Editorial

Cancer Surveillance Research

William F. Anderson

Biostatistics Branch (BB) DHHS/NIH/NCI/Division of Cancer Epidemiology and Genetics (DCEG) EPS, Executive Blvd Bethesda, Maryland

The journal *Cancer Epidemiology, Biomarkers & Prevention* (*CEBP*) has launched a new manuscript section entitled Cancer Surveillance Research (CSR). The CSR section will consist of original reports using cancer case and population data to examine, test, and develop hypotheses for cancer prevalence, incidence, and mortality. The scope of CSR includes descriptive epidemiology, public health statistics for time trends in cancer burden, genetic, behavioral, and environmental risk factors, cancer disparities and geographic variations, screening and diagnostic practice patterns, and methodological developments for assessing cancer data.

CSR studies are the "eyes and ears" for the monitoring and assessment of cancer burden through the examination of vital health statistics. Evaluation of demographic, temporal, and geographic variations in cancer rates can suggest clues to genetic or environmental exposures, cultural influences, health behaviors, geographic variations, and racial/ethnic variations for subgroups at unexpected risk for certain cancers. Cancer rates may be used to verify the consistency of existing cancer-related hypotheses and/or to generate new ideas for future analytic research. CSR can estimate the external validity of a randomized clinical trial. A well-designed and controlled clinical trial has strong internal validity. However, the generalizability (or external validity) of the randomized study cannot be assumed because subject participation depends upon selection and inclusion criteria, which might not reflect the population at large. Generalizability concerns are further heightened by the fact that clinical trial participants are generally healthier, wealthier, younger, Caucasian, and urban dwellers (1, 2). The merging of population-based and clinical trial evidence can help to determine if an efficacious clinical trial is effective in the general population (3, 4). Additionally, when a disease is rare, cancer surveillance data might be the most reliable source of information. Finally, as the population ages and the cost of cancer-related services increase (5), CSR can aid health care planners and policy makers manage and direct limited resources.

CSR has been enhanced through advancements in computer hardware and software, bioinformatics, and statistical methodologies, local, national and international databases (6). However, cancer surveillance data are underutilized, largely owing to two mistaken impressions (7). First, there is a lack of recognition of the available resources for CSR. Second, descriptive epidemiology-the core methodology for CSR-is often viewed as simplistic, uncertain, and/or unreliable.

Contrary to some mistaken views, population-based resources, databases, and statistical tools are readily available from public-use websites such as the National Cancer Institute's Surveillance, Epidemiology, and End Results program (SEER) (8), Cancer Mortality Maps & Graphs (9), Centers for Disease Control and Prevention (CDC) (10), North American Association of Central Cancer Registries (NAACCR) (11), and the International Agency for Research on Cancer (IARC) (12). SEER provides cancer incidence and survival data from 17 Tumor Registries, covering approximately 26% of the United States. The current SEER database has nearly 5 million cancer cases spanning more than 1 billion person-years from 1973 through 2005. The Cancer Mortality Maps & Graphs website has interactive charts, text, tables, and figures for more than 40 cancers from 150 through 1994. The CDC's National Program of Cancer Registries (NPCR) supports the maintenance of high quality tumor registries for states in the United States. The CDC's National Center for Health Statistics (NCHS) distributes cancer mortality information for the calendar period 1969 to 2005. The NAACCR promotes CSR as an umbrella organization for central cancer registries in the United States and Canada, government agencies, and professional organizations. IARC's CANCER *Mondial* website provides information on the occurrence of cancer world-wide through five programs: (1) *Cancer Incidence in Five Continents* volumes I to IX, (2) ACCIS, incidence and survival data of children and adolescents in Europe, (3) mortality data from the World Health Organization (WHO), (4) GLOBOCAN 2002 for the incidence, prevalence, and mortality from 27 cancers for all countries in the world in 2002, and (5) NORDCAN project from 41 major cancers in Nordic countries.

The simplistic view of descriptive epidemiology partly reflects the reality that descriptive studies are secondary and/or retrospective analyses, dependent upon the observational method. Observational results are cross-sectional, capturing a ''snap-shot'' in time and are subject to uncontrollable chance, bias, or confounding. Population-based descriptive studies begin with a rate matrix (sometimes referred to as a Lexis diagram) (13), as illustrated in Table 1 for female breast cancer from the SEER database (1974-2005). Indeed, it would be imprudent for CSR to

Table 1. Female breast cancer cases (in situ and invasive) in SEER's 9 + 13 + 17 Registries Databases, diagnosed during the years 1974 through 2005

Time (Years)	1974-1977	1978-1981	1982-1985	1986-1989	1990-1993	1994-1997	1998-2001	2002-2005	Birth Cohort
Age (Years)	Rate	Rate	Rate	Rate	Rate	Rate	Rate	Rate	
21-24	1	1	2	2	2	2	2	2	1981
25-28	6	6	7	8	7	7	8	8	1977
29-32	19	17	19	22	22	21	22	20	1973
33-36	39	39	42	48	47	48	49	44	1969
37-40	72	72	83	94	91	92	96	86	1965
41-44	116	117	137	166	160	158	165	152	1961
45-48	167	164	193	238	246	241	245	226	1957
49-52	202	194	223	277	290	306	308	277	1953
53-56	222	217	245	296	312	335	363	323	1949
57-60	260	254	292	345	356	383	421	403	1945
61-64	287	286	328	396	399	419	464	451	1941
65-68	328	317	371	446	459	472	509	497	1937
69-72	344	341	386	474	486	510	538	513	1933
73-76	357	353	410	487	510	534	575	536	1929
77-80	374	348	405	493	499	533	572	545	1925
81-84	352	348	381	464	482	498	547	519	1921
									1917
									1913
									1909
									1905
									1901
									1897
									1893

ignore the complex interactions associated with the Lexis figure.

For example, the two-dimensional geometry of the Lexis diagram shows that three fundamental descriptive variables (age, period, and cohort) are in a single plane and are linearly dependent (Table 1), i.e., age at diagnosis in rows, year of diagnosis (calendar-period) in columns, and year of birth (birth-cohort) in the diagonals. Given the relationship $C = P - A$ (birth-cohort = calendar-period − age at diagnosis), Table 1 has 23 birth-cohorts (1893, 1897, . . . 1981, referred to by mid-year of birth) that are derived from 16 four-year age groups (21-24, 25-28, . . . 81-84 years) and eight four-year time periods (1974-1977, 1978-1981, . . . 2002-2005). Birth-cohort reflects time trends that impact all age groups for a given generation, i. e., risk factor exposures. Calendar-period effects reflect secular trends that affect all age groups at a certain point in time, i.e., changing screening or diagnostic practice patterns. Age is a surrogate for age-related biological factors. Because age, period, and cohort are collinear, it is not possible to completely separate calendar-period effects from age effects or the birth-cohort effects from calendar-period effects, giving rise to the so-called "nonidentifiability" issue.

Given the uncertainty associated with these nonidentifiability issues, CSR requires a close interface between data resources and statistical techniques (14). Age standardization attempts to minimize the impact of different age distributions when comparing rates over time and across populations. Nonlinear regression models have been applied to a sophisticated analyses of time trends (15). Careful attention to plotting techniques facilitates temporal comparisons, fairly conveying the data without overemphasizing the results (16). Multivariate analyses allow for the simultaneous study of two or more dependent variables. Poisson regression can assess cancer-specific hazard rates that are adjusted for any number of covariates such as age at diagnosis, year of diagnosis, stage, grade, etc. (17). Age-period-cohort (APC) models estimate a number of identifiable parameters adjusted for age, period, and cohort effects (18).

For example, two very useful APC parameters for descriptive studies are the "drifts" (linear trends) (refs. 19, 20) and the "fitted" age at onset curve (21). Net drift equals the sum of the linear trends in period and cohort effects for all age groups (Fig. 1A). The net drift quantifies the average annual percentage change in the logarithm of the rates adjusted for period and cohort deviations. It is a summary measure of the overall
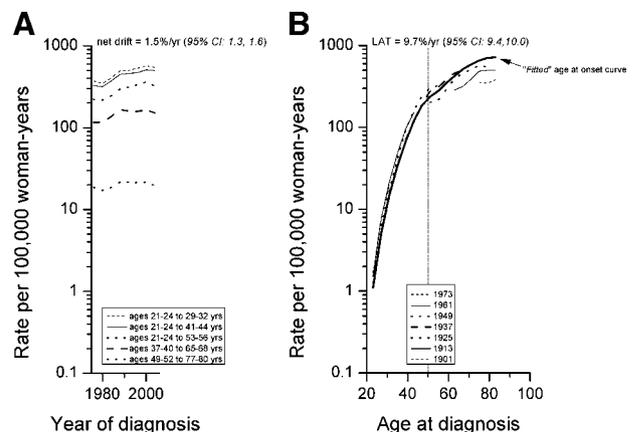

Figure 1. A. Age-specific incidence rate trends with the net drifts for female breast cancer. **B.** Age-period-cohort (APC) "fitted" age-specific curves with the longitudinal age trend for female breast cancer.

trend during the study period, and is closely related to the estimated annual percentage in the age standardized rates. Figure 1A shows a net drift of 1.5% per year (95% CI, 1.3-1.6). In other words, female breast cancer incidence rates rose at a rate of 1.5% for each yearly increment in the calendar time.

Another type of drift is the longitudinal age trend or LAT (Fig. 1B). The LAT quantifies the average annual percentage change in the logarithm of the rates adjusted for age and period deviations, and estimates the average percentage change of the "fitted" age at onset curve (Fig. 1B). The fitted curve is an extrapolation of the age-specific rates for the mid birth-cohort based upon the age-specific rates for all other cohorts in the study (16). Figure 1B shows a LAT of 9.7% per year of attained age (95% CI, 9.4-10.0). In other words, female breast cancer age-specific rates rose at a rate of nearly 10% for each yearly increment in the age at diagnosis. Note the collinearity (nonidentifiability) for the age, period, and cohort effects in Table 1 and Fig. 1. The 1973 birth-cohort in Fig. 1B corresponds to the age-specific temporal trend for ages 21-24 to 29-32 years in Fig. 1A, whereas the 1961 birth-cohort in Fig. 1B corresponds to the age-specific temporal trend for ages 21-24 to 41-44 in Fig. 1A, etc.

In launching the new CSR section, the journal hopes to publicize and highlight both the resources and methodology for cancer surveillance; and through CSR, to assess emerging cancer trends and cancer-related hypotheses. CSR manuscripts should be 3,000 words or less (not counting abstract, references, or legends), have a total of six or fewer tables and/or figures, a structured abstract of 250 words or less (with background, methods, results, and conclusion), and no more than 40 references. Supplemental data can be provided if needed.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C. Threats to applicability of randomised trials: exclusions and selective participation. J Health Serv Res Policy 1999;4:112–21. PubMed.
2. Hutchins LF, Unger JM, Crowley JJ, Coltman CA, Jr., Albain KS. Underrepresentation of patients 65 years of age or older in cancer_ treatment trials. N Engl J Med 1999;341:2061–7. PubMed doi:10.1056/NEJM199912303412706.
3. Cochrane AL. Effectiveness and Efficiency: Random Reflections on Health Services. London: Nuffield Provincial Hospitals Trust; 1972.
4. Berry DA, Ravdin PM. Breast cancer trends: a marriage between clinical trial evidence and epidemiology. J Natl Cancer Inst 2007;99:1139–41. PubMed doi:10.1093/jnci/djm080.
5. Meropol NJ, Schulman KA. Cost of cancer care: issues and implications. J Clin Oncol 2007;25:180–6. PubMed doi:10.1200/JCO.2006.09.6081.
6. Fraumeni JF, Jr., Rimer BK. Cancer surveillance series: inauguration. J Natl Cancer Inst 1999;91:1004. doi:10.1093/jnci/91.12.1004.
7. Glaser SL, Clarke CA, Gomez SL, O'Malley CD, Purdie DM, West DW. Cancer surveillance research: a vital subdiscipline of cancer epidemiology. Cancer Causes Control 2005;16:1009–19. PubMed doi:10.1007/s10552-005-4501-2.
8. Surveillance, Epidemiology, and End Results program (SEER) [homepage on the Internet]. Bethesda (MD): National Cancer Institute. Available from: http://seer.cancer.gov/.
9. Cancer Mortality Maps & Graphs [homepage on the Internet]. Bethesda (MD): National Cancer Institute. Available from: http://www3.cancer.gov/atlasplus/.
10. Centers for Disease Control and Prevention (CDC) [homepage on the Internet]. Atlanta (GA): Centers for Disease Control and Prevention (CDC). Available from: http://www.cdc.gov/.
11. North American Association of Central Cancer Registries (NAACCR) [homepage on the Internet]. Atlanta (GA): Centers for Disease Control and Prevention (CDC). Available from: http://www.naaccr.org/.
12. International Agency for Research on Cancer (IARC) [homepage on the Internet]. Lyons, France: International Agency for Research on Cancer (IARC). Available from: http://www.iarc.fr/.
13. Vandeschrick C. The Lexis diagram, a misnomer. Demogr Res 2001;4:97–124. doi:10.4054/DemRes.2001.4.3.
14. Hankey BF, Ries LA, Edwards BK. The surveillance, epidemiology, and end results program: a national resource. Cancer Epidemiol Biomarkers Prev 1999;8:1117–21. PubMed.
15. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19:335–51. PubMed doi:10.1002/(SICI)1097-0258(20000215)19:33.0.CO;2_Z.
16. Devesa SS, Donaldson J, Fears T. Graphical presentation of trends in rates. Am J Epidemiol 1995;141:300–4. PubMed.
17. Breslow NE, Lubin JH, Marek P, Langholz B. Multiplicative models and cohort analysis. J Am Stat Assoc 1983;78:1–12. doi:10.2307/2287093.
18. Holford TR. Age_period_cohort analysis. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics 1. Chichester, England: John Wiley & Sons; 1998. p. 82–99.
19. Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: Age_period and age_cohort models. Stat Med 1987;6:449–67. PubMed doi:10.1002/sim.4780060405.
20. Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: Age_period_cohort models. Stat Med 1987;6:469–81. PubMed doi:10.1002/sim.4780060406.
21. Anderson WF, Rosenberg PS, Menashe I, Mitani A, Pfeiffer RM. Age_related crossover in breast cancer incidence rates between Black and White Ethnic Groups. J Natl Cancer Inst 2008;100:1804–14. PubMed doi:10.1093/jnci/djn411.

Correction

Correction: Editorial on Cancer Surveillance Research

In the article (1) on Cancer Surveillance Research in the June 2009 issue of *Cancer Epidemiology, Biomarkers & Prevention*, there are errors in the first full paragraph on page 1671 and in the Fig. 1A legend. Reference 16 in that paragraph should be Reference 21. The last full sentence should read as follows: The 1973 birth-cohort in Fig. 1B corresponds to ages 21-24 to 29-32 in Table 1, whereas the 1961 birth-cohort corresponds to ages 21-24 to 41-44, etc. The correct Fig. 1A legend should read as follows from top to bottom: ages 77-80 yr, ages 65-68 yr, ages 53-56 yr, ages 41-44 yr, and ages 29-32 yr.

Reference

1. Anderson WF. Cancer Surveillance Research. Cancer Epidemiol Biomarkers Prev 2009;18:1669–71.

Updated version	Access the most recent version of this article at: http://cebp.aacrjournals.org/content/18/6/1669

Cited articles	This article cites 14 articles, 2 of which you can access for free at: http://cebp.aacrjournals.org/content/18/6/1669.full#ref-list-1
Citing articles	This article has been cited by 3 HighWire-hosted articles. Access the articles at: http://cebp.aacrjournals.org/content/18/6/1669.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://cebp.aacrjournals.org/content/18/6/1669. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.