

Overdiagnosis: epidemiologic concepts and estimation

Jong-Myon Bae

Department of Preventive Medicine, Jeju National University School of Medicine, Jeju, Korea

Overdiagnosis of thyroid cancer was propounded regarding the rapidly increasing incidence in South Korea. Overdiagnosis is defined as 'the detection of cancers that would never have been found were it not for the screening test', and may be an extreme form of lead bias due to indolent cancers, as is inevitable when conducting a cancer screening programme. Because it is solely an epidemiological concept, it can be estimated indirectly by phenomena such as a lack of compensatory drop in post-screening periods, or discrepancies between incidence and mortality. The erstwhile trials for quantifying the overdiagnosis in screening mammography were reviewed in order to secure the data needed to establish its prevalence in South Korea.

KEY WORDS: Cancer screening, Mammography, Thyroid neoplasms, Incidence, Disease progression

INTRODUCTION

According to the 2011 National Cancer Statistics in South Korea (hereafter Korea), thyroid cancer showed the highest annual percentage change (APC) at 23.3% [1]. Such an epidemic was suggested to be caused by overdiagnosis [2,3]. However, if the definition of overdiagnosis is not shared accurately, it can turn into a medical ethics issue that may lead to healthcare providers performing unnecessary treatment. Moreover, when the extent of overdiagnosis is not quantified, the public health policy implications cannot be properly addressed.

An overdiagnosis controversy already exists with respect to prostate, breast, renal, and thyroid cancer in Western countries [4], expressed as 'an epidemic of diagnosis' [5]. There has been suspicion of overdiagnosis, particularly on performing mammography for the early diagnosis of breast cancer, which has been performed since the 1980s [6]. This has greatly impacted the medical field after 2000 [7], including diverse studies focused on estimating the extent of overdiagnosis [8,9].

Correspondence: **Jong-Myon Bae**
Department of Preventive Medicine, Jeju National University School of Medicine, 102 Jejudaehak-ro, Jeju 690-756, Korea
Tel: +82-64-755-5567, Fax: +82-64-725-2593, E-mail: jmbae@jejunu.ac.kr

Received: Jan 17, 2015, Accepted: Feb 10, 2015, Published: Feb 10, 2015

This article is available from: <http://e-epih.org/>

© 2015, Korean Society of Epidemiology

© This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Therefore, the purpose of this study is to examine the implications of the definition of overdiagnosis, focusing on papers regarding mammography overdiagnosis. The study also aims to organize quantification methods to estimate the extent of overdiagnosis, as well as qualification criteria to confirm its presence. Based on this, another goal of the present study was to check the feasibility of quantification methods in Korean conditions and to further investigate which data should be secured in the future. Such attempts are expected to provoke evidence-based medical discussion regarding the overdiagnosis of thyroid cancer, which is a current controversy.

MAIN BODY

Change pattern in cancer incidence according to cancer screening

When a cancer screening program for the purpose of decreasing mortality via early diagnosis is newly initiated or more precise diagnostic devices are introduced, the incidence of corresponding cancer is increased [10], but the change pattern differs by the time of screening [11,12]. At the beginning of screening, incident cases as well as prevalent cases at lead time are detected and diagnosed, thereby exhibiting a rapid increase in incidence [13-16]. As the screening program is continued, only incident cases are detected, with the incidence manifesting as a stable increase [17,18]. Once a screening program is not provided, as with subjects approaching old age, it shows a complementary drop phenomenon whereby the incidence suddenly

decreases (because the incident cases at lead time have already been diagnosed), and then converges to the level of non-screening group incidence [9,19-21]. Biesheuvel et al. [9] schematized such a change pattern.

Meanwhile, change patterns vary depending on incidence levels and trends, implementation extent and accuracy of screening methods, and extent of overdiagnosis [22,23]. Accordingly, a cancer screening program inevitably results in overdiagnosis [24-28]. Therefore, the increase in incidence should not be simplified as an outcome of overdiagnosis [29].

Definition of overdiagnosis

Table 1 shows sentences with definitions found in papers relevant to the overdiagnosis of screening mammography [4,9,16,20,24,25,28-44]. Tracing back the year of publication, they are based on the definition by Prorok et al. [30] in 1999, which is ‘the detection of cancers that would never have been found were it not for the screening test.’

However, there are a couple of tiny differences. First, detec-

tion was used interchangeably with diagnosis [45,46] because the term ‘overdetection’ implies that it is the result of implementing unnecessary screening tests [9,47]. However, the definitions in Table 1 need to be differentiated by whether an object of detection or diagnosis is cancer or lesion. Although Zahl et al. [32] and Day [33] highlighted lesion, the remainder was predominantly specified as cancer, and cancer diagnosis should always accompany patho-histological judgments. Based on these, the term overdiagnosis seems more reasonable than overdetection. Second is the distinction as to whether it includes invasive cancer as well as carcinoma *in situ*. Zahl et al. [32] stated these cases as ‘low malignancy lesions’, whereas Moss [34] and Seigneurin et al. [39] emphasized the inclusion of invasive cancer. As the progression of diagnosed cancer cells can be different [4,48,49] and some cases undergo regression [50-53], it is reasonable to include invasive cancers as well [44,48]. Third, regarding conditions that are only detected with screening tools, Seigneurin et al. [39] emphasized their ‘non-progressive’ nature, and Cervera et al. [28] interpreted them as ‘not having resulted

Table 1. The related sentences for defining overdiagnosis in cancer screening

Author (published year) [reference]	Sentence
Prorok et al. (1999) [30]	The detection of cancers that would never have been found were it not for the screening test
Etzioni et al. (2003) [31]	The detection of cancers that would otherwise not have been detected in the remaining life span of the individuals
Zahl et al. (2004) [32]	The detection of low malignancy lesions that otherwise would not be detected in a patient's lifetime
Day (2005) [33]	Lesions that are detected at screening but which would not have surfaced clinically in the lifetime of the individual
Moss (2005) [34]	The diagnosis of invasive or <i>in situ</i> breast cancer that, in the absence of screening, would not have presented clinically during the woman's lifetime
de Koning et al. (2006) [35]	The detection of cancers by screening that would otherwise never have been clinically diagnosed but are now consequently treated
Paci & Duffy (2005) [29]	Diagnosis of cancer that would not have led to clinically manifested disease in a woman's lifetime, in the absence of screening
Warren & Eleti (2006) [36]	Diagnosis of cancers that would not present during the lifetime of the patient and is one of the downsides of screening
Biesheuvel et al. (2007) [9]	The detection with screening of cancer that would not have presented clinically during the ‘woman's lifetime (and therefore would not be diagnosed in the absence of screening)’
Mandelblatt et al. (2009) [37]	The proportion of cases in each strategy that would not have clinically surfaced in a woman's lifetime among all cases
Duffy et al. (2010) [38]	The diagnosis of a cancer as a result of screening that would not have been diagnosed in the ‘woman's lifetime had screening not taken place
Welch & Black (2010) [4]	The diagnosis of a “cancer” that would otherwise not go on to cause symptoms or death
Seigneurin et al. (2011) [39]	The proportion of non-progressive cancers among all cases of invasive cancer and carcinoma <i>in situ</i>
Gunsoy et al. (2012) [40]	The diagnosis of screen detected cancers that would not have presented clinically in a lifetime in the absence of screening
Hellquist et al. (2012) [16]	The excess of cancers diagnosed with screening compared with without screening that is not due to earlier diagnosis
Marmot (2012) [41]	The diagnosis of cancer as a result of screening which would not have been diagnosed in the patient's lifetime had screening not taken place
Puliti et al. (2012) [20]	The detection of a cancer at screening, histologically confirmed, that would never have been identified clinically in the lifetime of the woman
Yen et al. (2012) [42]	The diagnosis of cancer as a result of screening that would not have arisen in the lifetime of the host
Etzioni et al. (2013) [43]	Excess cases detected because of cancer screening
Miller (2013) [25]	The detection by screening of a cancer not destined to present clinically in the person's lifetime
Zahl et al. (2013) [24]	The detection of lesions that would never have been detected in a person in the absence of screening
Cervera et al. (2014) [28]	The detection of a disease that would have gone undetected without screening when that disease would not have resulted in morbimortality and was treated unnecessarily
Pace & Keating (2014) [44]	The detection of a tumor through screening that would not have become clinically evident in the absence of screening

in morbimortality'. As final results of subjects diagnosed by screening could be identified after death [43], it is therefore difficult to develop methods for overdiagnosis estimation [4,29,54,55]. Fourth, there is a claim that the term 'overtreatment' should be used in lieu of overdiagnosis. Prior to the Prorok et al. [30] statement, Hurley and Kaldor [56] had already been using the term 'overtreatment of abnormalities' as a harm of screening mammography in 1992, and Gur and Sumkin [57] suggested using the term 'overtreatment' when considering society as a whole, as well as the medical field. However, even in cases of overdiagnosis, treatment may not be received according to one's level of understanding of the disease and treatment preference based on potential gains or losses related to the outcome [58]. Thus, the term 'overdiagnosis' is more appropriate. However, overtreatment is a result of a preceding overdiagnosis [46], and overtreatment is obviously prevented once overdiagnosis is reduced [49].

Mechanism of overdiagnosis

Overdiagnosis occurs when non-progressive or regressive cancer is diagnosed, and when deaths are caused by other reasons

[4,20,29,36]. Among them, the first indicates an improper control of the interference of length time bias, which is more frequently diagnosed for indolent cancer that shows better prognosis due to slower-than-expected progression where changes in mortality cannot be made, even without treatment [9,49,54, 59-61]. In other words, unknown future progression at the time of diagnosis is a major reason of overdiagnosis [8,20,33]. The occurrence of overdiagnosis according to various cancer progressions was schematized by Welch and Black [4] and Zahl et al. [24] (Figure 1). Due to this background, Esserman et al. [62] proposed the term 'indolent lesion of epithelial origin (IDLE)'.

Judgment criteria for the presence of overdiagnosis and its estimation

When overdiagnosis occurs after the diagnosis of IDLE cancer, a complementary drop phenomenon with respect to incidence does not occur after screening termination [8-10,20,21, 63]; moreover, mortality does not change substantially compared to an increase in incidence [4]. The presence of overdiagnosis is determined based on these two phenomena (Table 2),

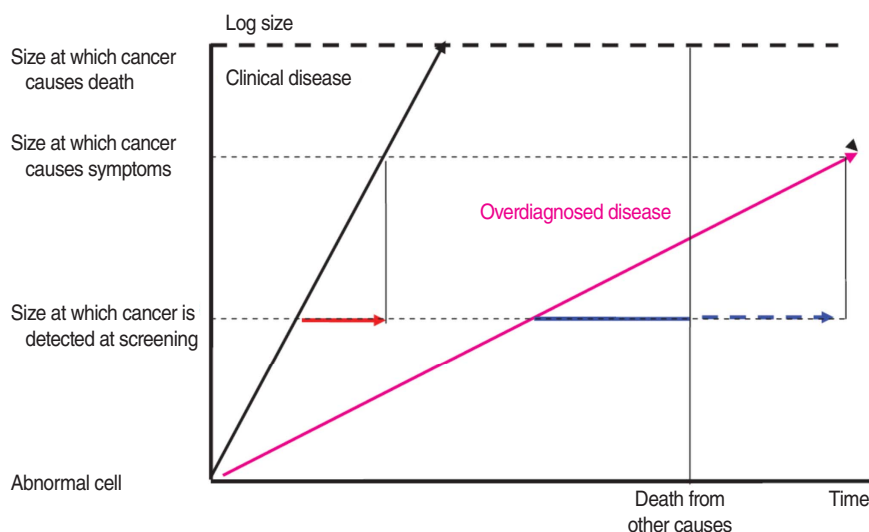


Figure 1. Two different types of tumor growth indicated by the black (clinical disease) and pink arrows (overdiagnosed disease). Clinically relevant lead time is indicated by the red arrow. When overdiagnosed tumors are included, some researchers defined lead time as time to death (solid blue arrow, so-called censored lead time), while others included the time to clinical appearance if there had been no competing causes of death (solid plus dotted blue arrow, so-called uncensored lead time). Source from Zahl, et al. Br J Cancer 2013;109:2014-2019, with permission from Oxford University Press [24].

Table 2. Decision criteria of identifying overdiagnosis

	Qualitative	Quantitative
Randomized trials	No compensatory drop after stopping screening program	Excess cases per screen-detected cases (Figure 1)
Observational studies	Rapid rising in cancer diagnosis without a following change in mortality	Incidence rate ratio between the observed in screen group and the expected in control group

and these form the base of study by Ahn et al. [2] that claims overdiagnosis as the main cause for the epidemic in thyroid cancer among Koreans.

Meanwhile, the definition of overdiagnosis is based on epidemiologic deduction [29] and therefore, the extent of overdiagnosis cannot be directly estimated unless cancer progression can be known in advance at the current medical level [4,10,34,53,55]. Still, a best guess can be made in the presence of long-term follow-up data with randomized controlled trials [4,25,64,65]. To more conveniently explain the estimation of the extent of overdiagnosis, contents described by Welch and Black [4] are schematized in Figure 2. Cancer identified in a screening group can be classified into cases diagnosed by screening (B) and cases diagnosed by clinical symptom appearance (As). Cancer detected in the control group that did not receive screening can be divided into cases diagnosed by clinical symptom appearance (An) and cases diagnosed after the screening completion (C). The difference between screening group and non-screening group, $d (= As+B-An-C)$, is the extent of overdiagnosis, and the percentage (%) of d with denominator of the cases diagnosed by screening in screening group (B) is the overdiagnosis proportion. While cases diagnosed in the screening group after the screening termination are not further reflected, catch-up cases in the control group (C) are added in order to differentiate over-incident cases by lead time bias and over-incident cases by overdiagnosis [34, 65-67]. The Malmö study had 15 years of additional observation after study completion [67]. If the incidence of the screening group is not reduced after screening termination and remains high compared to the non-screening group, a longer follow-up period is necessary [8,21,68].

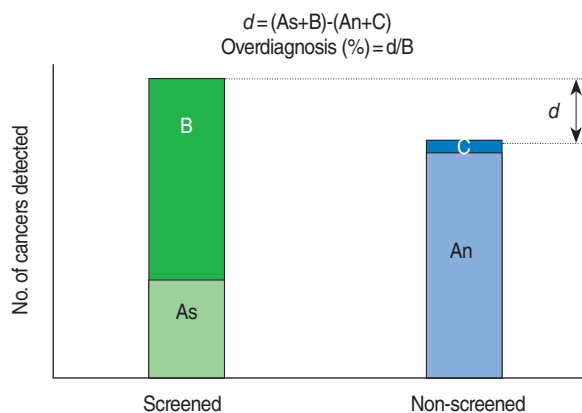


Figure 2. Extent of overdiagnosis in a randomized trial. As, clinically detected cancers in the screened group; An, clinically detected cancers in the non-screened group; B, screen-detected cancers in the screened group; C, catch-up cancers of post-screening periods in the non-screened group; d , extra cancers. Modified from Welch and Black. *J Natl Cancer Inst* 2010;102:605-613, with permission from Nature Publishing Group [4].

Instead of d that can be obtained by randomized tests (Figure 2), the extent of overdiagnosis in an observational study is considered as the difference between the observed cumulative incidence in the screening group and the expected cumulative incidence in the control group [15,18,42,43,63,69,70]. However, the following four major items should be reviewed in the study plan to secure the validity of the estimated outcomes [9,20]. First, the selection of the non-screening control group should be assessed [27,43,55]. Randomized trials can secure an assumption that the risk of cancer incidence is at the same level irrespective of control or screening group [38,42,67], whereas the best control group that satisfies this assumption should be selected in observational studies [71]. The non-screening group cohort is established based on screening through a prospective cohort study [72], or a historical national control group is proposed via retrospective cohort study prior to the introduction of the screening program [9,10,27,73]. Second, the validity of the methods used to estimate the predicted incidence in the control group should be determined [74,75]. Longstanding incidence rates are needed; therefore, it is important and necessary to have modeling that actively reflects the incidence trend via age-period-cohort (APC) analysis [15,18,39,65,69,75,76]. Third, the conceptual determination of lead time bias [24,77], acceptability of their control methods [4,8,21,53,72,73,78], and confinement of follow-up period ranges [9,10,43,53] are important to this process. Fourth, the possibility of obtaining information to differentiate As and B in Figure 2 for screening group incident cases is important [29]. Although it is easy to obtain information when population-based cancer registry information is utilized in a prospective cohort study [72], survey outcomes performed for other purposes can be derived in a retrospective cohort study [55].

Therefore, observational studies have a variety of estimation methods according to types of control group selection, estimation methods of expected incidence, and lead time control methods [8,20,32,42,43,79]. Biesheuvel et al. [9] classified these diverse study methods into incidence rate, cumulative incidence, and modeling with respect to utilized epidemiologic information. Based on this classification, estimating procedures of the expected incidence in control group, roles of the National Cancer Registry Program, and additionally requested information are summarized in Table 3 in order to identify which studies are plausible in domestic conditions and type of information required in the future. Utilization of cumulative incidence rather than incidence rate is primarily recommended [9,20], and study design differs depending on how to estimate the expected incidence of the control group in Table 3, while the National Cancer Registry Program plays an important role in all cases. Statistical modeling that can reflect the natural history of diseases [9,69] is possible only when such information, including inter-

Table 3. The required data for quantifying overdiagnosis of cancer screening

Categories	Study design	Required data			Reference
		For expected incidences in non-screening group	Population-based registry	Specific information	
Incidence	Cohort	Historical national control population before implementing the screening programme	Nationwide annual incidences of invasive cancer		32
Cumulative incidence	Randomized trials	Randomization	Linkage for follow-up	Catch-up cancers in control group after the end of trial	67
	Cohort	Construction of the non-screened cohort	Linkage for follow-up		72
	Population-based programme	Historical control population before implementation of the screening programme	Source of control group and/or Linkage for follow-up		73
Modeling	(Simulation)	Micro-simulation model	Annual age, stage, and screening-specific incidences	Interval cancer rates since the start of screening	69
	(Simulation)	Approximate Bayesian computation analysis with a stochastic simulation model	Lifetime probability of the cancer	The natural course of invasive cancer & carcinomas <i>in situ</i>	39

val cancer [69,80,81] is secured [22,35,39,53]. Considering these and domestic conditions lacking randomized trials for cancer screening, a cohort analysis of cumulative incidence can be performed first if it is known whether the diagnosis was made by screening or was made due to clinical findings [34,72]. If this is challenging, comparative analyses on cumulative incidences between pre- and post-program of cancer screening can be conducted. However, considering that National Cancer Statistics information is available from 2002, estimation methods applying APC can be implemented [55,82].

CONCLUSION AND SUGGESTION

Estimating the extent of overdiagnosis is a method of investigating the extent of the harms and benefits of screening [38,55], and this is an important evidence for determining public health policies [20,43]. Considering current conditions in Korea, the extent of overdiagnosis can be estimated by establishing a strategy that utilizes statistical results from the Korea Central Cancer Registry database and individual follow-up data as much as possible [65,83]. The barriers are that information regarding whether a diagnosis was made by screening has been collected only in recent years [29], and official statistics have been published since 2002 [34]. To overcome these limitations, national research institutes should promote and lead relevant studies in the optimal use of collectable information.

Despite efforts to validate the utilizing data, there are still limitations left to be addressed regarding the correct interpretation of the estimated extent of overdiagnosis [59,75]. For example, the extent of incidence increase can be differently interpreted according to screening participation rate [43], sensitivity may

alter depending on the introduction of new screening methods [54], and risk factors of breast cancer incidence can change (e.g., an increase of oral hormone administration or a change in maternity nursing history) [11,55], and effective treatment methods can be newly developed [55]. Because of these limitations, overdiagnosis estimation ranges of screening mammography are very broad [9,20], and thus there is still controversy on the overdiagnosis [43]. Although a checklist was proposed for the appraisal of related studies [43], an ultimate solution for this controversy is to overcome length time bias causing overdiagnosis [49,54,59-61]. It is therefore necessary to identify the percentage of cancer that has not progressed to invasive cancer in carcinoma *in situ* [84], the IDLE proportion of overall diagnostic cancer [62], and the percentage of autopsy-based diagnosis [85]. In addition, the Korea Central Cancer Registry has collected information of cancer stage in Surveillance, Epidemiology, and End Results (SEER) from 2009 incident cases [1,82], and it is necessary to obtain follow-up results according to the SEER stage regarding thyroid, breast, and prostate cancers, which have been currently controversial in overdiagnosis. Lastly, given the fact that subjects who are negatively impacted from overdiagnosis are healthy individuals [4,27], decision aids allowing for the understanding of losses and gains of correct screening need to be simultaneously developed [58,86].

ACKNOWLEDGEMENTS

This study was supported by 2013 cancer research support project from the Korea Foundation for Cancer Research (no. 2013-2).

CONFLICT OF INTEREST

The author has no conflicts of interest to declare for this study.

SUPPLEMENTARY MATERIAL

Supplementary material is available at <http://www.e-epih.org/>.

REFERENCES

1. Jung KW, Won YJ, Kong HJ, Oh CM, Lee DH, Lee JS. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2011. *Cancer Res Treat* 2014;46:109-113.
2. Ahn HS, Kim HJ, Welch HG. Korea's thyroid-cancer "epidemic"-screening and overdiagnosis. *N Engl J Med* 2014;371:1765-1767.
3. Lee JH, Shin SW. Overdiagnosis and screening for thyroid cancer in Korea. *Lancet* 2014;384:1848.
4. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;102:605-613.
5. Davies L, Welch HG. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg* 2014;140:317-322.
6. Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. *Int J Epidemiol* 2004;33:43-55.
7. Bae JM. On the benefits and harms of mammography for breast cancer screening in Korean women. *Korean J Fam Pract* 2014;4:1-6 (Korean).
8. Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
9. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer over-detection with mammography screening: a systematic review. *Lancet Oncol* 2007;8:1129-1138.
10. Njor SH, Olsen AH, Blichert-Toft M, Schwartz W, Vejborg I, Lynge E. Overdiagnosis in screening mammography in Denmark: population based cohort study. *BMJ* 2013;346:f1064.
11. Jonsson H, Johansson R, Lenner P. Increased incidence of invasive breast cancer after the introduction of service screening with mammography in Sweden. *Int J Cancer* 2005;117:842-847.
12. Glass AG, Lacey JV Jr, Carreon JD, Hoover RN. Breast cancer incidence, 1980-2006: combined roles of menopausal hormone therapy, screening mammography, and estrogen receptor status. *J Natl Cancer Inst* 2007;99:1152-1161.
13. Kopans DB. Arguments against mammography screening continue to be based on faulty science. *Oncologist* 2014;19:107-112.
14. Svendsen AL, Olsen AH, von Euler-Chelpin M, Lynge E. Breast cancer incidence after the introduction of mammography screening: what should be expected? *Cancer* 2006;106:1883-1890.
15. Martinez-Alonso M, Vilaprinyo E, Marcos-Gragera R, Rue M. Breast cancer incidence and overdiagnosis in Catalonia (Spain). *Breast Cancer Res* 2010;12:R58.
16. Hellquist BN, Duffy SW, Nyström L, Jonsson H. Overdiagnosis in the population-based service screening programme with mammography for women aged 40 to 49 years in Sweden. *J Med Screen* 2012;19:14-19.
17. Boer R, Warmerdam P, de Koning H, van Oortmarssen G. Extra incidence caused by mammographic screening. *Lancet* 1994;343:979.
18. Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH. Complexities in the estimation of overdiagnosis in breast cancer screening. *Br J Cancer* 2008;99:1176-1178.
19. Møller B, Weedon-Fekjaer H, Hakulinen T, Tryggvadóttir L, Storm HH, Talbäck M, et al. The influence of mammographic screening on national trends in breast cancer incidence. *Eur J Cancer Prev* 2005;14:117-128.
20. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen* 2012;19 Suppl 1:42-56.
21. Ciatto S. The overdiagnosis nightmare: a time for caution. *BMC Womens Health* 2009;9:34.
22. Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst* 2002;94:981-990.
23. Feuer EJ, Wun LM. How much of the recent rise in breast cancer incidence can be explained by increases in mammography utilization? A dynamic population model approach. *Am J Epidemiol* 1992;136:1423-1436.
24. Zahl PH, Jørgensen KJ, Gøtzsche PC. Overestimated lead times in cancer screening has led to substantial underestimation of overdiagnosis. *Br J Cancer* 2013;109:2014-2019.
25. Miller AB. Overdiagnosis of breast cancer. *Int J Cancer* 2013;133:2511.
26. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2006:CD001877.
27. Paci E, Warwick J, Falini P, Duffy SW. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? *J Med Screen* 2004;11:23-27.
28. Cervera Deval J, Sentís Crivillé M, Zulueta JJ. Overdiagnosis in cancer screening. *Radiologia* 2014. doi:10.1016/j.rx.2014.06.007.
29. Paci E, Duffy S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis and overtreatment in service screening. *Breast Cancer Res* 2005;7:266-270.
30. Prorok PC, Kramer BS, Gohagan JK. Screening theory and study design: the basics. In: Kramer BS, Gohagan JK, Prorok PC, editors. *Cancer screening: theory and practice*. New York: Marcel Dekker; 1999. p. 29-53.
31. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, et al. The case for early detection. *Nat Rev Cancer* 2003;3:243-252.
32. Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *BMJ* 2004;328:921-924.
33. Day NE. Overdiagnosis and breast cancer screening. *Breast Cancer Res* 2005;7:228-229.
34. Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Res* 2005;7:230-234.
35. de Koning HJ, Draisma G, Fracheboud J, de Bruijn A. Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. *Breast Cancer Res* 2006;8:202.
36. Warren R, Eleti A. Overdiagnosis and overtreatment of breast cancer: is overdiagnosis an issue for radiologists? *Breast Cancer Res* 2006;8:205.
37. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738-747.

38. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010;17:25-30.
39. Seigneurin A, François O, Labarère J, Oudeville P, Monlong J, Colonna M. Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data. *BMJ* 2011;343:d7017.
40. Gunsoy NB, Garcia-Closas M, Moss SM. Modelling the overdiagnosis of breast cancer due to mammography screening in women aged 40 to 49 in the United Kingdom. *Breast Cancer Res* 2012;14:R152.
41. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778-1786.
42. Yen AM, Duffy SW, Chen TH, Chen LS, Chiu SY, Fann JC, et al. Long-term incidence of breast cancer by trial arm in one county of the Swedish Two-County Trial of mammographic screening. *Cancer* 2012;118:5728-5732.
43. Etzioni R, Gulati R, Mallinger L, Mandelblatt J. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. *Ann Intern Med* 2013;158:831-838.
44. Pace LE, Keating NL. A systematic assessment of benefits and risks to guide breast cancer screening decisions. *JAMA* 2014;311:1327-1335.
45. Yaffé MJ, Pritchard KI. Overdiagnosing overdiagnosis. *Oncologist* 2014;19:103-106.
46. Hersch J, Barratt A, Jansen J, Houssami N, Irwig L, Jacklyn G, et al. The effect of information about overdiagnosis of breast cancer on women's decision-making about mammography screening: study protocol for a randomised controlled trial. *BMJ Open* 2014;4:e004990.
47. Suzuki A, Ishida T, Ohuchi N. Controversies in breast cancer screening for women aged 40-49 years. *Jpn J Clin Oncol* 2014;44:613-618.
48. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013;108:2205-2240.
49. Puliti D, Paci E. The other side of technology: risk of overdiagnosis of breast cancer with mammography screening. *Future Oncol* 2009;5:481-491.
50. Zahl PH, Maehlen J, Welch HG. The natural history of invasive breast cancers detected by screening mammography. *Arch Intern Med* 2008;168:2311-2316.
51. Dussan C, Zubor P, Fernandez M, Yabar A, Szunyogh N, Visnovsky J. Spontaneous regression of a breast carcinoma: a case report. *Gynecol Obstet Invest* 2008;65:206-211.
52. Ross MB, Buzdar AU, Hortobagyi GN, Lukeman JM. Spontaneous regression of breast carcinoma: follow-up report and literature review. *J Surg Oncol* 1982;19:22-24.
53. Etzioni R, Gulati R. Oversimplifying overdiagnosis. *J Gen Intern Med* 2014;29:1218-1220.
54. Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW. Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. *Breast J* 2006;12:338-342.
55. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* 2012;367:1998-2005.
56. Hurlley SF, Kaldor JM. The benefits and risks of mammographic screening for breast cancer. *Epidemiol Rev* 1992;14:101-130.
57. Gur D, Sumkin JH. Screening for early detection of breast cancer: overdiagnosis versus suboptimal patient management. *Radiology* 2013;268:327-328.
58. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2014;1:CD001431.
59. Al-Foheidi M, Al-Mansour MM, Ibrahim EM. Breast cancer screening: review of benefits and harms, and recommendations for developing and low-income countries. *Med Oncol* 2013;30:471.
60. Berry DA, Baines CJ, Baum M, Dickersin K, Fletcher SW, Gøtzsche PC, et al. Flawed inferences about screening mammography's benefit based on observational data. *J Clin Oncol* 2009;27:639-640.
61. Heywang-Köbrunner SH, Hacker A, Sedlacek S. Advantages and Disadvantages of Mammography Screening. *Breast Care (Basel)* 2011;6:199-207.
62. Esserman LJ, Thompson IM, Reid B, Nelson P, Ransohoff DF, Welch HG, et al. Addressing overdiagnosis and overtreatment in cancer: a prescription for change. *Lancet Oncol* 2014;15:e234-e242.
63. Zahl PH, Mæhlen J. Overdiagnosis of breast cancer after 14 years of mammography screening. *Tidsskr Nor Laegeforen* 2012;132:414-417.
64. Marmot MG. Sorting through the arguments on breast screening. *JAMA* 2013;309:2553-2554.
65. Duffy SW, Michalopoulos D, Sebuødegård S, Hofvind S. Trends in aggregate cancer incidence rates in relation to screening and possible overdiagnosis: a word of caution. *J Med Screen* 2014;21:24-29.
66. Duffy SW, Parmar D. Overdiagnosis in breast cancer screening: the importance of length of observation period and lead time. *Breast Cancer Res* 2013;15:R41.
67. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006;332:689-692.
68. Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. *JAMA* 2009;302:1685-1692.
69. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev* 2011;33:111-121.
70. Junod B, Zahl PH, Kaplan RM, Olsen J, Greenland S. An investigation of the apparent breast cancer epidemic in France: screening and incidence trends in birth cohorts. *BMC Cancer* 2011;11:401.
71. Peeters PH, Verbeek AL, Straatman H, Holland R, Hendriks JH, Mravunac M, et al. Evaluation of overdiagnosis of breast cancer in screening with mammography: results of the Nijmegen programme. *Int J Epidemiol* 1989;18:295-299.
72. Puliti D, Miccinesi G, Zappa M, Manneschi G, Crocetti E, Paci E. Balancing harms and benefits of service mammography screening programs: a cohort study. *Breast Cancer Res* 2012;14:R9.
73. Paci E, Miccinesi G, Puliti D, Baldazzi P, De Lisi V, Falcini F, et al. Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. *Breast Cancer Res* 2006;8:R68.
74. Jørgensen KJ, Zahl PH, Gøtzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. *BMC Womens Health* 2009;9:36.
75. Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. *Cancer Causes Control* 2010;21:275-282.
76. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006:19-25.
77. Zahl PH, Jørgensen KJ, Gøtzsche PC. Lead-time models should not be used to estimate overdiagnosis in cancer screening. *J Gen Intern Med* 2014;29:1283-1286.
78. Zahl PH, Jørgensen KJ, Maehlen J, Gøtzsche PC. Biases in estimates of overdiagnosis due to mammography screening. *Lancet Oncol* 2008;

- 9:199-201.
79. Puliti D, Zappa M, Miccinesi G, Falini P, Crocetti E, Paci E. An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence. *Eur J Cancer* 2009;45:3166-3171.
80. Caumo F, Vecchiato F, Strabbioli M, Zorzi M, Baracco S, Ciatto S. Interval cancers in breast cancer screening: comparison of stage and biological characteristics with screen-detected cancers or incident cancers in the absence of screening. *Tumori* 2010;96:198-201.
81. Heidinger O, Batzler WU, Krieg V, Weigel S, Biesheuvel C, Heindel W, et al. The incidence of interval cancers in the German mammography screening program: results from the population-based cancer registry in North Rhine-Westphalia. *Dtsch Arztebl Int* 2012;109:781-787.
82. Bleyer A. Were our estimates of overdiagnosis with mammography screening in the United States “based on faulty science”? *Oncologist* 2014;19:113-126.
83. Falk RS, Hofvind S, Skaane P, Haldorsen T. Overdiagnosis among women attending a population-based mammography screening program. *Int J Cancer* 2013;133:705-712.
84. Yen MF, Tabár L, Vitak B, Smith RA, Chen HH, Duffy SW. Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. *Eur J Cancer* 2003;39:1746-1754.
85. Thorat MA. What caused the decline in US breast cancer incidence? *Nat Clin Pract Oncol* 2008;5:314-315.
86. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. *JAMA Intern Med* 2014;174:448-454.