amarasingham R, Patel PC, Toto K, et al. Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study. BMJ Qual Saf. 2013;22:998-1005.
5. Shams I, Ajorlou S, Yang K. A predictive analytics approach to reducing 30-day avoidable readmissions among patients with heart failure, acute myocardial infarction, pneumonia, or COPD. Health Care Manage Sci. 2015;18:19-34.
6. Escobar GJ, Puopolo KM, Wi S, et al. Stratification of risk of early-onset sepsis in newborns ≥ 34 weeks' gestation. Pediatrics. 2014;133:30-36.
7. Christakis NA, Lamont EB, Parkes CM, et al. Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. BMJ. 2000;320:469-472.
8. Sborov K, Giaretta S, Koong A, et al. Impact of accuracy of survival predictions on quality of end-of-life care among patients with metastatic cancer who receive radiation therapy. J Oncol Pract. 2019;15:e262-e270.

9. Fong Y, Evans J, Brook D, et al. The Nottingham prognostic index: five- and ten-year data for all-cause survival within a screened population. Ann R Coll Surg Engl. 2015;97:137-139.

10. Alexander M, Wolfe R, Ball D, et al. Lung cancer prognostic index: a risk score to predict overall survival after the diagnosis of non-small-cell lung cancer. Br J Cancer. 2017;117:744-751.

11. Lakin JR, Robinson MG, Bernacki RE, et al. Estimating 1-year mortality for high-risk primary care patients using the "surprise" question. JAMA Intern Med. 2016; 176:1863-1865.

12. Morita T, Tsunoda J, Inoue S, et al. The Palliative Prognostic Index: a scoring system for survival prediction of terminally ill cancer patients. Support Care Cancer. 1999;7:128-133.

13. Chow R, Chiu N, Bruera E, et al. Inter-rater reliability in performance status assessment among health care professionals: a systematic review. Ann Palliat Med. 2016;5:83-92.

14. Burwell SM. Setting value-based payment goals: HHS efforts to improve U.S. health care. N Engl J Med. 2015;372:897-899.

15. Center for Medicare & Medicaid Innovation. Oncology care model. https://innovation.cms.gov/initiatives/oncology-care/. Accessed October 17, 2018.

16. Kline R, Adelson K, Kirshner JJ, et al. The Oncology Care Model: perspectives from the Centers for Medicare & Medicaid Services and participating oncology practices in academia and the community. Am Soc Clin Oncol Educ Book. 2017;37:460-466.

17. Kline RM, Bazell C, Smith E, et al. Centers for Medicare and Medicaid Services: using an episode-based payment model to improve oncology care. J Oncol Pract. 2015;11:114-116.

18. Ostrovsky A, O'Connor L, Marshall O, et al. Predicting 30- to 120-day readmission risk among Medicare fee-for-service patients using nonmedical workers and mobile technology. Perspect Health Inf Manag. 2016;13:1e.

19. Conn J. Predictive analytics tools help hospitals reduce preventable readmissions. Mod Healthc. 2014;44:16-17.

20. Elfiky AA, Pany MJ, Parikh RB, et al. Development and application of a machine learning approach to assess short-term mortality risk among patients with cancer starting chemotherapy. JAMA Netw Open. 2018;1:e180926.

21. Bertsimas D, Dunn J, Pawlowski C, et al. Applied informatics decision support tool for mortality predictions in patients with cancer. *JCO Clin Cancer Informatics*. 2018;2:1-11.

22. Brooks GA, Kansagra AJ, Rao SR, et al. A clinical prediction model to assess risk for chemotherapy-related hospitalization in patients initiating palliative chemotherapy. JAMA Oncol. 2015;1:441-447.

23. Yeo H, Mao J, Abelson JS, et al. Development of a nonparametric predictive model for readmission risk in elderly adults after colon and rectal cancer surgery. J Am Geriatr Soc. 2016;64:e125-e130.

24. Fieber JH, Sharoky CE, Collier KT, et al. A preoperative prediction model for risk of multiple admissions after colon cancer surgery. J Surg Res. 2018; 231:380-386.

25. Manning AM, Casper KA, Peter KS, et al. Can predictive modeling identify head and neck oncology patients at risk for readmission? Otolaryngol Head Neck Surg. 2018;159:669-674.

26. Vogel J, Evans TL, Braun J, et al. Development of a trigger tool for identifying emergency department visits in patients with lung cancer. Int J Radiat Oncol Biol Phys. 2017;99:S117.

27. Furlow B. Predictive analytics reduces chemotherapy-associated hospitalizations. Managed Healthcare Executive. https://www.managedhealthcareexecutive.com/mhe-articles/predictive-analytics-reduces-chemotherapy-associated-hospitalizations. Accessed March 13, 2019.

28. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. npj Digital Med. 2018;1.

29. Wong AJ, Kanwar A, Mohamed AS, et al. Radiomics in head and neck cancer: from exploration to application. Transl Cancer Res. 2016;5:371-382.

30. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J Clin. 2019;69:127-157.

31. Chan H-P, Hadjiiski L, Zhou C, et al. Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography-a review. Acad Radiol. 2008; 15:535-555.

32. Wang S, Burtt K, Turkbey B, et al. Computer aided-diagnosis of prostate cancer on multiparametric MRI: a technical review of current research. BioMed Res Int. 2014;2014:789561.

33. Song SE, Seo BK, Cho KR, et al. Computer-aided detection (CAD) system for breast MRI in assessment of local tumor extent, nodal status, and multifocality of invasive breast cancers: preliminary study. Cancer Imaging. 2015;15:1.

34. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach [published correction appears in Nat Commun. 2014;5:4644]. Nat Commun. 2014;5:4006.

35. Coroller TP, Grossmann P, Hou Y, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol. 2015;114:345-350.

36. Sorace AG, Wu C, Barnes SL, et al. Repeatability, reproducibility, and accuracy of quantitative MRI of the breast in the community radiology setting. J Magn Reson Imaging. 2018;48:695-707.

37. Yu K-H, Zhang C, Berry GJ, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun. 2016; 7:12474.

38. Sooriakumaran P, Lovell DP, Henderson A, et al. Gleason scoring varies among pathologists and this affects clinical risk in patients with prostate cancer. Clin Oncol (R Coll Radiol). 2005;17:655-658.

39. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al; CAMELYON16 Consortium. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA. 2017;318:2199-2210.

40. Hravnak M, Devita MA, Clontz A, et al. Cardiorespiratory instability before and after implementing an integrated monitoring system. Crit Care Med. 2011; 39:65-72.

41. Raza SA, Barreira CM, Rodrigues GM, et al. Prognostic importance of CT ASPECTS and CT perfusion measures of infarction in anterior emergent large vessel occlusions. J Neurointerv Surg. Epub 2018 Dec 7.

42. Parikh RB, Kakad M, Bates DW. Integrating predictive analytics into high-value care: the dawn of precision delivery. JAMA. 2016;315:651-652.

43. Burki TK. Predicting lung cancer prognosis using machine learning. Lancet Oncol. 2016;17:e421.

44. van den Akker J, Mishne G, Zimmer AD, et al. A machine learning model to determine the accuracy of variant calls in capture-based next generation sequencing. BMC Genomics. 2018;19:263.

45. Welsh JL, Hoskin TL, Day CN, et al. Clinical decision-making in patients with variant of uncertain significance in BRCA1 or BRCA2 genes. Ann Surg Oncol. 2017; 24:3067-3072.

46. Pashayan N, Morris S, Gilbert FJ, et al. Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer: a life-table model. JAMA Oncol. 2018;4:1504-1510.

47. Karanis TB, Bermudez Canta FA, Mitrofan L, et al. Research' vs 'real world' patients: the representativeness of clinical trial participants. Annals Oncol. 2016; 27:1570P.

48. O'Connor JM, Fessele KL, Steiner J, et al. Speed of adoption of immune checkpoint inhibitors of programmed cell death 1 protein and comparison of patient ages in clinical practice vs pivotal clinical trials. JAMA Oncol. 2018;4:e180798.

49. Flatiron Health Database. https://flatiron.com/real-world-evidence/. Accessed August 26, 2018.

50. ASCO. CancerLinQ Discovery data access toolkit. https://cancerlinq.org/sites/cancerlinq.org/files/ASC-1725-CancerLinQ-Discovery-Toolkit-Update-v4.pdf. Accessed August 26, 2018.

51. Parikh RB, Obermeyer Z, Navathe AS. Regulation of predictive analytics in medicine. Science. 2019;363:810-812.

52. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25:44-56.

53. U.S. Department of Health and Human Services. Digital Health Software Precertification (Pre-Cert) Program. https://www.fda.gov/MedicalDevices/DigitalHealth/UCM567265. Accessed February 28, 2019.

54. Wolff RF, Moons KGM, Riley RD, et al; PROBAST Group. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170:51-58.

55. Mullainathan S, Obermeyer Z. Does machine learning automate moral hazard and error? Am Econ Rev. 2017;107:476-480.

56. Manrai AK, Funke BH, Rehm HL, et al. Genetic misdiagnoses and the potential for health disparities. N Engl J Med. 2016;375:655-665.

57. Struewing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med. 1997;336:1401-1408.