

Observational Research Designs

7

7.1 An Overview of Observational Study Designs

Observational studies do not dictate the cancer screening regimens that their study subjects utilize. Instead, these studies collect data on individuals' cancer screening practices, cancer outcomes, and other factors if needed. Because no regimens are dictated, an observational study can capture information about and evaluate a variety of cancer screening practices, including use of different tests or cancer screening regimens. Observational studies can be retrospective or prospective in nature, with the distinction dependent on how and when individuals are chosen for study inclusion. Prospective observational studies of cancer screening track individuals as they move forward in time until the event of interest happens or the study is complete. Retrospective observational studies of cancer screening look at past experiences of individuals who have had the event of interest and others who have not. Prospective observational studies are said to sample based on exposure (cancer screening experience), while retrospective observational studies are said to sample based on outcome (death).

Observational studies provide weaker evidence than experimental studies because observational studies are subject to confounding. Confounding occurs when a third factor is associated

with both the cancer screening practice and cause-specific mortality, meaning that the third factor is not equally present among groups of individuals with different cancer screening practices and is not equally present among groups of individuals with different cancer outcomes. An example comes from observational studies of colorectal cancer screening and colorectal cancer mortality. Individuals who exercise are more likely to have colorectal cancer screening and also are less likely to die of colorectal cancer. If an observational study observes a reduction in colorectal cancer mortality with cancer screening, we cannot be sure what is responsible. Is it cancer screening, exercise, a combination of both, or some unknown protective factor that is more likely among individuals who receive cancer screening and who exercise?

Confounding is a type of bias and leads to an incorrect estimate of the true relationship of cancer screening and cause-specific mortality. It is present in varying degrees in all observational studies and while it can be dampened using statistical methods, these methods cannot eliminate all confounding because it is not possible to measure all confounders or measure them accurately.

Observational studies usually are less expensive and easier to perform than experimental studies. There are many reasons for that: the study usually does not administer or pay for the cancer screening test; existing databases often are used; and retrospective studies do not need to wait for time to pass since the data already have been collected. Some prospective studies can take as long or longer than a randomized controlled trial (RCT), however. Retrospective studies are often used as a first pass to examine a hypothesis about a cancer screening test, especially if use of that test is prematurely disseminating into community practice.

Observational study designs that are frequently used in cancer screening assessment will be discussed: cohort, case-control, and ecologic. Single-arm studies, sometimes known as case series, will be presented as well. Readers who wish to learn more about observational research can consult *Modern Epidemiology* (3rd edition) by Rothman, Greenland, and Lash [1].

7.2 Cohort Studies

7.2.1 Design Features

A cohort is a group of people with something in common, either by nature or design, who are followed through time for an event of interest. Research cohorts can be created in one of two ways. Prospective cohorts are created in real-time; data is collected as time passes. Retrospective cohorts, also known as historic cohorts, are created after data have been collected. These cohorts comprise extracted data from pre-existing data sources, such as Medicare or the medical records of health maintenance organization members. Retrospective cohorts usually are analyzed as if their data had been collected prospectively and generally are constructed for the purpose of addressing pre-determined research questions. In prospective cohort studies, individuals usually are recruited to actively participate in the study, but with retrospective cohort studies, individuals usually do not know that their data are being used to answer a specific research question.

Information on cohort experience can come from a variety of data sources. Prospective cohorts usually rely heavily on participant interviews and participant-completed questionnaires, and may use medical records to validate procedures and diagnoses. Retrospective cohorts typically have little-to-no additional information collected on them. In both instances, deaths can be confirmed with collection of death certificates, while death certificate cause of death can be verified with review of medical records that document clinical experiences prior to death. It is recommended that records for at least the 3–6 months prior to the date of death be considered [2].

No prospective cohorts have been established for the primary purpose of examining cancer screening effectiveness, though some have been established to collect other information about the cancer screening process in community settings. Some pre-existing prospective cohorts have been used to address effectiveness if cancer screening and death information are available. Many retrospective cohorts have been created to address a range of questions regarding cancer screening. Cohorts without

information on cancer screening can be repurposed by collecting the needed information if it is available.

7.2.2 Analysis Features

Cohort members choose their cancer screening regimens, which means that confounding is all but guaranteed. Therefore, outcome measures need to be calculated using statistical models that allow for adjustment for confounding variables. If timing of death (date of death or person-years of experience) is available, Cox proportional hazards regression or Poisson regression can be used, with the choice determined by assumptions regarding whether the hazard of death changes over time [3]. Logistic regression can be used if information on timing of death is unavailable.

Poisson regression produces a cause-specific mortality rate ratio, Cox proportional hazards regression produces a cause-specific hazard rate ratio, and logistic regression produces an odds ratio, which estimates a risk ratio (also known as a relative risk) in the case of a rare outcome like cancer. Each ratio represents a measure of disease burden in the individuals who received the cancer screening regimen of interest divided by those who did not. When assessing cancer screening data, the exact measure used (mortality rate, hazard rate, or odds) is of less importance than the ratio that they will produce. The three methods, when applied to a cancer screening cohort with typical experience, usually will produce ratios that lead to the same conclusion about the benefit of cancer screening.

Risk difference measures are sometimes used to describe how the absolute rather than relative magnitude of disease burden changes with cancer screening. To calculate a risk difference, the measure of interest (incidence rate, mortality rate, or hazard rate) in the presence of cancer screening is subtracted from the measure of interest in the absence of cancer screening. For example, a cause-specific mortality rate of 4 per 1000 person-years in the absence of cancer screening and a cause-specific mortality rate of 3 per 1000 person-years in the presence of cancer screening result in a risk difference of 1 per 1000 person-years. Difference measures are more useful than relative measures when considering health care resource allocation.

7.2.3 Strengths and Weaknesses

Cohort studies allow for evaluation of effectiveness, something of the utmost importance because the manner in which cancer screening is utilized in community settings is often quite different from the idealized regimens in RCTs. For example, an RCT might test an annual regimen, but the regimen that evolves in the community could have longer or varied cancer screening intervals, especially when the cancer screening test is not fully acceptable to community members. Cohort studies also can be used to examine uptake of a new cancer screening test or test performance measures.

The value of a cohort often depends on the extent of confounding and timing of data collection. The chance and possible impact of confounding must be discussed whenever cohort data are presented. Regarding timing, cohort data are analyzed as if data were collected with the passing of time, meaning that collection of information on exposure (cancer screening and confounding) occurs before the outcome (cause-specific mortality) has occurred or is known. The data for some cohorts are collected that way, but for cohorts that are repurposed, researchers often need to collect information on past events. These data may be affected by recall bias, which happens when the passage of time results in data errors that then lead to incorrect estimates of the true relationship between cancer screening and cause-specific mortality. For example, let's say a cohort is used to examine the ability of colonoscopy to reduce colorectal cancer mortality. When participants are asked about cancer screening use in the past 10 years, some may erroneously report a past colonoscopy when in fact their exam was a flexible sigmoidoscopy, which also reduces colorectal cancer mortality but to a lesser degree. Non-trivial error in reporting would lead to an observed association between colonoscopy and colorectal cancer that is weaker than the real association. The recall error led to measurement of the impact of receiving flexible sigmoidoscopy or colonoscopy, rather than only colonoscopy.

Needless to say, it's best to collect information as soon as possible after an exposure occurs. It is probably best to collect information on past cancer screening activities from medical records or health insurance claims rather than participant

interviews, although medical records can be lost, and laws may be enacted that make use of both sources more difficult. Also, medical records do not always provide complete or correct information. They are subject to human error and in some instances creative procedure coding to maximize insurance reimbursement.

To have adequate statistical power, cohort studies evaluating cancer screening usually need to be large and have a number of years of follow-up. Establishment of a new cohort and the infrastructure to track the experience of the participants will be expensive, although typically less than that of an RCT, as cancer screening activity and follow-up occurs as part of community care. Repurposing of an existing cohort can save money and time, but the need for additional data will lead to a reduction in resources saved.

7.2.4 Variations

A nested case-control study of cancer screening uses all cause-specific deaths in a cohort as cases, but only a subset of the rest of cohort members as controls. This design is used when additional data collection is needed and is expensive or time-consuming, as in the situation of needing to determine the indication for a medical test. Nested case-control studies of cancer screening are constructed and analyzed in the same manner as case-control studies of cancer screening; the only difference is that cases and controls are drawn from an established cohort rather than another source. Details of case-control studies of cancer screening, nested and otherwise, are presented later on in this chapter.

7.2.5 Examples of Cancer Screening Cohort Studies

The BCSC is a prospective cohort study of breast cancer screening. It is a cancer screening test registry: information on screening mammograms and other breast cancer screening imaging tests is collected, as well as information on the women who receive them.

The unit of analysis is often a test rather than a woman. The BCSC is not intended to evaluate cancer screening effectiveness; instead, it strives to “assess and improve the delivery and quality of breast cancer screening and related patient outcomes”. The cohort has been used to evaluate important issues in breast cancer screening, including screening adherence, test performance, and supplemental screening [4].

The Nurses’ Health Study (NHS) and Health Professionals Follow Up Study (HPFS) are on-going prospective cohort studies, each designed to explore causes of major health conditions in the US. The NHS began in 1976 and the HPFS in 1985. Both studies added questions to their self-administered questionnaires in 1990 regarding receipt of lower endoscopy (colonoscopy or sigmoidoscopy). The researchers published findings on the impact of lower endoscopy on colorectal cancer mortality in 2013 [5].

Kaiser Permanente of North California (KPNC) is an integrated health care delivery system with more than 4 million members. Their extensive electronic health databases have been used to address many questions in cancer etiology and prevention, including cancer screening. An example of how a health care organization’s databases can be used to conduct a retrospective cohort study of cancer screening can be found in KPNC’s report on the long-term risk of colorectal cancer death after a negative colonoscopy [6].

7.3 Case-Control Studies

7.3.1 Design Features

A case-control study is retrospective in nature, meaning that all exposures and events have occurred before the study begins. A case-control study includes cases, who are individuals who had the outcome of interest, and controls, who are individuals who did not have that outcome at a point in time that is determined by the case’s experience. The design has been used extensively in cancer etiology studies. A case-control study often aims to include the universe of cases: all individuals who experience the outcome of

interest during a specific time period. Controls are randomly sampled (usually within age strata) from the population that gave rise to the cases. In case-control studies of cancer etiology a population-based roster, such as a list of drivers' license holders, is used to sample controls.

In principle, case-control studies of cancer screening are the same as case-control studies of etiology. Cases are individuals who have died due to the cancer of interest. Controls are selected for a specific case, with random selection usually stratified on age and sex of the case. In addition, controls must not have been diagnosed with the cancer of interest prior to the case's diagnosis date; the reason is to ensure an equal and contemporaneous opportunity for cancer screening. Some case-control studies of cancer screening have required that controls be alive on the date of the case's death. Cancer screening experience during a specific period (as discussed below) is compared in cases and controls.

Case-control studies of cancer screening usually select their cases and controls from health system patient rosters because access to medical records is a necessity. Medical records are used to determine whether a test was done for cancer screening as opposed to diagnostic evaluation, and obtain details of cancer diagnoses. As was noted earlier in this chapter, case-control studies of cancer screening also can be constructed by selecting cases and controls from an established cohort.

7.3.2 Analysis Features

Case-control studies of cancer screening are designed and analyzed as matched case-control studies because exposure assignment for controls is defined by the experience of a case. Conditional logistic regression models are used to account for the matching and to adjust for other possible confounders. Logistic regression produces an odds ratio; in the instance of a case-control study of cancer screening, it is the ratio of the odds of receiving cancer screening among those who died of the cancer of interest divided by the odds of receiving cancer screening among those who did not die of the cancer of interest.

The primary challenge in analysis of case-control studies of cancer screening is assessing cancer screening exposure. An exposure window, one that reflects the period when cancer screening could have been beneficial to cases (Phase B as defined in Chap. 2), must be defined. The exposure window for cases ends no later than the date of diagnosis, and usually ends prior to the date of diagnosis to exclude the period when cases were undergoing diagnostic evaluation. Controls are given a reference date, which corresponds to the final date of their matched case's exposure window. Only cancer screening experience prior to that date is considered to be in the exposure window. Cancer screening tests that occurred in the distant past should be excluded if there is reason to believe that they were done prior to the time the case's cancer was in Phase B.

7.3.3 Strengths and Weaknesses

Case-control studies of cancer screening are retrospective research and can be done more quickly and inexpensively than cohort studies or RCTs. The number of cases is known at the start of the study, and controls are selected only if they match to a known case. Detailed information, such as that found in medical records, usually is needed to determine whether a test was for cancer screening and whether it occurred within the exposure window.

Confounding is a concern in case-control studies of cancer screening. Recall bias may be of concern if medical record abstractors are aware of the study hypothesis, or if medical records are systematically missing information, or are systematically unavailable. Because the exposure window must be inferred, it never will correctly capture the exact period in which cancer screening could have been of benefit to the cases. The exposure window must be thoughtfully chosen, and sensitivity analyses can be used to explore the impact of varying its definition.

The many methodologic challenges in design and analysis of case-control studies of cancer screening are discussed in Cronin et al. [7] and Weiss [8].

7.3.4 Example of Case-Control Studies of Cancer Screening

Using data on women residing in Saskatchewan, Pocobelli and Weiss conducted a case-control study of breast cancer mortality in relation to receipt of screening mammography [9]. Saskatchewan has a universal health care system funded by the government, with nearly all residents eligible for coverage. About 90% of residents also are eligible for province-funded outpatient prescription drug benefits. The cases and controls for the cancer screening study were sampled from a larger study that utilized the roster of women with drug benefits. Cases were women who died due to breast cancer at 50–79 years of age during the years 1990–2008. Controls were selected for each case and were women who had the same birth year as the case and were not diagnosed with breast cancer prior to the case's date of diagnosis. Additional methodologic considerations, including definition of the exposure window, are discussed in the paper.

7.4 Ecologic Studies

7.4.1 Design Features

An ecologic study is the observational equivalent of a cluster-level RCT: the experience of groups, usually geopolitical entities, rather than individuals, is examined. The outcome of interest is cause-specific mortality rates, and the exposure is a measure of cancer screening utilization in the entity. Ecologic studies of cancer screening often compare cause-specific mortality rates for countries with different degrees of cancer screening utilization.

7.4.2 Analysis Features

Data from ecologic studies of cancer screening often are presented using simple two-axis plots. Cause-specific mortality rates are plotted on one axis and their associated cancer screening use

metric is plotted on the other. Percentage of eligible individuals screened is an example of a metric that has been used in ecologic studies. If cancer screening reduces cause-specific mortality, a graph of cause-specific mortality rates on the y-axis and the cancer screening metric that measures use on the x-axis should produce a pattern of negative correlation. Figure 7.1 presents a fictional ecologic study in which utilization of cancer screening and cause-specific mortality are negatively correlated, as suggested by a fitted line that slopes downward. We cannot assume that cancer screening is the reason for the decrease in cause-specific mortality as other factors may be at play. In the instance of an ecologic study that suggests a reduction in cause-specific mortality, changes or regional differences in cancer treatment need to be carefully considered as confounders. Accuracy of the summary measures must be considered as well.

Ecologic studies can provide compelling evidence that cancer screening implementation has not led to reductions in cause-specific mortality for some cancer sites. In Fig. 7.2, the cause-specific mortality rate hovers between 3.5 and 4.0 per 10,000 person-years regardless of cancer screening uptake, and the fitted line suggests no negative correlation. It is unlikely that such a pattern would mask a benefit of cancer screening due to

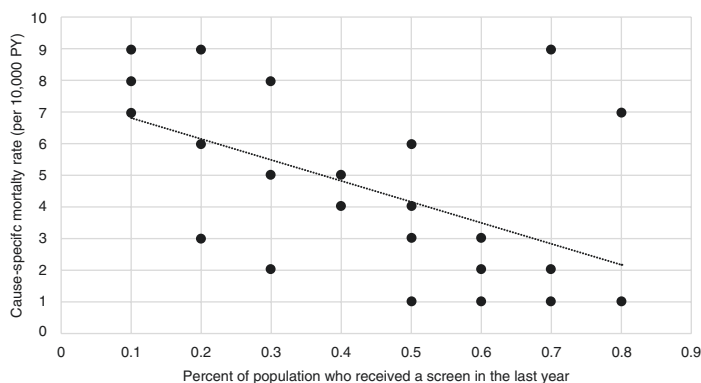


Fig. 7.1 Plot of data from an ecologic study that suggests a benefit of cancer screening. PY stands for person-years. Data are fictional

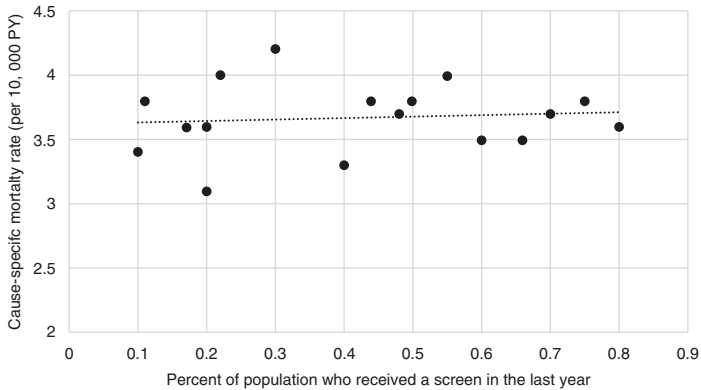


Fig. 7.2 Plot of data from an ecologic study that suggests no benefit of cancer screening. PY stands for person-years. Data are fictional

confounding, as the confounding factor would need to increase cause-specific mortality and increase cancer screening use. Though very high-risk individuals do tend to receive cancer screening more frequently, they are a small percentage of the individuals in a population, and cannot drive entity-level rates unless most deaths from cancer occur in the high risk group.

7.4.3 Strengths and Weaknesses

Ecologic studies of cancer screening are usually easier to undertake and less expensive than individual-level observational studies. Cause-specific mortality rates are publicly available for geographic entities in the US and elsewhere. Obtaining data on cancer screening use within a group is more challenging, but data sources already exist in the US for some cancer screening practices. Electronic medical records or administrative claims from a universal provider such as Medicare could be used as well. It can, however, be challenging to identify a set of entities to compare. To determine appropriateness, it is useful to return to the counterfactual principle and choose entities that are as comparable as possible except for cancer screening practices.

Ecologic studies have a number of shortcomings. There can be confounding at the entity level, although linear regression can be used to adjust for some degree of the influence of confounding factors if that information is available. The results may not be applicable to the individuals within the entities, as was discussed in the context of cluster-level RCTs. Measures of cancer screening utilization that are not calculated in conjunction with individual-level medical records often are overestimates, as they can reflect use of cancer screening modalities that can be used for diagnostic purposes. For example, a measure of colonoscopy utilization derived by counting all colonoscopies performed will include both screening colonoscopies and diagnostic colonoscopies.

7.4.4 Variations

A time trend study is a type of ecologic study that examines changes in cancer mortality rates as time passes, with time as a marker for changes in cancer screening practice. Time trend studies are useful for examining changes in cause-specific mortality rates after cancer screening is introduced or after cancer screening regimens change. A two-axis graph can be used, with cause-specific mortality rates on the y-axis and year on the x-axis. A metric of cancer screening utilization, if available, can be included by using a second (right-sided) y-axis. Otherwise, milestones in cancer screening practices, such as the year that cancer screening was recommended for the first time, can be annotated. Figure 7.3 is an example, again fictional, of such a graph. Rates of cause-specific mortality decrease soon after widespread recommendation of cancer screening (1993). We cannot, however, assume that the decrease in mortality is due to cancer screening; other concurrent changes, such as treatment improvements, that might explain the observed pattern need to be considered before drawing any conclusions.

In this fictional example, cause-specific mortality is stable prior to wide-spread recommendation of cancer screening in 1993. Cause-specific mortality begins to drop in 1994. It stabilizes

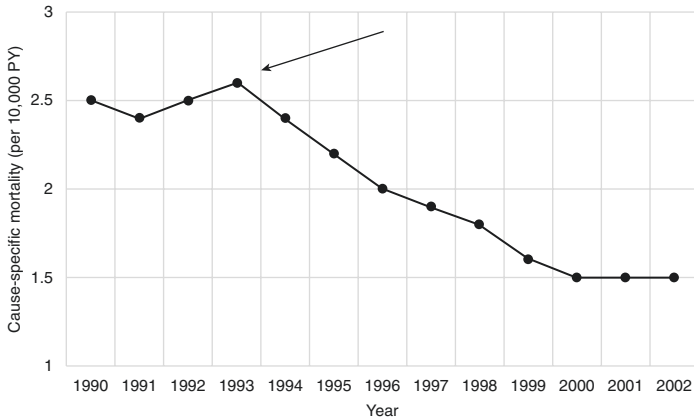


Fig. 7.3 Time trends in cancer mortality before and after recommendation of cancer screening (1993). PY stands for person-years. Data are fictional

around 2000, perhaps due to a leveling off of cancer screening utilization.

7.4.5 Examples of Ecologic Studies of Cancer Screening

A current controversy in breast cancer screening is whether reductions in breast cancer mortality are due primarily to screening or to improvements in treatment. To examine that question, Autier et al. examined breast cancer mortality rates in neighboring European countries with different histories of screening use but access to similar treatments [10]. Their ecologic analysis suggests that cancer screening has played only a minor role in improvements in breast cancer mortality.

The use of thyroid cancer screening in South Korea began to increase in 1999 when it was offered as a paid add-on test to the set of cancer screening tests offered for free through a national cancer screening program. No changes in thyroid cancer mortality were observed as utilization increased, and in 2013, use began to wane due to the evidence of no benefit and compelling evidence of overdiagnosis [11].

7.5 Single-Arm Studies

7.5.1 Design Features

In the context of cancer screening, a single-arm study refers to the experience, within a period of time, of a set of individuals who receive a screen in the context of a medical study. The screen is usually not standard of care. The test is considered to be experimental for cancer screening purposes, but a single-arm study is considered an observational study because it involves no randomization. Single-arm studies are a type of cohort study in which no participants are unscreened.

7.5.2 Analysis Features

Cancer screening single-arm studies are used to assess performance of proposed tests, most notably, the ability of a test to lead to cancer detection at an early stage. These studies tend to enroll a small number of participants. Because there is no study comparison group, results either are presented on their own or compared with those from published literature or a population-level database such as SEER.

7.5.3 Strengths and Weaknesses

Cancer screening single-arm studies are very limited in the information they can provide. The participants usually are a highly select group, suggesting that their experience is unlikely to be representative of what would occur in the general population. Participants typically are not chosen in a random fashion. They may be paid for their participation, or they may be required to pay to participate. Data collection usually does not include cause-specific mortality experience. Nevertheless, single-arm studies are a useful way to assess whether a proposed cancer screening test should receive further study.

7.5.4 Variations

A case series (or clinical series) is similar to a single-arm study. The difference is that case series include the clinical experiences of individuals not enrolled in a study. They are a culling of patients who have had the same exposure, in this case, cancer screening. Analyses examine their post-screening experience. The terms single-arm study and case/clinical series have been used interchangeably.

7.5.5 Examples of Cancer Screening Single-Arm Studies

The Mayo Clinic initiated a single-arm study in 1999 to evaluate the performance of lung cancer screening with low dose computed tomography (LDCT) [12]. At that time, there was evidence, though not definitive, that LDCT screening might reduce lung cancer mortality. The purpose of this study was to address some of the outstanding questions in LDCT screening, including the magnitude of false positive tests and the prevalence of adverse downstream effects. Participants were offered four annual LDCT screens. They were current or former cigarette smokers, with former smokers having quit less than 10 years ago. They also had at least a 20 pack-year history of smoking.

7.6 Two-in-One Single-Arm Studies

7.6.1 Design Features

In a two-in-one single-arm study, individuals are offered the chance to receive cancer screening with an experimental test in addition to and at the same time as the standard of care cancer screening test. Each participant receives both tests, and each test is evaluated without knowledge of the results of the other. Action is taken if either test result is positive.

Two-in-one single-arm studies have been used to determine if an experimental test has improved performance measures relative to the standard of care. They also have been used to examine whether an experimental screening test with a favorable feature (for example, lower cost, less invasiveness, or greater patient acceptability) has similar performance measures as the standard of care test. Two-in-one single-arm studies have been used to compare two tests already available in clinical settings. They also have been used to compare an experimental test with a diagnostic test, as diagnostic tests provide a definitive answer as to the presence of cancer.

A two-in-one single-arm study usually cannot be used to evaluate tests beyond diagnosis, although excessively optimistic speculation about the benefits of the experimental test is not uncommon.

7.6.2 Analysis Features

The analytic focus of a two-in-one cancer screening single-arm study is a comparison of the performance of the tests. Of most interest is how and when the two tests disagree. Tables 7.1 and 7.2 present data from a fictional two-in-one single-arm study with 1000 participants. Table 7.1 compares positivity rates and Table 7.2 compares cancer diagnoses.

In Table 7.1, we see that both tests returned positive results for 80 individuals. Twenty individuals, however, received a positive experimental test result and a negative standard of care test result.

Table 7.1 Comparison of results from the standard of care and experimental screening tests

		Standard of care test		Total
		Positive result	Negative result	
Experimental test	Positive result	80	20	100
	Negative result	0	900	900
	Total	80	920	1000

Data are fictional

Table 7.2 Cancer diagnoses by results of standard of care and experimental screening tests

		Standard of care test		Total
		Positive result	Negative result	
Experimental test	Positive result	35	10	45
	Negative result	2	3	5
	Total	37	13	50

Data are fictional

The experimental test had a higher positivity rate, which may or may not indicate improvement over the standard of care test. A higher positivity rate could lead to a higher false positive rate or to additional cancer diagnoses. The meaning of additional cancer diagnoses is uncertain as well. They may represent cancers that are curable due to early detection, not curable regardless of early detection, or overdiagnosed.

In Table 7.2, we see that 35 cancers were diagnosed after both a positive standard of care and experimental cancer screening test. An additional 10 cancers were diagnosed as a result of a positive experimental test, even though the standard of care test was negative. We assume that these cancers were false negatives for the standard of care test, though we will never know what would have happened in the absence of the experimental test.

7.6.3 Strengths and Weaknesses

A two-in-one single-arm study may provide useful information if the standard of care test is known to reduce cause-specific mortality and the experimental test appears to have increased sensitivity and positive predictive value (PPV). A demonstrated increase in those two performance measures is usually interpreted to mean that the new test is superior, but to make that leap, one must assume that more asymptomatic diagnoses will lead to a greater reduction in cause-specific mortality. The existence of overdiagnosis and, for some cancers, equally efficacious treatment at a later stage, challenge that assumption.

A cancer screening two-in-one single-arm study cannot provide definitive evidence of efficacy or effectiveness. In the instance of a test that has not yet disseminated, results are best used to make decisions regarding the need for an RCT.

7.6.4 Examples of Two-in-One Single-Arm Studies

Blood-based biomarker cancer screening tests are of particular interest in colorectal cancer screening as the available screening tests, lower endoscopy and fecal testing, are not palatable to many individuals. Testing for circulating methylated SEP9 DNA has been under consideration as a way to screen for colorectal cancer. To examine the performance of SEP9 testing, individuals who were scheduled for screening colonoscopy were invited to give blood plasma samples prior to their colonoscopy preparation regimen [13]. Performance measures for the SEP9 test were calculated using the results and ultimate outcome of colonoscopy screening, the gold standard in both colorectal cancer screening and diagnosis.

Screening mammography has evolved from the use of film-based to computer-based imaging. Film-based mammography only provides two-dimensional hard copy images. Digital mammography provides three-dimensional images that are read on a computer screen and can be manipulated to allow additional interpretation. The Digital Mammographic Imaging Screening Trial (DMIST) was designed to measure what were expected to be relatively small but potentially clinically important differences in diagnostic accuracy between digital and film mammography [14]. Women enrolled in the trial received both tests on the same day, and each test was read by a different radiologist. Diagnostic evaluation was performed if either test was positive. Performance measures were calculated assuming that neither test was definitive.

7.7 All Study Designs: Critical Data Elements

Most studies of cancer screening, regardless of study design, collect a large amount of data. Critical data elements are those that are necessary for proper assessment of screening performance, effectiveness, or efficacy. Other data elements may be collected for ancillary studies, or they may be collected for “what if” situations, but their collection must not jeopardize the collection of the critical data elements. Resources, including participant good will and staff time, always are limited.

Not every research endeavor will be able to collect every data element, even the critical elements. As a result, not every research endeavor will be able to answer every question. Even so, studies that do not collect every critical element may provide useful information, although the limitations of the research in the absence of such data must be clearly stated. The inability to collect most critical data elements should lead to questions about the value of the research.

Critical data elements for individual-level studies include date of birth; receipt, date and results of cancer screening tests; diagnosis of the cancer of interest; date of diagnosis; cancer characteristics (including stage, histology and location); age at death; date of death; cause of death. Indication for all relevant medical tests or procedures that are proximal to either the date of screen or the date of diagnosis should be collected to differentiate cancer screening from diagnostic evaluation. If that information is not available, any information that can be used to derive indication should be collected. Other valuable data elements include cancer treatment procedures, and adverse events of any medical procedure associated with cancer screening, diagnostic evaluation, or cancer treatment. Risk factors for the cancer of interest, as well as other potential confounders, should be collected, especially in observational studies.

Ecologic studies of cancer screening require entity-level cancer mortality rates and a metric of cancer screening use. One option for a test administered annually is to use the percent of residents who received a cancer screening test in the last

12 months. Other useful data elements include measures of cancer screening availability and characteristics of the entity that may predict cancer screening behavior, such as percent of residents with a college degree.

References

1. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 758.
2. Miller, Anthony B. (Dalla Lana School of Public Health, University of Toronto). Email to: Pamela Marcus (National Cancer Institute); 2018.
3. Fagerstrom, Richard M (National Cancer Institute). Conversation with: Pamela Marcus (National Cancer Institute); 2018.
4. Breast cancer surveillance consortium [Internet]. [Seattle]. [cited 2019 Oct 23]. Available from: <https://www.bcsr-research.org/>
5. Nishihara R, Wu K, Lochhead P, Morikawa T, Liao X, Qian ZR, Inamura K, Kim SA, Kuchiba A, Yamauchi M, Imamura Y, Willett WC, Rosner BA, Fuchs CS, Giovannucci E, Ogino S, Chan AT. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med*. 2013;19(369):1095–105.
6. Lee JK, Jensen CD, Levin TR, Zauber AG, Schottinger JE, Quinn VP, Udaltsova N, Zhao WK, Fireman BH, Quesenberry CP, Doubeni CA, Corley DA. Long-term risk of colorectal cancer and related deaths after a colonoscopy with normal findings. *JAMA Intern Med*. 2018;179(2):153–60.
7. Cronin KA, Weed DL, Connor RJ, Prorok PC. Case-control studies of cancer screening: theory and practice. *J Natl Cancer Inst*. 1998;90(7):498–504.
8. Weiss NS. *Clinical epidemiology: the study of the outcome of illness*. 2nd ed. New York: Oxford University Press; 1996. p. 163.
9. Pocobelli G, Weiss NS. Breast cancer mortality in relation to receipt of screening mammography: a case-control study in Saskatchewan, Canada. *Cancer Causes Control*. 2015;26(2):231–7.
10. Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ*. 2011;28(343):d4411.
11. Ahn HS, Kim HJ, Kim KH, Lee YS, Han SJ, Kim Y, Ko MJ, Brito JP. Thyroid cancer screening in South Korea increases detection of papillary cancers with no impact on other subtypes or thyroid cancer mortality. *Thyroid*. 2016;26(11):1535–40.

12. Swensen SJ, Jett JR, Hartman TE, et al. Lung cancer screening with CT: Mayo Clinic experience. *Radiology*. 2003;226:756–61.
13. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, Castañón-Vélez E, Blumenstein BA, Rösch T, Osborn N, Snover D, Day RW, Ransohoff DF, PRESEPT Clinical Study Steering Committee, Investigators and Study Team. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut*. 2014;63:317–25.
14. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LF, Bassett L, D'Orsi D, Jong R, Rebner M, for the Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic performance of digital versus film mammography for breast-cancer screening. *New Engl J Med*. 2005;353(17):1773–83.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

