癌症筛查中过度诊断的评测方法

夏昌发，陈万青
（国家癌症中心/ 国家肿瘤临床医学研究中心/ 中国医学科学院北京协和医学院肿瘤医院，北京 100021）

摘　要：过度诊断是评价癌症筛查利弊的关键问题，但对其准确评测存在一定难度，国内相关研究也十分有限。全文对评测过度诊断情况的指标进行了定义，全面介绍了国际上用于癌症筛查过度诊断情况评测的三类主要方法，结合相关研究设计总结了六种设置未筛查对照的方法，对比分析了各种方法的数据要求、应用条件、优势及缺点，提供了经典案例和应用建议，以期为正确评测筛查中的过度诊断情况提供方法思路。

关键词：癌症；筛查；过度诊断；方法学
中图分类号：R730.1；R73-31　　**文献标识码**：A　　**文章编号**：1004-0242(2022)03-0161-08
doi：10.11735/j.issn.1004-0242.2022.03.A001

Methods for Quantifying Overdiagnosis in Cancer Screening

XIA Chang-fa, CHEN Wan-qing
(National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China)

Abstract: Overdiagnosis is a key issue in evaluating the benefits and harms of cancer screening. However, there is a lack of studies on overdiagnosis in cancer screening from China, as it is difficult to quantify the overdiagnosis accurately. This article gives the definition of metric for measuring overdiagnosis in cancer screening; comprehensively introduces three major approaches used to quantify overdiagnosis internationally; and presents six methods to set unscreened control population according to study design. Data requirements, application conditions, strengths and limitations of above methods are compared and analyzed. The article also provides some reference cases and suggestions to guide the good practices for choosing appropriate methods to obtain the unbiased estimates of overdiagnosis from cancer screening.

Key words: cancer; screening; overdiagnosis; methodology

癌症是我国居民的首位死亡原因[1]。筛查和早诊早治是降低癌症死亡负担的有效方法[2]，近年来我国逐渐扩大癌症筛查和早诊早治的覆盖人群[3]。癌症筛查的目标是发现早期癌症或癌前病变，通过早诊早治提升患者预后、降低癌症负担。成功的筛查可以预防处于临床前期的癌症，但同时也可能将一部分终其一生也不会自然发病的潜在癌症提前探测出来，给“患者”造成过度诊断和后续的过度治疗[4]。过度诊断是癌症筛查带来的最主要危害之一[5]，科学评测筛查带来的过度诊断对合理权衡筛查的利弊至关重要[6]。

对癌症过度诊断的直接测量十分困难，因为尚无任何技术能在筛查阳性病例发现之时就确定该病例是否属于过度诊断，而是需要经过长期随访或采用对照比较的方法来间接确定[7]。目前国际上对癌症筛查的过度诊断问题已进行较为深入的研究，并建立起多种评测方法来量化过度诊断的情况。为促进我国全面认识癌症筛查的利与弊，本文对评测癌症筛查过度诊断的方法学进行系统总结和介绍，以期为我国的癌症筛查评估工作提供思路。

1　过度诊断的评测指标

一般采用过度诊断率来衡量癌症筛查中过度诊

收稿日期：2021-12-09；修回日期：2022-01-04
基金项目：国家自然科学基金(81974492)；京津冀基础研究合作专项(J200017)；深圳医疗卫生三名工程(SZSM201911015)；中国医学科学院中央级公益性科研院所基本科研业务费专项资金(2019PT320027)
通信作者：陈万青，E-mail：chenwq@cicams.ac.cn

断的情况。过度诊断率的分子通常被定义为：通过筛查发现并经组织病理确诊的、在患者死亡之前都不会被自然发现的癌症病例数(Figure 1)。而过度诊断率的分母的定义则有多种形式，如：一段时期内由筛查发现的癌症病例总数、一段时期内癌症累积发病数(包括筛查发现的或临床发现的)、接受筛查的总人数、符合筛查的总人数、受邀参与筛查的总人数等[4]。其中，将分母定义为一段时间内由筛查发现的癌症病例总数是最常见的做法[8]，即：

$$过度诊断率(\%)=\frac{筛查发现并经病理确诊的、在患者死亡之前都不会被自然发现的癌症病例数}{一段时间内由筛查发现的癌症病例总数}\times 100\%$$

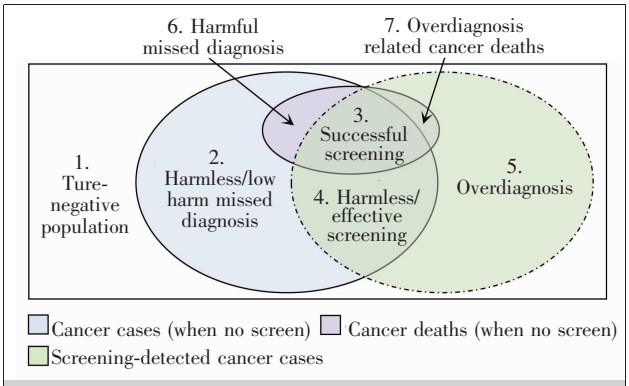

Figure 1　Venn diagram illustrating the overdiagnosis in cancer screening

另外，也有一些研究会报告衡量癌症筛查过度诊断情况的次要指标，如过度诊断例数与每预防1例死亡所需筛查人数的比值[9]和相对发病风险[10]等。

2　过度诊断的评估方法

目前已经建立起多种方法体系来评测癌症筛查项目中的过度诊断情况，它可以基于多种研究设计类型，如随机对照试验、队列和病例对照研究、生态学研究、模型研究、病理和影像学研究等[8,11]。而按评测思路和方法学类型，则可以将各种研究分为估算超额发病率、调整领先时间、构建自然史模型等三大类(Table 1)[4,9,12-20]。

2.1　超额发病率法

超额发病率法的核心思想是将筛查人群发现的癌症病例数与未筛查人群中的癌症发病数进行比较，用差值作为因为筛查过度诊断癌症病例数的一种近似替代[4,21-22]。可见，超额发病率法的前提假设是：在较长一段评测时期内(大于领先时间)，如果筛查的过度诊断率为0，则人群中的癌症累积发病率不会因为筛查而增加，即筛查人群和未筛查人群的累积发病率是相同的，筛查不过是将人群中陆续自然发病的癌症病例提前集中发现而已。

如图2(Figure 2)所示，假设人群自然发病率(无筛查情况下的发病率)恒定为100/10万，癌症筛查的领先时间为2年，共设置4种筛查和过度诊断的情景：情景1和情景2均进行一次性筛查，其中情景1过度诊断率为0，情景2过度诊断率为20%；情景3和情景4均进行持续性筛查，其中情景3过度诊断率为0，情景4过度诊断率为20%。筛查开始后，由于领先时间效应，发病率会形成一个高峰；而后，在一段时期内会形成一个发病率的低谷，因为这一时期内的部分病例已经被之前的筛查提前数年发现了。显然，领先时间的长短决定了筛查后发病率高峰与低谷的时间间隔，以及发病率高峰的大小。即使在持续筛查的情景下，如果不存在过度诊断，人群发病率最终会回归到筛查前的水平(Figure 2 Scenario 3)；而存在过度诊断时，人群发病率最终会高于筛查前的水平(Figure 2 Scenario 4)。

使用超额发病率法评测过度诊断率需要足够长的随访时间，因为该方法为了避免对领先时间进行估计，需要通过长期随访来释放领先时间效应[22]。对于有停止时间的筛查项目，如评价筛查技术的临床试验、基于出生队列而设置的停止筛查年龄等，需要筛查停止后的随访时间要足以超过最大领先时间；例如，在图2情景2中(Figure 2 Scenario 2)，筛查停止后2年以上评估过度诊断率才能得到正确的结果，即5/25=20%。对于持续筛查的项目，对过度诊断率的评测需要去除筛查开始后发病率经历高峰和低谷时期(领先时间效应期)，即在发病率稳定后再计算过度诊断率；例如，在图2情景4中(Figure 2 Scenario 4)，使用第1~6年的筛查情况计算的过度诊断率是45/105=42.9%，用第5~6年筛查结果计算的真实过度诊断率应该是6/30=20%。

2.2　领先时间调整法

领先时间调整法直接调整筛查人群中的领先时间，并将调整后的发病率与未筛查人群的发病率进行比较，进而得到过度诊断率[22]。而在实际筛查中，领先时间无法精确测量，只能通过各种方式进行近似推算[4]。按照对领先时间的近似处理方式，领先时

Table 1 Methods to quantify overdiagnosis in cancer screening

Item	Excess incidence approach	Lead time approach		Natural history modelling
		Average lead time	Lead time distribution	
Data required	Cancer incidence in screened population Cancer incidence in unscreened population	Cancer incidence in screened population Cancer incidence in unscreened population Mean lead time estimate	Cancer incidence in screened population Cancer incidence in unscreened population Distribution of lead time	Age-specific cancer incidence in unscreened population Competing mortality (or approximate by all-cause mortality) Clinical data (stage distribution) Screening performance (sensitivity and specificity by stage)
Application conditions	Long-term follow-up after cessation of screening (≥ lead time) or the screening strategy remains unchanged for a long time Excess and drop need to be estimated in individuals with equal participation rates and risk of cancer No opportunistic screening (no contamination)	Lead-time is a fixed estimate Lead time is estimated from symptomatic cancers First screening round (i.e. prevalent round) is often excluded from analyses, because the incidence of first round is not steady	Exponential distribution of lead time Lead time is estimated from symptomatic cancers No competing mortality	Define model structure on natural history and screen detectability Model needs to be calibrated using epidemiological data
Applicable cancers	Screenings that focus on detecting early cancer (e.g. lung, breast and prostate cancer)	Screenings that focus on detecting early cancer (e.g. lung, breast and prostate cancer)	Screenings that focus on detecting early cancer (e.g. lung, breast and prostate cancer)	All cancer screenings, especially those focus on detecting precancerous lesions (e.g. cervix and colon cancer)
Strengths	Proxy assessment, i.e. does not require estimate or distribution of lead-time	Direct assessment, i.e. does not require long-term follow-up	Direct assessment, i.e. does not require long-term follow-up	Parameters from different sources can be synthesised Screening scenarios can be set flexibly
Limitations	Needs long follow-up (while the exactly time not known)	Approximate estimates of natural history (lead time) are needed Underestimation due to the exclusion of the first screening round	Approximate estimates of natural history (lead time) are needed	Approximate estimates of natural history are needed Modelis complex and may lack transparency Many parameters and uncertainties may accumulate
Reference cases	Patz, 2014[9]; Miller, 2014[12]	Zackrisson, 2006[13]; Kalager, 2012[14]	Paci, 2006[15]; Beckmann, 2015[16]	Seigneurin, 2011[17]; Ten, 2020[18]; Brenner, 2015[19]; Sroczynski, 2020[20]

间调整法主要分为平均领先时间法、领先时间数学分布法等两种。

平均领先时间法将未筛查人群的发病率向前平移一段时间后再与筛查人群的发病率进行比较，平移量等于平均领先时间。由于筛查初期会形成一个发病率的高峰(提前集中发现了潜伏病例)，因此应将首次参与筛查者排除在外。实际操作中往往难以将首次参与筛查者识别并排除在外，作为一种替代，通常会在发病率稳定后再计算过度诊断率[22]。由于要排除首次参与筛查者，导致平均领先时间法会低估过度诊断率。

领先时间数学分布法假设领先时间不是一个固定值，而是服从指数分布。与平均领先时间法类似，通过调整领先时间，再比较筛查人群和未筛查人群的发病率，进而得到过度诊断率。显然，对领先时间的假设难以验证，因此领先时间数学分布法对过度诊断率的估计依赖对领先时间所做的假设。

2.3 自然史模型模拟法

使用自然史模型评测过度诊断率的方法较为复杂，它通过数学模型模拟整个癌症发生发展、筛查诊

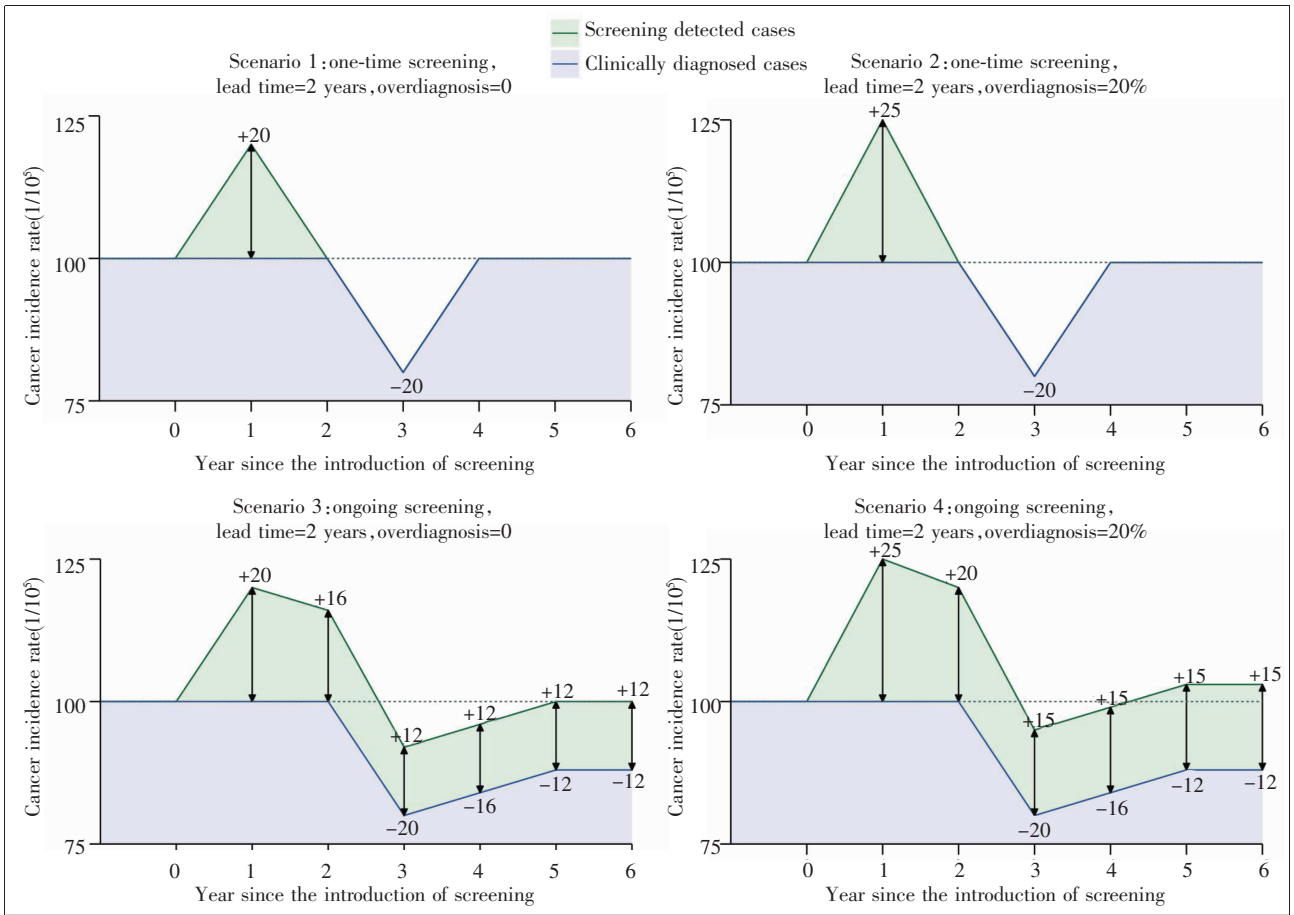

Figure 2　Schematic diagram illustrating the excess incidence approach(adapted from Etzioni,2013[21])

断的全过程,模拟对象可以是人群(静态固定模型),也可以是个体(微观模拟模型)。使用校准后的自然史模型，可以灵活模拟接受筛查和未接受筛查的人群的发病率,并推算领先时间。得到模拟结果后,既可以采用直接法调整领先时间，也可以采用超额发病率法计算过度诊断率。和领先时间一样,许多癌症的自然史难以观测,并且可能存在地区和种族差异[4],因此自然史模型必须经过流行病学数据全方位校准(如年龄别发病率、分期构成等)才能使用。即使如此，校准后自然史模型也可能与真实人群的情况存在差距，因为模型使用的大量参数间可能存在潜在的依赖性，不同的模型参数组合可能都给出与现有流行病学数据相吻合的结果。

3　未筛查对照的设置方法

无论使用何种方法评测癌症筛查的过度诊断情况，都需要准确衡量或预测筛查人群的癌症自然发病率(即无筛查时的发病率)[21]。自然史模型通过模拟人群不接受筛查的情景，可直接获得无筛查时癌症的自然发病率；而超额发病率法和领先时间调整法则需要设置未筛查对照来获得自然发病率。不同研究设计有不同的设置未筛查对照的方法[22]:随机对照试验的对照组是未筛查对照的最佳选择；队列研究和病例对照研究可使用未参与筛查者作为未筛查对照;生态学研究则可使用未筛查地区、未筛查前的趋势、非筛查目标人群和年龄-时期-队列(APC)模型调整筛查效应作为未筛查对照。

设置未筛查对照的各种方法都需要满足一定的应用条件，同时也具有一定的优势和缺点(Table 2)[9-10,12-14,23-30]。随机对照试验的对照组是衡量自然发病率的最佳选择，但与其他设置对照的方法一样,仍存在沾染的风险,即对照组人群有可能参与试验外机会性筛查，因为筛查试验往往无法实现对筛查对象的盲法处理。生态学研究可以较为灵活地设置未筛查对照，但是影响人群的自然发病率的因

Table 2 Methods to obtain unscreened reference population

Item	Randomized controlled trial	Cohort and case-control study	Ecological study			
Methods for control	Control group	Non attenders	Control region	Pre-screening trend	Not invited	APC model
Screened population	Intervention group	Persons actually screened (ever screened)	Persons invited for screening in screened region	Persons invited for screening	Persons invited for screening	Cancer incidence in screened population
Unscreened population	Control group	Non attenders (never screened)	Persons in unscreened region	Pre-screening trends extrapolate to the screening period	Persons not invited to screening (non-target birth cohorts or ages)	Expected cancer incidence in screened population by adjusting the screening effect
Application conditions	Individual or comparable cluster randomization No contamination and high adherence Long-term follow-up	No selection biases Long-term follow-up	Screened and unscreened regions are comparable No opportunistic screening in the unscreened region (no contamination)	Cancer incidence remains unchanged if no screening Diagnostic ability remains unchanged	No opportunistic screening in the unscreened population (no contamination) Cancer incidence in reference cohorts or ages are (or could adjust to) comparable	Cancer incidence in screened and unscreened population differs only by the effect of screening Screening participation Effect of different phases of screening on cancer incidence is constant, otherwise use phases of screening as an explaining variable
Strengths	Based on individual data RCT design represents the highest chance for a comparable reference	Based on individual data	Optional adjust for time trends simultaneously	Easy to perform	Easy to perform	Cancer incidence in screened and unscreened population differs only by the effect of screening
Limitations	Limited generalizability Screening contamination and non adherence Long-term follow-up data not available for most trials	Potential for selection bias Due to the lead time, person-years of screened population are slightly underestimated thus overdiagnosis may be slightly overestimated	Unadjusted factors may differ between screened and unscreened regions No longer possible to find an absolute unscreened region	Projected cancer incidence in unscreened population depends onmodel used Uncertainty in extrapolation of the pre-screening trend	Different birth cohorts may incomparable Potential for opportunistic screening (screening contamination)	Without consideration of many other contributed factors, the model is too simple to explain the cancer incidence, model estimates may not be robust
Reference cases	Patz, 2014[9]; Miller, 2014[12]; Zackrisson, 2006[13]	Johns, 2017[25]; Falk, 2013[26]	Njor, 2013[10]; Kalager, 2012[14]	Park, 2016[27]; Jørgensen, 2009[28]	Zahl, 2004[29]; Autier, 2017[30]	Ripping, 2015[23]; Heinävaara, 2014[24]

素复杂，可能存在地区差异，也可能随时间的变化而改变，即使是同一出生队列（在考虑了领先时间带来的检出年龄差距后）其发病风险和筛查参与率亦可能随年龄增加而改变，因此生态学研究设置未筛查对照时要格外注意控制偏倚。APC 模型法本质上也是使用未筛查人群（未筛查的出生队列或年龄）作为对照，通过广义线性模型调整筛查效应而获得期望发病率（相当于筛查人群的自然发病率）[23-24]。

4 小结与建议

癌症筛查的过度诊断情况通常用过度诊断率来衡量，常被定义为通过筛查确诊的、在患者死亡之前都不会被自然发现的癌症病例数占一段时间内由筛查发现的癌症病例总数的比例。用于癌症筛查过度诊断的评测方法可以分为三大类，分别是超额发病率法、领先时间调整法和自然史模型模拟法。除自然史模型模拟外，超额发病率法和领先时间调整法都需要设置未筛查对照来衡量或预测筛查人群的癌症自然发病率，不同的研究设计可以采用不同方法来合理设置未筛查对照。

三类评测过度诊断的方法各有优劣。领先时间调整法不需要停止筛查并进行长期随访，但是它依赖于对领先时间和对自然史的近似估计与假设，可能因此对过度诊断率的估计带来偏差。而超额发病率法则是一种理论上更为稳健的方法，因为它不需要对领先时间进行假设，但是它却需要长期随访并要求把握合适的评测时期。自然史模型模拟法虽然具有融合参数、灵活评估的优势，但是该方法较为复杂，其结果的不确定性使其往往被归为证据级别较低的一类方法。可见，各种评测方法对材料都有不同的要求，都存在一定的应用条件，拥有某些相对优势的同时也存在各种各样的局限性。目前还不存在用于过度诊断率评测的最佳方法[22]，因此，对筛查中过度诊断的评测有必要采用多种方法进行相互印证和补充。

除上述三种评测方法外，也有一些其他方法来间接或替代评测癌症筛查的过度诊断情况[8,22]。尽管这些方法有一些应用的例子，但其方法原理和可靠性仍存在巨大争议，因此本文未对其进行介绍。例如，采用体积倍增时间的长短人为定义过度诊断[31-32]、将十分早期的病例全部定为过度诊断[33-35]、将早期病例相对于晚期病例的超额发病定义为过度诊断[36-37]、将预期寿命超过人群预期寿命的病例全部判定为过度诊断[38]等。

适用于不同癌症的过度诊断评测方法有所不同。三类评测方法都适用于旨在发现早期癌症的筛查，如肺癌、乳腺癌、前列腺癌等；而对于旨在发现癌前病变的筛查，如宫颈癌和结直肠癌，则只能采用自然史模型模拟的方法[22]。由于自然史模型不用通过设置未筛查对照来获得自然发病率，因此对于那些无法获得有效未筛查对照的筛查项目或筛查技术，自然史模型模拟法也是唯一的选择。同种评测方法的优势与缺点对不同癌症的影响也不同，这与各种癌症本身的自然史和各种筛查所致的领先时间长短有关，例如，较短的随访时间对于肺癌等领先时间较短的癌症影响较小，而对于前列腺癌等领先时间较长的癌症则影响较大。

影响过度诊断评测可靠性的因素不仅包括研究设计类型，也包括评测方法和未筛查对照的设置方式[22]。即使是设计良好、长期随访的随机对照试验，如果不能准确计算超额发病率或不能合理调整领先时间，其准确性可能还不如分析方法合理的生态学研究。

癌症筛查是一把双刃剑，成功的筛查能降低癌症对生命健康的威胁，起到挽救生命、提高生命质量的作用；同时，筛查也可能带来过度诊断及后续的过度治疗问题，给“患者”带来不必要的身心健康损失及不必要的经济负担[5]。因此，对筛查项目进行评估时，除了关注筛查的有效性和安全性之外，也要将过度诊断问题放在重要的位置加以考虑，卫生经济学成本效益评估中也应将过度诊断带来的成本损失和健康损害考虑在内。全面客观地对筛查项目的利与弊进行卫生技术评估才能支撑卫生政策合理决策、支持筛查参与者做出最佳决定。

参考文献：

[1] 国家卫生健康委员会. 2020 中国卫生健康统计年鉴[M]. 北京：中国协和医科大学出版社，2020：280-281.
National Health Commission. 2020 China health statistics yearbook [M]. Beijing: China Peking Union Medical University Press，2020：280-281.

[2] Wender RC,Brawley OW,Fedewa SA,et al. A blueprint for cancer screening and early detection: advancing screening's contribution to cancer control[J]. CA Cancer J Clin,2019,69(1):50-79.

[3] 李霓,李江,陈万青,等. 癌症筛查指南及共识质量评价研究进展[J]. 中华流行病学杂志,2021,42(2):211-214.
Li N,Li J,Chen WQ,et al. Research progress of quality assessment of cancer screening guidelines and consensus [J]. Chinese Journal of Epidemiology,2021,42(2):211-214.

[4] Davies L,Petitti DB,Martin L,et al. Defining,estimating, and communicating overdiagnosis in cancer screening[J]. Ann Intern Med,2018,169(1):36-43.

[5] Gilbert N. The pros and cons of screening [J]. Nature, 2020,579(7800):S2-S4.

[6] Elmore JG. Solving the problem of overdiagnosis[J]. N Engl J Med,2016,375(15):1483-1486.

[7] Srivastava S,Koay EJ,Borowsky AD,et al. Cancer overdiagnosis: a biological challenge and clinical dilemma[J]. Nat Rev Cancer,2019,19(6):349-358.

[8] Carter JL,Coletti RJ,Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods[J]. BMJ,2015,350:g7773.

[9] Patz EF Jr,Pinsky P,Gatsonis C,et al. Overdiagnosis in low-dose computed tomography screening for lung cancer [J]. JAMA Intern Med,2014,174(2):269-274.

[10] Njor SH,Olsen AH,Blichert-Toft M,et al. Overdiagnosis in screening mammography in Denmark: population based cohort study[J]. BMJ,2013,346:f1064.

[11] Gulati R,Feuer EJ,Etzioni R. Conditions for valid empirical estimates of cancer overdiagnosis in randomized trials and population studies[J]. Am J Epidemiol,2016,184(2): 140-147.

[12] Miller AB,Wall C,Baines CJ,et al. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial[J]. BMJ,2014,348:g366.

[13] Zackrisson S,Andersson I,Janzon L,et al. Rate of overdiagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study [J]. BMJ, 2006,332(7543):689-692.

[14] Kalager M,Adami HO,Bretthauer M,et al. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program[J]. Ann Intern Med,2012,156(7):491-499.

[15] Paci E,Miccinesi G,Puliti D,et al. Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy[J]. Breast Cancer Res,2006,8(6):R68.

[16] Beckmann K,Duffy SW,Lynch J,et al. Estimates of overdiagnosis of breast cancer due to population-based mammography screening in South Australia after adjustment for lead time effects[J]. J Med Screen,2015,22(3):127-135.

[17] Seigneurin A,François O,Labarère J,et al. Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data[J]. BMJ,2011,343:d7017.

[18] Ten Haaf K,Bastani M,Cao P,et al. A comparative modeling analysis of risk-based lung cancer screening strategies[J]. J Natl Cancer Inst,2020,112(5):466-479.

[19] Brenner H,Altenhofen L,Stock C,et al. Prevention,early detection,and overdiagnosis of colorectal cancer within 10 years of screening colonoscopy in Germany [J]. Clin Gastroenterol Hepatol,2015,13(4):717-723.

[20] Sroczynski G,Esteban E,Widschwendter A,et al. Reducing overtreatment associated with overdiagnosis in cervical cancer screening—a model-based benefit-harm analysis for Austria[J]. Int J Cancer,2020,147(4):1131-1142.

[21] Etzioni R,Gulati R,Mallinger L,et al. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening[J]. Ann Intern Med,2013, 158(11):831-838.

[22] Ripping TM,Ten Haaf K,Verbeek ALM,et al. Quantifying overdiagnosis in cancer screening: a systematic review to evaluate the methodology[J]. J Natl Cancer Inst,2017,109 (10):djx060.

[23] Ripping TM,Verbeek AL,Fracheboud J,et al. Overdiagnosis by mammographic screening for breast cancer studied in birth cohorts in the Netherlands [J]. Int J Cancer, 2015,137(4):921-929.

[24] Heinävaara S,Sarkeala T,Anttila A. Overdiagnosis due to breast cancer screening: updated estimates of the Helsinki service study in Finland[J]. Br J Cancer,2014,111(7): 1463-1468.

[25] Johns LE,Coleman DA,Swerdlow AJ,et al. Effect of population breast screening on breast cancer mortality up to 2005 in England and Wales: an individual-level cohort study[J]. Br J Cancer,2017,116(2):246-252.

[26] Falk RS,Hofvind S,Skaane P,et al. Overdiagnosis among women attending a population-based mammography screening program[J]. Int J Cancer,2013,133(3):705-712.

[27] Park S,Oh CM,Cho H,et al. Association between screening and the thyroid cancer "epidemic" in South Korea:

evidence from a nationwide study[J]. BMJ,2016,355:i5745.

[28] Jørgensen KJ,Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends[J]. BMJ,2009,339:b2587.

[29] Zahl PH,Strand BH,Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study[J]. BMJ,2004,328(7445):921-924.

[30] Autier P,Boniol M,Koechlin A,et al. Effectiveness of and overdiagnosis from mammography screening in the Netherlands: population based study[J]. BMJ,2017,359:j5224.

[31] Veronesi G,Maisonneuve P,Bellomi M,et al. Estimating overdiagnosis in low-dose computed tomography screening for lung cancer: a cohort study[J]. Ann Intern Med,2012,157(11):776-784.

[32] Lindell RM,Hartman TE,Swensen SJ,et al. Five-year lung cancer screening experience: CT appearance,growth rate,location,and histologic features of 61 lung cancers[J]. Radiology,2007,242(2):555-562.

[33] Graif T,Loeb S,Roehl KA,et al. Under diagnosis and over diagnosis of prostate cancer[J]. J Urol,2007,178(1):88-92.

[34] Pelzer AE,Colleselli D,Bektic J,et al. Clinical and pathological features of screen vs non-screen-detected prostate cancers: is there a difference?[J]. BJU Int,2008,102(1):24-27.

[35] Malila N,Leinonen M,Kotaniemi-Talonen L,et al. The HPV test has similar sensitivity but more overdiagnosis than the Pap test—a randomised health services study on cervical cancer screening in Finland [J]. Int J Cancer,2013,132(9):2141-2147.

[36] Bleyer A,Welch HG. Effect of three decades of screening mammography on breast-cancer incidence [J]. N Engl J Med,2012,367(21):1998-2005.

[37] Welch HG,Prorok PC,O'Malley AJ,et al. Breast-cancer tumor size,overdiagnosis,and mammography screening effectiveness[J]. N Engl J Med,2016,375(15):1438-1447.

[38] Sone S,Nakayama T,Honda T,et al. Long-term follow-up study of a population-based 1996—1998 mass screening programme for lung cancer using mobile low-dose spiral computed tomography[J]. Lung Cancer,2007,58(3):329-341.