Journal of the American Medical Informatics Association, 27(9), 2020, 1443–1449
doi: 10.1093/jamia/ocaa152
Brief Communications

Brief Communications

Electronic health records contain dispersed risk factor information that could be used to prevent breast and ovarian cancer

Thomas H. Payne [1,2] Lue Ping Zhao,[3] Calvin Le,[1] Peter Wilcox,[1] Troy Yi,[1] Jesse Hinshaw,[1] Duncan Hussey,[1] Alex Kostrinsky-Thomas,[4] Malika Hale,[1] John Brimm,[1] Fuki M. Hisama[1,2]

[1]University of Washington School of Medicine, Seattle, Washington, USA, [2]Brotman Baty Institute for Precision Medicine, Seattle, Washington, USA, [3]Fred Hutchison Cancer Research Center, Seattle, Washington, USA, and [4]Pacific Northwest University of Health Sciences, Yakima, Washington, USA

Corresponding Author: Thomas H. Payne, MD, Department of Medicine, Box 359780, 325 Ninth Ave, Seattle, WA 98104-2499, USA (tpayne@u.washington.edu)

Received 8 May 2020; Revised 11 June 2020; Editorial Decision 16 June 2020; Accepted 23 June 2020

ABSTRACT

Objective: The genetic testing for hereditary breast cancer that is most helpful in high-risk women is underused. Our objective was to quantify the risk factors for heritable breast and ovarian cancer contained in the electronic health record (EHR), to determine how many women meet national guidelines for referral to a cancer genetics professional but have no record of a referral.

Methods and Materials: We reviewed EHR records of a random sample of women to determine the presence and location of risk-factor information meeting National Comprehensive Cancer Network (NCCN) guidelines for a further genetic risk evaluation for breast and/or ovarian cancer, and determine whether the women were referred for such an evaluation.

Results: A thorough review of the EHR records of 299 women revealed that 24 (8%) met the NCCN criteria for referral for a further genetic risk evaluation; of these, 12 (50%) had no referral to a medical genetics clinic.

Conclusions: Half of the women whose EHR records contain risk-factor information meeting the criteria for further genetic risk evaluation for heritable forms of breast and ovarian cancer were not referred.

Key words: oncology, genetics, electronic health records

INTRODUCTION

Over a quarter-million women are diagnosed with breast and ovarian cancer each year.[1] Of these, 5-10% have cancers linked to inherited pathogenic variants that, if identified before cancer develops, might prompt an intervention to avoid morbidity and fatal disease.[2] Yet the genetic testing for hereditary breast cancer that is most helpful in women at increased risk for heritable cancer is underused.[3] Even among women newly diagnosed with breast cancer, fewer than half with clinical indications receive a formal genetic risk assessment.[4] Several reasons that women do not receive formal genetic risk assessments have been identified, including older age at diagnosis, insurance status, distance from genetic services, and patient attitude about the value of genetic services. The single most important factor, however, is the lack of physician referrals to genetics services, even for patients who meet national guidelines for a formal genetic evaluation.[5] The identification of appropriate candidates for referral at those locations where women receive primary care is essential.

The use of risk assessment tools in the primary care setting has been shown to have moderate to high accuracy in guiding which patients should be referred to a cancer genetics professional, but the use of these tools requires additional time and effort to gather personal and family history information.[6]

Risk factor information is often available in the electronic health record (EHR), because it has been gathered and stored in the course of routine care. The full story of a patient's risk for heritable cancer within their record often does not exist in a single location. It is fragmented across entries created by many authors, over many years, in many locations and formats, and commonly from many different institutions in which women have received care over their lifetimes. As a result, what patients and providers might know from the full content of EHR records differs from what they are acting on today. The focus of our study is on whether or not providers referred patients who met the National Comprehensive Cancer Network (NCCN) criteria.

We define "unrecognized EHR risk-factor information" as information that exists within a patient's EHR record but is not known by current treating providers. If this unrecognized EHR risk factor information could be found and acted on, additional women at high genetic risk could be identified and referred for genetic counseling as a preventative measure, with the goal of improving their health outcomes. The objective of this study is to characterize and quantify the unrecognized EHR risk-factor information related to breast and ovarian cancer, to determine how many women meet national guidelines for referral to a cancer genetics professional based on information in their EHR but have no record of such a referral.

MATERIALS AND METHODS

Population

We identified 9573 women who were ≥30 years old and were seen ≥5 times or hospitalized ≥2 times in the University of Washington (UW) Medicine health system in western Washington state between April 2018 and April 2019, using the University of Washington Enterprise Data Warehouse, and then randomly selected patients for manual review. Given the time available for chart reviews, charts of the first 299 randomly selected patients were reviewed.

Chart review

Six medical students trained in EHR use reviewed UW Medicine inpatient (Cerner) and outpatient (Epic) EHR records of the random subset of this sample to ascertain the presence and location of breast and ovarian cancer risk-factor content within EHR records in notes, reports, orders, outside records, and scanned documents. The review included UW Medicine health system records, documents received from other institutions, handwritten questionnaires completed by patients, and records viewable from other institutions using vendor information exchange tools (CareEverywhere). It included structured data, such as encoded problem lists and the Epic family history tool, and unstructured data, such as the narrative text of progress notes, consultant notes, and imaging requisitions.

Data collected

Chart reviewers recorded all risk-factor information, as defined in the NCCN Guidelines version 3.2019 criteria,[7] using the REDCap[8] electronic data capture tool hosted at the University of Washington. NCCN Guidelines include personal and family histories of many types of cancer, along with age of onset and degree of relatedness; Ashkenazi Jewish ancestry; known pathogenic/likely pathogenic variants in a cancer susceptibility gene; and other factors. A complete list of the risk-factor information recorded is in the Table 1.

Table 1. Data collected during manual chart review

Patient ID
Zip code
Race
Ethnicity
Primary care provider
Clinic visits last year
Hospitalizations last year
Referred to genetics clinic (Y/N)
Date seen in genetics clinic
BMI
Age of menarche
Menopause reached (Y/N)
Menopause age
Gravida
Parity
Age at first childbirth
Hormone replacement therapy (Y/N)
Prior breast biopsy (Y/N)
Findings of breast biopsy
Breast density
BiRad
Breast cancer diagnosis (Y/N)
—For each: age at diagnosis, source of breast cancer diagnosis, triple negative (Y/N), lobular (Y/N)
Cowden Syndrome criteria (Y/N)
Personal history of pancreatic cancer (Y/N)
Personal history of ovarian, fallopian or primary peritoneal cancer (Y/N)
Ashkenazi Jewish ancestry (Y/N)
Founder mutation in relative (Y/N)
Known pathogenic/likely pathogenic variant in a cancer susceptibility gene found on tumor testing in the family
Family history of cancer (list)
—For each: information source, age of onset, relatedness, type (17 listed types + other)

Note: The list consists of data used in criteria for further genetic risk evaluations in the National Comprehensive Cancer Network Version 3.2019, Breast and/or Ovarian Cancer Genetic Assessment.[6] The italicized text indicates data that were extracted from the Enterprise Data Warehouse rather than from a manual chart review but that were included to confirm the patient's identity during the chart review.

BiRad: Breast Imaging Reporting and Data System; BMI, body mass index; N: no; Y: yes.

Determination of criteria for referral

The NCCN criteria for further genetic risk evaluation for breast and/or ovarian cancer were applied to determine which women met the criteria for referral for a genetic consultation. Of these, we noted what percentage had a record of a referral to a medical geneticist or genetic counselor. We randomly selected 10% of these charts to be reviewed a second time by another reviewer, to measure interviewer agreement. The number of charts reviewed was constrained by the time required for a manual EHR review and the project timeline. The project was approved by the University of Washington Institutional Review Board.

RESULTS

We reviewed the complete EHR records of 299 women. Each review required up to 1 hour and included a detailed review of UW Medi-

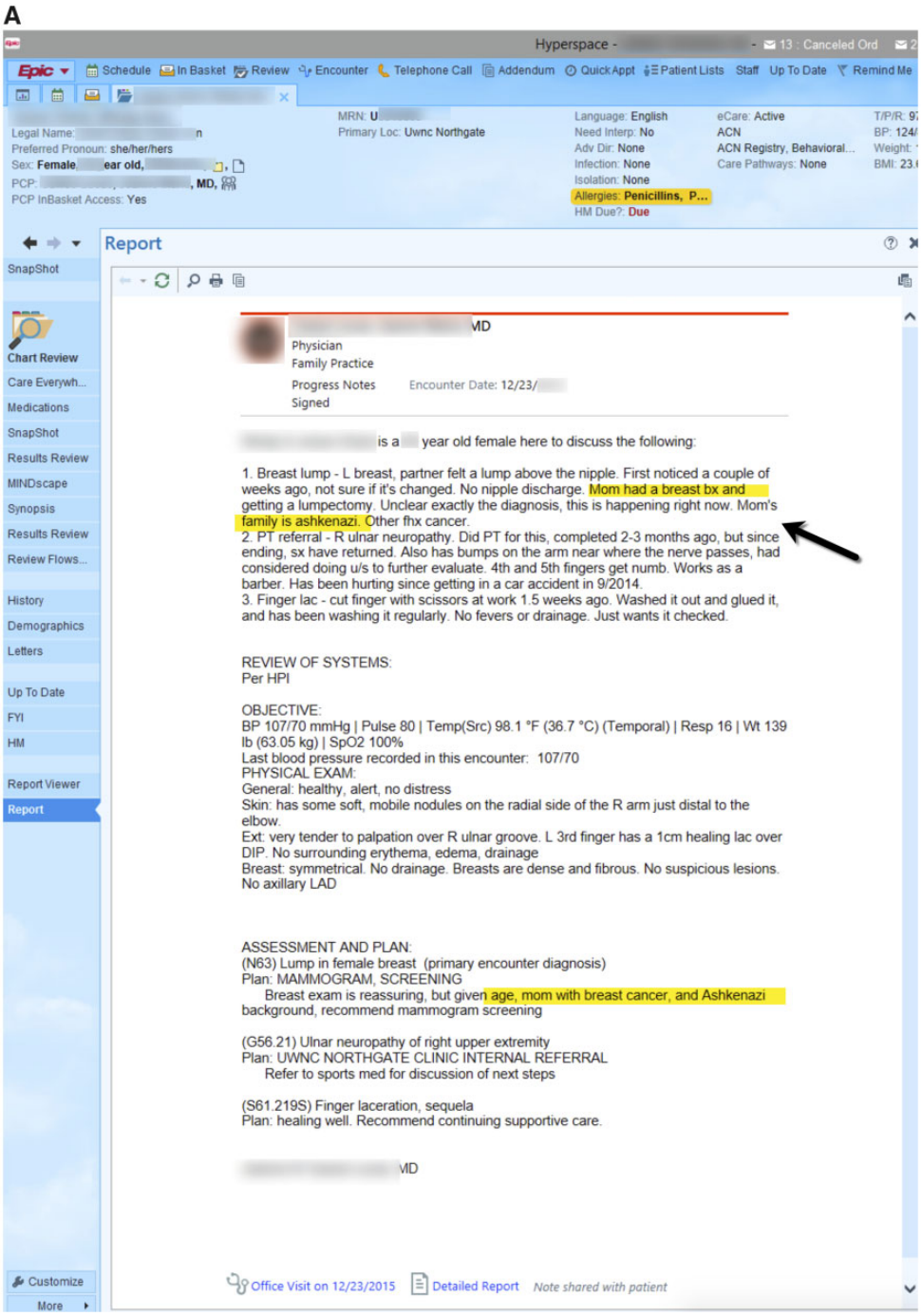

Figure 1. Examples of EHR documents (deidentified) containing risk-factor information listed in NCCN Guidelines Version 3.2019. (A) EHR note that includes both family history of breast cancer and Ashkenazi Jewish heritage in narrative text. (B) Outside EHR dermatology note that includes information on family history of breast and uterine cancer (arrow). (C) Scanned outside imaging requisition that includes family history of cancer (arrow). EHR: electronic health record.

cine EHRs and outside records. Risk-factor information was found in many locations within the EHRs, including scanned notes and imaging requisitions, outside notes, the family history module, and in note narrative text (Figure 1).

Using the risk-factor information in the EHR, 24 women (8%) met the NCCN criteria for referral for a further genetic risk evaluation; of these, 12 (50%) had no record of a referral to a medical genetics clinic. The location of risk-factor information for these 12 patients is shown in Figure 2. The most common risk factors were a family history or personal history of ovarian cancer or breast cancer. Family history information was frequently found in note narrative text rather than in the EHR family history tool. Agreement between 2 independent chart reviewers was high for quantitative risk factors, including body mass index, age at the first birth, and age at menarche (kappa values = 0.93, 0.99, and 0.98, respectively; *P*-values much smaller than .001), but was moderate for family history of cancer (kappa = 0.55; *P*-value = .0038).

B

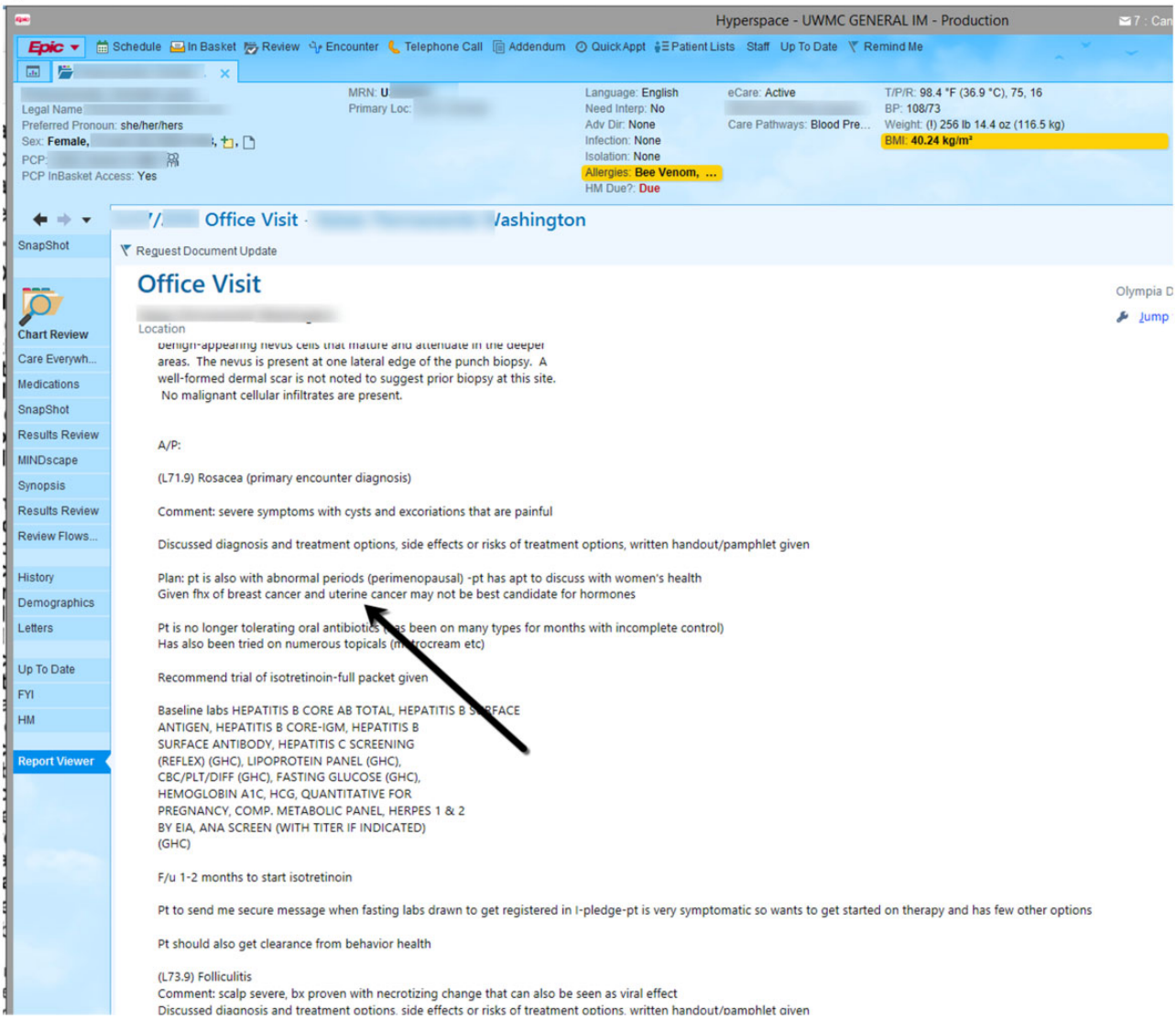

Figure 1. Continued

DISCUSSION

People at increased risk for heritable forms of breast and ovarian cancer benefit from referral to a cancer genetics professional. Comprehensive genetic counseling has been shown to reduce breast cancer–related worry and depression, increase patient understanding of risks, and reduce the intention for inappropriate genetic testing.[4] In this random sample of women selected from among those cared for in our health-care system, we were able to identify many women whose EHRs contain risk-factor information meeting national guidelines for further genetic risk evaluation, yet half of these women had no record of a referral in their EHR. Identifying these women did not require additional surveys, visits, or outreach efforts: the risk-factor information was already in their records, but it was dispersed in notes and other locations within the EHR, so current treating physicians may not have been aware of it. Finding this information took trained reviewers far more time than most busy clinicians can reasonably devote to a chart review. However, had the scattered risk factors for each patient been presented together to a treating provider with knowledge of NCCN guidelines, more women might have been referred to a medical geneticist or genetic counselor, and might have engaged appropriately in a discussion of the risks and benefits of genetic testing. For this reason, the dispersion of risk-factor information in the record may pose a barrier to recognizing an enhanced risk for cancer. It is a missed opportunity amenable to improvement by better methods to search and summarize EHR risk-factor information and prompt patients and physicians for additional information that could result in the identification of women who would benefit from an appropriate referral to a cancer genetics clinic for counseling and testing.

We chose to focus on hereditary breast ovarian cancer risk for 3 reasons: it is common,[9] there are widely accepted guidelines supporting genetic evaluation,[7] and diagnosis is linked to recommendations for treatment, such as the availability of poly-ADP ribose [PARP] inhibitors for the treatment of ovarian cancer with a BRCA1 or 2 pathogenic variant.[10] The same case could be made for

C

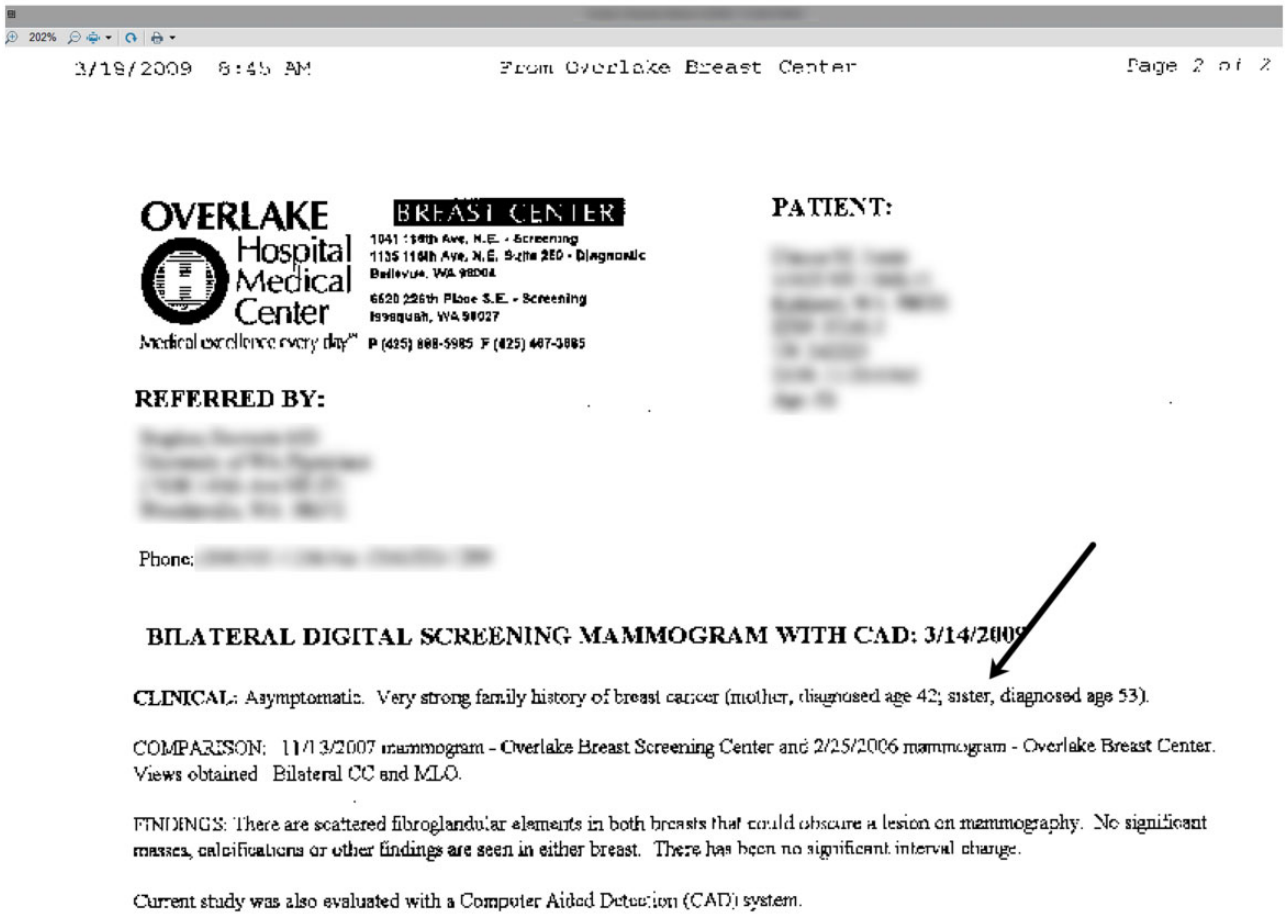

Figure 1. Continued

hereditary colorectal cancer. For other types of cancer, the evidence for a germline genetic cause is low (lung cancer) or the hereditary forms are rare and do not commonly affect treatment (sarcoma, clear cell renal carcinoma), so the benefits of this approach on health outcomes for these cancers are not clear, but may change either because of new genetic discoveries or because of changes in treatment recommendations.

Over the last decade, most US medical records have switched from paper to electronic form,[11] resulting in an enormous corpus of medical information in machine-readable form that did not exist when clinical cancer genetic testing was first available.

To our knowledge, this is the first report showing the amount of cancer risk-factor information in the EHR and seeking to determine its potential impact on patient referrals. Other work has focused on information in family history tools, including the use of natural language processing to find information in comments,[12] but this did not consider information in outside records, scanned documents, or in other parts of the record, nor other risks, such as Ashkenazi Jewish ancestry. As others have noted, primary care providers are overwhelmed with other clinical priorities that prevent the systematic documentation and use of family health history tools.[13,14] If providers gather risk-factor information in the course of care—and often they do not[15]—it may be entered in a fashion most expeditious for the pace of clinical practice; specialized structured data capture tools may not be used because of the time required to use them.[16]

There have been numerous efforts to improve and centralize the capture of family history information in the EHR, using both structured (checkbox, grid and other) and unstructured methods.[13,17,18] Though structured methods are often available to enter family history data, many providers use unstructured methods, such as narrative text, and the data entry task usually falls to the physician.[19] The focus of most of these methods is family history, which is an important risk factor, but other risk factors, such as personal cancer history and ancestry, may be recorded elsewhere in the EHR outside of a dedicated family history tool.

A limitation of our study is that we focused on referrals, and not completed counseling. Though referrals are the single most important factor,[5] there are other reasons women do not receive counseling and appropriate testing for pathogenic variants that increase the risk for cancer.

Methods exist to find this scattered risk-factor information automatically, without requiring the manual review of hundreds of pages, but are not yet a feature of most commercial EHRs. Our findings suggest that a different approach may be helpful: gathering risk-factor information from the wide range of locations and formats in which it is recorded may increase the number of women at risk who will be identified. Our results should lend impetus to apply these methods, which include image[20] and natural language processing,[21] to finding important, actionable information dispersed within the records.

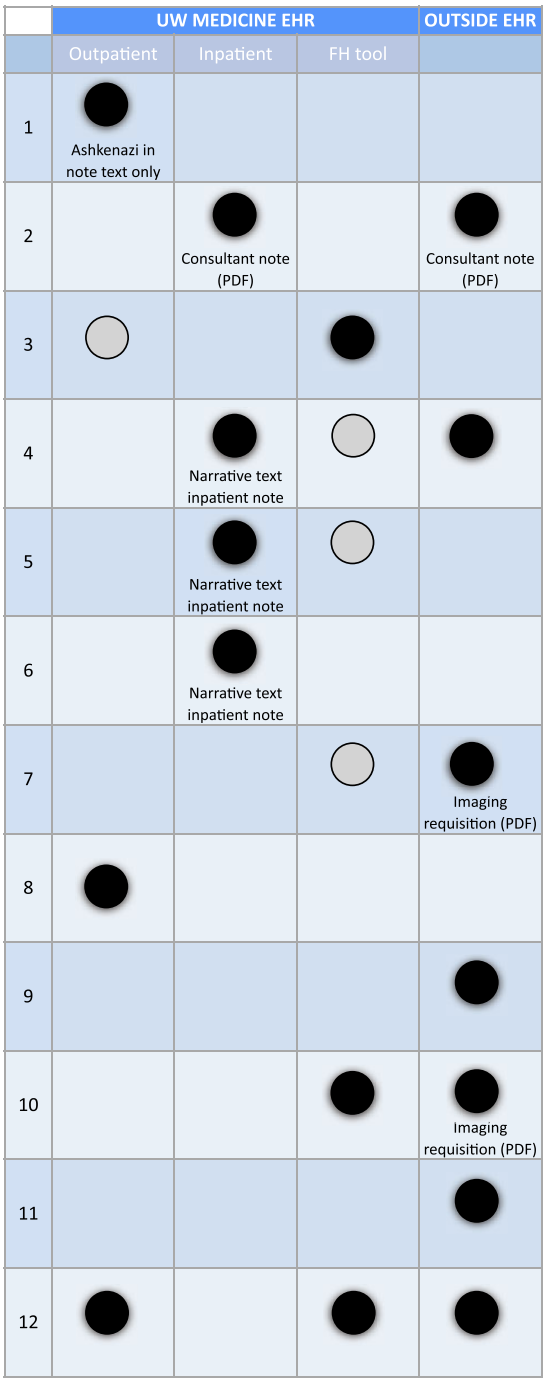

Figure 2. Location of risk-factor information within UW Medicine and outside EHR. Each row represents a patient whose record contained information meeting the NCCN 3.2019 criteria for referral to a cancer genetics professional but who had no record of a referral. The black circles indicate the source contained information meeting the NCCN 3.2019 criteria for referral. The gray circles indicate risk-factor information not meeting the NCCN criteria, such as a family history of cancer without a mention of age. EHR: electronic health record; FH: Family history; NCCN: National Comprehensive Cancer Network; UW: University of Washington.

CONCLUSIONS

We found that half of women whose EHR record contained risk-factor information meeting criteria for further genetic risk evaluation for heritable forms of breast and ovarian cancer were not referred, and that this risk-factor information was often dispersed in the EHR in locations other than the family history tools designed to collect it. If this were gathered and presented, it could lead to discussions between women and their providers that could lead to testing that might avert new incidence, morbidity, and mortality from heritable breast and ovarian cancer.

FUNDING

This work was supported by a Brotman Baty Institute for Precision Medicine Catalytic Collaborations award. REDcap is supported by UL1 TR002319, KL2 TR002317, and TL1 TR002318 from National Center for Advancing Translational Sciences/National Institutes of Health.

AUTHOR CONTRIBUTORS

All authors made substantial contributions to the conception or design of the work; contributed to the acquisition, analysis, or interpretation of data for the work; participated in drafting the work and revising it critically for important intellectual content; have finally approved the version to be published; and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICT OF INTEREST

None declared.

REFERENCES

1. National Cancer Institute Surveillance, Epidemiology, and End Results Program. Cancer stat facts. https://seer.cancer.gov/statfacts/html/breast.html Accessed 20 January 2020.
2. Breastcancer.org. US breast cancer statistics. http://www.breastcancer.org/symptoms/understand_bc/statistics Accessed 18 December 2019.
3. Chun DS, Berse B, Venne VL, *et al*. BRCA testing within the Department of Veteran Affairs: concordance with clinical guidelines. *Fam Cancer* 2017; 16 (1): 41–9.
4. Katz SJ, Ward KC, Hamilton AS, *et al*. Gaps in receipt of clinically indicated genetic counseling after diagnosis of breast cancer. *J Clin Oncol* 2018; 36 (12): 1218–24.
5. Swink A, Nair A, Hoof P, *et al*. Barriers to the utilization of genetic testing and genetic counseling in patients with suspected hereditary breast and ovarian cancers. *Proc (Bayl Univ Med Cent)* 2019; 32 (3):340–4.
6. Nelson HD, Pappas M, Cantor A, Haney E, Holmes R. Risk assessment, genetic counseling, and genetic testing for BRCA related cancer in women. USPSTF evidence report. *JAMA* 2019; 322 (7): 666–85.
7. National Cancer Comprehensive Network. Genetic/familial high-risk assessment: breast and ovarian (version 3.2019). Accessed 19 January 2019.
8. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; 42 (2): 377–81.
9. Manickam K, Buchanan AH, Schwartz MLB, *et al*. Exome sequencing-based screening for BRCA1/2 expected pathogenic variants among adult biobank participants. *JAMA Netw Open* 2018; 1 (5): e182140.
10. George A, Kaye S, Banerjee S. Delivering widespread BRCA testing and PARP inhibition to patients with ovarian cancer. *Nat Rev Clin Oncol* 2017; 14 (5): 284–96.
11. Blumenthal D. Launching HITECH. *N Engl J Med* 2010; 362 (5): 382–5.

12. Mowery DL, Kawamoto K, Bradshaw R, *et al*. Determining onset for familial breast and colorectal cancer from family history comments in the electronic health record. *AMIA Jt Summits Transl Sci Proc* 2019; 2019: 173–81.
13. Del Fiol G, Kohlmann W, Bradshaw RL, *et al*. Standards-based clinical decision support platform to manage patients who meet guideline-based criteria for genetic evaluation of familial cancer. *JCO Clin Cancer Inform* 2020; 4 (4): 1–9.
14. Guttmacher AE, Collins FS, Carmona RH. The family history–more important than ever. *N Engl J Med* 2004; 351 (22): 2333–6.
15. Scheuner MT, Sieverding P, Shekelle PG. Delivery of genomic medicine for common chronic adult diseases. A systematic review. *JAMA* 2008; 299 (11): 1320–34.
16. Polubriaginof F, Tatonetti NP, Vawdrey DK. An assessment of family history information captured in an electronic health record. *AMIA Annu Symp Proc* 2015; 2015: 2035–42.
17. Wilson BJ, Carroll JC, Allanson J, *et al*. Family history tools in primary care: does one size fit all? *Public Health Genom* 2012; 15 (3-4): 181–8.
18. Rich EC, Burke W, Heaton CJ, *et al*. Reconsidering the family history in primary care. *J Gen Intern Med* 2004; 19 (3): 273–80.
19. Scheuner MT, de Vries H, Kim B, Meili RC, Olmstead SH, Teleki S. Are electronic health records ready for genomic medicine? *Genet Med* 2009; 11 (7): 510–7.
20. Li X, Hu G, Teng X, Xie G. Building structured personal health records from photographs of printed medical records. MIA Annu Symp Proc 2015; 2015: 833–42.
21. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. *J Am Med Inform Assoc* 2011; 18 (5): 544–51.