

Residual confounding after adjustment for age: a minor issue in breast cancer screening effectiveness

Guido van Schoor · Ellen Paap · Mireille J. M. Broeders · André L. M. Verbeek

Received: 10 February 2011 / Accepted: 8 April 2011 / Published online: 26 April 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Residual confounding, after adjustment for age, is the major criticism of observational studies on breast cancer screening effectiveness. We developed realistic scenarios for the prevalence and strength of risk factors on screened and not screened groups, and explored the impact of residual confounding bias. Our results demonstrate that residual confounding bias is a minor issue in screening programme evaluations.

Keywords Bias · Breast cancer screening · (Residual) confounding · Effectiveness

Introduction

Breast cancer screening programmes are now an established part of the health care service of many countries [1]. The continuous evaluation of this practice is based on observational studies, leading to the possibility of confounding and self-selection bias.

To assess the effect of screening, breast cancer mortality in both screened and not screened women has to be compared; this can be looked upon as the relative risk (RR, or rate ratio) of breast cancer mortality. Confounding bias of the RR occurs when the prevalence of a risk factor (or set of risk factors) for

breast cancer death is imbalanced across the compared groups. To adjust for the confounding effect, the prevalence of the risk factor(s) has to become similar in both groups.

Usually age is the only risk factor measured when evaluating population-based breast cancer screening programmes, because information on date of birth and date of invitation of women is mostly available. Therefore, after age, residual confounding bias in the screening—mortality relation remains the major criticism of observational studies. This term covers both within-stratum confounding, for example too-broad age categories, and confounding due to unmeasured variables [2]. Self-selection bias can be regarded as a special form of residual confounding because participation may induce an imbalance in the risk factors for breast cancer death.

Having accounted for age, we clarified the influence of adjustment for residual confounding on the rate ratio of breast cancer death. We compared the mortality rate in the screened (M_s) with not screened women (M_{ns}). This results in an ‘apparent’ screening—mortality association (RR_a) that is seemingly real, but not necessarily so because of possible residual confounding bias. This effect of screening, RR_a , can be unravelled in the ‘specific’ screening effect RR_s , and a ‘non-specific’ effect of the potential confounding factor(s) C , which is reflected in the following formula.

$$\begin{aligned} RR_a &= M_s / M_{ns} \\ &= RR_s * C \\ &= RR_s * [p_1 RR_c + (1 - p_1)] / [p_2 RR_c + (1 - p_2)] \end{aligned}$$

The quantity C thus represents the effect of the potential confounder(s) among screened and the not screened women. The influence of C depends on the relative risk of breast cancer death RR_c , the proportion p_1 of screened women with the confounder present, and the proportion p_2

G. van Schoor (✉) · E. Paap · M. J. M. Broeders ·
A. L. M. Verbeek
Department of Epidemiology, Biostatistics and HTA, Radboud
University Nijmegen Medical Centre, PO Box 9101, Internal
Postal Code 133, 6500 HB Nijmegen, The Netherlands
e-mail: g.vanschoor@ebh.umcn.nl

M. J. M. Broeders · A. L. M. Verbeek
National Expert and Training Centre for Breast Cancer
Screening, Nijmegen, The Netherlands

of not screened women having the confounder. The formula is based on previous work by Cornfield and colleagues [3], Schlesselman [4] and Greenland [5].

Suppose, as shown in Fig. 1, that the apparent RR_a is 0.50, and a risk factor producing a twofold increase in risk of breast cancer death (RR_c) is present among 20% (p_1) of the screened group and 50% (p_2) of the not screened group. Then, the non-specific part of the apparent screening effect is $0.20 \cdot 2 + 0.80 \cdot 1 = 1.20$ among the screened women, and $0.50 \cdot 2 + 0.50 \cdot 1 = 1.50$ among the not screened women. The ratio of these non-specific effects is $1.2/1.5 = 0.80$, which is the influence of confounding (C) among the screened and not screened groups. Accordingly, the specific RR_s will become $0.50/0.80 = 0.63$.

In the above calculation we used the cohort approach and the risk ratio (or rate ratio) as a measure of effect. However, this same method can be applied when the odds ratio (OR) is the effect measure. The case–control design has been increasingly used for the evaluation of screening programmes [6–12]. In the case–control evaluation, the odds of having been screened versus not screened in the case group of breast cancer deaths is compared to the same odds in the control group of invited women from whom the cases originate. As such, the OR is the mortality in screened versus not screened women.

Example based on the Nijmegen Breast Cancer Screening Programme

As an example, we report on a case–control study conducted within the Nijmegen breast cancer screening

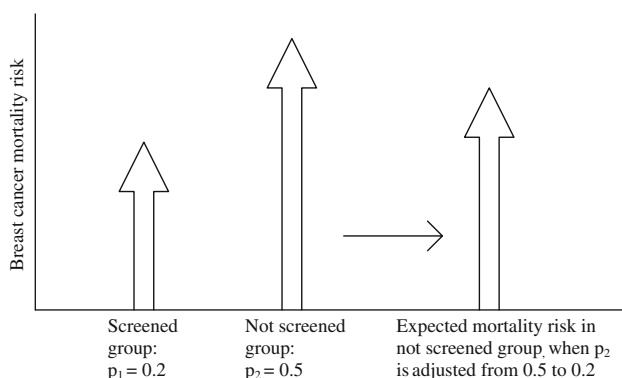


Fig. 1 A heuristic device to address residual confounding in the mortality effect of breast cancer screening. Both *arrows on the left* indicate the observed breast cancer mortality risk in the screened and not screened group, suggesting $RR_a = 0.50$. We assume that a confounder with a twofold relative risk on breast cancer death (RR_c), is present among 20% (p_1) of the women in the screened group and among 50% (p_2) in the not screened group. The *arrow on the right* indicates the expected breast cancer mortality risk in the not screened population when the presence of the risk factor in that group is adjusted from 50% to 20%. The adjusted RR_s becomes 0.63 (also demonstrated in Fig. 2)

programme which started in 1975. After adjustment for age, we found that the breast cancer mortality rate in the screened group was 65% lower than that of the not screened group: OR = 0.35 and 95% Confidence Interval (CI) = 0.19–0.64 [12]. What role could residual confounding have played in this finding?

Dense mammographic breast pattern, for which a high relative risk of 6 has been reported, is a likely candidate for being treated as a confounding factor [13]. Despite its strength, this factor is not common in postmenopausal women. Nevertheless, suppose its prevalence in all screened women is 5% ($p_1 = 0.05$) in contrast to a supposed 20% ($p_2 = 0.20$) prevalence in the not screened women, then, according to the formula, the apparent OR of 0.35 would be adjusted to an OR of 0.56 (see also Fig. 2, left upper diagram).

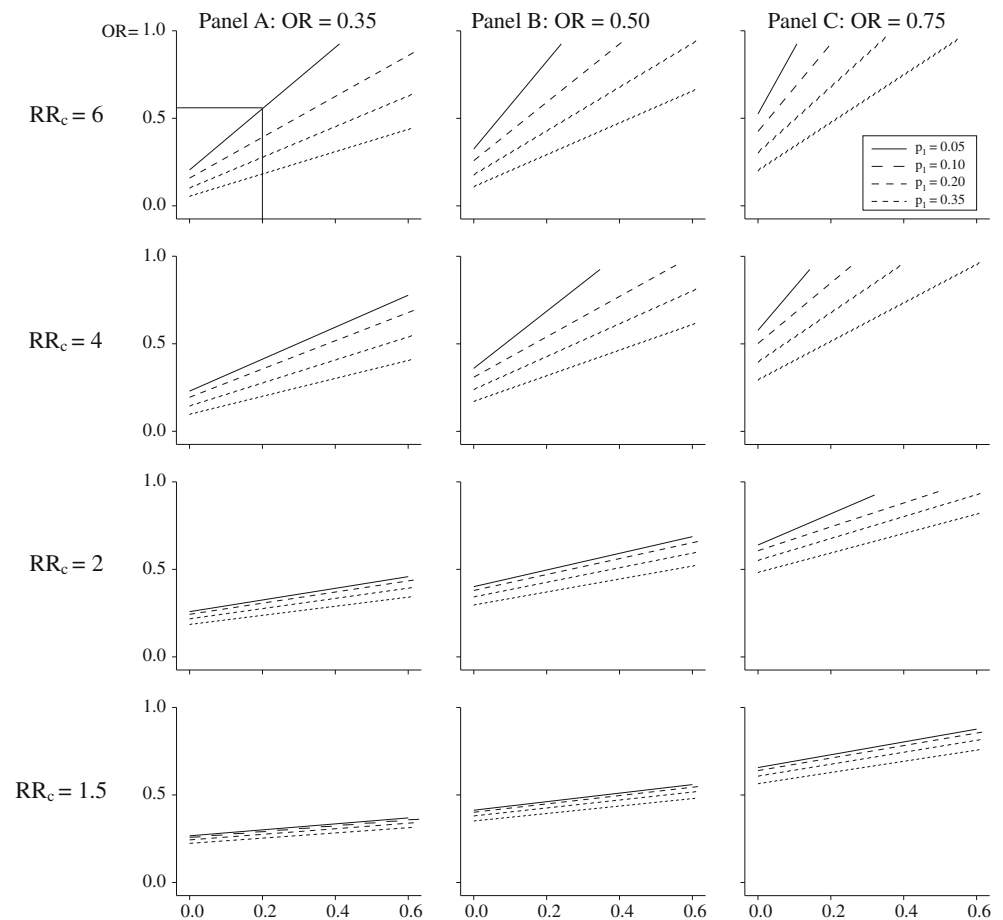
Other risk factors for breast cancer like obesity, socio-economic status, nulliparity, late age at menopause, early age at menarche, and family history show a 1.5 to fourfold relative risk of breast cancer at most [14]. We assume that the risk magnitude of the factors applies to the incidence and mortality alike. Fig. 2 illustrates the impact these risk factors may have as confounders. Panel A shows the baseline situation of an age-adjusted screening–mortality OR of 0.35; Panel B is for OR = 0.50 and Panel C for OR = 0.75. The expected values of the ORs in order of decreasing magnitude are displayed on the Y-axis in each figure: after adjustment for dense breast pattern $RR_c = 6$; late age at menopause $RR_c = 4$; nulliparity $RR_c = 2$; and serious overweight $RR_c = 1.5$. The X-axis shows the proportion (p_2) of the not screened population with the confounding factor. In each figure, the lines present the OR adjusted for the confounding factor with p_2 ranging from 0 to 0.6, and four different situations of the proportion (p_1) confounder in the screened group: the upper line is for a $p_1 = 0.05$, then $p_1 = 0.10$, $p_1 = 0.20$ and the lowest line for a $p_1 = 0.35$. In practice, the deviations between apparent and adjusted ORs are minor.

Discussion

Previous screening programme evaluations have qualitatively discussed the magnitude of residual confounding bias on their effectiveness estimate [6, 9, 10, 12] or estimated the amount of bias due to self-selection [7, 8, 11]. We present an educated and pragmatic method to quantify the potential impact of residual confounding, and to de-bias the comparison of screened with unscreened groups, a method originally introduced by Cornfield et al. [3]. Our results demonstrate that residual confounding has a minor influence on the observed screening effect.

Closely related to residual confounding is self-selection bias and healthy screenee bias. The difference between

Fig. 2 Diagrams of the adjustment for residual confounding in the effectiveness measurement of breast cancer service screening. Panel A shows the baseline situation of an age-adjusted screening—mortality OR = 0.35; panel B is for OR = 0.50 and panel C for OR = 0.75. From top to bottom, the figures represent the adjusted ORs for confounding factors with $RR_c = 6, 4, 2$ and 1.5 , respectively. The X-axis displays the proportion (p_2) of the not screened population with the confounding factor. The lines displayed in the figures present the adjusted OR for the confounding factor for p_2 ranging from 0.0 to 0.6, and four different points of departure for p_1 of the screened population (upper line at $p_1 = 0.05$, then $p_1 = 0.10$, $p_1 = 0.20$ and the lowest line $p_1 = 0.35$). The Y-axis in each figure depicts the expected ORs adjusted for residual confounding



these three biases is subtle; the nuances seem to lie in the clarification of definable confounding factors or a combination of indefinable confounding factors. Self-selection into screening may result in an imbalance of a combination of indefinable risk factors, causing a different background risk of dying from breast cancer in screened versus not screened women [15]. Healthy screenee bias may occur because some women in the not screened group, although invited for screening, may already have been diagnosed with cancer, while screened women were not diagnosed with breast cancer at the time of participation [16]. Both biases can be regarded as a form of residual confounding [17] since participation in screening may be correlated with the baseline risk of dying from breast cancer.

An estimate of the amount of self-selection can be obtained by calculating the ratio of the breast cancer deaths among not invited and not screened women [18]. This calculation is not possible in a steady state situation of population based screening since there is no uninvited group. By using the implementation period of screening, we [11] quantified a 0.84 lower background risk in not screened women compared with not yet invited women. A similar Italian study found a 1.11 higher risk in the not screened group [8]. Duffy et al. [18] proposed a factor

based on data from the Swedish and Canadian screening trials, showing a 1.36 higher risk for not screened women. With these factors, the difference in background risk between not screened and screened women can be calculated by taking the percentage uptake in a programme into account [18]. For instance, if we use Duffy's factor of 1.36 and if the screening uptake is 80%, which is in accordance with most European programmes, not screened women have a 1.42 higher background risk compared with screened women. This factor actually represents C in our formula, it is the difference in background risk $p_1 = 0$ and $p_2 = 1$. In this scenario an apparent OR of 0.35 would be adjusted to 0.51. However, using our factor of 0.84, not screened women have a 0.80 lower background risk compared with screened women. In our scenario an apparent OR of 0.35 would be adjusted to 0.28.

In Cornfield's original paper [3], he stated that a confounding factor completely explains an 'apparent' effect when the effect of confounding in the comparing groups equals the 'apparent' effect, then $RR_a = C$, and $RR_s = 1$.

In our example we applied this method to adjust ORs for combinations of p_2 between 0 and 0.6, and values of $p_1 = 0.05, 0.10, 0.20$ and 0.35 . These values were chosen based on the expected prevalence of the risk factors in the

female population, i.e. 5% for mammographic density, 10% for late age at menopause, 20% for nulliparity, and 35% for serious overweight. As, we aimed to challenge the age-adjusted screening effect, we developed scenarios where p_1 was smaller than p_2 .

Our calculation does not account for random error or uncertainties about the relation of risk factors and breast cancer. It is possible to correct for this by using more complex techniques based on a Monte Carlo and a Bayesian approach [19]. However, the aim of this study was to present a heuristic device to address residual confounding.

In conclusion, in studies on breast cancer screening the mortality reduction ranges from 38 to 70% [6–12]. As we have shown, residual confounding does not have a great effect on these estimates of screening effectiveness. After having addressed for age, future breast cancer screening programme evaluations can ignore residual confounding.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Shapiro S, Coleman EA, Broeders MJM, Codd MB, de Koning HJ, Fracheboud J, et al. Breast cancer screening programmes in 22 countries: Current policies, administration and guidelines. *Int J Epidemiol.* 1998;27:735–42.
- Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
- Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst.* 1959; 22:173–203.
- Schlesselman JJ. Assessing effects of confounding variables. *Am J Epidemiol.* 1978;108:3–8.
- Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol.* 1996;25:1107–16.
- Fielder HM, Warwick J, Brook D, Gower-Thomas K, Cuzick J, Monypenny I, et al. A case-control study to estimate the impact on breast cancer death of the breast screening programme in Wales. *J Med Screen.* 2004;11:194–8.
- Gabe R, Tryggvadottir L, Sigfusson BF, Olafsdottir GH, Sigurdsson K, Duffy SW. A case-control study to estimate the impact of the Icelandic population-based mammography screening program on breast cancer death. *Acta Radiol.* 2007;48: 948–55.
- Puliti D, Miccinesi G, Collina N, De Lisi V, Federico M, Ferretti S, et al. Effectiveness of service screening: a case-control study to assess breast cancer mortality reduction. *Br J Cancer.* 2008;99: 423–7.
- Roder D, Houssami N, Farshid G, Gill G, Luke C, Downey P, et al. Population screening and intensity of screening are associated with reduced breast cancer mortality: evidence of efficacy of mammography screening in Australia. *Breast Cancer Res Treat.* 2008;108:409–16.
- Allgood PC, Warwick J, Warren RM, Day NE, Duffy SW. A case-control study of the impact of the East Anglian breast screening programme on breast cancer mortality. *Br J Cancer.* 2008;98:206–9.
- Paap E, Holland R, den Heeten GJ, van Schoor G, Botterweck AA, Verbeek ALM, et al. A remarkable reduction of breast cancer deaths in screened versus unscreened women: a case-referent study. *Cancer Causes Control.* 2010;21:1569–73.
- van Schoor G, Moss SM, Otten JDM, Donders R, Paap E, den Heeten GJ, et al. Increasingly strong reduction in breast cancer mortality due to screening. *Br J Cancer.* 2011;104:910–4.
- Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med.* 2007;356:227–36.
- Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. *J Natl Cancer Inst.* 2010;102:680–91.
- Connor RJ, Prorok PC, Weed DL. The case-control design and the assessment of the efficacy of cancer screening. *J Clin Epidemiol.* 1991;44:1215–21.
- Weiss NS, Rossing MA. Healthy screening bias in epidemiologic studies of cancer incidence. *Epidemiology.* 1996;7:319–22.
- Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology.* 2004;15:615–25.
- Duffy SW, Cuzick J, Tabár L, Vitak B, Chen THH, Yen MF, et al. Correcting for non-compliance bias in case-control studies to evaluate cancer screening programmes. *J R Stat Soc Ser C Appl Stat.* 2002;51(2):235–43.
- Steenland K, Greenland S. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. *Am J Epidemiol.* 2004;160: 384–92.