ASCO special article

# Clinical Molecular Marker Testing Data Capture to Promote Precision Medicine Research Within the Cancer Research Network

Andrea N. Burnett-Hartman, MPH, PhD[1]; Natalia Udaltsova, PhD[2]; Lawrence H. Kushi, ScD[2]; Christine Neslund-Dudas, PhD[3]; Alanna Kulchak Rahm, PhD[4]; Pamala A. Pawloski, PharmD[5]; Douglas A. Corley, MD, PhD, MPH[2]; Sarah Knerr, PhD[6]; Heather Spencer Feigelson, MPH, PhD[1]; Jessica Ezzell Hunter, MS, PhD[7]; David C. Tabano, PhD[1]; Mara M. Epstein, ScD[8]; Stacey A. Honda, MD, PhD[9]; Monica Ter-Minassian, ScD[10]; Julie A. Lynch, PhD, RN, MBA[11]; and Christine Y. Lu, PhD, MSc[12]

abstract

**PURPOSE** To evaluate health care systems for the availability of population-level data on the frequency of use and results of clinical molecular marker tests to inform precision cancer care.

**METHODS** We assessed cancer-related molecular marker test data availability across 12 US health care systems in the Cancer Research Network. Overall, these systems provide care to a diverse population of more than 12 million people in the United States. We performed qualitative analyses of test data availability for five blood-based protein, nine germline, and 14 tissue-based tumor marker tests in each health care system's electronic health record and tumor registry using key informants, test code lists, and manual review of data types and output. We then performed quantitative analyses to estimate the proportion of patients with cancer with test utilization data and results for specific molecular marker tests.

**RESULTS** Health systems were able to systematically capture population-level data on all five blood protein markers, six of 14 tissue-based tumor markers, and none of the nine germline markers. Successful, systematic data capture was achievable for tests with electronic data feeds for test results (blood protein markers) or through prior manual abstraction by tumor registrars (select tumor-based markers). For test results stored in scanned image files (particularly germline and tumor marker tests), information on which test was performed and test results was not readily accessible in an electronic format.

**CONCLUSION** Even in health care systems with sophisticated electronic health records, there were few codified data elements available for evaluating precision cancer medicine test use and results at the population level. Health care organizations should establish standards for electronic reporting of precision medicine tests to expedite cancer research and facilitate the implementation of precision medicine approaches.



## INTRODUCTION

Recent advances in precision medicine have resulted in national guidelines that require the implementation of precision approaches in clinical care.[1,2] In cancer, precision medicine uses specific characteristics of patients and their tumors to develop tailored care plans and treatment regimens. Applications of precision medicine technologies include using blood-based protein markers to predict patient prognosis and monitor patients for recurrence,[3-6] identifying patients who are at a high risk for cancer using germline genetic tests,[7-9] and analyzing molecular markers in tumor tissue to characterize patients' risk of recurrence and tumor susceptibility to specific targeted therapies.[10-13]

Precision medicine has enormous potential to improve cancer care. However, there are many unanswered questions about the impact of precision medicine on survival, quality of life, patient costs, health system costs, and health disparities in real-world clinical settings (as opposed to in clinical trials). The population-level value, including long-term outcomes, of some precision medicine technologies, including those that are now standard of care, remains poorly understood.[14] These evidence gaps make it difficult to generate clinical practice guidelines and secure reimbursement for many molecular assays and tests,[15] and there is a need to incorporate real-world evidence to assess the effectiveness, value, and equitable use of these tests in clinical settings.[16-19]

There are also challenges to implementing recommended precision medicine technologies in clinical settings and evaluating relevant outcomes.[20] Specifically,

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

Accepted on July 22, 2019 and published at ascopubs.org/journal/cci on September 5, 2019: DOI https://doi.org/10.1200/CCI.19.00026

although there are clinical guidelines from professional societies about precision medicine,[21-23] many health care systems have not widely or comprehensively adopted standardized protocols governing when and how to use recommended genetic, molecular, and protein blood marker tests. Importantly, most new cancer therapies focus on specific aberrant molecular pathways. Given their narrow applications and expense, the use of novel targeted therapies should be guided by molecular tests to determine which patients are most likely to benefit.[24-26] Inappropriate use of targeted therapies in patients who have not undergone required or recommended testing or who do not have the correct therapeutic target(s) has been previously demonstrated.[27,28] Thus, it is important for health care systems to be able to evaluate the use and results of molecular assays, genetic testing, and targeted therapies, along with their impacts on patient outcomes.

Access to accurate, reliable, and longitudinal data about biomarker and genetic test use and results in large populations across diverse health care systems is required to gain a better understanding of the benefits, risks, costs, and value of precision medicine approaches in cancer care.[29] Here, we describe the current availability of population-level health system data on the use of, and results for, clinical tests of blood-based protein markers, germline markers, and tissue-based tumor markers across health systems in one of the largest cancer research collaborations in the United States, the Cancer Research Network (CRN).[30,31] We also outline the challenges of capturing population-level data on molecular marker testing in clinical practice and provide recommendations for increasing capacity to evaluate and improve the use of precision medicine in cancer care.

## METHODS

This analysis was conducted under data infrastructure activities within the CRN, which was approved by the primary CRN Institutional Review Board at Kaiser Permanente Northern California with ceding of institutional review board oversight of these activities from the other participating sites. We performed qualitative analyses of test data availability using key informants who were specialty-specific experts in the fields of pathology, oncology, and health informatics. We also assessed test code lists and performed a manual review of data types and output. We then performed quantitative analyses to estimate the number and proportion of patients with cancer that have test utilization data and results available for specific molecular marker tests, as described later in Methods.

### Population and Setting

The analysis involved 12 geographically distributed US health care organizations that provide care to a diverse population of more than 12 million people (Fig 1). All organizations are members of the CRN[30,31] and are part of the larger Health Care Systems Research Network, an established consortium of 20 research centers affiliated with large, integrated health care systems.[32] The CRN was established in 1998 to facilitate multisite collaborative research on cancer prevention, screening, treatment, survival, and palliation in diverse populations. The CRN comprises the National Cancer Institute and nonprofit research centers affiliated with these integrated health care delivery systems. These systems include both enrolled health plan members and aligned patients. Members are enrolled through employer-sponsored insurance, individual

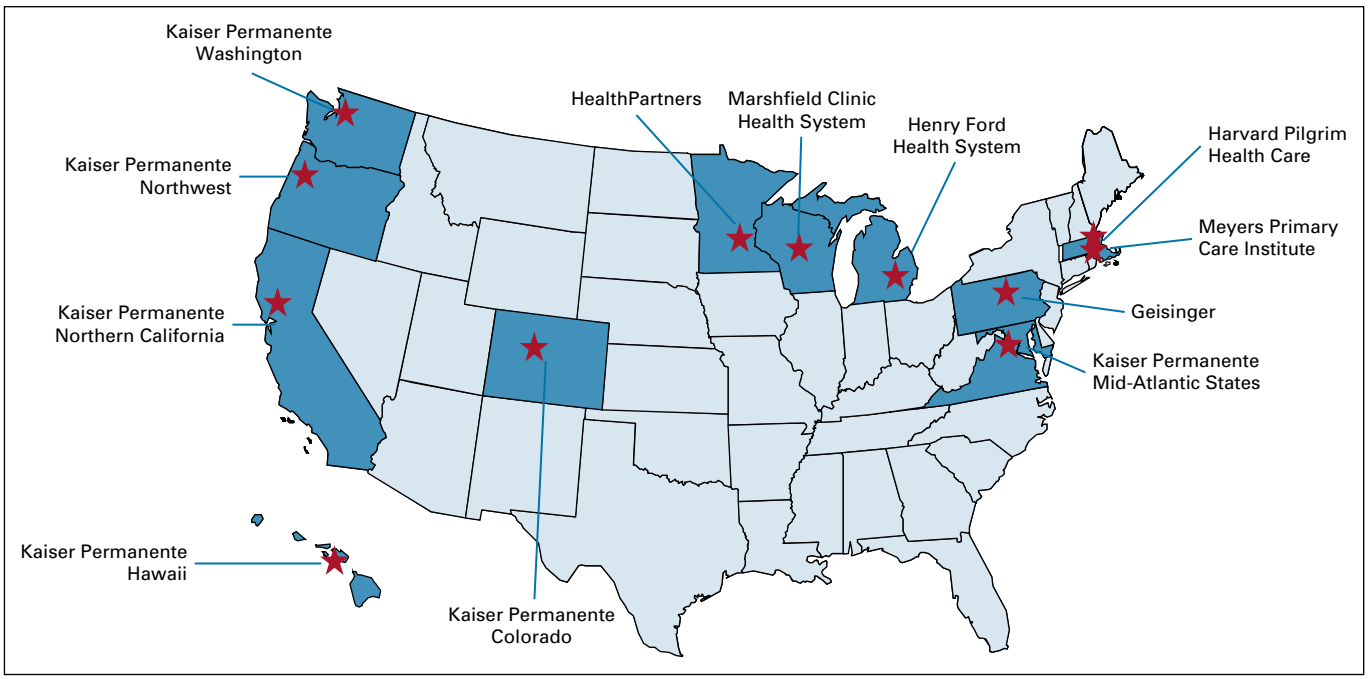


**FIG 1.** Cancer Research Network sites contributing data to assess capture of molecular marker information in electronic health records.

 

insurance plans, and capitated Medicare and Medicaid programs. Aligned patients use system medical group providers. Members served by these systems are generally representative of each system's geographic service area.[33]

### Data Sources

Data were obtained from the Health Care Systems Research Network Virtual Data Warehouse (VDW) from January 1, 2005, through December 31, 2015. The VDW is a series of data standards and automated processes that guide the generation of similarly constructed data tables at each organization to facilitate multisite research projects.[34] Developed in collaboration with CRN members, the VDW includes information on demographics, health plan enrollment, pharmacy fills, diagnoses from inpatient and outpatient care, utilization of procedures, laboratory values, and vital signs. At each site, data from health system or health plan administrative and claims databases and electronic health records (EHRs) are extracted, transformed, and loaded into the VDW common data model format.

The VDW also includes tumor registry data consistent with the standards of the North American Association of Central Cancer Registries and the National Cancer Institute's SEER program. These data include information on all cancer types, including each patient's cancer site, sequence, diagnosis date, stage, and site-specific factors (data items on tumor characteristics, prognosis, predicted treatment response, and clinical significance), based on the American Joint Committee on Cancer Collaborative Stage Data Collection System,[35] and receipt of primary chemotherapy, immunotherapy, hormone therapy, surgical treatment, and radiotherapy. Data contained within the tumor registries were collected by certified tumor registrars who review all elements of each patient's medical record, including pathology reports, laboratory test results, imaging reports, infusion session summaries, surgical operation reports, physician orders, and clinician progress notes.

### Assessment of Molecular Marker Data Availability

We used a multipronged approach to identify data sources for blood-based protein, germline, and tissue-based tumor marker tests. For blood-based protein marker tests, we focused on cancer antigen (CA) 125, prostate-specific antigen, CA 19-9, carcinoembryonic antigen, and α-fetoprotein. For germline tests, we assessed data on tests that included *BRCA1*, *BRCA2*, *MLH1*, *PMS2*, *MSH2*, *MSH6*, *EPCAM*, *FAP*, and *MUTYH*. For tissue-based tumor marker tests, we evaluated the availability of data on tests of relevance to breast (estrogen receptor [ER], progesterone receptor [PR], human epidermal growth factor receptor 2 [HER2], and multigene signature panel results), colorectal (*BRAF*, *KRAS*, and microsatellite instability status), and lung (*ALK*, *PDL1*, *EGFR*, *ROS1*, *RET*, *KRAS*, and *MET*) cancers.

These markers were selected via a consensus process and consultation with specialty-specific experts based on each marker's importance for driving decisions about patient care or because they were the focus of a current or planned CRN research project. We identified these markers within the VDW data using a combination of methods, including the use of standard Current Procedural Terminology procedure codes, International Classification of Diseases (ICD) Ninth Revision and ICD-10 diagnostic codes, and site-specific factor variables within the tumor registries. Outside the VDW, we used custom health system–specific codes that were identified via text searches of health system diagnosis and procedure code descriptions.

### *KRAS* Mutation Testing in Colorectal Cancer

In an exploratory analysis of *KRAS* mutation testing in colorectal cancer tissue biopsies, we used Kaiser Permanente Colorado as a test site to investigate the use of text-mining approaches in pathology reports to identify molecular tests and results. We identified all colorectal cancers occurring from 2010 to 2015 at Kaiser Permanente Colorado using the tumor registry data and colorectal cancer–specific ICD for Oncology, Third Edition, diagnosis codes. In this cohort of colorectal cancers, we electronically tagged the following terms in patient pathology reports: *KRAS*, mutation, and molecular. We then assessed the sensitivity and specificity of this text mining approach to identify *KRAS* mutation testing against the gold standard assessment of *KRAS* testing utilization and results from the tumor registry.

### Evaluation of Site-Specific Factor Variables for Breast Cancer

Because breast cancer is one of the best-annotated cancers for site-specific factor variables in tumor registry data, we used breast cancer as a case example to demonstrate how standardized, systematic documentation allows for population-level assessment of molecular marker test utilization and results across time and health systems. Relevant site-specific factors collected in relation to breast cancer include ER, PR, and HER2 status and multigene expression profiles. Possible values for site-specific factor results include the following: "positive/elevated," "negative/normal," "borderline," "not applicable/no information collected," "ordered/not interpretable," "test ordered, no results," "test not done," and "unknown or no information." Across 12 sites, we identified all individuals with a breast cancer diagnosis (C50.x based on ICD for Oncology, Third Edition, codes) between January 1, 2005, and December 31, 2015. We summed the number of individuals with available site-specific factor test results defined as the presence of structured results data (ie, positive, negative, normal, borderline) or the presence of a numeric value for a risk score associated with multigene expression profiles. A structured data reporting test not done, results

not in the chart, unknown, or not interpretable were considered unavailable test results. We calculated the proportion of available site-specific factor test results in the VDW tumor registry among identified patients with breast cancer by site and year.

## RESULTS

The ability of each CRN health system to electronically capture data on the use and results of clinical biomarker and genetic tests varied according to the type of test and the way in which the test results were stored in the VDW and EHR. Here, we detail the availability of biomarker and genetic testing data in the EHR at CRN sites by type of test. This information is also summarized in Table 1.

### Blood-Based Protein Markers

Information on blood-based protein markers resided in the following three locations: VDW tumor registry, VDW laboratory table, and the EHR's laboratory management system. Carcinoembryonic antigen, α-fetoprotein, and CA 19-9 test utilization and test values were available for all patients with specific GI cancers as site-specific factor variables in the VDW tumor registry. However, the VDW did not capture results of these tests for patients without cancer (Table 1). CA-125 and prostate-specific antigen test utilization and results were documented as site-specific factor variables for ovarian cancer and prostate cancer, respectively, in the VDW tumor registry. These tests were also standardly captured in the VDW laboratory table at most CRN sites,

**TABLE 1.** Clinical Biomarker and Genetic Marker Data Assessed in the Cancer Research Network

| Molecular Marker | Marker Type | Cancer Site | Data Source |
|---|---|---|---|
| CA-125 | Blood-based protein | Ovarian | VDW laboratory table; VDW tumor registry; EHR laboratory management system |
| PSA | Blood-based protein | Prostate | VDW laboratory table; VDW tumor registry; EHR laboratory management system |
| CA 19-9 | Blood-based protein | Pancreatic, gallbladder, bile duct, and gastric | VDW tumor registry; EHR laboratory management system |
| CEA | Blood-based protein | Multiple | VDW tumor registry; EHR laboratory management system |
| AFP | Blood-based protein | Liver | VDW tumor registry; EHR laboratory management system |
| *BRCA1* and *BRCA2* | Germline | Hereditary breast and ovarian cancer syndrome, breast, ovarian, prostate, pancreatic, melanoma | VDW procedure table; VDW diagnosis table; EHR scanned image reports |
| *MLH1*, *PMS2*, *MSH2*, *MSH6*, *EPCAM* | Germline | Lynch syndrome (hereditary nonpolyposis colorectal cancer), colorectal and other GI sites, endometrial, ovarian, brain, pancreatic, and skin | EHR scanned image reports |
| *FAP*, *MUTYH* | Germline | Familial adenomatous polyposis, colorectal | EHR scanned image reports |
| Estrogen receptor status, progesterone receptor status, HER2, multigene signature panel | Tissue-based tumor marker | Breast | VDW tumor registry; EHR scanned image reports; pathology reports |
| MSI status* | Tissue-based tumor marker | Colorectal | VDW tumor registry; EHR scanned image reports; pathology reports |
| *KRAS* | Tissue-based tumor marker | Colorectal | VDW tumor registry; EHR scanned image reports; pathology reports |
| *BRAF* | Tissue-based tumor marker | Colorectal, melanoma | EHR scanned image reports; pathology reports |
| *ALK*, *PDL1*, *EGFR*, *ROS1*, *RET*, *KRAS*, *MET* | Tissue-based tumor marker | Lung | EHR scanned image reports; pathology reports |

Abbreviations: AFP, α-fetoprotein; CA, cancer antigen; CEA, carcinoembryonic antigen; EHR, electronic health record; HER2, human epidermal growth factor receptor 2; MSI, microsatellite instability; PSA, prostate-specific antigen; VDW, Virtual Data Warehouse.

*Includes immunohistochemistry and DNA tests.

 

which made it possible to systematically analyze test utilization and results in the entire health system membership across CRN sites for these markers. Finally, we found that blood-based protein marker data were also stored in the laboratory management system of the EHR. These data elements were transferred directly from testing laboratories electronically, and the results of tests were captured in standard data fields within the EHR.

### Germline Genetic Markers

In each CRN health system, the results of germline tests were stored as scanned image reports in the EHR. There were no separate electronic data feeds from testing laboratories for test results. Some germline tests were noted as text in patients' clinic visit notes, but the results were not consistently or standardly documented. Gene-specific procedure codes were not used for either the individual gene or panel test results for *MLH1*, *PMS2*, *MSH2*, *MSH6*, *EPCAM*, *FAP*, and *MUTYH*. However, for *BRCA1* and *BRCA2* testing, standard procedure codes were first used in health systems in 2013. In some health systems, custom health system–generated procedure codes were used to identify specific genetic tests. However, even when we used the standard procedure codes and the custom health system–generated procedures codes, population-level data on use of genetic testing was not captured well by procedure codes.

### Tissue-Based Tumor Markers

Like germline tests, the results of tumor marker tests were stored in the EHR as scanned image reports at each health system. Results of tumor marker tests were inconsistently reported in the text of pathology reports, and these lacked standard documentation formats. Often, the pathology report referenced the scanned image results as an addendum, without describing the type of test performed or the results of the test. Neither the use nor results of tissue-based tumor marker tests could be captured in VDW procedure or diagnoses tables through standard codes.

Among the 14 tissue-based tumor markers that we assessed, the use and results of six markers were systematically captured in the VDW tumor registry as site-specific factor variables. These included all of the breast cancer tumor markers that we evaluated (ER, PR, HER2, and multigene signature panels) and microsatellite instability status and *KRAS* mutation status for colorectal cancer. However, *BRAF* mutation status in colorectal cancer and the lung cancer tumor markers (*ALK*, *PDL1*, *EGFR*, *ROS1*, *RET*, *KRAS*, and *MET*) were not in the VDW tumor registry as site-specific factor variables, likely because they were not required fields for collaborative staging at the time of this analysis.

### *KRAS* Mutation Testing in Colorectal Cancer

Exploratory analysis to use text-mining approaches in pathology reports to identify patients with colorectal cancer who received *KRAS* mutation tissue testing included a cohort of 1,567 patients with colorectal cancer diagnosed from 2010 to 2016 at Kaiser Permanente Colorado. Among 223 patients with *KRAS* testing confirmed via tumor registry data, text mining was able to identify 143 patients as having had a *KRAS* test (67% sensitivity). Among 1,344 patients with colorectal cancer who did not have *KRAS* testing, text mining was able to correctly classify 1,270 patients as not having *KRAS* testing (94% specificity).

### Site-Specific Factors for Breast Cancer

All health systems had standardized, structured documentation of somatic tumor marker testing and results for breast cancer in the tumor registry. Table 2 lists the number and yearly proportion of patients with breast cancer with test results available in the VDW tumor registry for ER, PR, and HER2 status and for multigene signature scores across all sites from 2005 through 2015. For ER and PR, all participating sites had high proportions of patients with available results from 2005 through 2015, and eight of 12 sites had relatively stable rates over time. HER2 and multigene results were available from 2010 onward for all sites.

Most health systems had ER status test results available as electronic data elements for more than 80% of patients with breast cancer between 2005 and 2015 (Fig 2). In contrast, data capture for the multigene signature panel began in 2010, and there was significant variation in the availability of test results for this panel over time and by health system (Fig 3).

## DISCUSSION

Among the markers investigated in this study, health systems within the CRN were able to capture population-level data from standard fields in the EHR or tumor registry on all five blood protein markers, six of 14 tumor-based markers, and none of the nine germline markers. The VDW is one of the most influential common data models and one of the most distributed health data networks in the United States for health care research among millions of patients for more than 20 years, and it preceded and served as a model for the US Food and Drug Administration's Sentinel and PCORnet data models.[36] But even with the VDW's depth and breadth of data capture, our results indicate that currently health systems in the CRN are unable to efficiently capture molecular testing use or results at the population level for the majority of cancer types. On a patient-by-patient basis, the scanned image of the test can be opened and read and is thus usable for clinical care. However, unlike other laboratory test results, these data are unable to be accessed and pooled into an electronic data file across hundreds or thousands of patients. This provides enormous challenges for evaluating the use of precision medicine approaches in real-world clinical settings.

Breast cancer was the only cancer type for which we were able to assess multiple molecular marker tests across time

**TABLE 2.** Breast Cancer Site–Specific Factor Data Availability in the Cancer Research Network From 2005 to 2015

| Year | Total No. of Patients With Breast Cancer | No. of Patients (%) | | | | |
|---|---|---|---|---|---|---|
| | | Estrogen Receptor Test Results Available | Progesterone Receptor Test Results Available | HER2 Test Results Available | Results Available From Combined Information From Estrogen Receptor, Progesterone Receptor, and HER2 Tests | Multigene Score Test Results Available |
| 2005 | 7,500 | 6,619 (99) | 6,572 (88) | 24 (< 1) | 21 (< 1) | 0 (0) |
| 2006 | 7,785 | 6,947 (99) | 6,902 (89) | 44 (1) | 35 (< 1) | 2 (< 1) |
| 2007 | 8,062 | 7,312 (99) | 7,282 (90) | 58 (1) | 45 (1) | 6 (< 1) |
| 2008 | 8,075 | 7,258 (99) | 7,222 (89) | 56 (1) | 45 (1) | 8 (< 1) |
| 2009 | 8,693 | 7,681 (99) | 7,627 (88) | 160 (2) | 181 (2) | 29 (< 1) |
| 2010 | 8,812 | 7,928 (99) | 7,890 (90) | 3,855 (44) | 4,255 (48) | 579 (7) |
| 2011 | 9,156 | 8,318 (99) | 8,275 (90) | 5,630 (61) | 4,801 (52) | 910 (10) |
| 2012 | 9,439 | 8,734 (99) | 8,672 (92) | 6,145 (65) | 5,362 (57) | 1,125 (12) |
| 2013 | 9,594 | 8,798 (99) | 8,732 (91) | 6,413 (67) | 5,794 (60) | 1,244 (13) |
| 2014 | 9,480 | 9,016 (99) | 8,957 (94) | 6,590 (70) | 5,951 (63) | 1,450 (15) |
| 2015 | 8,115 | 7,871 (99) | 7,827 (96) | 5,667 (70) | 5,413 (67) | 1,397 (17) |

NOTE. Table lists counts of patients with actual test results available, with test results defined as positive, negative, normal, or borderline, or numeric values.
Abbreviation: HER2, human epidermal growth factor receptor 2.

and across health systems. The data on breast cancer molecular markers were available because national cancer registry guidelines require standardized reporting of these markers, so cancer registrars in each health system manually abstract these data. Thus, successful, population-level data capture was only achievable for tests with electronic data feeds for test results (blood protein markers) or when the molecular marker was manually abstracted by health system tumor registrars (select tumor-based markers) from text-based results and converted into an electronic data element. For test results that were stored in scanned image files (particularly germline and tumor marker tests), information on which test was performed and the results of these tests were not readily accessible for population-level evaluation of precision medicine approaches in clinical settings.

Although health systems generally have care pathways and national guidelines available for patient care, deviations for these guidelines are commonly reported.[37-41] There are patient factors, including financial burden, that contribute to the care that a patient receives in the real world.[42,43] There are also provider factors, such as specialty and years in practice, that are associated with variations in care.[44,45] Characterizing the patient, provider, and health system factors that are associated with deviation from guideline-concordant care pathways for precision oncology is essential to ensuring high-quality, equitable care. To effectively

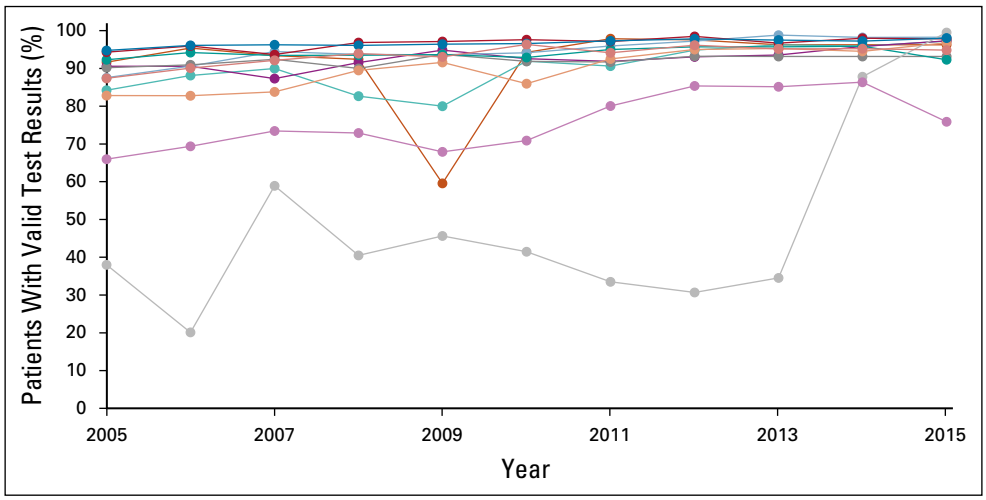


**FIG 2.** Percentage of patients with breast cancer with tumor estrogen receptor status results available in the Cancer Research Network Virtual Data Warehouse tumor registry by health system from 2005 to 2015.

 

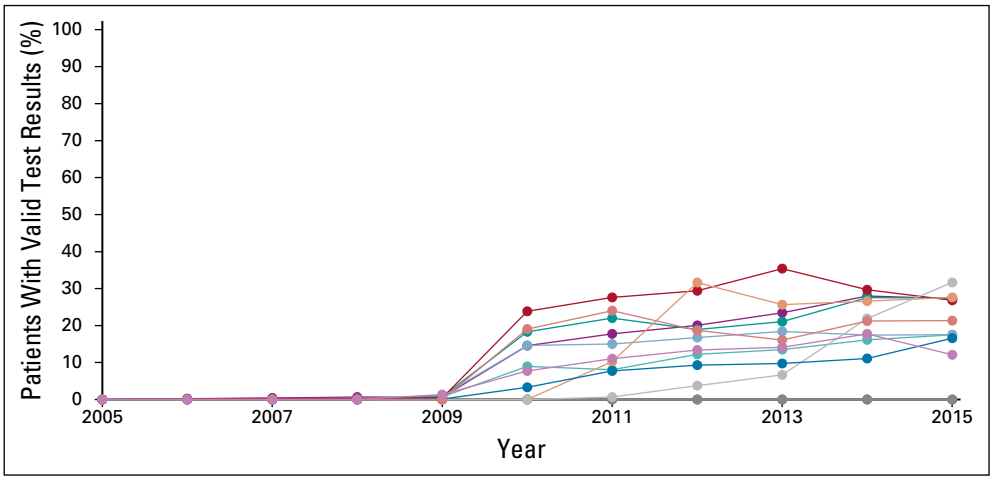


**FIG 3.** Percentage of patients with breast cancer with multigene signature results available in the Cancer Research Network Virtual Data Warehouse tumor registry by health system from 2005 to 2015.

evaluate precision medicine utilization, costs, and effectiveness within and across health systems, more comprehensive and efficient data capture of molecular assay and genetic testing results is needed.

Possible approaches to improve the population-level data capture of information needed to evaluate precision medicine in real-world clinical settings include the following: require that laboratories send electronic data to health care systems for genetic and other molecular marker test results that conform to Health Level Seven International (HL7) Fast Healthcare Interoperability Resources (FHIR) specifications for genomic data[46,47]; develop and use standard procedure codes for each molecular marker and genetic test; develop diagnosis codes for specific genetic variant carrier status; encourage EHR vendors to incorporate standardized tables to capture test status and results for these tests; and build infrastructure into common data models, such as the VDW, to incorporate tables that include structured data on molecular assay and genetic testing results. Implementing these approaches, which are not mutually exclusive, will require health systems to invest resources in data infrastructure, but these improvements will enhance the quality of care and allow health systems and researchers to identify ways to optimize the use of precision medicine approaches in clinical settings.

Several health systems, including Intermountain Healthcare and the US Department of Veterans Affairs, have begun to work with laboratories to obtain electronic data for genetic and other molecular marker testing on their members. In doing so, these health systems are able to evaluate and improve precision cancer care within their systems. Two studies at Intermountain Healthcare reported improved survival and similar or decreased costs associated with the use of tumor testing to identify targeted cancer therapies.[48,49] Within the US Department of Veterans Affairs, a recent study of patients with lung cancer evaluated the use of erlotinib, a drug that targets cancers with activating mutations in the epidermal growth factor receptor (*EGFR*) gene. This study found that 11% of patients whose lung cancer tested negative for *EGFR* mutations and 17% of patients with no *EGFR* test results were treated with erlotinib.[27] This was inappropriate because erlotinib use in patients with wild-type *EGFR* lung cancer is associated with poorer outcomes than platinum-based chemotherapy.[50,51] Thus, by obtaining and evaluating electronic data on molecular marker testing in their members, the Department of Veterans Affairs was able to identify a specific area for quality improvement initiatives that may lead to better patient outcomes.

To reduce inappropriate use of targeted therapies, preauthorization and provider attestation to the presence of the molecular target are often required by insurers, particularly those that operate outside of the integrated systems described in our article.[52] In addition, in 2018, the Centers for Medicare and Medicaid Services released a decision memo noting that next-generation sequencing as a diagnostic laboratory test is reasonable and necessary and is covered nationally for patients diagnosed with specific tumor types. However, this rule change did not come with new or additional procedure codes or diagnosis billing codes specific to gene mutations or other molecular markers of interest. Thus, codifiable requirements allowing for population-based identification of the use of tests or codifiable methods for describing test results are still lacking.

Some health systems have begun to track genetic and other molecular marker test use through system-generated modified procedure and diagnostic codes that are specific to a particular gene or genetic syndrome. This includes diagnostic codes for specific gene mutation carrier status. Unfortunately, system-generated codes are not used consistently and are not applied in a standardized manner across health systems. The result is missing utilization data

and the inability to conduct research across multiple health systems. This is important because any one system may not have sufficient power to examine the outcomes of precision medicine approaches. To improve data capture for these tests, unique standard procedure, diagnostic, and billing codes are necessary.[53] For some tests, gene-specific billing codes are available, and a recent study by Lynch et al[54] using Medicare claims data demonstrated that gene-specific billing codes facilitated population-level research in precision medicine. As new tests are introduced into clinical practice, it is important to implement new codes in a rapid fashion, rather than using a generic laboratory test code that lacks specification of the test used.

Additional data infrastructure is also needed to standardize the complex data elements that are reported in genetic and other molecular test results. Results from genetic and other molecular testing include more information and are reported in a different format than results from traditional laboratory tests (eg, cholesterol tests, WBC counts) for which EHR systems were designed to capture. Several recent studies have successfully used HL7 FHIR data standards to integrate electronic genomic data with clinical EHR data.[46,55] This integration was for the purpose of improving patient management, but using the HL7 FHIR data specifications also has the potential to facilitate the development and standardization of research-ready data that include the molecular marker data elements necessary for population-based research on the use of precision medicine in clinical practice.

Our assessment of the availability of population-level data to evaluate precision medicine approaches for cancer care in real-world clinical settings included 12 diverse health systems with sophisticated EHRs and multiple data sources. However, our findings should be considered with regard to several limitations. Most of the participating health systems use integrated care delivery models, so our results may not be generalizable to academic care settings or medical centers that are primarily fee-for-service models. In addition, there was significant variation in the depth and amount of data on molecular marker test use and the results available in standard fields between cancer types, and we were unable to assess guideline-concordant test or therapy use. Despite these limitations, our study highlights the current challenges systems face even in the environment of sophisticated EHR data to assess precision cancer care. In addition, we were able to identify specific areas for improvement in the way that molecular test use information and results are stored within the EHR.

Within the CRN health systems, several initiatives are now underway to improve population-level data capture of molecular marker testing and to better incorporate molecular marker data in to the VDW. Work is currently underway to include CA 19-9, carcinoembryonic antigen, and α-fetoprotein into the VDW laboratory table; this work will improve the electronic capture of these markers across the entire membership of CRN health systems. The Henry Ford Health System has begun to systematically capture the results of multimarker panel tumor tests in a discreet database that can be used for research to evaluate precision cancer care, and the Kaiser Permanente Center for Effectiveness and Safety Research is working with testing laboratories to obtain electronic data on molecular testing and results within Kaiser Permanente health plan members. These efforts will help to facilitate research aimed at evaluating and optimizing precision medicine in clinical practice and ultimately lead to better patient outcomes and improved efficiency by ensuring the implementation of guideline-based precision cancer care.


## AFFILIATIONS

[1]Institute for Health Research, Kaiser Permanente Colorado, Aurora, CO
[2]Kaiser Permanente Northern California, Oakland, CA
[3]Henry Ford Health System, Detroit, MI
[4]Genomic Medicine Institute, Geisinger, Danville, PA
[5]HealthPartners Institute, Bloomington, MN
[6]University of Washington and Kaiser Permanente Washington Health Research Institute, Seattle, WA
[7]Kaiser Permanente Northwest, Portland, OR
[8]University of Massachusetts Medical School, Worcester, MA
[9]Kaiser Permanente Hawaii, Honolulu, HI
[10]Kaiser Permanente Mid-Atlantic States, Rockville, MD
[11]Department of Veterans Affairs Salt Lake City Health System, Salt Lake City, UT
[12]Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA

## CORRESPONDING AUTHOR

Andrea N. Burnett-Hartman, Kaiser Permanente Colorado, 2550 S Parker Rd, Ste 200, Aurora, CO 80014; e-mail: andrea.n.burnett-hartman@kp.org.



## SUPPORT

Supported by National Cancer Institute, National Institutes of Health (NIH) Grant No. U24CA171524 (L.H.K.); an Ebert Career Development Award, Harvard Pilgrim Health Care Institute and Harvard Medical School (C.Y.L.); and National Center for Research Resources and the National Center for Advancing Translational Sciences, NIH, Grant No. KL2TR001454 (M.M.E.).


## AUTHOR CONTRIBUTIONS

**Conception and design:** Andrea N. Burnett-Hartman, Lawrence H. Kushi, Christine Neslund-Dudas, Pamala A. Pawloski, Douglas A. Corley, Jessica Ezzell Hunter, Stacey A. Honda, Julie A. Lynch, Christine Y. Lu

**Financial support:** Lawrence H. Kushi, Christine Neslund-Dudas, Douglas A. Corley

**Administrative support:** Lawrence H. Kushi

**Provision of study materials or patients:** Christine Neslund-Dudas, Pamala A. Pawloski, Douglas A. Corley, Monica Ter-Minassian

**Collection and assembly of data:** Andrea N. Burnett-Hartman, Natalia Udaltsova, Christine Neslund-Dudas, Alanna Kulchak Rahm, Douglas A. Corley, Jessica Ezzell Hunter, David C. Tabano, Mara M. Epstein, Monica Ter-Minassian, Christine Y. Lu

 

**Data analysis and interpretation:** Andrea N. Burnett-Hartman, Lawrence H. Kushi, Christine Neslund-Dudas, Alanna Kulchak Rahm, Pamala A. Pawloski, Douglas A. Corley, Sarah Knerr, Heather Spencer Feigelson, Jessica Ezzell Hunter, David C. Tabano, Mara M. Epstein, Stacey A. Honda, Monica Ter-Minassian, Julie A. Lynch, Christine Y. Lu
**Manuscript writing:** All authors
**Final approval of manuscript:** All authors
**Accountable for all aspects of the work:** All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/cci/author-center.

**Douglas A. Corley**
**Research Funding:** Pfizer (Inst)

**David C. Tabano**
**Employment:** Bristol-Myers Squibb

**Stacey A. Honda**
**Employment:** Hawaii Permanente Medical Group
**Stock and Other Ownership Interests:** Hawaii Permanente Medical Group

**Julie A. Lynch**
**Research Funding:** Genomic Health (Inst), AstraZeneca (Inst), Myriad Genetics (Inst), Boehringer Ingelheim (Inst), Astellas Pharma (Inst), CardioDx (Inst), Janssen (Inst)

No other potential conflicts of interest were reported.

## REFERENCES


1. Kalia M: Biomarkers for personalized oncology: Recent advances and future challenges. Metabolism 64:S16-S21, 2015 (suppl 1)
2. Borad MJ, LoRusso PM: Twenty-first century precision medicine in oncology: Genomic profiling in patients with cancer. Mayo Clin Proc 92:1583-1591, 2017
3. Vesely S, Jarolim L, Duskova K, et al: The use of early postoperative prostate-specific antigen to stratify risk in patients with positive surgical margins after radical prostatectomy. BMC Urol 14:79, 2014
4. Fakih MG, Padmanabhan A: CEA monitoring in colorectal cancer: What you should know. Oncology (Williston Park) 20:579-587, 2006
5. Loosen SH, Neumann UP, Trautwein C, et al: Current and future biomarkers for pancreatic adenocarcinoma. Tumour Biol 39: doi: 1010428317692231
6. Montagnana M, Danese E, Giudici S, et al: HE4 in ovarian cancer: From discovery to clinical application. Adv Clin Chem 55:1-20, 2011
7. Lynch HT, de la Chapelle A: Hereditary colorectal cancer. N Engl J Med 348:919-932, 2003
8. Bartosch C, Clarke B, Bosse T: Gynaecological neoplasms in common familial syndromes (Lynch and HBOC). Pathology 50:222-237, 2018
9. Shiovitz S, Korde LA: Genetics of breast cancer: A topic in evolution. Ann Oncol 26:1291-1299, 2015
10. Miyamoto Y, Suyama K, Baba H: Recent advances in targeting the EGFR signaling pathway for the treatment of metastatic colorectal cancer. Int J Mol Sci 18: E752, 2017
11. Győrffy B, Hatzis C, Sanft T, et al: Multigene prognostic tests in breast cancer: Past, present, future. Breast Cancer Res 17:11, 2015
12. Devji T, Levine O, Neupane B, et al: Systemic therapy for previously untreated advanced BRAF-mutated melanoma: A systematic review and network meta-analysis of randomized clinical trials. JAMA Oncol 3:366-373, 2017
13. Sacher AG, Gandhi L: Biomarkers for the clinical use of PD-1/PD-L1 inhibitors in non-small-cell lung cancer: A review. JAMA Oncol 2:1217-1222, 2016
14. Phillips KA, Deverka PA, Sox HC, et al: Making genomic medicine evidence-based and patient-centered: A structured review and landscape analysis of comparative effectiveness research. Genet Med 19:1081-1091, 2017
15. Tuckson RV, Newcomer L, De Sa JM: Accessing genomic medicine: Affordability, diffusion, and disparities. JAMA 309:1469-1470, 2013
16. Lu CY, Cohen JP: Can genomic medicine improve financial sustainability of health systems? Mol Diagn Ther 19:71-77, 2015
17. Salgado R, Solit DB, Rimm DL, et al: Addressing the dichotomy between individual and societal approaches to personalised medicine in oncology. Eur J Cancer 114:128-136, 2019
18. Fahr P, Buchanan J, Wordsworth S: A review of the challenges of using biomedical big data for economic evaluations of precision medicine. Appl Health Econ Health Policy 17:443-452, 2019
19. Agarwala V, Khozin S, Singal G, et al: Real-world evidence in support of precision medicine: Clinico-genomic cancer data as a case study. Health Aff (Millwood) 37:765-772, 2018
20. Roberts MC, Kennedy AE, Chambers DA, et al: The current state of implementation science in genomic medicine: Opportunities for improvement. Genet Med 19:858-863, 2017
21. Richards S, Aziz N, Bale S, et al: Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17:405-424, 2015
22. Li MM, Datto M, Duncavage EJ, et al: Standards and guidelines for the interpretation and reporting of sequence variants in cancer: A joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. J Mol Diagn 19:4-23, 2017
23. Birkeland ML, McClure JS: Optimizing the clinical utility of biomarkers in oncology: The NCCN Biomarkers Compendium. Arch Pathol Lab Med 139:608-611, 2015
24. Albaba H, Lim C, Leighl NB: Economic considerations in the use of novel targeted therapies for lung cancer: Review of current literature. Pharmacoeconomics 35:1195-1209, 2017
25. Maroun R, Fleury L, Nachbaur G, et al: Real-world costs and outcomes in metastatic renal cell carcinoma patients treated with targeted therapies: A cohort study from the French health insurance database. Curr Med Res Opin 33:1755-1762, 2017
26. Schwartzberg L, Kim ES, Liu D, et al: Precision oncology: Who, how, what, when, and when not? Am Soc Clin Oncol Educ Book 37:160-169, 2017
27. Lynch JA, Berse B, Chun D, et al: Epidermal growth factor receptor mutational testing and erlotinib treatment among veterans diagnosed with lung cancer in the United States Department of Veterans Affairs. Clin Lung Cancer 18:401-409, 2017
28. Efimova O, Berse B, Denhalter DW, et al: Clinical decisions surrounding genomic and proteomic testing among United States veterans treated for lung cancer within the Veterans Health Administration. BMC Med Inform Decis Mak 17:71, 2017

29. Chambers DA, Feero WG, Khoury MJ: Convergence of implementation science, precision medicine, and the learning health care system: A new model for biomedical research. JAMA 315:1941-1942, 2016
30. Wagner EH, Greene SM, Hart G, et al: Building a research consortium of large health systems: The Cancer Research Network. J Natl Cancer Inst Monogr 35:3-11, 2005
31. Hornbrook MC, Hart G, Ellis JL, et al: Building a virtual cancer research organization. J Natl Cancer Inst Monogr 35:12-25, 2005
32. Steiner JF, Paolino AR, Thompson EE, et al: Sustaining research networks: The twenty-year experience of the HMO Research Network. EGEMS (Wash DC) 2:1067, 2014
33. Nekhlyudov L, Greene SM, Chubak J, et al: Cancer research network: Using integrated healthcare delivery systems as platforms for cancer survivorship research. J Cancer Surviv 7:55-62, 2013
34. Ross TR, Ng D, Brown JS, et al: The HMO Research Network Virtual Data Warehouse: A public data model to support collaboration. EGEMS (Wash DC) 2:1049, 2014
35. American Joint Committee on Cancer: Collaborative Stage Data Collection System. https://cancerstaging.org/cstage/Pages/default.aspx
36. Weeks J, Pardee R: Learning to share health care data: A brief timeline of influential common data models and distributed health data networks in U.S. health care research. EGEMS (Wash DC) 7:4, 2019
37. Rocque GB, Williams CP, Kenzik KM, et al: Concordance with NCCN treatment guidelines: Relations with health care utilization, cost, and mortality in breast cancer patients with secondary metastasis. Cancer 124:4231-4240, 2018
38. Williams CP, Kenzik KM, Azuero A, et al: Impact of guideline-discordant treatment on cost and health care utilization in older adults with early-stage breast cancer. Oncologist 24:31-37, 2019
39. Kole AJ, Stahl JM, Park HS, et al: Predictors of nonadherence to NCCN guideline recommendations for the management of stage I anal canal cancer. J Natl Compr Canc Netw 15:355-362, 2017
40. Schwam ZG, Sosa JA, Roman S, et al: Receipt of care discordant with practice guidelines is associated with compromised overall survival in nasopharyngeal carcinoma. Clin Oncol (R Coll Radiol) 28:402-409, 2016
41. Nadpara PA, Madhavan SS, Tworek C, et al: Guideline-concordant lung cancer care and associated health outcomes among elderly patients in the United States. J Geriatr Oncol 6:101-110, 2015
42. Narang AK, Nicholas LH: Out-of-pocket spending and financial burden among Medicare beneficiaries with cancer. JAMA Oncol 3:757-765, 2017
43. Neugut AI, Subar M, Wilde ET, et al: Association between prescription co-payment amount and compliance with adjuvant hormonal therapy in women with early-stage breast cancer. J Clin Oncol 29:2534-2542, 2011
44. Kimmick GG, Camacho F, Mackley HB, et al: Individual, area, and provider characteristics associated with care received for stages I to III breast cancer in a multistate region of Appalachia. J Oncol Pract 11:e9-e18, 2015
45. Carpenter WR, Meyer AM, Wu Y, et al: Translating research into practice: The role of provider-based research networks in the diffusion of an evidence-based colon cancer treatment innovation. Med Care 50:737-748, 2012
46. Alterovitz G, Warner J, Zhang P, et al: SMART on FHIR genomics: Facilitating standardized clinico-genomic apps. J Am Med Inform Assoc 22:1173-1178, 2015
47. Mandel JC, Kreda DA, Mandl KD, et al: SMART on FHIR: A standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 23:899-908, 2016
48. Haslem DS, Van Norman SB, Fulde G, et al: A retrospective analysis of precision medicine outcomes in patients with advanced cancer reveals improved progression-free survival without increased health care costs. J Oncol Pract 13:e108-e119, 2017
49. Haslem DS, Chakravarty I, Fulde G, et al: Precision oncology in advanced cancer patients improves overall survival with lower weekly healthcare costs. Oncotarget 9:12316-12322, 2018
50. Garassino MC, Martelli O, Broggini M, et al: Erlotinib versus docetaxel as second-line treatment of patients with advanced non-small-cell lung cancer and wild-type EGFR tumours (TAILOR): A randomised controlled trial. Lancet Oncol 14:981-988, 2013
51. Lee CK, Brown C, Gralla RJ, et al: Impact of EGFR inhibitor in non-small cell lung cancer on progression-free and overall survival: A meta-analysis. J Natl Cancer Inst 105:595-605, 2013
52. Lu CY, Loomer S, Ceccarelli R, et al: Insurance coverage policies for pharmacogenomic and multi-gene testing for cancer. J Pers Med 8:E19, 2018
53. Liang SY, Phillips KA, Wang G, et al: Tradeoffs of using administrative claims and medical records to identify the use of personalized medicine for patients with breast cancer. Med Care 49:e1-e8, 2011
54. Lynch JA, Berse B, Dotson WD, et al: Utilization of genetic tests: Analysis of gene-specific billing in Medicare claims data. Genet Med 19:890-899, 2017
55. Warner JL, Rioth MJ, Mandl KD, et al: SMART precision cancer medicine: A FHIR-based app to provide genomic information at the point of care. J Am Med Inform Assoc 23:701-710, 2016