ELSEVIER

Journal of Clinical Epidemiology 67 (2014) 184−189

Journal of Clinical Epidemiology

Comparisons of colorectal cancer mortality between screening participants and the general population are strongly biased unless an incidence-based mortality approach is used

Hermann Brenner[a,b,*], Michael Hoffmeister[a], Lina Jansen[a]

[a]*Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany*
[b]*German Cancer Consortium (DKTK), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany*

Accepted 13 August 2013; Published online 5 November 2013

Abstract

Objectives: A common approach in the evaluation of screening for colorectal cancer (CRC) is comparing observed numbers of CRC deaths in screening participants with expected numbers derived from CRC mortality in the general population. We aimed to illustrate and quantify an often-overlooked bias that may occur in such studies if CRC mortality in the general population is not restricted by the date of diagnosis (whereas screening participants by definition do not have a prior CRC diagnosis).

Study Design and Setting: We illustrate and quantify the expected bias using cancer registry data from the United States.

Results: Unless an incidence-based mortality approach is used, expected numbers of CRC deaths in screening cohorts (and hence estimated screening effects) are substantially overestimated. Overestimation of expected CRC deaths is most severe (more than fivefold) during the first year of follow-up and rapidly decreases in the subsequent years. Nevertheless, overestimation of 5- and 10-year cumulative numbers of expected CRC deaths is still as high as 60−70% and 20−30%, respectively. Substantial bias even persists if the initial years of follow-up are excluded from the analyses.

Conclusion: Careful restriction of expected CRC deaths by an incidence-based mortality approach is indispensable for deriving valid screening effect estimates.

Keywords: Bias; Cohort study; Colorectal cancer; Incidence; Mortality; Screening

1. Introduction

Randomized trials have shown that screening by fecal occult blood test or sigmoidoscopy may substantially reduce mortality from colorectal cancer (CRC) [1−4]. Strong reduction of CRC mortality after sigmoidoscopy or colonoscopy has also been suggested by multiple case−control [5−9] and cohort studies [10−14]. In several cohort studies, the observed number of CRC deaths in screening participants was compared with the expected number derived from age- and sex-specific mortality rates in the general population [10,11,14]. Although this ''standardized mortality ratio'' (SMR) approach appears straightforward, it is prone to potential biases. A well-known potential source of bias is confounding by factors associated with screening participation, such as increased health consciousness, or higher prevalence of known risk factors, such as a positive family history of CRC, which could lead to over- or underestimation of screening effects, respectively. Furthermore, screening effects could be underestimated by ''contamination,'' that is, use of screening in some proportion of the general population.

An often unconsidered source of bias results from the fact that CRC screening is done in people without a prior diagnosis of CRC, whereas CRC mortality rates in the general population are not subject to such restrictions. To obtain true SMRs, the comparison group likewise needs to be restricted to individuals with no diagnosis of CRC before the start of the observation period. Else, screening effects may be overestimated because inclusion of individuals diagnosed with CRC before the start of the observation period will increase CRC death rates in the comparison group. Although methods exist [15] and have been applied [14] to partition mortality by the date of diagnosis in the presence of high-quality cancer registry data in an ''incidence-based mortality'' approach, use of such methods for deriving

Financial support: This work was supported in part by a grant from the German Cancer Aid (grant no. 108230).

Conflict of interest: None.

* Corresponding author. Tel.: +49-6221-421300; fax: +49-6221-421302.

E-mail address: h.brenner@dkfz.de (H. Brenner).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.08.017

What is new?

- To our knowledge, this is the first study quantifying potential bias in estimating screening effects in cohort studies comparing observed colorectal cancer (CRC) mortality among participants of CRC screening and expected CRC mortality using general population mortality data.
- Unless an incidence-based mortality approach that restricts mortality by time of diagnosis is used, expected numbers of CRC deaths in screening cohorts (and hence estimated screening effects) are substantially overestimated.
- Overestimation of expected CRC deaths is most severe (more than fivefold) during the first year of follow-up and rapidly decreases in subsequent years.
- Nevertheless, overestimation of 5- and 10-year cumulative numbers of expected CRC deaths is still as high as 60–70% and 20–30%, respectively.
- Careful restriction of expected CRC deaths by an incidence-based mortality approach is indispensable for deriving valid screening effect estimates.

unbiased estimates of expected numbers of deaths has not become common practice in CRC screening studies. In this article, we illustrate and quantify the bias in estimates of screening effects resulting from the use of ''unrestricted CRC mortality data'' for deriving expected CRC deaths in SMR analyses.

2. Methods

2.1. Database

Our analyses are based on cancer registry data from the US Surveillance, Epidemiology, and End Results (SEER) Program. The SEER-9 database issued in April 2012 includes data on incident cancer cases in 1973–2009 from population-based cancer registries in Connecticut, New Mexico, Utah, Iowa, Hawaii, Atlanta (from 1975 onward), Detroit, Seattle-Puget Sound (from 1974 onward), and San Francisco-Oakland, which together cover a population of around 30 million people [16]. Geographic areas are selected for inclusion in the SEER Program based on their ability to operate and maintain a high-quality population-based cancer reporting system and their epidemiologically significant population subgroups [17]. The SEER population is comparable with the general US population, although it is more urban and has a higher proportion of foreign-born persons than the latter, and for certain cancer sites, there is underrepresentation of US cancer mortality experience [18].

To assess the potential bias by deriving expected numbers of CRC deaths from a general population cohort rather than a cohort with no prior CRC diagnosis, we conducted model calculations for birth cohort 1941–1950, aged 50–59 years at the beginning of 2000, and followed during the 10-year period from the beginning of 2000 to the end of 2009. The age group 50–59 years was chosen as CRC screening in the average risk population is commonly recommended from age 50 years onward [19,20]. The 10-year period 2000–2009 was chosen as screening intervals of 10 years are commonly recommended for screening colonoscopy, and 2000–2009 is the most recent 10-year period for which cancer registry and mortality data were available at the time of analysis.

Let CRC_{total} be the total number of CRC deaths in the total (SEER-9) population POP_{total}, and let CRC_{free} be the number of CRC deaths in the initially CRC-free population POP_{free}. Note that ''initially CRC free'' in this context denotes the absence of a CRC diagnosis before the year 2000, that is, before the start of the follow-up but does not preclude the presence of undiagnosed preclinical CRC. When applying CRC mortality in the total population, the biased expected number of CRC deaths in the initially CRC-free population, denoted EXP_{biased}, is

$$EXP_{biased} = (CRC_{total}/POP_{total}) \times POP_{free} = CRC_{total} \times (POP_{free}/POP_{total}).$$

Note that EXP_{biased} is close to CRC_{total}, the total number of CRC deaths, because the proportion of people in the population with a prior CRC diagnosis is small, and the ratio (POP_{free}/POP_{total}) is therefore close to 1. When (correctly) applying CRC mortality in the initially CRC-free population, the true expected number of CRC cases, denoted EXP_{true}, is

$$EXP_{true} = (CRC_{free}/POP_{free}) \times POP_{free} = CRC_{free}.$$

CRC_{total}, CRC_{free}, POP_{total}, and POP_{free} were derived from the incidence, mortality, and population data included in the SEER-9 database. The SEER-9 incidence data includes multiple cancer incidence entries for some patients. The date of the first CRC diagnosis (International Classification of Diseases, 10th revision: C18, C19, or C20) was used as the date of CRC diagnosis in the analysis. CRC_{total} was determined from the SEER-9 incidence data by counting the number of patients from the birth cohort who died of CRC in the respective year (cause of death code: 21040 or 21050). To determine CRC_{free}, the same computation was performed, but patients with a CRC diagnosis before 2000 were excluded. POP_{total} was directly given by the SEER-9 population data. POP_{free} was derived by subtracting for each year the number of patients who had a CRC diagnosis before 2000 and were still living in this year.