



# When gold standards are not so golden: prevalence bias in randomized trials on endoscopic colorectal cancer screening

Hermann Brenner<sup>1,2,3</sup> · Thomas Heisser<sup>1,4</sup> · Rafael Cardoso<sup>1</sup> · Michael Hoffmeister<sup>1</sup>

Received: 20 May 2023 / Accepted: 17 July 2023 / Published online: 2 August 2023  
© The Author(s) 2023

## Abstract

Randomized trials on the effectiveness of screening endoscopy in reducing colorectal cancer (CRC) risk have reported statistically significant, but rather modest reduction of CRC risk by the screening offer. However, risk estimates in these trials included substantial proportions of prevalent CRC cases which were early detected, but could not possibly have been prevented by screening. Thereby, a key principle of randomized prevention trials is violated that only “at risk” persons who do not yet have the disease one aims to prevent should be included in measures of preventive effects. Using recently published data from the Nordic-European Initiative on Colorectal Cancer (NordICC) trial as an example, we illustrate that approaches aimed to account for “prevalence bias” lead to effect estimates that are substantially larger than those reported in the trial and more in line with results from observational studies and real life settings. More rigorous methodological work is needed to develop effective and user-friendly tools to prevent or adjust for prevalence bias in future screening studies.

**Keywords** Bias · Colorectal cancer · Endoscopy · Incidence · Prevalence · Trial

## Acronyms and Abbreviations

CRC	colorectal cancer
ITS	intention-to-screen
NordICC	Nordic-European Initiative on Colorectal Cancer
PP	per-protocol
RCT	randomized controlled trial

## Introduction

Randomized controlled trials (RCTs) are the gold standard approach to establish causality of preventive or therapeutic effects of medical interventions. A commonly employed endpoint for studies of preventive measures is the incidence of the disease one aims to prevent. In such studies, people who already have the disease should be excluded as the intervention can no longer prevent it. Obviously, this important prerequisite has not been fulfilled in the RCT-based estimates of the preventive effects of screening colonoscopy or sigmoidoscopy [1–6], in which a large proportion of presumed incident colorectal cancer (CRC) cases were already present at baseline. The obvious dilemma here is that identifying prevalent cases at baseline would have required a thorough large bowel exam, such as colonoscopy, among all participants, i.e. the very exam whose efficacy in preventing CRC by removing precancerous lesions would be the subject of investigation. In theory, in such a setting, the effectiveness of CRC prevention could still be assessed in a randomized design in which participants with findings of prevalent CRC at colonoscopy would be excluded and the remaining participants would be randomized in such a way that precancerous lesions would be removed in the

✉ Hermann Brenner  
h.brenner@dkfz.de

<sup>1</sup> Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), INF 581, Heidelberg 69120, Germany

<sup>2</sup> Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumour Diseases (NCT), Heidelberg, Germany

<sup>3</sup> German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

<sup>4</sup> Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany

intervention group only but not in the control group. Obviously, such an approach would be unethical and not be a viable option.

However, simply including prevalent cases in both the intervention and the control group and not accounting for the resulting bias in estimates of incidence reduction is not a good solution either, as it may lead to strongly misleading results. We use the recently reported first RCT estimates of screening colonoscopy effects on CRC incidence from the Nordic-European Initiative on Colorectal Cancer (NordICC) trial [1] as an example to illustrate this prevalence bias.

## Methods

In the NordICC trial, 84,585 participants aged 55–64 years from Poland, Norway and Sweden were randomly assigned in a 1:2 ratio to the offer of a single screening colonoscopy or usual care [1]. The offer was used by 42% of participants in the screening group. After 10 years of follow-up, the estimated reduction of CRC risk was 18% in intention-to-screen (ITS) analysis and 31% in per-protocol (PP) analysis, respectively.

The prevalence of cancers at recruitment was only known from participants who actually underwent screening colonoscopy. In the NordICC trial, 62 of 102 cancers (61%) observed within 10 years among 11,843 screened participants were already present and detected at screening colonoscopy, i.e., prevalence at screening was  $62/11,843 = 0.52\%$ . The baseline prevalence of CRC among the unscreened participants is unknown but the overall prevalence in the invited group and the usual-care group should have been approximately equal, given the randomization and the large sample size. It is therefore plausible to assume identical CRC prevalences in the invited group and the usual-care group. However, selective use of the screening offer might have led to some variation in prevalence between users and nonusers of screening within the invited group, and overall within-group prevalence could therefore be higher or lower than the observed 0.52%. To account for this, we assumed a prevalence of 0.52% in both the intervention group and the usual-care group in our base-case exemplary calculations,

and additionally conducted sensitivity analyses assuming a range of theoretically possible and plausible baseline prevalences.

Derivation of the minimum and maximum theoretically possible prevalence is illustrated in Table 1. They were obtained by assuming that all observed CRC cases in the unscreened subgroup of the invited group ( $n = 157$ ) were either truly incident or prevalent cases. While neither of these extreme scenarios is realistic, true prevalence in the invited group must have been somewhere between the resulting prevalence estimates, i.e. 0.22% and 0.78%.

To further narrow down the prevalence estimates to a plausible range, we derived expected prevalence from reported cancer incidence data in 2009–2014, the recruitment period, in the three countries [7, 8], and previously derived estimates of mean sojourn time (MST) of CRC in preclinical phase (which ranged from 3 to 6 years) [9–11]. Let  $I$ ,  $P$ , and  $T$  be the annual incidence, (preclinical) prevalence and annual clinical manifestation rate of preclinical CRC. Then incidence and prevalence can be expressed as.

$$I = P \times T, \text{ and}$$

$$P = I / T = I \times \text{MST},$$

Derivation of a plausible range of prevalences using this approach and taking country-, sex- and age-specific incidence rates and shares of the trial population into account is illustrated in Table 2 and yielded prevalence values between 0.29% and 0.58%.

We derived ranges of theoretically possible and plausible values of cumulative incidence of truly incident cases by subtracting the so-derived prevalences from the cumulative incidence metrics reported in the NordICC trial, which had included the prevalent cases, and we derived ranges of theoretically possible and plausible values of “prevalence-corrected” risk ratios for truly incident cases obtained after these subtractions for both the ITS and the PP analysis.

**Table 1** Reported CRC cases in the NordICC study, minimum and maximum theoretically possible baseline prevalence of CRC in the invited group

Scenario	Screened <i>N</i> = 11,843			Unscreened <i>N</i> = 16,377			Total <i>N</i> = 28,220
	Prevalent	Incident	Total	Prevalent	Incident	Total	Prevalent
Observed	62 (0.52%)	40	102	unknown	unknown	157	unknown
$T_{\min}$	62 (0.52%)	40	102	0	157	157	62 (0.22%)
$T_{\max}$	62 (0.52%)	40	102	157	0	157	219 (0.78%)

CRC, colorectal cancer;  $T_{\min}$  and  $T_{\max}$ , theoretical minimum and maximum of CRC prevalence, respectively

**Table 2** Expected CRC prevalence in the NordICC trial population derived from sex- and age-specific CRC incidence rates during the recruitment period<sup>a</sup> and estimates of mean CRC sojourn time from previous studies [9–11]

Metric		Country		
		Poland	Norway	Sweden
Trial population	Men	27,198	13,217	1,984
	Women	27,330	13,194	1,662
CRC incidence age 55–64 (per 100,000)	Men	108.4	133.3	93.2
	Women	66.6	103.0	73.6
		All countries and both sexes combined		
Weighted average CRC incidence <sup>b</sup>		96.9 per 100,000		
CRC prevalence if MST=3 years		0.29%		
CRC prevalence if MST=4 years		0.39%		
CRC prevalence if MST=5 years		0.48%		
CRC prevalence if MST=6 years		0.58%		

CRC, colorectal cancer; MST, mean sojourn time

<sup>a</sup> For Poland, CRC incidence rates were extracted for the years 2009–2013 from the European Cancer Information System (ECIS) database [7] (ECIS database did not include data for 2014); for Norway and Sweden, CRC incidence rates were extracted for the years 2009–2014 from the Nordcan database [8]

<sup>b</sup> weighted according to the country, sex and age distribution of the NordICC trial study population

## Results

Table 3 shows the reported results of the NordICC trial with the inclusion of cancers that were already present at baseline and the estimated results with the exclusion of prevalent cancers. Reported cumulative incidence of the ITS analysis was 0.98% for the intervention group and 1.20% for the

usual-care group, resulting in a risk ratio of 0.82 which corresponds to a risk reduction by 18%. Reported cumulative incidence of the PP analysis was 0.84% for the intervention group and 1.22% for the usual-care group, resulting in a risk ratio of 0.69 which corresponds to a risk reduction by 31%.

If prevalent cancers were excluded, the cumulative incidences in the intervention group and the usual-care group would decrease to 0.46 and 0.68, respectively, in the base-case ITS analysis, resulting in a risk ratio of 0.68, i.e. the estimated risk reduction would increase from 18 to 32%. The theoretically possible range of prevalence-corrected risk reduction would be from 22 to 52%, and a plausible range of risk reduction derived from cancer-registry data would be from 25 to 35%.

In the PP analysis, base-case exclusion of 0.52% prevalent cases would lead to cumulative incidences of 0.32% and 0.70% in the intervention group and the usual-care group, respectively, resulting in a risk ratio of 0.46, i.e. the estimated risk reduction would increase from 31 to 54%. The theoretically possible range of prevalence-corrected risk reduction would be from 38 to 86%, and a plausible range of risk reduction derived from cancer-registry data would be from 41 to 59%.

## Discussion

The exemplary calculations based on published results from the NordICC trial provided in this article suggest a much stronger preventive effect of screening colonoscopy than reflected in the reported RCT results which also included

**Table 3** Reported estimates of CRC risk reduction with inclusion of prevalent cases and estimated risk reduction after excluding prevalent cases in the NordICC trial on screening colonoscopy

Analysis	Scenario	Handling of prevalent cases (assumed prevalence)	Cumulative Incidence [%]		Risk ratio	Risk reduction
			Intervention Group	Usual-Care Group		
ITS	<b>Reported</b>	<b>Included<sup>a</sup></b>	<b>0.98</b>	<b>1.20</b>	<b>0.82</b>	<b>18%</b>
	T <sub>Min</sub>	Excluded (0.22%)	0.76	0.98	0.78	22%
	P <sub>Low</sub>	Excluded (0.29%)	0.69	0.88	0.75	25%
	<b>Base-case</b>	<b>Excluded (0.52%)<sup>b</sup></b>	<b>0.46</b>	<b>0.68</b>	<b>0.68</b>	<b>32%</b>
	P <sub>High</sub>	Excluded (0.58%)	0.40	0.62	0.65	35%
	T <sub>Max</sub>	Excluded (0.78%)	0.20	0.42	0.48	52%
	<b>Reported</b>	<b>Included<sup>a</sup></b>	<b>0.84</b>	<b>1.22</b>	<b>0.69</b>	<b>31%</b>
PP	T <sub>Min</sub>	Excluded (0.22%)	0.62	1.00	0.62	38%
	P <sub>Low</sub>	Excluded (0.29%)	0.55	0.93	0.59	41%
	<b>Base-case</b>	<b>Excluded (0.52%)<sup>b</sup></b>	<b>0.32</b>	<b>0.70</b>	<b>0.46</b>	<b>54%</b>
	P <sub>High</sub>	Excluded (0.58%)	0.26	0.64	0.41	59%
	T <sub>Max</sub>	Excluded (0.78%)	0.06	0.44	0.14	86%

CRC, colorectal cancer; ITS, intention-to-screen; P<sub>Low</sub> and P<sub>High</sub>, lower and upper end of plausible range of CRC prevalence based on a weighted average of incidence rates in the contributing countries and range of mean CRC sojourn time estimates; PP, per-protocol, T<sub>min</sub> and T<sub>max</sub>, theoretical minimum and maximum of CRC prevalence, respectively

<sup>a</sup> as reported by Bretthauer et al. [1]

<sup>b</sup> base-case analysis assuming equal CRC prevalence in screened and unscreened participants

prevalent cases that could no longer have been prevented by screening colonoscopy. Although these non-preventable CRC cases diminished reported screening effects, their earlier detection is an additional asset of screening as it enhances chances of cure. The presented patterns therefore imply the need for a more differentiated view on the evidence of the screening colonoscopy effects and may help to resolve some of the ongoing controversy regarding interpretation of the trial results [12] and some of the apparent discrepancy from findings from observational studies and real-life settings.

In the United States, CRC incidence has declined by approximately 50% in the screening age range in the past 30–40 years [13], despite adverse trends in the prevalence of key CRC risk factors such as obesity and increasing incidence at younger ages [14]. The most plausible explanation for this dramatic decline is the widespread use of colonoscopy for both screening and diagnostic purposes. With preventive effects of screening endoscopy in the order of magnitude of the reported NordICC trial results these real-life changes could not have been achieved. Our illustration may help to explain much of the apparent discrepancy between the reported RCT results and these real-life data.

Another factor to be considered in the interpretation of the RCT results is that diagnostic colonoscopies, which have similar preventive potential as screening endoscopies through detection and removal of CRC precursors, have meanwhile become more common also in European countries in which the NordICC trial was conducted [15]. Not accounting for diagnostic colonoscopies or, in general, colonoscopies outside the screening trials conducted during follow-up may further have attenuated the reported RCT effect estimates [16]. Whereas the relative importance of prevalence bias would be expected to gradually decrease with prolonged follow-up, such as 15- or 20-year follow-up of the trial cohort, the role of other biases such as contamination would be expected to further increase over time.

Although our numerical example focused on the first and so far only RCT results on the long-term impact of screening colonoscopy on CRC risk, the illustrated prevalence bias is expected to have similarly affected RCTs on CRC screening by other modalities, such as flexible sigmoidoscopy or fecal occult blood tests [2–6, 17, 18], whose preventive effects may likewise have been substantially underestimated. Although RCTs are less prone to various other biases than observational studies, the randomized design does not protect from prevalence bias which may be substantial in screening studies as illustrated in our article.

In summary, the preventive effects of screening endoscopy are likely to be stronger than suggested by the reported RCT results. Accounting for the prevalence bias leads to effect estimates that are much more in line with results from

observational studies and real-life settings [13, 14, 19, 20]. More rigorous methodological work is needed to develop effective and user-friendly tools to prevent or adjust for prevalence bias in future screening studies. The apparently low RCT-based effect estimates should not unduly discourage use of CRC screening, the likely most effective way to cope with the ongoing global CRC epidemic, which is expected to lead to an increase in case numbers from approximately 1.9 million in 2020 to 3.2 million in 2040 [21]. In the contrary, efforts of prevention need to be enhanced, and major efforts are needed to better disentangle true prevention of CRC occurrence, early detection of already prevalent CRC, and their combined contribution to lowering the CRC burden in CRC screening studies.

**Authors Contribution** The analyses were performed and the first draft of the manuscript was written by Hermann Brenner. All authors commented on previous versions of the manuscript and read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work was supported in part by grants from the German Federal Ministry of Education and Research (grant no. 01KD2104A) and the German Cancer Aid (grant no. 70114735).

## Declarations

**Competing Interests** The authors have no relevant financial or non-financial interests to disclose.

**Ethics approval** Not applicable. This study included only calculations based on previously published summary statistics of a multicenter randomized controlled trial that had obtained ethical approval by the ethics committees at all participating centers, the Swedish National Council on Medical Ethics, and the Health Council of the Netherlands, as outlined in detail in reference 1 from which the summary statistics were extracted.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Bretthauer M, Loberg M, Wieszczy P, et al. Effect of Colonoscopy Screening on Risks of Colorectal Cancer and related death. *N Engl J Med*. 2022;387(17):1547–56.

2. Holme Ø, Løberg M, Kalager M, Bretthauer M, Hernán MA, Aas E, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. *JAMA*. 2014;312(6):606–15.
3. Atkin W, Wooldrage K, Parkin DM, Kralj-Hans I, MacRae E, Shah U, et al. Long term effects of once-only flexible sigmoidoscopy screening after 17 years of follow-up: The UK Flexible Sigmoidoscopy Screening randomised controlled trial. *Lancet*. 2017;389(10076):1299–311.
4. Miller EA, Pinsky PF, Schoen RE, Prorok PC, Church TR. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: long-term follow-up of the randomised US PLCO cancer screening trial. *Lancet Gastroenterol Hepatol*. 2019;4(2):101–10.
5. Senore C, Riggi E, Armaroli P, Bonelli L, Sciallero S, Zappa M, et al. Long-term follow-up of the Italian flexible sigmoidoscopy screening trial. *Ann Intern Med*. 2022;175(1):36–45.
6. Juul FE, Cross AJ, Schoen RE, Senore C, Pinsky P, Miller E, et al. 15-Year benefits of Sigmoidoscopy Screening on Colorectal Cancer incidence and mortality: a pooled analysis of Randomized trials. *Ann Intern Med*. 2022;175(11):1525–33.
7. Larønningen S, Ferlay J, Beydagan H, Bray F, Engholm G, Ervik M et al. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 9.2 (23.06.2022). Association of the Nordic Cancer Registries. Cancer Registry of Norway. Available from: <https://nordcan.iarc.fr/>, accessed 26 June 2023.
8. Randi G, Zaro F, Carvalho RN, Martos C, Dimitrova N, Dyba T et al. The European Cancer Information System (ECIS) web application; JRC113106, <https://ecis.jrc.ec.europa.eu/index.php>, accessed 26 June 2023.
9. Launoy G, Smith TC, Duffy SW, Bouvier V. Colorectal cancer mass-screening: estimation of faecal occult blood test sensitivity, taking into account cancer mean sojourn time. *Int J Cancer*. 1997;73(2):220–4.
10. Prevost TC, Launoy G, Duffy SW, Chen HH. Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches. *Am J Epidemiol*. 1998;148(6):609–19.
11. Brenner H, Altenhofen L, Katalinic A, Lansdorp-Vogelaar I, Hoffmeister M. Sojourn time of preclinical colorectal cancer by sex and age: estimates from the German national screening colonoscopy database. *Am J Epidemiol*. 2011;174(10):1140–6.
12. Knopf KB, Gyawali B. Intention to probe into the colonoscopy trial: is it the procedure or the trial that failed? *EClinicalMedicine*. 2022;55:101793.
13. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin*. 2017;67(3):177–93.
14. Shah RR, Millien VO, da Costa WL Jr, Oluyomi AO, Gould Suarez M, Thrift AP. Trends in the incidence of early-onset colorectal cancer in all 50 United States from 2001 through 2017. *Cancer*. 2022;128(2):299–310.
15. Cardoso R, Guo F, Heisser T, Hoffmeister M, Brenner H. Utilisation of Colorectal Cancer Screening tests in European countries by type of screening offer: results from the European Health Interview Survey. *Cancers*. 2020;12(6):1409.
16. Brenner H, Stock C, Hoffmeister M. In the era of widespread endoscopy use, randomized trials may strongly underestimate the effects of colorectal cancer screening. *J Clin Epidemiol*. 2013;66(10):1144–50.
17. Lauby-Secretan B, Vilahur N, Bianchini F, Guha N, Straif K, International Agency for Research on Cancer Handbook Working Group. The IARC Perspective on Colorectal Cancer Screening. *N Engl J Med*. 2018;378(18):1734–40.
18. Knudsen AB, Rutter CM, Peterse EFP, Lietz AP, Seguin CL, Meester RGS, et al. Colorectal Cancer screening: an updated modeling study for the US Preventive Services Task Force. *JAMA*. 2021;325(19):1998–2011.
19. Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ*. 2014;348:g2467.
20. Cardoso R, Guo F, Heisser T, Hackl M, Ihle P, De Schutter H, et al. Colorectal cancer incidence, mortality, and stage distribution in European countries in the colorectal cancer screening era: an international population-based study. *Lancet Oncol*. 2021;22(7):1002–13.
21. Morgan E, Arnold M, Gini A, Lorenzoni V, Cabasag CJ, Laversanne M, et al. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut*. 2023;72(2):338–44.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.