Estimation of Overdiagnosis in Mammographic Screening: a Critical Assessment

doi: https://doi.org/10.32635/2176-9745.RBC.2021v67n2.1281

Estimação do Sobrediagnóstico no Rastreamento Mamográfico: uma Avaliação Crítica

Estimación del Sobrediagnóstico en el Cribado Mamográfico: una Evaluación Crítica

Arn Migowski[1]; Paulo Nadanovsky[2]; Cid Manso de Mello Vianna[3]

INTRODUCTION

Because of its magnitude and consequences for health, overdiagnosis is considered the most important harm associated with mammographic screening[1]. However, its definition is still controversial[2-4] and no consensus or consistency exists about the better form of calculation[3,5]. These discrepancies have been used to disqualify the estimates of overdiagnosis in general without, however, discussing in depth which are the most reliable[1,6].

Overdiagnosis is the diagnosis of cases of cancer which would never manifest clinically if not detected in screening[2,7] or that would not onset clinically because of competing causes of death[7,8]. This last definition gains relevance in women with comorbidities or elderly, but it is arguable in very long follow-up periods[5,9].

Overdiagnosis should include either *in situ* or invasive cancer cases[10]. The spontaneous regression of ductal carcinoma *in situ* (DCIS) is relatively common and there are situations documented of the same phenomenon in invasive cancer, which correspond to more than half of the cases of overdiagnosis[10-12].

The aim of the present article is to discuss the implications of the use of different study designs and calculation methods to estimate overdiagnosis of breast cancer.

DEVELOPMENT

Although the existence of overdiagnosis in mammography screening is practically consensual, its magnitude is still debatable[6]. One of the reasons of variation is the study design since it can be randomized clinical trials (RCT), observational studies or modeling.

The estimates of model-based lead-time (LT) varied from one to seven years, suggesting it would be necessary long follow-up to estimate overdiagnosis. However, these modellings overestimate LT by ignoring the overdiagnosis and causes of competing deaths, hypothesizing that all tumors progress, varying the velocity of growth[13]. Nevertheless, the actual LT would be one year only, inferred from the difference of diameters of tumors of the intervention groups of screening RCT[14]. Estimation errors about the duration of the potentially preclinical screen-detectable phase (sojourn time) are causes of variation of model-based estimates[15].

Another form of infer the actual LT is the analysis of the variations of the cancer incidence in screening because it is predicted that there is a drop of interval cancers compared to the basal incidences by the early detection of many cases. Had LT been longer as estimated in these models, this reduction would last for many years, but it returns to pre-screening levels in only two to three years after the last round[14]. The inclusion of overdiagnosis cases in LT estimates increases them artificially in until nine years depending on the age-range.

If on one side the inclusion of cancers detected years after the end of RCT dilutes the overdiagnosis, in the other, consider only the duration of RCT may overestimate it, defining the early detected cancer cases as overdiagnosis. The maximum time to include cases must be defined, estimating the LT and adding this time to the post-intervention period of RCT. A three-to-five years follow-up period after the end of RCT would be more than enough to estimate the overdiagnosis, controlling by LT and avoiding the dilution by the contamination of the control group (CG) along the years[14].

One of the evidences of the existence of overdiagnosis in screening was the exponential growth of cancer detected after its implementation without later compensation with reduction of the cases diagnosed in advanced stages[11].

Well conducted ecological and cohort studies are deemed as good options for quantification of overdiagnosis[16], producing estimates between 40% and

[1]National Institute of Cardiology (INC); National Cancer Institute José Alencar Gomes da Silva (INCA). Rio de Janeiro (RJ), Brazil. E-mail: arnmigowski@yahoo.com.br. Orcid iD: https://orcid.org/0000-0002-4861-2319
[2]Instituto de Medicina Social da Universidade do Estado do Rio de Janeiro (IMS/Uerj); Escola Nacional de Saúde Pública da Fundação Oswaldo Cruz (Ensp/Fiocruz). Rio de Janeiro (RJ), Brazil. E-mail: nadanovsky@ims.uerj.br. Orcid iD: https://orcid.org/0000-0003-3345-9873
[3]IMS/Uerj. Rio de Janeiro (RJ), Brazil. E-mail: cdvianna@gmail.com. Orcid iD: https://orcid.org/0000-0003-0252-1144
Corresponding author: Arn Migowski. Rua das Laranjeiras, 374, 5º andar – Laranjeiras. Rio de Janeiro (RJ), Brazil. CEP 22240-002. E-mail: arnmigowski@yahoo.com.br

60% although part can be explained by the increase of the incidence associated with the popularization of the hormone replacement therapy in the years 1990 and greater sensitivity of digital mammography when compared with those used in RCT[17,18].

An ecological study estimated that 31% of all the breast cancer cases diagnosed in the United States during three decades in women with 40 years of age or more would be overdiagnosis[11]. This proportion could have increased to 45% with change of denominator if accounted only the cases detected in screening[19].

In the Dutch screening program, the estimate after deducting the clinical LT is 14 cases of overdiagnosis for each death by breast cancer avoided in women between 50 and 74 years of age corresponding to 32% among cancers detected for the women invited for screening and 52% of the cancers detected in screening, results attributed in part to the extension of the screening beyond 69 years old, discounted the cases diagnosed as *in situ* or stage I which evolved to II to IV[20].

In Denmark, screening was introduced only in part of the country for decades creating a natural CG. In this country, the estimate of overdiagnosis tends to be lower, because of the age range screened and low adherence rates, recall and detection of DCIS, being 33% after adjustment reducing the excess of cases in the target-population (50-69) by the drop of the incidence occurred with women older than 70 years[21].

The existence of contemporaneous CG with random allocation of the intervention is an advantage of RCT in relation to observational studies in overdiagnosis estimates since it provides a baseline for the comparison of the excess-incidence[8]. The absence of CG makes necessary to predict which would be the incidence in the baseline using, for instance, historical series of non-screened age groups. The estimates tend to be similar utilizing data of RCT whether age range, periodicity, adherence to the treatment, contamination of the CG and calculation method are similar[7,22].

Shorter screening intervals tend to increase the overdiagnosis[7]. Low adherence to screening in the intervention group tends to underestimate it, the UK Age Trial being the most extreme example with only 68% of adherence[22].

The contamination of the CG can occur both during the study and in the follow-up after its ending, underestimating the overdiagnosis. The inclusion of all the cancers detected in long follow-ups can dilute the estimates in half or even more[15].

In the Cochrane systematic review, the estimates of overdiagnosis were 29% including all the RCT and of 33%, counting only the cases detected before screening was offered to the CG[19]. All included the 40-49 age range which presents more overdiagnosis. In only two RCT, mammographic screening was not offered to CG after the end: Malmö and the Canadian National Breast Screening Study (CNBSS), although they have suffered contamination during the intervention period[19]. In Malmö, only to a subgroup of the CG between 55 and 69 years screening was not offered after the period of intervention[2].

Another source of discrepancy between the estimates of overdiagnosis is the method of calculation, especially regarding the differences of denominator. The most used options of denominator are described in Chart 1. The form of calculation can create discrepancies easily corrigible which are not inconsistencies as many claim[6]. This is clear comparing the estimates of the Malmö and CNBSS trials using the same formula[23,24].

An independent panel proposed the use of denominators 2 and 4 (Chart 1), both including DCIS[1]. There are also others who advocate the use of all cancers of the screened group in the denominator because the RCT had different screening intervals which influences the incidence of interval cancers[16].

RCT are the most reliable source to estimate overdiagnosis[25]. However, they tend to underestimate the magnitude in the current clinical practice for having utilized greater screening intervals, allowing several forms of screening in CG and used less sensitive mammographies[5,6].

Apparently in the UK Age Trial, there is little overdiagnosis if considered the follow up period after the end of the trial where mammography screening was actively offered to CG[26]. However, based in the incidence of cancer in CG in the period of intervention, it is possible to infer that the proportion of overdiagnosis was 35%[27].

Malmö's authors estimated overdiagnosis in 10%, 15 years after its end in women in the age range of 55-69 years[28]. This figure is underestimated by the contamination of 24% during the study, reducing the contrast of the incidence among the groups and the estimate would rise to 20% with the same calculation[24] if maintained in these 15 years after the end. Analyzing only the cases detected in the experimental group during the study, instead of all in these 15 years, the proportion would go from 10% to 15%. In this calculation, the original numerator was kept (number 1 of the Chart), but the denominator started to include less cases. This estimate still underestimates the overdiagnosis because it counts all the cancers detected, including interval cancers which have worse prognosis. The denominator would reduce and the overdiagnosis would increase to 24% if only cancers detected in screening were included in the denominator[29].

Chart 1. Different numerators and denominators used in the calculation of the ratio of overdiagnosis in clinical trials of mammography screening

Denominator[a]
1. Mammography screening detected cancers during the intervention period
2. All cancers detected during the intervention period (including interval cancers)
3. All cancers detected during the intervention period plus an additional follow-up period
4. All cancers detected in the lifetime of the women from the date the screening began in the experimental group
5. Mammography screening-detected cancers during the intervention period plus an additional follow-up period
6. Mammography screening-detected cancers during the intervention period plus an additional follow-up period (non-palpable)
Numerator
1. Difference among the cancer cases in the intervention and control groups during the study
2. Difference among the cancer cases in the intervention and control groups during follow-up period

(a) all breast cancer cases in the experimental group of women invited for screening during the study.

Considering the contamination of the CG, this estimate would increase even more. A percent of 37% would be reached if this same form of calculation was used in the CNBSS II.

CNBSS is recognized as a reliable source to estimate overdiagnosis[30,31]. The calculation used was the ratio between the numerator 2 and denominator 1 of the Chart[23]. Because CNBSS used screening with breast clinical examination, cancers detected exclusively by mammography were identified. Its estimates of overdiagnosis can be underestimated by contamination of the CG which occurred in some provinces due to the implantation of screening programs, reducing the estimates in five-years follow-up[19]. Estimates of overdiagnosis soon after the end of the study tend to overestimate it for not considering the LT of approximately one year. The three-year estimates after the study end are the most reliable avoiding the contamination in longer follow-up periods (Table 1).

CONCLUSION

The best available evidences indicate that 25% to 30% of the cancers detected in screening between 50 and 69 years of age are overdiagnosed, and over 40% between 40 and 49 years of age. Several factors explain the discrepancies encountered in the literature such as adherence to screening, screening intervals, age-range, contamination of the CG, estimated LT and differences in the denominator utilized. The estimates of overdiagnosis of the main studies become consistent when these factors are harmonized.

CONTRIBUTIONS

Arn Migowski contributed for study conception and design; collection, analysis and interpretation of the data, wording and critical review with intellectual contribution. Paulo Nadanovsky and Cid Manso de Mello Vianna contributed for the interpretation of the data and critical review with intellectual contribution. All the authors approved the final version to be published.

Table 1. Estimates of overdiagnosis in clinical trials of Malmö I, CNBSS and UK Age Trial, including invasive and *in situ*[a] cancers

	40-49 years CNBSS I[23]	50-59 years CNBSS II[23]	55-69 years Malmö I[24]	40-49 years UK AgeTrial[28]
During the study	37%	37%	-	
1 year after the study	40%	38%	-	-
2 years after the study	43%	34%	-	-
3 years after the study	43%	30%	-	-
4 years after the study	41%	26%	-	-
5 years after the study	41%	25%	-	-
10 years after the study	52%	14%	-	-
15 years after the study	44%	14%	24%	-
20 years after the study	55%	16%	-	35%[c]
5 years after the study[b]	100%	44%	-	-

(a) The numerator is the difference between the number of cancers in the mammography arm and in the control group, the denominator consists of cancers detected in the mammography arm in the screening; (b) Only cancers detected in mammography screening; (c) Expected incidence in the control group without screening.

DECLARATION OF CONFLICT OF INTERESTS

There is no conflict of interests to declare.

FUNDING SOURCES

None.

REFERENCES

1. Brodersen J, Schwartz LM, Heneghan C, et al. Overdiagnosis: what it is and what it isn't. BMJ Evid Based Med. 2018;23(1):1-3. doi: http://doi.org/10.1136/ebmed-2017-110886
2. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review [Internet]. London: Cancer Research UK; 2012 [cited 2020 Oct 7]. Available from: https://www.cancerresearchuk.org/sites/default/files/breast-screening-review-exec_0.pdf
3. Migowski A, Stein AT, Ferreira CBT, et al. Guidelines for early detection of breast cancer in Brazil. I - Development methods. Cad Saude Publica. 2018;34(6):e00116317. doi: http://doi.org/10.1590/0102-311X00116317
4. Hofmann B. Back to basics: overdiagnosis is about unwarranted diagnosis. Am J Epidemiol. 2019;188(10):1812-7. doi: http://doi.org/10.1093/aje/kwz148
5. Davies L, Petitti DB, Martin L, et al. Defining, estimating, and communicating overdiagnosis in cancer screening. Ann Intern Med. 2018;169(1):36-43. doi: http://doi.org/10.7326/M18-0694
6. Pace LE, Keating NL. A systematic assessment of benefits and risks to guide breast cancer screening decisions. JAMA. 2014;311(13):1327-35. doi: http://doi.org/10.1001/jama.2014.1398
7. Myers ER, Moorman P, Gierisch JM, et al. Benefits and harms of breast cancer screening: a systematic review. JAMA. 2015;314(15):1615-34. doi: http://doi.org/10.1001/jama.2015.13183
8. International Agency for Research on Cancer, World Health Organization. IARC Handbooks of cancer prevention: preamble for screening [Internet]. Lyon: IARC; 2019 [cited 2020 Oct 7]. Available from: http://handbooks.iarc.fr/docs/HB-Preamble-Screening.pdf
9. Etzioni R, Gulati R, Mallinger L, et al. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. Ann Intern Med. 2013;158(11):831-8. doi: http://doi.org/10.7326/0003-4819-158-11-201306040-00008
10. Harris RP. Invited commentary: beyond overdiagnosis-diagnosis without benefit. Am J Epidemiol. 2019;188(10):1818-20. doi: http://doi.org/10.1093/aje/kwz145
11. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med. 2012;367(21):1998-2005. doi: http://doi.org/10.1056/NEJMoa1206809
12. Zahl PH, Andersen JM, Maehlen J. Spontaneous regression of cancerous tumors detected by mammography screening. JAMA. 2004;292(21):2579-80. doi: http://doi.org/10.1001/jama.292.21.2579
13. Welch HG, Gorski DH, Albertsen PC. Trends in metastatic breast and prostate cancer--lessons in cancer dynamics. N Engl J Med. 2015;373(18):1685-7. doi: http://doi.org/10.1056/NEJMp1510443
14. Zahl PH, Jørgensen KJ, Gøtzsche PC. Lead-time models should not be used to estimate overdiagnosis in cancer screening. J Gen Intern Med. 2014;29(9):1283-6. doi: http://doi.org/10.1007/s11606-014-2812-2
15. Zahl PH, Jørgensen KJ, Gøtzsche PC. Overestimated lead times in cancer screening has led to substantial underestimation of overdiagnosis. Br J Cancer. 2013;109(7):2014-9. doi: http://doi.org/10.1038/bjc.2013.427
16. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ. 2015;350:g7773. doi: https://doi.org/10.1136/bmj.g7773
17. Gøtzsche PC. On the benefits and harms of screening for breast cancer. Int J Epidemiol. 2004;33(1):56-64. doi: https://doi.org/10.1093/ije/dyh014
18. Zahl PH, Maehlen J. Ramifications of screening for breast cancer: definition of overdiagnosis is confusing in follow-up of Malmö trial. BMJ. 2006;332(7543):727-8. doi: https://doi.org/10.1136/bmj.332.7543.727-b
19. Gøtzsche PC, Jørgensen K. Screening for breast cancer with magrapmmohy. Cochrane Database Syst Rev. 2013;2013(6):CD001877. doi: https://doi.org/10.1002/14651858.CD001877.pub5
20. Autier P, Boniol M, Koechlin A, et al. Effectiveness of and overdiagnosis from mammography screening in the Netherlands: population based study. BMJ. 2017;359:j5224. doi: https://doi.org/10.1136/bmj.j5224
21. Jørgensen KJ, Zahl PH, Gøtzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. BMC Womens Health. 2009;9:36. doi: https://doi.org/10.1186/1472-6874-9-36
22. Jacklyn G, Glasziou P, Macaskill P, et al. Meta-analysis of breast cancer mortality benefit and overdiagnosis adjusted for adherence: improving information on the effects of attending screening mammography. Br J Cancer. 2016;114(11):1269-76. doi: https://doi.org/10.1038/bjc.2016.90
23. Baines CJ, To T, Miller AB. Revised estimates of overdiagnosis from the Canadian National Breast Screening Study. Prev Med. 2016;90:66-71. doi: https://doi.org/10.1016/j.ypmed.2016.06.033

24. Gøtzsche PC. Ramifications of screening for breast cancer: overdiagnosis in the Malmö trial was considerably underestimated. BMJ. 2006;332(7543):727. doi: https://doi.org/10.1136/bmj.332.7543.727-a

25. Lauby-Secretan B, Scoccianti C, Loomis D, et al. Breast-cancer screening -viewpoint of the IARC Working Group. N Engl J Med. 2015;372(24):2353-8. doi: https://doi.org/10.1056/NEJMsr1504363

26. Moss SM, Wale C, Smith R, et al. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. Lancet Oncol. 2015;16(9):1123-32. doi: https://doi.org/10.1016/S1470-2045(15)00128-X

27. Autier P. Breast cancer: doubtful health benefit of screening from 40 years of age. Nat Rev Clin Oncol. 2015;12(10):570-2. doi: https://doi.org/10.1038/nrclinonc.2015.162

28. Zackrisson S, Andersson I, Janzon L, et al. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ. 2006;332(7543):689-92. doi: https://doi.org/10.1136/bmj.38764.572569.7C

29. Welch HG, Schwartz LM, Woloshin S. Ramifications of screening for breast cancer: 1 in 4 cancers detected by mammography are pseudocancers. BMJ. 2006;332(7543):727. doi: https://doi.org/10.1136/bmj.332.7543.727

30. Baker SG, Prorok PC, Kramer BS. Lead time and overdiagnosis. J Natl Cancer Inst. 2014;106(12):dju346. doi: https://doi.org/10.1093/jnci/dju346

31. van den Ende C, Oordt-Speets AM, Vroling H, et al. Benefits and harms of breast cancer screening with mammography in women aged 40-49 years: a systematic review. Int J Cancer. 2017;141(7):1295-1306. doi: https://doi.org/10.1002/ijc.30794

Recebido em 3/11/2020
Aprovado em 4/12/2020