Published in final edited form as:

Epidemiology. 2015 May ; 26(3): 362–364. doi:10.1097/EDE.0000000000000272.

Cohort studies of the efficacy of screening for cancer

Noel S. Weiss

University of Washington, and the Fred Hutchinson Cancer Research Center, Seattle WA

Abstract

In most research settings, the absence of information regarding the presence of symptoms and signs of a cancer in BOTH screened and unscreened persons will limit the ability of cohort studies to provide a valid estimate of the efficacy of screening to prevent mortality from that cancer. Bias from this source generally will be small or absent in cohort studies that seek to measure the impact of those forms of cancer screening that have the potential to decrease cancer incidence, but the analysis of such studies must incorporate assumptions regarding the natural history of the cancer in question.

Cohort studies of the efficacy of screening to prevent mortality from cancer are characterized by a comparison of the observed number of deaths among persons screened for a given form of cancer with that expected in the absence of screening. Obtaining valid estimates of the efficacy of screening to prevent death from cancer by means of cohort studies requires data on the presence of symptoms and signs of the malignancy in question in cohort members, so that screening and non-screening exams can be distinguished from one another. Unfortunately, in many settings such data are not available, compromising the ability of these studies to further our understanding of the impact of screening on cancer mortality.

For screening modalities that have the potential to identify treatable cancer precursor lesions, cohort studies also can compare observed and expected cancer incidence. While the lack of high-quality data on the reason for receipt of a given test (screening, or a response to the presence of symptoms or signs) is not so great a threat to the validity of these studies – the tests in question are those that would have taken place well before the time of diagnosis, at a time when symptoms and signs from the cancer are not likely to have been present – they face other issues that could compromise their value.

The purpose of this Commentary is to describe the design and analytic strategies that can maximize the likelihood that cohort studies of cancer screening efficacy, both those of mortality and those of incidence, will provide valid results.

Mortality Studies

Some cohort studies of the causes of cancer compare the observed mortality in a group of individuals in whom a particular exposure has been present with that expected based on the

rate in the population in which cohort members reside. However, for the evaluation of a cancer screening examination that seeks to identify occult cancer (or a precursor lesion) in persons without a history of the disease, such an approach is not useful:

1. If the screening modality in question has been relatively commonly used, to the extent that screening is effective the population mortality rate will not reflect the rate in unscreened persons.
2. No matter what the prevalence of screening, screened persons and those in the underlying population will differ in two important respects: a) a prior diagnosis of the cancer in question is possible only in the latter group. (Tests done in persons with an earlier diagnosis of a given cancer to look for additional or recurrent cancer are a form of “surveillance”, not screening); and b) signs and symptoms of the cancer cannot, by definition, be present in truly screened individuals – tests in such persons would be “diagnostic” in nature, not screening - but they will be present in some proportion of members of the population being used as a basis for comparison. Therefore, even in the absence of effective treatment of early-stage cancer, a reduced rate of death from a given form of cancer in screened persons (relative to the rate in the population as a whole) would be expected due to the preferential exclusion from the screened group of those with symptomatic-but-undiagnosed cancer and those previously diagnosed with the cancer.

Some studies of disease etiology based on comparisons of “exposed” and “non-exposed” cohort members have also obtained information on screening history. For example, during 1988–2008, participants in the Nurses’ Health Study and Health Professionals Follow-up Study were queried every two years regarding receipt of screening colorectal endoscopy (1). Among persons with no history of colorectal cancer as of the completion of the initial questionnaire, mortality from colorectal cancer was compared between those who reported ever having been screened (person- time accrued from the midpoint of the 2-year questionnaire cycle in which the first screen had been reported) and those who had not (yet) been screened. While there were fewer deaths from colorectal cancer than expected following receipt of screening, two sources of potential bias hinder the interpretation of the observed decrease in mortality:

- a. To the extent that some endoscopic examinations done for signs or symptoms of colorectal cancer were incorrectly reported as screens, the decreased risk of colorectal cancer death due to screening would have been underestimated (i.e. there would have been a falsely high mortality ratio associated with receipt of screening); and
- b. To the extent that ascertainment of screening endoscopy was accurate, screened and unscreened cohort members would not be comparable for the presence of signs and symptoms of colorectal cancer; these could have been present only among unscreened persons. Because the presence of signs and symptoms of colorectal cancer would be expected to be associated with an increased likelihood of being diagnosed with (and therefore potentially dying from) colorectal cancer, the observed mortality ratio associated with

receipt of screening would be falsely low (i.e. providing an overestimate of screening efficacy).

There was considerable specificity of the associations observed in this study: receipt of screening sigmoidoscopy was associated with a reduction in mortality only from cancers that arose in the rectum and distal colon, whereas screening colonoscopy was associated with a reduction in mortality from cancers throughout the colon and rectum. Therefore, it seems likely that the results are an indication of genuine efficacy, even if the potential sources of bias interfere with our ability to quantify that efficacy very well.

The introduction of large-scale government-sponsored mammography screening programs in a number of European countries has led to the existence of screened and unscreened cohorts of women whose mortality from breast cancer can be ascertained. Breast cancer mortality among initially breast cancer-free women (according to registry records) who accepted the invitation has been compared to that among women who did not (2). Certainly, there is the possibility for confounding in these studies. The underlying risk of death from breast cancer may differ between women who do and do not accept an invitation to undergo screening, due to a difference in the prevalence of risk factors for breast cancer incidence and/or to the presence of signs or symptoms that may enhance participation in the program. However, several estimates of the collective magnitude of confounding from these sources are available, through a comparison of breast cancer mortality of invited non-participants and women who had not been invited (3). The increased mortality among women who declined an invitation – some 30%–40% above that in those who were not invited – argues that women who accept an invitation for screening mammography may well be at somewhat lower risk relative to women who decline screening. Methods for obtaining estimates of screening efficacy that account for this bias have been developed (3).

The problem of confounding in cohort studies of the efficacy of screening in reducing cancer mortality is a generic one, and not confined to screening mammography. Would the estimates of the magnitude of confounding derived from the experience obtained in population-wide breast cancer screening programs in Europe generalize more broadly, perhaps to other screening modalities evaluated in other parts of the world, and to settings in which a government-sponsored program has not issued invitations? I suspect the answer to that question is “no”: the balance of the different factors that bear on the receipt of a given form of cancer screening is likely to differ widely by type of test, geographic location, and type of health care system.

Incidence Studies

By identifying treatable precursor lesions to cancer (eg cervical intraepithelial neoplasia, colon polyps), some tests (eg cervical cytology, colonoscopy) have the potential to decrease cancer incidence. In assessing the degree to which that potential is fulfilled, the future incidence of the cancer in question in screened persons must be compared to that expected on the basis of the incidence in persons who have not been screened. Malignancies found at the time of an initial screen cannot be considered; that initial screen did not have the potential to prevent these invasive lesions. But the determination of the rate of cancer in screen-negative persons cannot begin immediately following their negative test result. The

incidence of cancer in these persons will be relatively low for a period of time, due to the removal from the cohort of prevalent cases of cancer. No corresponding “removal” of prevalent cases can be done in unscreened persons. As a result of this form of “healthy screenee bias” (4), even if the identification of precursor lesions does not lead to treatment that can prevent cancer, if follow-up were to begin immediately following receipt of a negative screening exam the ratio of observed to expected cases would be less than one.

Minimizing healthy screenee bias requires delaying the accrual of cancer cases and person-time in the screened group until prevalent cases in unscreened persons would become clinically evident, i.e. until the remaining duration of the occult invasive phase of the cancer has elapsed. Possibly with this in mind, the accrual of cancer cases and person-time in the Nurses’ Health Study and Health Professionals Follow-up Study (1) was based on information on receipt of screening that had been obtained at least two years in the past. Of course, the duration of the occult invasive phase of a malignancy varies from tumor to tumor, and cannot be known in any specific instance. Nonetheless, the duration of the occult invasive phase may be short relative to that of the pre-invasive detectable phase of the tumor, the period of time during which screening has the potential to lead to a decreased incidence of cancer. (Such a difference in these durations likely is present for cervical cancer - the average duration of cervical precancerous lesions considerably exceeds the duration of occult invasive cancer (5), and also for colorectal cancer, as polyps typically are present for much longer than occult cancer (6).) In such instances, errors in estimating the duration of the occult invasive phase of the cancer would not detract to any great extent from the validity of the estimate of the ability of screening to give rise to decreased cancer incidence.

In a cohort study with very long-term follow-up, a cancer may be diagnosed towards the end of the study whose pre-invasive state emerged only after the performance of a negative screening test near the beginning of follow-up. Because of when it was performed, that test did not have the ability to lead to treatment to prevent the occurrence of this particular malignancy. The analysis of studies in which the duration of follow-up of cohort members exceeds the estimated duration of the tumor’s pre-invasive detectable phase needs to accommodate this scenario, by re-categorizing person-time from “screened” to “unscreened” once a certain number of years following the last screen has elapsed.

Consideration of the respective durations of the pre-invasive detectable phase and occult invasive phase of a malignancy in assessing the impact of screening on cancer incidence is not unique to cohort studies. In case-control studies, screening histories of persons with incident invasive cancer and controls are compared for the period beginning at the time prior to diagnosis that corresponds to the combined duration of the two phases and ending x months prior to diagnosis, where x equals the estimated duration of the occult invasive phase (7).

The potential source of bias in cohort studies of the efficacy of screening against mortality from cancer – arising in most such studies from there being the possibility of symptoms or signs of the cancer only in unscreened persons – would not be expected to be present to any appreciable degree in cohort studies of screening against cancer incidence. The pre-invasive lesions being sought in such screening, e.g. cervical intraepithelial neoplasia or colon

polyps, in most instances produce no symptoms or signs (5, 6). Of course, there can be differences between screened and unscreened persons with regard to characteristics that predispose to the occurrence of the cancer in question, e.g. genetic characteristics that have led to a positive family history. To the extent possible given the data available, confounding from these sources should be controlled.

In the report from the Nurses' Health Study/Health Professionals Follow-up Study on screening endoscopy (1), the observations pertaining to the incidence of colorectal cancer were stratified according to test result: negative exam, versus polyp(s) identified and removed. Such an analytic approach addresses the ability of a negative exam (or a positive one with polyp removal) to predict the subsequent risk of colorectal cancer, relative to the risk in unscreened persons. But in order to assess the magnitude of the impact of screening endoscopy in preventing cancer, it is necessary to compare the incidence in unscreened persons to all of those who were screened, irrespective of the test results.

Conclusion

There is considerable opportunity for bias in a cohort study of cancer screening efficacy. Awareness of the most important potential sources of bias can help in deciding whether to embark on such a study in a particular setting, and in the structuring of the comparisons to be made between screened and unscreened persons.

Acknowledgments

Aruna Kamineni, PhD, Paul Doria-Rose, PhD

Financial support – Grant K05 CA092002 from the National Cancer Institute

Grant support: This work was supported in part by Grant number 2 K05 CA092002 from the National Cancer Institute

The author thanks Dr. Paul Doria-Rose and Dr. Aruna Kamineni for their suggestions.

References

1. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *NEJM*. 2013; 369(12):1095–1105. [PubMed: 24047059]
2. Hofvind S, Ursin G, Tretli S, et al. Breast cancer mortality in participants of the Norwegian Breast Cancer Screening Program. *Cancer*. 2013; 119(17):3106–3112. [PubMed: 23720226]
3. Duffy SW, Cuzick J. Correcting for non-compliance bias in case-control studies to evaluate cancer screening programmes. *Appl Statist*. 2002; 51:235–243.
4. Weiss NS, Rossing MA. Healthy screenee bias in epidemiologic studies of cancer incidence. *Epidemiology*. 1996; 7(3):319–322. [PubMed: 8728451]
5. Schiffman, MH.; Hildesheim, A. Cervical cancer, in *Cancer Epidemiology and Prevention*. 3. Schottenfeld, D.; Fraumeni, JF., editors. Oxford; New York: 2006.
6. Giovannucci, E.; Wu, K. Cancers of the colon and rectum. In: *Schottenfeld, D.; Fraumeni, JF., editors. Cancer Epidemiology and Prevention*. 3. Oxford; New York: 2006.
7. Weiss NS. Case-control studies of the efficacy of screening tests designed to prevent the incidence of cancer. *Am J Epidemiol*. 1999; 149(1):1–4. [PubMed: 9883787]