

1
2
3
4
5
6
7

<연구방법>

과진단: 역학적 개념과 산출법

Overdiagnosis : Epidemiologic Concepts and Estimation

Epidemiology and Health

8

Abstract

9

10 Overdiagnosis of thyroid cancer was propounded about the rapidly increasing
11 incidence in Korean. Overdiagnosis is defined as 'the detection of cancers that would never
12 have been found were it not for the screening test'. It may be an extreme form of length
13 bias due to indolent cancers so that it is inevitable with conducting cancer screening
14 programme. Because of only an epidemiological concept, it can be guessed indirectly by
15 the phenomena such as no compensatory drop in post-screening periods or the wider
16 discrepancy between incidence and mortality. The erstwhile trials for quantifying the
17 overdiagnosis in screening mammography were reviewed in order to secure the data needed
18 to establish the overdiagnosis in Korea.

19

20

21 Key words: Cancer screening, Mammography, Thyroid neoplasms, Incidence, Disease
22 progression

23

24 **1. 서론**

25

26 2011년도 국가암통계 발표에 따르면 한국인에서 암발생 1위인 갑상선암은 연간 변
27 동률 (annual percentage change, APC) 23.3%로 단연 높다[1]. 이처럼 급격한 증가
28 (epidemic)를 보인 현상을 과진단 (overdiagnosis)¹에 의한 것이라는 주장이 제기되었다
29 [2,3]. 그런데 과진단의 정의를 정확히 공유하지 않으면 의료진들이 불필요한 치료를 한다는
30 의료윤리적 문제로 비화될 수 있으며, 과진단의 크기를 계량화하지 않을 경우 과진단의 보
31 건의료 정책적 함의를 제대로 도출할 수 없다.

32

33 과진단의 논란은 이미 서구 나라에서 갑상선을 포함하여 전립선, 유방, 신장 등에서
34 있었으며[4], 해당 암에 대한 진단의 폭증 (an epidemic of diagnosis)이란 용어로 표현하고
35 있다[5]. 특히 1980년대부터 시행되어 온 유방암 조기검진을 위한 유방촬영술
36 (mammography)에 대하여[6], 과진단의 주장으로 2000년도 이후 의학계에 큰 파장을 일으
37 키고 있으며[7], 과진단의 크기를 산출하려는 다양한 시도들이 보고되었다[8,9].

38

39 따라서 본 원고는 유방촬영술 과진단에 관한 논문들을 중심으로 과진단의 정의에
40 함축된 의미들을 살펴보고, 과진단의 유무를 확인하는 정성적 기준뿐만 아니라 그 크기도
41 산출하는 정량적 방법들을 정리하고자 한다. 이를 통해 한국의 여건상 산출법의 현실 가능
42 성을 점검하고, 나아가 앞으로 어떤 자료들을 확보해야 할 것인가를 알아보하고자 한다. 이런
43 시도가 현재 논란이 되고 있는 갑상선암의 과진단에 대한 근거 기반의 의학적 논의들을 촉

¹ Overdiagnosis 용어 중 'over-' 번역을 '과다, 과잉'으로 할 경우 객관적 사실을 넘어서 사회적 판단 가치까지 함의될 여지가 있어, '과(過)-'란 사실적 의미만을 강조하기 위하여 "과진단"이라 함.

44 발시키기를 기대한다.

45

46

Epidemiology and Health

47 **2. 본론 :**

48

49 **가. 암검진 시행에 따른 암발생률 변화양상**

50

51 조기 진단으로 사망률을 감소시킬 목적으로 시행되는 암검진사업을 새로이 시작하
52 거나, 보다 정밀한 진단기구가 도입되면 해당 암의 발생률이 증가하는데[10], 검진시행 시기
53 별로 변화 양상이 다르다[11,12]. 검진시행 초기에는 조기발견기간 (lead time)에 있던 유병
54 자 (prevalent case) 뿐만 아니라, 새로이 생긴 발생자 (incident case)까지 검출하여 진단함
55 으로서 급격한 발생률 증가 현상을 보인다[13-16]. 검진사업이 지속되면 발생자만 검출되면
56 서 발생률은 안정적인 증가를 보인다[17,18]. 고령이 되면서 검진사업을 받지 않게 되면 그
57 동안 조기발견기간에 생기는 발생자를 이미 진단해 버렸기 때문에 발생률이 갑자기 감소하
58 는 보상감소 (complementary drop) 현상을 보인 다음, 미검진군의 발생 수준으로 수렴한다
59 [9,19-21]. Biesheuvel et al. [9] 은 이러한 변동양상을 도식화하여 제시하였다.

60 한편 변동양상은 발생률 수준 및 추세, 검진법의 시행 정도, 검진술의 정확성과 함
61 께 과진단의 크기에 따라 달리 나타난다[22,23]. 따라서 암검진 사업은 과진단의 문제를 불
62 가피하게 (inevitable)으로 야기시킨다[24-28]. 따라서 발생률 증가 현상을 과진단의 결과로
63 단순화해서는 안된다[29].

64

65 **나. 과진단의 정의**

66

67 유방촬영술의 과진단에 관한 논문들[4,9,16,20,24,25,28-44]에 있어 정의를 하고 있는
68 문구를 찾아 Table 1로 정리해 보았다. 발표 연도에 따라 거슬러 올라갔을 때, 1999년도
69 Prorok et al. [30]의 '평생동안 몰랐을 암을 검진으로 알아낸 경우 (the detection of cancers

70 that would never have been found were it not for the screening test)'란 정의를 근간으로 하고 있
71 다.

72

73 그렇지만 몇 가지 점에서 미세한 차이를 발견할 수 있다. 첫째는 검출(detection)과
74 진단(diagnosis)의 차이를 두지 않고 혼용해서 사용하고 있다[45,46]. 과검출 (overdetection)
75 이란 용어에는 불필요한 검진 검사를 시행한 결과임을 강조한다는 의미를 갖기 때문이다
76 [9,47]. 그런데, 표 1의 정의에서 검출/진단의 목적어가 암 (cancer)와 병변 (lesion)인가에 따
77 라 구분할 필요가 있다. Zahl et al. [32]와 Day [33]는 lesion을 강조한 반면, 나머지 대부분
78 은 cancer로 명시하고 있으며, 병리조직학적 (patho-histological) 판정을 받아야만 암으로
79 진단할 수 있다. 이런 점에서 과검출보다는 과진단이란 용어가 보다 타당하겠다. 둘째는 상
80 피내암(carcinoma in situ) 뿐만 아니라 침윤(invasive)암 포함 여부이다. Zahl et al. [32]은
81 'low malignancy lesions' 으로 제시한 반면, Moss [34]와 Seigneurin et al. [39]은 침윤암도
82 포함한다는 것을 강조하고 있다. 진단받은 암세포의 진행속도가 다를 수 있고[4,48,49], 어떤
83 경우에는 퇴화 (regression)하는 경우도 있기에[50-53], 침윤암도 포함하는 것이 타당하다
84 [44,48]. 셋째, 검진이 없었으면 알지 못한다는 조건에 대하여 Seigneurin et al. [39]은 'non-
85 progressive'를 강조하였고, Cervera et al. [28]은 'not have resulted in morbimortality'로 해석
86 하고 있다. 그러나 해당 조건은 검진으로 진단된 대상자를 사후에야 알아낼 수 있기 때문에
87 [43], 과진단의 산출법 개발에 어려움을 갖게 한다[4,29,54,55]. 넷째, 과진단 말고도 과치료
88 (over-treatment)란 용어를 사용해야 한다는 주장이 있다. Prorok et al. [30] 발표 이전인
89 1992년도에 Hurley & Kaldor [56]은 유방촬영술 유해의 하나로 'overtreatment of
90 abnormalities'란 용어를 이미 사용했으며, Gur & Sumkin [57]은 의료계 뿐만 아니라 사회
91 전반을 고려할 때 과치료란 용어를 사용하기를 제안하였다. 그러나 과진단을 해도 당사자의
92 질병이해 수준과 결과 득실의 선호도에 따라 치료받지 않을 수 있다는 주장[58]이 있어 과

93 진단이란 용어가 더 적합하다. 다만 과치료는 과진단이 선행한 결과에 해당하므로[46], 과진
94 단을 줄이면 과치료를 막는 것은 분명하다[49].

95

96 **다. 과진단의 발생 기전**

97

98 **과진단이 일어나는 경우로는 진행을 하지 않거나 (non-progressive) 저절로 퇴화하**
99 **는 (regressive) 암을 진단했거나, 다른 원인으로 사망한 경우로 이분할 수 있다[4,20,29,36].**
100 **이중 첫 번째 경우는 예상보다 늦은 진행으로 예후가 더 좋고 치료를 안해도 사망률 변동**
101 **을 만들지 못하는 지체하는 (indolent) 암[43,53]일수록 반복되는 암검진에서 더 잘 진단되**
102 **는 기간차이 바이어스 (length time bias)가 개입되는 것을 제대로 통제하지 못함을 뜻한다**
103 **[9,49,54,59-61]. 다시 말해서 진단 당시 향후의 진행여부를 미리 알아낼 수 없다는 점이 과**
104 **진단의 주된 이유인 것이다[8,20,33]. 암의 다양한 진행에 따른 과진단의 발생 현상을**
105 **Welch & Black [4] 과 Zahl et al. [24] 에서 도식화하여 제시하였다 (Fig 1). 이런 배경 때문에**
106 **Esserman et al. [62]은 indolent lesion of epithelial origin (IDLE)이란 용어를 제안하였다.**

107

108 **라. 과진단의 유무 판단기준과 산출법**

109

110 IDLE에 해당하는 암을 진단하여 과진단이 생겼다면 검진 중단 후에 나타나는 발생
111 률의 보상감소 현상이 나타나지 않거나[8-10,20,21,63], 발생률 증가에 비해 사망률은 별 변
112 동이 없는 현상이 벌어지게 된다[4]. 이런 두 가지 현상에 따라 과진단 유무를 판단하며
113 (Table 2), Ahn et al. [2]이 한국의 갑상선암 증가에 있어 과진단을 주장한 근거인 것이다.

114

115 한편, 과진단의 정의 자체가 역학적 추론에 기반하고 있기 때문에[29], 현대 의학수

116 준으로 암진행여부를 미리 알아낼 수 없는 한 과진단의 크기를 직접 산출할 수 없다
117 [4,10,34,53,55]. 그나마 무작위 시험 연구로 장기간 관찰한 자료가 있다면 최선의 예측
118 (best guess)을 해 낼 수 있다[4,25,64,65]. 과진단의 크기 산출 과정을 보다 쉽게 설명하기
119 위하여, Welch & Black [4]가 기술한 내용을 Fig. 2로 도식화하였다. 검진을 받은 군에서 발
120 견된 암은 검진으로 인해 진단된 경우 (B)와 임상증상 발현에 따라 진단된 경우 (As)로 나
121 늘 수 있다. 검진을 받지 않는 비교군에서 발견된 암은 임상증상 발현에 따라 진단된 경우
122 (An)와 검진시행 종료이후에 진단된 경우 (C)로 나눌 수 있다. 여기서 검진군과 비검진군간
123 의 진단자 차이인 $d (=As+B-An-C)$ 가 과진단 크기가 되며, 검진군에서 검진으로 인해 진단
124 된 경우 (B)를 분모로 한 d 의 비율 (%)을 과진단율이라 한다. 검진이 종료된 이후에 검진군
125 에서 진단된 경우는 더 이상 반영하지 않지만, 비교군에서 추가로 진단된 (catch-up) 경우
126 (C)는 추가하는 것은, 조기발견 바이어스 (lead time bias)에 의한 과잉 발생자와 과진단에
127 의한 과잉 발생자를 구별하기 위한 것이다[34,65-67]. Malmö 연구에서는 종료 후 15년의
128 추가 관찰기간을 가졌다[67]. 만약 검진 중단 후에도 검진군의 발생률이 감소하지 않고 비
129 검진군에 비해 계속해서 높게 유지될 경우에는 예상보다 더 긴 추적기간을 가져야 한다
130 [8,21,68].

131
132 관찰연구에서 과진단의 크기는 무작위 시험으로 얻어낼 수 있는 Fig. 1의 d 를 대신
133 하여, 검진군의 관찰된 (observed) 누적 발생률과 비교군의 예상된 (expected) 누적 발생률
134 간의 차이 크기로 삼는다[15,18,42,43,63,69,70]. 그렇지만 산출된 결과의 타당성을 확보하기
135 위해서는 연구기획 단계에서 다음 4가지 주요 사항을 검토해야 한다[9,20]. 첫째, 비검진 비
136 교군의 선정 여부이다[27,43,55]. 무작위 시험은 비교군과 검진군간의 암발생 위험 수준이
137 동일하다는 가정을 확보할 수 있는 반면에[38,42,67], 관찰연구는 이 가정을 만족하는 최선
138 의 비교군을 선정해야만 한다[71]. 전향적 코호트 연구를 통해 검진여부에 맞추어 비검진자

139 코호트를 구축하거나[72], 후향적 코호트 연구를 통해 검진사업 도입 전의 비교군
140 (historical national control group)를 상정한다[9,10,27,73]. 둘째, 비교군에서 예상 발생률을
141 추정하는 방법의 타당성 여부이다[74,75]. 최대한 오랜 연도의 발생률 수치가 필요한데, 연
142 령-기간-코호트 분석 (age-period-cohort analysis) 등으로 발생 추세를 적극 반영하는 모델
143 링이 필요하다[15,18,39,65,69,75,76]. 셋째, 조기발견 바이어스의 개념 설정[24,77], 이의 통
144 제 방법의 수용성 여부[4,8,21,53,72,73,78], 그리고 추적 관찰기간의 범위 산정[9,10,43,53]이
145 다. 넷째, 검진군의 발생자에 있어 Fig. 1의 As 와 B를 구분하는 정보의 확보 가능성 여부가
146 다[29]. 전향적 코호트 연구로 암등록 (Registry)의 정보를 활용한다면 정보 확보가 용이하나
147 [72], 후향적 코호트 연구를 한다면 다른 목적으로 시행한 조사(survey) 결과를 차용할 수도
148 있다[55].

149
150 따라서 관찰연구는 비교군 선정 종류, 예상 발생률의 산출법, 조기발견 기간의 통제
151 법 등에 따라 여러 산출 방법들이 생긴다[8,20,32,42,43,79]. 이렇게 다양한 연구방법들을
152 Biesheuve et al. [9]은 활용하는 역학정보에 맞추어 시점 발생률 (incidence rate), 누적 발생
153 률 (cumulative incidence), 모델링의 3가지로 분류하였다. 이 분류법에 따라 비교군의 예상
154 발생률 산출과정, 국가단위의 암등록사업의 역할, 추가적인 요구 정보를 검토한 내용을
155 Table 3으로 정리해 보았다. 이는 국내 여건상 어떤 연구가 가능한지, 앞으로 어떤 정보가
156 필요한가를 살펴보기 위한 것이다. 시점 발생률 보다는 누적 발생률을 활용하는 방법을 우
157 선 추천하고 있는 가운데[9,20], Table 3에서 비교군의 예상 발생률을 어떻게 산출할 것인가
158 에 따라 연구설계가 달라지지만 국가 암등록사업은 모든 경우에 있어서 중요한 역할을 한
159 다는 것을 알 수 있다. 또한 질병의 자연사를 반영할 수 있는 통계 모델링[9,69]을 위해서는
160 간격암(interval cancer) [69,80,81]를 포함한 질병의 자연사에 관한 정보를 확보해야만 가능
161 하다[22,35,39,53]. 이런 점들을 고려하고 암검진에 대한 무작위 시험이 없는 국내 여건을

162 감안할 때, 진단을 하게 된 계기가 검진에 의한 것인지 아니면 임상에 의한 것인지를 알 수
163 있다면 누적발생률의 코호트 분석을 우선 시도할 수 있겠다[34,72]. 이것이 어렵다면 누적발
164 생률의 사업시점 전후의 대조 분석을 할 수 있겠지만, 국가암통계가 2002년부터 가능하다
165 는 점을 감안할 때 APC 를 적용한 추정 방법 등을 강구해 볼 수 있겠다[55,82].

166 .

167

168 **3. 결론 및 제언**

169

170 과진단의 크기를 산출하는 것은 검진의 이득에 비하여 유해가 얼마나 큰가를 알아
171 내는 것으로[38,55], 보건의료 정책 결정에 중요한 근거가 된다[20,43]. 한국의 현실을 감안
172 할 때, 과진단의 크기 산출을 위해서는 한국중앙암등록본부의 데이터베이스에서 통계 결과
173 치 뿐만 아니라 개인(individual) 수준의 추적 자료를 최대한 활용하는 전략을 우선 강구할
174 수 있다[65,83]. 문제는 진단시 검진여부에 의한 것인가를 알아내는 정보가 최근어야 수집되
175 고 있다는 점[29]과, 2002년부터 공식 통계가 발표되어 왔다는 점이다[34]. 이런 한계들을
176 극복하기 위해서는 수집가능한 정보들을 최대한 사용할 수 있도록 국가연구기관에서 관련
177 연구를 주도해야 할 것이다.

178

179 연구에 활용할 자료의 타당성 확보 노력에도 불구하고, 산출된 과진단 크기를 올바
180 로 해석하기 위해서는 극복해야 할 한계들이 여전히 있다[59,75]. 예를 들어 검진 수검률에
181 따라 발생률 증가 폭의 해석이 달라질 수 있고[43], 새로운 검사법 도입에 따라 민감도 변
182 동이 생기며[54], 경구호르몬제 복용 증가나 출산수유력 변화처럼 유방암 발생 위험요인이
183 변동할 수 있고[11,55], 효과적인 치료법이 새로이 개발될 수 있다[55]. 이런 한계들 때문에,
184 유방촬영술의 과진단 산출 제시 범위는 매우 넓어[9,20] 과진단 여부에 대한 논쟁이 여전히
185 것이다[43]. 관련 논문들의 평가 (appraisal)를 위한 점검표도 제시되었지만[43], 이런 논쟁의
186 궁극적인 해결방안은 과진단을 야기시킨 기간차이 바이어스를 극복토록 하는 것이다
187 [49,54,59-61]. 따라서 상피내암 중 침윤암으로 진행하지 않는 분율[84]이나 전체 진단암 중
188 IDLE 분율[62], 부검으로 진단하게 된 분율[85] 등을 알아내는 것도 필요하다. 또한 한국중
189 앙암등록본부는 2009년도 발생자부터 Surveillance, Epidemiology, and End Results (SEER) 병

190 기 자료를 수집하고 있는데[1,82], 현재 과진단 논란이 되고 있는 갑상선, 유방, 전립선 암에
191 대항 SEER 병기에 따른 추적조사 결과를 확보할 필요가 있다. 마지막으로 과진단의 해를
192 입는 당사자는 일반 건강인이란 점에서[4,27], 검진의 득실을 올바르게 이해할 수 있는 결정조
193 력도구 (decision aids)를 개발하는 것도 같이 이루어져야 한다[58,86].

194

195

196

197

198 감사의 글

199

200 본 연구는 대한암연구재단의 2013년도 암연구지원사업 지원을 받음(No. 2013-2).

201

202 **References**

203

204 1. Jung KW, Won YJ, Kong HJ, Oh CM, Lee DH, Lee JS. Cancer statistics in Korea: incidence, mortality,
205 survival, and prevalence in 2011. *Cancer Res Treat* 2014;46:109-123.

206 2. Ahn HS, Kim HJ, Welch HG. Korea's thyroid-cancer "epidemic"--screening and overdiagnosis. *N Engl J*
207 *Med* 2014;371:1765-1767.

208 3. Lee JH, Shin SW. Overdiagnosis and screening for thyroid cancer in Korea. *Lancet* 2014;384:1848.

209 4. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;102:605-613.

210 5. Davies L, Welch HG. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck*
211 *Surg* 2014;140:317-322.

212 6. Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. *Int J Epidemiol*
213 2004;33:43-55.

214 7. Bae JM. On the benefits and harms of mammography for breast cancer screening in Korean women.
215 *Korean J Fam Pract* 2014;4;1-6. (Korean)

216 8. Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes:
217 systematic review of incidence trends. *BMJ* 2009;339:b2587.

218 9. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on
219 estimates of invasive breast cancer overdiagnosis with mammography screening: a systematic review.
220 *Lancet Oncol* 2007;8:1129-1138.

221 10. Njor SH, Olsen AH, Blichert-Toft M, Schwartz W, Vejborg I, Lynge E. Overdiagnosis in screening
222 mammography in Denmark: population based cohort study. *BMJ* 2013;346:f1064.

223 11. Jonsson H, Johansson R, Lenner P. Increased incidence of invasive breast cancer after the introduction of
224 service screening with mammography in Sweden. *Int J Cancer* 2005;117:842-847.

225 12. Glass AG, Lacey JV Jr, Carreon JD, Hoover RN. Breast cancer incidence, 1980-2006: combined roles of
226 menopausal hormone therapy, screening mammography, and estrogen receptor status. *J Natl Cancer*
227 *Inst* 2007;99:1152-1161.

228 13. Kopans DB. Arguments against mammography screening continue to be based on faulty science.

- 229 Oncologist. 2014;19:107-112.
- 230 14. Svendsen AL, Olsen AH, von Euler-Chelpin M, Lynge E. Breast cancer incidence after the introduction of
231 mammography screening: what should be expected? *Cancer* 2006;106:1883-1890.
- 232 15. Martinez-Alonso M, Vilapriño E, Marcos-Gragera R, Rue M. Breast cancer incidence and overdiagnosis
233 in Catalonia (Spain). *Breast Cancer Res* 2010;12:R58.
- 234 16. Hellquist BN, Duffy SW, Nyström L, Jonsson H. Overdiagnosis in the population-based service screening
235 programme with mammography for women aged 40 to 49 years in Sweden. *J Med Screen*
236 2012;19:14-19.
- 237 17. Boer R, Warmerdam P, de Koning H, van Oortmarssen G. Extra incidence caused by mammographic
238 screening. *Lancet* 1994;343:979.
- 239 18. Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH. Complexities in the estimation of overdiagnosis in
240 breast cancer screening. *Br J Cancer* 2008;99:1176-1178.
- 241 19. Møller B, Weedon-Fekjaer H, Hakulinen T, Tryggvadóttir L, Storm HH, Talbäck M, et al. The influence of
242 mammographic screening on national trends in breast cancer incidence. *Eur J Cancer Prev*
243 2005;14:117-128.
- 244 20. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic
245 screening for breast cancer in Europe: a literature review. *J Med Screen* 2012;19 Suppl 1:42-56.
- 246 21. Ciatto S. The overdiagnosis nightmare: a time for caution. *BMC Womens Health* 2009;9:34.
- 247 22. Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, et al. Overdiagnosis due to prostate-
248 specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst*
249 2002;94:981-990.
- 250 23. Feuer EJ, Wun LM. How much of the recent rise in breast cancer incidence can be explained by
251 increases in mammography utilization? A dynamic population model approach. *Am J Epidemiol*
252 1992;136:1423-1436.
- 253 24. Zahl PH, Jørgensen KJ, Gøtzsche PC. Overestimated lead times in cancer screening has led to substantial
254 underestimation of overdiagnosis. *Br J Cancer* 2013;109:2014-2019.
- 255 25. Miller AB. Overdiagnosis of breast cancer. *Int J Cancer* 2013;133:2511.

- 256 26. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev*
257 2006;(4):CD001877.
- 258 27. Paci E, Warwick J, Falini P, Duffy SW. Overdiagnosis in screening: is the increase in breast cancer
259 incidence rates a cause for concern? *J Med Screen* 2004;11:23-27.
- 260 28. Cervera Deval J, Sentís Crivillé M, Zulueta JJ. Overdiagnosis in cancer screening. *Radiologia* 2014. [PMID:
261 25174786]
- 262 29. Paci E, Duffy S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis and overtreatment in
263 service screening. *Breast Cancer Res* 2005;7:266-270.
- 264 30. Prorok PC, Kramer BS, Gohagan JK. Screening theory and study design: The basics. In: Kramer BS,
265 Gohagan JK, Prorok PC, editors. *Cancer Screening: Theory and Practice*. New York: Marcel
266 Dekker; 1999.p.29-53.
- 267 31. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, et al. The case for early detection. *Nat*
268 *Rev Cancer* 2003;3:243-252.
- 269 32. Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction
270 of nationwide screening: prospective cohort study. *BMJ* 2004;328:921-924.
- 271 33. Day NE. Overdiagnosis and breast cancer screening. *Breast Cancer Res* 2005;7:228-229.
- 272 34. Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled
273 trials of breast cancer screening. *Breast Cancer Res* 2005;7:230-234.
- 274 35. de Koning HJ, Draisma G, Fracheboud J, de Bruijn A. Overdiagnosis and overtreatment of breast cancer:
275 microsimulation modelling estimates based on observed screen and clinical data. *Breast Cancer Res*
276 2006;8:202.
- 277 36. Warren R, Eleti A. Overdiagnosis and overtreatment of breast cancer: is overdiagnosis an issue for
278 radiologists? *Breast Cancer Res* 2006;8:205.
- 279 37. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography
280 screening under different screening schedules: model estimates of potential benefits and harms. *Ann*
281 *Intern Med* 2009;151:738-747.
- 282 38. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, et al. Absolute numbers of lives saved and

- 283 overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening
284 Programme in England. *J Med Screen* 2010;17:25-30.
- 285 39. Seigneurin A, François O, Labarère J, Oudeville P, Monlong J, Colonna M. Overdiagnosis from non-
286 progressive cancer detected by screening mammography: stochastic simulation study with calibration
287 to population based registry data. *BMJ* 2011;343:d7017.
- 288 40. Gunsoy NB, Garcia-Closas M, Moss SM. Modelling the overdiagnosis of breast cancer due to
289 mammography screening in women aged 40 to 49 in the United Kingdom. *Breast Cancer Res*
290 2012;14:R152.
- 291 41. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening:
292 an independent review. *Lancet* 2012;380:1778-1786.
- 293 42. Yen AM, Duffy SW, Chen TH, Chen LS, Chiu SY, Fann JC, et al. Long-term incidence of breast cancer by
294 trial arm in one county of the Swedish Two-County Trial of mammographic screening. *Cancer*
295 2012;118:5728-5732.
- 296 43. Etzioni R, Gulati R, Mallinger L, Mandelblatt J. Influence of study features and methods on overdiagnosis
297 estimates in breast and prostate cancer screening. *Ann Intern Med* 2013;158:831-838.
- 298 44. Pace LE, Keating NL. A systematic assessment of benefits and risks to guide breast cancer screening
299 decisions. *JAMA* 2014;311:1327-1335.
- 300 45. Yaffe MJ, Pritchard KI. Overdiagnosing overdiagnosis. *Oncologist* 2014;19:103-106.
- 301 46. Hersch J, Barratt A, Jansen J, Houssami N, Irwig L, Jacklyn G, et al. The effect of information about
302 overdetection of breast cancer on women's decision-making about mammography screening: study
303 protocol for a randomised controlled trial. *BMJ Open* 2014;4:e004990.
- 304 47. Suzuki A, Ishida T, Ohuchi N. Controversies in breast cancer screening for women aged 40-49 years. *Jpn*
305 *J Clin Oncol* 2014;44:613-618.
- 306 48. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of
307 breast cancer screening: an independent review. *Br J Cancer* 2013;108:2205-2240.
- 308 49. Puliti D, Paci E. The other side of technology: risk of overdiagnosis of breast cancer with mammography
309 screening. *Future Oncol* 2009;5:481-491.

- 310 50. Zahl PH, Maehlen J, Welch HG. The natural history of invasive breast cancers detected by screening
311 mammography. *Arch Intern Med* 2008;168:2311-2316.
- 312 51. Dussan C, Zubor P, Fernandez M, Yabar A, Szunyogh N, Visnovsky J. Spontaneous regression of a breast
313 carcinoma: a case report. *Gynecol Obstet Invest* 2008;65:206-211.
- 314 52. Ross MB, Buzdar AU, Hortobagyi GN, Lukeman JM. Spontaneous regression of breast carcinoma: follow-
315 up report and literature review. *J Surg Oncol* 1982;19:22-24.
- 316 53. Etzioni R, Gulati R. Oversimplifying overdiagnosis. *J Gen Intern Med* 2014;29:1218-1220.
- 317 54. Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW. Overdiagnosis, sojourn time, and sensitivity in the
318 Copenhagen mammography screening program. *Breast J* 2006;12:338-342.
- 319 55. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N*
320 *Engl J Med* 2012;367:1998-2005.
- 321 56. Hurley SF, Kaldor JM. The benefits and risks of mammographic screening for breast cancer. *Epidemiol*
322 *Rev* 1992;14:101-130.
- 323 57. Gur D, Sumkin JH. Screening for early detection of breast cancer: overdiagnosis versus suboptimal
324 patient management. *Radiology* 2013;268:327-328.
- 325 58. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health
326 treatment or screening decisions. *Cochrane Database Syst Rev* 2014;1:CD001431.
- 327 59. Al-Foheidi M, Al-Mansour MM, Ibrahim EM. Breast cancer screening: review of benefits and harms, and
328 recommendations for developing and low-income countries. *Med Oncol* 2013;30:471.
- 329 60. Berry DA, Baines CJ, Baum M, Dickersin K, Fletcher SW, Gøtzsche PC, et al. Flawed inferences about
330 screening mammography's benefit based on observational data. *J Clin Oncol* 2009;27:639-640.
- 331 61. Heywang-Köbrunner SH, Hacker A, Sedlacek S. Advantages and Disadvantages of Mammography
332 Screening. *Breast Care*. 2011;6:199-207.
- 333 62. Esserman LJ, Thompson IM, Reid B, Nelson P, Ransohoff DF, Welch HG, et al. Addressing overdiagnosis
334 and overtreatment in cancer: a prescription for change. *Lancet Oncol* 2014;15:e234-e242.
- 335 63. Zahl PH, Mæhlen J. Overdiagnosis of breast cancer after 14 years of mammography screening. *Tidsskr*
336 *Nor Laegeforen* 2012;132:414-417.

- 337 64. Marmot MG. Sorting through the arguments on breast screening. *JAMA* 2013;309:2553-2554.
- 338 65. Duffy SW, Michalopoulos D, Sebuødegård S, Hofvind S. Trends in aggregate cancer incidence rates in
339 relation to screening and possible overdiagnosis: a word of caution. *J Med Screen* 2014;21:24-29.
- 340 66. Duffy SW, Parmar D. Overdiagnosis in breast cancer screening: the importance of length of observation
341 period and lead time. *Breast Cancer Res* 2013;15:R41.
- 342 67. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years
343 after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006;332:689
- 344 68. Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. *JAMA*
345 2009;302:1685-1692.
- 346 69. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting
347 overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev* 2011;33:111-
348 121.
- 349 70. Junod B, Zahl PH, Kaplan RM, Olsen J, Greenland S. An investigation of the apparent breast cancer
350 epidemic in France: screening and incidence trends in birth cohorts. *BMC Cancer* 2011;11:401.
- 351 71. Peeters PH, Verbeek AL, Straatman H, Holland R, Hendriks JH, Mravunac M, et al. Evaluation of
352 overdiagnosis of breast cancer in screening with mammography: results of the Nijmegen programme.
353 *Int J Epidemiol* 1989;18:295-299.
- 354 72. Puliti D, Miccinesi G, Zappà M, Manneschi G, Crocetti E, Paci E. Balancing harms and benefits of service
355 mammography screening programs: a cohort study. *Breast Cancer Res* 2012;14:R9.
- 356 73. Paci E, Miccinesi G, Puliti D, Baldazzi P, De Lisi V, Falcini F, et al. Estimate of overdiagnosis of breast
357 cancer due to mammography after adjustment for lead time. A service screening study in Italy. *Breast*
358 *Cancer Res* 2006;8:R68.
- 359 74. Jørgensen KJ, Zahl PH, Gøtzsche PC. Overdiagnosis in organised mammography screening in Denmark.
360 A comparative study. *BMC Womens Health* 2009;9:36.
- 361 75. Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of
362 invasive breast cancer associated with screening mammography. *Cancer Causes Control* 2010;21:275-
363 282.

- 364 76. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. J
365 Natl Cancer Inst Monogr 2006;(36):19-25.
- 366 77. Zahl PH, Jørgensen KJ, Gøtzsche PC. Lead-time models should not be used to estimate overdiagnosis in
367 cancer screening. J Gen Intern Med 2014;29:1283-1286.
- 368 78. Zahl PH, Jørgensen KJ, Maehlen J, Gøtzsche PC. Biases in estimates of overdiagnosis due to
369 mammography screening. Lancet Oncol 2008;9:199-201.
- 370 79. Puliti D, Zappa M, Miccinesi G, Falini P, Crocetti E, Paci E. An estimate of overdiagnosis 15 years after the
371 start of mammographic screening in Florence. Eur J Cancer 2009;45:3166-3171.
- 372 80. Caumo F, Vecchiato F, Strabbioli M, Zorzi M, Baracco S, Ciatto S. Interval cancers in breast cancer
373 screening: comparison of stage and biological characteristics with screen-detected cancers or incident
374 cancers in the absence of screening. Tumori 2010;96:198-201.
- 375 81. Heidinger O, Batzler WU, Krieg V, Weigel S, Biesheuvel C, Heindel W, et al. The incidence of interval
376 cancers in the German mammography screening program: results from the population-based cancer
377 registry in North Rhine-Westphalia. Dtsch Arztebl Int 2012;109:781-787.
- 378 82. Bleyer A. Were our estimates of overdiagnosis with mammography screening in the United States
379 "based on faulty science"? Oncologist 2014;19:113-126.
- 380 83. Falk RS, Hofvind S, Skaane P, Haldorsen T. Overdiagnosis among women attending a population-based
381 mammography screening program. Int J Cancer 2013;133:705-712.
- 382 84. Yen MF, Tabár L, Vitak B, Smith RA, Chen HH, Duffy SW. Quantifying the potential problem of
383 overdiagnosis of ductal carcinoma in situ in breast cancer screening. Eur J Cancer 2003;39:1746-1754.
- 384 85. Thorat MA. What caused the decline in US breast cancer incidence? Nat Clin Pract Oncol 2008;5:314-
385 315
- 386 86. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. JAMA Intern
387 Med 2014;174:448-454.

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

Author (Year) [Reference]	Sentence
Prorok (1999) [30]	the detection of cancers that would never have been found were it not for the screening test
Etzioni (2003) [31]	the detection of cancers that would otherwise not have been detected in the remaining life-span of the individuals
Zahl (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 in situ breast cancer that, in the absence of screening, would not have presented clinically during the woman's lifetime.
de Koning (2005) [35]	the detection of cancers by screening that would otherwise never have been clinically diagnosed but are now consequently treated
Paci (2005) [29]	diagnosis of cancer that would in the absence of screening not have led to clinically manifest disease in the woman's lifetime
Warren (2005) [36]	diagnosis of cancers that would not present within the life of the patient and is one of the downsides of screening
Biesheuvel (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 (2009) [37]	the proportion of cases in each strategy that would not have clinically surfaced in a woman's life time among all cases
Duffy (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 (2010) [4]	the diagnosis of a "cancer" that would otherwise not go on to cause symptoms or death
Seigneurin (2011) [39]	the proportion of non-progressive cancers among all cases of invasive cancer and carcinoma in situ
Gunsoy (2012) [40]	the diagnosis of screen detected cancers that would not have presented clinically in a lifetime in the absence of screening

Hellquist (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 (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 (2012) [42]	the diagnosis of cancer as a result of screening that would not have arisen in the lifetime of the host
Etzioni (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 (2013) [24]	the detection of lesions that would never have been detected in a person's lifetime in the absence of screening
Cervera (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 (2014) [44]	the detection of a tumor through screening that would not have become clinically evident in the absence of screening

Epidemiology and Health

Table 2. Decision criteria of identifying overdiagnosis

	Qualitative	Quantitative
Randomized trials	No compensatory drop after stopping screening program	Excess cases per screen-detected cases (Fig 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

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

Categories	Study Design	Needed 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 in situ	39

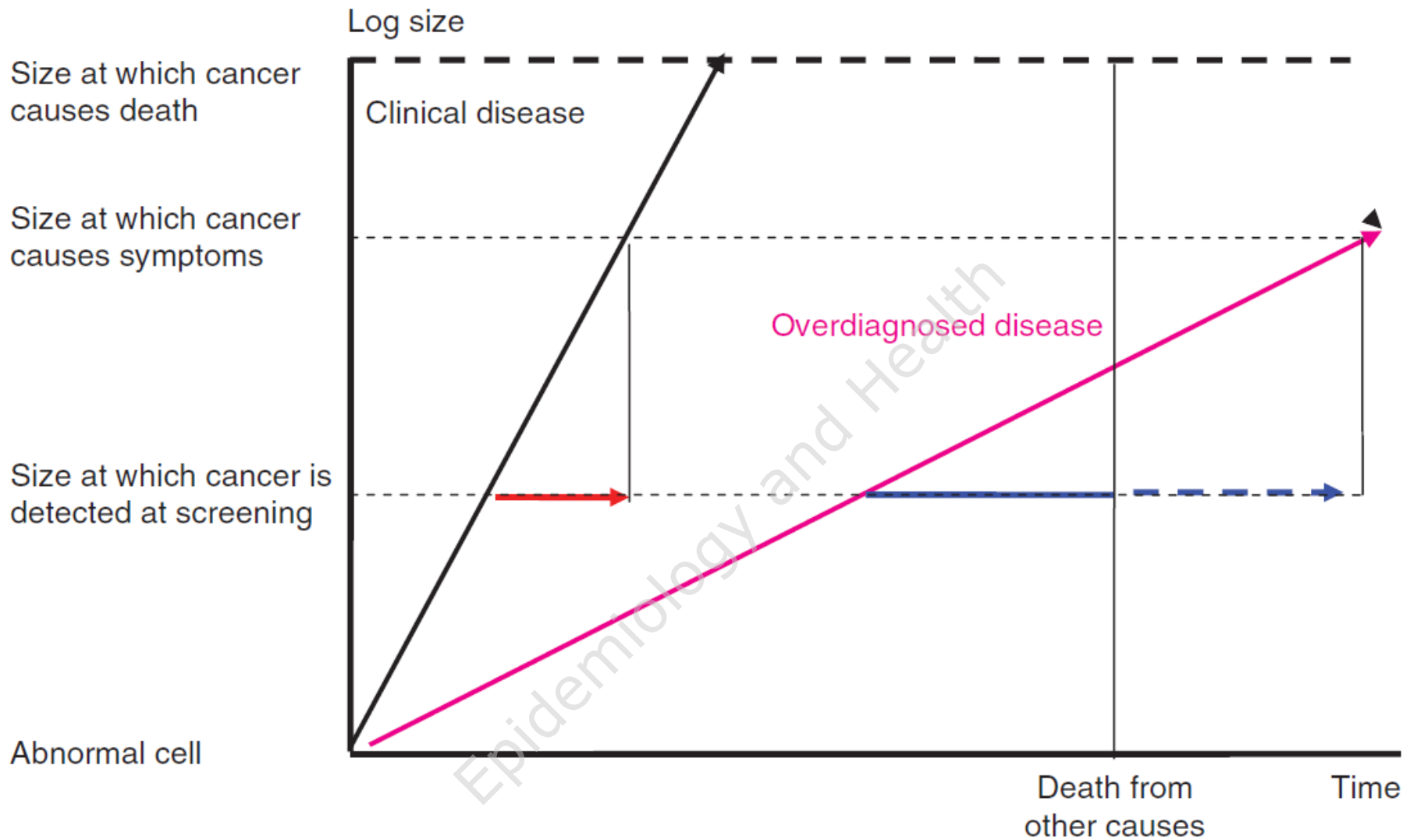


Fig 1. Two different types of tumor growth indicated by the black (clinical disease) and pink arrows (overdiagnosed disease), respectively. Clinically relevant lead time is indicated by the red arrow. When overdiagnosed tumors are included, some researchers have defined lead time as time to death (solid blue arrow, so-called censored lead time), other have 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). Cited from reference [24] under permission.

$$d = (As + B) - (An + C)$$

$$\text{Overdiagnosis}(\%) = d/B$$

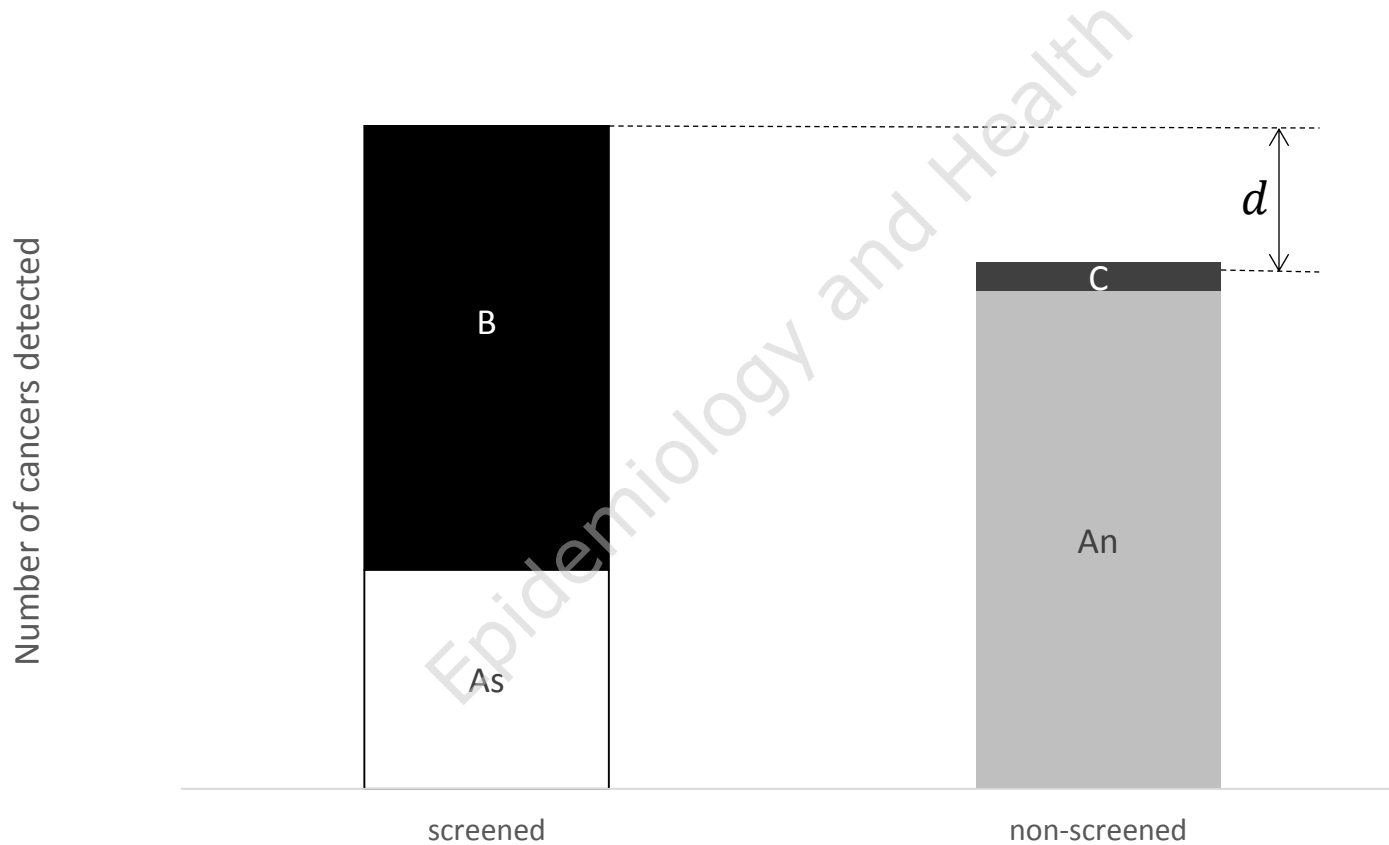


Fig 2. Size of overdiagnosis in randomized trial. (As: clinically detected cancers in screened group; An: clinically detected cancers in non-screened group; B: screen-detected cancers in screened group; C: catch-up cancers of post-screening periods in non-screened group; d: extra cancers). Modified from reference [4] with permission.