

# UC Davis

## UC Davis Previously Published Works

### Title

Collaborating on Data, Science, and Infrastructure: The 20-Year Journey of the Cancer Research Network

### Permalink

<https://escholarship.org/uc/item/2hd0473s>

### Journal

Healthcare, 7(1)

### ISSN

2213-0764

### Authors

Doria-Rose, V Paul  
Greenlee, Robert T  
Buist, Diana SM  
et al.

### Publication Date

2019-03-01

### DOI

10.5334/egems.273

Peer reviewed

## REVIEW

# Collaborating on Data, Science, and Infrastructure: The 20-Year Journey of the Cancer Research Network

V. Paul Doria-Rose<sup>1</sup>, Robert T. Greenlee<sup>2</sup>, Diana S. M. Buist<sup>3</sup>, Diana L. Miglioretti<sup>3,4</sup>, Douglas A. Corley<sup>5</sup>, Jeffrey S. Brown<sup>6,7</sup>, Heather A. Clancy<sup>5</sup>, Leah Tuzzio<sup>3</sup>, Lisa M. Moy<sup>5</sup>, Mark C. Hornbrook<sup>8,\*</sup>, Martin L. Brown<sup>1,\*</sup>, Debra P. Ritzwoller<sup>9</sup>, Lawrence H. Kushi<sup>5</sup> and Sarah M. Greene<sup>10</sup>

<sup>1</sup> Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, US

<sup>2</sup> Marshfield Clinic Research Institute, Marshfield, WI, US

<sup>3</sup> Kaiser Permanente Washington Health Research Institute, Seattle, WA, US

<sup>4</sup> University of California Davis School of Medicine, Davis, CA, US

<sup>5</sup> Division of Research, Kaiser Permanente Northern California, Oakland, CA, US

<sup>6</sup> Department of Population Medicine, Harvard Medical School, Boston, MA, US

<sup>7</sup> Harvard Pilgrim Health Care Institute, Boston, MA, US

<sup>8</sup> Center for Health Research, Kaiser Permanente Northwest, Portland, OR, US

<sup>9</sup> Institute for Health Research, Kaiser Permanente Colorado, Aurora, CO, US

<sup>10</sup>Health Care Systems Research Network, Seattle, WA, US

\* Retired

Corresponding author: V. Paul Doria-Rose ([doriarop@mail.nih.gov](mailto:doriarop@mail.nih.gov))

The Cancer Research Network (CRN) is a consortium of 12 research groups, each affiliated with a nonprofit integrated health care delivery system, that was first funded in 1998. The overall goal of the CRN is to support and facilitate collaborative cancer research within its component delivery systems. This paper describes the CRN's 20-year experience and evolution. The network combined its members' scientific capabilities and data resources to create an infrastructure that has ultimately supported over 275 projects. Insights about the strengths and limitations of electronic health data for research, approaches to optimizing multidisciplinary collaboration, and the role of a health services research infrastructure to complement traditional clinical trials and large observational datasets are described, along with recommendations for other research consortia.

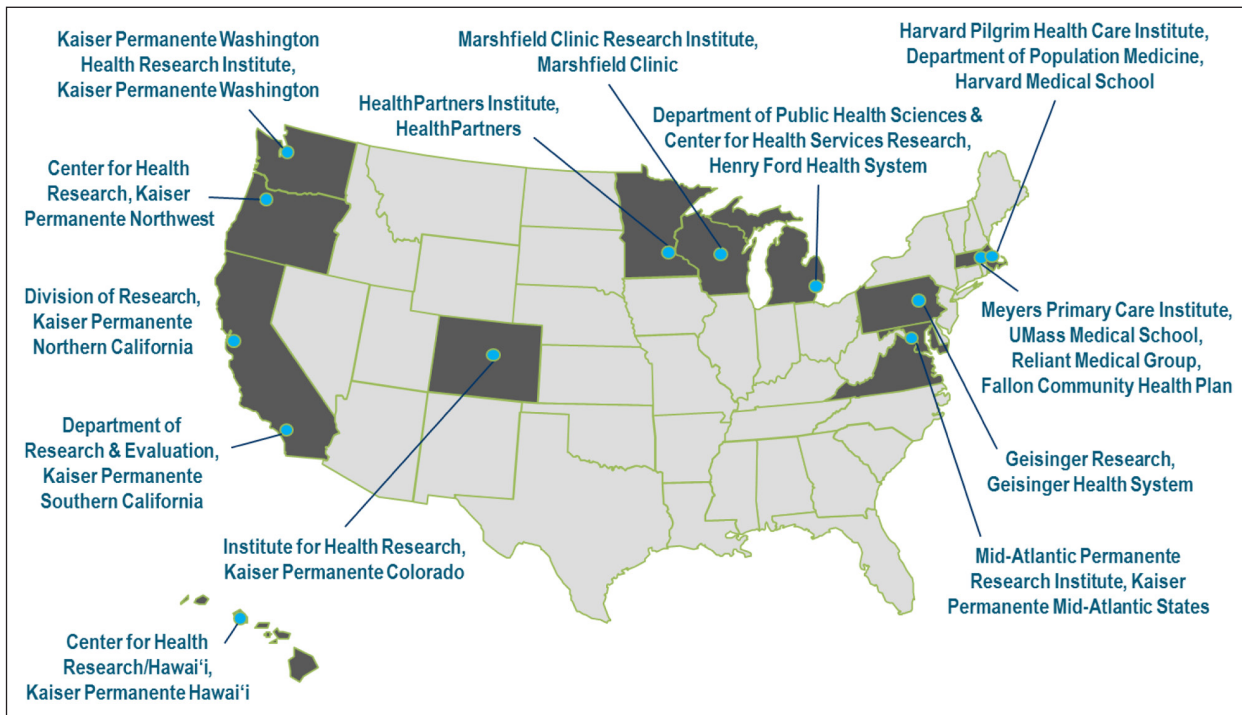
**Keywords:** Electronic Health Records; Delivery of Health Care; Quality Improvement; Health Services Research; Research Network

## Introduction

The Cancer Research Network (CRN) was founded in 1998 to increase the effectiveness of preventive, curative, and supportive interventions for major cancers, through a program of collaborative research based in community health care settings [1–3]. The CRN currently includes 12 research groups, each affiliated with a nonprofit integrated health care delivery system (**Figure 1**); collectively, these systems provide care to more than 12 million members. Member sites also belong to a larger organization of 18 health care delivery systems with in-house research centers known as the Health Care Systems Research Network (HCSRN; formerly Health Maintenance Organization Research Network, HMORN) [4, 5]. HCSRN members have key features in common, including: a cadre of research scientists, comprehensive data resources mapped to a common data model (the Virtual Data Warehouse [VDW]), and a defined health care member population that generally reflects the demographics of their underlying communities. These attributes combine to provide a research resource to support multisite studies. This infrastructure review describes the rationale for, evolution of, and insights derived from the CRN over its 20-year history.

## Rationale for Establishing a Cancer Research Network in Integrated Health Care Delivery Systems

With one out of every three people in the United States developing cancer in their lifetime [6], we need a broad spectrum of approaches to counter the clinical, economic, and psychosocial effects of this pervasive illness. Historically, basic science has driven diagnostic and prognostic advances, and national clinical trials infrastructure (chiefly



**Figure 1:** Cancer Research Network (CRN) Member Sites.

the National Cancer Institute [NCI] National Clinical Trials Network, previously the Cooperative Group Program) [7, 8] has led to rapid progress toward optimal treatments. However, basic science and clinical trials do not fully address the array of questions about primary prevention, screening, treatment, long-term outcomes, and psychosocial sequelae of cancer, especially in community settings. Moreover, many adult cancer clinical trials have dismal accrual rates [7, 9, 10]. Thus, population studies that encompass epidemiologic and health services research, including observational studies and intervention studies outside an oncology clinical trials context, provide critical information about cancer's broader impact and can assess the full range of determinants and consequences of this disease, as well as the experience of cancer patients who do not take part in clinical trials.

Before the CRN was established, the only major source of information about cancer-related health care utilization in the U.S. population was the Surveillance, Epidemiology, and End Results (SEER) cancer registry data linked with Medicare fee-for-service claims [11, 12]. The SEER-Medicare database continues to be a widely-used resource; however, it has several limitations. Notably, SEER-Medicare provides information primarily for Medicare beneficiaries ages 65 years and older. Also, since health care utilization data are mostly based on Medicare claims, SEER-Medicare does not capture many details of care (e.g., results of tissue examinations and laboratory tests, specific chemotherapeutic agents administered and their doses, patient vital signs) or contextual information about the health system setting.

Cognizant of this, the NCI, in partnership with HMORN leaders, created the CRN to generate and provide standardized, interoperable sources of data covering cancer prevention, screening, diagnosis, treatment, survivorship, and end-of-life care across the entire age spectrum for defined populations. Doing research within a health care environment allowed not only for basic descriptive studies, but also for design and testing of interventions to improve how care is delivered. Conducting multisite research in a network also facilitated studying rare cancers by pooling data across multiple health care systems, and natural experiments to assess how differences in patient populations or health system factors influence care delivery.

The organizations that formed the original CRN have remained relatively stable in their structure, defined populations, and delivery system arrangements over the last two decades. Additionally, patients diagnosed with cancer in these health systems tend not to leave the health system to seek care elsewhere upon receiving a cancer diagnosis [13, 14]. Along with population stability, the CRN has exhibited organizational stability, in that the sites developed and continue to maintain integral assets that comprise a robust infrastructure for cancer research. These assets include mature, comprehensive data resources, heterogeneous patient populations, administrative features that facilitate multisite collaboration such as reciprocity between Institutional Review Boards (IRBs) and subcontracting templates, and the opportunity to diffuse results from research into practice more effectively (since researchers are part of the parent health care system). Critically for cancer research, these organizations also have direct access to tumor registry data, either through their own internal registries, or through long-standing partnerships with SEER or state registries. This infrastructure enables projects that could not be readily undertaken in other settings.

## Evolution: The Changing Landscape of Cancer Care Delivery Research

Since the initial funding of the CRN, many advances have altered the landscape of cancer care delivery research. Basic science discoveries have yielded a better understanding of the processes by which genomic damage accumulates during the development of cancer. Subsequently, translational research has leveraged this knowledge to develop new cancer therapies, including drugs which target specific genetic mutations and therapies that activate a patient's immune system to fight cancer [15, 16]. Simultaneously, the development and widespread use of increasingly sophisticated tests for body imaging, cancer screening, and cancer staging have led to major changes in how and when cancers are diagnosed [17]. While in some instances these new screening, diagnostic, and therapeutic procedures have reduced cancer morbidity and mortality, in other cases technologies disseminate into routine care before there is evidence of efficacy (e.g., prostate specific antigen screening [18], robot-assisted surgery [19]). Furthermore, many new therapies are expensive, imposing significant financial burden even among insured patients, and exacerbating health disparities [20, 21]. This trend is likely to continue as development of immunotherapy and biologics matures in the emergent precision medicine era.

These advances in cancer diagnosis and treatment are occurring in a rapidly changing health insurance and health care financing environment. In the last twenty years available health insurance products have become increasingly complex. The development of cost-sharing through high-deductible plans, for example, has increased the absolute costs paid by patients [22, 23]. The passage of the Patient Protection and Affordable Care Act (ACA) of 2010 increased the number of insured individuals in the US by approximately 20 million [24], providing access to care for many who were previously uninsured, and mandated that certain preventive practices and screening tests be available without copayments [25]. The debate around potential modifications or repeal of the ACA, coupled with other proposed changes to federal health insurance programs, have led to a volatile landscape for both health care *and* health coverage, creating uncertainty and unpredictability for many patients undergoing cancer treatment. Consequently, especially for those undergoing cancer treatment, risks of financial toxicity and medical bankruptcy are increasing [26].

Evaluating the impact of the above changes is critical to continued progress in cancer control, particularly through research in community settings where most cancer patients are treated. Networks like the CRN are uniquely poised to examine these changes and evaluate the dissemination and comparative effectiveness of new screening/diagnostic tests and therapies and their impact on patient outcomes, including morbidity and mortality from the disease itself and impact on financial burden and other issues that affect quality of life.

Along with these changes in the health care landscape, concurrent advances in health information technology have facilitated research that taps into the huge volume of health data generated during routine clinical care. When the CRN was initially funded, research often included labor- and resource-intensive medical record abstraction to obtain information from handwritten chart notes. By 2006, however, all participating systems in the CRN had implemented electronic health record systems (EHRs) [27], which facilitated easier access to patient records for research purposes. Rapid growth of EHR systems has facilitated new types of clinical data collection, including patient-reported outcomes obtained via patient portals, and enabled the use of "big data" techniques such as natural language processing and other types of machine learning to extract data from unstructured clinical records in a more automated fashion, though validation of these methods can be complex and time-consuming [28, 29]. More recently, digital health tools including smartphone and tablet apps, social media platforms, and wearable devices represent new frontiers in health care, especially consumer-centered care. These tools have the potential for novel data collection and better patient care, yet would benefit from robust evidence [30].

## Insights: Lessons Learned from 20 Years of the CRN

When the CRN was first created, few other consortia were conducting research in community-based health care settings, and most health services research was based on administrative claims data available from Medicare, Medicaid, or other health insurers. Over time, with increasing availability of health care data, many other research networks have been established. This section describes key insights from the CRN relevant to others who work in large, multicenter/team science settings. These insights apply to three areas: data, science, and culture.

### Data

One key innovation of the CRN was the development of the VDW, a distributed data model in which a set of common data standards (including variable names, definitions, formats, and data structures) is used across participating health systems to facilitate multisite research [1, 31]. This allows each health system to maintain data locally and control data use, but enables a programmer at one site to write an analytic program that can be run at other sites, thereby minimizing or eliminating the need for duplicative effort and ensuring that analyses are consistent between sites. Each system's VDW is populated with provider-, patient-, and encounter-level data extracted from EHRs and administrative and claims databases, as well as information from cancer registries. The VDW contains information about patient demographics, cancer and other diagnoses, health plan enrollment, vital signs, pharmacy dispensing, laboratory orders, utilization, and Census data [1, 31]. With availability of data from EHRs, the VDW now also includes details of cancer infusion therapy data, not previously available in structured formats in community settings. The VDW provides the foundation for much

of CRN's research and has been the basis for data models used in other distributed networks such as the U.S. Food and Drug Administration's Sentinel Initiative [32] and the Patient-Centered Outcomes Research Institute's PCORnet [33–36].

Development of additional tools has enhanced VDW functionality and efficiency. The CRN has a query tool, CRNnet, to facilitate sharing programs and results for research preparation and quality assurance. CRNnet is the CRN's installation of PopMedNet [37, 38], an open-source platform to facilitate multisite research using distributed data. CRNnet's secure distributed querying approach allows authorized users to issue standardized requests for summary-level data; these queries do not require development of SAS code, allowing for rapid responses to simple questions. Another innovation to enhance multisite research was the Cancer Counter [1], a data utility that allows approved users to obtain counts of primary tumors and numbers of cancer patients with specific tumor types using selected variables extracted from tumor registry files maintained by CRN health systems.

Despite these advances, there are important limitations of the distributed data approach. There was a persistent hope (and even misperception) that the CRN would become a turn-key, queryable data resource leveraging a defined population that any approved user could access. The appeal of a point-and-click tool to answer a myriad of health care delivery-related questions was compelling, yet the complexity of curating the data and collating and interpreting results remained a formidable challenge. Underlying differences between health systems—even those on the same EHR platform—necessitate deep local expertise about the system's data, including contextual, temporal, and structural differences. Consequently, initial data extractions must be viewed carefully by users to avoid conducting analyses that yield invalid results.

For example, research on lung cancer screening at Kaiser Permanente Colorado (KPCO) was affected by difficulties in determining the denominator of screen-eligible patients. While documentation of smoking history is considered a vital sign that should be updated at every KPCO outpatient visit, researchers noted inconsistent and incomplete recording of this variable for potential study participants. Almost 20 percent had potentially inaccurate smoking status noted in both the EHR and VDW. Smoking status (including pack-years and time since quitting for former smokers) had not been updated for several outpatient visits, sometimes spanning years. Researchers determined that during clinic visits, clinic staff had not verbally re-assessed smoking status. Instead, EHR variables were “auto-filled” based on values derived from a previous visit. In response to this finding, KPCO researchers worked with clinical and operational leaders to quantify this issue within and across primary care medical offices, then worked with leadership to implement reminders and training within clinics. In this way, researchers' discovery of data errors led to an operational, clinically relevant improvement, while also preventing use of incorrect data for research.

Another example comes from a CRN project evaluating trends in medical imaging utilization over time at seven CRN health care systems [39, 40]. Quality checking of initial data pulled from participating health systems' VDW files using a common SAS program identified several data issues. Undercounting of imaging exams occurred due to variations in local coding practices such as use of local radiology codes that did not reflect the tabulated Current Procedural Terminology (CPT), Healthcare Common Procedure Coding System (HCPCS), or International Classification of Diseases, ninth and tenth revisions (ICD-9 and ICD-10) codes, and/or variations and errors in how sites populated VDW fields. Additionally, overcounting occurred when one examination had different codes with nonmatching dates. (For example, inpatient procedures had ICD codes associated with hospitalization date and CPT codes associated with exam date.) Sites needed to resubmit data multiple times before quality issues were resolved, which was time-consuming. Accounting for the above issues greatly reduced variation in imaging utilization rates between sites. For example, utilization of head CT among enrollees aged 65 years or older initially varied 2.8- to 4.7-fold between sites with the highest versus lowest utilization; after addressing these issues, variation was reduced to a 1.3- to 2.0-fold difference (D. Miglioretti, unpublished data).

A third example comes from a current study by two CRN sites funded by a National Institutes of Health research project grant (R01CA207375). The study is assessing the comparative effectiveness of different clinical work-up strategies for incidentally-detected lung nodules, and includes an assessment of death following nodule detection and workup. On initial analysis, investigators noticed a more than 2-fold difference in mortality rates between two site cohorts using the same distributed VDW code. On additional review of VDW and medical records data, this discrepancy was determined to be due to different definitions of disenrollment dates among deceased health plan members. At one site, death dates generally corresponded to disenrollment dates, whereas at the other site, administrative processing resulted in death dates varying from disenrollment dates by up to 30 days. Initial code written at the first site inadvertently excluded many deceased members from the study cohort at the other site because the program falsely concluded that their disenrollment was not due to death. Adjusting the code to use death date in place of the recorded disenrollment date resulted in equivalent management of subject eligibility at both sites, and corrected the spurious discrepancy between site-specific mortality rates.

These examples highlight why point-and-click querying and automated epidemiology are insufficient for evidence generation and decision-making. Local variations in how data are collected, stored, and processed means that the validity of results from any project depends on repeated cycles of quality assurance and quality improvement, which ultimately requires local data and clinical expertise and validation from medical records. Even with significant investments in data curation, continued maintenance and quality checks are essential for any consortium [41]. Ideally, lessons learned through these quality checks are fed back to VDW site data managers to improve data quality and curation. A

long-time CRN analyst once remarked, “Data get better with use,” which is particularly true in a federated data network. Though the VDW can be an efficient starting point for standardizing analytics (e.g., cohort extraction and aggregation), there is a need to confirm the accuracy of the analytic variables and approach (**Figure 2**).

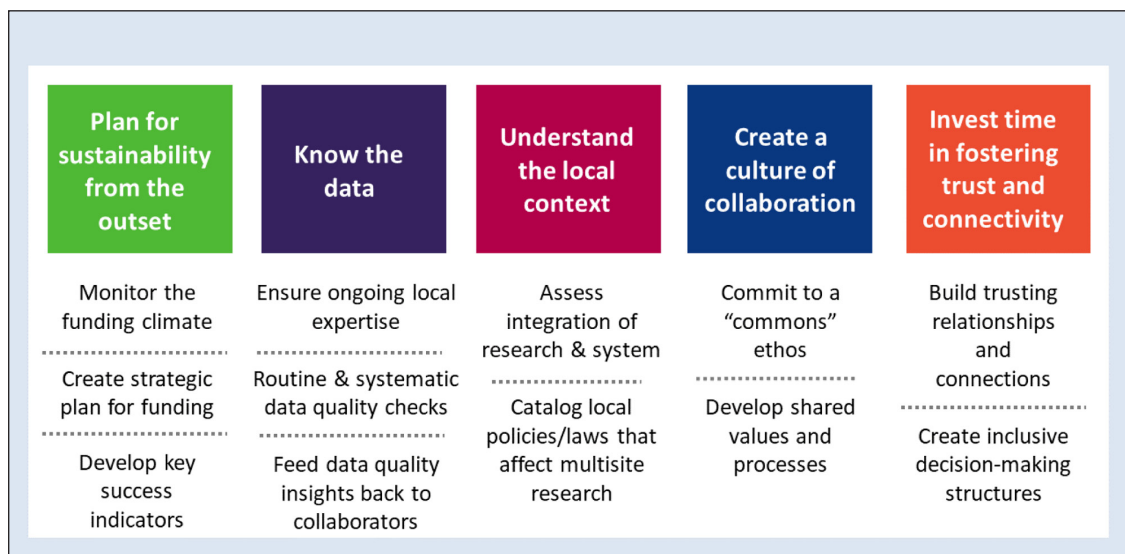
**Science and Clinical Translation**

Differentiating features of a health systems network linked through a collaborative data infrastructure include the ability to analyze a topic from patient, clinical, and system perspectives, and use the infrastructure to examine questions related to methodology, processes, and implementation. For instance, early CRN funding cycles were anchored by multisite studies with multilevel approaches. “HMOs Investigating Tobacco” (HIT), and “Detecting Early Tumors Enhances Cancer Therapy” (DETECT) studied, respectively, the organizational, clinician-level and patient-level factors that contribute to tobacco cessation and late-stage breast and cervical cancer occurrence. The HIT team sought to understand how nationally established tobacco cessation guidelines were implemented in nine CRN health systems, whether clinicians had adequate support for deploying the guidelines, and whether clinician-delivered tobacco cessation information affected patients’ satisfaction with care. The DETECT team sought to identify the extent to which late-stage breast and cervical cancers were attributable to breakdowns during the screening, detection, or treatment phases across six CRN sites. DETECT, in particular, was instrumental not only for clinically relevant findings [42–44] but also for promulgating a conceptual model [45] for studying and intervening on factors that influence cancer screening at multiple levels. As the CRN matured, its portfolio expanded to include affiliated grants and supplemental methodologic studies, and the CRN infrastructure was also leveraged to create new consortia [46–49].

Other studies leveraged the CRN’s large defined population, longitudinal data capabilities, and heterogeneity of the participating health systems to study the experience of cancer care. The CRN is unique in having sufficient data to study special populations, such as older women with breast cancer, a group for whom the most appropriate breast cancer care is still under-examined [50–55]. Similarly, a study of long-term psychosocial outcomes following prophylactic mastectomy benefited from the ability to identify women who received care from their health system for years after their procedure [56–59]. Diverse approaches to diagnosing and treating cancer patients has also enabled comparative effectiveness research evaluating deployment of advanced cancer therapies and associated outcomes [60, 61], and utilization of KRAS and Lynch Syndrome testing to inform metastatic colorectal cancer treatment [62–65].

In a multicenter research environment, the potential exists for inefficiencies and duplicative processes. To alleviate this, CRN investigators documented inefficient processes and created a scientific infrastructure based on a cooperative ethos across the parent HCSRN. An integral example pertained to IRB review of study protocols. The aforementioned study of prophylactic mastectomy included a survey of women at six sites. Review of the same survey protocol across sites resulted in different decisions by the IRBs (e.g., determination of whether the study required full review, expedited review, or was exempt), and also introduced requirements for participant contact that varied by site [66]. Based partly on this experience, the CRN and HCSRN developed a reciprocal reliance model for multisite IRB review designed to mitigate variation and enhance efficiency [66–68].

A key takeaway from the CRN’s experiences addressing pragmatic questions in cancer control is that uptake of findings remains challenging. This is not unique to the CRN, as the research-to-practice gap is a well-worn topic [69–71]. However, since the CRN was in the unique position of being able to study these gaps and pursue potential interventions within health system settings, we hoped that our research would have widespread traction with health system leaders. This was not always the case, however. For example, the aforementioned DETECT study identified multiple opportunities



**Figure 2:** Key considerations for collaborative multisite research.

for reducing late-stage diagnoses of breast and cervical cancer by enhancing patient follow-up and improving screening outreach, yet the health systems were not necessarily resourced or prepared to mobilize robust and rapid programs to address these opportunities. Similarly, the prophylactic mastectomy study identified informational needs for women considering this procedure, but the health systems' competing priorities prevented them from deploying decision aids to address this unmet specialized need. Thus, despite proximity to health systems and applicable findings, research insights were unevenly translated at the system level. Thus, future consortia might consider aligning research questions with system leaders' needs (**Figure 2**), which is consonant with the learning health system model [72, 73].

To increase the chances of translating system-level interventions, planning for implementation and sustainability from the outset is imperative. For example, a study on cancer care quality from the perspective of oncologists, primary care clinicians, health policy researchers, and patients showed that patient navigation—already demonstrated to be effective in the cancer *screening* context—helped attenuate distress and support treatment decision-making during the diagnosis and treatment phases of a cancer patient's journey [74]. This finding prompted a pragmatic trial on the role of nurse navigators for newly diagnosed breast and lung cancer patients. The positive results [75, 76] led to permanent deployment of a nurse navigator in the health system after the study.

### **Culture and Relationships**

These CRN research successes depended on developing a culture of collaboration (**Figure 2**). While critical for any multidisciplinary, team-science endeavor, this culture has unique features for research conducted within health care systems. First, health care system researchers must incorporate studies into the clinical care context in which they are conducted. This requires establishing and maintaining relationships with clinical and operations leaders, both to design study protocols that can be incorporated into clinical workflows with minimal disruptions, and ideally, to ensure that research questions are selected to inform care delivery within the system. Matching research studies to clinical needs can be challenging. Although all health system stakeholders share the goal of improving care, research occurs at a slower pace than most quality improvement efforts by clinical and operations staff. Obtaining research funding is a slow process, hindering ability to address key clinical questions rapidly. Nevertheless, open dialogue between clinical, operations, and research staff can create an environment in which research informs quality improvement efforts and vice versa.

The CRN has used several approaches to enhance relationships between clinical, operations, and research staff. To address the issue of differing time horizons, in 2015, the CRN initiated a set of Rapid Analysis Projects (RAPs) that were intended to be conducted over a 6-month timeframe. To select these projects, a survey was distributed to practicing clinicians within the CRN health systems, asking about evidence gaps that impeded their daily practice. Survey respondents suggested research questions and a prioritization process classified the questions according to both importance and feasibility. A panel of research and clinical stakeholders selected projects based on questions viewed to be both important and addressable using VDW or other readily available data. While IRB approval and other challenges prevented RAPs from being conducted as quickly as desired, processes developed in response to delays are now embedded in CRN practice and should enable future RAP cycles to be conducted at a faster pace.

Another approach to creating an environment that stimulates relationships between clinical, operations, and research staff was the formation of the Translational Research in Oncology (TRIO) group by CRN Kaiser Permanente sites. TRIO's goals include: providing a forum for Kaiser Permanente groups interested in improving cancer care and outcomes, minimizing duplication of analytic activities and bringing varied perspectives to addressing questions of interest, and identifying questions and projects of mutual interest to clinical, operational, and research staff. The group meets through quarterly webinars. Clinicians identify questions that impact patient care and may be amenable to EHR analysis, operations leaders prioritize questions to align with organizational goals and identify challenges and facilitators to improving care across the cancer spectrum, and researchers shape the questions and use them to design studies and analytic approaches.

Secondly, the CRN has cultivated collaborations with researchers at universities and cancer centers outside the CRN network. While CRN health systems have some clinician-researchers, they tend to be the exception rather than the rule, since clinical staff practicing within CRN health systems generally spend the majority of their time on patient care. External collaborations incorporate additional physician-scientists into CRN research programs, expanding the depth and breadth of expertise on research teams. Collaborations were also fostered through the CRN's external advisory committee, which was comprised of leading cancer researchers based in academic institutions whose expertise included behavioral science, geriatric oncology, and clinical epidemiology. External investigators were key contributors to the medical imaging, breast cancer in older women, and nurse navigator projects described above. However, it is important to note that these partnerships can be labor-intensive; orienting outside investigators to the complexities of health systems and EHR data is time-consuming. Thus, strategically engaging investigators who complemented the CRN's internal expertise proved to be the most effective approach to external collaboration.

Finally, building relationships with and training junior investigators has been critical for expanding the CRN's research capacity. The CRN infrastructure provided new opportunities to conduct large population-based research studies within integrated health care systems, but required skills and levels of understanding that are not acquired in current scientific training programs. In response to this training gap, the CRN established the CRN Scholars Program [77, 78] to assist

junior investigators in: developing skills to collaborate with stakeholders in integrated health care systems, developing and using complex multisite data, and working in collaborative research teams anchored in delivery systems. The Scholars Program originated in 2007 as a 22-month training activity designed to help junior investigators at CRN sites develop research independence using CRN data and scientific resources. More recently, the Scholars Program extended to 26 months and broadened its reach to include junior investigators from internal *and* external institutions who are committed to conducting population-based cancer research.

During the program, each Scholar devoted 20 percent effort, receiving one-on-one and peer mentoring, participating in bimonthly webinars, and learning the analytic tools, methods, procedures, and processes for efficiently running multisite CRN studies. The training focus remains on preparing investigators to develop cutting-edge skills for conducting research in integrated health care delivery systems, fulfilling the goal of expanding the cadre of independent investigators committed to performing cancer research within the CRN.

Ultimately, successful collaborative relationships, whether among clinical, operations, and research staff, internal and external investigators, or senior and junior scientists, are enhanced by a commitment to shared scientific goals. A diverse network that encompasses a broad array of interests may experience times where investigators need to support others' efforts on scientific questions that are not directly within their own research scope, applying a "commons" mindset. Adopting shared goals, even broad ones such as increasing adherence to screening or treatment guidelines, enhancing patient satisfaction with their care, or improving efficiency in care delivery, will help assure that priorities among researchers and between researchers and health systems are well-aligned. The CRN's broad mission helped establish such an ethos.

## Conclusion

After continuous funding by the NCI since 1998, the CRN's direct federal grant support will wind down in 2019, although numerous topic-specific collaborations will continue to flourish and perpetuate the CRN's infrastructure and processes. Over 20 years of collaboration, the CRN has accrued important insights related to productive partnerships, opportunities and pitfalls in EHR data, and leveraging health system capabilities for research on improving cancer prevention, care, and control (**Figure 2**). Although we grouped our insights under the headings of data, science, and culture/relationships, these are clearly interdependent in a research network environment. Moreover, we believe many of these insights are pertinent to other multidisciplinary research collaborations, regardless of topical focus.

In the CRN's history, given that studies encompassed prevention and early detection through treatment and end of life care, it was appealing to try to be "all things to all people". We realized over time that this did not play to the CRN's core strengths. For example, we had a diverse patient population cared for in community settings. Furthermore, CRN sites are substantial contributors to participation in oncology clinical trials, including the NCI's Community Oncology Research Program [79]. However, we discerned through trial and error that our researchers' priorities did not always align neatly with health system priorities, which meant that some research questions, despite their broad importance to the oncology community, were not suited to investigation within the CRN. In contrast, we were able to rapidly evaluate natural experiments, for example assessing diffusion of new therapies from controlled clinical trial settings to community practice [80, 81]. We urge other research networks to be thoughtful about the attributes of the participating partners, and align the research agenda to the strengths and priorities of the participants.

While EHRs have facilitated "big data" research, they do not automatically produce research-ready data. The low marks that EHRs have earned for usability and interoperability [82, 83] should also serve as cautions for researchers who use health record data. Our examples illuminate the importance of having project team members with intimate understanding of the underlying data sources who can apply knowledge about data idiosyncrasies in the research context, and compare differences across sites in a research consortium. A recent publication from the PCORnet data operations team [36] elaborates on the remarkable curation, validation and quality checking needed to ensure that data are research-ready. Laudably, the PCORnet [84], Sentinel [85], and HCSRN [86] data teams make their specifications and documentation public to help the research community, and have regular interactions to share lessons learned. However, since data sources are dynamic, it still behooves project teams to include subject matter experts who maintain current knowledge as data evolves.

At the time the CRN was first funded by the NCI in 1998, few others were conducting multi-site research based within health care delivery systems, and capacity-building funds were essential to develop the CRN's research infrastructure. Over time, however, the advent of EHRs and development of new techniques to facilitate data extraction from clinical records for research purposes increased the depth and breadth of research that could be done, and further enabled numerous other health systems to develop their own research capabilities and shared infrastructures. Consequently, the CRN infrastructure grant evolved to emphasize support for collaboration, dissemination of knowledge, and investigator training over data development. In parallel, the need grew for the CRN's component health systems, including their research arms, to explore other ways to maintain support for their research infrastructure. Sustainability of networks such as the CRN is dependent on organizational, sociopolitical, and financial exigencies, and ultimately requires the support and alignment of multiple stakeholders within each health care system. Furthermore, long-term network viability depends on continued evolution, such as promoting stronger connections to data science and informatics, inclusion of patients as part of the fabric of research teams, and attention to dissemination and implementation such



that successful research has a higher potential for adoption. The CRN's experience in developing data resources, creating a research program to improve patient care, and nurturing a culture of collaboration offers a practical roadmap for other research networks and health care organizations who want to use their collective resources and capabilities to improve health outcomes.

### Acknowledgements

The authors wish to thank Dr. Ed Wagner for his many years of leadership and vision in advancing the mission of the CRN. We also thank Chris Tachibana for her editorial assistance in preparing this manuscript.

Support for this publication was provided by the National Cancer Institute of the National Institutes of Health through award number U24CA171524. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### Competing Interests

The authors have no competing interests to declare.

### References

1. **Hornbrook, MC, Hart, G, Ellis, JL**, et al. Building a virtual cancer research organization. *J Natl Cancer Inst Monogr.* 2005; 35: 12–25. DOI: <https://doi.org/10.1093/jncimonographs/lgi033>
2. **Wagner, EH, Greene, SM, Hart, G**, et al. Building a research consortium of large health systems: The Cancer Research Network. *J Natl Cancer Inst Monogr.* 2005; 35: 3–11. DOI: <https://doi.org/10.1093/jncimonographs/lgi032>
3. **Welcome to the Cancer Research Network.** <https://crn.cancer.gov/>. Accessed June 18, 2018.
4. **Vogt, TM, Elston-Lafata, J, Tolsma, D**, et al. The role of research in integrated healthcare systems: The HMO Research Network. *Am J Manag Care.* 2004; 10(9): 643–648.
5. **Welcome to the HCSRN Website (formerly HMORN).** <http://www.hcsrnl.org/en/>. Accessed June 18, 2018.
6. **Siegel, RL, Miller, KD and Jemal, A.** Cancer statistics, 2018. *CA Cancer J Clin.* 2018; 68(1): 7–30. DOI: <https://doi.org/10.3322/caac.21442>
7. **IOM (Institute of Medicine).** 2010. A National Cancer Clinical Trials System for the 21st Century: Reinvigorating the NCI Cooperative Group Program. Washington, DC: The National Academies Press.
8. **An Overview of NCI's National Clinical Trials Network.** <https://www.cancer.gov/research/areas/clinical-trials/nctn>. Accessed June 18, 2018.
9. **Sateren, WB, Trimble, EL, Abrams, J**, et al. How sociodemographics, presence of oncology specialists, and hospital cancer programs affect accrual to cancer treatment trials. *J Clin Oncol.* 2002; 20(8): 2109–2117. DOI: <https://doi.org/10.1200/JCO.2002.08.056>
10. **Somkin, CP, Altschuler, A, Ackerson, L**, et al. Organizational barriers to physician participation in cancer clinical trials. *Am J Manag Care.* 2005; 11(7): 413–421.
11. **Warren, JL, Klabunde, CN, Schrag, D**, et al. Overview of the SEER-Medicare data: Content, research applications, and generalizability to the United States elderly population. *Med Care.* 2002; 40(8 Suppl): IV-3–18. DOI: <https://doi.org/10.1097/00005650-200208001-00002>
12. **SEER-Medicare: Brief Description of the SEER-Medicare Database.** <https://healthcaredelivery.cancer.gov/seermedicare/overview/>. Accessed June 18, 2018.
13. **Chubak, J, Ziebell, R, Greenlee, RT**, et al. The Cancer Research Network: A platform for epidemiologic and health services research on cancer prevention, care, and outcomes in large, stable populations. *Cancer Causes Control.* 2016; 27(11): 1315–1323. DOI: <https://doi.org/10.1007/s10552-016-0808-4>
14. **Field, TS, Cernieux, J, Buist, D**, et al. Retention of enrollees following a cancer diagnosis within health maintenance organizations in the Cancer Research Network. *J Natl Cancer Inst.* 2004; 96(2): 148–152. DOI: <https://doi.org/10.1093/jnci/djh010>
15. **The Future of Cancer Research: Accelerating Scientific Innovation.** President's cancer Panel Annual Report 2010–2011. Bethesda (MD): President's Cancer Panel; 2012 November. A web-based version of this report is available at: <https://deainfo.nci.nih.gov/advisory/pcp/annualReports/pcp10-11rpt/FullReport.pdf>.
16. **Lowy, DR and Collins, FS.** Aiming High—Changing the Trajectory for Cancer. *N Engl J Med.* 2016; 374(20): 1901–1904. DOI: <https://doi.org/10.1056/NEJMp1600894>
17. **Iglehart, JK.** Health insurers and medical-imaging policy—a work in progress. *N Engl J Med.* 2009; 360(10): 1030–1037. DOI: <https://doi.org/10.1056/NEJMp0808703>
18. **Wilt, TJ, Scardino, PT, Carlsson, SV**, et al. Prostate-specific antigen screening in prostate cancer: Perspectives on the evidence. *J Natl Cancer Inst.* 2014; 106(3). DOI: <https://doi.org/10.1093/jnci/dju010>
19. **Barbash, GI and Glied, SA.** New technology and health care costs—the case of robot-assisted surgery. *N Engl J Med.* 2010; 363(8): 701–704. DOI: <https://doi.org/10.1056/NEJMp1006602>
20. **Promoting Value, Affordability, and Innovation in Cancer Drug Treatment.** A Report to the President of the United States from the President's Cancer Panel. Bethesda (MD): President's Cancer Panel; 2018 March. A web-based version of this report is available at: <https://PresCancerPanel.cancer.gov/report/drugvalue>.

21. **Altice, CK, Banegas, MP, Tucker-Seeley, RD**, et al. Financial Hardships Experienced by Cancer Survivors: A Systematic Review. *J Natl Cancer Inst.* 2017; 109(2). DOI: <https://doi.org/10.1093/jnci/djw205>
22. **“Health Policy Brief: High-Deductible Health Plans.”** *Health Affairs*, February 4, 2016.
23. **Cohen, RA and Zammitti, EP.** High-deductible health plans and financial barriers to health care: Early release of estimates from the National Health Interview Survey, 2016. *National Center for Health Statistics*; 2017. Available from: <https://www.cdc.gov/nchs/nhis/releases.htm>.
24. **Cohen, RA, Zammitti, EP and Martinez, ME.** Health insurance coverage: Early release of estimates from the National Health Interview Survey, 2016. *National Center for Health Statistics*, May 2017. Available from: <https://www.cdc.gov/nchs/nhis/releases.htm>.
25. **Sabik, LM and Adunlin, G.** The ACA and Cancer Screening and Diagnosis. *Cancer J.* 2017; 23(3): 151–162. DOI: <https://doi.org/10.1097/PPO.0000000000000261>
26. **Ramsey, SD, Bansal, A, Fedorenko, CR**, et al. Financial Insolvency as a Risk Factor for Early Mortality Among Patients With Cancer. *J Clin Oncol.* 2016; 34(9): 980–986. DOI: <https://doi.org/10.1200/JCO.2015.64.6620>
27. **Scientific & Data Resources.** <https://crn.cancer.gov/resources/>. Accessed June 19, 2018.
28. **Carrell, DS, Halgrim, S, Tran, DT**, et al. Using natural language processing to improve efficiency of manual chart abstraction in research: The case of breast cancer recurrence. *Am J Epidemiol.* 2014; 179(6): 749–758. DOI: <https://doi.org/10.1093/aje/kwt441>
29. **Carrell, DS, Schoen, RE, Leffler, DA**, et al. Challenges in adapting existing clinical natural language processing systems to multiple, diverse health care settings. *J Am Med Inform Assoc.* 2017; 24(5): 986–991. DOI: <https://doi.org/10.1093/jamia/ocx039>
30. **Improving Cancer-Related Outcomes with Connected Health: A Report to the President of the United States from the President’s Cancer Panel.** Bethesda (MD): President’s Cancer Panel; 2016. A web-based version of this report is available at: <https://PresCancerPanel.cancer.gov/report/connectedhealth>.
31. **Ross, TR, Ng, D, Brown, JS**, et al. The HMO Research Network Virtual Data Warehouse: A Public Data Model to Support Collaboration. *EGEMS (Wash DC).* 2014; 2(1): 1049. DOI: <https://doi.org/10.13063/2327-9214.1049>
32. **Curtis, LH, Weiner, MG, Boudreau, DM**, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiol Drug Saf.* 2012; 21(Suppl 1): 23–31. DOI: <https://doi.org/10.1002/pds.2336>
33. **PCORnet Common Data Model (CDM).** <http://pcornet.org/pcornet-common-data-model/>. Accessed June 19, 2018.
34. **Califf, RM.** The Patient-Centered Outcomes Research Network: a national infrastructure for comparative effectiveness research. *N C Med J.* 2014; 75(3): 204–210. DOI: <https://doi.org/10.18043/ncm.75.3.204>
35. **Fleurence, RL, Curtis, LH, Califf, RM**, et al. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc.* 2014; 21(4): 578–582. DOI: <https://doi.org/10.1136/amiajnl-2014-002747>
36. **Qualls, LG, Phillips, TA, Hammill, BG**, et al. Evaluating Foundational Data Quality in the National Patient-Centered Clinical Research Network (PCORnet(R)). *EGEMS (Wash DC).* 2018; 6(1): 3. DOI: <https://doi.org/10.5334/egems.199>
37. **PopMedNet.** <https://www.popmednet.org/>. Accessed June 19, 2018.
38. **Davies, M, Erickson, K, Wyner, Z**, et al. Software-Enabled Distributed Network Governance: The PopMedNet Experience. *EGEMS (Wash DC).* 2016; 4(2): 1213. DOI: <https://doi.org/10.13063/2327-9214.1213>
39. **Miglioretti, DL, Johnson, E, Williams, A**, et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. *JAMA Pediatr.* 2013; 167(8): 700–707. DOI: <https://doi.org/10.1001/jamapediatrics.2013.311>
40. **Smith-Bindman, R, Miglioretti, DL, Johnson, E**, et al. Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996–2010. *JAMA.* 2012; 307(22): 2400–2409. DOI: <https://doi.org/10.1001/jama.2012.5960>
41. **Brown, JS, Kahn, M and Toh, S.** Data quality assessment for comparative effectiveness research in distributed data networks. *Med Care.* 2013; 51(8 Suppl 3): S22–29. DOI: <https://doi.org/10.1097/MLR.0b013e31829b1e2c>
42. **Leyden, WA, Manos, MM, Geiger, AM**, et al. Cervical cancer in women with comprehensive health care access: attributable factors in the screening process. *J Natl Cancer Inst.* 2005; 97(9): 675–683. DOI: <https://doi.org/10.1093/jnci/dji115>
43. **Taplin, SH, Ichikawa, L, Yood, MU**, et al. Reason for late-stage breast cancer: Absence of screening or detection, or breakdown in follow-up? *J Natl Cancer Inst.* 2004; 96(20): 1518–1527. DOI: <https://doi.org/10.1093/jnci/djh284>
44. **Zapka, JG, Puleo, E, Taplin, SH**, et al. Processes of care in cervical and breast cancer screening and follow-up—the importance of communication. *Prev Med.* 2004; 39(1): 81–90. DOI: <https://doi.org/10.1016/j.ypmed.2004.03.010>
45. **Zapka, JG, Taplin, SH, Solberg, LI**, et al. A framework for improving the quality of cancer care: The case of breast and cervical cancer screening. *Cancer Epidemiol Biomarkers Prev.* 2003; 12(1): 4–13.
46. **Coleman, KJ, Stewart, C, Waitzfelder, BE**, et al. Racial-Ethnic Differences in Psychiatric Diagnoses and Treatment Across 11 Health Care Systems in the Mental Health Research Network. *Psychiatr Serv.* 2016; 67(7): 749–757. DOI: <https://doi.org/10.1176/appi.ps.201500217>

47. **Magid, DJ, Gurwitz, JH, Rumsfeld, JS**, et al. Creating a research data network for cardiovascular disease: The CVRN. *Expert Rev Cardiovasc Ther.* 2008; 6(8): 1043–1045. DOI: <https://doi.org/10.1586/14779072.6.8.1043>
48. **Malin, JL, Ko, C, Ayanian, JZ**, et al. Understanding cancer patients' experience and outcomes: Development and pilot study of the Cancer Care Outcomes Research and Surveillance patient survey. *Support Care Cancer.* 2006; 14(8): 837–848. DOI: <https://doi.org/10.1007/s00520-005-0902-8>
49. **Tisminetzky, M, Bayliss, EA, Magaziner, JS**, et al. Research Priorities to Advance the Health and Health Care of Older Adults with Multiple Chronic Conditions. *J Am Geriatr Soc.* 2017; 65(7): 1549–1553. DOI: <https://doi.org/10.1111/jgs.14943>
50. **Bosco, JL, Lash, TL, Prout, MN**, et al. Breast cancer recurrence in older women five to ten years after diagnosis. *Cancer Epidemiol Biomarkers Prev.* 2009; 18(11): 2979–2983. DOI: <https://doi.org/10.1158/1055-9965.EPI-09-0607>
51. **Buist, DS, Chubak, J, Prout, M**, et al. Referral, receipt, and completion of chemotherapy in patients with early-stage breast cancer older than 65 years and at high risk of breast cancer recurrence. *J Clin Oncol.* 2009; 27(27): 4508–4514. DOI: <https://doi.org/10.1200/JCO.2008.18.3459>
52. **Enger, SM, Thwin, SS, Buist, DS**, et al. Breast cancer treatment of older women in integrated health care settings. *J Clin Oncol.* 2006; 24(27): 4377–4383. DOI: <https://doi.org/10.1200/JCO.2006.06.3065>
53. **Field, TS, Doubeni, C, Fox, MP**, et al. Under utilization of surveillance mammography among older breast cancer survivors. *J Gen Intern Med.* 2008; 23(2): 158–163. DOI: <https://doi.org/10.1007/s11606-007-0471-2>
54. **Geiger, AM, Thwin, SS, Lash, TL**, et al. Recurrences and second primary breast cancers in older women with initial early-stage disease. *Cancer.* 2007; 109(5): 966–974. DOI: <https://doi.org/10.1002/cncr.22472>
55. **Lash, TL, Fox, MP, Buist, DS**, et al. Mammography surveillance and mortality in older breast cancer survivors. *J Clin Oncol.* 2007; 25(21): 3001–3006. DOI: <https://doi.org/10.1200/JCO.2006.09.9572>
56. **Altschuler, A, Nekhlyudov, L, Rolnick, SJ**, et al. Positive, negative, and disparate—women's differing long-term psychosocial experiences of bilateral or contralateral prophylactic mastectomy. *Breast J.* 2008; 14(1): 25–32. DOI: <https://doi.org/10.1111/j.1524-4741.2007.00521.x>
57. **Geiger, AM, Nekhlyudov, L, Herrinton, LJ**, et al. Quality of life after bilateral prophylactic mastectomy. *Ann Surg Oncol.* 2007; 14(2): 686–694. DOI: <https://doi.org/10.1245/s10434-006-9206-6>
58. **Geiger, AM, West, CN, Nekhlyudov, L**, et al. Contentment with quality of life among breast cancer survivors with and without contralateral prophylactic mastectomy. *J Clin Oncol.* 2006; 24(9): 1350–1356. DOI: <https://doi.org/10.1200/JCO.2005.01.9901>
59. **Rolnick, SJ, Altschuler, A, Nekhlyudov, L**, et al. What women wish they knew before prophylactic mastectomy. *Cancer Nurs.* 2007; 30(4): 285–291, quiz 292–283. DOI: <https://doi.org/10.1097/01.NCC.0000281733.40856.c4>
60. **Carroll, NM, Delate, T, Menter, A**, et al. Use of Bevacizumab in Community Settings: Toxicity Profile and Risk of Hospitalization in Patients With Advanced Non-Small-Cell Lung Cancer. *J Oncol Pract.* 2015; 11(5): 356–362. DOI: <https://doi.org/10.1200/JOP.2014.002980>
61. **Ritzwoller, DP, Carroll, NM, Delate, T**, et al. Comparative effectiveness of adjunctive bevacizumab for advanced lung cancer: The cancer research network experience. *J Thorac Oncol.* 2014; 9(5): 692–701. DOI: <https://doi.org/10.1097/JTO.0000000000000127>
62. **Cross, DS, Rahm, AK, Kauffman, TL**, et al. Underutilization of Lynch syndrome screening in a multisite study of patients with colorectal cancer. *Genet Med.* 2013; 15(12): 933–940. DOI: <https://doi.org/10.1038/gim.2013.43>
63. **Feigelson, HS, Zeng, C, Pawloski, PA**, et al. Does KRAS testing in metastatic colorectal cancer impact overall survival? A comparative effectiveness study in a population-based sample. *PLoS One.* 2014; 9(5): e94977. DOI: <https://doi.org/10.1371/journal.pone.0094977>
64. **Harris, JN, Liljestrand, P, Alexander, GL**, et al. Oncologists' attitudes toward KRAS testing: a multisite study. *Cancer Med.* 2013; 2(6): 881–888. DOI: <https://doi.org/10.1002/cam4.135>
65. **Webster, J, Kauffman, TL, Feigelson, HS**, et al. KRAS testing and epidermal growth factor receptor inhibitor treatment for colorectal cancer in community settings. *Cancer Epidemiol Biomarkers Prev.* 2013; 22(1): 91–101. DOI: <https://doi.org/10.1158/1055-9965.EPI-12-0545>
66. **Greene, SM, Geiger, AM, Harris, EL**, et al. Impact of IRB requirements on a multicenter survey of prophylactic mastectomy outcomes. *Ann Epidemiol.* 2006; 16(4): 275–278. DOI: <https://doi.org/10.1016/j.annepidem.2005.02.016>
67. **Greene, SM, Braff, J, Nelson, A**, et al. The process is the product: a new model for multisite IRB review of data-only studies. *IRB.* 2010; 32(3): 1–6.
68. **Paolino, AR, Lauf, SL, Pieper, LE**, et al. Accelerating regulatory progress in multi-institutional research. *EGEMS (Wash DC).* 2014; 2(1): 1076. DOI: <https://doi.org/10.13063/2327-9214.1076>
69. **Glasgow, RE and Chambers, D.** Developing robust, sustainable, implementation systems using rigorous, rapid and relevant science. *Clin Transl Sci.* 2012; 5(1): 48–55. DOI: <https://doi.org/10.1111/j.1752-8062.2011.00383.x>
70. **Sung, NS, Crowley, WF, Jr., Genel, M**, et al. Central challenges facing the national clinical research enterprise. *JAMA.* 2003; 289(10): 1278–1287. DOI: <https://doi.org/10.1001/jama.289.10.1278>

71. **Westfall, JM, Mold, J and Fagnan, L.** Practice-based research—“Blue Highways” on the NIH roadmap. *JAMA*. 2007; 297(4): 403–406. DOI: <https://doi.org/10.1001/jama.297.4.403>
72. **National Academy of Medicine (formerly Institute of Medicine).** Integrating research and practice: Health system leaders working toward high-value care: Workshop summary. Washington, DC: The National Academies Press; 2015.
73. **Greene, SM, Reid, RJ and Larson, EB.** Implementing the learning health system: From concept to action. *Ann Intern Med*. 2012; 157(3): 207–210. DOI: <https://doi.org/10.7326/0003-4819-157-3-201208070-00012>
74. **Aiello Bowles, EJ, Tuzzio, L, Wiese, CJ, et al.** Understanding high-quality cancer care: A summary of expert perspectives. *Cancer*. 2008; 112(4): 934–942. DOI: <https://doi.org/10.1002/cncr.23250>
75. **Horner, K, Ludman, EJ, McCorkle, R, et al.** An oncology nurse navigator program designed to eliminate gaps in early cancer care. *Clin J Oncol Nurs*. 2013; 17(1): 43–48. DOI: <https://doi.org/10.1188/13.CJON.43-48>
76. **Wagner, EH, Ludman, EJ, Aiello Bowles, EJ, et al.** Nurse navigators in early cancer care: A randomized, controlled trial. *J Clin Oncol*. 2014; 32(1): 12–18. DOI: <https://doi.org/10.1200/JCO.2013.51.7359>
77. **Buist, DSM, Field, TS, Banegas, MP, et al.** Training in the Conduct of Population-Based Multi-Site and Multi-Disciplinary Studies: The Cancer Research Network’s Scholars Program. *J Cancer Educ*. 2017; 32(2): 283–292. DOI: <https://doi.org/10.1007/s13187-015-0925-x>
78. **Nichols, HB, Sprague, BL, Buist, DS, et al.** Careers in cancer prevention: What you may not see from inside your academic department—a report from the American Society of Preventive Oncology’s Junior Members Interest Group. *Cancer Epidemiol Biomarkers Prev*. 2012; 21(8): 1393–1395. DOI: <https://doi.org/10.1158/1055-9965.EPI-12-0664>
79. **Kaiser Permanente NCORP.** <https://ncorp.cancer.gov/findasite/profile.php?org=1645>. Accessed October 26, 2018.
80. **Aiello, EJ, Buist, DS, Wagner, EH, et al.** Diffusion of aromatase inhibitors for breast cancer therapy between 1996 and 2003 in the Cancer Research Network. *Breast Cancer Res Treat*. 2008; 107(3): 397–403. DOI: <https://doi.org/10.1007/s10549-007-9558-z>
81. **Bowles, EJ, Wernli, KJ, Gray, HJ, et al.** Diffusion of Intraperitoneal Chemotherapy in Women with Advanced Ovarian Cancer in Community Settings 2003–2008: The Effect of the NCI Clinical Recommendation. *Front Oncol*. 2014; 4: 43. DOI: <https://doi.org/10.3389/fonc.2014.00043>
82. **The Office of the National Coordinator for Health Information Technology.** Connecting Health and Care for the Nation: A 10-Year Vision to Achieve an Interoperable Health IT Infrastructure. Available at: <https://www.healthit.gov/sites/default/files/ONC10yearInteroperabilityConceptPaper.pdf>.
83. **Miller, H and Johns, L.** Interoperability of Electronic Health Records: A Physician-Driven Redesign. *Manag Care*. 2018; 27(1): 37–40.
84. **PCORnet Common Data Model (CDM) Specification, Version 4.1.** [http://pcornet.org/wp-content/uploads/2018/05/PCORnet-Common-Data-Model-v4-1-2018\\_05\\_15.pdf](http://pcornet.org/wp-content/uploads/2018/05/PCORnet-Common-Data-Model-v4-1-2018_05_15.pdf). Accessed June 19, 2018.
85. **Sentinel Common Data Model.** <https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model/sentinel-common-data-model>. Accessed June 19, 2018.
86. **The Health Care Systems Research Network Virtual Data Warehouse.** <http://www.hcsrn.org/en/Tools%20&%20Materials/VDW/VDWDataModel/VDWSpecifications.pdf>. Accessed June 19, 2018.

**How to cite this article:** Doria-Rose, VP, Greenlee, RT, Buist, DSM, Miglioretti, DL, Corley, DA, Brown, JS, Clancy, HA, Tuzzio, L, Moy, LM, Hornbrook, MC, Brown, ML, Ritzwoller, DP, Kushi, LH and Greene, SM. 2019 Collaborating on Data, Science, and Infrastructure: The 20-Year Journey of the Cancer Research Network. *eGEMs (Generating Evidence & Methods to improve patient outcomes)*, 7(1): 7, pp. 1–11. DOI: <https://doi.org/10.5334/egems.273>

**Submitted:** 21 June 2018

**Accepted:** 16 October 2018

**Published:** 29 March 2019

**Copyright:** © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

**] *eGEMs (Generating Evidence & Methods to improve patient outcomes)* is a peer-reviewed open access journal published by Ubiquity Press.**

**OPEN ACCESS** 