The Challenge of Conducting Comparative Effectiveness Research in Cancer: The Impact of a Fragmented US Health-Care System

Paul A. Fishman, Mark C. Hornbrook, Debra P. Ritzwoller, Maureen C. O’Keeffe-Rosetti, Jennifer Elston Lafata, Ramzi G. Salloum

Comparative effectiveness research (CER) can make important contributions to the transformation of US health care by filling gaps left by tightly controlled clinical trials. However, without comprehensive and comparable data that reflect the diversity of the US health-care system, CER’s value will be diminished. We document the limits of observational CER by examining the age at diagnosis, disease stage, and select measures of health-care use among individuals diagnosed with incident cancer aged 65 or older from four large health maintenance organizations (HMOs) relative to seniors identified through the linked Surveillance, Epidemiology, and End Results (SEER)–Medicare data for the period 1999–2007. Aged individuals in the HMOs were younger, diagnosed at earlier stages, and more likely to receive care in inpatient settings than individuals in the linked SEER–Medicare data. These differences highlight the need for comprehensive and comparable datasets that reflect the diversity of US health care to support CER that can inform health-care reform in the United States.


Comparative effectiveness research (CER) provides evidence about the benefits and harms of different interventions and strategies to prevent, diagnose, treat, and monitor health conditions in “real world” settings (1) and is increasingly seen as a critical component in support of health reform in the United States (1,2). Valid inferences from observational CER depends on high-quality and consistent data on alternative treatments provided to diverse populations, but the variety of ways in which American health care is organized, delivered, financed, and recorded inhibits the availability of comprehensive and comparable necessary data to support this work. We describe the challenges for conducting observational CER for assessing cancer care with data on older adults diagnosed and treated for their cancer in the fee-for-service (FFS) US market compared with seniors treated in a capitated US market. We review the challenges that arise from the fragmented US health-care system and propose a research and a policy agenda for improving the prospects of CER to identify improvements in cancer care.

Background

Cancer is the second leading cause of death in the United States (3,4) and a substantial driver of total health-care costs (5). Therefore, improvements in the detection, treatment, and management of cancer are at the core of efforts to improve population health outcomes and lower total national health-care spending in the United States. However, despite significant resources devoted to cancer research over the past several decades, fundamental questions about effective approaches to care remain along the spectrum from screening (6–11) to palliation (12,13). Further, unexplained variation in rates of prevention, treatment, outcomes, and cost remains across regions and among specific segments of the population, highlighting the potential benefit of identifying and disseminating evidence-based approaches to cancer control and prevention.

CER can identify best practices and provide evidence of the sources of variation in care that affect outcomes; however, the fragmented nature of US health care impacts both the quality and availability of the data upon which CER depends. This fragmentation manifests in several ways:

1. Health insurance: Americans’ access to insurance and the completeness of covered services differ by age, employment, income, military status, and ethnicity, with almost one-fifth of the US population currently without insurance. These different contact points create parallel but different access to services and thus service utilization and data availability.

2. Finance: Most American health-care providers are paid on a FFS basis, reimbursed only when they deliver services, but 22% (14) of Americans are insured through prepaid, capitated insurance that pays providers prospectively regardless of what, if any, services are provided. FFS and capitation create different incentives for care delivery that may impact the type and scope of services used as well as the availability of data, because capitation reduces the incentive to document the provision of every billable service or procedure.

3. Organization of care: Most Americans are forced to bundle their care from providers that practice independently of one another, but some Americans receive care from integrated systems that coordinate patient care across providers and care settings. Care integration supports more comprehensive information on health-care use, as electronic information systems can be used to follow patients within and across episodes of care.
4. Capture of health information: Although there is increasing use of electronic medical records (EMRs), the majority of US providers continue to rely on paper records. Thus, although EMRs can improve access to patient-level information about personal and environmental risk factors, little is known about how efficiently EMRs from independent providers can be linked.

Coordinated efforts by US federal and state health agencies have led to data resources that document the incidence and prevalence of cancer, addressing many of the challenges created by fragmented US health care. Leading this work are the National Program of Cancer Registries (NPCR) managed by the Centers for Disease Control and Prevention, and the SEER program of the National Cancer Institute (NCI) (15), which together capture cancer incidence data for 96% of the entire country through 45 state, Washington, DC, and territorial cancer registries that report to the NPCR and 17 regional population-based cancer registries that report through SEER. SEER registries provide clinical information (tumor site, morphology and diagnosis stage, first treatment course, and survival), whereas NPCR clinical data are more limited.

Several projects have linked NPCR and SEER with information on health-care use to support CER, with the linkage created and maintained by the NCI between SEER and claims data for Medicare beneficiaries maintained by the Centers for Medicare and Medicaid Services (CMS) being the most widely used (16). The Medicare program insures approximately 50 million Americans (17)—almost all individuals over age 65 and a smaller number of younger adults and children with permanent disabilities. The link between SEER and Medicare allows for cancer-specific CER among Medicare beneficiaries, but two critical gaps remain in its use for comprehensive research. First, because the link is with Medicare, which is primarily an insurance program for persons aged 65 and over, younger adults and children are (for the most part) excluded from the dataset. Although the majority (55.9%) of incident cancers in the United States in the last decade were among individuals aged 65 and over (18), the linked SEER–Medicare data exclude the large minority of Americans with incident cancers younger than age 65. Second, the linked SEER–Medicare data exclude the one-quarter of older Americans enrolled in the Medicare Advantage program (17), an insurance option offered by CMS that allows seniors to receive care from health maintenance organizations (HMOs) that are paid on a capitated, or per-person, basis, rather than FFS. The exclusion of these seniors is an artifact of different data reporting requirements CMS imposes on FFS and capitated providers because HMOs serving seniors are not required to submit detailed, itemized information on health service use similar to the claims submitted by FFS providers.

Methods

We documented the challenge of conducting cancer-specific observational CER using linked SEER–Medicare data by examining differences in cancer incidence and stage of illness and select measures of health service use among seniors in the linked SEER–Medicare data relative to seniors enrolled in four large HMOs. We focused on older adults for two reasons. First, as indicated above, more than half of all individuals diagnosed with incident cancer in the United States are aged 65 or older. Second, the linked SEER–Medicare data provided by the NCI have resulted in this resource being the leading source of information about cancer care and outcomes in the United States.

We examined the differences in populations diagnosed with cancer captured by the linked SEER–Medicare data with seniors diagnosed with cancer in four large nonprofit US HMOs. The four HMOs are Group Health Cooperative, based in Seattle, WA; the Health Alliance Plan (HAP)/Henry Ford Medical Group (HFMG), based in Detroit, MI; and the Northwest and Colorado regions of Kaiser Permanente, based in Portland, OR, and Denver, CO, respectively. Each of these HMOs provides comprehensive health-care services through, primarily, closed-panel delivery models with salaried physicians. The four health plans are members of the HMO Cancer Research Network (CRN), created by the NCI as a population-based laboratory to conduct research on cancer prevention, early detection, treatment, long-term care, and surveillance. The CRN is the largest research effort of the HMO Research Network (HMORN), a consortium of 19 health-care organizations with both defined patient populations and formal, recognized research capabilities.

The four health plans participating in the current study provide health service and capitated health insurance through a wide range of private and public health insurance programs, including employer-sponsored and individual and family plans, Medicaid programs for low-income Americans, as well as the Medicare Advantage for older adults and disabled individuals. Each health plan serves a population that generally represents their local communities (16,17). Group Health and HAP/HFMG are each located within SEER catchment areas and follow SEER protocols to abstract and provide clinical information on all cancer diagnoses among plan members to the local SEER registrars. Individuals diagnosed with cancer at the two Kaiser sites were identified from health plan–specific tumor registries, which follow SEER-compatible protocols. For HMO enrollees, we used enrollment and tumor registry data to identify all incident cancers for individuals whose diagnosis was made as of their 65th birthday and were enrolled in the health plan for at least 30 days prior to their diagnosis. We imposed no requirement for enrollment in the health plan following the cancer diagnosis to allow for individuals who may have died shortly following their diagnoses. Each incident cancer diagnosis was counted independently, so individuals may have multiple primary cancers. Demographic, diagnostic, and stage-of-illness information were obtained from the tumor registries for the years 1999–2007.

Data on patients whose cancer status and health-care use were available from the linked SEER–Medicare data were obtained from Information Management Services Inc (IMS) for the years 1999–2007. The study team obtained the complete set of data for all individuals identified in the 17 participating regional SEER registries for whom a link with CMS files was made, which reflects 93% of all older adults insured through FFS Medicare whose incident cancer is captured in a SEER registry. Demographic, diagnosis, and stage-of-illness information were derived from the Patient Entitlement and Diagnosis Summary File (PEDSF), which contains the relevant data for both SEER and CMS person-level data. All Medicare beneficiaries whose health-care use and cancer status were in the linked SEER–Medicare data during these years were included in
the analysis file with no requirements on pre- or postenrollment in Medicare or survival following the cancer diagnosis.

The stage of disease at diagnosis is based on the SEER Summary Stage, which is available in several variables reported on the PEDSF provided by IMS. The derived SEER Summary Stage 2004 is available for incident cancers diagnosed from 2004; the Summary Stage 2000 captures stage for cancers diagnosed between 2001 and 2003; and the Summary Stage 1977 identifies stage of disease for cancers diagnosed from 1995 to 2000. The SEER Summary Stage reports incident cancers as in situ, localized, regional, or distant/metastasis.

This study was approved by the Institutional Review Boards of the four HMO sites.

Results

There were 16,079 seniors with incident cancer from the HMOs and 405,166 in the linked SEER–Medicare data that met inclusion criteria during the period 1997–2007. Seniors in the HMOs were generally younger than in the linked SEER–Medicare data at time of diagnosis (Figure 1) and were more likely to be diagnosed with localized disease than were seniors whose experience was captured in the SEER-Medicare data (Figure 2). Among HMO enrollees diagnosed with any cancer, 48% had either in situ or localized disease, 19% had regional, and 17% had distant metastasis. Among seniors in SEER–Medicare, 38% had either in situ or localized disease, 36% had regional, and 20% had distant metastatic disease.

Age and stage for individuals with incident cancer for the HMO and linked SEER–Medicare data for colorectal, prostate, breast, and lung cancers are reported in Table 1. Although seniors in the HMOs and linked SEER–Medicare data had similar age distributions for lung and colorectal cancers, women in the HMOs with breast cancer and those diagnosed with colorectal cancer were younger than seniors in the linked SEER–Medicare data. The most notable result regarding stage is the large percentage of men in the HMOs with localized prostate cancer relative to the linked SEER–Medicare data.

Unadjusted rates per 1000 person-months for select measures of health service use for the 6 months before and 6 months after a cancer diagnosis are reported in Figure 3. Inpatient admissions (Figure 3A) and days per admission (Figure 3B) are higher among individuals identified in the linked SEER–Medicare data before and after their diagnosis relative to individuals in the HMOs. The relative gap in use rates is smaller in the post-diagnostic period but still almost twice as great among individuals in FFS settings relative to HMOs for both measures. Outpatient visits (Figure 3C) are higher among individuals in the linked SEER–Medicare data prior to diagnosis but substantially higher among the aged in HMOs in the 6 months following diagnosis.

Discussion

We have demonstrated opportunities for policy-relevant and clinically meaningful CER studies using pooled HMO and SEER–Medicare data that arise from variations in cancer incidence and patterns of care across these systems. Seniors in the HMOs were, on average, younger, diagnosed at earlier stages, and more likely to receive post-diagnosis care in outpatient settings relative to seniors in FFS Medicare. Several factors may explain these differences. HMOs may generally enroll younger seniors and the earlier diagnostic stage may reflect greater emphases on prevention and screening within the HMOs that result in earlier detection.

![Figure 1.](image-url) Cancer incidence by age and market segment, all cancers.
note the potentially adverse consequences of this outcome as earlier detection may result in treatment initiation that is premature or perhaps not medically indicated (6,8) and also acknowledge previous research that has examined these factors (19–23). Our results confirm the critical role of observational CER that compares cancer care and outcomes in FFS and HMO Medicare in identifying best practices for older adults in the United States and elsewhere.

Our finding of the different mix and intensity of service use among seniors diagnosed with cancer in the HMOs and FFS Medicare programs has not been previously documented and highlights both the challenge of and opportunity for conducting CER in cancer care in the United States. The evidence we report suggests that the correlated demographic, clinical, and service mix/intensity factors among seniors with cancer in HMOs are a significant methodological challenge to the conduct of comprehensive, integrated, and comparable CER on cancer care for Americans over age 65. The fragmented structure of the US health-care delivery system creates challenges in assessing the impact of variations in financing, sources of care, and patient preferences on observed treatment patterns and outcomes. This fragmentation also makes it

<table>
<thead>
<tr>
<th>Cancer site</th>
<th>N</th>
<th>SEER-Medicare</th>
<th>HMO</th>
<th>SEER-Medicare</th>
<th>HMO</th>
<th>SEER-Medicare</th>
<th>HMO</th>
<th>SEER-Medicare</th>
<th>HMO</th>
<th>SEER-Medicare</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3206</td>
<td>8993</td>
<td>4210</td>
<td>113551</td>
<td>4225</td>
<td>90234</td>
<td>4335</td>
<td>111388</td>
<td></td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65–69</td>
<td>18.1%</td>
<td>20.4%</td>
<td></td>
<td>29.0%</td>
<td>26.6%</td>
<td>28.1%</td>
<td>22.2%</td>
<td>20.7%</td>
<td>20.1%</td>
<td></td>
</tr>
<tr>
<td>70–74</td>
<td>22.6%</td>
<td>23.0%</td>
<td></td>
<td>29.9%</td>
<td>28.6%</td>
<td>26.2%</td>
<td>23.8%</td>
<td>28.9%</td>
<td>25.2%</td>
<td></td>
</tr>
<tr>
<td>75–79</td>
<td>25.1%</td>
<td>20.6%</td>
<td></td>
<td>23.3%</td>
<td>23.5%</td>
<td>22.0%</td>
<td>23.4%</td>
<td>24.1%</td>
<td>25.4%</td>
<td></td>
</tr>
<tr>
<td>80–84</td>
<td>19.1%</td>
<td>19.8%</td>
<td></td>
<td>12.0%</td>
<td>13.4%</td>
<td>14.3%</td>
<td>17.3%</td>
<td>17.4%</td>
<td>17.8%</td>
<td></td>
</tr>
<tr>
<td>&gt;85</td>
<td>16.1%</td>
<td>20.4%</td>
<td></td>
<td>5.9%</td>
<td>8.0%</td>
<td>9.5%</td>
<td>13.3%</td>
<td>8.9%</td>
<td>11.4%</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In situ</td>
<td>2.1%</td>
<td>5.2%</td>
<td></td>
<td>0.00%</td>
<td>0.03%</td>
<td>17.3%</td>
<td>15.4%</td>
<td>0.1%</td>
<td>0.1%</td>
<td></td>
</tr>
<tr>
<td>Localized</td>
<td>30.4%</td>
<td>37.7%</td>
<td></td>
<td>71.45%</td>
<td>2792%</td>
<td>52.6%</td>
<td>54.7%</td>
<td>17.0%</td>
<td>17.0%</td>
<td></td>
</tr>
<tr>
<td>Regional</td>
<td>36.5%</td>
<td>34.6%</td>
<td></td>
<td>6.58%</td>
<td>62.01%</td>
<td>16.2%</td>
<td>22.2%</td>
<td>21.6%</td>
<td>22.4%</td>
<td></td>
</tr>
<tr>
<td>Distant metastasis</td>
<td>14.4%</td>
<td>16.1%</td>
<td></td>
<td>6.53%</td>
<td>4.82%</td>
<td>2.6%</td>
<td>4.5%</td>
<td>43.1%</td>
<td>50.6%</td>
<td></td>
</tr>
<tr>
<td>Missing/unknown</td>
<td>16.6%</td>
<td>6.4%</td>
<td></td>
<td>15.44%</td>
<td>5.22%</td>
<td>11.4%</td>
<td>3.1%</td>
<td>18.2%</td>
<td>9.9%</td>
<td></td>
</tr>
</tbody>
</table>

* HMO = health maintenance organization; SEER = Surveillance, Epidemiology, and End Results.

Figure 2. Stage of disease for incident cancers.
Figure 3. Unadjusted rates per 1000 patients for selected measures of health service use before and after diagnosis by market segment. A) Inpatient admissions per 1000 person-months. B) Inpatient days per 1000 person-months. C) Outpatient visits per 1000 person-months.
difficult to isolate best practices for wider dissemination throughout the United States. Continued efforts to improve data quality and harmonization across the United States, paired with creative multivariable statistical modeling tools, will provide deeper insights and new enigmas. These incremental insights will provide feedback to policy makers, clinicians, and patients on how our health systems are working (and failing), as well as a more rational basis to redirect and/or refine health policy initiatives.

Our study uses the experiences of older Americans to highlight the implications for CER caused by the fragmented US health-care sector (24). The NIH has created programs designed to address some of these gaps and to support collaborative research that bridges differences across geographic regions, care delivery models, and insurance markets. NIH initiatives such as the Roadmap for Medical Research and the Clinical and Translational Science Award program are examples of efforts designed to reduce barriers to conducting collaborative and translational research, but these efforts have yet to produce comprehensive health information resources to support the CER on which health reform depends.

There are examples of successful investments made by federal agencies in coordinated population-based research and health information technology that have the potential to support the type of CER needed to support US health-care reform. One success is the HMORN, of which the four health plans that participated in this study are members, whose research infrastructure has been primarily supported by the NCI. The power of combining HMO datasets across the network creates the opportunity for direct comparisons of the otherwise fragmented elements of US health care (25). The successful investment of the NCI into cancer-specific research within the HMORN has led to subsequent investments into mental health and cardiovascular disease, but each of these efforts is limited to one market segment.

The NIH has invested in several nationally representative panel data series, some of which, such as the Health and Retirement Survey and the observational panel developed for the Women’s Health Initiative, also link to Medicare data as the NCI has done with SEER. Although these efforts have the same limitations regarding seniors enrolled in HMOs, they are examples of how CER can be supported through coordinated data collection efforts over time that link detailed primary and clinical data with information on health service use.

An example of an explicit investment in health information technology is the DARTNet program, co-supported by the Agency for Healthcare Research and Quality (AHRQ) and the NIH. DARTNet is a federated network of electronic health record data and other clinical information from typically smaller clinical settings across the county linked through a secure web-based system that can be searched and queried as one large database while maintaining privacy and confidentiality of patient data (http://www.dartnet.info/). AHRQ has also long supported primary care practice–based research networks (PBRNs), which are, as a group of primary care practices, affiliated in their mission to investigate questions related to community–based practice and to improve the quality of primary care (26). PBRNs are often limited by their ability to easily share health records and clinical information but hold promise as a way of conducting CER to reflect actual care settings.

Without data that allow for analyses of the differences in populations, care processes, and outcomes throughout the United States, public and private policy leaders cannot make informed decisions about how to evaluate and implement best practices. The investment made in the data maintained by NPCR and linked SEER–Medicare data is a strong platform on which to build a multisetor and multiregional comprehensive dataset that can fully capture the entire population’s experience with cancer care and outcomes.

The most effective outcome in support of observational CER in support of improved cancer care and outcomes is the completion of a comprehensive cancer data system, as called for in a 1999 US Institute of Medicine report on improving the quality of cancer care (27). The investments made by the NPCR and NCI to provide comprehensive data on cancer incidence in the United States, and to link this data with health-care use for many older Americans, have supported critical research efforts. The next step is development of a comprehensive data resource that captures health-care use and outcomes for all Americans with cancer.

Conclusion

Health-care reform in the United States requires research that identifies and disseminates evidence of effective care outside of rigorously controlled clinical trials that can reduce the overall cost of providing services. The need for this research has been identified in the most important pieces of federal health legislation passed in recent years, which have created and funded programs of research to support the CER on which health-care reform will depend. What is missing from the current plan for supporting research is the creation of comprehensive data that capture the full diversity of US health-care delivery and finance that will allow researchers to isolate the sources of variations in health-care delivery and determine best practices. We highlighted the need for the creation of such data showing significant differences among those diagnosed with cancer in the FFS Medicare program and those served by HMOs in the Medicare Advantage program.

References


AHRQ Publication No. 01-P020.


Funding
This work was supported by the National Cancer Institute at the National Institutes of Health (Grant numbers R01 CA114204 and R01 CA 114204 to MCH, RC2 CA148185 to DPR, R25 CA116339 to RGS, and Cooperative Agreement numbers U19 CA79689 and SU2CA148471).

Note
The authors thank Erin Keast and Jenny Staab of the Kaiser Permanente Center for Health Research for critical contributions to this paper.

Financial Disclosures: The authors have no conflicts to disclose.

Affiliations of authors: The Group Health Research Institute, Seattle, WA (PAF); The Center for Health Research, Kaiser Permanente Northwest, Portland, OR (MCH, MCO-R); The Institute for Health Research, Kaiser Permanente Colorado, Denver, CO (DPR); Massey Cancer Center, Virginia Commonwealth University, Richmond, VA (JEL); The Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI (JEL, RGS); Gillings School of Global Public Health, The University of North Carolina, Chapel Hill, NC (RGS).