Published in final edited form as:

Cancer Epidemiol. 2016 December ; 45: 26–31. doi:10.1016/j.canep.2016.09.003.

Active follow-up versus passive linkage with cancer registries for case ascertainment in a cohort

PF Pinsky¹, K Yu¹, A Black², WY Huang², and PC Prorok¹

¹Division of Cancer Prevention, National Cancer Institute

²Division of Cancer Epidemiology and Genetics, National Cancer Institute

Abstract

Background—Ascertaining incident cancers is a critical component of cancer-focused epidemiologic cohorts and of cancer prevention trials. Potential methods for cancer case ascertainment include active follow-up and passive linkage with state cancer registries. Here we compare the two approaches in a large cancer screening trial.

Methods—The Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial enrolled 154,955 subjects at ten U.S. centers and followed them for all-cancer incidence. Cancers were ascertained by an active follow-up process involving annual questionnaires, retrieval of records and medical record abstracting to ascertain and confirm cancers. For a subset of centers, linkage with state cancer registries was also performed. We assessed the agreement of the two methods in ascertaining incident cancers from 1993–2009 in 80,083 subjects from six PLCO centers where cancers were ascertained both by active follow-up and through linkages with 14 state registries.

Results—The ratio (times 100) of confirmed cases ascertained by registry linkage compared to active follow-up was 96.4 (95% CI: 95.1–98.2). Of cancers ascertained by either method, 86.6% and 83.5% were identified by active follow-up and by registry linkage, respectively. Of cancers missed by active follow-up, 30% were after subjects were lost to follow-up and 16% were reported but could not be confirmed. Of cancers missed by the registries, 27% were not sent to the state registry of the subject's current address at the time of linkage.

Conclusion—Linkage with state registries identified a similar number of cancers as active follow-up and can be a cost-effective method to ascertain incident cancers in a large cohort.

The authors report no conflicts of interest.

Authorship Contribution

Paul Pinsky - Project conception, data analysis, writing of the manuscript, collection of data, editing of the manuscript, final approval

Kelly Yu - Project conception, data analysis, collection of data, editing of the manuscript, final approval

Amanda Black - Data analysis, collection of data, editing of the manuscript, final approval

Wen-Yi Huang - Data analysis, collection of data, editing of the manuscript, final approval

Philip Prorok - Data analysis, collection of data, editing of the manuscript, final approval

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

cancer registries; linkage; active-follow-up; case ascertainment

Introduction

In cohort studies designed to assess etiologic risk factors for cancer, as well as in cancer prevention or screening trials, determining and verifying the endpoints of cancer incidence is a critical component. In many studies, this process of identifying incident cancers is quite complex and resource intensive. For example, in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial, which also functions as a cancer etiology cohort study, the process involves the following: 1) sending out regular questionnaires to all subjects asking about recent cancer diagnoses, 2) determining the location of relevant medical records (e.g., physician or hospital name) and obtaining subject consent for the study to receive such records, 3) obtaining the relevant medical records from providers and 4) performing medical record abstraction to verify cancer diagnosis, site, histology and other characteristics ^{1,2}.

A potentially less resource intensive approach to ascertaining incident cancers in a U.S.-based cohort involves linkage with state cancer registries. Essentially all state cancer registries accept applications for linkages with defined cohorts. The process is complicated logistically, however, by the fact that there is no centralized application mechanism across states or even regions; specific applications for linkage must be sent to each state registry for which linkage is desired. For a nationwide cohort, linkage would thus involve around fifty applications; however, many cohorts are more geographically concentrated and would only need applications for a subset of registries, making the approach more logistically feasible. In addition to the logistics issues of a registry linkage approach, a key question concerns the relative completeness of case ascertainment of this approach as compared to the more standard active approach, for example, as described above for PLCO.

In this paper we compare the findings of the two approaches, active follow-up versus passive linkage with state cancer registries, in PLCO. After many years using the above active approach, PLCO recently transitioned to a hybrid approach that combined active follow-up with passive linkages to state registries. For a time period of about 15 years, incident cancers were comprehensively assessed with both approaches in a subset of PLCO subjects. In this manuscript we assess the relative yield of identified cancers with the two approaches, and compute various metrics of agreement between them. We also analyze the reasons for discrepant results between the approaches.

Design of PLCO Trial and Cohort

The design of PLCO has been described previously ^{1,2}. Enrollment of men and women aged 55–74 was performed from 1993–2001 at ten screening centers nationwide. Exclusion criteria included history of a PLCO cancer and current cancer treatment. At study entry, participants completed a baseline questionnaire that inquired about demographics, medical history, smoking history, and past screenings.

Participants randomized to the intervention arm were offered screening for lung, colorectal, prostate (men only) and ovarian (women only) cancers for up to six years. Control arm participants received usual care. Intervention arm subjects with positive screens were advised to seek diagnostic evaluation; such evaluation was decided by subjects and their primary physicians, not by trial protocol. PLCO screening center staff obtained medical records related to diagnostic follow-up of positive screens and medical record abstractors recorded information on relevant diagnostic procedures and cancer diagnoses.

In addition to screen detected trial-related cancers ascertained as described above, the trial attempted to ascertain all diagnosed cancers by means of a self-reported annual study update (ASU) questionnaire that asked about the type and date of any cancers diagnosed in the prior year or since the last time the questionnaire was completed. Participants not returning the questionnaire were contacted by repeat mailing or telephone. Study center personnel sought and abstracted medical records to attempt to confirm self-reported cancers.

The trial ascertained deaths and causes of death through various means, including the annual questionnaire and National Death Index (NDI) searches, with death certificates obtained where possible; cancers first identified through death certificates or NDI also went through the confirmation process. The ICDO-2 was utilized for cancer topography and histology coding. Clinical and pathologic stage was determined using the TNM staging system and categorized according to the fifth edition of the American Joint Committee on Cancer's Cancer Staging Manual. Follow-up of subjects continued until December 31, 2009, 13 years from randomization, or a subject's known death, whichever came first.

Linkage with State Registries

For efficiency purposes, only registries with substantial numbers of expected linkages were included in the linkage effort. Of these, 13 were "home-state" registries, i.e., registries of a state in which a PLCO screening center was located (one center had a satellite center in another state and for one center, Georgetown University, DC, MD and VA were all considered home-state registries). Six other state registries were chosen on the basis of the most frequent current addresses of PLCO subjects outside of the home states of the centers as of 2012. Note, only subjects' addresses (primary address and alternate address) as of 2012 or their last known date alive were available. Each state required its own registry linkage application and approval process. For the actual linkage, PLCO subjects' PII, including name, date of birth, gender and SSN were sent to the registries. Some registries used a deterministic matching algorithm and others a probabilistic one; for the latter, the default cutoff set by the registry was generally used to determine matches.

The registry linkage effort was a two-phase process. The first phase, carried out in 2012–2013, involved 14 state registries, including the home registries of six screening centers, and covered the period from 1993 through the end of 2009, with complete coverage for this period. The second phase is ongoing and involves five additional registries; it will cover the period through 2014. This report focuses on the first phase, and thus examines active follow-up versus registry linkage for subjects enrolled at the six screening centers whose home registries were included in phase one.

In 2011, PLCO transitioned from active follow-up at the ten screening centers to a three-tiered model where subjects re-consented to either centralized direct contact follow-up, passive follow-up (through linkages) only, or refused extended follow-up. The first two groups were eligible for this linkage effort with the state registries. Therefore, this analysis includes all subjects in these two follow-up groups at the above six screening centers.

Data Analysis

An active follow-up identified (AF-identified) cancer was defined as a cancer identified and confirmed by medical records through the trial's active follow-up cancer ascertainment process, excluding any information obtained from cancer registries; a registry-identified (RG-identified) cancer was defined as a cancer identified through linkages to state registries. AF- and RG-identified cancers were generally categorized by cancer site according to the SEER site recode classifications, with some related categories aggregated (see Appendix for details)³. Jointly-identified cancers were those that were both AF- and RG- identified and matched on the cancer site (they weren't required to match on diagnosis date but the concordance of dates was analyzed). In addition, some cancers identified by each source in a different but related cancer site were also considered jointly-identified in some instances (see Appendix for details).

Only invasive cancers diagnosed during the PLCO active follow-up period (through Dec, 31, 2009 or 13 years of follow-up, whichever came first) were included in our analysis; non-melanoma skin cancers were excluded. We included the first primary cancer diagnosis for each subject, but also subsequent primaries if they were at distinct anatomic sites; for multiple primaries, each counted as an event (note PLCO did not systematically collect data on same-site second primaries, due to the difficulty in distinguishing these from recurrences of the original tumor). As stated above, for PLCO, ICDO-2 was utilized. For registries, ICDO-3 codes were supplied for all cancers; a subset also had ICDO-2 codes supplied. For matching purposes, we used the ICDO-2 codes from the registries if available, otherwise, ICDO-3 codes were used.

We computed the ratio of the number of RG-identified versus AF-identified cancers. Additionally, of all cancers identified by either method, the proportion identified by each method (AF, RG) was computed; this gives an upper bound on the completeness of case ascertainment of each method since some cancers could have been missed by both methods. 95% confidence intervals on these ratios and proportions were computed using bootstrapping. The above analyses were stratified by demographic factors, screening center, time period and cancer site. For jointly-identified cancers, we assessed agreement on date of diagnosis and histology. To assess the relative yield of home state registry versus non-home state registry linkages, we calculated the percentage of attempted linkages that resulted in a match for each. For non-home state linkages, a subject's PII sent to multiple state registries was counted multiple times in the denominator.

As a potential reason for RG-identified cancers that were not AF-identified, we examined how many of these cases occurred subsequent to loss to follow-up in the active process. Loss to follow-up was defined as the time following the last ASU cancer self-report form received from the subject.

Results

The six PLCO centers with completed linkage data from their home state registries randomized 88,072 subjects. Of these, 80,083 (90.9%) had their PII sent to at least one registry based on having either active or passive follow-up status. Compared to subjects at these six centers not sent to any registry, registry-included subjects were more likely to be male, college educated, in the intervention arm and less likely to be non-Hispanic white (Table 1). Compared to subjects at the other four centers, registry-included subjects were more likely to be college educated and less likely to be non-Hispanic white.

Table 2 shows registry linkage submissions and cancer matches for each screening center. Essentially all subjects were submitted to their home-state registry (ies). In general, subjects were submitted to non-home state registries if either their last address was in that state or if the state was neighboring to the home state; however, sometimes logistical considerations precluded an otherwise appropriate non-home state registry linkage. The vast majority of matches were from centers' home state registries (96.2%–98.7%). The yield (matches over attempted linkages) from home state registries ranged from 13.6 to 18.0%, while the yield from non-home state registries was under 1% except for the center in Hawaii (2.6%).

A total of 16,838 cancers were identified by either source, with 14,583 (86.6%) AF-identified, 14,058 (83.5%) RG-identified and 11803 (70.1%) jointly identified. The ratio (times 100) of RG-to AF- identified cancers ($RAT_{RG/AF}$) was 96.4 (95% CI: 95.5–97.4) (Table 3). $RAT_{RG/AF}$ varied little by subject age or calendar year of diagnosis. Across centers, $RAT_{RG/AF}$ ranged from 79.2 to 108.6.

$RAT_{RG/AF}$ ranged from 88.9 to 107.9 for major cancers, with the exception of leukemia, which had a $RAT_{RG/AF}$ value of 115 (Table 4). The jointly-identified cancers had generally close agreement on date of diagnosis as determined from the registry versus active follow-up. About two thirds (64.7%) matched exactly on date, 77.7% were within one week, 90.6% were within one month, 95.4% were within 6 months and 98.8% were within one year.

Overall, 78% of jointly-identified cancers had an exact match on histology (4 digit morphology code) (Table 4). For most cancer types, the match percentage ranged from 60 to 90%; exceptions included lymphoma, kidney and thyroid cancer under 60% and prostate cancer and multiple myeloma over 90%. In general, there was little difference in match rate according to whether the registry provided an ICDO-2 histology code or only provided an ICDO-3 code, with match rates of 79.5% and 77.6%, respectively (recall PLCO utilized ICDO-2 for all cancers).

For the RG-identified cancers that were not AF-identified, Table 5 displays potential reasons for non-ascertainment by active follow-up. The slight majority (54%) were diagnosed before any loss to follow-up (i.e., before the last ASU cancer self-report form) but had no report of the cancer of interest; 16% were diagnosed during that period but did have a report of the cancer (without confirmation). Additionally, 23% had diagnosis during the loss to follow-up period (i.e., after the last ASU) and no report of the cancer of interest, while 7% had diagnosis during loss to follow-up but did have an unconfirmed report of the cancer (some from death certificates).

To examine why AF-identified cancers were not identified by the registries, we analyzed subjects' current state of residence and whether the corresponding registries were queried. Of AF-identified cancers that were not RG-identified (N=2780), 27% were not queried to the registry of the subject's current state of residence (or state of residence when last known alive if deceased), primarily because that state was not among the 14 included in the current linkage round. This compares to only 4% not queried to the state registry corresponding to their current residence among RG-identified cancers.

Discussion

In this analysis comparing active follow-up for all-cancer incidence to a passive process utilizing state cancer registry linkages, we found that the yield of confirmed cancers was similar by both methods, with the ratio (times 100) of cancers found by the registry linkage to those found by the active process being 96.4. Each method missed a substantial fraction of cancers, with an upper bound on complete case ascertainment of 86.6% (active follow-up) and 83.5% (registries).

In terms of efficiency, the linkage process has large up-front costs associated with the application process for each state registry, especially when many states are involved. However, once that is done, the costs are similar for a 1,000 versus a 100,000 person cohort. Additionally, the costs are the similar for one particular cancer type versus all-cancer incidence. In contrast, with the active follow-up approach, the costs are roughly proportional to cohort size; further, costs increase along with the number of cancer types that are of interest. Therefore, for a modest-sized, say, breast cancer prevention trial where the only concern is breast cancer incidence, an active process might be more cost-effective. On the other hand, for a large epidemiologic cohort tracking all-cancer incidence, the linkage approach probably would be more cost-effective.

A related issue is the time lag in obtaining linked cancer data from the registries. In general across the states, registry data are available with a lag of about 18–24 months. In other words, complete data for (say) 2013 would be first available towards the end of 2015. In contrast, for active follow-up as carried out in PLCO, with study update forms sent out annually and another six months or so for obtaining and abstracting medical records, the corresponding lag was about 18 months.

It is recognized that having to apply separately to each state registry is a substantial barrier to being able to perform linkage studies efficiently. In response to this, pilot efforts, funded by the National Cancer Institute and with the cooperation of the North American Association of Central Cancer Registries (NAACCR), are underway to try to streamline this process, with an ultimate goal of a single nation-wide application, similar to the NDI process⁴.

In our analysis we linked with several state registries beyond those where the trial subjects originally resided, i.e., beyond the home-state registries. However, even though the follow-up period was rather long, median of 11 years, we found relatively few case linkages from these non-home state registries, implying that inter-state residential mobility in this middle-

aged and older cohort may have been limited. Thus, for similar aged cohorts, linking only with the registries of the states in which the enrolling centers are located may be sufficient, with perhaps a few neighboring state registries included as well. However, it should be noted that only a limited number of non home-state registries were included here; including more might have increased modestly the number of resulting linkages.

There were several reasons why cancers found by linkage were missed by the active process. These included cancers diagnosed after subjects were lost to follow-up and with no reported cancer (23%) and subjects with reported but not confirmed cancer (23%); still, over half (54%) had not even a report of the cancer even though it was diagnosed during subjects' period of follow-up. As PLCO committed substantial resources to the active follow-up process, with numerous QC checks and repeated contact efforts, this may represent near the limit of what can be accomplished by an effective active follow-up process. Note also that since the follow-up period is generally known for each subject, incidence rates can be computed based on the number of events and the total person time during subjects' follow-up periods. Therefore, the missed cancer rate for active follow-up would be somewhat lower with respect to computed incidence rates than with respect to case ascertainment during the entire study period.

Of cancers missed by the registry, some were due to not linking with the subject's state of residence, though this was only a minority (27%). Beyond that, cancer registries, especially non-SEER registries, do not capture all cancers due to various factors. For example, state registries generally cannot disclose cancers that are diagnosed at Veterans Administration hospitals.

A number of epidemiologic studies have utilized state registry linkages for cancer case ascertainment, including the AARP Diet and Health Study, a study of cancer in Gulf War Veterans, and a study of cancer among organ transplant patients⁵⁻⁷. To our knowledge, only one other U.S. study has compared cancer case ascertainment using state registry linkage with that of self-report (and medical record confirmation) in the same subjects. A pilot study conducted within the AARP Diet and Health Study mailed questionnaires to 12,000 subjects asking about self-reported cancers, with 8326 (69.4%) responding to the questionnaire⁸. The study authors estimated that the registry ascertained 89% (239 of 268) of the self-reported and confirmed cancers and that 86% (239 of 278) of registry ascertained cases (among the 8326 respondents) were self-reported. For comparison, in our analysis, 81% of active follow-up confirmed cancers were identified by the registry, 84% of registry-ascertained cases were confirmed with active follow-up and 87% of registry cases were either confirmed or reported by active follow-up.

In addition to case ascertainment, other characteristics of cancers, for example, stage and treatment, may be desirable. It is beyond the scope of this manuscript to assess the completeness of such data in the various state registries. Researchers needing detailed stage or treatment data should check with the registry(ies) about the completeness of these data during the time periods of interest.

Conclusion

Using state cancer registry linkages can be an effective and efficient method to ascertain incident cancers in a large cohort.

Acknowledgments

This research was supported by contract number HHSN261201100008C from the National Cancer Institute. The funding agency had no role in the research, other than the fact that the authors were employees of the funding agency.

Cancer incidence data have been provided by the following registries: Arizona, California, Colorado, D.C., Hawaii, Idaho, Maryland, Michigan, Nevada, Ohio, Pennsylvania, Texas, Utah, Virginia. All are supported in part by funds from the Centers for Disease Control and Prevention, National Program for Central Registries, local states or by the National Cancer Institute, Surveillance, Epidemiology, and End Results Program. The results reported here and the conclusions derived are the sole responsibility of the authors.

References

1. Prorok P, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000; 21:273S–309S. [PubMed: 11189684]
2. Schoen RE, Pinsky PF, Weissfeld JL, et al. Colorectal cancer incidence and mortality with screening flexible sigmoidoscopy. *New Engl J Med*. 2012; 366:2345–2357. [PubMed: 22612596]
3. Surveillance, Epidemiology and End Results, National Cancer Institute. [Accessed Nov, 2015] SEER – Site-recode ICD-03 to WHO. 2008. http://seer.cancer.gov/siterecode/icdo3_dwho/home/index.html
4. Deapen, D. Advancing cancer research through a virtual pooled registry. NAACR Annual Meeting; <http://www.naacr.org/AC2015/Presentations/Advancing%20Cancer%20Research%20Through.pdf>
5. Schatzkin A, Subar AF, Thompson FE, et al. design and serendipity in establishing a large cohort with wide dietary intake distributions. *Am J Epidemiol*. 2001; 154:1119–1125.
6. Young HA, Maillard JD, Levine PH, et al. Investigating the risk of cancer in 1990–1991 US Gulf War veterans with the use of state cancer registry data. *Ann Epidemiol*. 2010; 20:265–272. [PubMed: 20116284]
7. Engels EA, Pfeiffer RM, Fraumeni JF, et al. Spectrum of cancer risk among US solid organ transplant recipients. *JAMA*. 2011; 306:1891–1901. [PubMed: 22045767]
8. Michaud DS, Midthune D, Hermansen S, et al. Comparison of cancer registry case ascertainment with SEER estimates and self-reporting in a subset of the NIH-AARP Diet and Health Study. *J Registry Management*. 2005; 32:70–75.

Appendix SEER Site Coding

We utilized the SEER ICD-03 to WHO 2008 site re-coding scheme to classify AF- and RG-identified cancers, with a few exceptions³. The following sets of cancers were classified together as a single entity for our analysis: all oral cavity cancers, including pharynx, nose, nasal cavity and larynx, colon and rectal cancer, all lymphomas, and all types of leukemia.

In addition, in some instances, if an AF-identified and a RG-identified cancer in the same subject were in different but related categories, and their dates matched to within a given pre-specified limit, the cancer was considered to be jointly-identified. Specifically, these pairs of related cancer classes included colorectal cancer versus small intestine cancer, stomach cancer versus esophageal cancer, lymphoma versus leukemia, and other digestive cancer versus a specific digestive cancer (i.e., colorectal, pancreatic, liver, stomach or esophageal). For these pairs, if the dates were within 30 days it was considered jointly-

identified. Additionally, if one source specified an unknown primary site and the other source specified a known primary site, and the dates matched exactly, the cancer was considered jointly-identified. Note if both sources specified an unknown primary site (C809), the cancer was also considered Jointly-identified (regardless of the date match). Brain and central nervous system benign and borderline tumors (e.g, meningiomas) were excluded. Note these adjustments had little overall effect on the percentage of cancers that were jointly identified; without these adjustments the jointly identified percentage decreased from 70.1% to 68.9%.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

Two methods of cancer case ascertainment were evaluated in a large cohort

Active follow-up of subjects was compared to linkage with state cancer registries

The case ascertainment rate was similar in the two methods

Each method missed about 15% of all ascertained cases

Table 1

PLCO Subjects

	All (N= 154898)	Six centers - Sent to Registry (N= 80083)	Six centers – not sent to registry ^I (N=7989)	Other four centers (not sent) (N=66826)
	N(%)	N(%)	N(%)	N(%)
Age 65+ (at enrollment)	36.0	36.2	37.1	35.5
Male	49.5%	49.0%	38.4	51.4
Enrolled 1993–1996	44.7	45.6	40.4	44.2
College Education	35.1	39.2%	30.2%	30.8
Non-Hispanic White	85.6%	80.8%	91.4	90.6
Intervention Arm	50.0%	50.8%	41.8	50.0

^I Subjects who refused extended follow-up

Table 2

Registry Linkages by Screening Center

Center (location)	Subjects included in current analysis	# sent to home state registry (ies)	Matches from home state registry(ies) ¹	# sent to non-home state registries ²	Matches from non-home registries ¹	% of all matches that were from home state registry (ies)
	N	N	N(N per 100 sent)	N	N(N per 100 sent)	
Univ of Colorado (Denver, CO)	11,368	11,357	2043 (18.0)	12545	55 (0.4)	97.4
Georgetown Univ (Washington, DC)	8034	8027	1094 (13.6)	4658	43 (0.9)	96.2
Pacific Health (Honolulu, HI)	10657	10640	1871 (17.6)	2746	72 (2.6)	96.3
Henry Ford (Detroit, MI)	23980	23974	3835 (16.0)	23468	81 (0.3)	97.9
Univ of Pittsburgh (Pittsburgh, PA)	15018	14719	2491 (16.9)	26686	31 (0.1)	98.7
Univ of Utah (Salt Lake City, UT & Boise, ID)	11026	11022	1903 (17.3)	23327	47 (0.2)	97.6

¹Matches are per cancer type, not per subject. A subject could have more than one match if they had diagnoses of more than one cancer type.

²Subjects sent to multiple non-home state registries are counted multiple times (one for each registry sent to). Non-home state registries were AZ, CA, NV, OH, and TX and any other registry not the home-state registry for the given center. Note subjects at Georgetown and University of Utah had multiple home state registries (DC/MD/VA & UT/ID, respectively), but number sent reflects number sent to at least one home state registry (i.e., these are not counted multiple times for multiple home-state registries).

Table 3

Cancers identified by active follow-up and by registries – overall and by center, year and age

	Total Cancers	Ratio (times 100) of RG-identified to AF-identified cancers (95% CI)	AF-identified as % of all cancers	RG-identified as % of all cancers
	N(%)			
All invasive cancers	16838	96.4 (95.5–97.4)	86.6	83.5
Center				
Univ of Colorado	2490	93.8 (92.0–96.0)	90.9	85.3
Georgetown Univ	1841	79.3 (76.3–82.3)	86.5	68.6
Pacific Health	2172	107.7 (105.3–109.9)	86.3	93.0
Henry Ford	4781	108.6 (106.4–111.3)	79.4	86.3
Univ of Pittsburgh	3263	84.0 (82.3–86.2)	92.8	78.0
Univ of Utah	2291	97.8 (95.6–100.2)	88.6	86.6
Year 1993–2000	5363	97.7 (96.1–99.1)	88.6	86.5
Year 2001–2009	11475	95.8 (94.6–97.0)	85.7	82.1
Men	10372	93.7 (92.5–94.8)	87.9	82.3
Women	6466	101.0 (99.3–102.6)	84.6	85.4
Age < 75	13067	95.5 (94.5–96.6)	87.7	83.7
Age 75+	3771	99.8 (97.4–102.0)	82.8	82.7

AF-identified: ascertained by active follow-up; RG-identified: ascertained by registry linkage.

Table 4

Cancers identified by active follow-up and by registries by cancer type

Cancer Type	Total cancers identified	Ratio (times 100) of RG-identified to AF-identified	AF-identified as % of total cancers	RG-identified as % of total cancers	% Exact Histology Match ¹
Bladder	624	89.5 (85.1–93.8)	90.4	80.9	70
Brain	175	100.0 (91.0–110.1)	84.2	84.2	89
Breast	2078	99.6 (97.3–102.0)	88.8	88.5	83
Colorectal	1373	102.8 (99.5–106.3)	84.4	86.7	76
Endometrium	437	99.5 (93.8–105.2)	86.3	85.8	70
Esophagous	173	100.7 (92.5–109.0)	87.9	88.4	86
Leukemia	648	114.6 (106.2–124.2)	68.7	78.7	63
Lymphoma	749	95.1 (90.6–99.8)	85.2	81.0	52
Kidney	480	97.6 (91.9–103.0)	85.8	83.8	53
Liver	186	107.9 (98.8–119.0)	81.2	86.7	86
Lung	2056	93.3 (91.1–95.4)	92.1	85.9	68
Melanoma	628	104.1 (97.3–111.6)	74.5	77.5	70
Multiple Myeloma	242	89.7 (82.0–97.5)	88.0	78.9	94
Ovary	252	92.4 (84.9–100.0)	88.5	81.7	65
Oral cavity	377	104.3 (96.5–112.3)	79.8	83.3	82
Pancreas	491	95.1 (90.2–100.2)	88.0	83.7	75
Prostate	4526	88.9 (87.5–90.3)	92.8	82.5	95
Stomach	230	107.4 (98.5–117.1)	82.2	88.3	67
Thyroid	177	104.4 (92.3–118.8)	76.8	80.2	46

AF-identified: ascertained by active follow-up; RG-identified: ascertained by registry linkage.

¹Exact match on four digit ICD-O morphology code for jointly-identified cancers.

Note: All cancer types with at least 100 cases are included in the table.

Table 5

Reasons for non-matches of registry-only identified cancers

PLCO Status	N (%)
All	2255
Diagnosis date after last Annual Study Update (ASU) form, not reported	524 (23)
Diagnosis date after last Annual Study Update (ASU) form, reported but not confirmed	162 (7)
Diagnosis within ASU coverage period, reported but not confirmed	351 (16)
Diagnosis within ASU coverage period, not reported	1218 (54)

Note: Reported implies cancer of the same type as ascertained from the registry