pcori

PCORI RCR FINAL REPORT

Date (mm/dd/yyyy): 09/17/2018	
Title of Project: Molecular Testing Rapid Cycle Project (RCP)	
Type of Progress Report: ☒ Status Report	
Period Covered by this Report: 06/13/2018 (mm/dd/yyyy) to 09/17/2018 (mm/dd/yyyy) *Record the dates of the reporting period per the fully executed Milestone Schedule.*	
Principal Investigator & Institution Updated Contact Information:	
PI First Name:	Russ
PI Last Name:	Waitman
Office Phone:	913-945-7087
Institution Legal Name:	University of Kansas Medical Center
Address (street, city, state, zip code):	3901 Rainbow Blvd, MS 3065, Kansas City, KS. 66160
Telephone:	913-945-7087
Study Coordinator & Institution Updated Contact Information:	
First Name:	Steve
Last Name:	Fennel
Office Phone:	913-588-9809
Institution Legal Name:	University of Kansas Medical Center
Address (street, city, state, zip code):	3901 Rainbow Blvd, MS 3065, Kansas City, KS. 66160
Telephone:	913-588-9809

MILESTONES UPDATE

Record each milestone label, name, description, and projected completion date (columns A-D), as shown in your fully executed Milestone Schedule. Complete Columns E, F, and G for milestones due or completed during the current reporting period.

Column E: Check appropriate box indicating milestone completion status during reporting period. Additional information on milestones that were not completed is required and should be provided in the section below this table.
Column F: Select actual date of milestone completion.
Column G: If applicable, select appropriate reason for delay/non-completion of projected milestone during the specified reporting period.

Column A	Column B	Column C	Column D	Column E	Column F	Column G
Milestone Label *(e.g., B-1, etc.)*	**Milestone Name**	**Description**	**Projected Completion Date**	**Completed? (Yes/No)**	**Date Completed**	**If Not Completed, Reason for Delay**
A	Effective Date 09/13/17	--	08/17/2018	--	--	--
B1	Protocol Complete	--	08/09/2017	☑Yes ☐No	08/08/2017	
B2	IRB Approval Obtained	Obtain IRB Approval for study and submit approval letter to PCORI.	11/13/2017	☑Yes ☐No	11/21/2017	
C1	Aim 1 Data Preparation	Identify cancer-relevant data sources, conduct detailed review of codes (diagnosis, procedure, labs, meds) and ETL cancer-relevant data to CDM	11/13/2017	☑Yes ☐No	01/24/2018	
C2	Aim 1 Query Execution	Preliminary and final Aim 1 queries executed	01/27/2018	☑Yes ☐No	05/05/2018	
C3	Aim 2 Query Development and Testing	Develop Aim 2 algorithms and query; test Aim 2 SAS code	12/13/2017	☑Yes ☐No	04/22/2018	
C4	Aim 2 Query 1 - Extract Cohort	Extract cohort with test results	01/13/2018	☑Yes ☐No	01/26/2018	
C5	Aim 3 Query Development and Testing	Develop and test Aim 3 SAS code	11/13/2017	☑Yes ☐No	08/13/2018	
C6	Aim 3 - Alignment of GROUSE Data Repository	Fixes to GROUSE CDM based on Aim 1 query	01/13/2018	☑Yes ☐No	08/06/2018	

--	**Report Submission**	**Submit Progress Report, Using Interim Progress Report Template**	**02/13/2018**	☑Yes ☐No	02/13/2018	
--	**PCORI Feedback - Go/No-Go**	**Project checkpoint based on final Aim 1 query**	**02/27/2018**	--	--	--
D1	Aim 1 Data Analysis	Complete analysis of Aim 1 Data	05/12/2018	☑Yes ☐No	06/01/2018	
D2	Aim 2 Query Execution	Retrieve treatment data from CDM	03/13/2018	☑Yes ☐No	05/01/2018	
D3	Aim 2 Data Analysis	Complete analysis of Aim 2 Data	05/12/2018	☑Yes ☐No	08/10/2018	
D4	Aim 3 Data Analysis	Build analysis dataset using linked claims data and complete analysis of Aim 3 data	05/12/2018	☑Yes ☐No	08/30/2018	
D5	Aim 4 Data Quality Evaluation	Evaluation of data quality (qualitative evaluation)	02/27/2018	☑Yes ☐No	09/04/2018	
D6	Aim 3 Data Validation	Build CONSORT-type diagram indicating inclusions and exclusions and examine data completeness by study site by year. Finalize cohort numbers and re-run test and treatment use statistics.	9/4/2018	☑Yes ☐No	09/04/2018	
D7	Aim 4 Share Results with DRN OC	Discuss Data Quality Evaluation results with DRN OC and discuss scalability of the pooled data analysis approach, strategies for supporting tumor data in the CDM, and potential targeted data source investigations to evaluate variation in CDM medication tables	8/28/2018	☑Yes ☐No	09/06/2018	
G	**Final Progress Report**	**Submit Final Progress Report, Using Final Progress Report Template**	**06/12/2018**	☑Yes ☐No	**09/17/2018**	
H	**Project Period End Date**	**--**	**06/12/2018**	☑Yes ☐No	**09/17/2018**	

Table of Contents

ACCOMPLISHMENTS, CHALLENGES AND NEXT STEPS

Executive Summary

PROJECT OBJECTIVE: To document the patterns of use of molecular biomarkers and molecular-guided cancer therapies for patients with solid tumors in real-world settings and to test and expand the capacity of PCORnet to describe the details of cancer pathology, results of testing, cancer treatment and outcomes.

STUDY POPULATION: Patients with a single primary solid tumor diagnosed during 2013-2016 from 11 PCORnet member sites within three Clinical Data Research Networks (CDRNs). PCORnet Distributed Research Network Operations Center (DRN OC) additionally distributed descriptive queries to all PCORnet sites.

STUDY DESIGN: Retrospective database study with chart review for a sample of patients

METHODS: Oral and infused molecular-guided for which corresponding molecular tests were required in FDA indications labeling were included. Tests included those used to identify ALK fusion, BRAFV660E/V600K, EGFR Exon 19 deletion, $EGFR^{L858R}$, $EGFR^{T790M}$, ERBB2 amplification, KRAS/NRAS/HRAS, MSI/MMR, PD-1, and BRCA1/2. The pooled dataset consisted of tumor registry data from each site and linked Common Data Model (CDM) diagnosis, procedures, prescribing, dispensing, medication administration, and death data tables. For the pooled dataset, tests and treatments were described by patient and tumor characteristics and study site. In addition, three focused studies were conducted. A chart review study of patients with advanced colorectal cancer at 2 participating sites examined whether molecular-guided therapy use was in accordance with test results and described completeness of the electronic medical record (EMR) data for identifying in and out-of-system tests and treatments and described completeness of CDM for identifying tests and treatments. A record linkage study of 7 sites linked CDM data with Medicare claims to examine completeness of CDM data and how much claims data add to the ability of the CDM to capture tests, treatments and outcomes. A third study focused on advanced non-small cell lung cancer patients treated with immune checkpoint (PD-L1) inhibitors in order to examine overall survival and other real-world endpoints in the 11 participating PCORnet sites.

FINDINGS: The dataset included linked registry billing and EMR data for 86,154 patients. There was variability among molecular tests for how easily they were found with queries of CPT codes in the CDM. Those molecular markers that are primarily interrogated via Fluorescence in situ hybridization (FISH) or immunohistochemistry (IHC) anatomic pathology tests have only a nonspecific CPT code and are particularly under-reported whereas molecular pathology tests have CPT codes that identify the specific gene being examined and the CDM captures more of these tests. Molecular-guided therapies that were prescribed or administered in the system were well-captured across 4 distinct CDM tables. Testing rates varied across cancer types, were most common with stage IV disease, and varied across study sites (from 36% to 57%). The molecular-targeted therapies were received by 5% of patients and use rates were lowest among the oldest patients and highest for patients with breast, esophagus, oral cavity and pharynx, lung and bronchus, and melanoma skin cancers. When a sample of patients with metastatic colorectal cancer was electronically sampled and subjected to chart review, test results and molecular-guided therapies were found to be completely ascertained from the EMR and therapy was concordant with treatment guidelines in all cases. When a cohort of breast cancer patients was linked to Medicare

claims, the testing rates detected increased from 49% to 60% and treatment rates from 4.4% to 6.7% with the addition of claims data.

CONCLUSION: This project demonstrated several key attributes of PCORnet for studies of cancer molecular tests and associated targeted therapies, including:

- Data were efficiently shared in CDM format
- The CDM data were combined with other existing data in the EMR (pathology, mortality, text notes) and tumor registries (North American Association of Central Cancer Registries standardized format) to extend their value
- CDM data for colorectal cancer patients were linked at two participating institutions to charts and a statewide cancer registry which verified that most testing and treatment was captured in the EMRs of the PCORnet data partners, though sometimes in unstructured format.
- CDM data were linked to Medicare claims for breast cancer patients to support completeness of data capture - this verified the value added by claims data linkage to ascertain structured testing and treatment data.

While there were some limitations to completeness of cancer-related data capture in structured and interoperable form in the PCORnet CDM, it was reassuring that charts fully captured testing and treatment data for patients seeing medical oncologists at PCORnet institutions.

Specific Aims

The goal of this rapid cycle project was to document the patterns of use of molecular biomarkers and molecular-guided cancer therapies for patients with new-onset solid tumors (all single primary solid tumors) cared for in a variety of community and academic care settings. The project was also intended to test and expand the capacity of PCORnet to describe the details of cancer pathology, results of testing, cancer treatment and outcomes. Eleven PCORnet member sites participated in the Rapid Cycle Project (RCP) from three Clinical Data Research Networks (CDRNs) (**Table 1**). We additionally conducted descriptive queries distributed to all PCORnet sites by the PCORnet Distributed Research Network Operations Center (DRN OC).

Table 1. Participating sites

Site Summary	Study Aims Participated	CDRN
University of Iowa (UIOWA)	1,2,3	GPC
University of Kansas Medical Center (KUMC)	1,3	GPC
Marshfield Clinic Research Institute (MCRI)	1,3	GPC
University of Nebraska Medical Center (UNMC)	1,2,3	GPC
University of Texas Southwestern (UTSW)	1,3	GPC
University of Minnesota (UMN)	1	GPC
University of Missouri (MU)	1,3	GPC
Medical College of Wisconsin (MCW)	1,3	GPC
University of Florida (UFL)	1	OneFlorida
Vanderbilt University Medical Center (VAND)	1	MidSouth
Medical University of South Carolina (MUSC)	1	MidSouth
Duke/ Distributed Research Network Operations Center (DRN OC)	Descriptive Queries	CC

We had four specific aims:

Aim 1. **(Use Characteristics)** In a cohort of patients with an invasive single primary solid tumor, describe the use of common molecular tumor and, in some cases, germline biomarker testing and associated targeted cancer therapies.

Aim 2. **(Test Results)** In a subcohort of patients who had molecular biomarker testing and for whom a test result was available, determine whether the selected treatment was in accordance with the test result.

Aim 3. **(Completeness and Outcomes)** Using the cohort from Aim 1 in sites with linked claims data, assess the completeness of the EHR-derived data for identifying cancer treatments, molecular tests and outcomes.

Aim 4. **(Network Capacity)** Using the results from Aims 1-3, from the descriptive queries, and surveys of network partner data, synthesize and report lessons learned about network capacity for conducting pragmatic outcomes research in cancer including how well networks are able to capture the diagnoses, therapies, tests, test results, and outcomes of treatment and what issues must be overcome.

The approach and findings for each aim are summarized and detailed below.

Aim 1. (Use Characteristics) In a cohort of patients with an invasive single primary solid tumor, describe the use of common molecular tumor and, in some cases, germline biomarker testing and associated targeted cancer therapies.

Overview

The 11 participating sites are based in healthcare systems within three PCORnet networks (GPC, OneFlorida, MidSouth) across 10 US states and include 10 academic medical centers. These sites were selected from 80 PCORnet partner sites because they could rapidly provide hospital tumor registry data from their site along with linked electronic health records in PCORnet Common Data Model (CDM) format. The three non-GPC participating sites rapidly executed data use agreements for sharing de-identified individual level data; GPC sites already had a network data sharing agreement executed. The pooled dataset consisted of tumor registry data from each site and linked CDM diagnosis, procedures, prescribing, dispensing, medication administration, and death data tables. Data sources for the CDM include institutional billing and electronic health record data. The study cohort included patients with a single primary solid tumor diagnosed during 2013-2016, without regard to age or sex. We included molecular-guided therapies from the FDA Table of Pharmacogenomic Biomarkers in Drug Labeling that had a targetable genomic variant and cancer type defined in the indications and usage section of the drug label during 2013-2016 (**Table 2**). New indications for these therapies have been added in recent years. For instance, in 2014, pembrolizumab was approved for patients with advanced NSCLC whose patients express PD-L1. By 2017 it was approved for advanced NSCLC irrespective of PD-L1 expression and approved for patients with any solid tumor with microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR).

Table 2. Molecular-guided therapies included in the rapid cycle project

Molecular therapy group	Drug names	FDA Approved Cancer Type	Genomic variant
ALK inhibitor	Alectinib, ceritinib, crizotinib	NSCLC	ALK fusion
EGFR kinase inhibitor	Afatinib, erlotinib, gefitinib, osimertinib	NSCLC	EGFR Exon 19 deletion; $EGFR^{L858R}$; $EGFR^{T790M}$
EGFR MAB inhibitor	Cetuximab, panitumumab	Colorectal	EGFR amplification
ERBB2 kinase inhibitor	Lapatinib	Breast	ERBB2 amplification
ERBB2 MAB inhibitor	Ado-trastuzumab, pertuzumab, trastuzumab	Breast	ERBB2 amplification
PARP inhibitor	Olaparib, rucaparib	Ovarian	BRCA1/BRCA2
BRAF inhibitor	Dabrafenib, vemurafenib, cobimetinib, trametinib	Melanoma	BRAFV600E/V600K
PD-1 inhibitor	Pembrolizumab, nivolumab	NSCLC	PD-1

Abbreviations: NSCLC, non-small-cell lung cancer

We describe the approach and detailed findings for this aim in seven sections below:

1. Cancer cohort characteristics and experience linking hospital oncology registries with the PCORnet CDM compared with CDM-only cohort extraction
2. Frequency of molecular test order use by patient characteristics

3. Experience identifying molecular tests in the PCORnet CDM PROCEDURES table
4. Frequency of molecular-targeted therapy use by patient characteristics
5. Experience identifying molecular-targeted therapy in PCORnet CDM tables
6. Frequency of targeted therapy use among patients with test orders and frequency of test orders among patients who received targeted therapy
7. Evaluation of real-world outcome data in the CDM

Cancer cohort characteristics and experience linking hospital oncology registries with the PCORnet CDM compared with CDM-only cohort extraction

The DRN OC query was based on ICD9/10 Codes on encounters among patients during a three-year time period (2015-2017) who had two prior years free of the diagnosis codes for the specified cancer. This identified 1,492,871 patients with breast, colorectal, esophageal, lung, melanoma, ovarian, pancreatic, or prostate cancer in 67 Network Partners ("all-PCORnet" cohort)(**Table 3**). When restricted to the 11 Network Partners, 251,294 patients were identified with these tumors. In comparison, the number of these tumors identified with linked registry data during 2013-2016, in these 11 Network Partners totaled 47,629 (**Table 3**). Across all tumor types, linkage with tumor registries identified 86,154 patients with single primary solid tumors with linked CDM data. To optimize data completeness, cases were limited to 'analytic' cases, meaning that the tumor registry classified them as receiving all or part of their first course of treatment from the participating medical center.

Characteristics of the all-PCORnet patients are displayed in **Additional Documents Table 1a** and the tumor registry-linked patients in **Additional Documents Table 2**. In the tumor registry-linked study population from the 11 RCP Network Partner sites, one-fourth of included patients were age 70 or above. Race and ethnicity were relatively homogeneous in the RCP study population (86% White, 3.4% Hispanic) compared with the all-PCORnet cohort (71-79% White for non-melanoma cancers, 7% Hispanic). Among the 79% of RCP patients with a known tumor stage at diagnosis, 7% were in situ, 55% stage I-II, 18% stage III, and 21% were metastatic at the time of their initial diagnosis.

When applied to the RCP-participating datamarts, the ICD code-based data for a 3-year period classified from 2.6 to 5.6 times as many patients as having a new cancer as did the tumor registry linked data over a 4-year period (**Table 3**). This is consistent with prior comparisons that the PORTAL CDRN (not an RCP participant) has made, when they examined concordance between their tumor registry and ICD code-based ascertainment of cases by DRN OC query. The magnitude of difference suggests that some combination of the following may be occurring with the ICD code-based cohort identification: individuals may not have had new diagnoses (e.g. they were diagnosed earlier and codes represent follow-up visits or evaluation and management of disease progression); cases diagnosed and treated elsewhere would not be included in the tumor registry; patients with more than one cancer would not be included in the registry cohort (we included only patients with single primary cancer); and some may not have had cancer at all (for example, 'rule-out' coding).

Table 3. Cancer cohort sizes by CDM and registry cohort identification

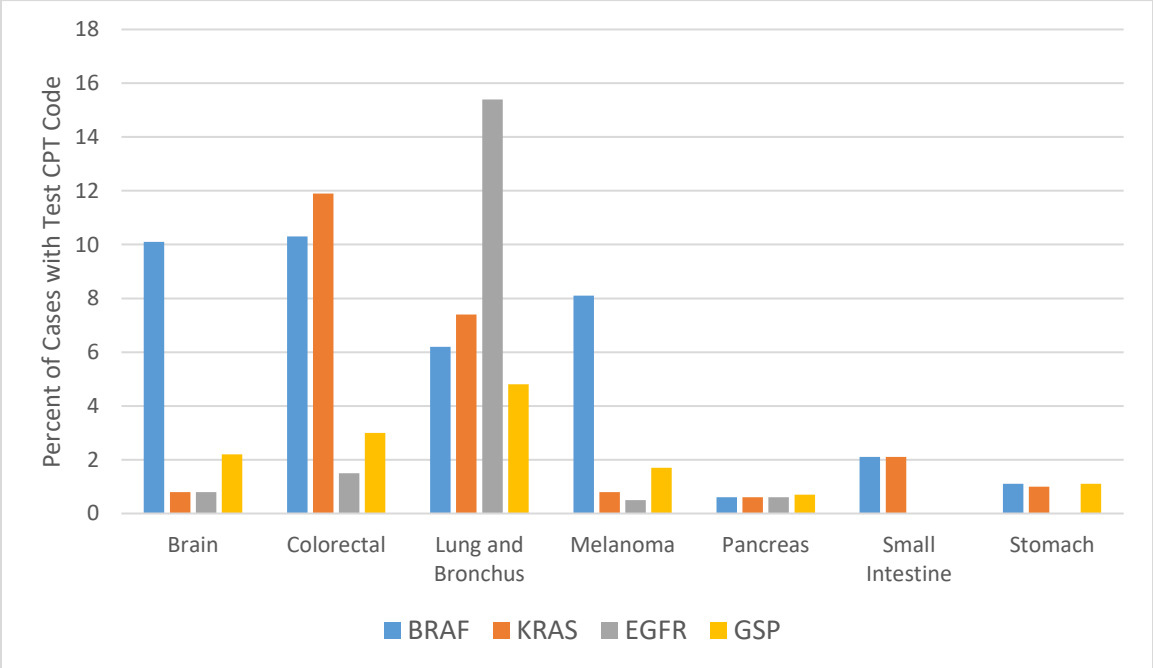
Cancer Site	All PCORnet datamarts, CDM query used for cohort identification	RCP Participating Datamarts, CDM query used for cohort identification	RCP Participating Datamarts, tumor registries used for cohort identification
Breast	404,319	72,030	14,049
Colorectal	181,872	31,930	5,350

Cancer Site	All PCORnet datamarts, CDM query used for cohort identification	RCP Participating Datamarts, CDM query used for cohort identification	RCP Participating Datamarts, tumor registries used for cohort identification
Esophageal	336,45	6,207	1,120
Lung	217,778	43,332	8,789
Melanoma	150,016	20,513	4,711
Ovarian	46,944	9,779	1,366
Pancreatic	67,215	12,629	3,365
Prostate	39,1082	54,874	8,879

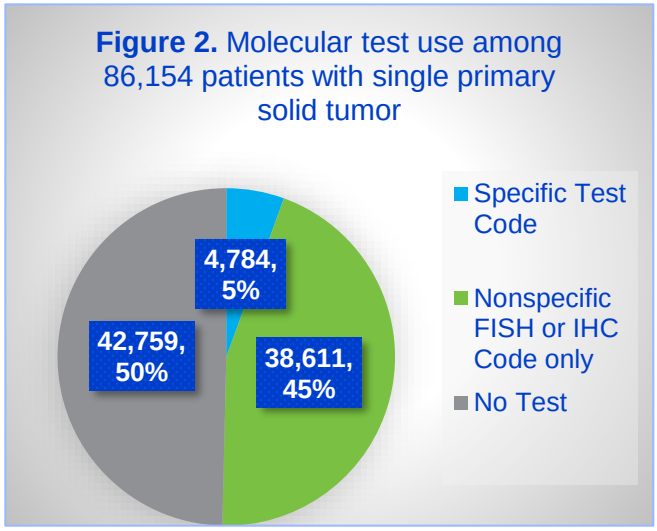
Frequency of molecular test procedures by patient characteristics

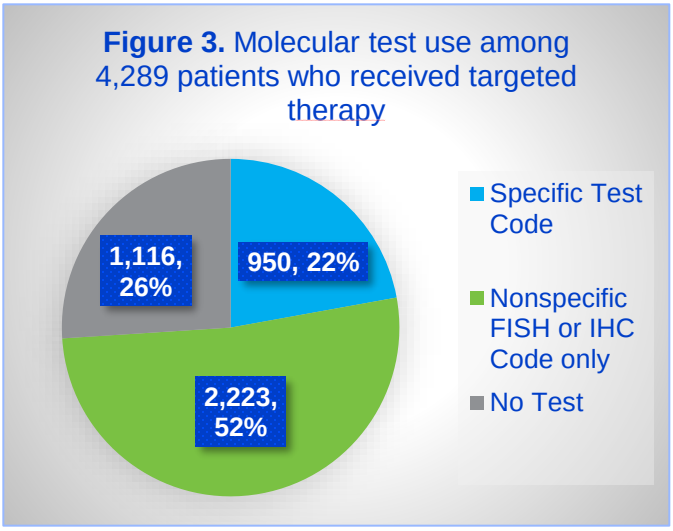
In the RCP tumor-registry linked study population, the prevalence of any molecular tumor testing detected via CPT code varied by tumor site, from a low of 26% of endocrine tumors to a high of 73% of breast cancers (**Additional Documents Table 3**). Tests with CPT codes that identify specific molecular targets were used much less commonly. Of the three specific tests used for more than 1% of patients (BRAF 2.4%, KRAS 1.8%, EGFR 2.0%), testing rates ranged from 0 to 10.3%, 0 to 15.4%, and 0 to 11.9% across cancer types. Genome sequence panels were associated with 1.1% of tumors and testing rates ranged from 0% to 4.8% across cancer types. Cancer types most commonly associated with these four most common molecular tests (other than FISH or IHC) are displayed in **Figure 1**. Test orders were most prevalent among patients diagnosed with stage IV disease (**Additional Documents Table 3**), consistent with indications for use of most molecular test-guided therapies. Testing rates also varied across study sites from a low of 36% (site E) to a high of 57% (site J), and with diagnosis year (47% in 2013 and 53% in 2016/17).

Figure 1. BRAF, KRAS, EGFR, and GSP CPT-coded tests, by cancer type



Testing rates among the cohort selected using only diagnosis codes from the CDM (the 'all-PCORnet' cohort) were substantially lower (**Additional Documents Table 1b**) than in the RCP tumor-registry linked study population. For instance, IHC tests were ordered among 48% of tumors in RCP sites compared with 15% in the all-PCORnet data. Similarly, FISH tests were found for 9.5% in the RCP sites and 3% in the all-PCORnet sites. BRAF, EGFR, and KRAS were found for 1% each and GSP for 0.3% in the all-PCORnet sites. The all-PCORnet data included only 8 cancer types. Four of the eight were cancer types associated with more frequent testing (colorectal, lung, melanoma, and pancreas). In addition to different cohort selection, the all-PCORnet query included only test codes that appeared on day 0-365 after the index date. In contrast, the RCP data included test codes occurring before or after the diagnosis date. Because cancer diagnosis is a process that occurs over time, including tests only after a specific date when a diagnosis code appears may be underestimate testing. However, including test codes before and after a diagnosis date may be too liberal and include more procedures used for diagnosis and classification rather than to identify molecular targets. Finally, the denominator may be inflated by rule-out diagnosis codes in the all-PCORnet data.





Experience identifying molecular tests in the PCORnet CDM PROCEDURES table

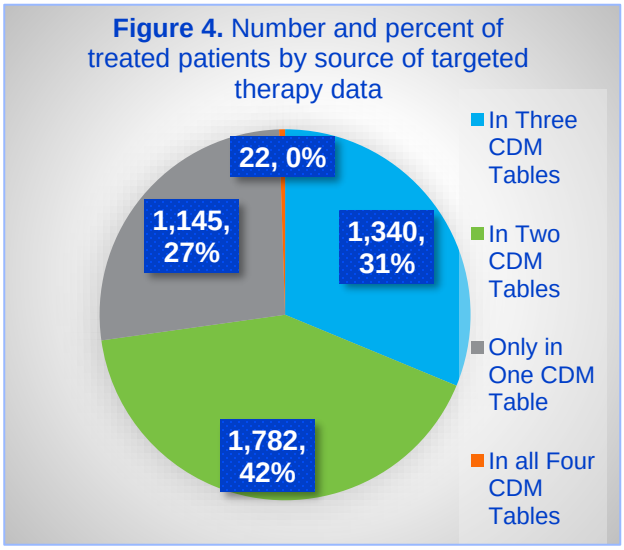
Molecular tests are identified with billing data in the CDM PROCEDURES table and half of all patients had a billing code for one of the testing procedures (**Figure 2**). However many molecular tests use procedures for which the billing codes do not specify the exact molecular marker being examined and the same billing code is used for an array of purposes in addition to identifying drug targets. The dominant billing codes used were for nonspecific fluorescence in situ hybridization (FISH) and immunohistochemistry (IHC) which are procedures used to detect many types of mutations. For instance, in addition to identifying specific drug targets, IHC is also used to diagnose a cancer as benign or malignant, determine the stage and grade of a tumor, and identify the cell type and origin of a metastasis to find the site of the primary tumor. Among patients who received targeted therapy, the proportion of target-specific tests was greater, but still only 22% (**Figure 3**). Among patients who received targeted therapy, CPT codes identified no potentially corresponding test order for 26% of treated patients. Billed test orders appear to under-identify use of molecular tests. Natural language processing or a structured pathology data source that includes target-specific names should be added to the data model.

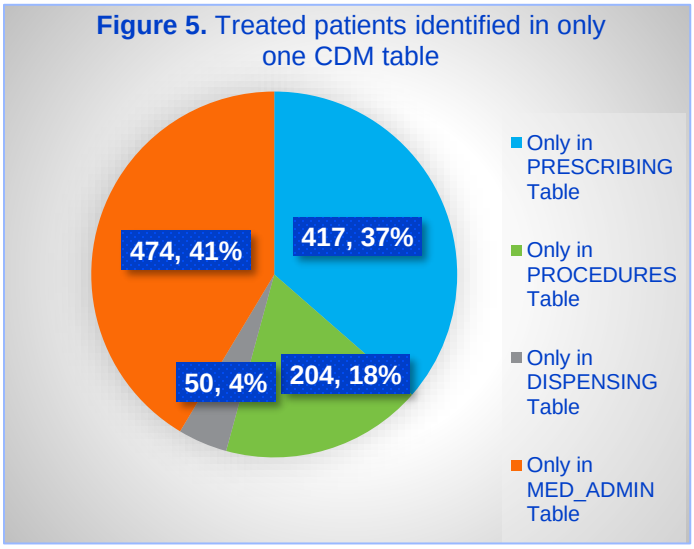
Frequency of molecular-targeted therapy use by patient characteristics

Molecular targeted therapy was identified among 4,289 (5%) cancer patients in the RCP tumor-registry linked study population (**Additional Documents Table 3**). Use rates were lowest among the oldest patients (4.7% 60-69, 3.9% 70-79, 2.5% 80+). Targeted therapy use was highest for patients with breast (10.5%), esophagus (6%), oral cavity and pharynx (7.9%), lung and bronchus (12.3%), and melanoma skin cancers (5.9%). Targeted therapy was used most commonly among patients diagnosed with stage IV disease, consistent with indications for use of most molecular test-guided therapies. Targeted therapy frequency varied across study sites. Sites B and D reported rates of 1% or less whereas the remainder of sites recorded targeted therapies for 3.8% to 8.4% of patients. Targeted therapy rates increased only modestly over the included diagnosis years, with 5.2% of patients classified as treated in 2015-2017 compared with 4.6% in 2013 and 4.9% in 2014. Indications for testing may vary with tumor histology (e.g. squamous vs adenocarcinoma). Although we did not match testing and treatment to the scenarios they would be most likely to be used (e.g. advanced stage non-small-cell lung adenomas), these types of follow-up investigations are possible with the pooled dataset.

The DRN OC also developed a query request to obtain frequency of use for six of the 21 medications of interest. The query selected patients age 21-85+ with a diagnosis of cancer in time periods: 1/1/2015 – 12/31/2017. Aggregate results including 1,596,945 patients with cancer ICD-9 or 10 codes were obtained from 54 Data Marts (**Additional Documents Table 1.c.**). Of this cohort, a total of 2% (31,756) of patients had a medication of interest. This prevalence is not comparable with the estimated prevalence of use in the RCP tumor registry-linked estimates because only a select group of medications were examined. The purpose of the query was only to examine how tables were populated across DataMarts as a first step toward asking data partners to compare their counts with the source data.

Experience identifying molecular-targeted therapy in PCORnet CDM tables





Therapies were identified from one or more of four CDM tables (**Figures 4-6**). Each table contributed additional treated patients. It might be expected that there would be no medication administrations unaccompanied by a prescribing event, no billed infusion procedures unaccompanied by a medication administration event, and no un-billed medication administrations. However, treated patients were frequently (27%) identified in only one of the four CDM tables.

This project was conducted using CDM version 3 tables supplemented with a pre-release MED_ADMIN table that RCP sites pilot-tested. It should be noted that some Network Partners felt that the preliminary guidance for CDM version 3 was ambiguous about how to populate the CDM with medication ordering, dispensing, and administration data. Version 4.1 guidance is substantially improved and, once the curated CDM version 4.1 tables are in production, replication of the Aim 1 analyses may indicate a cleaner separation of ordering, prescribing, and administration events. Further examination may be warranted to better understand the extraction of source data and how they are populated in the various CDM tables.

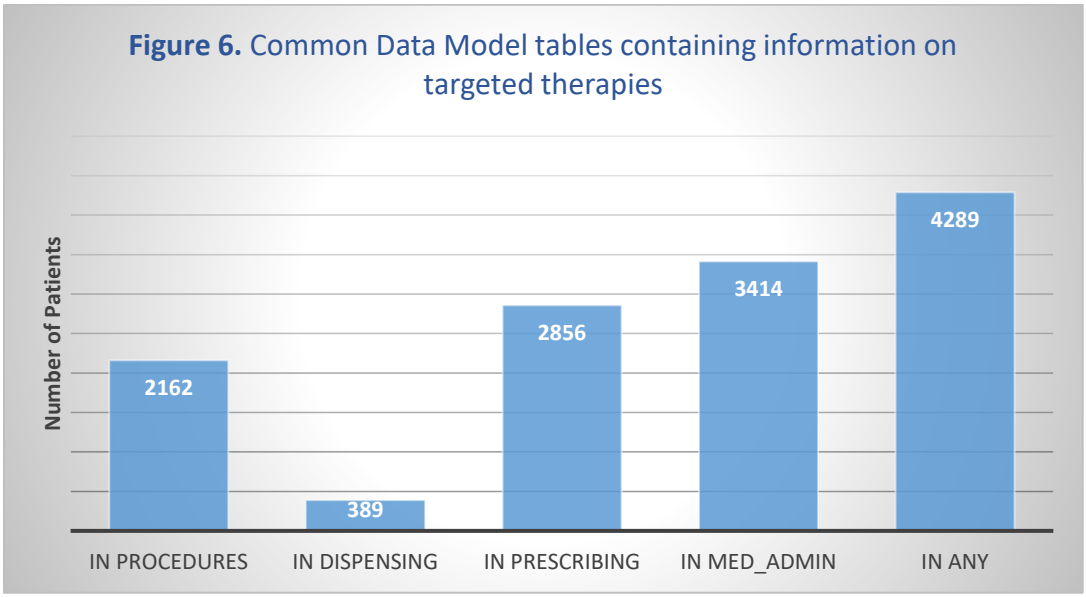


Figure 6. Common Data Model tables containing information on targeted therapies

Frequency of targeted therapy use among patients with test orders and frequency of test orders among patients who received targeted therapy

The relationship among testing and treatment was examined overall (**Figure 7**) and among those who received targeted therapy (**Table 4**). The highest rate of targeted therapy initiation (27.1% of patients) was seen among patients who had a genomic sequence analysis panel (GSP) (**Figure 7**). Full details of types of targeted therapy received among patients tested are displayed in **Additional Documents Table 3**. Among patients who received one of the targeted therapies investigated (**Table 4**), the most common therapies were the ERBB2 monoclonal antibody (MAB) inhibitors ado-trastuzumab, pertuzumab, and trastuzumab (n= 1,571). Patients who received a targeted therapy were more likely to have a test order than patients who did not receive a therapy (74.0% vs. 49.1%, respectively). A more detailed display of the particular tests received among therapy recipients is presented in **Additional Documents Table 4**. Because these data represent tests ordered by the participating medical centers, tests ordered by other health systems would not be captured in the CDM PROCEDURES table that was used to identify testing. In Aim 2 (below) we examined tests and test results using structured and unstructured data to determine whether the electronic medical record notes can provide a more complete picture and to assess on- and off-label use of therapies.

Figure 7. Percent of patients tested who received a targeted therapy, by test type

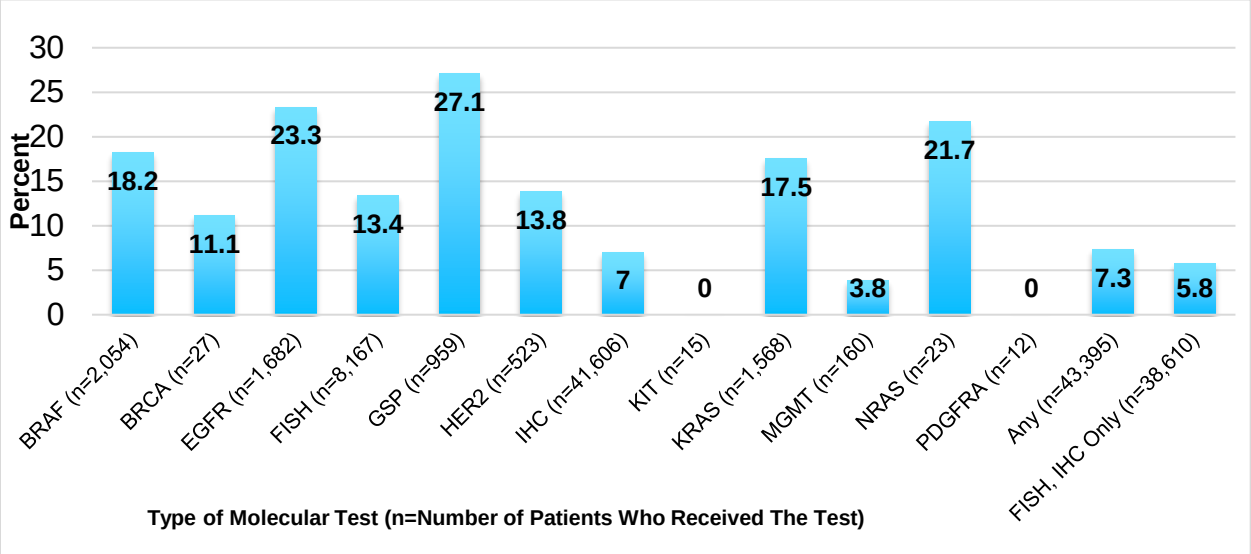


Table 4. Summary of tests among targeted therapy recipients

Molecular therapy group	Drug names	n	No Test, n (%)	Any Test, n (%)	Specific Test, n (%)	FISH or IHC Code Only, n (%)
ALK inhibitor	Alectinib, ceritinib, crizotinib	114	27 (23.7)	87 (76.3)	44 (38.6)	43 (37.7)
EGFR kinase inhibitor	Afatinib, erlotinib, gefitinib, osimertinib	453	118 (26.0)	335 (73.9)	184 (40.6)	151 (33.3)
EGFR MAB inhibitor	Cetuximab, panitumumab	586	213 (36.3)	373 (63.7)	104 (17.8)	269 (45.9)
ERBB2 kinase inhibitor	Lapatinib	56	8 (14.3)	48 (85.7)	4 (7.1)	44 (78.6)
ERBB2 MAB inhibitor	Ado-trastuzumab, pertuzumab, trastuzumab	1,571	390 (24.8)	1,181 (75.2)	93 (5.9)	1,088 (69.3)
PARP inhibitor	Olaparib, rucaparib	44	9 (20.5)	35 (79.5)	2 (4.5)	33 (75.0)
BRAF inhibitor	Dabrafenib, vemurafenib, cobimetinib, trametinib	203	29 (14.3)	174 (85.7)	108 (53.2)	66 (32.5)
PD(L)1 inhibitor	Pembrolizumab, nivolumab	1,547	370 (23.9)	1177 (76.1)	503 (32.5)	674 (43.6)
Any molecular therapy		4,289	1,116 (26.0)	3,173 (74.0)	950 (22.2)	2,223 (51.8)
No molecular therapy		81,865	41,643 (50.9)	40,222 (49.1)	3,835 (4.7)	36,387 (44.4)

Evaluation of Real-World Outcome data

RCP sites participated in the Friends of Cancer Research Real-World Endpoints (RWE) pilot project (https://www.focr.org/events/future-use-real-world-evidence). As part of this project we were able to examine the completeness of death data in the CDM for patients with advanced NSCLC treated with immune checkpoint inhibitors. We found that local sources of vital status are incomplete and insufficient for quantifying overall survival. Estimated median real-world overall survival was 8.58

months (7.56, 10.26) when only sites that had external data (social security death index and/or state death data) were included. If sites that had only local EMR data were included, the estimated median survival increased by over 80% to 15.78 months (12.2, 24.6). We also had to query sites to determine when they last updated their external data. PCORnet sites should secure external death data and CDM specifications should require documentation of the last update of all sources of death data.

The RWE pilot project also demonstrated that several extractable endpoints from the EMR correlate with overall survival. Further validation is required to determine whether these endpoints are reliable surrogates for overall survival and/or whether they can support regulatory and payer decision-making. In the PCORnet sites that used external death data, time to treatment discontinuation was moderately correlated with overall survival (Spearman correlation coefficient 0.62 (0.54, 0.69)). This was somewhat lower than correlations reported by other pilot project participating data systems (ranged from 0.77 to 0.89) (https://www.focr.org/sites/default/files/pdf/RWE%20Meeting%20Slide%20_7.9.18_Final.pdf). Time to next treatment was another endpoint most data systems were able to examine and correlations ranged from 0.36 to 0.70. Because PCORnet uses RxNorm codes which were not included in the systemic chemotherapy computable phenotype, the RCP sites did not compute this endpoint. Recurrence and progression were also not examined because these require chart review.

Aim 2. (Test Results) In a subcohort of patients who had molecular biomarker testing and for whom a test result was available, determine whether the selected treatment was in accordance with the test result.

Overview

In the subgroup of patients who had molecular tests for a genomic variant, it is of interest to know how frequently the selected treatment accorded with results of the tests. Because each therapy and test differs in terms of the important markers and associated molecular-guided therapies of interest, and because cancer care delivery may vary by cancer type, one specific use case was selected in order to provide very close scrutiny. We chose to ascertain results of target-specific tests for metastatic colorectal cancer because both anatomic and molecular pathology tests are used for this cancer and because we could take advantage of a statewide tumor registry to assess completeness of capture.

Key findings include:

1. Cohort extraction: A combination of hospital tumor registry and EMR data had high predictive value for identifying patients with metastatic colorectal cancer (CRC) but chart review was required to identify the subset of patients whose systemic therapy was being managed by the institution. A total of 213 chart-confirmed metastatic CRC cases (138 from UIOWA, 75 from UNMC) contributed data to the study aim.
2. Targeted therapy data: Use of molecularly guided therapies for patients with metastatic colorectal cancer was low (16% at UIOWA and 18.7% at UNMC). Of those who did receive these therapies, between 86% (UIOWA) and 100% (UNMC) were received at the participating institution.
3. Molecular testing: Outside testing was rare when treatment occurred at UIOWA/UNMC (1-6%) and results of outside testing were often found in clinic notes. In comparison with registry data for these patients (statewide registry for Iowa, hospital registry for Nebraska), only 5 of 213 patients had test results found only in the registry. In contrast, 74 of 213 patients had test results found through chart review at UIOWA/UNMC that were not found in the registries, likely reflecting that the therapies are often considered after a tumor progresses or recurs (tumor registries only include information about the initial diagnosis and treatment course). CPT codes under-identified testing, especially for MSI/MMR where there is no specific CPT code for MMR via immunohistochemistry.
4. Concordance with test results: All 30 patients who received cetuximab or panitumumab received it in concordance with KRAS test results (wild-type). All 6 patients who received pembrolizumab received it in concordance with MSI/MMR test results (high or abnormal).
5. Ability to identify computable pathology and genomic data: Emerging abilities to encode structured pathology data happened too recently to be fully employed at the participating sites. UIOWA and UNMC differ substantially in how they are incorporating test results in their clinical research data warehouses.

We describe the approach and detailed findings in five steps below: (1) cohort extraction; (2) targeted therapy data retrieval; (3) molecular test data retrieval; (4) concordance of molecular testing and molecular-guided therapy; and (5) ability to identify computable pathology and genomic data.

Cohort extraction

Both sites identified a cohort of patients age 18 or older diagnosed with stage IV CRC between January 1, 2013 and December 31, 2016, or patients diagnosed with stages I-III colorectal cancer between January 1, 2013 and December 31, 2016 and progressed to metastatic disease by December 31, 2016. Patients were excluded if they had prior or concurrent tumors. After cohort extraction, we added a final inclusion criterion requiring patients to have had at least one face-to-face visit with a medical oncologist at UIOWA or UNMC. This criterion was added because of the large number of patients who received only surgery at the two study sites and therefore had no reason to have complete records of molecular testing or targeted therapy; inclusion of these patients in the denominator would have led to an underestimation of rates of use of molecular testing and targeted therapy. The sites did not attempt to apply this criterion of visiting a medical oncologist computationally (it was done via chart review), but it is important for future work to determine how best to implement it systematically. Descriptions of how the inclusion criteria were applied at each site are specified in **Aim 2 – Table 1** below. A descriptive table of patient characteristics is provided in **Additional Documents Table 5**.

Aim 2 – Table 1. Application of inclusion criteria by study site

Inclusion criteria	UIOWA cohort	UNMC cohort
Patients diagnosed with stage IV CRC or stage I-III CRC and had metastatic ICD-9/10 code, 2013-2016	- Iowa Cancer Registry used to identify patients diagnosed with stages I-IV disease (analytic cases only) - ICD 9/10 codes used to identify metastatic disease among those diagnosed with stages I-III disease, confirmed by chart review to identify distant metastatic disease (did not include patients with positive lymph nodes only) - **N = 163**	- North American Association of Central Cancer Registries (NAACCR) codes used to identify patients with Stage UNK, I-IV CRC. (analytic cases only) - SNOMED CT, ICD-10-CM and CEA levels > 25 used to identify metastatic disease among those diagnosed with stages I-III disease (did include patients with positive lymph nodes only). Confirmed by chart review - **N = 221**
Met with UIOWA/UNMC medical oncologist	- Chart review confirmation - **N = 138**	- Chart review confirmation - **N = 75**

Candidate patient cohort identification using computational means is a driving factor behind the PCORnet CDM. Many molecular-guided cancer therapeutics are primarily indicated for metastatic disease, often after other therapies have been tried and often among patients who progressed after initially being diagnosed at a local or regional stage. While the NAACCR-coded tumor registries assisted in identification of CRC patients, encounter diagnoses (ICD 9 or 10 CM), problem lists (SNOMED CT), and clinical laboratory data (CEA levels) were used to identify patients with potentially metastatic disease. This approach is promising. At UIOWA for example, chart review for patients selected in this way could not find evidence of metastatic disease in only 9 patients with a remaining 163 confirmed via chart review (PPV 94.8%).

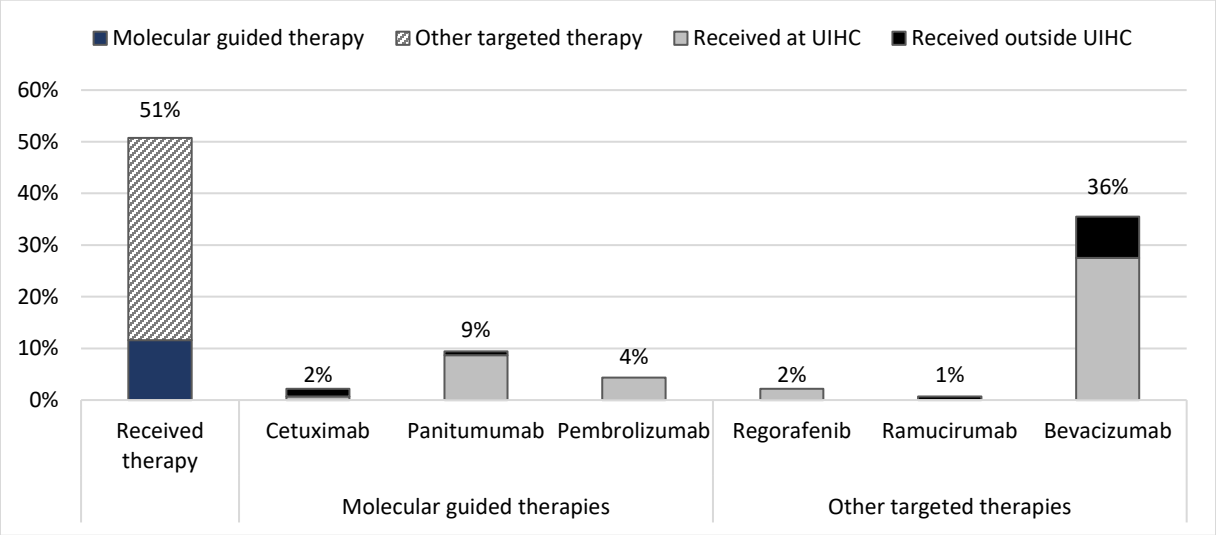
Targeted therapy data retrieval

Targeted therapies for metastatic CRC include the following: cetuximab, panitumumab, pembrolizumab, nivolumab, Ziv-aflibercept, regorafenib, ramicirumab and bevacizumab. Only cetuximab, panitumumab, pembrolizumab and nivolumab are considered 'molecular-guided therapies' because they are only

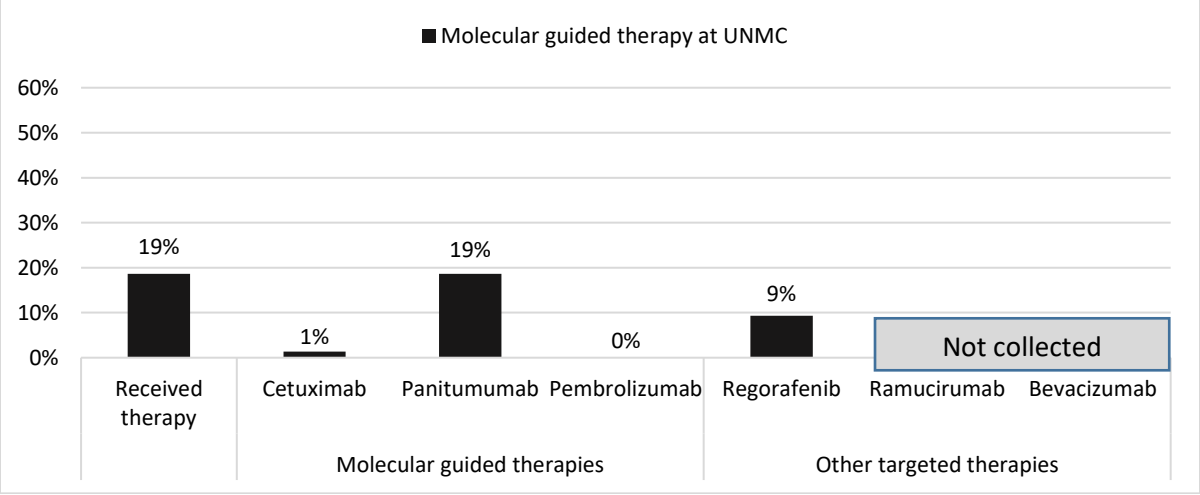
indicated for patients who have specific results to molecular tests. These four therapies were extracted by each study site. Both sites also searched for off-label uses of vemurafenib and debrafanib but did not find any use of these therapies in our study cohorts. In addition, the UIOWA site extracted the other targeted therapies that do not require molecular testing prior to administration (Ziv-aflibercept, regorafenib, ramicirumab and bevacizumab).

Among patients in the UIOWA cohort, 51% received one of the targeted therapies listed above (**Aim 2-Figure 1a**). However, only 16% (22/138) received at least one of the 4 molecular-guided therapies. Likewise, 18.7 % (14/75) of UNMC patients received molecular-guided therapies (**Aim 2-Figure 1b.**). The proportion of patients who received molecular-guided therapies outside of the study sites at UIOWA was 14% (3/22) and 0% at UNMC. Of the 22 UIOWA patients who received molecular-guided therapy, 4 (8.2%) were not identified in the CDM; 3 patients received these therapies outside the UI Healthcare system.

Aim 2- Figure 1a. Percentage of patients receiving molecular-guided vs. targeted therapy at UIOWA



Aim 2- Figure 1b. Percentage of patients receiving molecular-guided vs. targeted therapy at UNMC

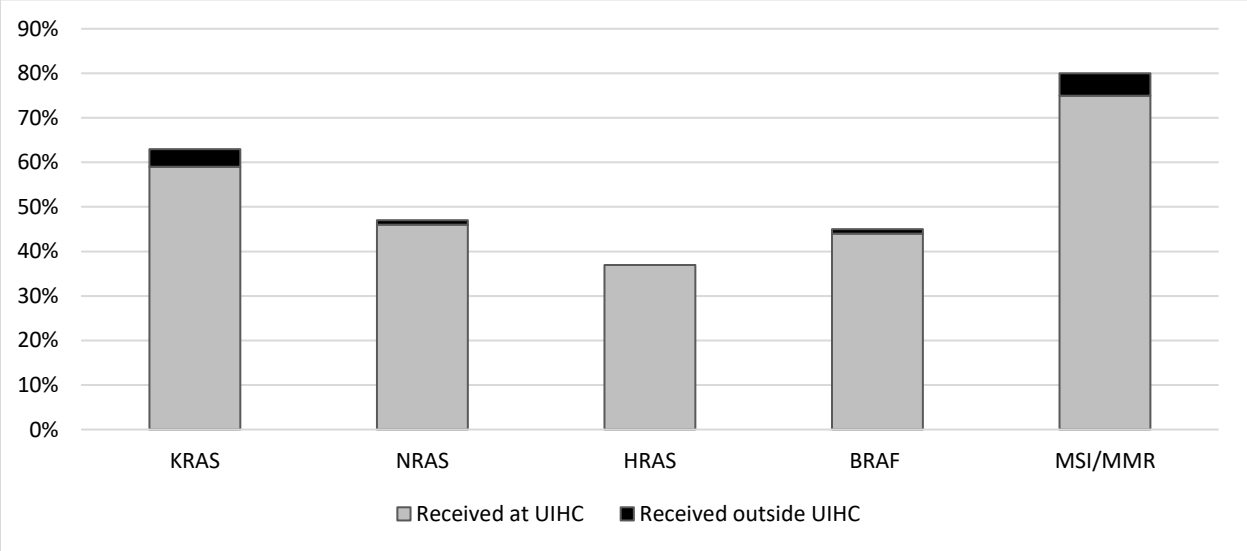


UNMC compared medications identified by chart review with those identified computationally. Computational identification of medication was determined by queries against EHR data using RxNorm and National Drug Code (NDC) for each medication of interest. Database queries identified 12 distinct patients who were prescribed panitumumab while chart review identified 14. The difference in patient counts reflected the differing time frames considered. The database query did not interrogate the EHR for medication data after 12/31/2016 while the manual chart review considered records into 2017. The medication administration dates for each of the two patients were after 12/31/2016 and outside of the query date range. Comparison of regorafenib use as determined by chart review and computational methods were more interesting. The net difference between patients prescribed regorafenib was one. Six were identified by database query and seven were identified by chart review. The patient identified by chart review was not identified by database query because regorafenib was given as part of a clinical trial. Clinical trial medications are not identified by RxNorm or NDC.

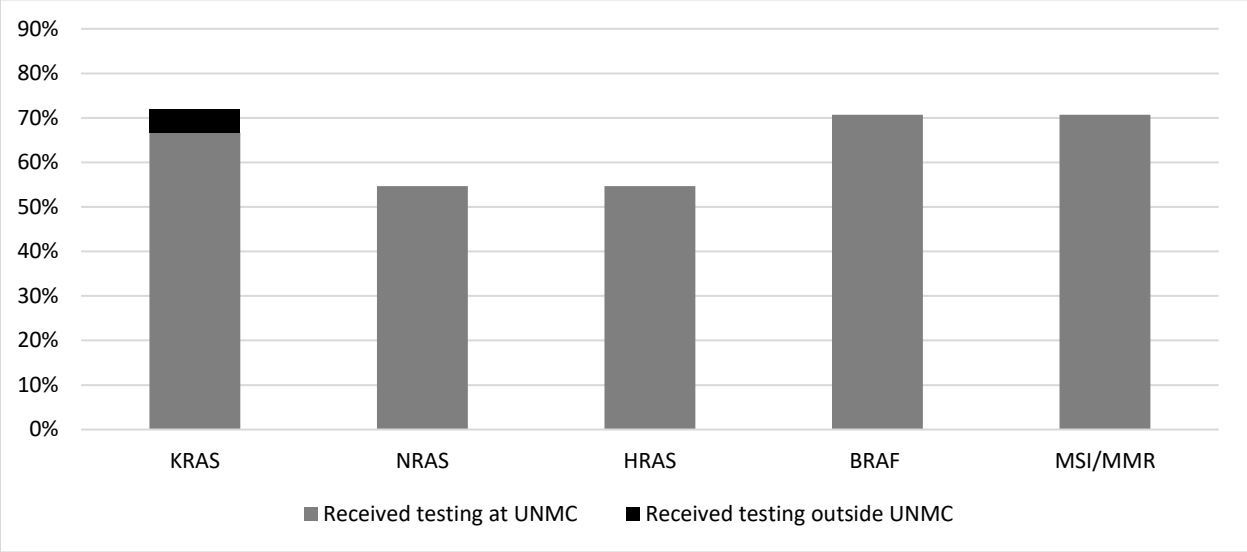
Molecular testing data retrieval

Treatment guidelines indicate that epidermal growth factor receptor (EGFR) inhibitors (cetuximab, panitumumab) should only be given to patients with no *RAS* or *BRAF* tumor gene mutations. In addition, pembrolizumab or nivolumab are recommended only for patients with metastatic mismatch repair (MMR)-deficient colorectal cancer in second- or third-line therapy. MMR deficiency can be assessed by a microsatellite instability (DNA) test typically performed in molecular pathology labs, or by a MMR immunohistochemistry test typically performed in anatomic pathology labs. We therefore extracted results for the following molecular tests: KRAS, NRAS, HRAS, BRAF, MSI and MMR. We further examined the proportion of testing that occurred outside of the study sites (see **Aim 2- Figures 2a-b**). Over half of patients received KRAS testing, which is the oldest of the *RAS* tests, roughly half received NRAS and BRAF testing, and less than half received HRAS testing, which is the newest *RAS* test. The majority of patients at both sites received MMR/MSI testing.

Aim 2- Figure 2a. Molecular testing received: UIOWA



Aim 2- Figure 2b. Molecular testing received: UNMC



Outside testing was rare when treatment occurred at UIOWA/UNMC (1-5%), but more common when the patient came only for consultation. The results of outside tests were often found in clinic notes. In order to confirm this finding, we compared KRAS testing found via record review to the KRAS site specific factor collected by the Iowa Statewide Cancer Registry which incorporates information from all hospitals where patients receive diagnostic or treatment services (**Aim 2 – Tables 3a-b**). Although statewide registry data were not available for Nebraska, comparisons with UNMC registry data were conducted as it is possible that hospital tumor registrars incorporate such information when noted in the EHR. From these comparisons, we found only 3 patients at Iowa and 2 at Nebraska who had KRAS testing performed that was not documented in the medical record. Conversely, 38 Iowa patients and 36 Nebraska patients had testing found by medical record review that was not captured by the registries. This is likely due in large part because KRAS testing was often performed after the typical registry abstracting window has closed (63% had KRAS testing performed >6 months after diagnosis).

In addition to the UIOWA and UNMC comparisons, all participating RCP sites were requested to manually examine a small number of cases for which evidence of therapy was present in the CDM, but no evidence of molecular test data was found in the CDM. Of the 59 patient charts that were examined across all sites, 47 (80%) contained evidence of test results. For the vast majority of these cases, the test orders/results were found in some sort of clinical note or scanned document (i.e., unstructured data). Of these, 12 were documents from an external source.

Aim 2 – Table 3a. Comparison of KRAS testing captured by EMR review vs. statewide Iowa Cancer Registry

		Iowa Statewide Tumor Registry KRAS Test Results			
		Abnormal	Normal	Test not done/ unknown	TOTAL
UIOWA EMR KRAS Test Results	Abnormal	25	0	16	41
	Normal	1	26	22	49
	Test not done/unknown	2	1	45	48
	TOTAL	28	27	83	138

Aim 2 – Table 3b. Comparison of KRAS testing captured by EMR review vs. UNMC Tumor Registry data

		UNMC Tumor Registry KRAS Test Results			
		Abnormal	Normal	Test not done/ unknown	TOTAL
UNMC EMR KRAS Test Results	Abnormal	10	1	12	23
	Normal	0	11	24	35
	Test not done/unknown	2	0	15	17
	TOTAL	12	12	51	75

Numerous cancer research studies using large healthcare databases, such as SEER-Medicare or the Cancer Research Network Virtual Data Warehouse, rely on billing codes to identify procedures and treatments. For instance, in settings in which test *results* cannot yet be extracted from the EMR computationally, there may be interest in efficiently identifying a cohort of patients who were known to have had a particular test performed for targeted curation of EMR data. Thus it was of interest to examine how completely CPT codes used for billing could identify patients who had had KRAS, BRAF, and MSI/MMR testing.

In the UIOWA cohort, we examined what percent of testing could have been captured by CPT alone using the PCORnet CDM data. There are not specific CPT codes available for MMR testing so it cannot be identified using billing/claims data. Compared to chart review, use of CPT codes performed reasonably for KRAS and BRAF testing, which have their own CPT codes (**Aim 2 – Table 4**). However, use of CPT codes performed very poorly for MSI/MMR testing. MMR (vs. MSI) testing was used in the overwhelming majority of patients (72%) at UIOWA, and there are not specific CPT codes available for MMR immunohistochemistry testing. This underscores the importance of initiatives to identify *computable* pathology and genomic data.

Aim 2 – Table 4. Comparison of molecular testing at UIOWA based on EMR review vs. CDM CPT codes

	Molecular Testing	EMR Data			Percent Tested	
Data	**KRAS (CPTs: 81275, 81276)**	Tested	Not Tested	TOTAL	**EMR Review**	**CPT Only**
	Tested	80	0	80	**65%**	**58%**
	Not Tested	10	48	58		
	TOTAL	90	48	138		
	BRAF (CPT: 81210)	Tested	Not Tested	TOTAL		
	Tested	56	0	56	**46%**	**41%**
	Not Tested	7	75	82		
	TOTAL	63	75	138		
	MSI/MMR (CPT: 81301)	Tested	Not Tested	TOTAL		
	Tested	39	0	39	**80%**	**28%**
	Not Tested	71	28	99		
	TOTAL	110	28	138		

Concordance of molecular testing and molecular-guided therapy

All 16 patients at UIOWA and all 14 patients at UNMC who received an EGFR inhibitor (cetuximab or panitumumab) received KRAS testing, and all had normal (wild-type) KRAS which is in accordance with treatment guidelines (**Aim 2 – Table 5**). KRAS testing was recommended prior to EGFR inhibitors for the entire duration of the study period, whereas the other RAS and BRAF tests were not recommended until after the start of the study period. This may at least partially explain why there were a few patients (1 at UIOWA and 4 at UNMC) who were found to have NRAS, HRAS or BRAF mutations but still received cetuximab or panitumumab.

All 6 patients who received pembrolizumab at UIOWA had high or abnormal MSI/MMR test results (indicating MMR deficiency) and were therefore in accordance with treatment guidelines (**Aim 2 – Table 6**). No patients at UNMC received pembrolizumab, and no patients at either site received nivolumab.

Aim 2 – Table 5. Use of EGFR inhibitors and recommended molecular testing

		Iowa (n=138)				UNMC (n=75)			
		Cetuximab (N=3)		Panitumumab (N=13)		Cetuximab (N=1)		Panitumumab (N=14)	
Molecular tests and results		N	%	N	N	%	%	N	%
KRAS	Normal/Negative	3	100%	12	100%	1	100%	14	100%
	Mutated/Positive	-	-			-	-		
	Not done	-	-			-	-		
NRAS	Normal/Negative	1	33%	7	58%	-	-	8	57%
	Mutated/Positive	-	-	2	17%	-	-	2	14%
	Not done	2	67%	3	25%	-	-	4	29%
HRAS	Normal/Negative	-	-	9	75%	-	-	10	71%
	Mutated/Positive	1	33%	0	-	-	-	0	-
	Not done	2	67%	3	25%	-	-	4	29%
BRAF	Normal/Negative	1	33%	9	76%	1	100%	11	79%
	Mutated/Positive	-	-	2	17%	-	-	2	14%
	Not done	2	67%	1	8%	-	-	1	7%

Aim 2 – Table 6. Use of pembrolizumab and recommended molecular testing

		Iowa (N=6)		UNMC (N=0)	
Molecular tests		N	%	N	%
MSI/MMR	Stable/Normal	-	-	-	-
	High/Abnormal	6	100%	-	-
	Not done	-	-	-	-

Ability to identify *computable* pathology and genomic data

The ability to computationally identify pathology and genomic data pertaining to patients and cancer was challenging. Anatomic pathology data describing histologies, histologic grade and tumor staging is routinely recorded in natural language. Even when pathology data is recorded in synoptic format (i.e. using structured checklists to produce standardized clinical documentation consistent with the College of American Pathologists cancer worksheets), it is still in natural language and not computable. UIOWA anatomic and molecular pathologists began entering discrete test result data into the Epic EMR which can be added to the CDM. We attempted to computationally extract molecular testing data from the

UIOWA EMR to determine how closely this data would compare to the manual chart review data for the patients who had their testing performed at UIOWA (**Table 7**). We could computationally extract information from the EMR for all KRAS, NRAS, HRAS, MSI/MMR and all but one BRAF tests that were done at UIOWA. However, we could only extract the actual KRAS test results for 65 (76%) of patients tested, NRAS results for 42 (67%) of patients tested, HRAS results for 29 (57%) of patients tested, and BRAF results for 38 (60%) of patients tested. We were able to extract information about MSI/MMR results in all 103 patients tested. These results suggest that overall, we can electronically pull information about having the test done in >98% of people who had testing at UIOWA, but can only extract the results for roughly two-thirds of the patients who received these particular tests. The issues in identifying results were largely related to multiple mutation (panel) testing which made it difficult to identify which mutation was present when pulling the data electronically. In addition to mutation testing results, there is interest in the grade and staging data for many studies – a follow-up activity could detail more fully the difficulties encountered, and estimate costs for extracting, transforming and loading such data so studies could appropriately budget for this.

Aim 2 – Table 5. Identification of molecular tests and results of molecular tests by electronic extract of UIOWA EMR vs. manual chart review

Test Type	Tested at UIOWA according to manual chart review	Test identified via electronic extraction of UIOWA EMR	Test result identified via electronic extraction of UIOWA EMR
KRAS	85	85 (100%)	65 (76%)
NRAS	63	63 (100%)	42 (67%)
HRAS	51	51 (100%)	29 (57%)
BRAF	61	60 (98%)	38 (62%)
MSI/MMR	103	103 (100%)	103 (100%)

UNMC employs computable, SNOMED CT encoded cancer worksheets for CRC. Thus, at UNMC, data created at the time of pathology assessment is available for computation. However, this process did not begin at UNMC until 2015 and no cases included in this study had such data available. Future studies with more recent timeframes of interest would expose pathology information to computational identification at UNMC. These data include mismatch repair data and BRAF assessments as determined by immunohistochemistry.

Genetic sequence information pertaining to the oncogenes evaluated in this study (i.e., BRAF, KRAS, NRAS, HRAS) are encoded in SNOMED CT and sent to UNMC biobanks in computable format. However, only data from recent cases are currently stored in the biobank in this encoded fashion. While data was not available in UNMC biorepositories, molecular data and variant results were nevertheless available and computably accessible in the molecular pathology file system. Processing of these molecular tests and results into the biorepository and clinical data warehouse (CDW) is now tractable at UNMC and supports future research projects addressing molecular studies and cancer.

The readiness of CDWs to support studies similar to that of Aim 2 using computational methods in lieu of chart review or which minimize the need for chart review is promising. Future work is certainly required in the areas of pathology and genomics. Future studies with time frames of interest that correspond to the more recent implementation of UNMC pathology and genomic data capture will further elucidate the extent to which CDWs can adequately address research questions involving pathology and molecular pathology.

Aim 3. (Completeness and Outcomes) Using the cohort from Aim 1 in sites with linked claims data, assess the completeness of the EHR-derived data for identifying cancer treatments, molecular tests and outcomes.

Overview

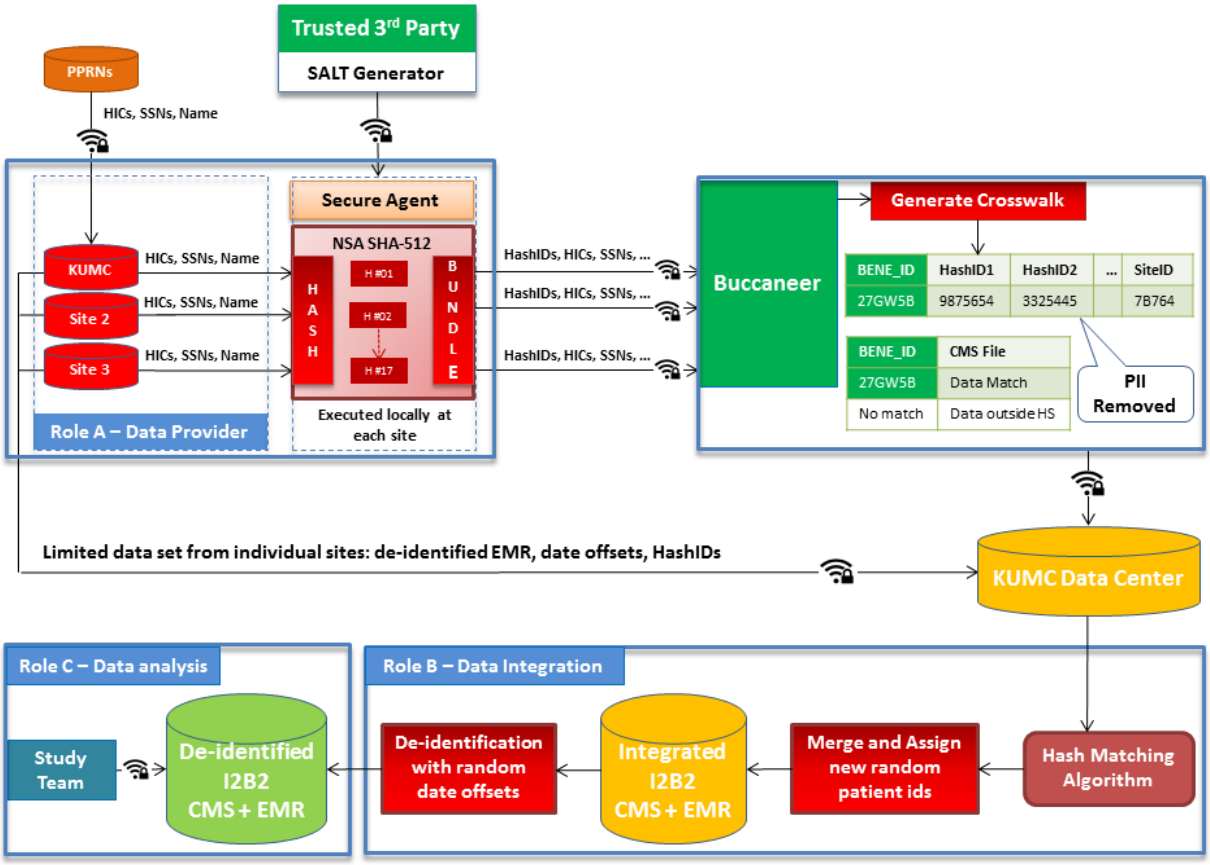
In Aim 3, we built upon the breast cancer cases from the Aim 1 findings to compare testing and treatments from the CDM against linked Medicare claims in the Greater Plains Collaborative Reusable Observable Study Environment (GROUSE) (https://informatics.gpcnetwork.org/trac/Project/wiki/GROUSE) that integrates site CDMs and i2b2 repositories (e.g. Tumor Registry elements) with Medicare/Medicaid claims data for their entire state populations. Seven GPC sites fully participated in this Aim:

- University of Iowa (UIOWA)
- University of Kansas Medical Center (KUMC)
- Marshfield Clinic Research Institute (MCRI)
- University of Nebraska Medical Center (UNMC)
- University of Texas Southwestern (UTSW)
- University of Missouri (MU)
- Medical College of Wisconsin (MCW)

To achieve this aim we first successfully incorporated all claims and refreshed crosswalks with sites against the entire 2011 through 2015 Medicare Claims (20,505,798 distinct beneficiaries). Patients were included who, at the time of diagnosis, had Part A, B and D coverage and who were not in enrolled in Medicaid Managed Care as defined by the Claims Medicare Beneficiary Summary File.

Entrusted with both Medicare/Medicaid claims and complete copies of the sites' PCORnet Common Data Models and i2b2 repositories, GROUSE uses a series of steps to protect health information (**Aim 3 - Figure 1**). Two final representations of the PCORnet Common Data Model were created that are linked on patid: one containing regular CDMs and the other CDM contained CMS claims for the cohort.

Aim 3 – Figure 1. Process for loading, cross-walking, and integrating CDM data from sites with Medicare data.



Cancer cohort characteristics and experience linking hospital oncology registry defined PCORnet CDMs against a CMS claims CDM

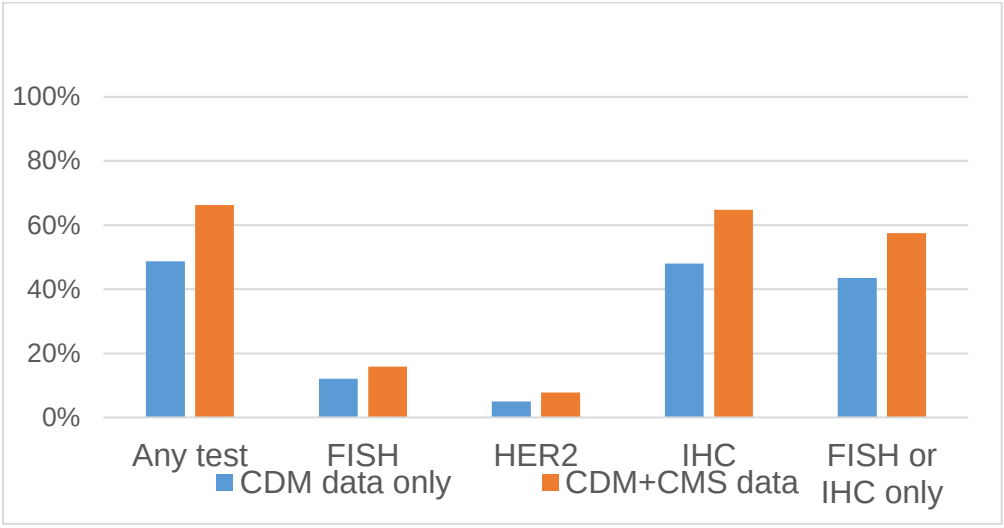
Characteristics of the all-PCORnet patients are displayed in **Additional Documents Table 6** with 11,124 patients in the replicated Aim 1 cohort that crosswalked to 2,154 patients in the Medicare breast cancer cohort. Note that approximately half of the cohort were level 2 class of case, "Initial diagnosis and treatment at reporting facility (NAACCR|610:10-14)", which may influence subsequent findings in terms of complete capture of diagnostics and treatment.

Molecular testing comparisons

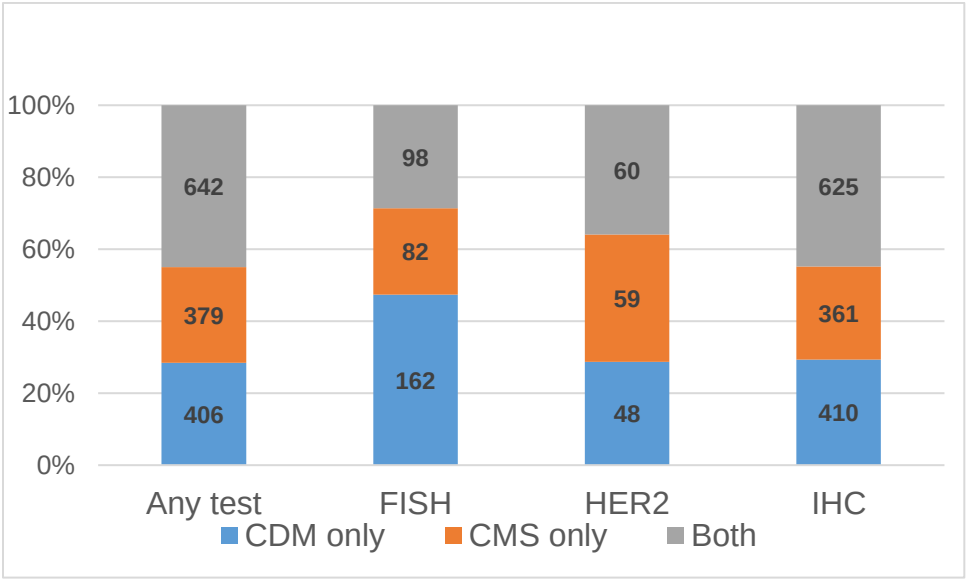
Aim 3 – Figure 2 illustrates the molecular test procedures identified when using the CDM only vs. when claims were added. CMS data detected more BRAF (.05% vs 0%), BRCA (5.16% vs 0%), and KRAS (.05% vs 0%) testing but less EGFR (0 vs .05%), HER2 (5.11 vs. 5.41%), and FISH or IHC (37.3% vs. 41.5%) testing. Although overall testing rates appear similar in aggregate, when examined on the individual level (**Aim 3 – Figure 3**), concordance is lower than would be expected. We would expect that molecular testing may sometimes precede treatment at a tertiary academic medical center and were not surprised by the approximately 20 to 30 percent of testing seen only in the claims as that may be due to testing by the provider of initial diagnosis and screening. However we were surprised at the number of tests detected only in the CDM and fairly reduced overlap between the two sources. We will work to check

both linkage methods and conduct more detailed analysis at the case level. Overall, results support the value added by linkage to claims data as seen in **Aim 3 – Figure 4**, where overall use rates were higher when EMR and CMS data were combined (49% vs 60%, respectively).

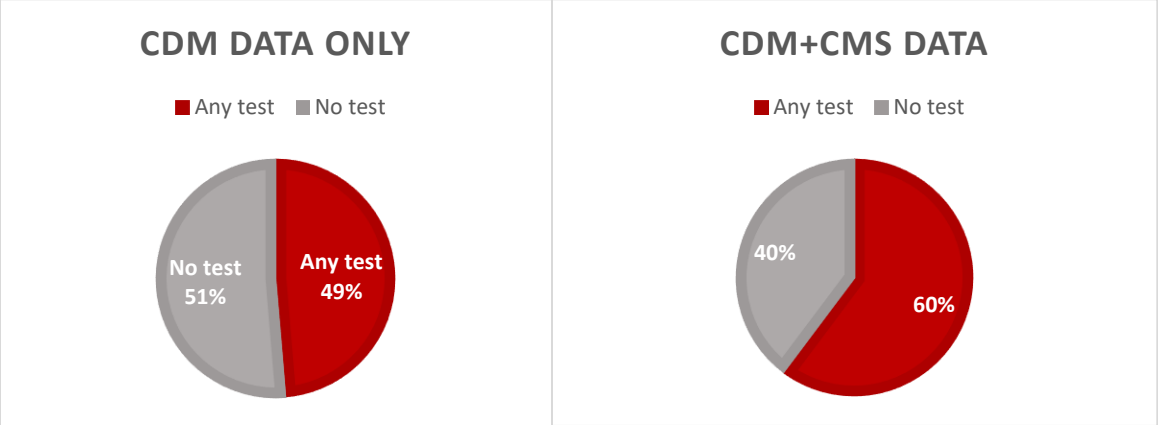
Aim 3 - Figure 2. Percentage of breast cancer patients receiving molecular testing by whether supplemented by Medicare claims



Aim 3 – Figure 3. Molecular testing detected by source of data when using CDM supplemented with CMS claims



Aim 3 – Figure 4. Total molecular test usage in the site data versus the site data augmented with Medicare claims



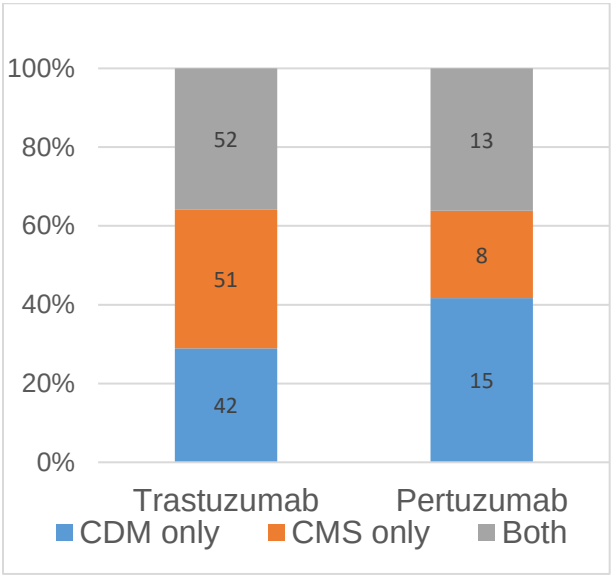
Targeted therapy comparisons

Similar comparisons were made for treatment with targeted therapies with **Aim 3 – Table 1** illustrating the total detected therapies. Several sites had incomplete dispensing data for earlier years and these will need to be removed for future comparisons so that treatment rates are not artificially low. In spite of this, CMS claims seldom identified more treatment than CDM tables for the CMS-linked patients – 4.41% received any of the interrogated therapies according to EHR-derived CDM data compared with 4.78% in the CMS-derived CDM. However, the merger of CMS and site CDM had a significant additive effect as illustrated in **Aim 3 – Figure 5**, continuing to note that the site CDM versus the CMS claims each identified different sources of medication information. The CDM identified a few patients who received lapatinib whereas the CMS claims did not. Shown in **Aim 3 – Figure 6**, overall use rates were higher when EMR and CMS data were combined. Finally, we examined treatment use for the full breast cancer cohort and the Medicare sample and noted that treatment rates were lower in the latter sample (4.36% vs 7.23%; data not shown), which was expected due to the older age in this group and patterns noted in the literature of lower treatment rates among older patients.

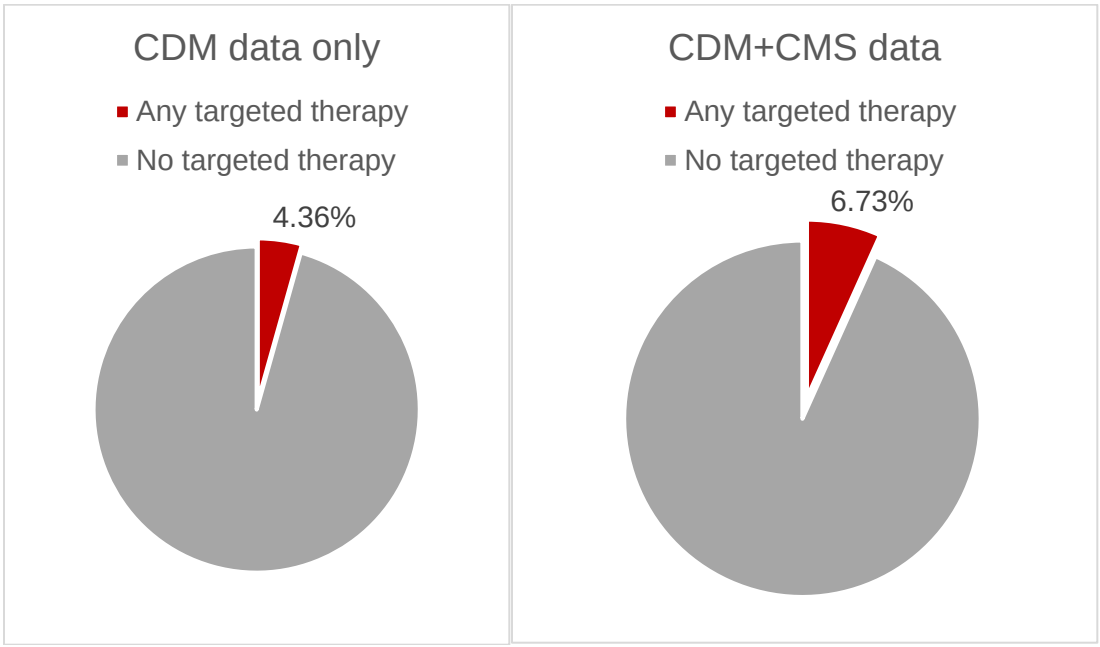
Aim 3 – Table 1. Targeted therapy percentages for site CDMs versus Medicare claims

Data	Diagnosis year	n	Any targeted molecular therapy	lapatinib	trastuzumab	ado-trast	pertuzumab
GROUSE EMR CDM, Aim 3	2011-2015	2,154	4.41	<11	4.36	<11	1.3
GROUSE CMS CDM, Aim 3	2011-2015	2,154	4.78	0	4.78	<11	0.97

Aim 3 – Figure 5. Breast cancer targeted therapy use contributions by PCORnet CDM and Medicare claims



Aim 3 – Figure 6. Total targeted therapy usage in the site data versus the site data augmented with Medicare claims



Preliminary comparisons of vital status data sources

Vital status was compared in EHR data and Medicare enrollment file data. During 2011-2015, a total of 174 deaths were identified in either of the two sources. Of these, 76.4% were identified in both sources, 13.2% in Medicare data only and 10.3% in EHR data only.

Aim 4. (Network Capacity) Using the results from Aims 1-3, from the descriptive queries, and surveys of network partner data, synthesize and report lessons learned about network capacity for conducting pragmatic outcomes research in cancer including how well networks are able to capture the diagnoses, therapies, tests, test results, and outcomes of treatment and what issues must be overcome.

To address this aim, we summarize findings from four categories of issues pertinent to ability of a data resource to support research studies:

1. Can we identify and follow cases through time and different data files?
2. Are the data we need complete and valid?
3. Can we identify gaps in the data?
4. How much variability is there across study sites?

Identifying and following cases

Case ascertainment. In spite of some inefficiencies caused by some differences in formatting and duplicate records, it was possible to extract tumor characteristics from linked registry data and merge these with CDM data. This adds value for selecting eligible patients based not only on having a confirmed cancer diagnosis but also characteristics such as histology, grade, stage, and other characteristics. The CDM does not presently include a tumor table and this study demonstrates the potential value that would be added by creating a way for tumor characteristics to be incorporated in the CDM, including associated guidance and regular data quality evaluations with each CDM refresh.

Adoption of a common tumor table may make data pooling and distributed analyses easier in the future. The PCORnet Cancer Collaborative Research Group has proposed specifications for a CDM tumor table. Despite the fact that all sites provided NAACCR-compliant data, there were inconsistencies in formatting across sites that needed to be reconciled. Some sites, for example, provided numbers in text format which preserved leading zeros specified by data dictionaries. Other sites provided such data in numeric formats that did not allow this. Similarly, some sites coded dates as datetime values (i.e., seconds) where others coded them as date values (i.e., days). This makes pooling of data unnecessarily complex. Updates to tumor registry data files are often provided as duplicate records, and it is difficult to determine which record contains the most accurate information. Tracking patients over a length of time only compounds this problem as updates accumulate. In this study, we included only those cases with a single record, and that resulted in the exclusion of a large number of cases. This affected data from some participating institutions more than others. Once a decision about how tumor data will be represented in the CDM, guidance should be developed about loading tumor data and should require addressing multiple records before loading a single record per tumor.

Limitations of tumor registries include that the ascertainment and completion lag limits their utility for prospective studies that need to recruit patients soon after diagnosis. EMRs provide a valuable complement for rapidly identifying potential new diagnoses to subject to verification through chart review and for ascertaining disease progression.

Follow-up time. The ideal patient identifier is unique, universal, and permanent. This project involved two data pulls. Patient ID numbers and date shifting changed between the data pulls at one site. This was able to be fixed but illustrates that vigilance is needed to maintain consistent linkages over time.

Although it is not crucial that patient ID number stays the same across data pulls, it is essential that a mapping be preserved. Ideally, most projects would not involve multiple data cuts.

Algorithms to identify proxies for enrollment in the EMR system are needed in order to be able to determine how much follow-up we have for people. EMR systems typically use encounter-based enrollment algorithms applied to EMR data. However, the classification of ambulatory visits vs. other ambulatory visits in the CDM is not clear. A number of people were identified who only had "other ambulatory visits" and some of these were encounters after death, suggesting some kind of administrative event. Removing "other ambulatory" encounters did seem to mostly fix the problem of encounters after death dates, but this may have excluded some real encounters too as some "other ambulatory" encounters can potentially be an indicator of "follow-up" (e.g., telephone refills, lab-only encounters).

Data completeness and validity

Several analyses suggest that EMR data at PCORnet sites are highly complete for capturing molecular tumor tests and associated therapies for patients receiving medical oncology care at these large, mostly academic, institutions. In our manual chart reviews at two sites we were able to identify colorectal cancer patients who only came to a PCORnet site for surgery and who would not be expected to have received targeted therapy at the site. Among patients who did receive medical oncology care at PCORnet sites, testing was identified in their PCORnet EMR with rare exception and treatments were completely ascertained. Physician specialty has been added in CDM 4.1 so it should be possible to include in future algorithms to select patients who see a medical oncologist at the PCORnet facility. There is a gap, however, between the information about tests, test results, and treatments noted in various locations in the EMR and the structured/computable data extracted to CDM tables (see below).

Data availability in structured/interoperable form

Case definition. Tumor registries contain the critical information needed to identify cancer-related cohorts and characterize the characteristics of their cancer at the time of diagnosis. Whereas NAACCR-structured variables from hospital tumor registry are valuable for case definition, many targeted therapies are indicated for patients whose cancer has progressed. These patients may have had their initial diagnosis elsewhere or long ago, before the earliest date of available tumor registry records. ICD 9/10 CM codes may be useful to identify such patients but then critical tumor information (e.g. stage, grade, histological type) is not easily available. **These tradeoffs could potentially be quantified by examining the sensitivity and specificity of a strategy that combines tumor registry data and ICD9/10 codes for identifying metastatic cancer cases and treatments.**

Molecular tests ordered. There is variability among molecular tests for how easily they are found with queries of CPT codes in the CDM. For KRAS and BRAF testing, an additional 5-7% of cases with the test were found with manual review of records, but for MSI/MMR testing manual review indicated an additional 52% of total cases had the test, because most of these tests are coded with the nonspecific CPT for immunohistochemistry. When we linked CDM to claims data for a cohort of breast cancer patients, the percent of tested patients increased, indicating the potential to expose more test orders to structured queries through claims data linkage. However this would not solve the challenge of non-specific CPT codes. It should also be noted that we did not evaluate the LAB_RESULT_CM table as for molecular testing information because preliminary analyses suggested test results were rarely incorporated. **Natural language processing or structured pathology data (such as that developed by**

the College of American Pathologists) that includes target-specific names should be added to the data model and the LAB_RESULT_CM table evaluated for fitness of use.

Test results. Although molecular laboratory results are present in the EMR, extraction from the system is complicated by different reporting structures (in-house single gene versus multi-gene panel vs testing performed at outside reference laboratory) but was easily detected through manual record review. The completeness of test result data was reassuring as described above and we were able to determine that all patients who received cetuximab, panitumumab or pembrolizumab did so in accordance with test-result-based treatment guidelines. There is hope on the horizon that discrete test results could be made available for computational extraction. Unless encoded in an accepted terminology (e.g. SNOMED CT as at UNMC), it is likely that computation will be locally variable. At UIOWA for instance, although we were able to computationally extract information from the EMR for all but one colorectal cancer patient, actual KRAS test results were extracted for 65 (76%) of patients tested, NRAS results for 42 (67%) of patients tested, HRAS results for 29 (57%) of patients tested, and BRAF results for 38 (60%) of patients tested. We were able to extract information about MSI/MMR results in all 103 patients tested. The issues in identifying results were largely related to multiple mutation (panel) testing which made it difficult to identify which mutation was present when pulling the data computationally (although this could be solved with text searches for specific variants given more time). Building a systematic query to retrieve this data should be possible with sufficient programming resources and must be considered up front. Without common standards and a CDM table location for storage, these queries would be institution-specific.

Targeted therapies, infused and oral. As described above, targeted therapies were completely ascertained in the EMR for colorectal cancer patients who received medical oncology care. However, some treatment data is not available in structured/interoperable form. With inclusion of medication administration in the new CDM table, it was possible to increase ascertainment of infused therapies in the CDM. With the exception of out-of-system treatment (3 of 22 colorectal cases at UIOWA) which were ascertained via chart review, infused therapies were well-captured by CDM tables in the metastatic colorectal cancer case study. Database query identified 6 of 7 oral regorafenib users.

When we linked CDM to claims data for a cohort of breast cancer patients, the percent of treated patients increased, indicating the potential to expose more treatments to structured queries through claims data linkage. The unexpected finding that both CDM and claims contributed uniquely to treatment capture suggests further investigation to trace individual medications across the various CDM tables and claims would be a worthwhile effort.

Most medications in the CDM were coded with RXCUI codes from RxNorm. **Whereas these could be readily compiled to represent the small number of targeted therapies of interest, there is an urgent need to develop a mapping between RXCUIs and established systemic therapy computable phenotypes based on NDC, ICD-9/10 and CPT/HCPCS codes.** This will be essential for determining line of therapy and time-to-next-treatment measures.

Medications data were frequently found in only one CDM table. Researchers should be advised to build their query to examine every table in which the medications can be found.

Medications were sometimes found in the CDM PRESCRIBING or billed PROCEDURES tables with no record of being administered (not in MED_ADMIN or DISPENSING tables). At other times, medications

were indicated as having been administered (in MED_ADMIN or DISPENSING table) with no record of having been prescribed (not included in the CDM PRESCRIBING table). **Additional work is recommended to examine the individual-level data across these four CDM tables – as well as claims data - to better understand this pattern and help inform future PCORnet studies.**

Outcomes. Local sources of vital status were found to be incomplete. Including PCORnet sites that did not have external sources of survival data would have nearly doubled the estimated median survival time for patients with advanced NSCLC treated with immune checkpoint inhibitors to almost 16 months, considerably outside the range observed in clinical trials. **When possible, PCORnet sites should secure external death data and CDM specifications should require documentation of the last update of all sources of death data.** For Medicare beneficiaries, the addition of Medicare vital status data contributed to identifying an additional 10% of total deaths.

Variability across participating sites

We anticipated that ability to identify tests, treatments, and outcomes would vary across network sites and that opportunities to improve capabilities would be identified.

- Testing rates varied across study sites from a low of 36% (site E) to a high of 57% (site J), and with diagnosis year (47% in 2013 and 53% in 2016/17).
- Targeted therapy frequency varied across study sites. Sites B and D reported rates of 1% or less whereas the remainder of sites recorded targeted therapies for 3.8% to 8.4% of cancer patients. There were no major differences in cancer type distribution across sites that would explain this, although it is possible that there is a difference in stage, e.g. when someone becomes advanced they refer out, or an accumulation of small differences that could contribute.

A focused follow-up project could conduct targeted data source review to provide deeper insights about medication table usage patterns.

Future Research and Next Steps

This project demonstrated several attributes of PCORnet for studies of cancer molecular tests and associated targeted therapies, including:

- Streamlined data governance allowed data to efficiently be shared in CDM format
- The CDM data were combined with other existing data in the EMR (pathology, mortality, text notes) and tumor registries to extend their value
- CDM data for colorectal cancer patients were linked at two participating institutions to charts and a statewide cancer registry which verified that most testing and treatment was captured in the EMRs of the PCORnet data partners, though sometimes in unstructured format. Test results and molecular-guided therapies were completely ascertained from the EMR and therapy was concordant with treatment guidelines.
- CDM data were linked to Medicare claims for breast cancer patients to support completeness of data capture - this verified the value added by claims data linkage to ascertain structured testing and treatment data.
- The pooled dataset can continue to be analyzed as preliminary data for future research

Recommendations for next steps include:

1. Adopt a common tumor table in the Common Data Model.
2. Add natural language processing or structured pathology data (such as that developed by the College of American Pathologists) that includes target-specific names to the data model and evaluate the LAB_RESULT_CM table for fitness for use.
3. Develop a mapping between RXCUIs and established systemic therapy computable phenotypes based on NDC, ICD-9/10 and CPT/HCPCS codes.
4. Advise researchers to build their query to examine every table in which the medications can be found. Medications data were frequently found in only one CDM table.
5. Examine the individual-level data across the four CDM tables that contain medication information as well as claims data to better understand patterns of agreement among tables.
6. Conduct targeted data source review to provide deeper insights about medication table usage patterns.
7. Examine the sensitivity and specificity of a strategy that combines tumor registry data and ICD9/10 codes for identifying metastatic cancer cases and treatments.
8. Advise PCORnet Network Partners to secure external death data.
9. Require documentation of the last update of all sources of death data in the CDM.

While some limitations were found in completeness of cancer-related data capture in structured and interoperable form in the PCORnet CDM, it is reassuring that the data were quite complete in the underlying EMR when examined in the metastatic colorectal cancer case study and that the proportion of tests and treatments identifiable in structured interoperable form can be extended through claims data linkage. This affords great promise for computer-aided, chart confirmation-supported pragmatic trials and observational research studies.

ADDITIONAL DOCUMENTS

All documents relevant and noteworthy, or required based on your milestone schedule, should be shared.
Specifically, please provide a lay abstract for your Rapid Cycle Research project for sharing to PCORnet.org and Commons.

Lay Abstract

PCORnet has created a Common Data Model so researchers can combine data for many patients from across PCORnet healthcare settings. The objective of this study was to see if the databases created for PCORnet contain genetic lab tests that should be used by doctors to guide treatment selection. Eleven healthcare systems efficiently pooled de-identified data from electronic medical records, hospital bills and hospital cancer registries. Data were combined for 86,154 cancer patients. Testing rates varied across cancer types, were most common with stage IV disease, and varied across participating PCORnet healthcare institutions (from 36% to 57%). Five percent of patients received treatments that the tests are used to select. Chart reviews confirmed that most testing and treatment was captured in the medical record, though sometimes this was in unstructured format that is not searchable by computer. Linking to Medicare showed that adding insurance claims information can increase the amount of testing and treatment data available for computers to search. While there were some limitations to completeness of cancer-related data capture in structured and interoperable form in the PCORnet CDM, it was reassuring that charts fully captured testing and treatment data for patients seeing medical oncologists at PCORnet institutions.

Research Team Members

Distributed Research Network Operations Center		
Michaela Dunn	Shaila Murji	Jessica Sturtevant
Darcy Louzao	Kshema Nagavedu	
Marshfield Clinic Research Institute		
Robert Greenlee	Steffani Roush	Laurel Verhagen
Eric LaRose		
Medical College of Wisconsin		
Kristen Osinski	Alex Stoddard	Brad Taylor
Medical University of South Carolina		
Andrea Abbott	Amanda Cameron	Katie Kirchoff
University of Florida		
Jiang Bian	Marcia Hodge	Kathryn McAuliffe
Tajuana Chisholm	Gigi Lipori	Elizabeth Shenkman
University of Iowa		
Aaron Bossler	Brian Gryzlak	Mary Schroeder
Ryan Carnahan	Amanda Kahl	Nick Smith
Mary Charlton	Brad McDowell	Michael Wright
Elizabeth Chrischilles		
University of Kansas Medical Center		
Daryl Budin	Steve Fennel	Hillary Sandoval
Sravani Chandaka	Tamara McMahon	Russ Waitman
Dan Connolly	Lav Patel	
University of Minnesota		
Bryan Akkerman	Tim Meyer	
University of Missouri		
Vasanthi Mandhadi	Abu Mosa	Danny Myers
Todd McNeeley		
University of Nebraska Medical Center		
W. Scott Campbell	Carol Geary	Yeshwanth Narayana
James R. Campbell	James C. McClay	Jay Pedersen
Apar Ganti		
University of Texas Southwestern Medical Center		
Shiby Antony	Lisha Huang	Phillip Reeder
Lindsay Cowell		
Vanderbilt University Medical Center		
Rebecca Johnston	Christine Micheel	Lucy Wang
Michele Lenoue-Newton	Joey Schneider	Shu Xu
Mia Levy		

Additional Documents Tables

pcori

Additional Documents Table 1a-c. DRN OC Common Data Model queries 1-3

Table 1a. Patient Characteristics for DRN OC Common Data Model Query#1 Based on individuals who received care during 1/1/2015 to 12/31/2017 at any of the 67 Network Partner institutions that contributed data available as of March 29, 2018.[a]

	Breast		Colorectal		Esophageal		Lung		Melanoma		Ovarian		Pancreatic		Prostate	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Total N	404,319		181,872		33,645		217,778		150,016		46,944		67,215		391,082	
By Sex (N, % of Condition)																
Male	4,451	1%	103,073	57%	27,033	80%	119,952	55%	87,231	58%	50	0%	37,894	56%	390,346	100%
Female	399,775	99%	78,779	43%	6,574	20%	97,776	45%	62,730	42%	46,806	100%	29,275	44%	564	0%
Other*	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
By Age (N, % of Condition)																
21-44	41,946	10%	12,640	7%	980	3%	5,827	3%	17,949	12%	7,315	16%	2,566	4%	1,695	0%
45-64	189,700	47%	71,447	39%	12,682	38%	73,765	34%	54,431	36%	21,190	45%	23,802	35%	112,383	29%
65-84	155,144	38%	81,884	45%	17,959	53%	123,855	57%	67,183	45%	16,851	36%	36,166	54%	242,244	62%
85+	17,494	4%	15,825	9%	1,791	5%	14,240	7%	10,353	7%	1,384	3%	4,555	7%	34,630	9%
Hispanic (N, % of Condition)																
Yes	27,266	7%	12,712	7%	1,357	4%	8,878	4%	2,942	2%	3,702	8%	4,025	6%	20,132	5%
No	273,216	68%	128,710	71%	24,928	74%	159,236	73%	106,513	71%	31,959	68%	48,600	72%	294,798	75%
Other*	103,802	26%	40,381	22%	7,143	21%	49,588	23%	40,386	27%	11,098	24%	14,380	21%	76,079	19%
Race (N, % of Condition)																
American Indian or Alaska Native	1,290	0%	714	0%	83	0%	719	0%	154	0%	101	0%	171	0%	1,270	0%
Asian	16,562	4%	5,693	3%	497	1%	6,643	3%	453	0%	1,636	3%	1,911	3%	6,200	2%
Black/African American	42,527	11%	20,455	11%	2,368	7%	23,768	11%	1,149	1%	4,330	9%	7,589	11%	57,822	15%
Native Hawaiian or Pacific Islander	1,127	0%	338	0%	12	0%	395	0%	48	0%	71	0%	90	0%	490	0%
White	287,951	71%	130,330	72%	26,663	79%	162,345	75%	133,560	89%	33,946	72%	49,024	73%	277,080	71%
Other*	54,534	13%	23,975	13%	3,558	11%	23,513	11%	14,105	9%	6,403	14%	7,965	12%	47,837	12%

[a] Includes 'new' diagnoses in the query period 1/1/2015 - 12/31/2017 defined as no diagnosis code for the specified cancer in the 1 to 730 days prior to index date. Individuals who received care at more than one Network Partner during the period would be counted once per Network Partner visit, leading to the potential for double-counting.

Table 1b (Query #2): Molecular Tests Among Patients with New Cancer Diagnosis (Eight Cancer Types in Table 1a) from 1/1/2015 – 12/31/2017 based on individuals who received care at any of 62 Network Partner institutions that contributed data by June 19, 2018.		
	Number of patients	%
Total N	1,096,991	100%
By Sex (N, % of patients)		
Female	594,700	54%
Male	502,219	46%
Other*	17	0%
By Age (N, % of patients)		
21-44	78,022	7%
45-64	442,384	40%
65-84	514,917	47%
85+	61,650	6%
By Hispanic (N, % of patients)		
Yes	59,245	5%
No	771,750	70%
Other*	265,956	24%
By Race (N, % of patients)		
Black/African American	108,271	10%
White	788,034	72%
Other*	200,669	18%
By Tests (N, % of patients) - No Cancer 2 years prior and presence of test 1 year after index event		
BRCA	3,908	0%
BRAF	6,663	1%
EGFR	7,899	1%
FISH	33,757	3%
Genetic	4,112	0%
Genetic Breast	126	0%
GSP	3,763	0%
GSP Hema	86	0%
HER2	2,650	0%
IHC	159,685	15%
IHC Breast	331	0%
IHC Tumor	67,686	6%
KIT	188	0%
KRAS	6,064	1%
MGMT	57	0%
NOS	9,869	1%
NRAS	459	0%
PDGFRA	157	0%
Urovysion	527	0%

Table 1c: Aggregate Counts of Any Cancer and Medications Across All Responding Sites	Cancer Diagnosis from 1/1/2015-12/31/2017	
	N	%
Total N	1,596,945	100 %
By Sex (N, % of patients)		
Female	884,899	56%
Male	711,967	45%
Other*	29	0%
By Age (N, % of patients)		
21-44	102,568	7%
45-64	632,610	40%
65-84	771,349	49%
85+	90,393	6%
By Hispanic (N, % of patients)		
Yes	79,246	5%
No	1,198,596	76%
Other*	319,092	20%
By Race (N, % of patients)		
Black/African American	167,999	11%
White	1,195,558	76%
Other*	233,381	15%
By Medication (N, % of patients) - Cancer and presence of medication during 1/1/2015 - 12/31/2017		
Afantinib - Prescribing	1,271	0%
Cetuximab - Prescribing	1,216	0%
Erlotinib - Prescribing	3,547	0%
Nivolumab - Prescribing	2,934	0%
Pertuzumab - Prescribing	2,083	0%
Trastuzumab - Prescribing	4,888	0%
Afantinib - Dispensing	174	0%
Cetuximab - Dispensing	116	0%
Erlotinib - Dispensing	912	0%
Nivolumab - Dispensing	508	0%
Pertuzumab - Dispensing	334	0%
Trastuzumab - Dispensing	758	0%
Cetuximab - Procedure	1,036	0%
Nivolumab - Procedure	2,496	0%
Pertuzumab - Procedure	3,012	0%
Trastuzumab - Procedure	6,471	0%

*"Other" includes all remaining PCORnet CDM values not specified by this request

Additional Documents Table 2. Patient characteristics (Aim 1)

Patient Characteristic	Value	Count	Percent
Age at Diagnosis	<20 years	1575	1.8
	20-29 years	1802	2.1
	30-39 years	4275	5.0
	40-49 years	9300	10.8
	50-59 years	21116	24.5
	60-69 years	26362	30.6
	70-79 years	15696	18.2
	80+ years	6028	7.0
Sex	Female	44588	51.8
	Male	41538	48.2
	Other/Unknown	28	0.0
Race	Black	8337	9.7
	White	74449	86.4
	Other	2524	2.9
	Unknown	844	1.0
Ethnicity	Spanish/Hispanic	2925	3.4
	Not Hispanic	81972	95.1
	Unknown	1257	1.5
Stage at Diagnosis	Stage 0	4336	5.0
	Stage 1	20942	24.3
	Stage 2	13708	15.9
	Stage 3	11111	12.9
	Stage 4	13196	15.3
	Stage Occult	55	0.1
	Stage Unknown	18104	21.0
	Not applicable	4702	5.5
Tumor Site	Bones and Joints	674	0.8
	Brain and Other Nervous System		
	Brain	2598	3.0
	Cranial Nerves Other Nervous System	193	0.2
	Breast	14049	16.3
	Digestive System		
	Anus, Anal Canal and Anorectum	415	0.5
	Colorectal	5350	6.2
	Esophagus	1120	1.3
	Gallbladder	270	0.3
	Intrahepatic Bile Duct	411	0.5

Patient Characteristic	Value	Count	Percent
	Liver	2839	3.3
	Other Biliary	518	0.6
	Other Digestive Organs	104	0.1
	Pancreas	3365	3.9
	Peritoneum, Omentum and Mesentery	165	0.2
	Retroperitoneum	168	0.2
	Small Intestine	532	0.6
	Stomach	1210	1.4
	Endocrine System		
	Other Endocrine including Thymus	276	0.3
	Thyroid	3656	4.2
	Eye and Orbit	570	0.7
	Female Genital System		
	Cervix Uteri	1141	1.3
	Corpus Uteri	3861	4.5
	Other Female Genital Organs	221	0.3
	Ovary	1366	1.6
	Uterus, NOS	112	0.1
	Vagina	64	0.1
	Vulva	460	0.5
	Male Genital System		
	Other Male Genital Organs	29	0.0
	Penis	182	0.2
	Prostate	8879	10.3
	Testis	544	0.6
	Oral Cavity and Pharynx		
	Floor of Mouth	294	0.3
	Gum and Other Mouth	754	0.9
	Hypopharynx	165	0.2
	Lip	167	0.2
	Nasopharynx	160	0.2
	Oropharynx	143	0.2
	Other Oral Cavity and Pharynx	63	0.1
	Salivary Gland	414	0.5
	Tongue	1548	1.8
	Tonsil	815	0.9
	Respiratory System		
	Larynx	1142	1.3

Patient Characteristic	Value	Count	Percent
	Lung and Bronchus	8789	10.2
	Nose, Nasal Cavity and Middle Ear	405	0.5
	Pleura	137	0.2
	Trachea, Mediastinum and Other Respiratory Organs	82	0.1
	Skin excluding Basal and Squamous		
	Melanoma of the Skin	4711	5.5
	Other Non-Epithelial Skin	843	1.0
	Soft Tissue including Heart	1767	2.1
	Urinary System		
	Kidney and Renal Pelvis	4477	5.2
	Other Urinary Organs	95	0.1
	Ureter	136	0.2
	Urinary Bladder	2579	3.0
	Other	1126	1.3
Molecular Tests Ordered	BRAF	2054	2.4
	BRCA	27	0.0
	EGFR	1682	2.0
	FISH	8167	9.5
	Genomic Sequence Analysis Panel	951	1.1
	Genetic	1376	1.6
	HER2	523	0.6
	Immunohistochemistry	41606	48.3
	KIT	15	0.0
	KRAS	1568	1.8
	MGMT	160	0.2
	NRAS	23	0.0
	PDGFRA	12	0.0
	FISH or immunohistochemistry, no specific molecular test	38611	44.8
	Any test except "Genetic"	43395	50.4
Targeted Therapies [a]	ALK inhibitor	114	0.1
	EGFR kinase inhibitor	453	0.5
	EGFR MAB inhibitor	586	0.7
	ERBB2 kinase inhibitor	56	0.1
	ERBB2 mab inhibitor	1571	1.8
	PARP inhibitor	44	0.1
	BRAF inhibitor	203	0.2
	PDL1 inhibitor	1547	1.8
	Any targeted molecular therapy, all data sources	4289	5.0

Patient Characteristic	Value	Count	Percent
	None	81865	

[a] Targeted therapies include: ALK inhibitors alectinib, ceritinib, crizotinib; EGFR kinase inhibitors afatinib, erlotinib, gefitinib, osimertinib; EGFR monoclonal antibody (MAB) inhibitors cetuximab, panitumumab; ERBB2 kinase inhibitor lapatinib; ERBB2 MAB inhibitors ado-trastuzumab, pertuzumab, trastuzumab; PARP inhibitors olaparib, rucaparib; BRAF inhibitors dabrafenib, vemurafenib, cobimetinib, trametinib; PD(L)1 inhibitors pembrolizumab, nivolumab.

pcori

Additional Documents Table 3. Frequency of molecular biomarker testing for actionable targets among patients with single primary solid tumors (Aim 1)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
All Patients	2054 (2.38%)	27 (0.03%)	1682 (1.95%)	8167 (9.48%)	959 (1.11%)	1376 (1.60%)	523 (0.61%)	41606 (48.29%)	15 (0.02%)	1568 (1.82%)	160 (0.19%)	23 (0.03%)	12 (0.01%)	38610 (44.82%)	43395 (50.37%)
Age at Diagnosis															
<20 years	61 (3.87%)	0 (0.00%)	16 (1.02%)	222 (14.10%)	17 (1.08%)	15 (0.95%)	0 (0.00%)	776 (49.27%)	1-10 patients	17 (1.08%)	1-10 patients	1-10 patients	0 (0.00%)	723 (45.90%)	800 (50.79%)
20-29 years	66 (3.66%)	1-10 patients	24 (1.33%)	176 (9.77%)	11 (0.61%)	27 (1.50%)	1-10 patients	923 (51.22%)	0 (0.00%)	36 (2.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	859 (47.67%)	954 (52.94%)
30-39 years	129 (3.02%)	1-10 patients	38 (0.89%)	443 (10.36%)	42 (0.98%)	75 (1.75%)	30 (0.70%)	2257 (52.80%)	0 (0.00%)	70 (1.64%)	13 (0.30%)	1-10 patients	0 (0.00%)	2132 (49.87%)	2364 (55.30%)
40-49 years	273 (2.94%)	1-10 patients	129 (1.39%)	1078 (11.59%)	115 (1.24%)	189 (2.03%)	96 (1.03%)	5039 (54.18%)	1-10 patients	198 (2.13%)	16 (0.17%)	1-10 patients	1-10 patients	4658 (50.09%)	5248 (56.43%)
50-59 years	498 (2.36%)	1-10 patients	427 (2.02%)	2073 (9.82%)	225 (1.07%)	410 (1.94%)	150 (0.71%)	10379 (49.15%)	1-10 patients	434 (2.06%)	51 (0.24%)	1-10 patients	1-10 patients	9537 (45.16%)	10770 (51.00%)
60-69 years	568 (2.15%)	1-10 patients	558 (2.12%)	2311 (8.77%)	329 (1.25%)	395 (1.50%)	154 (0.58%)	12172 (46.17%)	1-10 patients	485 (1.84%)	47 (0.18%)	1-10 patients	1-10 patients	11288 (42.82%)	12720 (48.25%)
70-79 years	338 (2.15%)	1-10 patients	385 (2.45%)	1396 (8.89%)	185 (1.18%)	203 (1.29%)	65 (0.41%)	7351 (46.83%)	1-10 patients	258 (1.64%)	18 (0.11%)	1-10 patients	1-10 patients	6842 (43.59%)	7697 (49.04%)
80+ years	121 (2.01%)	1-10 patients	105 (1.74%)	468 (7.76%)	35 (0.58%)	62 (1.03%)	22 (0.36%)	2709 (44.94%)	1-10 patients	70 (1.16%)	1-10 patients	1-10 patients	1-10 patients	2571 (42.65%)	2842 (47.15%)
Sex															
Female	923 (2.07%)	26 (0.06%)	824 (1.85%)	5379 (12.06%)	445 (1.00%)	883 (1.98%)	520 (1.17%)	24841 (55.71%)	1-10 patients	727 (1.63%)	72 (0.16%)	1-10 patients	1-10 patients	23229 (52.10%)	25744 (57.74%)
Male	1130 (2.72%)	1-10 patients	856 (2.06%)	2786 (6.71%)	514 (1.24%)	491 (1.18%)	1-10 patients	16749 (40.32%)	1-10 patients	840 (2.02%)	88 (0.21%)	15 (0.04%)	1-10 patients	15368 (37.00%)	17635 (42.46%)
Other/Unknown	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	16 (57.14%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	13 (46.43%)	16 (57.14%)
Ethnicity															
Not Hispanic	1957 (2.39%)	25 (0.03%)	1624 (1.98%)	7827 (9.55%)	939 (1.15%)	1325 (1.62%)	510 (0.62%)	39430 (48.10%)	14 (0.02%)	1474 (1.80%)	159 (0.19%)	23 (0.03%)	11 (0.01%)	36543 (44.58%)	41146 (50.20%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
Spanish/Hispanic	65 (2.22%)	1-10 patients	29 (0.99%)	246 (8.41%)	11 (0.38%)	37 (1.26%)	1-10 patients	1705 (58.29%)	1-10 patients	66 (2.26%)	1-10 patients	0 (0.00%)	1-10 patients	1622 (55.45%)	1741 (59.52%)
Unknown	32 (2.55%)	1-10 patients	29 (2.31%)	94 (7.48%)	1-10 patients	14 (1.11%)	1-10 patients	471 (37.47%)	0 (0.00%)	28 (2.23%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	445 (35.40%)	508 (40.41%)
Race															
Black	179 (2.15%)	1-10 patients	192 (2.30%)	933 (11.19%)	97 (1.16%)	130 (1.56%)	46 (0.55%)	4215 (50.56%)	1-10 patients	213 (2.55%)	1-10 patients	1-10 patients	1-10 patients	3900 (46.78%)	4386 (52.61%)
Other	51 (2.02%)	0 (0.00%)	50 (1.98%)	255 (10.10%)	23 (0.91%)	32 (1.27%)	1-10 patients	1272 (50.40%)	1-10 patients	37 (1.47%)	1-10 patients	0 (0.00%)	1-10 patients	1204 (47.70%)	1330 (52.69%)
Unknown	11 (1.30%)	1-10 patients	11 (1.30%)	56 (6.64%)	1-10 patients	1-10 patients	1-10 patients	289 (34.24%)	0 (0.00%)	11 (1.30%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	278 (32.94%)	306 (36.26%)
White	1813 (2.44%)	25 (0.03%)	1429 (1.92%)	6923 (9.30%)	835 (1.12%)	1210 (1.63%)	466 (0.63%)	35830 (48.13%)	11 (0.01%)	1307 (1.76%)	155 (0.21%)	18 (0.02%)	1-10 patients	33228 (44.63%)	37373 (50.20%)
Major Cancer Site															
Bones and Joints	1-10 patients	0 (0.00%)	1-10 patients	103 (15.28%)	1-10 patients	0 (0.00%)	0 (0.00%)	298 (44.21%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	306 (45.40%)	317 (47.03%)
Brain and Other Nervous System	269 (9.64%)	0 (0.00%)	20 (0.72%)	293 (10.50%)	57 (2.04%)	13 (0.47%)	0 (0.00%)	1822 (65.28%)	0 (0.00%)	20 (0.72%)	153 (5.48%)	0 (0.00%)	0 (0.00%)	1445 (51.77%)	1900 (68.08%)
Breast	14 (0.10%)	21 (0.15%)	14 (0.10%)	3489 (24.83%)	13 (0.09%)	23 (0.16%)	519 (3.69%)	9950 (70.82%)	0 (0.00%)	14 (0.10%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	9645 (68.65%)	10212 (72.69%)
Digestive System	647 (3.93%)	1-10 patients	157 (0.95%)	703 (4.27%)	226 (1.37%)	807 (4.90%)	1-10 patients	7417 (45.04%)	13 (0.08%)	731 (4.44%)	1-10 patients	1-10 patients	12 (0.07%)	6828 (41.46%)	7916 (48.07%)
Endocrine System	50 (1.27%)	0 (0.00%)	12 (0.31%)	78 (1.98%)	17 (0.43%)	1-10 patients	0 (0.00%)	992 (25.23%)	0 (0.00%)	19 (0.48%)	1-10 patients	0 (0.00%)	0 (0.00%)	964 (24.52%)	1034 (26.30%)
Eye and Orbit	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	152 (26.67%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	142 (24.91%)	158 (27.72%)
Female Genital System	30 (0.42%)	0 (0.00%)	13 (0.18%)	88 (1.22%)	28 (0.39%)	426 (5.90%)	1-10 patients	3902 (54.01%)	0 (0.00%)	17 (0.24%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	3864 (53.48%)	3925 (54.33%)
Male Genital System	1-10 patients	0 (0.00%)	1-10 patients	92 (0.95%)	1-10 patients	1-10 patients	0 (0.00%)	1670 (17.33%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1697 (17.61%)	1711 (17.76%)
Miscellaneous	31 (2.82%)	1-10 patients	42 (3.81%)	112 (10.17%)	16 (1.45%)	12 (1.09%)	1-10 patients	772 (70.12%)	0 (0.00%)	30 (2.72%)	1-10 patients	0 (0.00%)	0 (0.00%)	707 (64.21%)	783 (71.12%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
Oral Cavity and Pharynx	16 (0.35%)	0 (0.00%)	11 (0.24%)	545 (12.05%)	1-10 patients	1-10 patients	0 (0.00%)	2065 (45.66%)	0 (0.00%)	12 (0.27%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	2072 (45.81%)	2097 (46.36%)
Other	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	17 (68.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	17 (68.00%)	17 (68.00%)
Respiratory System	565 (5.35%)	1-10 patients	1358 (12.87%)	2017 (19.11%)	439 (4.16%)	45 (0.43%)	1-10 patients	5800 (54.95%)	0 (0.00%)	661 (6.26%)	1-10 patients	11 (0.10%)	0 (0.00%)	4404 (41.72%)	6265 (59.36%)
Skin excluding Basal and Squamous	392 (7.06%)	1-10 patients	26 (0.47%)	45 (0.81%)	82 (1.48%)	22 (0.40%)	0 (0.00%)	3232 (58.19%)	1-10 patients	37 (0.67%)	0 (0.00%)	1-10 patients	0 (0.00%)	2868 (51.64%)	3336 (60.06%)
Soft Tissue including Heart	12 (0.68%)	0 (0.00%)	1-10 patients	338 (19.13%)	19 (1.08%)	1-10 patients	0 (0.00%)	1075 (60.84%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1109 (62.76%)	1141 (64.57%)
Urinary System	11 (0.15%)	1-10 patients	11 (0.15%)	251 (3.44%)	29 (0.40%)	1-10 patients	0 (0.00%)	2442 (33.51%)	0 (0.00%)	11 (0.15%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	2542 (34.88%)	2583 (35.45%)
Detailed Cancer Site															
Anus, Anal Canal and Anorectum	1-10 patients	0 (0.00%)	1-10 patients	17 (4.10%)	1-10 patients	1-10 patients	0 (0.00%)	165 (39.76%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	163 (39.28%)	175 (42.17%)
Bones and Joints	1-10 patients	0 (0.00%)	1-10 patients	103 (15.28%)	1-10 patients	0 (0.00%)	0 (0.00%)	298 (44.21%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	306 (45.40%)	317 (47.03%)
Brain	263 (10.12%)	0 (0.00%)	20 (0.77%)	284 (10.93%)	56 (2.16%)	13 (0.50%)	0 (0.00%)	1688 (64.97%)	0 (0.00%)	20 (0.77%)	152 (5.85%)	0 (0.00%)	0 (0.00%)	1318 (50.73%)	1765 (67.94%)
Breast	14 (0.10%)	21 (0.15%)	14 (0.10%)	3489 (24.83%)	13 (0.09%)	23 (0.16%)	519 (3.69%)	9950 (70.82%)	0 (0.00%)	14 (0.10%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	9645 (68.65%)	10212 (72.69%)
Cervix Uteri	1-10 patients	0 (0.00%)	1-10 patients	26 (2.28%)	1-10 patients	1-10 patients	0 (0.00%)	484 (42.42%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	483 (42.33%)	485 (42.51%)
Colorectal	553 (10.34%)	1-10 patients	82 (1.53%)	65 (1.21%)	158 (2.95%)	692 (12.93%)	0 (0.00%)	2748 (51.36%)	0 (0.00%)	638 (11.93%)	0 (0.00%)	1-10 patients	0 (0.00%)	2139 (39.98%)	3034 (56.71%)
Corpus Uteri	1-10 patients	0 (0.00%)	1-10 patients	35 (0.91%)	11 (0.28%)	413 (10.70%)	1-10 patients	2192 (56.77%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	2178 (56.41%)	2197 (56.90%)
Cranial Nerves Other Nervous System	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	134 (69.43%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	127 (65.80%)	135 (69.95%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
Esophagus	12 (1.07%)	0 (0.00%)	1-10 patients	174 (15.54%)	1-10 patients	14 (1.25%)	0 (0.00%)	570 (50.89%)	0 (0.00%)	12 (1.07%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	599 (53.48%)	619 (55.27%)
Eye and Orbit	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	152 (26.67%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	142 (24.91%)	158 (27.72%)
Floor of Mouth	0 (0.00%)	0 (0.00%)	0 (0.00%)	27 (9.18%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	102 (34.69%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	103 (35.03%)	103 (35.03%)
Gallbladder	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	98 (36.30%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	93 (34.44%)	99 (36.67%)
Gum and Other Mouth	1-10 patients	0 (0.00%)	1-10 patients	65 (8.62%)	1-10 patients	0 (0.00%)	0 (0.00%)	255 (33.82%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	253 (33.55%)	260 (34.48%)
Hypopharynx	0 (0.00%)	0 (0.00%)	0 (0.00%)	14 (8.48%)	1-10 patients	0 (0.00%)	0 (0.00%)	66 (40.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	66 (40.00%)	67 (40.61%)
Intrahepatic Bile Duct	1-10 patients	0 (0.00%)	1-10 patients	12 (2.92%)	1-10 patients	11 (2.68%)	0 (0.00%)	213 (51.82%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	210 (51.09%)	223 (54.26%)
Kidney and Renal Pelvis	1-10 patients	1-10 patients	1-10 patients	86 (1.92%)	13 (0.29%)	1-10 patients	0 (0.00%)	1768 (39.49%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1779 (39.74%)	1800 (40.21%)
Larynx	1-10 patients	0 (0.00%)	1-10 patients	105 (9.19%)	1-10 patients	1-10 patients	1-10 patients	385 (33.71%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	384 (33.63%)	394 (34.50%)
Lip	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	25 (14.97%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	26 (15.57%)	26 (15.57%)
Liver	1-10 patients	0 (0.00%)	1-10 patients	62 (2.18%)	1-10 patients	1-10 patients	0 (0.00%)	799 (28.14%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	814 (28.67%)	822 (28.95%)
Lung and Bronchus	541 (6.16%)	1-10 patients	1349 (15.35%)	1848 (21.03%)	425 (4.84%)	38 (0.43%)	0 (0.00%)	5049 (57.45%)	0 (0.00%)	651 (7.41%)	1-10 patients	11 (0.13%)	0 (0.00%)	3673 (41.79%)	5496 (62.53%)
Melanoma of the Skin	379 (8.05%)	1-10 patients	25 (0.53%)	37 (0.79%)	82 (1.74%)	21 (0.45%)	0 (0.00%)	2684 (56.97%)	1-10 patients	36 (0.76%)	0 (0.00%)	1-10 patients	0 (0.00%)	2330 (49.46%)	2785 (59.12%)
Miscellaneous	31 (2.82%)	1-10 patients	42 (3.81%)	112 (10.17%)	16 (1.45%)	12 (1.09%)	1-10 patients	772 (70.12%)	0 (0.00%)	30 (2.72%)	1-10 patients	0 (0.00%)	0 (0.00%)	707 (64.21%)	783 (71.12%)
Nasopharynx	1-10 patients	0 (0.00%)	1-10 patients	78 (48.75%)	0 (0.00%)	1-10 patients	0 (0.00%)	102 (63.75%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	106 (66.25%)	108 (67.50%)
Nose, Nasal Cavity and Middle Ear	19 (4.69%)	0 (0.00%)	1-10 patients	45 (11.11%)	1-10 patients	1-10 patients	0 (0.00%)	225 (55.56%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	205 (50.62%)	230 (56.79%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
Oropharynx	0 (0.00%)	0 (0.00%)	0 (0.00%)	18 (12.59%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	82 (57.34%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	84 (58.74%)	84 (58.74%)
Other	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	17 (68.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	17 (68.00%)	17 (68.00%)
Other Biliary	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	203 (39.19%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	200 (38.61%)	213 (41.12%)
Other Digestive Organs	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	77 (74.04%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	68 (65.38%)	78 (75.00%)
Other Endocrine including Thymus	1-10 patients	0 (0.00%)	1-10 patients	43 (15.58%)	1-10 patients	1-10 patients	0 (0.00%)	201 (72.83%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	195 (70.65%)	205 (74.28%)
Other Female Genital Organs	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	128 (57.92%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	125 (56.56%)	130 (58.82%)
Other Male Genital Organs	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	12 (41.38%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	14 (48.28%)	14 (48.28%)
Other Non-Epithelial Skin	13 (1.54%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	548 (65.01%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	538 (63.82%)	551 (65.36%)
Other Oral Cavity and Pharynx	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	31 (49.21%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	31 (49.21%)	31 (49.21%)
Other Urinary Organs	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	42 (44.21%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	38 (40.00%)	44 (46.32%)
Ovary	1-10 patients	0 (0.00%)	1-10 patients	18 (1.32%)	1-10 patients	1-10 patients	0 (0.00%)	853 (62.45%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	842 (61.64%)	862 (63.10%)
Pancreas	19 (0.56%)	0 (0.00%)	20 (0.59%)	29 (0.86%)	22 (0.65%)	21 (0.62%)	0 (0.00%)	1132 (33.64%)	0 (0.00%)	19 (0.56%)	1-10 patients	0 (0.00%)	0 (0.00%)	1113 (33.08%)	1160 (34.47%)
Penis	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	39 (21.43%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	37 (20.33%)	40 (21.98%)
Peritoneum, Omentum and Mesentery	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	110 (66.67%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	106 (64.24%)	111 (67.27%)
Pleura	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	98 (71.53%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	97 (70.80%)	98 (71.53%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
Prostate	1-10 patients	0 (0.00%)	1-10 patients	77 (0.87%)	1-10 patients	1-10 patients	0 (0.00%)	1416 (15.95%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1438 (16.20%)	1449 (16.32%)
Retroperitoneum	1-10 patients	0 (0.00%)	1-10 patients	41 (24.40%)	1-10 patients	0 (0.00%)	0 (0.00%)	111 (66.07%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	113 (67.26%)	117 (69.64%)
Salivary Gland	1-10 patients	0 (0.00%)	1-10 patients	25 (6.04%)	1-10 patients	1-10 patients	0 (0.00%)	192 (46.38%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	191 (46.14%)	201 (48.55%)
Small Intestine	11 (2.07%)	0 (0.00%)	1-10 patients	17 (3.20%)	1-10 patients	15 (2.82%)	1-10 patients	397 (74.62%)	1-10 patients	11 (2.07%)	0 (0.00%)	0 (0.00%)	1-10 patients	383 (71.99%)	406 (76.32%)
Soft Tissue including Heart	12 (0.68%)	0 (0.00%)	1-10 patients	338 (19.13%)	19 (1.08%)	1-10 patients	0 (0.00%)	1075 (60.84%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1109 (62.76%)	1141 (64.57%)
Stomach	13 (1.07%)	0 (0.00%)	1-10 patients	259 (21.40%)	13 (1.07%)	24 (1.98%)	0 (0.00%)	794 (65.62%)	1-10 patients	12 (0.99%)	0 (0.00%)	0 (0.00%)	1-10 patients	827 (68.35%)	859 (70.99%)
Testis	0 (0.00%)	0 (0.00%)	0 (0.00%)	11 (2.02%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	203 (37.32%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	208 (38.24%)	208 (38.24%)
Thyroid	48 (1.31%)	0 (0.00%)	1-10 patients	35 (0.96%)	12 (0.33%)	1-10 patients	0 (0.00%)	791 (21.64%)	0 (0.00%)	16 (0.44%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	769 (21.03%)	829 (22.68%)
Tongue	1-10 patients	0 (0.00%)	1-10 patients	186 (12.02%)	1-10 patients	1-10 patients	0 (0.00%)	704 (45.48%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	705 (45.54%)	708 (45.74%)
Tonsil	1-10 patients	0 (0.00%)	0 (0.00%)	123 (15.09%)	1-10 patients	0 (0.00%)	0 (0.00%)	506 (62.09%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	507 (62.21%)	509 (62.45%)
Trachea, Mediastinum and Other Respiratory Organs	1-10 patients	0 (0.00%)	1-10 patients	15 (18.29%)	1-10 patients	0 (0.00%)	0 (0.00%)	43 (52.44%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	45 (54.88%)	47 (57.32%)
Ureter	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	37 (27.21%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	40 (29.41%)	42 (30.88%)
Urinary Bladder	1-10 patients	0 (0.00%)	1-10 patients	156 (6.05%)	1-10 patients	1-10 patients	0 (0.00%)	595 (23.07%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	685 (26.56%)	697 (27.03%)
Uterus, NOS	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	77 (68.75%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	75 (66.96%)	77 (68.75%)
Vagina	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	25 (39.06%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	23 (35.94%)	29 (45.31%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
Vulva	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	143 (31.09%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	138 (30.00%)	145 (31.52%)
Stage at Diagnosis															
Stage 0	1-10 patients	1-10 patients	0 (0.00%)	121 (2.79%)	1-10 patients	1-10 patients	65 (1.50%)	2054 (47.37%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	2030 (46.82%)	2103 (48.50%)
Stage 1	196 (0.94%)	1-10 patients	206 (0.98%)	1998 (9.54%)	57 (0.27%)	307 (1.47%)	201 (0.96%)	9519 (45.45%)	1-10 patients	157 (0.75%)	0 (0.00%)	1-10 patients	1-10 patients	9348 (44.64%)	9888 (47.22%)
Stage 2	237 (1.73%)	1-10 patients	133 (0.97%)	1375 (10.03%)	67 (0.49%)	182 (1.33%)	170 (1.24%)	5626 (41.04%)	1-10 patients	146 (1.07%)	0 (0.00%)	1-10 patients	0 (0.00%)	5353 (39.05%)	5883 (42.92%)
Stage 3	393 (3.54%)	1-10 patients	234 (2.11%)	909 (8.18%)	118 (1.06%)	235 (2.12%)	58 (0.52%)	5307 (47.76%)	1-10 patients	272 (2.45%)	1-10 patients	0 (0.00%)	0 (0.00%)	4803 (43.23%)	5543 (49.89%)
Stage 4	676 (5.12%)	1-10 patients	841 (6.37%)	1609 (12.19%)	179 (1.36%)	241 (1.83%)	26 (0.20%)	7206 (54.61%)	1-10 patients	768 (5.82%)	1-10 patients	1-10 patients	1-10 patients	6136 (46.50%)	7604 (57.62%)
Stage Occult	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	16 (29.09%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	16 (29.09%)	17 (30.91%)
Stage Unknown	312 (1.72%)	1-10 patients	194 (1.07%)	1687 (9.32%)	476 (2.63%)	370 (2.04%)	1-10 patients	8858 (48.93%)	1-10 patients	162 (0.89%)	43 (0.24%)	13 (0.07%)	1-10 patients	8229 (45.45%)	9244 (51.06%)
Not applicable	236 (5.02%)	0 (0.00%)	73 (1.55%)	464 (9.87%)	60 (1.28%)	37 (0.79%)	1-10 patients	3020 (64.23%)	0 (0.00%)	62 (1.32%)	113 (2.40%)	0 (0.00%)	0 (0.00%)	2695 (57.32%)	3113 (66.21%)
Study Site															
A	231 (1.95%)	0 (0.00%)	216 (1.82%)	838 (7.08%)	0 (0.00%)	55 (0.46%)	523 (4.42%)	5958 (50.34%)	0 (0.00%)	132 (1.12%)	1-10 patients	0 (0.00%)	0 (0.00%)	5117 (43.23%)	6094 (51.49%)
B	62 (1.38%)	12 (0.27%)	61 (1.36%)	233 (5.19%)	0 (0.00%)	26 (0.58%)	0 (0.00%)	2090 (46.56%)	0 (0.00%)	38 (0.85%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	2000 (44.55%)	2141 (47.69%)
C	39 (0.49%)	1-10 patients	241 (3.06%)	1296 (16.43%)	11 (0.14%)	401 (5.08%)	0 (0.00%)	4130 (52.36%)	1-10 patients	47 (0.60%)	43 (0.55%)	1-10 patients	0 (0.00%)	3976 (50.41%)	4343 (55.06%)
D	17 (0.59%)	0 (0.00%)	41 (1.41%)	261 (9.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1403 (48.40%)	0 (0.00%)	13 (0.45%)	1-10 patients	0 (0.00%)	0 (0.00%)	1366 (47.12%)	1430 (49.33%)
E	44 (0.60%)	0 (0.00%)	51 (0.69%)	512 (6.93%)	243 (3.29%)	0 (0.00%)	0 (0.00%)	2561 (34.65%)	0 (0.00%)	71 (0.96%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	2375 (32.14%)	2691 (36.41%)
F	66 (0.89%)	0 (0.00%)	72 (0.98%)	623 (8.44%)	20 (0.27%)	24 (0.33%)	0 (0.00%)	4093 (55.46%)	0 (0.00%)	58 (0.79%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	4054 (54.93%)	4168 (56.48%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
G	300 (3.50%)	1-10 patients	294 (3.43%)	640 (7.48%)	150 (1.75%)	146 (1.71%)	0 (0.00%)	3876 (45.28%)	0 (0.00%)	263 (3.07%)	109 (1.27%)	0 (0.00%)	0 (0.00%)	3329 (38.89%)	4116 (48.08%)
H	47 (0.54%)	1-10 patients	1-10 patients	1319 (15.19%)	227 (2.61%)	63 (0.73%)	0 (0.00%)	4538 (52.27%)	1-10 patients	44 (0.51%)	1-10 patients	1-10 patients	0 (0.00%)	4365 (50.28%)	4700 (54.13%)
I	0 (0.00%)	0 (0.00%)	0 (0.00%)	67 (1.69%)	0 (0.00%)	1-10 patients	0 (0.00%)	2085 (52.66%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	2088 (52.74%)	2088 (52.74%)
J	246 (2.30%)	1-10 patients	136 (1.27%)	802 (7.51%)	0 (0.00%)	1-10 patients	0 (0.00%)	5980 (56.00%)	0 (0.00%)	289 (2.71%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	5697 (53.35%)	6087 (57.00%)
K	1002 (8.09%)	1-10 patients	560 (4.52%)	1576 (12.72%)	308 (2.49%)	654 (5.28%)	0 (0.00%)	4892 (39.48%)	13 (0.10%)	613 (4.95%)	0 (0.00%)	12 (0.10%)	12 (0.10%)	4243 (34.24%)	5537 (44.69%)
Year of Diagnosis															
2013	482 (2.36%)	11 (0.05%)	395 (1.93%)	1994 (9.74%)	55 (0.27%)	260 (1.27%)	401 (1.96%)	9104 (44.49%)	1-10 patients	372 (1.82%)	19 (0.09%)	1-10 patients	0 (0.00%)	8319 (40.65%)	9574 (46.78%)
2014	647 (2.95%)	1-10 patients	568 (2.59%)	2131 (9.73%)	82 (0.37%)	352 (1.61%)	122 (0.56%)	10363 (47.32%)	1-10 patients	516 (2.36%)	25 (0.11%)	1-10 patients	1-10 patients	9620 (43.93%)	10826 (49.44%)
2015	589 (2.57%)	1-10 patients	521 (2.28%)	2091 (9.14%)	280 (1.22%)	367 (1.60%)	0 (0.00%)	11402 (49.84%)	1-10 patients	476 (2.08%)	54 (0.24%)	1-10 patients	1-10 patients	10626 (46.45%)	11835 (51.73%)
2016	269 (1.51%)	1-10 patients	191 (1.07%)	1707 (9.55%)	418 (2.34%)	328 (1.84%)	0 (0.00%)	9194 (51.46%)	1-10 patients	177 (0.99%)	50 (0.28%)	12 (0.07%)	1-10 patients	8645 (48.38%)	9550 (53.45%)
2017	67 (2.20%)	1-10 patients	1-10 patients	244 (8.01%)	124 (4.07%)	69 (2.27%)	0 (0.00%)	1543 (50.67%)	1-10 patients	27 (0.89%)	12 (0.39%)	0 (0.00%)	1-10 patients	1400 (45.98%)	1610 (52.87%)
Among those getting each targeted therapy[a]															
ALK inhibitor	14 (12.28%)	1-10 patients	31 (27.19%)	49 (42.98%)	11 (9.65%)	1-10 patients	0 (0.00%)	82 (71.93%)	0 (0.00%)	16 (14.04%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	43 (37.72%)	87 (76.32%)
EGFR kinase inhibitor	42 (9.27%)	0 (0.00%)	148 (32.67%)	143 (31.57%)	35 (7.73%)	1-10 patients	0 (0.00%)	302 (66.67%)	0 (0.00%)	47 (10.38%)	1-10 patients	1-10 patients	0 (0.00%)	151 (33.33%)	335 (73.95%)
EGFR MAB inhibitor	55 (9.39%)	0 (0.00%)	21 (3.58%)	71 (12.12%)	20 (3.41%)	39 (6.66%)	0 (0.00%)	341 (58.19%)	0 (0.00%)	80 (13.65%)	0 (0.00%)	1-10 patients	0 (0.00%)	269 (45.90%)	373 (63.65%)
ERBB2 kinase inhibitor	1-10 patients	0 (0.00%)	1-10 patients	23 (41.07%)	0 (0.00%)	1-10 patients	1-10 patients	46 (82.14%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	44 (78.57%)	48 (85.71%)

Patient Characteristic	BRAF	BRCA	EGFR	FISH	GSP	Genetic	HER2	Any IHC	KIT	KRAS	MGMT	NRAS	PDGFRA	FISH or IHC only, no specific test	Any test except "Genetic"
ERBB2 mab inhibitor	12 (0.76%)	1-10 patients	14 (0.89%)	539 (34.31%)	1-10 patients	13 (0.83%)	72 (4.58%)	1118 (71.16%)	0 (0.00%)	12 (0.76%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1088 (69.26%)	1181 (75.18%)
PARP inhibitor	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	35 (79.55%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	33 (75.00%)	35 (79.55%)
BRAF inhibitor	91 (44.83%)	0 (0.00%)	14 (6.90%)	16 (7.88%)	23 (11.33%)	1-10 patients	0 (0.00%)	147 (72.41%)	0 (0.00%)	15 (7.39%)	1-10 patients	0 (0.00%)	0 (0.00%)	66 (32.51%)	174 (85.71%)
PDL1 inhibitor	211 (13.64%)	0 (0.00%)	196 (12.67%)	296 (19.13%)	188 (12.15%)	44 (2.84%)	0 (0.00%)	1083 (70.01%)	0 (0.00%)	124 (8.02%)	1-10 patients	1-10 patients	0 (0.00%)	674 (43.57%)	1177 (76.08%)
Any targeted molecular therapy, all sources	374 (8.72%)	1-10 patients	392 (9.14%)	1074 (25.04%)	260 (6.06%)	111 (2.59%)	72 (1.68%)	2931 (68.34%)	0 (0.00%)	274 (6.39%)	1-10 patients	1-10 patients	0 (0.00%)	2223 (51.83%)	3173 (73.98%)

Abbreviations: FISH, fluorescence in situ hybridization; IHC, immunohistochemistry

[a] Targeted therapies include: ALK inhibitors alectinib, ceritinib, crizotinib; EGFR kinase inhibitors afatinib, erlotinib, gefitinib, osimertinib; EGFR monoclonal antibody (MAB) inhibitors cetuximab, panitumumab; ERBB2 kinase inhibitor lapatinib; ERBB2 MAB inhibitors ado-trastuzumab, pertuzumab, trastuzumab; PARP inhibitors olaparib, rucaparib; BRAF inhibitors dabrafenib, vemurafenib, cobimetinib, trametinib; PD(L)1 inhibitors pembrolizumab, nivolumab.

Additional Documents Table 4. Targeted therapies among patients with single primary solid tumors (Aim 1)

	Targeted Molecular Therapy Received [a]								
Value	**Any targeted molecular therapy**	**ALK inhibitor**	**EGFR kinase inhibitor**	**EGFR MAB inhibitor**	**ERBB2 kinase inhibitor**	**ERBB2 mab inhibitor**	**PARP inhibitor**	**BRAF inhibitor**	**PDL1 inhibitor**
All Patients	4289 (4.98%)	114 (0.13%)	453 (0.53%)	586 (0.68%)	56 (0.06%)	1571 (1.82%)	44 (0.05%)	203 (0.24%)	1547 (1.80%)
Age at Diagnosis									
<20 years	28 (1.78%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients
20-29 years	78 (4.33%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	30 (1.66%)	0 (0.00%)	1-10 patients	25 (1.39%)
30-39 years	288 (6.74%)	1-10 patients	1-10 patients	26 (0.61%)	1-10 patients	175 (4.09%)	1-10 patients	25 (0.58%)	68 (1.59%)
40-49 years	643 (6.91%)	21 (0.23%)	39 (0.42%)	74 (0.80%)	11 (0.12%)	351 (3.77%)	1-10 patients	32 (0.34%)	149 (1.60%)
50-59 years	1246 (5.90%)	29 (0.14%)	109 (0.52%)	210 (0.99%)	20 (0.09%)	471 (2.23%)	14 (0.07%)	48 (0.23%)	427 (2.02%)
60-69 years	1245 (4.72%)	32 (0.12%)	153 (0.58%)	169 (0.64%)	12 (0.05%)	382 (1.45%)	16 (0.06%)	45 (0.17%)	521 (1.98%)
70-79 years	607 (3.87%)	1-10 patients	108 (0.69%)	86 (0.55%)	1-10 patients	123 (0.78%)	1-10 patients	26 (0.17%)	286 (1.82%)
80+ years	154 (2.55%)	1-10 patients	27 (0.45%)	11 (0.18%)	1-10 patients	39 (0.65%)	0 (0.00%)	1-10 patients	66 (1.09%)
Sex									
Female	2621 (5.88%)	61 (0.14%)	256 (0.57%)	158 (0.35%)	48 (0.11%)	1473 (3.30%)	36 (0.08%)	84 (0.19%)	631 (1.42%)
Male	1667 (4.01%)	53 (0.13%)	197 (0.47%)	428 (1.03%)	1-10 patients	98 (0.24%)	1-10 patients	119 (0.29%)	915 (2.20%)
Other/Unknown	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Ethnicity									
Not Hispanic	4115 (5.02%)	110 (0.13%)	428 (0.52%)	566 (0.69%)	52 (0.06%)	1489 (1.82%)	42 (0.05%)	196 (0.24%)	1507 (1.84%)
Spanish/Hispanic	126 (4.31%)	1-10 patients	19 (0.65%)	17 (0.58%)	1-10 patients	63 (2.15%)	1-10 patients	1-10 patients	23 (0.79%)
Unknown	48 (3.82%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	19 (1.51%)	0 (0.00%)	1-10 patients	17 (1.35%)
Race									
Black	441 (5.29%)	13 (0.16%)	44 (0.53%)	81 (0.97%)	1-10 patients	185 (2.22%)	1-10 patients	1-10 patients	132 (1.58%)
Other	174 (6.89%)	1-10 patients	39 (1.55%)	16 (0.63%)	1-10 patients	77 (3.05%)	1-10 patients	1-10 patients	37 (1.47%)
Unknown	28 (3.32%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	1-10 patients	1-10 patients	1-10 patients	1-10 patients
White	3646 (4.90%)	95 (0.13%)	368 (0.49%)	485 (0.65%)	44 (0.06%)	1301 (1.75%)	40 (0.05%)	192 (0.26%)	1368 (1.84%)
Major Cancer Site									
Bones and Joints	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Brain and Other Nervous System	62 (2.22%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	13 (0.47%)	37 (1.33%)
Breast	1473 (10.48%)	1-10 patients	0 (0.00%)	0 (0.00%)	44 (0.31%)	1442 (10.26%)	1-10 patients	0 (0.00%)	18 (0.13%)
Digestive System	494 (3.00%)	1-10 patients	49 (0.30%)	202 (1.23%)	1-10 patients	103 (0.63%)	1-10 patients	15 (0.09%)	145 (0.88%)
Endocrine System	17 (0.43%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients

	Targeted Molecular Therapy Received [a]								
Value	**Any targeted molecular therapy**	**ALK inhibitor**	**EGFR kinase inhibitor**	**EGFR MAB inhibitor**	**ERBB2 kinase inhibitor**	**ERBB2 mab inhibitor**	**PARP inhibitor**	**BRAF inhibitor**	**PDL1 inhibitor**
Eye and Orbit	19 (3.33%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	16 (2.81%)
Female Genital System	64 (0.89%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	30 (0.42%)	1-10 patients	26 (0.36%)
Male Genital System	21 (0.22%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients
Miscellaneous	50 (4.54%)	0 (0.00%)	1-10 patients	16 (1.45%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	26 (2.36%)
Oral Cavity and Pharynx	355 (7.85%)	0 (0.00%)	11 (0.24%)	294 (6.50%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	98 (2.17%)
Other	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Respiratory System	1188 (11.26%)	102 (0.97%)	366 (3.47%)	64 (0.61%)	1-10 patients	1-10 patients	1-10 patients	15 (0.14%)	702 (6.65%)
Skin excluding Basal and Squamous	307 (5.53%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	128 (2.30%)	240 (4.32%)
Soft Tissue including Heart	49 (2.77%)	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	42 (2.38%)
Urinary System	183 (2.51%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	175 (2.40%)
Detailed Cancer Site									
Anus, Anal Canal and Anorectum	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Bones and Joints	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Brain	60 (2.31%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	12 (0.46%)	36 (1.39%)
Breast	1473 (10.48%)	1-10 patients	0 (0.00%)	0 (0.00%)	44 (0.31%)	1442 (10.26%)	1-10 patients	0 (0.00%)	18 (0.13%)
Cervix Uteri	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Colorectal	210 (3.93%)	0 (0.00%)	0 (0.00%)	186 (3.48%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	25 (0.47%)
Corpus Uteri	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients
Cranial Nerves Other Nervous System	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Esophagus	67 (5.98%)	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	49 (4.38%)	0 (0.00%)	0 (0.00%)	19 (1.70%)
Eye and Orbit	19 (3.33%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	16 (2.81%)
Floor of Mouth	13 (4.42%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Gallbladder	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients
Gum and Other Mouth	46 (6.10%)	0 (0.00%)	0 (0.00%)	39 (5.17%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	16 (2.12%)

	Targeted Molecular Therapy Received [a]								
Value	**Any targeted molecular therapy**	**ALK inhibitor**	**EGFR kinase inhibitor**	**EGFR MAB inhibitor**	**ERBB2 kinase inhibitor**	**ERBB2 mab inhibitor**	**PARP inhibitor**	**BRAF inhibitor**	**PDL1 inhibitor**
Hypopharynx	22 (13.33%)	0 (0.00%)	0 (0.00%)	18 (10.91%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients
Intrahepatic Bile Duct	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Kidney and Renal Pelvis	155 (3.46%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	147 (3.28%)
Larynx	54 (4.73%)	0 (0.00%)	1-10 patients	40 (3.50%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	20 (1.75%)
Lip	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Liver	36 (1.27%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	34 (1.20%)
Lung and Bronchus	1084 (12.33%)	102 (1.16%)	363 (4.13%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	14 (0.16%)	650 (7.40%)
Melanoma of the Skin	280 (5.94%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	120 (2.55%)	216 (4.59%)
Miscellaneous	50 (4.54%)	0 (0.00%)	1-10 patients	16 (1.45%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	26 (2.36%)
Nasopharynx	19 (11.88%)	0 (0.00%)	0 (0.00%)	14 (8.75%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Nose, Nasal Cavity and Middle Ear	36 (8.89%)	0 (0.00%)	1-10 patients	16 (3.95%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	20 (4.94%)
Oropharynx	12 (8.39%)	0 (0.00%)	1-10 patients	11 (7.69%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Other	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Other Biliary	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Other Digestive Organs	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Other Endocrine including Thymus	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients
Other Female Genital Organs	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients
Other Male Genital Organs	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)
Other Non-Epithelial Skin	27 (3.20%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	24 (2.85%)
Other Oral Cavity and Pharynx	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Other Urinary Organs	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Ovary	38 (2.78%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	27 (1.98%)	1-10 patients	1-10 patients
Pancreas	60 (1.78%)	1-10 patients	39 (1.16%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients	19 (0.56%)

	Targeted Molecular Therapy Received [a]								
Value	**Any targeted molecular therapy**	**ALK inhibitor**	**EGFR kinase inhibitor**	**EGFR MAB inhibitor**	**ERBB2 kinase inhibitor**	**ERBB2 mab inhibitor**	**PARP inhibitor**	**BRAF inhibitor**	**PDL1 inhibitor**
Penis	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Peritoneum, Omentum and Mesentery	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients
Pleura	13 (9.49%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	12 (8.76%)
Prostate	14 (0.16%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients
Retroperitoneum	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Salivary Gland	15 (3.62%)	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients
Small Intestine	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients
Soft Tissue including Heart	49 (2.77%)	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	42 (2.38%)
Stomach	69 (5.70%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	48 (3.97%)	0 (0.00%)	0 (0.00%)	25 (2.07%)
Testis	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Thyroid	12 (0.33%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Tongue	135 (8.72%)	0 (0.00%)	1-10 patients	116 (7.49%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	34 (2.20%)
Tonsil	86 (10.55%)	0 (0.00%)	1-10 patients	79 (9.69%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	15 (1.84%)
Trachea, Mediastinum and Other Respiratory Organs	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Ureter	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Urinary Bladder	21 (0.81%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	21 (0.81%)
Uterus, NOS	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
Vagina	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Vulva	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
Stage at Diagnosis									
Not applicable	132 (2.81%)	1-10 patients	20 (0.43%)	26 (0.55%)	0 (0.00%)	1-10 patients	1-10 patients	16 (0.34%)	76 (1.62%)
Stage 0	14 (0.32%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	1-10 patients
Stage 1	466 (2.23%)	1-10 patients	21 (0.10%)	1-10 patients	1-10 patients	372 (1.78%)	1-10 patients	1-10 patients	61 (0.29%)
Stage 2	624 (4.55%)	1-10 patients	22 (0.16%)	26 (0.19%)	1-10 patients	456 (3.33%)	1-10 patients	12 (0.09%)	107 (0.78%)
Stage 3	630 (5.67%)	1-10 patients	47 (0.42%)	68 (0.61%)	1-10 patients	203 (1.83%)	12 (0.11%)	49 (0.44%)	277 (2.49%)
Stage 4	1440 (10.91%)	72 (0.55%)	277 (2.10%)	371 (2.81%)	31 (0.23%)	163 (1.24%)	19 (0.14%)	88 (0.67%)	573 (4.34%)
Stage Occult	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)

	Targeted Molecular Therapy Received [a]								
Value	**Any targeted molecular therapy**	**ALK inhibitor**	**EGFR kinase inhibitor**	**EGFR MAB inhibitor**	**ERBB2 kinase inhibitor**	**ERBB2 mab inhibitor**	**PARP inhibitor**	**BRAF inhibitor**	**PDL1 inhibitor**
Stage Unknown	982 (5.42%)	24 (0.13%)	65 (0.36%)	84 (0.46%)	1-10 patients	362 (2.00%)	1-10 patients	35 (0.19%)	450 (2.49%)
CDRN Site									
A	996 (8.42%)	16 (0.14%)	86 (0.73%)	146 (1.23%)	1-10 patients	425 (3.59%)	1-10 patients	30 (0.25%)	346 (2.92%)
B	44 (0.98%)	1-10 patients	31 (0.69%)	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)
C	457 (5.79%)	14 (0.18%)	47 (0.60%)	73 (0.93%)	1-10 patients	159 (2.02%)	12 (0.15%)	1-10 patients	165 (2.09%)
D	22 (0.76%)	1-10 patients	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	1-10 patients
E	314 (4.25%)	1-10 patients	1-10 patients	53 (0.72%)	0 (0.00%)	107 (1.45%)	1-10 patients	1-10 patients	155 (2.10%)
F	316 (4.28%)	1-10 patients	36 (0.49%)	63 (0.85%)	1-10 patients	92 (1.25%)	1-10 patients	23 (0.31%)	130 (1.76%)
G	405 (4.73%)	1-10 patients	34 (0.40%)	51 (0.60%)	1-10 patients	80 (0.93%)	1-10 patients	36 (0.42%)	235 (2.75%)
H	417 (4.80%)	22 (0.25%)	47 (0.54%)	38 (0.44%)	1-10 patients	142 (1.64%)	12 (0.14%)	23 (0.26%)	153 (1.76%)
I	207 (5.23%)	1-10 patients	20 (0.51%)	25 (0.63%)	0 (0.00%)	89 (2.25%)	0 (0.00%)	1-10 patients	63 (1.59%)
J	404 (3.78%)	15 (0.14%)	78 (0.73%)	54 (0.51%)	14 (0.13%)	227 (2.13%)	1-10 patients	22 (0.21%)	1-10 patients
K	707 (5.71%)	21 (0.17%)	57 (0.46%)	83 (0.67%)	1-10 patients	246 (1.99%)	1-10 patients	46 (0.37%)	289 (2.33%)
Year of Diagnosis									
2013	946 (4.62%)	28 (0.14%)	137 (0.67%)	182 (0.89%)	20 (0.10%)	393 (1.92%)	1-10 patients	52 (0.25%)	190 (0.93%)
2014	1064 (4.86%)	25 (0.11%)	136 (0.62%)	157 (0.72%)	16 (0.07%)	395 (1.80%)	15 (0.07%)	61 (0.28%)	331 (1.51%)
2015	1186 (5.18%)	30 (0.13%)	107 (0.47%)	138 (0.60%)	1-10 patients	381 (1.67%)	15 (0.07%)	47 (0.21%)	541 (2.36%)
2016	933 (5.22%)	27 (0.15%)	66 (0.37%)	96 (0.54%)	1-10 patients	327 (1.83%)	1-10 patients	41 (0.23%)	422 (2.36%)
2017	160 (5.25%)	1-10 patients	1-10 patients	13 (0.43%)	0 (0.00%)	75 (2.46%)	0 (0.00%)	1-10 patients	63 (2.07%)
Among those with each molecular test type									
BRAF	374 (18.21%)	14 (0.68%)	42 (2.04%)	55 (2.68%)	1-10 patients	12 (0.58%)	1-10 patients	91 (4.43%)	211 (10.27%)
BRCA	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)
EGFR	392 (23.31%)	31 (1.84%)	148 (8.80%)	21 (1.25%)	1-10 patients	14 (0.83%)	1-10 patients	14 (0.83%)	196 (11.65%)
FISH	1074 (13.15%)	49 (0.60%)	143 (1.75%)	71 (0.87%)	23 (0.28%)	539 (6.60%)	1-10 patients	16 (0.20%)	296 (3.62%)
GSP	260 (27.34%)	11 (1.16%)	35 (3.68%)	20 (2.10%)	0 (0.00%)	1-10 patients	1-10 patients	23 (2.42%)	188 (19.77%)
Genetic	111 (8.07%)	1-10 patients	1-10 patients	39 (2.83%)	1-10 patients	13 (0.94%)	1-10 patients	1-10 patients	44 (3.20%)
HER2	72 (13.77%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	72 (13.77%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
IHC	2931 (7.04%)	82 (0.20%)	302 (0.73%)	341 (0.82%)	46 (0.11%)	1118 (2.69%)	35 (0.08%)	147 (0.35%)	1083 (2.60%)
KIT	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)

	Targeted Molecular Therapy Received [a]								
Value	**Any targeted molecular therapy**	**ALK inhibitor**	**EGFR kinase inhibitor**	**EGFR MAB inhibitor**	**ERBB2 kinase inhibitor**	**ERBB2 mab inhibitor**	**PARP inhibitor**	**BRAF inhibitor**	**PDL1 inhibitor**
KRAS	274 (17.47%)	16 (1.02%)	47 (3.00%)	80 (5.10%)	1-10 patients	12 (0.77%)	1-10 patients	15 (0.96%)	124 (7.91%)
MGMT	1-10 patients	0 (0.00%)	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients	1-10 patients
NRAS	1-10 patients	0 (0.00%)	1-10 patients	1-10 patients	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1-10 patients
PDGFRA	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Any test except "Genetic"	3173 (7.31%)	87 (0.20%)	335 (0.77%)	373 (0.86%)	48 (0.11%)	1181 (2.72%)	35 (0.08%)	174 (0.40%)	1177 (2.71%)
FISH or IHC only, no specific test	2223 (5.76%)	43 (0.11%)	151 (0.39%)	269 (0.70%)	44 (0.11%)	1088 (2.82%)	33 (0.09%)	66 (0.17%)	674 (1.75%)

Abbreviations: FISH, fluorescence in situ hybridization; IHC, immunohistochemistry

[a] Targeted therapies include: ALK inhibitors alectinib, ceritinib, crizotinib; EGFR kinase inhibitors afatinib, erlotinib, gefitinib, osimertinib; EGFR monoclonal antibody (MAB) inhibitors cetuximab, panitumumab; ERBB2 kinase inhibitor lapatinib; ERBB2 MAB inhibitors ado-trastuzumab, pertuzumab, trastuzumab; PARP inhibitors olaparib, rucaparib; BRAF inhibitors dabrafenib, vemurafenib, cobimetinib, trametinib; PD(L)1 inhibitors pembrolizumab, nivolumab.

pcori

Additional Documents Table 5. Demographic table of UIOWA and UNMC cohorts (Aim 2)

		IOWA						NEBRASKA (Through5/31/2018)					
		All Cases (N=138)		Stage IV (N=91)		Stage I-III with metastasis (N=47)		All Cases (N=75)		Stage IV (N=45)		Stage I-III with metastasis (N=30)	
		N	%	N	%	N	%	N	%	N	%	N	%
Age at Diagnosis	<50	20	14%	14	15.4	6	12.8	19	25%	15	33%	4	13%
	50-69	86	62%	55	60.4	31	66.0	40	53%	23	51%	17	57%
	70+	32	23%	22	24.2	10	21.3	16	21%	9	20%	7	23%
Sex	Male	81	59%	53	58.2	28	59.6	46	61%	26	58%	20	67%
	Female	57	41%	38	41.8	19	40.4	29	39%	19	42%	10	33%
Race	White	125	91%	82	90.1	43	91.5	66	88%	42	93%	24	80%
	Other	13	9%	9	9.9	4	8.5	9	12%	3	7%	6	20%
Marital Status	Married	74	54%	39	42.9	35	74.5	45	60%	29	64%	16	53%
	Unmarried/Unknown	64	46%	52	57.1	12	25.5	30	40%	16	36%	14	47%
Insurance Status	Private	58	42%	37	40.7	21	44.7	38	51%	24	53%	14	47%
	Medicaid	30	22%	22	24.2	8	17.0	26	35%	12	27%	14	47%
	Medicare	33	24%	23	25.3	10	21.3	3	4%	3	7%	0	0%
	VA/Tricare	-	-	-	-	-	-	3	4%	2	4%	1	3%
	No insurance/Unknown	17	12%	9	9.9	8	17.0	5	7%	4	9%	1	3%
Grade	Well differentiated	6	4%	2	2.2	4	8.5	10	13%	7	16%	3	10%
	Moderately differentiated	74	54%	44	48.4	30	63.8	45	60%	28	62%	17	57%
	Poorly differentiated	26	19%	18	19.8	8	17.0	8	11%	2	4%	6	20%
	Undifferentiated	4	3%	2	2.2	2	4.3	0	0%	0	0%	0	0%
	Unknown	28	20%	25	27.5	3	6.4	12	16%	8	18%	4	13%
Histology	Adenocarcinoma	115	83%	78	85.7	37	78.7	66	88%	40	89%	26	87%
	Adenocarcinoma with mucinous and/or serrated, signet cell, signet ring features	12	9%	5	5.5	7	14.9	6	8%	4	9%	2	7%
	Neuroendocrine	8	6%	6	6.6	2	4.3	2	3%	1	2%	1	3%
	Other	3	2%	2	2.2	1	2.1	1	1%	1	2%	1	3%
Surgery	Yes	71	51%	45	49.5	26	55.3	48	64%	40	89%	8	27%
	No/Unknown	67	49%	46	50.5	21	44.7	27	36%	5	11%	22	73%
Radiation	Yes	32	23%	19	20.9	13	27.7	18	24%	12	27%	6	20%
	No/Unknown	106	77%	72	79.1	34	72.3	57	76%	33	73%	24	80%
Chemotherapy	Yes	111	80%	68	74.7	43	91.5	70	93%	41	91%	29	97%
	No/Unknown	27	20%	23	25.3	4	8.5	5	7%	4	9%	1	3%

Additional Documents Table 6. Patient characteristics in the full breast cancer cohort and in the cohort with linked Medicare claims (Aim 3)

		Full cohort[a]		Medicare-linked cohort[b]	
		N	%	N	%
Sample size		11,124		2,154	
Sex					
	Male	75	0.7	22	1.0
	Female	11,043	99.3	2,131	98.9
Median age (SD)		58 (12.9)		71 (8.9)	
Age group					
	<50 years	2,822	25.4	72	3.3
	50-59 years	3,085	27.7	92	4.3
	60-65 years	1,768	15.9	203	9.4
	66-69 years	1,127	10.1	574	26.6
	70-79 years	1,652	14.9	868	40.3
	80+ years	664	6.0	345	16.0
Site					
	A	4,038	36.3	665	30.9
	B	1,505	13.5	380	17.6
	C	1,805	16.2	439	20.4
	D	716	6.4	171	7.9
	G	772	6.9	122	5.7
	I	1,024	9.2	214	9.9
	J	1,264	11.4	163	7.6
Stage					
	stage 0	1,811	16.3	360	16.7
	stage1	4,353	39.1	1,024	47.5
	stage 2	2,848	25.6	497	23.1
	stage 3	975	8.8	177	8.2
	stage 4	419	3.8	79	3.7

	stage unknown	718	6.5	17	0.8
Class of case					
	Level 1	322	2.9	68	3.2
	Level 2	4,327	38.9	995	46.2
	Level 3	6475	58.2	1091	50.6
Level 1: Initial diagnosis at reporting facility, treatment received elsewhere (NAACCR|610:0)					
Level 2: Initial diagnosis and treatment at reporting facility (NAACCR|610:10-14)					
Level 3: Initial diagnosis elsewhere, treatment at reporting facility (NAACCR|610:20-22)					

[a] Full cohort includes patients with a single primary breast tumor diagnosed during 2011-2015.
[b]Medicare-linked cohort includes patients in the full cohort with A, B, D and no-HMO eligibility.

pcori

CERTIFICATION

This document must be certified by the Principal Investigator and the designated Administrative Official (AO).

Principal Investigator:

☒ I certify that I, as the Principal Investigator, have reviewed and approved this document (and any associated attachments, if applicable).

PI First Name: Russ
PI Last Name: Waitman
Date: September 13, 2018

Administrative Official:

☒ I certify that I, as the designated Administrative Official, have reviewed and approved this document (and any associated attachments, if applicable).

AO First Name: Deborah
AO Last Name: Maloney
Date: September 17, 2018