

COMMENTARY

Leveraging Epidemiology and Clinical Studies of Cancer Outcomes: Recommendations and Opportunities for Translational Research

Joanne W. Elena, Lois B. Travis, Naoko I. Simonds, Christine B. Ambrosone, Rachel Ballard-Barbash, Smita Bhatia, James R. Cerhan, Patricia Hartge, Rebecca S. Heist, Lawrence H. Kushi, Timothy L. Lash, Lindsay M. Morton, Kenan Onel, John P. Pierce, Leslie L. Robison, Julia H. Rowland, Deborah Schrag, Thomas A. Sellers, Daniela Seminara, Xiao Ou Shu, Nancy E. Thomas, Cornelia M. Ulrich, Andrew N. Freedman

Manuscript received December 13, 2011; revised September 26, 2012; accepted October 1, 2012.

Correspondence to: Joanne Watters Elena, PhD, MPH, Clinical and Translational Epidemiology Branch, DCCPS 6130 Executive Blvd, rm 5136 Bethesda, Maryland 20892 (email: elenajw@mail.nih.gov).

As the number of cancer survivors continues to grow, research investigating the factors that affect cancer outcomes, such as disease recurrence, risk of second malignant neoplasms, and the late effects of cancer treatments, becomes ever more important. Numerous epidemiologic studies have investigated factors that affect cancer risk, but far fewer have addressed the extent to which demographic, lifestyle, genomic, clinical, and psychosocial factors influence cancer outcomes. To identify research priorities as well as resources and infrastructure needed to advance the field of cancer outcomes and survivorship research, the National Cancer Institute sponsored a workshop titled “Utilizing Data from Cancer Survivor Cohorts: Understanding the Current State of Knowledge and Developing Future Research Priorities” on November 3, 2011, in Washington, DC. This commentary highlights recent findings presented at the workshop, opportunities to leverage existing data, and recommendations for future research, data, and infrastructure needed to address high priority clinical and research questions. Multidisciplinary teams that include epidemiologists, clinicians, biostatisticians, and bioinformaticists will be essential to facilitate future cancer outcome studies focused on improving clinical care of cancer patients, identifying those at high risk of poor outcomes, and implementing effective interventions to ultimately improve the quality and duration of survival.

J Natl Cancer Inst ;2013;105:85–94

The number of cancer survivors in the United States has grown from 3 million in 1971 to an estimated 12 million in 2012 (1), in part because of advances made in earlier diagnosis, supportive care, and more effective treatments. As patients survive longer, disease recurrence and the late effects of cancer treatments become of increasing importance, not only to patients, but also to their families, health-care providers, and cost reimbursement systems. Survivors are at increased risk of second malignant neoplasms, cardiovascular disease, and other chronic conditions, including pulmonary disease, osteoporosis, metabolic syndrome and obesity (2–4). In addition to life-threatening late effects such as second malignant neoplasms, pulmonary compromise, and cardiovascular disease, patients’ functional status and quality of life can be severely impaired by long-term conditions, such as cognitive decline, permanent hearing loss, and tinnitus (5).

Although numerous epidemiologic studies have investigated factors that affect cancer risk, far fewer have addressed the extent to which demographic, lifestyle, genomic, clinical, and psychosocial factors—as well as interactions among these factors—influence cancer outcomes among people diagnosed with cancer (6). The ultimate goal of delineating the influence of these factors is to improve the medical management of cancer patients. One recent review of the survivorship literature reported that quality of life was the research area most commonly addressed, comprising 62%

of studies (7). In recent years, however, there has been an increased focus on molecular, genetic, and predictive factors that affect cancer recurrence and other outcomes (6,8).

To address these questions and identify opportunities for future cancer outcomes research, the National Cancer Institute sponsored a workshop titled “Utilizing Data from Cancer Survivor Cohorts: Understanding the Current State of Knowledge and Developing Future Research Priorities” in Washington, DC, on November 3, 2011. The goal of the meeting was to discuss how to optimize research strategies, leverage available survivor data sources, and determine scientific research priorities. More than 90 scientists with expertise in diverse disciplines participated in the workshop. This report provides an overview of selected recent research findings presented at the workshop, opportunities using existing data, and suggestions for research, data, and infrastructure that are needed to advance the field of cancer outcomes and survivorship research. Although the term “cancer outcomes” encompasses a variety of cancer-related endpoints, including patient-reported outcomes, quality of care, and health service information, we focus our discussion on treatment-related toxicities, recurrence, second malignant neoplasms, mortality, and survival. However, much of the information presented here about identifying data sources and optimizing research strategies is relevant across cancer outcomes.

Research Opportunities: Construction of Scientific Hypotheses and Strategic Leveraging of Existing Data

Dr Lois Travis (University of Rochester) opened the workshop emphasizing the necessity of optimizing cancer outcomes research by describing some of the barriers to this research and her experience in devising and implementing solutions. The first challenge is to develop research questions that are novel, scientifically important, and clinically meaningful and then construct a diverse, multi-disciplinary team of investigators with the appropriate expertise. For example, Dr Travis and colleagues recently assembled a consortium to investigate the long-term effects of platinating agents, radiotherapy, and surgical approaches in testicular cancer survivors. Given the typically young age at diagnosis of this cancer and the approximate 95% cure rate, testicular cancer survivors can now expect to live many decades following diagnosis but may experience late effects secondary to their cancer and its treatment. The multidisciplinary team included expertise in medical oncology, pharmacogenomics, statistical genetics, radiation oncology, biology, psychosocial oncology, cardiology, neurology, nephrology, pathology, epidemiology, metal toxicology, bioinformatics, and biostatistics (5).

The next hurdle is locating and accessing the appropriate data, determining what databases might already be available, what variables are accessible and then planning for the collection of additional information needed to address the scientific hypotheses. When an existing infrastructure for cancer outcomes research does not exist, it becomes necessary first to construct the required database, ensuring not only that the foundation will allow for rigorous investigation of current hypotheses but also that it is sufficiently flexible and broad-based to facilitate investigations of future hypotheses that emerge (5). Throughout all study steps, the following overarching goals should be kept in mind: 1) the development of a high-impact, high-quality research program; 2) the translation of findings into clinical care guidelines to optimize the quality and duration of survival with cost-effective follow-up; and 3) the efficient utilization of research funds.

Pooling studies is one approach to maximizing efficiency by increasing statistical power. Dr Xiao Ou Shu (Vanderbilt University) presented work from the After Breast Cancer Pooling Project, an ongoing pooling analysis of four prospective studies of lifestyle factors and breast cancer prognosis: the Shanghai Breast Cancer Survivors Study, Women's Healthy Eating and Living Study, Life After Cancer Epidemiology Study, and Nurses' Health Study. Together, these studies include more than 18,000 breast cancer survivors, ranging in age from 20 to 83 years at diagnosis and with a variety of tumor subtypes (including 1020 triple-negative tumors) (9). They have been able to investigate several questions, including the effect of physical activity and prediagnosis body mass index on breast cancer survival (10,11).

The considerations and indications for pooling studies have been widely discussed (12). The advantages include capturing a wide range of exposures in a diverse (in this case, multi-ethnic) population, increased sample size, and greater efficiency. These pooling analyses, however, are limited by a number of factors, including heterogeneity across studies in terms of exposure and outcome assessment, eligibility criteria and enrollment procedures,

differences in treatments and completeness of treatment data among locations, potential marked differences in year of diagnosis of cases that may result in very different pathologic characterization and treatment experiences, cohort effects, and patterns of missing data. Although stratification of results can ameliorate the effects of heterogeneity in populations or disease diagnoses, a decrease in statistical power typically results. Advanced modeling techniques (eg, imputation) can also be useful. Harmonizing cohort data is a requisite and time-consuming part of pooling data that requires in-depth knowledge of each study and the within-study heterogeneity; but it may result in using the simplest response option (eg, yes/no variables) across studies, which raises the possibility of exposure misclassification and reduction in the power to detect differences in outcomes related to diverse exposures. Therefore, it is important to have a clear understanding of the extent of heterogeneity in study design and exposure assessment because extensive heterogeneity may necessitate the elimination of many studies from pooled analysis, thus reducing the power that can be achieved from a large pooled sample. Despite difficulties in pooling data, this is an efficient method to maximize power and detect associations otherwise not seen. For example, results were inconclusive about tamoxifen use for breast cancer treatment until a meta-analysis of randomized controlled trials was published in 1998, which substantially impacted clinical practice (13).

Multiple study designs, including cohort studies, clinical trials, and case-control studies, can address questions relevant to cancer outcomes; the advantages and limitations of each design have been discussed extensively (12,14,15). Investigators must decide which design best addresses the posed research question and whether the question involves discovery or confirmation of previous research findings, while considering available resources. No matter the choice, the design should also anticipate developing a resource for ongoing survivorship research. Many of the workshop presenters highlighted research that utilized available data in creative ways to answer high-priority questions. A brief description of the research and data that were discussed at the workshop are presented below, organized by study design.

Clinical Trials

Randomized controlled trials can provide the strongest answers to selected questions, but are typically limited by cost, sample size, the duration of available follow-up, and the select nature of patients who enroll in clinical trials, which often exclude older cancer patients because of preexisting comorbid medical conditions (16). However, data collected in these trials can be used for research questions beyond the original hypothesis by resourceful investigators who give careful thought to methodological concerns. Dr John Pierce (University of California–San Diego) presented an overview of the Women's Healthy Eating and Living Randomized Trial, a dietary intervention trial testing the effects of a diet high in fruits, vegetables, and fiber and lower in fat among more than 3000 early-stage breast cancer survivors aged 18 to 70 years (17). Although no difference in breast cancer recurrence or mortality was observed between the two arms of the trial in an ancillary study (18,19), investigators were able to use archived blood samples to address several research questions, including whether tamoxifen

metabolites and *CYP2D6* polymorphisms were related to breast cancer recurrence (18). In addition, based on data from food frequency questionnaires, they observed no adverse effects of soy foods on breast cancer prognosis (19). Biological samples increase the utility to ask additional questions beyond those anticipated at the time the study was designed. However, it is important to note that ancillary studies using trial data are not randomized for additional research questions and are not necessarily stronger evidence than other observational studies.

Data from the National Cancer Institute–sponsored Clinical Trials Cooperative Group Program, consisting of researchers, cancer centers, and community physicians throughout the United States, Canada, and Europe, have also been used to examine numerous cancer outcomes (20). Dr Smita Bhatia (City of Hope) presented data from the Children’s Oncology Group that reported black and Hispanic children with acute lymphoblastic leukemia experienced worse outcomes than white and Asian children (21). To further understand these findings, they investigated whether nonadherence to oral 6-mercaptopurine varied by race because oral antimetabolite therapy for 2 years during the maintenance phase is critical to ensure durable remissions. Using electronic monitoring (microprocessor chips in the caps of the pill bottles), the investigators found that adherence to prescribed oral 6-mercaptopurine helped explain the ethnic difference in survival seen in pediatric acute lymphoblastic leukemia patients (22). This study utilized an existing funded resource to construct a novel trial and collect both clinical data and biospecimens from patients.

Dr Christine Ambrosone (Roswell Park Cancer Institute) described a similar experience using data from the Southwest Oncology Group SWOG-8897 trial. This breast cancer cooperative trial banked lymph nodes to test hypotheses about genetics, treatment-related toxicity, and disease-free survival among women receiving cyclophosphamide-containing adjuvant therapy for breast cancer (23–26) and DNA repair pathways (27). One disadvantage of utilizing existing studies is that optimal data for answering the new questions may not have been collected. In this case, no blood specimens were available and thus only a limited number of single nucleotide polymorphisms were examined because of the limited amounts of available DNA. Ideally, at a relatively small incremental cost, clinical trials could collect high quality DNA to enable comprehensive assessment of variability across genes in multiple pathways to study survival in relation to pharmacogenetics or biomarkers of toxicities (28). In addition, questionnaires could be added to conduct nontherapeutic, hypothesis-driven studies of etiology and survival. An advantage to using data and samples from cooperative group trials is that the patient populations and treatments received are somewhat homogeneous and toxicities and recurrences are recorded. However, the generalizability of results may be limited by the highly selected patient population in trials (29).

Cohort Studies

Although cancer epidemiology cohorts have most often been used to evaluate the development of cancer in healthy individuals, there are a growing number of longitudinal studies of cancer survivors (<http://epi.grants.cancer.gov/survivor-cohort-resources/>). These include studies originally designed to investigate outcomes in survivors as well as cohorts created to study risk that may be adapted to answer specific questions on outcomes (30). One of the best known cohort

studies of cancer survivors is the Childhood Cancer Survivor Study (CCSS), initiated in 1993. Dr Leslie Robison (St Jude Children’s Research Hospital) presented an overview of the CCSS and the group’s recent findings. With more than 20 years of follow-up data, the CCSS has documented that 73% of survivors have at least one chronic health condition 30 years post-treatment (31), 44% of long-term childhood cancer survivors report markedly diminished health status (32), and the cumulative incidence of subsequent neoplasms 30 years after a childhood cancer diagnosis is 20.5% (33). There have been almost 200 publications in the last 10 years using the CCSS, with the scope of research spanning genetic risk factors, late effects from treatment, comorbidities, second malignant neoplasms, reproductive health, psychosocial outcomes, long-term health, and lifestyle behaviors (34).

Data collected by health maintenance organizations (HMOs) are another potential source of in-depth treatment information. Dr Lawrence Kushi (Kaiser Permanente Northern California) described use of the Pathways study, which studies the influence of lifestyle factors and molecular markers on breast cancer recurrence and survival (35), to investigate the patterns of complementary and alternative medicine use before and after diagnosis (36), the association between physical activity and quality of life during active treatment (37), and the association between tumor size and DNA methylation profiles (38). An advantage of using integrated HMO data, compared with medical records that are not administratively related, is that HMO data from all sources of medical care for an individual are more readily available and complete, including data from both outpatient and inpatient care, pharmacy, radiology, and other sources relevant to tracking cancer screening, diagnosis, and treatment, than data from cohort studies, where patients are likely to have been treated in diverse settings. These advantages within HMO medical record systems occur in part because the HMOs centralize record-keeping systems and many HMOs are at the vanguard of implementation of electronic medical records and databases. For example, Kaiser Permanente Northern California implemented the Beacon oncology module, part of KP HealthConnect, the Epic-based electronic medical record system (Epic Systems, Verona, WI). Beacon allows detailed documentation of patient consult and actions surrounding medical oncology visits, allowing physicians to create and manage treatment plans. Such detail on chemotherapy use was previously available only in clinical trial settings, and availability of these types of data will open up areas of research not previously achievable.

Dr Kushi also gave an overview of the HMO Cancer Research Network, a National Cancer Institute–supported resource of more than 10 million enrollees with data from the mid-1990s on enrollment, demographics, tumors, prescribed drugs, encounters with providers, vital signs, census, procedures, diagnoses, and lab values (<http://crn.cancer.gov/about/>). Tumor blocks are often available as well. A patient cohort designed to address a specific question can be assembled rapidly. For example, Bowles et al. used electronic administrative data from the HMO Cancer Research Network Virtual Data Warehouse to assemble a cohort in less than 1 year that consisted of more than 13,000 women diagnosed with invasive breast cancer from 1999 to 2007 and found that the combination of anthracycline plus trastuzumab, in particular, was associated with elevated heart failure risk (39).

Another alternative to the development of de novo recruitment at multiple sites, such as the CCSS, or coordination and pooling across multiple diverse health care settings, such as the HMO Cancer Research Network, is to use population-based registries. Dr Timothy Lash (Wake Forest University and Aarhus University Hospital) described several of the Danish medical and cancer registries, which record incident cancers in the Danish population and permit linkage to other relevant population-based registries (40). By combining data from a population-based cohort of Danish women diagnosed with stage I to stage III breast carcinoma registered to the Danish Breast Cancer Cooperative Group and prescription data from the National Registry of Medicinal Products, they found that use of simvastatin, a highly lipophilic statin, was associated with a reduced risk of breast cancer recurrence, whereas there was no association observed between hydrophilic statin use and breast cancer recurrence (41). Similarly, they found no association between CYP2D6 inhibition and recurrence in tamoxifen-treated patients using a nested case-control design within an analogous cohort (42). The advantage to using existing data is that the analyses for this work were completed before investigators would have been able to enroll the first participant, had they needed to design a new study. In addition, these projects were conducted for a small fraction of the cost that would be required to initiate a new study requiring de novo data collection. Other investigators have been able to use existing population-based cancer registries in Scandinavian countries and the United States to efficiently describe increased second malignant neoplasm risks among patients with various cancers (43–53). However, the ability to link patient data in one registry to data in another registry or clinical record system is much more feasible in countries that have universal health care, a unique patient identifier, and more centralized medical record systems for evaluating care within their countries. It is much more complex and costly to do such data linkages in countries, such as the United States, that do not have such medical record systems except in the context of integrated health-care delivery systems.

Another invaluable resource is the Surveillance Epidemiology and End Results (SEER)-Medicare data, which was presented by Dr Deborah Schrag (Dana Farber Cancer Institute). SEER-Medicare data (<http://healthservices.cancer.gov/seermedicare>) contains sociodemographic characteristics and health service claims billed to Medicare beneficiaries aged 65 years and older in SEER areas, including more than 100,000 cancer survivors. This dataset is useful for studying a variety of outcomes, including patterns and quality of care, postdiagnostic surveillance, treatment-related complications, and cost of care, particularly those that require procedures or hospitalizations. For example, Baxter et al. used SEER-Medicare data and found an association with increased hip fracture in elderly women who had undergone pelvic irradiation for anal, cervical, or rectal cancers (54). Similar research has been conducted using SEER-Medicare data to study late effects of treatments [eg, the risk of rectal cancer after prostate radiation (55)] and health-care utilization by prostate cancer survivors (56). One benefit of using SEER-Medicare data, as well as HMO data, is that the data reflect information experienced in the broader community, in contrast to a controlled clinical trial where groups such as cancer patients aged older than 65 are generally not included. This database can also be used to evaluate the comparative effectiveness of different cancer

therapies (57). The size of these databases allows examination of differential response by clinical characteristics such as tumor stage, comorbidity, and other factors relevant to outcomes. For example, recent studies from SEER-Medicare confirm that findings from clinical trials in selected patient samples showing modest benefits in colon cancer survival with the addition of bevacizumab to fluorouracil-based chemotherapy (58) and the addition of oxaliplatin to adjuvant 5-fluorouracil (59) are also observed in the more representative patient population in SEER-Medicare who are treated in the general community setting. Electronic medical records are being increasingly deployed, with several vendors dominating the marketplace. Augmenting administrative datasets with electronic medical record data is another important strategy for building epidemiologic research capacity. To the greatest extent feasible, epidemiologists should interface with software builders and designers to ensure that risk factor data are captured consistently across electronic medical records. Furthermore, linkage of administrative claims data to electronic medical records offers ever greater potential for leveraging routinely collected data for research purposes and for assembling large cohorts quickly and efficiently.

Case-Control Studies

Case-control studies are particularly cost efficient when information on an exposure or covariate is expensive to obtain, such as assays of biomarkers or supplemental interviews. One concern with observational studies in general is that the data may not be as systematically collected as in clinical trials, which generally measure fewer variables using a tightly controlled protocol, but often over a shorter time period, in fewer participants, and in selected subgroups of patients. Dr Rebecca Heist (Massachusetts General Hospital) discussed work from her team that evaluated the quality of data for a range of prognostic and outcome variables in a case series derived from a large case-control study of lung cancer, finding that data about overall survival, resection rates, postoperative complications, and follow-up for early-stage lung cancer outcomes were of reasonable quality using standard retrospective methods (60). By contrast, data about other late-stage lung cancer outcomes, such as symptoms, toxicity, response rates, progression-free survival, and disease-free survival, using standard retrospective data collection were found to be of poor quality. However, this information could be improved by implementing a rapid, prospective outcomes ascertainment system protocol, which allows for the quality of the outcomes data to be tested and revised in real time (60).

The number of molecular prognostic and pharmacogenomic studies using data from case-control studies is increasing because many such studies now routinely collect biological samples. Dr Heist described how she evaluated the association between lung cancer survival and polymorphisms of *MDM2*, a negative regulator of p53, using blood samples collected from a case series of patients in a large case-control study of lung cancer risk. Because they collected date of death (or last known date alive) and date of progression (or last known date without progression) they were able to determine that the *MDM2* G/G genotype was associated with poorer survival among early-stage non-small cell lung cancer patients (61). Such studies can potentially be used to individualize cancer therapies to find the least toxic and most effective treatments, predict cancer

susceptibility for subsequent cancers and other late sequelae, and identify new molecular targets for novel therapeutics.

A nested case-control study combines the prospective nature of a cohort investigation with the efficiency of the case-control approach (16). Dr Kenan Onel (University of Chicago) discussed his work within the CCSS to perform a genome-wide association study in pediatric Hodgkin lymphoma survivors. They identified two variants at chromosome 6q21 that were associated with radiation therapy-induced second malignancies in pediatric Hodgkin lymphoma patients (62). Dr Onel also presented work from a genome-wide association study of therapy-related acute myeloid leukemia, which found an association between three single nucleotide polymorphisms in the subset of therapy-related acute myeloid leukemia patients with acquired abnormalities on chromosomes 5 and/or 7, which are associated with prior exposure to alkylating agents (63). By conditioning on the shared exposure of radiation in these studies, these investigators were able to reduce the non-genetic heterogeneity among cases and controls, increasing their power to detect genetic associations with smaller sample sizes.

Converting existing population-based case-control studies designed to address etiologic hypotheses to prognostic cohort studies through follow-up of the cases can offer potential to study cancer outcomes. Such studies can take advantage of large sample sizes, population-based sampling, extensive environmental and lifestyle data, and available banked biospecimens. Dr James Cerhan (Mayo Clinic) presented work from the National Cancer Institute-SEER Survival Study, which used cases from a population-based, case-control study of non-Hodgkin lymphoma and obtained selected clinical data, treatment (eg, chemotherapy, radiation), and survival data (date and cause of death) from the SEER cancer registries. They identified four single nucleotide polymorphisms using germline genotyping data on immune function genes that, in combination with clinical factors, were statistically significantly associated with overall survival in follicular lymphoma patients (64). In addition, non-Hodgkin lymphoma patients who smoked, consumed alcohol, or were obese before diagnosis had poorer overall and lymphoma-specific survival (65). They also found that germline single nucleotide polymorphisms in the *LMO2* gene were able to better predict overall survival than immunohistochemical expression of the LMO2 protein, a demonstrated prognostic marker, in tumor tissue (66). These examples highlight the important contribution that epidemiologic studies can provide in understanding prognosis through combining clinical and tumor factors with host genetic and lifestyle factors. In considering which research questions can best be addressed in such studies, it is important to recognize that SEER cancer registry treatment data includes data only on the initial 4 months of treatment, has incomplete data on radiation therapy, does not include data on chemotherapy, and has limited data on biomarkers that may be used to guide therapy. If data on the specifics of chemotherapy or radiation therapy are needed, additional resources would be required to obtain the data retrospectively. Challenges of utilizing existing, population-based, case-control studies to identify case subjects for cancer patient cohort studies include selection bias due to loss of case subjects with early mortality (who are not recruited into the study); lack of detailed clinical and treatment data and pathology samples, which can be difficult to collect retrospectively; lifestyle factors that are

generally measured only at baseline and not available post-treatment; general inability to obtain serum samples before the onset of treatment; and difficulty in obtaining data on disease progression as a study endpoint. In addition, the cost and feasibility of collecting treatment and clinical data retrospectively must be considered. This may be a specific concern in the case of medical record systems that may have a time limit for retention of clinical data.

Dr Cerhan provided an example of a large prospective cohort study of newly diagnosed non-Hodgkin lymphoma patients that were simultaneously used for a clinic-based, case-control study and a prospective cohort study of non-Hodgkin lymphoma outcomes. The prognostic study was used to show that vitamin D deficiency was associated with poor event-free and overall survival (67), free-light chains are a prognostic biomarker (68), and historical and concomitant use of statins did not negatively impact R-CHOP (rituximab-cyclophosphamide, doxorubicin, vincristine, and prednisone) therapy in diffuse large B-cell patients (69). The investigators also showed the event-free and overall survival experience and the association for the free-light chain biomarker from their observational cohort was strikingly similar to data from a controlled clinical trial, highlighting that well-conducted observational studies can obtain valid results congruent with controlled trials. Nevertheless, there are limitations to observational studies of outcomes, including difficulty systematically assessing treatment responses, treatment toxicities, and other disease markers. Partnering with clinical trials groups, as previously discussed, may be a particularly effective approach to addressing this limitation. Another major limitation of many studies is the lack of assessment of health behaviors and other relevant exposures after a cancer diagnosis, which could result in substantial misclassification because many cancer patients change behaviors following diagnosis (70,71).

Research Opportunities: Creating New Cohorts

New molecular categorization technologies can inform treatment strategies that improve patient care. For example, treatment of metastatic colorectal cancer now includes genetic testing for KRAS mutations, and lung cancer patients with EGFR mutation may benefit most from erlotinib treatment (72,73). Observational studies can contribute in meaningful ways to discovering patient subsets and effective care strategies, especially for newly introduced medications and treatments. Dr Thomas Sellers (H. Lee Moffitt Cancer Center & Research Institute) presented an overview of Total Cancer Care, a prospective patient cohort study of 75,000 newly diagnosed cancer patients that includes patient-reported questionnaire data on risk factors, diet, quality of life, and complementary and alternative medicine use; clinical data from multiple sources (eg, pathology reports, treatment records, and laboratory tests); and biospecimens (eg, blood, snap-frozen tumor, adjacent “normal” tissue) at a cost of approximately \$150 million over the first 5 years. All data are available electronically and managed through an integrated data warehouse, greatly accelerating access to researchers. There is a portal to the warehouse for referring physicians and another portal for patients to permit access to their medical record and treatment plan. Recruitment is ongoing; as of March 2011, 72,188 patients had consented, 23,404 tumors had been collected, and 16,393 tumors have been profiled

for gene expression. By collecting a wide variety of information, this platform facilitates research on the molecular analysis of the tumor, lifestyle factors, host genetics, costs, and treatment decisions, among other topics. Patients will be followed throughout life, allowing for add-on studies, as needed, including recruitment to clinical trials for which they may be uniquely qualified based on the molecular characteristics of their tumor.

Workshop Recommendations

During the workshop, participants divided into smaller working groups to discuss the research issues, gaps, priorities, and resources needed to facilitate cancer outcome studies (Boxes 1–3). These boxes are not intended to express specifically defined priorities for any of the institutions represented at the meeting but rather serve to summarize the collective deliberations of the group.

Data Collection

Determining key relevant variables depends upon the research questions, but certain baseline data should be included in studies designed to evaluate outcomes in cancer patients or survivors. Critical factors in addition to treatment that have been documented to influence outcomes in these groups include age, gender, race/ethnicity, comorbidities, and some lifestyle factors. Collecting biospecimens for molecular and other type of biomarker measures that may predict outcomes is a key need in many areas. Box 1 lists variables that the discussion groups considered essential and additional variables that may be relevant depending on the research question. Moreover, data are only as rigorous as the methods used to collect them and thus depend upon the standardization and rigor of the tools used (reliability, validity) and the quality control procedures that are implemented to capture and code this information. Although we do not discuss the issues involved in collection and measurement here, thorough discussions are provided elsewhere (14,15).

Scientific Priorities

The prioritization of cancer outcomes research areas is challenging because the field encompasses many diverse and critical areas. The Institute of Medicine report, *From Cancer Patient to Cancer Survivor: Lost in Transition*, identified the following priority topics: 1) understudied (neglected) sites, 2) risk and type of late effects by age at exposure, and 3) the aging and cancer interface (74). Rather than identifying specific research questions, the group identified research that will provide clinically actionable information for treatment and screening interventions that inform evidence-based clinical care guidelines through the identification of clinical, genomic, and modifiable behavioral risk factors that influence cancer progression, survival, adverse events, and quality of life. Similar to the Institute of Medicine report, workshop participants also emphasized the need for additional studies in special populations such as long-term survivors, survivors with comorbidities, adolescent and young adult cancer survivors, and patients with rare cancers (Box 2). Incorporating new technologies has the potential to heighten the depth and accuracy of research in terms of patient and tumor characterization and also provides for cost-effective means of patient recruitment and data collection.

Box 1. Workshop recommendations for data collection in new studies of cancer outcomes

Information that should be included whenever possible

- Patient demographics (age, race/ethnicity, sex, education, family history, prior cancer history, reported and measured)
- Basic diagnostic information (cancer site, date of diagnosis, diagnostic method, location)
- Disease characteristics (histology, size, stage, grade, pathologic stage, tumor biomarkers)
- Patient clinical characteristics (functional status, clinical biomarkers)
- Lifestyle (diet, physical activity, body mass index, smoking, alcohol use), repeated measures before and after cancer diagnosis
- Comorbidities (cardiovascular disease, diabetes, hypertension, chronic obstructive pulmonary disease, metabolic syndrome, depression) at diagnosis and new onset after diagnosis.
- Type of treatment received: initial treatment, adjuvant therapies, surgery, radiation (fields and tumor dose), chemotherapy (regimens, drug names, cumulative doses), hormonal therapy, bone marrow/stem cell transplant (document differences in what is ordered vs actually received)
- Treatment-related toxicities (acute, chronic, and late effects)
- Biospecimens (blood samples, tumor specimens, normal tissue, germline DNA) collected at diagnosis (ie, pretherapy) and consider repeated measurements
- Clinical endpoints (treatment response, disease, progression/recurrence, survival [overall and cause-specific], quality of life, second malignant neoplasms)
- Additional information that should be considered, depending on hypotheses
- Additional demographic information (occupation, income, insurance, geographic data)
- Interruptions in treatment or failure to complete treatments
- Additional anthropometrics (body composition, weight change)
- Additional lifestyle factors (supplement use, over-the-counter drugs, mind/body stress, coping methods)
- Drug interactions
- Transition from oncologist to postcare
- Psychosocial functioning and coping
- Complementary/integrated medicine
- Cost of treatment
- Health services utilization
- Phenotypic assays
- Survivorship care plans (impact)
- Cognitive function (including baseline)
- Quality-of-care studies
- Additional patient-reported outcomes: screening behaviors, sun/sunscreen use, performance status, menopausal status, pregnancy/breast-feeding history

Thus, studies that utilize various new technologies to facilitate research in terms of identifying optimal methods for recruitment and data collection were encouraged.

Box 2. Clinical and translational research areas of interest identified during workshop for cancer survivor studies

Design studies that will provide clinically actionable information, which include clinical, molecular, and modifiable behavioral risk factors (eg, diet, physical activity, alcohol use, smoking, weight) that influence:

- Prognosis, recurrence, and survival
- Second primary neoplasms, cardiovascular and pulmonary disease, and other adverse health outcomes
- Quality-of-life and psychosocial issues
- Prognostic studies with biospecimens and health behaviors

Design studies that address issues in special populations such as:

- Long-term survivors
- Survivors with comorbidities
- Adolescent and young adult cancer survivors
- Understudied and/or rare cancers

Design studies that specifically examine:

- Screening interventions that inform evidence-based clinical care guidelines
- The impact of cancer on family members and caregivers
- Factors that influence compliance to cancer therapies
- The cost effectiveness of cancer therapies
- Reproductive potential and outcomes after childhood cancer

Design studies that evaluate and promote technology to:

- Establish consistent data elements for electronic medical records
- Identify innovative approaches for patient recruitment into studies
- Identify and test new technologies for data collection

Ensure that investments in research studies are able to inform clinical care guidelines throughout the cancer trajectory

Resources and Infrastructure Needed

Most of the recommended resources and infrastructure necessary to support research span cancer sites and hypotheses (Box 3). Generally, the recommendations fit into one of four categories: 1) funding opportunities, 2) data harmonization and pooling of data, 3) study coordination and implementation, and 4) enhanced data resources. Many of the recommendations offered here, especially those related to funding, will ultimately be decided by the funding agencies; however, this list serves as a comprehensive “wish list” for the workshop participants. The greatest articulated need was support for establishing new studies with comprehensive data collection (including biospecimens) and linking existing studies to high-quality treatment information and epidemiologic data. Standard collection of blood specimens in the context of cooperative group trials and availability to the broader research community would be highly advantageous. There was also great interest in tools to connect the research community through online resources, in-person meetings, and working groups to share best practices, promote pooling data, and facilitate general collaboration. Several of the suggestions were directed toward specific concerns, including streamlining the institutional review board process, establishing

Box 3. Workshop recommendations: resources and infrastructure needed to support clinical and translational research in cancer outcomes

Funding opportunities that support:

- Infrastructure for adequate follow-up of survivors in new and existing cohort studies
- Relatively inexpensive correlative studies that utilize existing cohort studies or clinical trials
- Data linkage to electronic medical and claims records and disease registries for additional comorbidities and treatment data
- Bioinformatics support including coordinating centers as data warehouses
- Interdisciplinary review of applications with expertise across clinical, epidemiologic, genomic, basic, and behavioral sciences

Harmonization of data across studies

Establish uniform definitions for outcomes (eg, recurrence) and integrate into clinical care

Create an online resource to direct researchers to available cancer outcomes data and resources (eg, PhenX–National Institute of Health supported tools, abstraction forms, questionnaires, biospecimen best practices),

Establish a repository of uniform self-report measurement tools (eg, depression, alcohol use, smoking, physical activity, diet, and other survivorship-relevant questionnaires) to enable data design/collection consistency and promote collaboration across studies

Ongoing workshops and/or working groups for new and existing cohorts with repeat follow-up and biospecimens

Request that cohorts have a list of common searchable terms (eg, MeSH terms)

Facilitate pooling studies and the development of new multi-center endeavors for uncommon cancers and rare exposures

Detailed, annotated samples available for genotyping

Improving study conduct/implementation: recruitment, consent, data collection, and analysis

Revise the consenting process to minimize time constraints.

Include central institutional review board review for multisite projects, incorporation of research consent into clinical care, and expanded use of video or verbal consent. Possible revisions in the National Cancer Institute’s Common Rule present an opportunity for the National Cancer Institute to advocate reformations in consent

Develop online technology for self-reported data collection, patient consent, data transfer, communication between investigators for data use, and data-sharing agreement templates

Establish secure site-specific portal system for patients seeking to access their information

Capitalize on existing infrastructure to better utilize coordinating centers as data warehouses after studies are concluded

Utilize patient advocacy groups for recruitment and follow-up of cancer patients

Box 3. (Continued)

Development of enhanced data resources

Leverage Surveillance Epidemiology and End Results data to construct linkages with the National Death Index, state cancer registries, and create sentinel Surveillance Epidemiology and End Results registries charged with collecting data on cancer recurrence as well as additional treatment, clinical, and other data

Better integration of epidemiologic information into clinical/preclinical trials to facilitate the translational continuum of drug development

Enhanced quality of treatment information in cancer registries Infrastructure (possibly a nationwide cancer registry) for following people regularly and reliably for cancer outcomes

standard definitions for cancer outcomes (eg, recurrence), and making detailed, annotated samples available for genotyping.

Summary

This report not only highlights the optimal uses of existing data and technologies using epidemiological methods to address critical research questions but also provides recommendations to facilitate future cancer outcome studies. Confronted with limited resources, it is essential that the highest priority scientific questions be addressed in the most cost-effective way through the application of carefully constructed designs. Epidemiologists, clinicians, biostatisticians, and bioinformaticists need to work together to apply rigorous methods across study designs with the focused intent to improve clinical care of cancer patients and identify those at high risk of poor outcomes to implement effective interventions to improve prognosis, quality of life, and overall health. As the number of cancer survivors continues to grow and survival improves, understanding what demographic, lifestyle, genomic, clinical, and psychosocial factors affect cancer outcomes and overall health of cancer survivors is critical to improve their quality and duration of survival.

References

1. Parry C, Kent EE, Mariotto AB, Alfano CM, Rowland JH. Cancer survivors: a booming population. *Cancer Epidemiol Biomarkers Prev.* 2011;20(10):1996–2005.
2. Yeh ET, Bickford CL. Cardiovascular complications of cancer therapy: incidence, pathogenesis, diagnosis, and management. *J Am Coll Cardiol.* 2009;53(24):2231–2247.
3. Tichelli A, Socie G. Considerations for adult cancer survivors. *Hematology.* 2005;2005(1):516–522.
4. Mariotto AB, Rowland JH, Ries LAG, Scoppa S, Feuer EJ. Multiple cancer prevalence: a growing challenge in long-term survivorship. *Cancer Epidemiol Biomarkers Prev.* 2007;16(3):566–571.
5. Travis LB, Beard C, Allan JM, et al. Testicular cancer survivorship: research strategies and recommendations. *J Natl Cancer Inst.* 2010;102(15):1114–1130.
6. Ambrosone CB, Rebbeck TR, Morgan GJ, et al. New developments in the epidemiology of cancer prognosis: traditional and molecular predictors of treatment response and survival. *Cancer Epidemiol Biomarkers Prev.* 2006;15(11):2042–2046.
7. Harrop JP, Dean JA, Paskett ED. Cancer survivorship research: a review of the literature and summary of current NCI-designated cancer center projects. *Cancer Epidemiol Biomarkers Prev.* 2011;20(10):2042–2047.
8. Bhatia S. Role of genetic susceptibility in development of treatment-related adverse outcomes in cancer survivors. *Cancer Epidemiol Biomarkers Prev.* 2011;20(10):2048–2067.
9. Nechuta S, Caan B, Chen W, et al. The After Breast Cancer Pooling Project: rationale, methodology, and breast cancer survivor characteristics. *Cancer Causes Control.* 2011;22(9):1319–1331.
10. Beasley J, Kwan M, Chen W, et al. Meeting the physical activity guidelines and survival after breast cancer: findings from the after breast cancer pooling project. *Breast Cancer Res Treat.* 2012;131(2):637–643.
11. Kwan M, Chen W, Kroenke C, et al. Pre-diagnosis body mass index and survival after breast cancer in the After Breast Cancer Pooling Project. *Breast Cancer Res Treat.* 2012;132(2):1–11.
12. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology.* 3rd ed. Philadelphia, PA: Lippincott-Raven; 2008.
13. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet.* 1998;351(9114):1451–1467.
14. Travis L, Bhatia S, Allan J, Oeffinger K, Ng A. *Second Primary Cancers.* Philadelphia, PA: Lippincott Williams & Wilkins; 2011.
15. Oeffinger KC, van Leeuwen FE, Hodgson DC. Methods to assess adverse health-related outcomes in cancer survivors. *Cancer Epidemiol Biomarkers Prev.* 2011;20(10):2022–2034.
16. Adami HO, Hunter DJ, Trichopoulos D. *Textbook of Cancer Epidemiology.* New York: Oxford University Press; 2002.
17. Pierce JP, Natarajan L, Caan BJ, et al. Influence of a diet very high in vegetables, fruit, and fiber and low in fat on prognosis following treatment for breast cancer: the Women's Healthy Eating and Living (WHEL) randomized trial. *JAMA.* 2007;298(3):289–298.
18. Madlensky L, Natarajan L, Tchu S, et al. Tamoxifen metabolite concentrations, CYP2D6 genotype, and breast cancer outcomes. *Clin Pharmacol Ther.* 2011;89(5):718–725.
19. Caan BJ, Natarajan L, Parker B, et al. Soy food consumption and breast cancer prognosis. *Cancer Epidemiol Biomarkers Prev.* 2011;20(5):854–858.
20. Nass SJ, Balogh E, Mendelsohn JM. A National Cancer Clinical Trials Network: recommendations from the Institute of Medicine. *Am J Ther.* 2011;18(5):382–391.
21. Bhatia S, Sather HN, Heerema NA, Trigg ME, Gaynon PS, Robison LL. Racial and ethnic differences in survival of children with acute lymphoblastic leukemia. *Blood.* 2002;100(6):1957–1964.
22. Bhatia S, Landier W, Shangguan M, et al. Nonadherence to oral mercaptopurine and risk of relapse in Hispanic and non-Hispanic white children with acute lymphoblastic leukemia: a report from the Children's Oncology Group. *J Clin Oncol.* 2012;30(17):2094–2101.
23. Ambrosone CB, Barlow WE, Reynolds W, et al. Myeloperoxidase genotypes and enhanced efficacy of chemotherapy for early-stage breast cancer in SWOG-8897. *J Clin Oncol.* 2009;27(30):4973–4979.
24. Choi J-Y, Barlow WE, Albain KS, et al. Nitric oxide synthase variants and disease-free survival among treated and untreated breast cancer patients in a Southwest Oncology Group clinical trial. *Clin Cancer Res.* 2009;15(16):5258–5266.
25. Yao S, Barlow WE, Albain KS, et al. Gene polymorphisms in cyclophosphamide metabolism pathway, treatment-related toxicity, and disease-free survival in SWOG 8897 clinical trial for breast cancer. *Clin Cancer Res.* 2010;16(24):6169–6176.
26. Yao S, Barlow W, Albain K, et al. Manganese superoxide dismutase polymorphism, treatment-related toxicity and disease-free survival in SWOG 8897 clinical trial for breast cancer. *Breast Cancer Res Treat.* 2010;124(2):433–439.
27. Sucheston L, Zhao H, Yao S, et al. Genetic predictors of taxane-induced neurotoxicity in a SWOG phase III intergroup adjuvant breast cancer treatment trial (S0221). *Breast Cancer Res Treat.* 2011;130(3):993–1002.
28. Freedman AN, Sansbury LB, Figg WD, et al. Cancer pharmacogenomics and pharmacoepidemiology: setting a research agenda to accelerate translation. *J Natl Cancer Inst.* 2010;102(22):1698–1705.
29. Ganz PA, Land SR, Antonio C, et al. Cancer survivorship research: the challenge of recruiting adult long term cancer survivors from a cooperative clinical trials group *J Cancer Surviv.* 2009;3(3):137–147.

30. Berrington de González A, Morton LM. Converting epidemiologic studies of cancer etiology to survivorship studies: approaches and challenges. *Cancer Epidemiol Biomarkers Prev.* 2012;21(6):875–880.
31. Oeffinger KC, Mertens AC, Sklar CA, et al. Chronic health conditions in adult survivors of childhood cancer. *N Engl J Med.* 2006;355(15):1572–1582.
32. Hudson MM, Mertens AC, Yasui Y, et al. Health status of adult long-term survivors of childhood cancer. *JAMA.* 2003;290(12):1583–1592.
33. Friedman DL, Whitton J, Leisenring W, et al. Subsequent neoplasms in 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. *J Natl Cancer Inst.* 2010;102(14):1083–1095.
34. Robison LL, Armstrong GT, Boice JD, et al. The Childhood Cancer Survivor Study: a National Cancer Institute–supported resource for outcome and intervention research. *J Clin Oncol.* 2009;27(14):2308–2318.
35. Kwan M, Ambrosone C, Lee M, et al. The Pathways Study: a prospective study of breast cancer survivorship within Kaiser Permanente Northern California. *Cancer Causes Control.* 2008;19(10):1065–1076.
36. Heather Greenlee, Kwan ML, Ergas IJ, et al. Complementary and alternative therapy use before and after breast cancer diagnosis: the Pathways Study. *Breast Cancer Res Treat.* 2009;117(3):653–665.
37. Mandelblatt J, Luta G, Kwan M, et al. Associations of physical activity with quality of life and functional ability in breast cancer patients during active adjuvant treatment: the Pathways Study. *Breast Cancer Res Treat.* 2011;129(2):521–529.
38. Christensen BC, Kelsey KT, Zheng S, et al. Breast cancer DNA methylation profiles are associated with tumor size and alcohol and folate intake. *PLoS Genet.* 2010;6(7):e1001043.
39. Bowles EA, Wellman RD, Delate T, et al. C-B2-01: cardiotoxic chemotherapy is associated with increased heart failure risk among women with breast cancer in the Cancer Research Network. *Clin Med Res.* 2011;9(3–4):148.
40. Gjerstorff ML. The Danish Cancer Registry. *Scand J Public Health.* 2011;39(7 Suppl):42–45.
41. Ahern TP, Pedersen L, Tarp M, et al. Statin prescriptions and breast cancer recurrence risk: a Danish nationwide prospective cohort study. *J Natl Cancer Inst.* 2011;103(19):1461–1468.
42. Lash TL, Cronin-Fenton D, Ahern TP, et al. CYP2D6 inhibition and breast cancer recurrence in a population-based study in Denmark. *J Natl Cancer Inst.* 2011;103(6):489–500.
43. Travis LB, Curtis RE, Glimelius B, et al. Second cancers among long-term survivors of non-Hodgkin's lymphoma. *J Natl Cancer Inst.* 1993;85(23):1932–1937.
44. Dores GM, Cote TR, Travis LB. New malignancies following Hodgkin lymphoma, non-Hodgkin lymphoma, and myeloma. In: Curtis R, Freedman D, Ron E, et al. ed. *New Malignancies Among Cancer Survivors: SEER Cancer Registries 1973–2000*. Bethesda, MD: National Institutes of Health; 2006:397–435.
45. Hodgson DC, Gilbert ES, Dores GM, et al. Long-term solid cancer risk among 5-year survivors of Hodgkin's lymphoma. *J Clin Oncol.* 2007;25(12):1489–1497.
46. Metayer C, Lynch CF, Clarke EA, et al. Second cancers among long-term survivors of Hodgkin's disease diagnosed in childhood and adolescence. *J Clin Oncol.* 2000;18(12):2435–443.
47. Dores GM, Metayer C, Curtis RE, et al. Second malignant neoplasms among long-term survivors of Hodgkin's disease: a population-based evaluation over 25 years. *J Clin Oncol.* 2002;20(16):3484–3494.
48. Hisada M, Biggar RJ, Greene MH, Fraumeni JF, Jr., Travis LB. Solid tumors after chronic lymphocytic leukemia. *Blood.* 2001;98(6):1979–1981.
49. Travis LB, Curtis RE, Boice JD Jr., Platz CE, Hankey BF, Fraumeni JF Jr. Second malignant neoplasms among long-term survivors of ovarian cancer. *Cancer Res.* 1996;56(7):1564–1570.
50. Travis LB, Fossa SD, Schonfeld SJ, et al. Second cancers among 40,576 testicular cancer patients: focus on long-term survivors. *J Natl Cancer Inst.* 2005;97(18):1354–1365.
51. Travis LB, Curtis RE, Storm H, et al. Risk of second malignant neoplasms among long-term survivors of testicular cancer. *J Natl Cancer Inst.* 1997;89(19):1429–1439.
52. Chaturvedi AK, Kleinerman RA, Hildesheim A, et al. Second cancers after squamous cell carcinoma and adenocarcinoma of the cervix. *J Clin Oncol.* 2009;27(6):967–973.
53. Brown LM, Chen BE, Pfeiffer RM, et al. Risk of second non-hematological malignancies among 376,825 breast cancer survivors. *Breast Cancer Res Treat.* 2007;106(3):439–451.
54. Baxter NN, Habermann EB, Tepper JE, Durham SB, Virnig BA. Risk of pelvic fractures in older women following pelvic irradiation. *JAMA.* 2005;294(20):2587–2593.
55. Baxter NN, Tepper JE, Durham SB, Rothenberger DA, Virnig BA. Increased risk of rectal cancer after prostate radiation: a population-based study. *Gastroenterology.* 2005;128(4):819–824.
56. Snyder CF, Frick KD, Herbert RJ, et al. Preventive care in prostate cancer patients: following diagnosis and for five-year survivors. *J Cancer Surviv.* 2011;5(3):283–291.
57. Friends of Cancer Research. Improving Medical Decisions Through Comparative Effectiveness Research: Cancer as a Case Study. http://focr.org/files/CER_REPORT_FINAL.pdf. Accessed October 16, 2012
58. Meyerhardt JA, Li L, Sanoff HK, Carpenter W, Schrag D. Effectiveness of bevacizumab with first-line combination chemotherapy for Medicare patients with stage IV colorectal cancer. *J Clin Oncol.* 2012;30(6):608–615.
59. Sanoff HK, Carpenter WR, Martin CF, et al. Comparative effectiveness of oxaliplatin vs non-oxaliplatin-containing adjuvant chemotherapy for stage III colon cancer. *J Natl Cancer Inst.* 2012;104(3):211–227.
60. Bradbury PA, Heist RS, Kulke MH, et al. A rapid outcomes ascertainment system improves the quality of prognostic and pharmacogenetic outcomes from observational studies. *Cancer Epidemiol Biomarkers Prev.* 2008;17(1):204–211.
61. Heist RS, Zhou W, Chirieac LR, et al. MDM2 polymorphism, survival, and histology in early-stage non-small-cell lung cancer. *J Clin Oncol.* 2007;25(16):2243–2247.
62. Best T, Li D, Skol AD, et al. Variants at 6q21 implicate PRDM1 in the etiology of therapy-induced second malignancies after Hodgkin's lymphoma. *Nat Med.* 2011;17(8):941–943.
63. Knight JA, Skol AD, Shinde A, et al. Genome-wide association study to identify novel loci associated with therapy-related myeloid leukemia susceptibility. *Blood.* 2009;113(22):5575–5582.
64. Cerhan JR, Wang S, Maurer MJ, et al. Prognostic significance of host immune gene polymorphisms in follicular lymphoma survival. *Blood.* 2007;109(12):5439–5446.
65. Geyer SM, Morton LM, Habermann TM, et al. Smoking, alcohol use, obesity, and overall survival from non-Hodgkin lymphoma. *Cancer.* 2010;116(12):2993–3000.
66. Cerhan JR, Natkunam Y, Morton LM, et al. LIM domain only 2 protein expression, LMO2 germline genetic variation, and overall survival in diffuse large B-cell lymphoma in the pre-rituximab era. *Leuk Lymphoma.* 2012;53(6):1105–1112.
67. Drake MT, Maurer MJ, Link BK, et al. Vitamin D insufficiency and prognosis in non-Hodgkin's lymphoma. *J Clin Oncol.* 2010;28(27):4191–4198.
68. Maurer MJ, Micallef INM, Cerhan JR, et al. Elevated serum free light chains are associated with event-free and overall survival in two independent cohorts of patients with diffuse large B-cell lymphoma. *J Clin Oncol.* 2011;29(12):1620–1626.
69. Nowakowski GS, Maurer MJ, Habermann TM, et al. Statin use and prognosis in patients with diffuse large B-cell lymphoma and follicular lymphoma in the rituximab era. *J Clin Oncol.* 2010;28(3):412–417.
70. Holmes RS, Zheng Y, Baron JA, et al. Use of folic acid-containing supplements after a diagnosis of colorectal cancer in the Colon Cancer Family Registry. *Cancer Epidemiol Biomarkers Prev.* 2010;19(8):2023–2034.
71. Velicer CM, Ulrich CM. Vitamin and mineral supplement use among US adults after cancer diagnosis: a systematic review. *J Clin Oncol.* 2008;26(4):665–673.
72. Dahabreh I, Terasawa T, Castaldi P, Trikalinos T. Systematic review: Anti-epidermal growth factor receptor treatment effect modification by KRAS mutations in advanced colorectal cancer. *Ann Intern Med.* 2011;154(1):37–49.
73. Yachida S, Jones S, Bozic I, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature.* 2010;467(7319):1114–1117.
74. Hewitt M, Greenfield S, Stovall E. *From Cancer Patient to Cancer Survivor: Lost in Transition*. Washington, DC: National Academies Press; 2005.

Funding

The National Cancer Institute sponsored the 2011 workshop.

Notes

The funder had no role in the study. The authors report no conflicts of interest.

Affiliations of authors: Division of Cancer Control and Population Science (JWE, NIS, RB-B, DAS, ANF) and Division of Cancer Epidemiology and Genetics (PH, LMM), National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD; Rubin Center for Cancer Survivorship and Department of Radiation Oncology, James P. Wilmot Cancer Center, University of Rochester Medical Center, Rochester, NY (LBT); Department of Cancer Prevention & Control, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY (CBA); Department of Population Sciences, City of Hope, Duarte, CA (SB); Department of Health Sciences Research, Mayo Clinic, Rochester, MN (JRC); Division of Hematology/Oncology, Massachusetts General Hospital, Boston, MA (RSH);

Division of Research, Kaiser-Permanente of Northern California, Oakland, CA (LHK); Department of Clinical Epidemiology, Aarhus University Hospital, Denmark (TLL); Department of Epidemiology & Prevention, Wake Forest School of Medicine, Winston-Salem, NC (TLL); Section of Hematology/Oncology, Department of Pediatrics, University of Chicago, Chicago, IL (KO); Moores UCSD Cancer Center, University of California, San Diego, La Jolla, CA (JPP); Department of Epidemiology and Cancer Control, St Jude Children's Research Hospital, Memphis, TN (LLR); Department of Adult Oncology, Center for Outcomes and Policy Research, Dana-Farber Cancer Institute, Boston, MA (DES); H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL (TAS); Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center and Vanderbilt-Ingram Cancer Center, Nashville, TN (XOS); Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC (NET); Fred Hutchinson Cancer Research Center, University of Washington, Seattle, WA (CMU); German Cancer Research Center and National Center for Tumor Diseases, Heidelberg, Germany (CMU).