State-Level Cancer Quality Assessment and Research: 
Building and Sustaining the Data Infrastructure

Joseph Lipscomb, PhD* and Theresa W. Gillespie, PhD†‡
*Department of Health Policy and Management, Rollins School of Public Health, and the Winship Cancer Institute, Emory University, Atlanta, GA
†Department of Surgery, School of Medicine, and the Winship Cancer Institute, Emory University, Atlanta, GA
‡Health Services Research & Development Program, Atlanta Veterans Affairs Medical Center, Atlanta, GA

Abstract
Over a decade ago, the Institute of Medicine called for a national cancer data system in the United States to support quality-of-care assessment and improvement, including research on effective interventions. Although considerable progress has been achieved in cancer quality measurement and effectiveness research, the nation still lacks a population-based data infrastructure for accurately identifying cancer patients and tracking services and outcomes over time. For compelling reasons, the most effective pathway forward may be the development of state-level cancer data systems, in which central registry data are linked to multiple public and private secondary sources. These would include administrative/claims files from Medicare, Medicaid, and private insurers. Moreover, such a state-level system would promote rapid learning by encouraging adoption of near-real-time reporting and feedback systems, such as the Commission on Cancer’s new Rapid Quality Reporting System. The groundwork for such a system is being laid in the state of Georgia, and similar work is advancing in other states. The pace of progress depends on the successful resolution of issues related to the application of information technology, financing, and governance.

Keywords
National Cancer Data System; quality-of-care assessment and improvement; cancer quality measurement and effectiveness research; Commission on Cancer’s new Rapid Quality Reporting System

Over a decade ago, the Institute of Medicine’s (IOM’s) National Cancer Policy Board issued companion reports calling for creation of enhanced data systems to monitor and improve the quality of cancer care in the United States.1,2 The reports concluded that many Americans do not receive adequate quality cancer care, although the extent of the problem is “unknown.” Moreover, in many instances, there was not an evidence-based consensus on...
what constitutes quality care. In response, the IOM Board defined the attributes of an “ideal” cancer care data system (as paraphrased below):

- well-established quality of care measures;
- use of cutting-edge information technology to capture data on patient care and outcomes;
- widely known and accepted standards for reporting clinical, service delivery, and outcomes data;
- national, population-based selection of cases;
- repeated cross-sectional analyses to monitor national trends in care and outcomes;
- creation of benchmarks for quality improvement;
- data systems that promote quality improvement at the individual practice level, while supporting public reporting of selected aggregate measures of quality;
- flexibility within the system to adapt to new evidence relating services to outcomes, advances in medical technology, and changes in the broader health care delivery system; and
- privacy protections to ensure that patient data are used only for intended, legitimate purposes.

Over time, these IOM reports have served to stimulate important, constructive policy responses from the National Cancer Institute (NCI), the American Society of Clinical Oncology (ASCO), the American College of Surgeons’ Commission on Cancer (CoC), and many others (as discussed here subsequently).

Then, in 2010, the IOM’s National Cancer Policy Forum issued a report articulating a new perspective and accompanying approach to understanding and improving cancer care quality—a transformative Weltanschaung calling for creation of a “rapid learning system for cancer care.” Such a system “uses advances in information technology to continually and automatically collect and compile from clinical practice, disease registries, clinical trials, and other sources of information, the evidence needed to deliver the best, most up-to-date care that is personalized for each patient.” That evidence is to be made available as rapidly as possible to users, which include patients, providers, payers, and public agencies. Critically important is that such a system learns routinely and interactively: data captured at the patient level would flow into a databank or coordinated databases, generating new evidence on the impact of specific interventions on outcomes that matter to decision makers at all levels.

Embracing the concept of a learning health care system, the Brookings Institution’s Engelberg Center has pursuasively called for the enhanced use of electronic health information collected in the course of care delivery—including electronic health records (EHRs), claims data, and registries—to support a range of complementary activities. These include quality measurement and reporting, evidence development for coverage decisions, medical product safety surveillance, and comparative effectiveness research. Of note, the
examples of “learning” cited in this Brookings report highlight the potential for progress at the level of the medical care practice, the local community, the region, and the nation. Statewide or state-oriented initiatives or opportunities for health care learning are not discussed.

With these developments as backdrop, this article argues that in the context of the US health care system, the state is a natural focal point—and perhaps the most suitable geopolitical arena—for cultivating a learning health care system for cancer care. Assessing and improving the quality of cancer care would be the central focus, but not the only one. In addition, the type of state-level data infrastructure proposed, and illustrated, in the following sections could support parallel work to evaluate the effectiveness, safety, and economic value of cancer care. The aim would be to inform not only patient-provider choices on the frontlines, but also decision making by third-party payers, private industry, regulators, and other standards-setting organizations. As the Brookings report implies, these multiple clinical and policy uses of a strong cancer data infrastructure are complementary endeavors.

This article is organized as follows. In the next section, we address several topics that are pivotal to the larger argument, specifically (1) the importance of identifying evidence-based, consensus-supported measures of cancer care quality; (2) why high-quality, population-based cancer registries remain central to cancer quality assessment at all levels, but why even the best registries need to be augmented by additional data sources to realize their potential; and (3) why the state offers an appropriate platform for organizing the functions of a rapid learning system for cancer care. The section after that summarizes ongoing work in the state of Georgia to develop a cancer care data system for quality assessment, with an eye toward future developments to promote rapid learning and multiway communication among providers, patients, and other decision makers. Building on the Georgia example and referencing recent public- and private-sector efforts to stimulate effective application of health information technology, the concluding section briefly discusses technical, financial, and governance issues related to building and sustaining a state-level data infrastructure.

TOWARD A STATE-LEVEL CANCER DATA SYSTEM FOR RAPID LEARNING

In this section, we take up 3 matters that constitute important desiderata for the remainder of the article.

Identifying Evidence-Based Measures of Cancer Care Quality

After more than 2 decades, the IOM’s 1990 definition of health care quality continues to provide a foundation for measure development: the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.\(^8\) For the individual, the ideal is receipt of evidence-based care along the disease continuum (prevention, detection, diagnosis, treatment, and palliation), delivered in a timely and technically competent manner, with good patient-provider communications and shared decision making. For populations, the ideal is credible evidence of appropriate care as indexed by predefined performance indicators of quality. A useful perspective on “appropriate” care is that it reflects (obversely) the absence of overuse, underuse, or misuse of services.\(^1\) Operationally, any measure of
quality requires a denominator (the set of individuals eligible for assessment, as defined by
inclusion/exclusion criteria) and a numerator (the number of those eligible who score
favorably); the measure is generally reported as a fraction or percent.

From a conceptual standpoint, the structure–process–outcome paradigm for quality-of-care
assessment defined by Donabedian nearly half a century ago remains a touchstone (Fig. 1).
Although improving health care outcomes (leading to a longer life or a better quality life)
may be the ultimate objective, much quality-of-care assessment for chronic diseases such a
cancer focuses, in fact, on “process” measures: Will the care being delivered lead, on
average, to desired outcomes? As a practical matter, there is frequently a considerable time
lag between the care rendered and the emergence of final outcomes. Moreover, in a given
instance, multiple person-and system-level factors, in addition to the specific treatments
received, may influence outcome, so that sometimes “appropriate” care may be associated
with poor outcomes and conversely.

This underscores the importance of identifying quality-of-care measures that are strongly
evidence based—meaning there is a statistically robust predictive relationship between
specific services (the processes of care) and outcomes that matter to patients and other
decision makers. Hence, the central importance of a learning health care system is to support
comparative effectiveness research that can, in turn, inform quality measurement
development. The vital interconnections between outcomes measurement, process–outcomes
evaluation, quality-of-care improvement, the subsequent impact on populations, and the
dynamic growth of the evidence base are depicted in Figure 2.

In the wake of the IOM reports and guided generally by the precepts above, federal
agencies and major cancer professional organizations, including ASCO, launched separate
(and ultimately complementary) initiatives to measure and improve cancer care quality. For
purposes of this article, we note in particular the Cancer Quality Measures Project,
established at the National Quality Forum (NQF) in 2002 under the guidance and
sponsorship of a coalition of federal agencies led by the NCI and involving also the Agency
for Healthcare Research and Quality (AHRQ), the Centers for Medicare & Medicaid
Services (CMS), and the Centers for Disease Control and Prevention (CDC). In parallel,
ASCO and the National Comprehensive Cancer Network (NCCN) initiated a cancer
measures consensus process. By 2007, the NQF and the ASCO–NCCN measure sets, both
focusing on breast cancer and colorectal cancer, had been reconciled to a single measure set
through a synchronization process involving also the American College of Surgeons’ CoC
and hosted by the NQF.

These measures—which are the focus of cancer quality assessment in the Georgia-based
project discussed in the next section—are summarized in Table 1. This initial measure set is
also being deployed now in the first large-scale effort in the United States to
“institutionalize” rapid case ascertainment—the Rapid Quality Reporting System (RQRS),
being launched by the CoC, with initial support from the NCI. As will be seen, RQRS is a
Web-based reporting and feedback system with the potential to be one cornerstone of a state-
level rapid learning system for cancer care.
The Central Cancer Registry and Quality Assessment: Pivotal, Yet Alone Not Enough

In its 2000 report on enhancing data systems, the IOM’s National Cancer Policy Board emphasized the pivotal role of central cancer registry systems in population-level assessment of quality. These systems include the NCI’s Surveillance, Epidemiology and End Results (SEER) program; the CDC’s National Program of Cancer Registries (NPCR); and the CoC’s National Cancer Data Base (NCDB). The SEER registries cover just under 30% of the US population currently, including 12 states in their entirety, whereas the NPCR supports central cancer registries in virtually every state and US territory. Both SEER and NPCR seek to include every incident cancer case within their geographic boundaries and are thus “population based.” The NCDB includes only those patients—estimated now to represent about 70% of all incident cases in the United States—treated at 1 or more of the roughly 1500 CoC-approved hospital-based cancer programs across the country. Although not therefore population based, the NCDB includes a very large and diverse patient base. In the prototype state-level cancer data system proposed and illustrated here subsequently, both the NPCR and NCDB play central roles (with SEER also of import in states included in that program).

Although the information collected through these registry systems differs in some important ways, all 3 routinely require reporting entities (which largely comprise hospital-based, not office practice–based, providers) to submit a core set of data elements on each newly diagnosed case. These data elements include demographics (including age, race/ethnic status, sex, marital status, address, and primary payer at time of diagnosis), primary type of cancer, stage at diagnosis, histology, grade, hormonal status (starting with 2008 cases), first course of treatment, codes for treating physicians (those involved with case management, surgery, and follow-up care), and patient survival status. Not reported, however, is detailed and confirmed information on all adjuvant therapies (chemotherapy, radiation therapy, hormonal therapy, or palliative care) administered over the treatment period. Nor is there information on disease recurrence, patient-reported outcomes (PROs), and other data possibly relevant for quality-of-care assessment.

For that reason, the IOM strongly recommended that federal agencies expand support for studies to link central registry data with administrative data (including insurance claims files), medical records information, external databases with information on the characteristics of the patient’s treating hospitals and physicians, and also data from “special studies of cases sampled from the registry” to obtain PROs and other information not available from secondary sources. Whether or not in direct response to such recommendations, there has been impressive progress in recent years, particularly in the last decade, to augment cancer registry data with external sources to support research into patterns of care, quality of care, intervention effectiveness, and also cost. Examples of generic types of data linkages include the following:

- Registry—Administrative/Claims: SEER–Medicare (generating >600 publications to this point), SEER–Medicaid, and SEER–Commercial Insurance Data
Hence, there is an abundance of proof of concept—we clearly know how to augment central cancer registry data to support the multiple types of population-based analyses that are important to a learning system for cancer care. But at least 3 important questions arise:

- How can we further promote the linking of registry data to multiple external sources (e.g., administrative and clinical) to enrich the empirical base? For example, far and away the most frequently used of today’s linked databases is SEER–Medicare, but it is missing some important data on adjuvant therapies and includes only patients 65 years or older who are enrolled in fee-for-service Medicare.

- How can we put the linking of registry data to multiple external sources on a financially and administratively sustainable course? At present, the SEER–Medicare database is the only ongoing linked data resource supporting cancer quality and effectiveness research in the United States. The data infrastructure for the other projects cited above was created, in each case, expressly for the enterprise at hand, with no commitment to (nor resources for) sustainability. Yet, for a learning cancer care system, sustainability is essential.

- What, then, is the most appropriate organizational or geopolitical unit for a population-based data system that supports quality assessment and a host of complementary cancer research activities?

We address these questions as the article progresses, with the third one taken up now.

The State as a Natural Arena for a Rapid Cancer Learning System

Several factors jointly support this contention:

- Virtually all states now have “comprehensive cancer control” plans, as strongly encouraged and actively promoted by the CDC, NCI, and the American Cancer Society. Indeed, a state must have a CDC-approved plan to be eligible for CDC-awarded cancer control funding. Most plans include specific benchmarks for statewide performance in primary prevention (e.g., smoking cessation), screening (e.g., percent of age-appropriate women undergoing mammography), early detection (e.g., stage at diagnosis), and access to diagnosis and treatment services. Performance in each of these areas can be, and frequently is, evaluated through what amounts to “quality-of-care” metrics, each with a denominator for the number eligible to receive the indicated service and a numerator for the number who get the service.
At least 1 state, Georgia, has pursued this strategy to its logical end. In the current version of the state plan (updated in 2008 and now in the implementation phase), specific quality-of-care measures have been included in each of the plan’s major performance areas: prevention and education; screening and detection; diagnosis, staging, and treatment; and palliation and survivorship. The measures are drawn from the 52 quality metrics proposed for Georgia in a 2005 IOM report commissioned by the Georgia Cancer Coalition (a nonprofit organization, supported largely by Tobacco Master Settlement Agreement funds) that works in partnership with state government on the cancer plan. (The quality measures used in the Georgia-based study in the next section constitute de facto a small subset of the 52 IOM-recommended metrics for the state.)

In summary, state cancer plans are pervasive and hold much promise as platforms for cancer learning and quality improvement. But plans are largely underexploited and underfunded. We return to this matter of sustainability in the final section.

- A second factor is the ever-improving quality of state-level cancer registries, in terms of completeness, accuracy, and timeliness of data submission. Virtually all NPCR-supported state registries have now achieved Gold or Silver Certification from the North American Association of Central Cancer Registries, with performance indicators that approach and sometimes match those for NCI’s SEER registries. To meet these standards requires regular, timely, productive 2-way communication between the state registry and the state’s reporting sites, which typically include all major (and smaller) facility-based cancer treatment programs.

- Although the CoC accreditation process is national in scope, every state has a CoC state chair. There is increasing emphasis on building strong, state-level communications and quality-improvement networks, including greater participation by CoC leadership in state cancer plans and coordinated work with regional and local American Cancer Society staff.

- The ability to effectively link state registry data (and not only SEER) to administrative/claims files in support of cancer quality and effectiveness research is being increasingly demonstrated in the literature. Adams et al. linked Georgia registry data to Medicaid files to evaluate the Breast and Cervical Cancer Prevention and Treatment Act in Georgia. Fleming et al. linked breast and prostate incident case data from 7 NPCR registries to Medicare files to compare claims-based versus medical records-based comorbidity measurement. Most recently, Boscoe et al. showed how New York State cancer registry data can be accurately linked to the state’s enormous Medicaid files for future quality-of-care assessment. And in ongoing research, Edge et al. are demonstrating how to link private claims data from United HealthCare and Anthem Blue Cross–Blue Shield of Ohio with registry data that draw from both the NCDB reporting sites in Ohio and the Ohio state cancer registry in support of quality-of-care assessment.

- With the exception of SEER–Medicare, the creation of population-based data sets that link detailed information about incident cases with additional data on services and outcomes starts naturally, and perhaps necessarily, at the state level. There is no regional or national source for this incident case information (beyond SEER).
the same token, efforts to link claims databases that may be multistate or even national in scope to incident cancer case information must be carried out at the state level, for that is where the incident case data reside in an accessible fashion.

- The state may be the right-size “laboratory” for population-based cancer learning—large enough to generate statistically robust findings on quality, effectiveness, and outcomes, while small enough to organize and promote effective, multiway communication among cancer programs and providers within its domain. Over the long term, national-level estimates of cancer care quality would be derived by aggregating across the states, taking advantage of the fact that cancer registries, administrative/claims files, and other relevant external data sources already operate, by and large, with highly standardized data items and common vocabularies. We return to this important consideration in the final section.

AUGMENTING STATE CANCER REGISTRY DATA: THE GEORGIA PROJECT

Developing the Capacity for Statewide Quality Assessment Through Multiple Concurrent Linkages

In September 2009, the Association of Schools of Public Health and the CDC, in partnership, awarded Emory University what has now become a 3-year, $500,000 grant to support, “Using Cancer Registry and Other Data Sources to Track Measures of Care in Georgia.” Funded substantially by the NCI through a federal Interagency Agreement with the CDC and supported additionally by the Georgia Cancer Coalition, the project has these specific aims:

1. For incident cases of breast and colorectal cancer in Georgia diagnosed over 1999–2005, link case-specific data from the Georgia Comprehensive Cancer Registry (GCCR) with the following external information sources:
   - Medicare files
   - Medicaid files
   - State Health Benefit Plan (SHBP) files, which cover all employees of the State of Georgia, including public school teachers, and their dependents (individuals were enrolled in a variety of private plans over this period, including Blue Cross–Blue Shield and United HealthCare)
   - Kaiser Permanente of Georgia (KPG) administrative and clinical files
   - Georgia State Hospital Discharge data (which also allows capture of inpatient and hospital outpatient-based care for the uninsured)
   - data from medical chart reviews (where this can feasibly fill gaps left by administrative files)
   - facility-specific descriptive data from the American Hospital Directory
   - physician-specific descriptive data from the CMS Medicare Physician Identification and Eligibility files
other, area-level secondary data sources (US Census and Area Resource File)

The aim is to create a set of “bilateral” linked data sets that can be deployed for quality-of-care assessment in Georgia, as depicted in Figure 3. At the “hub of the wheel” is the GCCR, working closely with Emory’s Georgia Center for Cancer Statistics, which for more than a decade has managed all GCCR data operations under contract. A compressed summary of the data elements available from each of the component data sets involved in the linkages is provided in Table 2. In Table 3, we show initial estimates of the incident breast and colorectal cancer cases in Georgia across 1999–2005 from the GCCR, Medicare, Medicaid, SHBP, and KPG.

2 Subject each bilateral linked data set to rigorous quality checks. The focus is on incident cases in the GCCR that are not found in the claims data; on patients who appear to be incident cases based on the claims data, but who are not in the GCCR; and on identifying factors that predict positive matches (and, by the same token, either type of mismatch).

3 Apply each bilateral linked data set to assess the quality of cancer care, focusing on the breast and colorectal measures shown in Table 1.

4 Design the initial (α) version of a “Consolidated Georgia Cancer Data Resource.” The basic idea is to create a linkage of linked data sets that will eventually (α) allow the GCCR to “follow” cancer patients—and survivors—over time regardless of health plan and (β) enable the creation of de-identified analytic data sets that would be tailored to specific cancer care analyses by virtue of drawing data elements from one, some, or all of the bilateral linked data sets.

Pivotal in accomplishing (α) is that, for each bilateral linkage (e.g., GCCR–SHBP), the project is creating both a de-identified linked data set available to researchers and, in parallel, an identified linked data set available only to the GCCR. With the latter, if and when a patient leaves one health plan (and thus one bilateral linked data set) and joins another plan (and thus another linked data set), the registry can continue to track health care service utilization. Clearly, the greater the proportion of the state’s cancer patients encompassed within the collection of bilateral linked data sets, the greater the capacity to carry out such longitudinal tracking.

Finally, to address patient privacy and confidentiality concerns, and to ensure full compliance with the Health Insurance Portability and Accountability Act compliance, either a memorandum of understanding or data exchange agreement has been put in place in connection with each bilateral linkage. The shape and content of the agreement vary with the linkage. For the GCCR–Medicare linked data, the approval process was precisely that for obtaining the linked SEER–Medicare data (because the GCCR–Medicare linkage is produced concurrently with, and under the same conditions as, the general SEER–Medicare linkage). For the GCCR–Medicaid linkage, a detailed data exchange agreement has been created that establishes the responsibilities and protections for all signatories: the Georgia Department of Community Health (which houses and manages both the GCCR and the
Medicaid data), Thomson Reuters HealthCare (which manages the Medicaid data under contract to the state and has now carried out the GCCR–Medicaid linkage), and Emory University.

By May 2011, all planned bilateral linkages had been successfully executed, and work on specific aims (2) and (3) was underway. The project expects to report its findings, including quality-of-care assessments for breast and colorectal cancer in Georgia, by July 2012. The project team is listed in the Appendix.

Future Steps: More Linkages, Faster Learning, a Broader Set of Quality Measures

The agenda previously mentioned, although ambitious, represents only the first steps toward a rapid learning system for cancer in Georgia—a system that can support population-based quality assessment and improvement, as well as research on the comparative effectiveness and costs of interventions. The steps immediately ahead include the following:

- **Securing the “Missing Links.”** Assuming we can execute (and maintain into the future) the set of bilateral linkages depicted in Figure 3, at least 60% of the incident breast and colorectal cancer cases in Georgia would be covered (it can be inferred from Table 3). To secure the remaining 40%—and to create, in general, a credible population-based cancer data system—requires linking the GCCR with files from the remaining 6 major commercial payers in Georgia (Aetna, Blue Cross–Blue Shield, Cigna, Coventry, Humana, and United HealthCare). A key lesson already from the current “augmenting” project is that one size does not fit all: For these new bilateral linkages to come about, “bilateral negotiations” will likely be required with each major payer. And adequate financial support will be needed to get the ball rolling and keep it in motion. (In the final section, we return to this crucial matter of incentives.)

- **Building a Rapid Learning System for Cancer Care.** Even if the GCCR were successfully linked to every major public and private payer in the state, a major barrier to rapid learning would persist: the multimonth lag between the cancer patient’s diagnosis and receipt of care and the availability of her data (cleaned and consolidated) in the state registry. Add to that the additional months before the registry and corresponding claims data are available for linkage, and one is talking about delays of 2 years or more between diagnosis and the possibility of quality-of-care assessment. Although databases with this time structure can well support important comparative effectiveness and economic evaluations (as SEER–Medicare alone amply demonstrates), the possibility of timely, influential feedback to providers—so they can learn from particular successes and failures—is effectively precluded.

There is a compelling response to this dilemma on the horizon: the CoC’s RQRS.\(^\text{14}\) The Rapid Quality Reporting System is intended to give feedback to providers on a near-real-time basis regarding their performance on selected quality-of-care measures. For each patient meeting inclusion criteria, a parsimonious set of data items is reported as rapidly as feasible to the NCDB through a specially designed Web portal. At any point in time, providers or administrators at the participating CoC facility can query the RQRS system for

*Cancer J. Author manuscript; available in PMC 2016 May 25.*
a variety of reports, including the status of individual patients (e.g., to identify who has failed or is about to fail the quality measure) and comparisons of how well the facility is performing compared with others. Thus, RQRS can provide each facility with rolling year-to-date estimates of its own performance on the quality measures, whether in comparison to the past or to similar CoC facilities.

The Rapid Quality Reporting System underwent beta testing from late 2009 through summer 2011 at 65 volunteering CoC-approved cancer programs across the United States—including at 30 of Georgia’s 39 CoC programs. The cancer quality measures for the RQRS beta test were precisely those summarized in Table 1. In a survey of all beta sites conducted by the CoC in mid-2010, about 93% of cancer registrars, 79% of cancer committee chairs and cancer liaison physicians, and 62% of cancer program administrators rated RQRS as a “very positive” or “somewhat positive” addition to their facility. Across all respondents, about 88% would recommend RQRS to other cancer programs. The Rapid Quality Reporting System is expected to be rolled out nationally by the end of 2011, with all of the CoC’s 1500 approved programs invited to participate on a voluntary basis. In parallel, it is expected that the CoC will expand apace the number of quality metrics in RQRS, including additional cancer disease sites, but no timetable has been established yet.

As the beta test numbers suggest, RQRS implementation is now a high priority in Georgia, with the state’s Comprehensive Cancer Control Plan Implementation Team committed to encouraging 100% participation by CoC-approved programs in the state. (By way of acknowledgement, one of the authors, J.L., cochairs the Implementation Team’s Data & Metrics workgroup, which has championed the RQRS commitment.) Yet, there remains a significant limitation at present: RQRS is structured for hospital-based, CoC-approved facilities that submit cases to the NCDB. There is no corresponding system, yet, geared to non-CoC facilities or to other, office-based cancer providers in the community. Hence, the emerging RQRS system has the potential to provide rapid, accurate measurement of cancer care quality for the roughly 70% of all incident cases treated at CoC-approved facilities, but not population-based estimates, strictly speaking.

There is, finally, the matter of how data on RQRS performance at the facility level might be effectively marshaled by the GCCR or by Georgia cancer plan leaders to provide a timely picture of the quality of cancer care within the state. As a practical matter, it is unlikely the GCCR could obtain such information centrally from the NCDB because the current (and long-standing) business associate agreements between the CoC and each facility do not provide for central distribution of the NCDB data in any way that might identify individual reporting facilities. A possible alternative approach to building RQRS-generated performance data into the state’s quality evaluation apparatus is to require facilities to report to the GCCR summary performance statistics on adherence to RQRS-evaluated cancer quality measures when such data are available. Hence, RQRS participants would be reporting data they already generate, whereas nonparticipants would effectively be exempt. Such an approach (which is not yet under active consideration in Georgia) could provide timely, rolling snapshots of cancer care quality across the state.
Broadening the Scope of Quality Measurement. By design, the Georgia project has focused thus far on a narrowly defined set of treatment measures (Table 1). The overall charge within the new state cancer plan—and from the IOM—is to measure quality of care across the cancer continuum: primary prevention, detection, diagnosis, treatment, posttreatment survivorship, and end of life. Although it is beyond the scope of this article to provide a comprehensive appraisal of how this might be accomplished at the state level, we emphasize for now the potentially important role of self-reported data. Specifically, we mean (1) PROs by those diagnosed with cancer and (2) survey-derived information on health-affecting behaviors, prevention activities, and screening participation by those at risk for cancer. Substantial enhancements should be pursued in both broad domains.

First, as demonstrated by the Prostate Cancer Outcomes Study, Cancer Care Outcomes Research & Surveillance Consortium, and a host of important although less ambitious investigations, it is feasible and fruitful to use population-based cancer registries as a sampling frame for the collection of PRO data on dimensions of health-related quality of life, perceptions and satisfaction with care, and economic burden. Such data would provide the basis for quality-of-care assessments that are inherently outcomes oriented, such as evaluating success at pain management and symptom control, whether as recorded at the site of care or through periodic interview-based surveillance. In addition, the systematic and strategic collection of PRO data over time will enrich the evidence base for appraising process–outcome relationships (Fig. 1), as the IOM has emphasized.

Although the application of PRO data to quality-of-care assessment remains the exception, the use of survey data to capture health behaviors and the use of prevention and screening services are largely the rule in quality evaluation along the pre-diagnostic portions of the cancer continuum. In particular, the CDC’s population-based Behavioral Risk Factor Surveillance System, which uses a state-based sampling frame, is a principal source of data for several prevention (e.g., quit smoking efforts) and screening (e.g., mammography use) measures in the Georgia comprehensive cancer control plan. But there are inherent limitations with such data. These include sample sizes often too small for robust area-based, covariate-adjusted calculations and, more fundamentally, the inability to validate self-reports with data on actual utilization of services. Within the Georgia state plan implementation committee, this has led to discussions about expanding the sample sizes of such surveys and also using administrative/claims files to obtain population-representative data on coded utilization of prevention and screening services. As with virtually all proposals to improve the quality of quality-of-care assessment, additional resources are required on a sustained basis. In the following section, we make a start at addressing such matters.

TOWARD A SUSTAINABLE STATE-LEVEL DATA INFRASTRUCTURE FOR CANCER QUALITY ASSESSMENT AND RESEARCH

The challenges and the potential facilitators for building and sustaining a state-level cancer data system can each be mapped into 1 of 3 broad domains, we believe. These might be characterized as technical and methodological (“information engineering”); financial (how will the infrastructure be supported?), and management and governance (how do data
contributors, sponsors, and users work together productively over time?). These are complex topics in their own right, and the brief discussion in the following section is best seen as a gateway to further inquiry:

**Technical and Methodological Issues**

It is premature at this stage to propose an “optimal” architecture for a state-level cancer data system, in Georgia or elsewhere. That said, we have moved well beyond proof-of-concept for many, if not most, of the key ingredients needed for an effective platform for data collection, exchange, and analysis:

- Decision processes for identifying evidence-based quality-of-care measures that capitalize on the types of data available from a state-level system have been developed and successfully applied in the field.\(^{11-13}\)

- There are already common vocabularies, widely in use, for virtually all of the concepts and variables required for cancer quality measurement, comparative effectiveness research, costing, and economic evaluations. Consensus definitions for all registry-based variables are maintained (and updated over time) by the North American Association of Central Cancer Registries,\(^ {25}\) setting common standards for SEER, NPCR, and NCDB registries. Complementing this are the standardized terminology and variable definitions from the CAP (College of American Pathology) Protocol and from LOINC (the Logical Observation Identifiers, Names and Codes) for laboratory and clinical observations.\(^ 2\) For more than 30 years, standardized variable definitions and coding for administrative/claims data have been published by the National Uniform Billing Committee, created by the American Hospital Association and including now about 20 member organizations representing all public and private payers in the United States, as well as the American National Standards Institute.\(^ {34}\) To be sure, there are not yet consensus definitions for such constructed variables as “race/ethnic” status or “insurance status,” which may be important in quality analyses or in comparative effectiveness research. But even there, both the North American Association of Central Cancer Registries and UB-04 record layouts contain standardized codes for the “atomistic” data elements needed for such constructed variables.

Most of the variable definitions that are fundamental to the conduct of cancer population sciences research have now been submitted for inclusion as Common Data Elements within the NCI’s cancer Biomedical Informatics Grid (caBIG).\(^ {35}\) In reality, consensus definitions on these concepts and variables, and well-defined processes for revising and updating over time, long preceded caBIG and will continue apace, whatever the path forward for this NCI initiative. Ongoing work by the caBIG Population Sciences Special Interest Group can only strengthen efforts to standardize all important measurement concepts.\(^ {36}\)

- There are standardized approaches for linking registry and administrative/claims data, using either deterministic or probabilistic methods.\(^ {29,37}\)

- Procedures for deidentifying and reidentifying sensitive data and for transmitting information securely have been demonstrated in multiple studies, including the current Georgia project.
The development of EHR systems that are compatible with cancer quality measurement—and the parallel development of measures that are compatible with the data collected by EHRs—is a dynamically evolving work in progress. But positive forces are at play. First, to the extent that we move toward a consensus set (or sets) of well-specified core measures of quality, we ipso facto map out the scope of the data elements that need to be collected for computing all the required numerators and denominators. By summer 2011, there were 67 NQF-endorsed cancer quality measures\(^3\) (far more now than the original core set shown in Table 1). Second, work is underway at the NQF to develop a “Quality Data Model”\(^3\) to ensure that approved quality measure specifications use the type of data that can be obtained from EHRs. In turn, this can only enhance the likelihood that EHR vendors will incorporate the type of data items often required by quality measures.

Although eventual broad and wide adoption of EHRs will promote efficiency and accuracy in cancer quality measurement, we note that the CoC’s RQRS does not depend on EHRs, but has been designed flexibly to accommodate any combination of electronic and paper-based data sources. Indeed, RQRS—and “RQRS-like” reporting and feedback systems—can become a centerpiece for a state-based rapid learning system for cancer care. As noted earlier, the 65 CoC cancer programs participating in the RQRS beta test have strongly applauded the effort. The challenge is how this type of data collection and exchange system can be extended to all significant cancer provider sites across a state (or the nation).

Office-based oncology providers remain a particular challenge. Over time, it will be interesting to see whether ASCO’s QOPI (Quality Oncology Practice Initiative)\(^4\)—which currently has more than 900 oncology practices across the country reporting voluntarily on multiple cancer quality measures, in conjunction with a new practice certification program\(^4\)—could evolve in a way that also promotes interpractice learning within a state or across states. One practical barrier, compared with the hospital-based RQRS, is that very few office-based oncology practices routinely report to the state cancer registry or have close communication ties. At this stage, both RQRS and QOPI actively promote only “bilateral communication” between the reporting oncology provider and the accrediting organization. In a state-level rapid learning system, one would also want to promote information sharing and cross-talk among providers. Effective ways to accomplish this remain to be explored.

**Financing the Enterprise**

To build and sustain a state-level data system for cancer quality assessment and research requires, in essence, some form of financial management plan and resource development strategy. Broadly speaking, we see 3 general approaches, which vary in terms of the assumed public funding available for data infrastructure development:

1. **Largely Self-sustaining, Without Significant Start-Up Funding.** At present, this perhaps aptly characterizes the Georgia project above, which has moved forward entirely on the basis of a competitive award Association of Schools of Public Health/CDC grant, funded largely by the NCI with additional support from the Georgia Cancer Coalition. For this funding “model” to succeed, there must be a significant, sustained demand for the research and policy-informing products
emerging from the data system. Specifically, support could come directly or indirectly as a result of: research grants and contracts from federal agencies (e.g., NCI, CDC, AHRQ) for a range of cancer health services research and outcomes analyses, including on quality of care; public agencies, private insurers, and self-insured purchasers that intend to use performance on designated quality metrics as one basis for provider payment; industry-supported projects on the safety, effectiveness, and the value-added of pharmaceutical and medical device products, including phase IV (post–Food and Drug Administration approval) studies; health care management and consulting firms that might find value in large-sample, population-based cancer data for studies being undertaken for public- or private-sector clients; and in-state users of cancer data, including health agencies and comprehensive cancer control planning operations.

Pay-for-performance efforts are already well underway, most notably in CMS’s “Physician Quality Reporting System” initiative that adjusts physician payments under Medicare according to performance on a range quality-of-care measures, including (in 2011) at least 14 in the cancer care domain. Such pay-for-performance mechanisms, public or private, would not be expected to channel dollars directly into a state cancer data system. Rather, cancer providers and their respective professional organizations would be supportive of such data infrastructure development because it is instrumental to generating credible information on performance.

All this said, a recurring challenge with this “build it and they will come” approach is that most grants and contracts supporting investigator-initiated, hypothesis-driven research provide little support for data infrastructure development. Rather, they are budgeted to cover roughly the incremental cost of the proposed research plan (and frequently fail to do that, given current funding constraints).

2 **Eventually Self-sustaining, But With Significant Start-Up Funding.** A prime example is the new state-level “Integrated Cancer Information and Surveillance System” being assembled by cancer prevention and control leaders at the University of North Carolina’s Lineberger Comprehensive Cancer Center. Intended eventually to support analyses of service delivery and outcomes across the cancer continuum, the Integrated Cancer Information and Surveillance System is underwritten during the current, build-up phase by the University Cancer Research Fund—created by the State of North Carolina expressly to grow the cancer research and application enterprise at the University of North Carolina. Similarly, the Georgia Cancer Coalition created several years ago a “Georgia Cancer Quality Information Exchange,” whose demonstration projects at selected cancer centers in the state will eventually inform extensions of the Georgia project described in this article.

3 **Sustained Public Support for Infrastructure Maintenance and Growth.** This approach recognizes the “public goods” nature of a large, population-based data system that can support multiple analyses of cancer quality, effectiveness, and costs. Specifically, no individual investigator could garner the resources and
expertise to build such a data system. Once built, use of the system by one investigator does not deter use by others. Indeed, investigators learn from each other over time as experience with the data system accumulates. The intended payoff, in return for public support, is a stream of high-quality research products that can inform cancer policy. This is, in essence, the de facto rationale behind NCI’s landmark SEER–Medicare database.16

Of great potential relevance in this regard is the ongoing development and evolution of state-level all-payer claims databases (APCDs), being created primarily to promote value-based purchasing and greater efficiency in local health care markets.45 As defined by the National Association of Health Data Organizations and the Regional All Payer Healthcare Information Council, an APCD is a state-level information system created by state mandate that typically includes data from medical and pharmacy claims, insurance plan eligibility files, and provider files from public and private payers operating in the state. At least 10 states presently have in place or are launching APCDs, including Kansas, Maine, Maryland, Massachusetts, Minnesota, New Hampshire, Oregon, Tennessee, Utah, and Vermont; 3 other states (Louisiana, Washington, and Wisconsin) have started voluntary claims reporting and consolidation initiatives. At present, only the Maine APCD includes virtually all participating payers in the state: Medicare, Medicaid, commercial insurers and third-party administrators (for self-insured client firms), but most other participating states are moving in that direction. All-payer claims databases are funded generally through some combination of general state appropriations and industry fee assessments.

No state (to our knowledge) has yet linked central cancer registry data to an APCD system. But doing so would appear to be a potentially feasible, affordable pathway to creating and sustaining a state-level cancer data system of the type being pursued in the Georgia project above.

Whatever the specific funding approach, the sustainability of a state-level cancer data system will likely be enhanced by the new federal program to adjust payments to Medicare and Medicaid providers based on the degree of their “meaningful use” of qualifying EHR systems.46 This is the case for 2 reasons. First, encouraging hospitals and physicians (and thus cancer centers and office-based oncologists) to purchase and use EHRs meeting national criteria for data collection and exchange will promote the accurate, timely, and complete reporting of data to state cancer registries and also facilitate rapid learning at the provider level. Second, the meaningful use criteria for both hospitals and professionals call for reporting on quality-of-care metrics that include cancer measures already endorsed variously by the NQF, the American Medical Association’s Physician Consortium for Performance Improvement, and/or the National Committee for Quality Assurance.46 These are precisely the kind of evidence-based quality measures that a state-level cancer data system—built around a strong central registry linked to multiple administrative/claims data sources—can generate on a sustained basis. Hence, the data required to demonstrate meaningful use at the practice level are also the data a state cancer data system needs reported to it for population-level monitoring of cancer care quality.
Management and Governance

Although clearly important, questions about how a state-level cancer data system would be managed and governed lie in uncharted territory—in good part, because no such system yet exists.

Under the public goods model for financing the system, as exemplified by SEER–Medicare, management and governance reside principally with the sponsoring federal or state agency or agencies. Input may be sought periodically from the scientific and policy communities, but technical and administrative decisions rest with agencies.

The fundamental problem with applying this approach to a state-level cancer data system that intends to be population based is that many of the required administrative/claims data sources are owned by commercial insurers and managed care organizations. There may be good reasons in a given instance for a state health agency to manage or govern the entire cancer data system enterprise—but there would seem to be no legal, administrative, or even logical reason why this needs to be the case. Indeed, with data from public and private sources being used to create an array of bilateral linked data sets to support a range of applications (Fig. 3), the more natural model for management and governance would be one of shared responsibilities and shared decision making among all key parties. It would not be unreasonable for a commercial insurer to want a say in how the data it contributes will be combined with other data and used by analysts over time to assess quality, effectiveness, health outcomes, and economic consequences. As recent conversations with some insurers operating in Georgia inform us, the desire for participatory decision making within such a state cancer data system is not lessened by assurances that the data sets distributed to future investigators will be deidentified.

All the while, APCDs have been structured in a way that frankly asserts the practical advantages of centralized management and mandatory reporting by payers. Specifically, these state data systems are legislatively authorized and publicly controlled, with a state agency given the legislative authority to collect and disseminate data. Whether, on the other hand, a more consensual approach to management and governance might evolve for a state-level cancer data system—with state government, private payers, cancer providers, researchers, and others around the table—remains to be seen.

The outcome of these decisions about financing, management, and governance, perhaps more than the resolution of technical and methodological issues, will shape the development and sustainability of state-level cancer data systems that can support quality assessment and research, as recommended a decade ago by the IOM.1,2

Acknowledgments

Sources of Funding: This study was supported by the Association of Schools of Public Health and the Centers for Disease Control and Prevention, PEP award 2008-R-08, and the National Cancer Institute, P30 CA138292-01.

References


Cancer J. Author manuscript; available in PMC 2016 May 25.


43. Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill. [Accessed May 25, 2011] Welcome to ICISS!. Available at: http://iciss.unc.edu


**FIGURE 1.**
The classic paradigm for assessing quality of care. Source: Donabedian.⁹
FIGURE 2.
Cancer quality of care (QOC) improvement cycle. Adapted from Lipscomb J. Enhancing cancer registry data for quality-of-care assessment. Presented at the annual meeting of the International Association of Cancer Registries, Yokohama, Japan, October 2010.
FIGURE 3.
Linking Georgia cancer registry data to public and private sources.
<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Breast and Colorectal Cancer Quality Measures Emerging From Public-Private NQF Deliberations</strong>°</td>
</tr>
</tbody>
</table>

**Breast cancer**
- Radiation therapy is administered within 1 y (365 d) of diagnosis for women younger than 70 y receiving breast-conserving surgery †
- Combination chemotherapy is considered or administered within 4 mo (120 d) of diagnosis for women younger than 70 y with American Joint Committee on Cancer (AJCC) T1cN0M0, or stage II or III hormone receptor–negative breast cancer †
- Tamoxifen or third-generation aromatase inhibitor is considered or administered within 1 y (365 d) of diagnosis for women with AJCC T1cN0M0, or stage II or III hormone receptor–positive breast cancer †

**Colon cancer**
- Adjuvant chemotherapy is considered or administered within 4 mo (120 d) of diagnosis for patients younger than 80 y with AJCC stage III (lymph node positive) colon cancer †
- At least 12 regional lymph nodes are removed and surgically examined for resected colon cancer ‡

**Rectal cancer**
- Radiation therapy is considered or administered within 6 mo (180 d) of diagnosis for patients younger than 80 y with clinical or pathological AJCC T4N0M0 or stage III receiving surgical resection for rectal cancer

° Each of these 6 measures is operationalized by the corresponding implied numerator–denominator specification, after application of appropriate inclusion/exclusion criteria. The measures were all submitted by the CoC to the NQF in its original “call for measures” in 2004. All measures are intended for performance assessment at the hospital or system level, not at the individual physician level. The technical specifications and operational definitions of the 6 measures were subsequently synchronized in a collaborative process hosted by the NQF and involving CoC, ASCO, and NCCN. The 3 breast cancer and 2 colon cancer measures were ultimately endorsed by the NQF in April 2007. This effectively concluded the NQF’s project to develop “National Voluntary Consensus Standards for Diagnosis and Treatment of Breast and Colon Cancer,” as supported jointly by NCI, AHRQ, CMS, and CDC. The rectal cancer measure here was developed independently by the CoC in conjunction with ASCO and NCCN.

† Endorsed by the NQF as an “accountability measure,” meaning the measure is deemed adequate for use in public reporting of performance, payment incentive programs, and selection of providers by consumers, health plans, or purchasers.

‡ Endorsed by the NQF as a “quality-improvement measure,” meaning the measure is to be used for internal monitoring of performance within an organization or group.
<table>
<thead>
<tr>
<th>Data Elements</th>
<th>GCCR File</th>
<th>MCARE Administration Files</th>
<th>MCAID Eligibility File</th>
<th>MCAID Claim Files</th>
<th>SHBP Plan Files</th>
<th>SHBP Enrollment Files</th>
<th>SHBP Claim Files</th>
<th>KPG Enrollment Files</th>
<th>KPG Clinical Data</th>
<th>GA HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient ID</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Date of birth</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sex</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Race</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Address</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MCAID ID</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Social security number</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Health plan type</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Primary cancer site</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>No. cancer sites</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Date of diagnosis</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cancer stage</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Method of diagnosis</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Date of death</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cause of death</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Insurance plan detail</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCAID eligibility</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCARE eligibility</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dates of coverage</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of coverage</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health maintenance organization enrollment dates</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnoses</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICD-9 codes</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CPT codes</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Input, output, and provider services</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Provider ID</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Pharmacy services</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skilled nursing facility services</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charges</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amount paid</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dates of service</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Revenue center codes</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Data Elements</td>
<td>GCCR File</td>
<td>MCARE Administration Files</td>
<td>MCAB Eligibility File</td>
<td>MCAB Claim File</td>
<td>SHBP Plan File</td>
<td>SHBP Enrollment File</td>
<td>SHBP Claims File</td>
<td>KPG Enrollment File</td>
<td>KPG Clinical Data</td>
<td>GA HDD</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Diagnosis related group</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provider codes</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

*Zip code level area variable.

ICD-9 indicates *International Classification of Diseases, Ninth Revision*. 
**TABLE 3**

Incident Breast and Colorectal Cancer Cases in Georgia, Over 1999–2005 From Selected Project Data Sources *

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Total Enrolled Population (e.g., for 2004)</th>
<th>Breast Cancer</th>
<th>Colorectal Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCCR (total GA incidence cases)</td>
<td>N/A</td>
<td>35,835</td>
<td>25,190</td>
</tr>
<tr>
<td>GCCR–Medicare</td>
<td>504,000</td>
<td>10,622</td>
<td>11,461</td>
</tr>
<tr>
<td>GCCR–Medicaid</td>
<td>362,390</td>
<td>7137</td>
<td>2235</td>
</tr>
<tr>
<td>SHBP</td>
<td>678,751</td>
<td>2863</td>
<td>1085</td>
</tr>
<tr>
<td>KPG</td>
<td>285,000</td>
<td>1203</td>
<td>456</td>
</tr>
</tbody>
</table>

* Incident cancer cases in the Georgia Hospital Discharge Data Set, not enumerated here, will include a substantial portion of the cases already included in the counts by payer source shown here.

N/A indicates not applicable.