The Expanding Potential for Cohort Studies to Inform Priorities for Cancer Prevention

Graham A Colditz* and Su-Hsin Chang

Department of Surgery, Division of Public Health Sciences, Washington University School of Medicine, USA

Abstract

While the contributions of cohorts to informing cancer prevention have grown, areas that have been omitted from recent discussions are identified and implications for training discussed. In particular, greater familiarity with and collaboration in modeling of disease can inform prevention building from cohort data to synthesis of behavior, cohort, and population evidence.

The evaluation and appreciation of the contributions of cohorts and their role in cancer prevention continues to be debated as the rate of growth in resources for cancer research slows and future return on investment becomes an explicit evaluation criterion through assessment of the impact of proposed studies on public health or clinical programs.

The discovery, development, delivery, paradigm of research by National Institutes of Health (NIH) has been used to summarize the impact of large initiatives such as funding of cohort studies [1]. In previous writing, the approach to evaluation of large cohorts has focused on these three summary measures:

- **Discovery**: to explain the etiology of disease and health conditions
- **Development**: to provide a basis for developing control measures and prevention procedures for groups and populations at risk (e.g., determination of causes, public health guidelines, risk models)
- **Delivery**: implementation, use of findings, evidence-based public health policy and clinical guidelines (e.g., public awareness, policy applications).

Examples from the Nurses’ Health Study and Childhood Cancer Survivors Study show that these metrics may summarize the contributions show that measures of output grow over time, but the summary of return on investment is poorly presented [1-4]. We might do well to expand the classification further by better understanding how data from cohorts are used and how these data form an essential basis for the growing demand of modeling disease at population levels [5-8].

A lead series of cohorts has been run by the American Cancer Society (ACS) which has led U.S. prospective studies documenting the link between smoking cigarettes and lung cancer from the first study of over 188,000 men [9,10], to the Cancer Prevention Study 1, follow-up of 1 million men and women [11] and also documenting the benefits of stopping smoking where after more than one year the risk was lower than current smokers and took more than 10 years to return to the risk of never smokers [12]. Subsequent follow-up data informed the estimates of tobacco smoking to cancer mortality in the USA providing essential input to the report by Doll and Peto on the potential to prevent cancer [13]. Further updates of the ACS cohorts refined our understanding of the burden of tobacco across decades [14]. The Cancer Prevention Study cohorts have also contributed leadership to documenting the burden due to overweight and obesity [15] setting the stage for the International Agency for Research on Cancer report on this topic and global estimates [16]. Like other cohorts studying lifestyle and diet data [17], the ACS also contributed major data on mortality due to alcohol [18].

Beyond the reporting of these associations, investigators at ACS have collaborated with Centers for Disease Control and Prevention (CDC) to provide necessary inputs to their programs for estimating the population burden of tobacco summarized as morbidity, mortality and economic costs [19].

CDC applications build from ACS prospective data (tobacco) to generate estimates of cancer burden for the nation and for states. Thus the ACS cohorts, like other cohorts, have contributed substantially to the discovery of public health and biologically important relations between smoking and cancer [20,21]. Likewise, the early report by Thun on the relation of aspirin to reduced colon cancer mortality [22] opened a field of study that has included subsequent randomized trials [23,24] and further observational data [25,26] to document the substantial reduction in risk with regular long term use of aspirin.

Refining Assessment of Public Health Impact

In the report by Colditz and Winn [1] applying these measures of impact of large initiatives to the Nurses’ Health Study, a final consideration of the overall scientific context was the development and expansion of collaborations. The early collaborations among cohorts included the Oxford hormone combined analysis (OCs [27] and PMH [28]) in the early 1990s, the diet and cancer collaborative individual patient data analysis (beginning in 1991) [17], then the Breast and Prostate Cancer Cohort Consortium linking cohorts to evaluate pathways (IGF and hormonal pathways for breast and prostate cancer) [29,30] and further expanding to apply new technologies and include these cohorts in genome wide association studies for breast [31] and prostate cancer [32]. Many more collaborations have evolved across cohorts as the resource of diet, lifestyle, body mass index, and cancer
outcomes have matured over time. Leadership of these collaborative efforts varies and now many are led by investigators outside of the National Cancer Institute. Some address factors like weight change, BMI, and mortality [33].

In some cohorts, collaboration has taken the form of the link to local or national laboratory investigators with expertise in particular assays, or pathologists with expertise in a particular malignancy (breast, colon). More recently, expanding links to laboratory scientist developing assays [34] or tissue markers [35], exemplify the broadening role of cohorts to address public health issues in cancer from etiology to estimation of population burden, and evaluation of the impact of dietary/lifestyle guidelines [36]. These approaches for epidemiologic data are summarized in a recent discussion of transforming epidemiology for the 21st Century [37].

However, beyond these broader applications by epidemiologists, cohort data can inform policy through incorporation into cost of illness studies [38,39], cost-effectiveness models that require natural history models for disease and outcomes [40-44], and disease models to evaluate screening and other prevention interventions [8,45]. Thus to fully quantify the contribution of prospective data from cohort studies, we will need improved tracking of the use of such epidemiologic data as well as broader training of epidemiologists to engage in these collaborative studies.

The return on investment in prospective data grows when these data are used for development and delivery, not just discovery and generation of research papers explaining etiology of cancer. The expanding fields of disease modeling and population burden of exposures from tobacco to obesity and physical activity, benefits of screening and mathematical models of disease development and progression after diagnosis all require grounding in issues of exposure assessment, as well as disease biology, and modeling techniques. For example, to evaluate the benefits of obesity treatment, a statistical model appropriately controlling for intention-to-treat (or any potential confounding) can be constructed to analyze the data and to estimate the effect of the treatment [46]. Alternatively, agent-based modeling or micro simulations can be used to compare different treatment regimens when the data do not support comparisons. The data, however, can be analyzed to inform the parameters in the agent-based models.

To more usefully draw on the cohort data, stronger collaborations between those generating prospective data and those incorporating such data into population modeling will be needed. Thus, expanded transdisciplinary training programs and funding for collaborative projects must be a high priority. Identifying new ways to share cohort data and bring evolving quantitative modeling methods to evaluate population burden of diseases or societal gains from interventions should be given top priority. These modeling techniques quantify both public health burden and benefits in terms of life years, relative risks, or economic costs. For example, regression models are used to measure association of key variables and the outcome variable keeping all other variables fixed [47]. The measured association is either a marginal effect or an average difference ceteris paribus [47]. Another example is agent-based models. These models can be used to simulate the life history of individuals. [48]. Under reasonable assumptions, any disease impact or policy/treatment intervention can be inputted to the model with the individual loss or gain as the output. The societal loss or gain can be derived by aggregating the individual level of loss or gain.

Conclusion

The growing sophistication of modeling and the integration of biologic as well as population level data will call for new skills among epidemiologists and disease modelers to maximize our use of existing data resources and better inform communities and policy makers how changes at multiple levels from policy to individual behaviors and therapies can reduce the cancer burden across all sectors of society, now and in the future.

References

1. Colditz GA, Winn DM (2008) Criteria for the evaluation of large cohort studies: an application to the nurses’ health study. J Natl Cancer Inst 100: 918-925.
2. Colditz GA (2007) Cohort studies of etiology and survival after cancer: the unique needs for uninterrupted funding. Cancer Causes Control 18: 235-241.
3. Colditz GA (2010) Ensuring long-term sustainability of existing cohorts remains the highest priority to inform cancer prevention and control. Cancer Causes Control 21: 649-656.
4. Colditz GA, Taylor PR (2010) Prevention trials: their place in how we understand the value of prevention strategies. Annu Rev Public Health 31: 105-120.
5. Habbema JD, Schechter CB, Cronin KA, Clarke LD, Feuer EJ (2006) Modeling cancer natural history, epidemiology, and control: reflections on the CISNET breast group experience. J Natl Cancer Inst Monogr : 122-126.
6. Zaubier AG, Landsorp-Vogelar I, Knudsen AB, Wilschut J, van Ballegooijen M, et al. (2009) Evaluating Test Strategies for Colorectal Cancer Screening-Age to Begin, Age to Stop, and Timing of Screening Intervals: A Decision Analysis of Colorectal Cancer Screening for the U.S. Preventive Services Task Force from the Cancer Intervention and Surveillance Modeling Network (CISNET). Rockville (MD).
7. Chang Y, Schechter CB, van Ravesteyn NT, Near AM, Heijnsdijk EA, et al. (2012) Collaborative modeling of the impact of obesity on race-specific breast cancer incidence and mortality. Breast Cancer Res Treat 136: 823-835.
8. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM (2000) Cost-effectiveness of screening for colorectal cancer in the general population. JAMA 284: 1954-1961.
9. HAMMOND EC, HORN D (1954) The relationship between human smoking habits and death rates: a follow-up study of 187,766 men. J Am Med Assoc 155: 1316-1326.
10. HAMMOND EC (1954) Smoking in relation to lung cancer. Conn State Med J 18: 3-9.
11. HAMMOND EC (1964) SMOKING IN RELATION TO MORTALITY AND MORBIDITY. FINDINGS IN FIRST THIRTY-FOUR MONTHS OF FOLLOW-UP IN A PROSPECTIVE STUDY STARTED IN 1959. J Natl Cancer Inst 32: 1161-1188.
12. HAMMOND EC (1965) EVIDENCE ON THE EFFECTS OF GIVING UP CIGARETTE SMOKING. Am J Public Health Nations Health 55: 682-691.
13. Doll R, Peto R (1981) The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst 66: 1191-1308.
14. Thun MJ, Carter BD, Feskanich D, Freedman ND, Prentice R, et al. (2013) 50-year trends in smoking-related mortality in the United States. N Engl J Med 368: 351-364.
15. Carle EE, Thun MJ, Petrelli JM, Rodríguez C, Heath CW Jr (1999) Body-mass index and mortality in a prospective cohort of U.S. adults. N Engl J Med 341: 1097-1105.
16. Vainio H, Kaaks R, Blanchini F (2002) Weight Control and Physical Activity. Eur J Cancer Prev 2: 894-100
17. Smith-Warner SA, Spiegelman D, Ritz J, Albanes D, Beeson WL, et al. (2006) Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. Am J Epidemiol 163: 1053-1064.
18. Thun MJ, Peto R, Lopez AD, Monaco JH, Henley SJ, et al. (1997) Alcohol consumption and mortality among middle-aged and elderly U.S. adults. N Engl J Med 337: 1705-1714.
19. Shults JM, Novotny TE, Rice DP (1991) Quantifying the disease impact of cigarette smoking with SAMMEC II software. Public Health Rep 106: 326-333.
20. Thun MJ, Henley SJ, Carle EE (2002) Tobacco use and cancer: an epidemiologic perspective for geneticists. Oncogene 21: 7307-7325.
21. Thun MJ, Calle EE, Rodriguez C, Wingo PA (2000) Epidemiological research at the American Cancer Society. Cancer Epidemiol Biomarkers Prev 9: 861-868.

22. Thun MJ, Namboodiri MM, Heath CW Jr (1991) Aspirin use and reduced risk of fatal colon cancer. N Engl J Med 325: 1593-1596.

23. Baron JA, Cole BF, Sandler RS, Halle RW, Ahnen D, et al. (2003) A randomized trial of aspirin to prevent colorectal adenomas. N Engl J Med 348: 891-899.

24. Burn J, Gerdes AM, Macrae F, Medklin JP, Moeslein G et al. (2011) Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. Lancet 378: 2081-2087.

25. Giovannucci E, Egan KM, Hunter DJ, Stampfer MJ, Colditz GA, et al. (1995) Aspirin and the risk of colorectal cancer in women. N Engl J Med 333: 609-614.

26. Flossmann E, Rothwell PM; British Doctors Aspirin Trial and the UK-TIA Aspirin Trial (2007) Effect of aspirin on long-term risk of colorectal cancer: consistent evidence from randomised and observational studies. Lancet 369: 1603-1613.

27. Collaborative Group on Hormonal Factors in Breast Cancer (1996) Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Lancet 347: 1713-1727.

28. [No authors listed] (1997) Breast cancer and hormone replacement therapy. Combined reanalysis of data from 51 epidemiological studies involving 52,705 women with breast cancer and 108,411 women without breast cancer. Lancet 350: 1047-1059.

29. Hunter DJ, Riboli E, Haiman CA, Albanes D, Altshuler D, et al. (2005) A candidate gene approach to searching for low-penetrance breast and prostate cancer genes. Nat Rev Cancer 5: 977-985.

30. Patel AV, Cheng I, Canzian F, Le Marchand L, Thun MJ, et al. (2008) IGF-1, IGFBP-1, and IGFBP-3 polymorphisms predict circulating IGF levels but not breast cancer risk: findings from the Breast and Prostate Cancer Cohort Consortium (BPC3). PLoS One 3: e2578.

31. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, et al. (2007) A genome-wide association study identifies alleles in FGFFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 39: 870-874.

32. Yeager M, Chatterjee N, Ciampa J, Jacobs KB, Gonzalez-Bosquet J, et al. (2009) Identification of a new prostate cancer susceptibility locus on chromosome 8q24. Nat Genet 41: 1055-1057.

33. Barrington de Gonzalez A, Hartge P, Cerhan JR, Flint AJ, Hannan L, et al. (2010) Body-mass index and mortality among 1.46 million white adults. N Engl J Med 363: 2211-2219.

34. Sutcliffe S, Neace C, Magnuson NS, Reeves R, Aaldere JE (2012) Trichomonosis, a curable UTI, and prostate carcinogenesis—a proposed molecular mechanism. PLoS Pathog 8: e1002801.

35. Ogino S, Lochhead P, Chan AT, Nishihara R, Cho E, et al. (2013) Molecular pathological epidemiology of epigenetics: emerging integrative science to analyze environment, host, and disease. Modern pathol 26: 465-484.

36. McCullough ML, Patel AV, Kushi LH, Patel R, Willett WC, et al. (2011) Following cancer prevention guidelines reduces risk of cancer, cardiovascular disease, and all-cause mortality. Cancer Epidemiol Biomarkers Prev 20: 1098-1097.

37. Khoury MJ, Lam TK, Ioannidis JP, Hartge P, Spitz MR, et al. (2013) Transforming epidemiology for 21st century medicine and public health. Cancer Epidemiol Biomarkers Prev 22: 508-516.

38. Oster G, Colditz GA, Kelly NL (1984) The economic costs of smoking and benefits of quitting for individual smokers. Prev Med 13: 377-389.

39. Huse DM, Oster G, Kilen AR, Lacey MJ, Colditz GA (1989) The economic costs of non-insulin-dependent diabetes mellitus. JAMA 262: 2708-2713.

40. Oster G, Delea TE, Huse DM, Regan MM, Colditz GA (1996) The benefits and risks of over-the-counter availability of nicotine polacrilex (“nicotine gum”). Med Care 34: 389-402.

41. Oster G, Huse DM, Delea TE, Colditz GA (1986) Cost-effectiveness of nicotine gum as an adjunct to physician’s advice against cigarette smoking. JAMA 256: 1315-1318.

42. Oster G, Huse DM, Delea TE, Savage DD, Colditz GA (1987) Cost effectiveness of labetalol and propranolol in the treatment of hypertension among blacks. J Natl Med Assoc 79: 1049-1055.

43. Oster G, Tuden RL, Colditz GA (1987) A cost-effectiveness analysis of prophylaxis against deep-vein thrombosis in major orthopedic surgery. JAMA 257: 203-208.

44. Oster G, Tuden RL, Colditz GA (1987) Prevention of venous thromboembolism after general surgery. Cost-effectiveness analysis of alternative approaches to prophylaxis. Am J Med 82: 889-899.

45. Goldie SJ, Gaffikin L, Goldhaber-Fiebert JD, Gordillo-Tobar A, Levin C, et al. (2005) Cost-effectiveness of cervical-cancer screening in five developing countries. N Engl J Med 353: 2158-2168.

46. Heckman J, Navarro-Lozano S (2004) Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models. Review of Economics and Statistics 861: 30-57.

47. Woolridge JM (2009) Introductory econometrics: a modern approach. (4th edn) Mason, OH: South Western, Cengage Learning.

48. Bonabeau E (2002) Agent-based modeling: methods and techniques for simulating human systems. Proc Natl Acad Sci U S A 99 Suppl 3: 7280-7287.

Citation: Colditz GA, Chang SH (2013) The Expanding Potential for Cohort Studies to Inform Priorities for Cancer Prevention. Epidemiol 3: e112. doi:10.4172/2161-1165.1000e112