# Considerations for Observational Research using Large Datasets in Radiation Oncology

**Dr. Reshma Jagsi, MD, DPhil [Associate Professor of Radiation Oncology] [Associate Chair for Faculty Affairs]**,
Department of Radiation Oncology, University of Michigan, Ann Arbor, MI, USA

**Dr. Justin E. Bekelman, MD [Assistant Professor]**,
Departments of Radiation Oncology and Medical Ethics and Health Policy, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

**Dr. Aileen Chen, MD [Assistant Professor]**,
Department of Radiation Oncology, Harvard Medical School, Boston, MA, USA

**Dr. Ronald C. Chen, MD, MPH [Assistant Professor]**,
Department of Radiation Oncology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA

**Dr. Karen Hoffman, MD [Assistant Professor]**,
Department of Radiation Oncology, Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

**Dr. Ya-Chen Tina Shih, PhD [Assistant Professor]**,
Department of Medicine, Section of Hospital Medicine, The University of Chicago, Chicago, IL, USA

**Dr. Benjamin D. Smith, MD [Assistant Professor]**, and
Department of Radiation Oncology, Division of Radiation Oncology, and Department of Health Services Research, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

**Dr. James B. Yu, MD [Assistant Professor of Therapeutic Radiology]**
Yale School of Medicine, New Haven, CT, USA

## Abstract

The radiation oncology community has witnessed growing interest in observational research conducted using large-scale data sources such as registries and claims-based datasets. With the

growing emphasis on observational analyses in health care, the radiation oncology community must possess a sophisticated understanding of the methodological considerations of such studies in order to evaluate evidence appropriately to guide practice and policy. Because observational research has unique features that distinguish it from clinical trials and other forms of traditional radiation oncology research, the Red Journal assembled a panel of experts in health services research to provide a concise and well-referenced review, intended to be informative for the lay reader, as well as for scholars who wish to embark on such research without prior experience. This review begins by discussing the types of research questions relevant to radiation oncology that large-scale databases may help illuminate. It then describes major potential data sources for such endeavors, including information regarding access and insights regarding the strengths and limitations of each. Finally, it provides guidance regarding the analytic challenges that observational studies must confront, along with discussion of the techniques that have been developed to help minimize the impact of certain common analytical issues in observational analysis. Features characterizing a well-designed observational study include clearly defined research questions, careful selection of an appropriate data source, consultation with investigators with relevant methodological expertise, inclusion of sensitivity analyses, caution not to overinterpret small but significant differences, and recognition of limitations when trying to evaluate causality. This review concludes that carefully designed and executed studies using observational data that possess these qualities hold substantial promise for advancing our understanding of many unanswered questions of importance to the field of radiation oncology.

## INTRODUCTION

The radiation oncology community has witnessed growing interests in observational research conducted using large-scale data sources such as registries and claims-based datasets. Given the low barriers to accessing certain sources of such data and the recent emphasis on understanding the "real-world" outcomes of medical interventions, studies relying on registry and/or claims data have proliferated in the Red Journal and the general oncology literature in recent years.

Although the literature of health services research contains detailed resources to guide the design and interpretation of such studies,[12345678910111213] and some reviews have considered specific topics of interest to the radiation oncologist, [141516171819] we are aware of no single, comprehensive overview specifically targeted towards the radiation oncologist who seeks to conduct or interpret such studies. Because observational research has unique features that distinguish it from clinical trials and other forms of traditional radiation oncology research, the Red Journal assembled a panel of experts in health services research to provide a concise and well-referenced overview, intended to be informative for the lay reader of the Red Journal, as well as for scholars who wish to embark on such research without prior experience.

In this manuscript, we begin by discussing the types of research questions relevant to radiation oncology that large-scale registry data may help illuminate. We then describe major potential data sources for such endeavors, including information regarding access and insights regarding the strengths and limitations of each. Finally, we provide guidance

regarding the analytic dilemmas that observational studies must confront, along with discussion of the techniques that have been developed to help minimize the impact of certain common challenges in observational analysis.

## RESEARCH QUESTIONS

Many different kinds of studies can be performed with registry and/or claims data. These data have two distinctive advantages over clinical trials data: 1) the large sample size of the database and 2) the "real world" nature of the data. This section outlines several types of studies that can be explored with such data. This is by no means an exhaustive list, and other novel and creative uses are certainly possible.

### Rare cancers and rare events

The characterization of the incidence and survival of patients diagnosed with rare cancers is one potential project that can be performed with large observational datasets. In particular, cancers which arise in a rare site[20] or a rare histology[21] benefit from the characterization of relatively straightforward concepts such as survival and patterns of presentation that would otherwise be difficult to obtain in single institution experiences. Large observational databases may also be useful in studies exploring the potential role for radiotherapy in the management of rare diseases. However, it is important to recognize that these studies are limited by variations in the expertise of the diagnosing pathologist, as there may be significant limitations for both sensitivity and specificity when attempting to accurately determine a rare diagnosis. In addition, these data are useful for characterizing rare events like radiation-induced second malignancies.[22]

### Changes in practice patterns

As radiation oncology is a rapidly changing field, the investigation of changes in radiation practice patterns over time is a particularly fertile area of research. For example, whether treatment patterns have changed in response to clinical evidence,[23] randomized trials,[2425] or clinical guidelines[26] is important in order to identify areas where further education is needed.

In particular for radiotherapy, investigating the dissemination of new technologies is also important, though dependent on the ability of the data to record whether new technology has been delivered. Dissemination of intensity modulated radiation therapy (IMRT)[27] and adoption of computed tomographic (CT) treatment planning[28] for lung cancers are two examples of areas where registry data have led to useful insights to the sociodemographic factors associated with the diffusion of important technology. Often, studies of new technology dissemination are limited by whether the data collected by registries or present in administrative claims are specific enough to identify the treatment to be studied. For example, studies of the use of accelerated partial breast irradiation using claims data have had to focus on brachytherapy utilization because the receipt of external beam partial breast irradiation cannot reliably be differentiated from incomplete courses of whole breast irradiation in claims data alone.

## Prediction models and nomograms

The creation of predictive models and nomograms is an area of potential research that is also dependent on a data source with diverse patient information. Some of the earliest pioneering work using cancer registry data was the characterization of survival of men following conservative management of prostate cancer, the so-called "Albertson tables."[29] This work has subsequently been updated,[30] and as the manner and stage in which patients present over time continues to change, updating of these tables will be needed.

Nomograms predicting benefit from post lumpectomy radiotherapy for older patients with breast cancer have been developed[31] to define subpopulations of patients who would have larger benefit from treatment.[32] These types of studies are valuable and give patients more specific insights into the outcomes from patients similar to them.

## Comparative effectiveness

Perhaps the most interesting area of current research using observational data in radiation oncology is to explore the comparative effectiveness of new technologies in the treatment of cancer. Observational studies using registry and/or claims data are particularly important where clinical trials are not forthcoming, will not be available during a time of rapid technology adoption, or will not be conducted at all. Furthermore, randomized trials can sometimes prioritize internal validity (the measurement of cause and effect with minimized bias within the study sample) over external validity (the application of study results to patients in the source population who are not included in the study sample). The external validity of randomized studies may be limited if the treatment arms, outcomes, or follow up periods tested are different or unrelated to current clinical practice[33] or if the participants are not representative of the patient population, a concern frequently voiced for elderly cancer patients.[34] Large observational studies are advantaged by their ability to draw information from a diverse array of patients, providers, and treating facilities in real world practice.

Examples of comparative effectiveness research where clinical trials already exist include IMRT compared to 3D conformal radiotherapy for head and neck cancers.[35] As for randomized patients, survival of patients with IMRT for head and neck cancers in community practice has been shown equivalent to 3D conformal RT[36] though continued diligence as to complications from treatment is needed.[37] Examples of comparative effectiveness where relevant randomized trials are ongoing include the comparative effectiveness of different types of radiotherapy for prostate cancer, including the comparison of 3D CRT and IMRT[38] and proton radiotherapy[39][40][41] and the comparison of breast brachytherapy to standard breast radiotherapy.[42][43] It is notable that these examinations of comparative effectiveness are limited by how up-to-date the data are, whether specific treatment codes are available, and whether complications are adequately captured by administrative claims. In particular, it is important to note that the diffusion of certain technologies may be limited in the early years of use, in which case even the experiences in a multi-institutional registry may reflect treatment at only one or a few select centers.

In summary, there are a wide variety of potential studies to be performed using registry or claims data. These studies typically take advantage of the large sample size and treatment

diversity available in such data. These studies are helpful in delineating the treatment and presentation of cancers both rare and common, understanding changes and variations in treatment patterns, and particularly helpful in investigating the comparative effectiveness of new technologies both where randomized trials have not yet been completed, or where external validation is needed.

## DATA SOURCES

Numerous potential datasets exist that are suitable for the investigation of questions relevant to the practice of radiation oncology (Table 1). In this section, we describe several commonly used data sources, with attention to the relative strengths and limitations of each.

### The Surveillance, Epidemiology, and End Results (SEER) Data

The Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute is a collection of population-based cancer registries which began collecting data on January 1, 1973, in Connecticut, Iowa, New Mexico, Utah, Hawaii, Detroit, and San Francisco/Oakland.[44] Over time, additional registries were added and SEER program now includes a total of 20 registries covering approximately 28% of the US population.[44] The data collected include: demographics (race and ethnicity, age at diagnosis, month and year of diagnosis, gender, marital status, state and county of residence at time of diagnosis), tumor site, tumor morphology, stage at diagnosis, first course of treatment (information on surgery and radiation but not systemic therapy), and survival (including cause of death from the death certificate). SEER data can be obtained free-of-charge; a signed data use agreement form is required, and data request can be found on the SEER website (http://seer.cancer.gov/data/). The website also provides data analytic tools.

The population-based nature of SEER makes it a powerful data source to examine temporal changes in cancer incidence, patterns of care, and survival, with results generalizable to the US population. The limitations mainly relate to the lack of certain data elements, including: systemic therapy, treatment after the first course, disease recurrence, comorbidities and functional status, and health related quality of life (HRQOL). Further, recent data suggest that SEER may underascertain post-operative radiation therapy.[4546] This can lead to spurious findings of underutilization or variation in the use of radiation therapy that are in fact artifacts of incorrect data.[47] Therefore, there has been interest in combining the detailed pathologic information in SEER with other data, such as claims or surveys, as described in subsequent sections.

### SEER-Medicare

Data from the SEER cancer registries have been linked with Medicare administrative claims data to provide enriched information about Medicare beneficiaries with cancer.[4849] The linked SEER-Medicare dataset contains information from the SEER registries (patient demographics, tumor characteristics, initial treatment and vital status) as well as billing claims for fee-for-service Medicare beneficiaries. The claims contain diagnostic and billing codes generated for services covered by Medicare including tests, procedures, office visits, admissions, durable medical equipment, home health, hospice care, and prescription drugs.

The most recent linkage, performed in 2012, includes cancer diagnoses through 2009 and Medicare claims through 2010.[49] Medicare covers individuals younger than 65 if they are disabled or have end-stage renal disease, however, most investigators use SEER-Medicare to study individuals age 65 and older. Since Medicare insures nearly all individuals age 65 and older, the data are fairly representative of national cancer care for older cancer patients enrolled in fee-for-service Medicare in the United States.

The linked SEER-Medicare claims provide insight about patient medical conditions, pretreatment evaluation, initial treatment and subsequent medical care that is not available in SEER data alone. Medical diagnosis codes in claims generated the year prior to cancer diagnosis are commonly used to characterize patient comorbid medical conditions.[48] Claims for imaging services can be used to evaluate assessment at diagnosis and surveillance after treatment. For example, procedure codes have been used to study the impact of breast MRI on type of breast cancer surgery and to evaluate the frequency of surveillance colonoscopy in colorectal cancer survivors.[50,51] Claims can identify treatments not captured in SEER including chemotherapy administration, androgen deprivation therapy administration, and type of reconstructive surgery.[52,53,54] Although SEER collects data regarding administration of radiation during the first course of treatment, the linked Medicare claims can identify radiation administration not captured in SEER[46] and provide additional granularity regarding type of radiation administered.[55] The linked SEER-Medicare database has been used to compare the use and outcomes of 3D-conformal radiotherapy, IMRT, proton radiotherapy, and stereotactic radiosurgery, which would not be possible using SEER alone.[28,40,56,57] Since multiple years of claims data are available, beneficiaries can be followed longitudinally over time. The longitudinal nature of claims data facilitates the study of events that occur after initial diagnosis and treatment such as procedures for post-treatment complications, surveillance follow up evaluation, cancer treatments for recurrence, and end-of-life care.[55,58,59] Because claims data reflect billed medical services, they are also uniquely suited to study the cost of care.[2]

The linked Medicare claims provide additional information about the health, treatment, and outcomes of cancer patients; however, research using the linked claims has distinct limitations as well. First, it is important to remember that claims data are administrative data not clinical data. Claims only contain services covered by and billed to the insurer. Medical care not billed to Medicare, such as services provided to beneficiaries by the Veterans Administration, Military Medical System, or community-sponsored screening programs are not captured. Second, claims are generated for billing. Information needed to obtain payment, such as the procedure performed, is of higher quality than information not required to obtain payment, such as secondary diagnoses. Third, claims reflect care that patients received rather than care that patients may have needed. Therefore, studies that use administration of salvage cancer treatment as a surrogate for cancer recurrence may miss patients whose cancer recurred but who declined salvage treatment. Fourth, there is limited information on why a patient did or did not receive a medical service. Fifth, although post-treatment complications can be inferred from diagnostic and billing codes, there are no patient-reported outcomes in claims data. Sixth, data for individual claims are only available for fee-for-service Medicare beneficiaries. Therefore, most studies using SEER-Medicare are limited to fee-for-service beneficiaries with continuous enrollment, despite the fact that

patients and their care may differ between the fee-for-service and managed care settings.[60] Seventh, prescription drug data are available only since 2006 onwards and for those beneficiaries with Part D Medicare coverage. Eighth, there is no follow up for cancer recurrence from the SEER registry data. Ninth, researchers can identify receipt of tests covered by Medicare (e.g., PSA, KRAS) but do not know the results of the tests. Finally, Medicare claims provide information about older patients with cancer but patterns of care and outcomes may be different for younger patients.[61]

Despite these limitations, SEER-Medicare data offer a unique opportunity to evaluate medical care for older cancer patients prior to diagnosis, during initial treatment, and after treatment, as well as the cost and quality of care. The data can be linked to other sources as well, including the Area Resource File to study the impact of health resources on medical care, the American Medical Association Physician Masterfile to study physician characteristics, and the American Hospital Association Annual Survey Database to study hospital characteristics. [3] Of note, physician characteristic linkages are imperfect, [4] and one study identified potential misclassification in particular for radiation oncologists.[62] Restricted geographic identifiers, i.e., census tract of residence, also are available for linkage to US Census tract-level data. A random sample of Medicare beneficiaries who do not have cancer is also available for comparative analyses.

Researchers interested in conducting studies utilizing SEER-Medicare data must obtain approval from the National Cancer Institute for specific research questions and complete a data use agreement (DUA). It should be noted that the DUA is strictly tied to the specific questions approved by the NCI and researchers planning to use the same data to study questions outside the scope of the DUA have to submit a new proposal to seek approval for a new DUA. In addition, manuscripts using SEER-Medicare data have to be approved by the NCI prior to submitting to a journal. The cost of the data depends on the number of files requested, and a separate level of approval is necessary to obtain access to data like restricted geographic identifiers. Separate files contain data on hospital inpatient services, hospital outpatient services, physician payments, home health, hospice, skilled nursing facilities and outpatient prescription drugs. The data files are large and complex and require computing power and advanced data management skills to create an analytic dataset. Given the complexities of the data, it is helpful to work with someone who has experience with the data when first using SEER-Medicare data. Of note, new measures are added to SEER-Medicare over time (e.g., AJCC M1 substaging for site of metastasis, HER2 status in breast cancer) to permit more granular sample definition; typically, the year in which the measure is first made available is excluded when conducting analyses using the measure. The NCI periodically offers videocasts of training workshops that may be particularly useful resources for beginners interested in working with this data source.[63]

### Claims-based datasets

Medical claims data that have not been linked to cancer registry information encompass a broader population. In contrast to the linked SEER-Medicare database which is limited to beneficiaries diagnosed in a SEER registry (approximately 26% of the population), the Chronic Conditions Warehouse is a nationally comprehensive database that contains 100%

of all Medicare fee-for service claims for all beneficiaries. Claims for younger patients can be obtained from several private sources, such as the MarketScan database which contains inpatient, outpatient and pharmaceutical claims from more than 100 payors in all 50 states.[60] Other proprietary commercial claims data include IMS LifeLink data, United Healthcare claims data, Humana claims data, among others. Although the commercial claims data provide information on a younger population, only privately insured individuals are included and long-term follow up can be limited due to frequent changes in insurance plans in the younger population.

Claims-based datasets are subject to the limitations of SEER-Medicare data as outlined above. Additionally, claims-based algorithms must be applied to these datasets to identify incident cancer cases.[64] While some such algorithms have been validated, they inevitably have certain limitations and thus will omit some patients with the cancer to be studied and also include certain patients without a true diagnosis of cancer or mistaken prevalent cases as incident cases.[65] Accordingly, great care should be applied when posing study questions with claims-only data to ensure that findings will not be sensitive to this inherent limitation. Additionally, there is no information about cancer stage or cancer histology. Algorithms to derive stage from claims data have been developed for breast cancer, but are limited in their predictive power.[66] In certain cases, biologic information from the tumor may be derived, for example if a prescription for endocrine therapy is identified, it may be reasonable to infer that the patient's breast cancer was estrogen receptor positive. Another limitation of some claims-based datasets is that they may not include race/ethnicity data as a matter of company policy, and this precludes those data from use in studies of race disparities.

Certain limitations of claims-based algorithms can pose particular challenges for those investigating issues in radiation oncology. For example, algorithms to identify delivery of palliative radiation to the bone have also been developed using claims but have yet to be validated.[67,68,69] Studies of skeletal-related events using claims data are limited by absence of clinically confirmed diagnosis of bone metastasis, and prior literature documents concerns with misclassification when using billing codes to identify cancer recurrence or metastasis. [70,71] In addition, although claims data do allow for the identification of certain events relative to diagnosis or another index date (such as timing of face-to-face contact with a radiation oncologist, timing of receipt of radiation therapy), other measures that may be of interest to radiation oncologists cannot be determined, including site of radiation delivery (e.g., to the bone or organ) and length of radiation therapy (e.g., 4 week course or 6 week course—although the total number of treatment fractions billed can be useful in this regard). Researchers using claims data will often create episodes of care in order to properly characterize utilization and costs associated with treatment for a given event or purpose. For example, a patient experiencing a fracture may generate claims related to the initial fracture event over the course of several weeks or months as they receive treatment and return to visit their physicians for follow-up care. While claims data present opportunities to investigate real-world utilization patterns, it is not often straightforward to isolate utilization and costs associated with specific events or conditions using claims data.

Medicare claims data can be obtained from Centers for Medicare and Medicaid Services (CMS) through the Research Data Assistance Center (www.resdac.org) and require an

extensive data security plan and approval of the CMS Privacy Board. Commercial claims data can be obtained from their respective data vendor (e.g., Truven Health Analytics (www.truvenhealth.com) for MarketScan database) and licensing fees typically differ between researchers in academic and non-academic settings.

### Survey Datasets

Some researchers have sought to enhance population-based sources of data with surveys that elicit information from patients themselves. Strengths of survey data include the ability to collect information reflecting patients' own perceptions and experiences with care, as well as data relating to domains such as quality of life that are best evaluated using patient-reported outcomes measures. Growing interest in patient-reported outcomes[72][73][74] has led many registries to begin exploring ways in which to incorporate patient-reported data in their datasets. Limitations relate to the selection bias that can be introduced by non-response, and studies relying on survey datasets with low response rates should be viewed with considerable caution.[75]

The SEER-MHOS (Medical Health Outcomes Survey) linked database results from a collaboration between the National Cancer Institute and the Centers for Medicare & Medicaid Services (CMS).[76][77] MHOS is a survey administered to a sample of Medicare managed care beneficiaries: 1,200 beneficiaries are randomly selected per managed care plan each year (1,000 per year from 1998-2006). The latest dataset available at the time of press includes cases diagnosed through 2009, with follow-up through 2011. For each selected beneficiary, a baseline survey is mailed and followed up with a phone call if no response was received; a follow-up survey two years later is administered the same way. Data collected include participant-reported demographics (age, gender, race and ethnicity, marital status, household income, education, smoking status), comorbid conditions, functional status and general HRQOL (assessed using the Short-Form 36 in years 1998-2005, and Veterans RAND 12-item Health Survey from 2006 and later). To obtain linked SEER-MHOS data, a data request and proposal need to be submitted for review (http://appliedresearch.cancer.gov/surveys/seer-mhos/obtain/req.docs.html). If approved, there is a cost to obtain the dataset and other procedures that must be followed.

SEER-MHOS is a unique, population-based data source to examine HRQOL of cancer patients, and also includes a number of self-reported comorbid conditions and functional status measures – thus, this linked dataset improves on some important limitations of the SEER alone. However, SEER-MHOS contains only patients 65 years or older, and overall has a relatively limited sample size. Because MHOS was not designed specifically to study HRQOL in cancer patients, sampled participants may have no history of cancer; in participants with cancer, survey administration is often not coincident with the time of cancer diagnosis or treatment. Further, no disease-specific HRQOL measures are used and long-term longitudinal assessment of HRQOL changes is not possible. In addition, data quality for longitudinal assessment of HRQOL changes is poor because a follow-up survey is administered to the same cohort two years after the baseline survey and the follow-up survey is limited to beneficiaries who are still enrolled in Medicare managed care plans.

Other sources of SEER-data linked to surveys also exist, and access to these may be available through request to the investigators who have developed these initiatives. For example, the Cancer Surveillance Outcomes Research Team (CanSORT) is a collaborative group funded by the National Cancer Institute to survey cohorts of newly diagnosed breast cancer patients in selected SEER regions in order to evaluate their decision-making experiences, satisfaction, and quality of life.[78] Detailed patient data on treatments received are linked to SEER data on pathologic and clinical disease characteristics and have been used to illuminate patterns of radiation therapy receipt as well as potential causes for variation in care.[79]

The Cancer Care Outcomes Research and Surveillance Consortium (CanCORS), funded by the National Cancer Institute in collaboration with the Department of Veterans Affairs, represents another valuable source of patient survey data. It focuses on newly diagnosed patients with colorectal or lung cancer from geographically diverse populations and health care delivery systems nationwide and includes data from medical records, as well as physician and patient surveys.[80] CanCORS data have been used for a number of purposes, including comparison of practice patterns for palliative radiation to clinical evidence and evaluation of metastatic cancer patients' perceptions of the intent of radiation therapy.[81,82,83]

### National Cancer Data Base (NCDB)

The National Cancer Data Base is a joint program of the Commission on Cancer (CoC) of the American College of Surgeons and the American Cancer Society, and collects data from over 1,400 CoC-accredited cancer hospitals throughout the US. The NCDB contains data on approximately 70% of incident US cancer patients.[84] The NCDB was started in 1989, and by 1996, reporting of cancer cases was required of all CoC-accredited hospitals, which include teaching, community, and Veterans Health Administration hospitals. At each hospital, certified tumor registrars abstract data from patient medical records, and registrars are required to obtain and submit patient treatment and follow-up data even if part of the care is received at another (e.g. non-CoC-accredited) hospital. Annually, registrars upload data to the NCDB on incident cancer cases as well as follow-up information on existing patients. Data contained in the NCDB include: demographics (age, gender, race, marital status, medical insurance), comorbidity status, stage, treatments received, recurrence and survival.

NCDB data can be obtained by investigators from CoC-accredited programs through an application process (see http://www.facs.org/cancer/ncdb/participantuserfiles.html). If approved, data are provided free of charge.

The NCDB is a powerful source of data for research given its large size, and inclusion of information on use of systemic therapy, comorbidities, and recurrence. However, an important limitation is that the NCDB is not population-based thus limiting generalizability – data are entered from accredited hospitals, and this selection may bias findings toward higher quality care and patient outcomes.[85] Specifically, although cancer diagnoses at hospitals approved by the Commission on Cancer (CoC) constitute 70% of all new cancer diagnoses in the U.S., CoC-approved hospitals are larger, more frequently situated in urban locations, and have more cancer-related services available to patients. Further, patients who

are diagnosed and receive all of their cancer care in the outpatient setting without entering a CoC-accredited hospital are not included in the NCDB.

## Centers of Excellence

Many large institutions maintain their own hospital registries, and some have joined together in collaborative registries, such as the one maintained by the National Comprehensive Cancer Network (NCCN).[86] Registries from centers of excellence are often restricted to investigators with an affiliation with a member site. The strengths of such datasets include the potential for access to highly detailed clinical information that may even extend to the level of dosimetric or other potentially relevant granular data. For example, the NCCN database contains more than 300 data elements that track the continuum of care longitudinally, including information on all clinical interventions. Their primary limitation is their failure to include a population-based sample. Selection effects can be highly problematic; estimates based on registries that include primarily centers of excellence are vulnerable towards significant bias.

## Radiation oncology-specific registries

In recent years, radiation oncologists have begun to assemble registries that include both academic and community sites to gather detailed radiation treatment planning information that may be particularly relevant for the investigation of the comparative effectiveness of radiation therapy and that other registries have traditionally lacked. ASTRO and the Radiation Oncology Institute are embarking on the nascent National Radiation Oncology Registry (NROR), which aims to monitor and improve the quality of care for patients treated with radiation.[87][88] In the state of Michigan, Blue Cross Blue Shield has funded a Collaborative Quality Initiative known as the Michigan Radiation Oncology Quality Consortium (MROQC). MROQC collects detailed dosimetric, clinical, and patient-reported outcomes data from multiple academic and community sites across the state, on all patients treated with adjuvant whole breast irradiation after lumpectomy and radiation therapy for lung cancer with curative intent.[89]

As these endeavors grow and mature, they will serve as unique sources of information for those wishing to compare different forms of radiation treatment. Their primary limitation is the especially resource-intensive nature of collecting detailed dosimetric information relating to radiation treatment plans and the fact that they are still works in progress.

## Canadian, European, and Asian Oncology Databases

Although the preceding sections have focused on data sources that have been assembled in the United States, many countries outside the United States have robust registries or administrative claims data at the provincial or even national level, and these may be promising sources for the investigation of key research questions. A listing of some of these non-US oncology databases can be found at http://www.ispor.org/OncologyORResources/SearchOcologyResources.aspx.

# ANALYTIC ISSUES

## Nonrandom assignment

In oncology, we are often interested in comparing the effect of different interventions or treatments. Randomized studies are the gold standard for determining whether one treatment is superior to another. However, for practical or ethical reasons, randomized studies are not always possible. In these situations, well-conducted observational studies can provide useful information to influence decision-making. One common challenge in analyzing observational data is that treatments are not randomly assigned. Therefore, it is possible that characteristics of patients, providers, or healthcare systems, both observed and unobserved, can influence not only treatment assignment, but also outcomes. For example, sicker patients might be less likely to receive aggressive treatment, and they may also be more likely to have a poor outcome, regardless of the treatment received. Analyses not accounting for such favorable selection into the more aggressive treatment will thus overestimate the benefit of the treatment. For example, one study [90] of patients with prostate and colon cancer in the SEER-Medicare database found that analysis of these observational data produced improbable results. They found that men with locally advanced prostate cancer who underwent androgen deprivation had higher prostate cancer mortality (hazard ratio, 1.5; 95% confidence interval, 1.29-1.92). Clinical trials have provided strong evidence that androgen deprivation improves cancer mortality, so this finding almost certainly reflects selection bias.

## Traditional regression-based methods to address confounding

To deal with the issue of confounding due to nonrandom treatment assignment, several statistical and econometric techniques are commonly employed, including multivariable regression analyses that attempt to control for potential confounders in observational studies. Researchers often start with a series of descriptive bivariate analyses to explore the unadjusted correlation between each explanatory variable and the outcomes of interest, then follow with multivariable models to include the treatment variable (quantified as a binary variable with 1 indicating the treatment received by patients in the case group and 0 otherwise) plus other potential confounders in the same model. Variations on multivariable regression models, including stepwise selection algorithms and survival models for time-dependent effects, are also frequently used. In theory, if all important confounders were observed and controlled for, multiple regression techniques could produce similar results from an observational study that would be observed in a randomized study.

## Propensity score models

Standard multiple regression techniques are limited, however, by specific assumptions about the relationship between the covariates and outcome and also by the fact that not all confounders may be observable. In these situations, propensity score models and instrumental variable analysis can play an important role in improving the accuracy of statistical inferences.

Propensity score methods assign a probability of receiving the new treatment or the intervention received by patients in the case group to each individual based on observed

covariates, and can be implemented in several ways, including covariate adjustment and matching.[91][92] In covariate adjustment, propensity scores (predicted from a logistic regression) are simply entered in the regression model as an additional explanatory variable for the outcome. In propensity matching, strata are created by grouping individuals by propensity scores and calculating the treatment effect within those strata. Unlike standard multiple regression models, propensity scores do not assume a specific relationship between the covariates and outcome (e.g. linear or log linear), but rather assign a probability or "propensity" for each individual to be assigned to a treatment group, based on observed characteristics. Therefore, like multiple regression models, propensity score methods are limited by the assumption that all factors affecting treatment assignment are accounted for by observable covariates. Furthermore, propensity score matched models do not include patients who remain unmatched, and therefore require large datasets with significant overlap in patient characteristics. Such approaches have been used previously in numerous studies of topics relevant to radiation oncology.[28][40][57]

Of note, in the study described above, in which locally advanced prostate cancer patients who received androgen deprivation fared worse, controlling for comorbidity, extent of disease, and other characteristics by multivariate analyses or by propensity analyses had remarkably small impact on the improbable results.[90] Thus, caution must be employed when attempting to use observational data to evaluate the impact of treatments on patient outcomes, even when techniques to address measured confounding factors have been applied. Other methods that arise from the causal inference literature, including instrumental variable analysis, inverse probability treatment weights,[93] and marginal structural models[94] have also been proposed for observational comparative effectiveness research.

### Instrumental variable analysis

Instrumental variable analysis is an econometric technique to remove the effects of hidden bias[95] and, unlike propensity scores, does not assume that all potential confounders are observed. The key assumption is that the instrumental variable does not affect the outcome directly, except through treatment assignment. Therefore, a good instrumental variable should be strongly correlated with treatment assignment (referred to as "instrument relevancy" in the econometric literature), yet lacking an independent effect on the outcome (referred to as "instrument exogeneity"). If an instrument is weak, (i.e. the proportion of treatment assignments in the instrument groups are about the same) then the estimate of treatment effect becomes unstable.[96] In practice, it is often very difficult to identify a good instrumental variable for a particular treatment and outcome, which can limit the use of this technique. In certain cases, use of geographic variables, such as local area treatment rate[97] or differential distance to a treating facility,[43][98] have been appropriate instruments. Upon identification of a set of candidate instrumental variables, it is important to check the validity of these instruments. Once a set of valid instruments are identified, it is also important to verify whether the treatment variable is indeed "endogenous" (i.e., the treatment variable is correlated with the error term). If treatment variable is in fact exogenous, the use of instrumental variable method is unnecessary and can lead to efficiency loss.

Instrumental variable approaches do not estimate the average effect of treatment in a population, but rather the average effect for patients whose treatment was induced by the instrument. The treatment effect produced by instrumental variable analysis applies to what has been called the "marginal" or "complying" patient population, defined as patients whose treatment status depends strongly on the instrument. Research should describe which patients comprise the marginal patient population to inform the generalizability of findings from instrumental variable methods.[99] Several examples of recent articles in radiation oncology using these techniques exist.[40][97][98][100][101][102]

### Handling Missing Data

Missing data is a common problem in randomized and non-randomized studies. Data can be missing because patients are lost to follow-up, centers are unable to collect certain variables for all patients, or myriad other reasons. The main concern is whether missing data could bias study results; this concern is heightened in studies of comparative effectiveness when the level of missing data is different between study groups. Therefore, all studies must address how missing data were handled in analyses.

Many approaches exist to handle missing data; we mention three here. The first approach involves omitting everyone without complete data, and is known as complete case analysis. However, if a large proportion of patients have missing data, their exclusion could result in a substantial loss of statistical power and biased results, especially if the discarded cases differ systematically from those retained. A second approach involves entering missing data into regression models as dummy variables. The downsides of this approach are similar to complete case analysis and, when levels of missing data are greater than 3 to 5% (as a general rule of thumb), this approach has also been shown to yield biased results.[103] Multiple imputation, the third approach, is commonly available in statistical packages and is accepted as the most valid method for handling missing data. Multiple imputation can be thought of as a method to represent the uncertainty of missing data in regression models, and involves the creation and combination of multiple datasets produced through the imputation process. As with any statistical approach, the validity of the results of multiple imputation depends on careful and appropriate modeling.

In summary, we recommend the following best practices for handling missing data in observational research.[104] First, reports should describe the number of patients excluded for missing data and the number and proportion of missing values for each variable of interest. Second, reports should examine differences between patients with complete or incomplete data across exposure and outcome variables. Third, reports should describe the methods used to handle missing data and the assumption or rationale for the approach. Lastly, reports should discuss whether and how missing data might bias study results, even in the context of valid approaches used to handle missing data (particularly if missingness may differ between study groups in comparative analyses). Several articles can be referred to for further background.[105][106][107]

## Analysis of Medical Costs Data

Studies examining the economic impact of an illness or a new treatment often are interested in estimating medical costs associated with the illness or treatment. When preparing the data for cost analysis, it is important to normalize the cost data to the same year of currency (e.g., in 2013 US dollars) by applying the appropriate inflation adjustment index, such as the use of Medicare Prospective Payment System adjuster and the Medicare Economic Index for US costs data extracted from Medicare Part A and B claims, respectively. It should also be noted that the type of cost information available varies by datasets. Some provide information on payments as well as charges (e.g., Medicare claims), while other record only payment (e.g., MarketScan) or charges (e.g., Nationwide Inpatient Sample or NIS). When both charges and payments are reported in the data, it is recommended that researchers use the payment variable for medical costs since charges are often highly inflated from the actual costs. If only charge information is provided, cost-to-charge ratio should be applied to obtain better cost measure.

Medical cost data tend to be highly skewed, with a small proportion of very ill patients incurring excessively high costs; therefore, when comparing costs between two interventions, the commonly used t-test for continuous variables is often inappropriate because cost variable is unlikely to be normally distributed. The recommended approach to compare mean medical costs when cost data are highly skewed is the nonparametric bootstrap method suggested by Barber and Thompson.[108] The violation of normality assumption also complicates the multivariable regression models since estimates based on the standard ordinary least squares (OLS) technique can be severely biased if the dependent variable is highly skewed. Historically, researchers applied the log-transformed model to deal with this issue. However, more recent econometric literature has pointed out that despite its ease of implementation, the log-transformed model must take into consideration of the role of heteroscedasticity when retransforming the estimate from log(cost) back to cost.[109] Generalized linear models (GLM) offer a nice alternative to the log-transformed model and can avoid the retransformation issue. The challenge for the GLM model is that one must specify a mean and variance function and the best model to use depends on the characteristics of the data.[110] More advanced methods, such as the extended estimating equations (EEE) approach,[111] have been developed to relax the requirement of GLM models. In general, applied econometricians agree that no one model universally dominates the others under all circumstances.[111][112] Given these complexities, it is highly advisable that studies of medical costs include individuals with relevant econometric expertise among the investigator team. Furthermore, medical costs data can be censored due to reasons such as loss of follow-up (e.g., switch from a fee-for-service plan to a HMO) or end of data collection period, resulting in partial observation of costs for patients whose data are censored. Methods such as the inverse probability weighted least squares method,[113] Lin's regression method,[114][115] and Carides' two-stage method [116]have been proposed to address the censoring issue in medical costs data. A comprehensive review can be found in Huang.[117]

**Other Biases**

Finally, no discussion of the analytic challenges of observational research using large datasets would be complete without mention of the risk of spurious findings when many exploratory analyses are performed without pre-specification or theoretical justification. A statistically significant p value does not guarantee that the relationship observed is not due to chance, and the risk of finding chance associations increases when there is a "greater number and lesser preselection of tested relationships."[118] It is also important to note that the large size of observational datasets allows for the detection of relatively small differences; statistical significance does not imply clinical significance, and these issues must be addressed in the interpretation of any large-scale study, including observational studies.

## CONCLUSIONS

In summary, carefully designed and executed studies using observational data hold substantial promise for advancing our understanding of many unanswered questions of importance to the field of radiation oncology. Several common features characterize a well-designed observational study. These include clearly defined research questions, careful selection of an appropriate data source, collaboration or at least consultation with investigators with relevant methodological expertise, inclusion of sensitivity analyses, caution not to overinterpret small but statistically significant differences as clinically meaningful differences, and recognition of limitations when trying to evaluate causality.

Experts have published guidelines for the conduct of comparative effectiveness research, and radiation oncologists who seek to use registry and/or claims data for such studies should consider their advice. For example, Table 2 reproduces a checklist for study design of an observational comparative effectiveness research protocol from the Agency for Healthcare Research and Quality.[1] Other checklists are also available and may prove useful in evaluating the rigor of a study's approach.[6,7,8,119,120,121] For those seeking to evaluate the strengths and limitations of the increasing volume of observational research in radiation oncology, we hope that the information in this manuscript and the references provided will prove useful and instructive.

Understanding these issues is critical because ongoing developments in the social and policy environment are likely to lead to new opportunities and interest in observational research in radiation oncology in the coming years. For example, the creation of health information exchanges (secure repositories of patient health records) could support the development of a true "learning healthcare system"[122,123,124] in oncology care. The introduction of new automated techniques for data extraction from medical records is likely also to introduce new challenges and the need for ongoing evaluation of data quality and limitations. Other changes, such as modifications of payment structures, will create opportunities for leveraging the methods of observational research to the evaluate impact of policy changes on outcomes. With the growing emphasis on observational analyses in health care, a sophisticated understanding of the issues discussed herein are of critical importance to the radiation oncology community if we are to evaluate evidence appropriately to guide practice and policy.

## Acknowledgments

# REFERENCES

1. Velentgas, P.; Dreyer, NA.; Nourjah, P.; Smith, SR.; Torchia, MM., editors. Developing a protocol for observational comparative effectiveness research: A user's guide. Agency for Healthcare Research and Quality; Rockville, MD: Jan. 2013 AHRQ Publication No. 12(13)-EHC099http://www.effectivehealthcare.ahrq.gov/ehc/products/440/1166/User-Guide-to-Observational-CER-1-10-13.pdf

2. Brown ML, Riley GF, Schussler N, et al. Estimating health care costs related to cancer treatment from SEER-Medicare data. Med Care. 2002; 40(8 Suppl):IV-104–117.

3. Schrag D, Bach PB, Dahlman C, et al. Identifying and measuring hospital characteristics using the SEER-Medicare data and other claims-based sources. Med Care. 2002; 40(8 Suppl):IV-96–103.

4. Baldwin LM, Adamache W, Klabunde CN, et al. Linking physician characteristics and medicare claims data: issues in data availability, quality, and measurement. Med Care. 2002; 40(8 Suppl):IV-82–95.

5. Earle CC, Nattinger AB, Potosky AL, et al. Identifying cancer relapse using SEER-Medicare data. Med Care. 2002; 40(8 Suppl):IV-75–81.

6. Berger ML, Mamdani M, Atkins D, et al. Good research practices for comparative effectiveness research: defining, reporting, and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force report—part I. Value Health. 2009; 12:1044–1052. [PubMed: 19793072]

7. Cox E, Martin BC, Staa TV, et al. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report—part II. Value Health. 2009; 12:1053–1061. [PubMed: 19744292]

8. Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—part III. Value Health. 2009; 12:1062–1073. [PubMed: 19793071]

9. Korn EL, Freidlin B. Methodology for comparative effectiveness research: potential and limitations. J Clin Oncol. 2012; 30:4185–4187. [PubMed: 23071226]

10. Lyman GH, Levine M. Comparative effectiveness research in oncology: an overview. J Clin Oncol. 2012; 30:4181–4184. [PubMed: 23071249]

11. Hershman DL, Wright JD. Comparative effectiveness research in oncology methodology: observational data. J Clin Oncol. 2012; 30:4215–4222. [PubMed: 23071228]

12. Lyman GH. Comparative effectiveness research in oncology. Oncologist. 2013; 18:752–759. [PubMed: 23697601]

13. Ramsey SD, Sullivan SD, Reed SD, et al. Oncology comparative effectiveness research: a multistakeholder perspective on principles for conduct and reporting. Oncologist. 2013; 18:760–767. [PubMed: 23650020]

14. Virnig BA, Warren JL, Cooper GS, et al. Studying radiation therapy using SEER-Medicare-linked data. Med Care. 2002; 40(8 Suppl):IV-49–54.

15. Aneja S, Yu JB. Comparative effectiveness research in radiation oncology: stereotactic radiosurgery, hypofractionation, and brachytherapy. Semin Radiat Oncol. Jan.2014 24:35–42. [PubMed: 24314340]

16. Chen AB. Comparative effectiveness research in radiation oncology: assessing technology. Semin Radiat Oncol. 2014; 24:25–34. [PubMed: 24314339]

17. Meyer AM, Wheeler SB, Weinberger M, et al. An overview of methods for comparative effectiveness research. Semin Radiat Oncol. 2014; 24:5–13. [PubMed: 24314337]

18. Chen RC. Comparative effectiveness research in oncology: the promise, challenges, and opportunities. Semin Radiat Oncol. 2014; 24:1–4. [PubMed: 24314336]

19. Bekelman JE, Shah A, Hahn SM. Implications of comparative effectiveness research for radiation oncology. Pract Radiat Oncol. 2011; 1:72–80. [PubMed: 24673918]

20. Urdaneta AI, Yu JB, Wilson LD. Population based cancer registry analysis of primary tracheal carcinoma. American journal of clinical oncology. 2011; 34:32–37. [PubMed: 20087156]

21. Yu JB, Blitzblau RC, Patel SC, et al. Surveillance, Epidemiology, and End Results (SEER) database analysis of microcystic adnexal carcinoma (sclerosing sweat duct carcinoma) of the skin. American journal of clinical oncology. 2010; 33:125–127. [PubMed: 19675445]

22. Beard CJ, Travis LB, Chen MH, et al. Outcomes in stage I testicular seminoma: a population-based study of 9193 patients. Cancer. 2013; 119:2771–2777. [PubMed: 23633409]

23. Bekelman JE, Epstein AJ, Emanuel EJ. Single-vs multiple-fraction radiotherapy for bone metastases from prostate cancer. JAMA. 2013; 310:1501–1502. [PubMed: 24104375]

24. Soulos PR, Yu JB, Roberts KB, et al. Assessing the impact of a cooperative group trial on breast cancer care in the medicare population. J Clin Oncol. 2012; 30:1601–1607. [PubMed: 22393088]

25. Hoffman KE, Nguyen PL, Chen MH, et al. Recommendations for post-prostatectomy radiation therapy in the United States before and after the presentation of randomized trials. J Urol. 2011; 185:116–120. [PubMed: 21074194]

26. Kuykendal AR, Hendrix LH, Salloum RG, et al. Guideline-discordant androgen deprivation therapy in localized prostate cancer: patterns of use in the medicare population and cost implications. Ann Oncol. 2013; 24:1338–1343. [PubMed: 23277483]

27. Shirvani SM, Jiang J, Gomez DR, et al. Intensity modulated radiotherapy for stage III non-small cell lung cancer in the United States: predictors of use and association with toxicities. Lung cancer. 2013; 82:252–259. [PubMed: 24018022]

28. Chen AB, Neville BA, Sher DJ, et al. Survival outcomes after radiation therapy for stage III non-small-cell lung cancer after adoption of computed tomography-based simulation. J Clin Oncol. 2011; 29:2305–2311. [PubMed: 21537034]

29. Albertsen PC, Fryback DG, Storer BE, et al. Long-term survival among men with conservatively treated localized prostate cancer. JAMA. 1995; 274:626–631. [PubMed: 7637143]

30. Lu-Yao GL, Albertsen PC, Moore DF, et al. Outcomes of localized prostate cancer following conservative management. JAMA. 2009; 302:1202–1209. [PubMed: 19755699]

31. Albert JM, Liu DD, Shen Y, et al. Nomogram to predict the benefit of radiation for older patients with breast cancer treated with conservative surgery. J Clin Oncol. 2012; 30:2837–2843. [PubMed: 22734034]

32. Breast Cancer Nomogram to Predict Benefit of Radiation for Older Patients With Breast Cancer Treated With Conservative Surgery. www.mdanderson.org/RadiationBenefitPredictor

33. Presley CJ, Soulos PR, Herrin J, et al. Reply to L.W. Cuttino et al. J Clin Oncol. 2013; 31:2227–2229. [PubMed: 23901421]

34. Levit, LA.; Balogh, EP.; Nass, SJ.; Ganz, PA., editors. Delivering high-quality cancer care: Charting a new course for a system in crisis. Institute of Medicine; Washington, DC: 2013.

35. Nutting CM, Morden JP, Harrington KJ, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol. 2011; 12:127–136. [PubMed: 21236730]

36. Yu JB, Soulos PR, Sharma R, et al. Patterns of care and outcomes associated with intensity-modulated radiation therapy versus conventional radiation therapy for older patients with head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2012; 83:e101–107. [PubMed: 22342095]

37. Beadle BM, Liao KP, Chambers MS, et al. Evaluating the impact of patient, tumor, and treatment characteristics on the development of jaw complications in patients treated for oral cancers: a SEER-Medicare analysis. Head Neck. 2013; 35:1599–1605. [PubMed: 23150453]

38. Bekelman JE, Mitra N, Efstathiou J, et al. Outcomes after intensity-modulated versus conformal radiotherapy in older men with nonmetastatic prostate cancer. Int J Radiat Oncol Biol Phys. 2011; 81:e325–334. [PubMed: 21498008]

39. Kim S, Shen S, Moore DF, et al. Late gastrointestinal toxicities following radiation therapy for prostate cancer. Eur Urol. 2011; 60:908–916. [PubMed: 21684064]

40. Sheets NC, Goldin GH, Meyer AM, et al. Intensity-modulated radiation therapy, proton therapy, or conformal radiation therapy and morbidity and disease control in localized prostate cancer. JAMA. 2012; 307:1611–1620. [PubMed: 22511689]

41. Yu JB, Soulos PR, Herrin J, et al. Proton versus intensity-modulated radiotherapy for prostate cancer: patterns of care and early toxicity. J Natl Cancer Inst. 2013; 105:25–32. [PubMed: 23243199]

42. Smith GL, Xu Y, Buchholz TA, et al. Association between treatment with brachytherapy vs whole-breast irradiation and subsequent mastectomy, complications, and survival among older women with invasive breast cancer. JAMA. 2012; 307:1827–1837. [PubMed: 22550197]

43. Presley CJ, Soulos PR, Herrin J, et al. Patterns of use and short-term complications of breast brachytherapy in the national medicare population from 2008-2009. J Clin Oncol. 2012; 30:4302–4307. [PubMed: 23091103]

44. Surveilance Epidemiology and End Results Program. [Accessed on January 3, 2014] http://seer.cancer.gov

45. Jagsi R, Abrahamse P, Hawley ST, et al. Underascertainment of radiotherapy receipt in Surveillance, Epidemiology, and End Results registry data. Cancer. 2012; 118:333–341. [PubMed: 21717446]

46. Walker GV, Giordano SH, Williams M, et al. Muddy water? Variation in reporting receipt of breast cancer radiation therapy by population-based tumor registries. Int J Radiat Oncol Biol Phys. 2013; 86:686–693. [PubMed: 23773392]

47. Smith GL, Eifel PJ. In Regard to Han et al. Int J Radiat Oncol Biol Phys. 2014; 88:459–460. [PubMed: 24411614]

48. 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:IV-3–18.

49. SEER-Medicare Fact Sheet. [Accessed Jan 3, 2014] Dec. 2013 Available at: http://appliedresearch.cancer.gov/seermedicare

50. Fortune-Greeley AK, Wheeler SB, Meyer AM, et al. Preoperative breast MRI and surgical outcomes in elderly women with invasive ductal and lobular carcinoma: a population-based study. Breast Cancer Res Treat. 2014; 143:203–212. [PubMed: 24305978]

51. Brawarsky P, Neville BA, Fitzmaurice GM, et al. Surveillance after resection for colorectal cancer. Cancer. 2013; 119:1235–1242. [PubMed: 23184361]

52. In H, Jiang W, Lipsitz SR, et al. Variation in the utilization of reconstruction following mastectomy in elderly women. Ann Surg Oncol. 2013; 20:1872–1879. [PubMed: 23263733]

53. Shahinian VB, Kuo YF, Freeman JL, et al. Characteristics of urologists predict the use of androgen deprivation therapy for prostate cancer. J Clin Oncol. 2007; 25:5359–5365. [PubMed: 18048816]

54. Warren JL, Harlan LC, Fahey A, et al. Utility of the SEER-Medicare data to identify chemotherapy use. Med Care. 2002; 40:IV-55–61.

55. Smith BD, Pan IW, Shih YC, et al. Adoption of intensity-modulated radiation therapy for breast cancer in the United States. J Natl Cancer Inst. 2011; 103:798–809. [PubMed: 21525437]

56. Guadagnolo BA, Huo J, Liao KP, et al. Changing trends in radiation therapy technologies in the last year of life for patients diagnosed with metastatic cancer in the United States. Cancer. 2013; 119:1089–1097. [PubMed: 23132206]

57. Shirvani SM, Jiang J, Chang JY, et al. Comparative effectiveness of 5 treatment strategies for early-stage non-small cell lung cancer in the elderly. Int J Radiat Oncol Biol Phys. 2012; 84:1060–1070. [PubMed: 22975611]

58. Butler Nattinger A, Schapira MM, Warren JL, et al. Methodological issues in the use of administrative claims data to study surveillance after cancer treatment. Med Care. 2002; 40:IV-69–74.

59. Potosky AL, Warren JL, Riedel ER, et al. Measuring complications of cancer treatment using the SEER-Medicare data. Med Care. 2002; 40:IV-62–68.

60. Jagsi R, Jiang J, Momoh AO, et al. Trends and variation in use of breast reconstruction in breast cancer patients undergoing mastectomy in the United States. J Clin Oncol. 2014 Epub ahead of print.

61. Pan IW, Smith BD, Shih YC. Factors contributing to underuse of radiation among younger women with breast cancer. J Natl Cancer Inst. 2014; 106:djt340. [PubMed: 24317177]

62. Pollack LA, Adamache W, Eheman CR, et al. Enhancement of identifying cancer specialists through the linkage of Medicare claims to additional sources of physician specialty. Health Serv Res. 2009; 44(2 Pt 1):562–576. [PubMed: 19207588]

63. SEER-Medicare Training. National Cancer Institute; Oct. 2013 Updatedhttp://appliedresearch.cancer.gov/seermedicare/considerations/training.html

64. Schulman KL, Berenson K, Shih YCT, et al. A checklist for ascertaining study cohorts in oncology health services research using secondary data: report of the ISPOR oncology good outcomes research practices working group. Value in Health. 2013; 16:655–669. [PubMed: 23796301]

65. Nattinger AB, Laud PW, Bajorunaite R, et al. An algorithm for the use of Medicare claims data to identify women with incident breast cancer. Health Serv Res. 2004; 39:1733–1749. [PubMed: 15533184]

66. Smith GL, Shih YC, Giordano SH, et al. A method to predict breast cancer stage using Medicare claims. Epidemiol Perspect Innov. 2010; 7:1. [PubMed: 20145734]

67. Sathiakumar N, Delzell E, Morrisey MA, et al. Mortality following bone metastasis and skeletal-related events among men with prostate cancer: a population-based analysis of US Medicare beneficiaries, 1999-2006. Prostate Cancer Prostatic Dis. 2011; 14:177–183. [PubMed: 21403668]

68. Sathiakumar N, Delzell E, Morrisey MA, et al. Mortality following bone metastasis and skeletal-related events among women with breast cancer: a population-based analysis of U.S. Medicare beneficiaries, 1999-2006. Breast Cancer Res Treat. 2012; 131:231–238. [PubMed: 21842243]

69. Lage MJ, Barber BL, Harrison DJ, et al. The cost of treating skeletal-related events in patients with prostate cancer. Am J Manag Care. 2008; 14:317–322. [PubMed: 18471035]

70. Hassett MJ, Ritzwoller DP, Taback N, et al. Validating billing/encounter codes as indicators of lung, colorectal, breast, and prostate cancer recurrence using 2 large contemporary cohorts. Med Care. 2012 Epub ahead of print.

71. Nordstrom BL, Whyte JL, Stolar M, et al. Identification of metastatic cancer in claims data. Pharmacoepidemiol Drug Saf. 2012; 21(Suppl 2):21–28. [PubMed: 22552976]

72. Basch, E.; Abernethy, A.; Mullins, CD., et al. Recommendations for incorporating patient-reported outcomes into the design of clinical trials in adult oncology. Center for Medical Technology Policy; 2010. http://cmtpnet.org/wp-content/uploads/downloads/2011/12/PRO-EGD-FINAL.pdf

73. Bruner DW, Hanisch LJ, Reeve BB, et al. Stakeholder perspectives on implementing the National Cancer Institute's patient-reported outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). Transl Behav Med. 2011; 1:110–122. [PubMed: 24073038]

74. Abernethy AP, Ahmad A, Zafar SY, et al. Electronic patient-reported data capture as a foundation of rapid learning cancer care. Med Care. 2010; 48(6 Suppl):S32–S38. [PubMed: 20473201]

75. Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. JAMA. 2012; 307:1805–1806. [PubMed: 22550194]

76. Ambs A, Warren JL, Bellizzi KM, et al. Overview of the SEER--Medicare Health Outcomes Survey linked dataset. Health Care Financ Rev. 2008; 29:5–21. [PubMed: 18773611]

77. [Accessed on January 3, 2014] SEER-Medicare Health Outcomes Survey (SEER-MHOS) linked database. http://appliedresearch.cancer.gov/surveys/seer-mhos/

78. CanSORT (Cancer Surveillance & Outcomes Research Team) website. http://www.med.umich.edu/cansort/

79. Jagsi R, Abrahamse P, Morrow M, et al. Patterns and correlates of adjuvant radiotherapy receipt after lumpectomy and after mastectomy for breast cancer. J Clin Oncol. 2010; 28:2396–2403. [PubMed: 20351324]

80. Ayanian JZ, Chrischilles EA, Fletcher RH, et al. Understanding Cancer Treatment and Outcomes: The Cancer Care Outcomes Research and Surveillance Consortium. J Clin Oncol. 2004; 22:2992–2996. [PubMed: 15284250]

81. Chen AB, Cronin A, Weeks JC, et al. Expectations about the effectiveness of radiation therapy among patients with incurable lung cancer. J Clin Oncol. 2013; 31:2730–2735. [PubMed: 23775958]

82. Chen AB, Cronin A, Weeks JC, et al. Palliative radiation therapy practice in patients with metastatic non-small-cell lung cancer: a Cancer Care Outcomes Research and Surveillance Consortium (CanCORS) Study. J Clin Oncol. 2013; 31:558–564. [PubMed: 23295799]

83. National Cancer Institute. Cancer Care Outcomes Research & Surveillance Consortium (CanCORS): Applied Research Program. Dec. 2013 http://appliedresearch.cancer.gov/cancors/cancors_fact_sheet.pdf

84. Bilimoria KY, Stewart AK, Winchester DP, et al. The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. Ann Surg Oncol. 2008; 15:683–690. [PubMed: 18183467]

85. Bilimoria KY, Bentrem DJ, Stewart AK, et al. Comparison of commission on cancer-approved and -nonapproved hospitals in the United States: implications for studies that use the National Cancer Data Base. J Clin Oncol. 2009; 27:4177–4181. [PubMed: 19636004]

86. NCCN Oncology Outcomes Database. http://www.nccn.org/network/business_insights/outcomes_database/outcomes.aspx

87. Palta JR, Efstathiou JA, Bekelman JE, et al. Developing a national radiation oncology registry: From acorns to oaks. Practical Radiation Oncology. 2012; 2:10–17. [PubMed: 24674031]

88. Efstathiou JA, Nassif DS, McNutt TR, et al. Practice-based evidence to evidence-based practice: building the National Radiation Oncology Registry. J Oncol Pract. 2013; 9:e90–e95. [PubMed: 23942508]

89. Vainshtein JM, Hayman JA, Moran JM, et al. Collaborative quality initiative in the treatment of breast and lung cancer: an important step toward high quality cost-effective care. Int J Radiat Oncol Biol Phys. 2013; 87(2 Suppl):S498–S499.

90. Giordano SH, Kuo YF, Duan Z, et al. Limits of observational data in determining outcomes from cancer therapy. Cancer. 2008; 112:2456–2466. [PubMed: 18428196]

91. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika. 1983; 70:41–55.

92. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat Med. 2004; 23:2937–2960. [PubMed: 15351954]

93. Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. Stat Methods Med Res. 2004; 13:17–48. [PubMed: 14746439]

94. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000; 11:550–560. [PubMed: 10955408]

95. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. Annu Rev Public Health. 1998; 19:17–34. [PubMed: 9611610]

96. Staiger D, Stock J. Instrumental variables regression with weak instruments. Econometrica. 1997; 65:557–586.

97. Bekelman JE, Handorf EA, Guzzo T, et al. Radical cystectomy versus bladder-preserving therapy for muscle-invasive urothelial carcinoma: examining confounding and misclassification bias in cancer observational comparative effectiveness research. Value Health. 2013; 16:610–618. [PubMed: 23796296]

98. Punglia RS, Saito AM, Neville BA, et al. Impact of interval from breast conserving surgery to radiotherapy on local recurrence in older women with breast cancer: retrospective cohort analysis. BMJ. 2010; 340:c845. [PubMed: 20197326]

99. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 2010; 19:537–554. [PubMed: 20354968]

100. Beadle BM, Liao KP, Elting LS, et al. Improved survival using intensity-modulated radiation therapy in head and neck cancers: A SEER-Medicare analysis. Cancer. 2014 Epub ahead of print.

101. Zeliadt SB, Potosky AL, Penson DF, et al. Survival benefit associated with adjuvant androgen deprivation therapy combined with radiotherapy for high- and low-risk patients with nonmetastatic prostate cancer. Int J Radiat Oncol Biol Phys. 2006; 66:395–402. [PubMed: 16904843]

102. Hadley J, Polsky D, Mandelblatt JS, et al. An exploratory instrumental variable analysis of the outcomes of localized breast cancer treatments in a medicare population. Health Econ. 2003; 12:171–186. [PubMed: 12605463]

103. Jones MP. Indicator and stratification methods for missing explanatory variables in multiple linear regression. JASA. 1996; 91:222–230.

104. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009; 338:b2393. [PubMed: 19564179]

105. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res. 1999; 8:3–15. [PubMed: 10347857]

106. Barladi AN, Enders CK. An introduction to modern missing data analyses. Journal of School Psychology. 2010; 48:5–37. [PubMed: 20006986]

107. Little, RJA.; Rubin, DB. Statistical analysis with missing data. Wiley; New York (NY): 1987.

108. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. Stat Med. 2000; 19:3219–3236. [PubMed: 11113956]

109. Manning WG. The logged dependent variable, heteroscedasticity, and the retransformation problem. Journal of Health Economics. 1998; 17:283–295. [PubMed: 10180919]

110. Manning WG, Mullahy J. Estimating log models: To transform or not to transform? Journal of Health Economics. 2001; 20:461–494. [PubMed: 11469231]

111. Basu A, Rathouz PJ. Estimating incremental and marginal effects on health outcomes using flexible link and variance function models. Biostatistics. 2005; 6:93–109. [PubMed: 15618530]

112. Basu A, Manning WG, Mullahy J. Comparing alternative models: Log vs. Cox proportional hazard? Health Econ. 2004; 13:749–765. [PubMed: 15322988]

113. Ba er O, Gardiner JC, Bradley CJ, et al. Longitudinal analysis of censored medical cost data. Health Econ. 2006; 15:513–525. [PubMed: 16389670]

114. Lin DY. Proportional means regression for censored medical costs. Biometrics. 2000; 56:775–778. [PubMed: 10985215]

115. Lin DY. Regression analysis of incomplete medical cost data. Stat Med. 2003; 22:1181–1200. [PubMed: 12652561]

116. Carides GW, Heyse JF, Iglewicz B. A regression-based method for estimating mean treatment cost in the presence of right-censoring. Biostatistics. 2000; 1:299–313. [PubMed: 12933511]

117. Huang Y. Cost analysis with censored data. Med Care. 2009; 47(7 Suppl 1):S115–S119. [PubMed: 19536024]

118. Ioannidis JP. Why most published research findings are false. PLoS Med. 2005; 2:e124. [PubMed: 16060722]

119. The STROBE checklist. http://www.equator-network.org

120. Motheral B, Brooks J, Clark MA, et al. A checklist for retrospective database studies--report of the ISPOR Task Force on Retrospective Databases. Value Health. 2003; 6:90–97. [PubMed: 12641858]

121. Garrison LP Jr, Neumann PJ, Erickson P, et al. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. Value Health. 2007; 10:326–335. [PubMed: 17888097]

122. Institute of Medicine. The Learning Health System Series. http://www.iom.edu/Activities/Quality/~/media/85DAF51E84634210B05C1317FFF94D22.pdf

123. Okun, S.; McGraw, D.; Stang, P., et al. Making the case for continuous learning from routinely collected data. Discussion Paper, Institute of Medicine; Washington, DC: 2013. http://www.iom.edu/Global/Perspectives/2013/MakingtheCaseforContinuousLearning.aspx

124. Abernethy AP, Etheredge LM, Ganz PA, et al. Rapid-learning system for cancer care. J Clin Oncol. 2010; 28:4268–4274. [PubMed: 20585094]

**Table 1**

Data Sources[*]

| Data Source | Strengths | Weaknesses |
|---|---|---|
| Surveillance Epidemiology and End Results (SEER) | Large, NCI-funded; population-based; regional cancer registries that cover 26% of US population (reasonably representative, though more urban, racial differences) | Limited information about key factors; inaccuracy, such as underascertainment of outpatient treatments, including RT; migration/loss to follow-up |
| SEER-Medicare | Claims data may better capture treatments received | Generalizability (age), misclassification errors, failure to capture all morbidity |
| Medicare | Claims represent entire population of Medicare beneficiaries rather than subset included in SEER | Generalizability (age), lack of information on pathology and staging, misclassification, failure to capture all morbidity |
| Private Insurer Claims Databases (e.g. Marketscan) | Includes claims data on younger patients than Medicare | Costly, selection of only privately insured patients |
| Survey Datasets | Detailed measures; some have population-based target population | Costly to fund, some selection bias due to non-response |
| National Cancer Database (NCDB) | Large, joint project of ACS/ACoS, captures 70% of all cancers diagnosed in US | Generalizability (CoC accredited centers) |
| Registries from Centers of Excellence (e.g. National Comprehensive Cancer Network) | More detailed information, potentially greater accuracy | Generalizability (patient selection effects), more limited numbers |
| Radiation Oncology-specific registries (e.g. National Radiation Oncology Registry, Michigan Radiation Oncology Quality Consortium) | Very detailed dosimetric and other data | Costly to establish, still in developmemnt |

[*] Table includes only those sources commonly used to study U.S. patients. Additional valuable sources exist in other countries, including robust nationwide, population-based registries.

**Table 2**

Checklist: Guidance and key considerations for study design for an observational CER protocol[*]

| Guidance | Key Considerations | Check |
|---|---|---|
| Provide a rationale for study design choice and describe key design features. | - Cohort study proposals should clearly define cohort entry date (baseline date), employ a new user design (or provide rationale for including prevalent users), and include plans for reporting losses to followup.<br><br>- Case-control study proposals should clearly describe the control sampling method, employ a new user design (or provide a rationale for assessing confounders at index date), and assess potential for recall bias (if applicable).<br><br>- Case-cohort study proposals should include how the sampling scheme will be accounted for during analysis.<br><br>- Case-crossover study proposals should discuss the potential for confounding by time-varying factors and clearly state how the resulting effect estimate can be interpreted.<br><br>- Case-time controlled study proposals should clearly weigh the pros and cons of accounting for calendar trends in the prevalence of exposure. | ☐ |
| Define start of followup (baseline). | - The time point for start of followup should be clearly defined and meaningful, ideally anchored to the time of a medical intervention (e.g., initiation of drug use).<br><br>- If alternative approaches are proposed, the rationale should be provided and implications discussed. | ☐ |
| Define inclusion and exclusion criteria at start of followup. (baseline). | - Exclusion and inclusion criteria should be defined at the start of followup (baseline) and should be based solely on information available at this point in time (i.e., ignoring potentially known events after baseline).<br><br>- The definition should include the time window for assessment (usually the same for all cohort members). | ☐ |
| Define exposure (treatments) of interest at start of followup. | | ☐ |
| Define outcome(s) of interest. | - Information should be provided on measures of accuracy if possible. | ☐ |
| Define potential confounders. | - Potential confounders known to be associated with treatment and outcome should be prespecified when possible.<br><br>- Confounders should be assessed prior to exposure or treatment initiation to ensure they are not affected by the exposure.<br><br>- Approaches to empirical identification of confounders should be described if planned. | ☐ |

[*] Reprinted with permission from: Velentgas P, Dreyer NA, Nourjah P, et al. (eds.) Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. AHRQ Publication No. 12(13)-EHC099. Rockville, MD: Agency for Healthcare Research and Quality; January 2013. P. 31. [Available at http://www.effectivehealthcare.ahrq.gov/ehc/products/440/1166/User-Guide-to-Observational-CER-1-10-13.pdf].