Review Article on Implementation of CT-based Screening of Lung Cancer

Overdiagnosis in lung cancer screening

David F. Yankelevitz, Claudia I. Henschke

Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Contributions: (I) Conception and design: DF Yankelevitz; (II) Administrative support: CI Henschke; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: David F. Yankelevitz, MD. Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
Email: DFYank@gmail.com.

Abstract: There have been hundreds of publications about overdiagnosis during the past decade describing concern regarding its potential for harms in lung cancer screening. However, the definition and frequency with which it occurs in screening trials remains unclear. This lack of clarity stems from its current definition which is not based on a clinical grounds but instead on an epidemiologic definition that depends on the cause of death. Thus, with the current definition an overdiagnosed cancer can only be diagnosed if the person does not die from the cancer, regardless of whether or not the cancer is aggressive or the treatment successful. Using a definition based on epidemiology rather than the clinical presentation is highly unusual. Furthermore, the frequency of overdiagnosis has also been a cause of great confusion. Prior to the results from the National Lung Screening trial (NLST), concerns were expressed that virtually all CT screen detected cancers would be overdiagnosed, yet the extended follow-up study of the National Lung Screening Trial shows that in essence there were virtually no overdiagnosis. Even more confusing is that it was previously suspected that there was a high rate of overdiagnosis when using chest radiographic screening and therefore as CT is a more sensitive imaging test and finds cancers even earlier, it would be presumed that the overdiagnosis rate for CT would be even be higher. A proposed change in the definition would focus more on the clinical manifestation of the cancer as to its aggressiveness as this can be diagnosed while the patient is alive. Using a definition that is based on clinical features, a cancer that manifests as a nonsolid nodule would be considered overdiagnosed if instead of being recognized as relatively indolent was instead thought to be an aggressive cancer. The concept of overtreatment arises if this nonaggressive cancer were treated aggressively.

Keywords: Overdiagnosis; screening harms; lung cancer

Submitted Jun 12, 2020. Accepted for publication Nov 05, 2020.
doi: 10.21037/tlcr-20-736
View this article at: http://dx.doi.org/10.21037/tlcr-20-736

Overdiagnosis of lung cancer and concerns about it have become an important topic in lung cancer screening. A PUBMED search targeting the key words "lung cancer" and "overdiagnosis" during the decade of 2010–2019 found nearly 400 published articles in total, an average of 41 articles per year. Similar numbers of publications are occurring for other cancers showing that the topic is of great concern. In the United States, the Centers for Medicare and Medicaid (CMS) require a shared decision-making session between the physician prescribing the screening and the person who is to be screened (1). During this conversation, potential risks and benefits are described so that the person can decide, along with their health care provider, whether to enroll in a screening program. One of the explicit requirements of CMS in regard to explaining harms is a discussion about overdiagnosis.

Since, much is being written about this topic and it is a required discussion point prior to enrolling in a screening program in the United States, it becomes critical to understand overdiagnosis, that is, what is its definition and how frequently does it occur?

The "NCI Dictionary of Cancer Terms" defines

overdiagnosis as follows (2):

- Finding cases of cancer with a screening test (such as a mammogram or PSA test) that will never cause any symptoms. These cancers may just stop growing or go away on their own. Some of the harms caused by overdiagnosis are anxiety and having treatments that are not needed.

In the article "Overdiagnosis in Cancer" (3), Welch provides two definitions for overdiagnosis: "(I) The cancer never progresses (or, in fact, regresses) or (II) the cancer progresses slowly enough that the patient dies of other causes before the cancer becomes symptomatic." Within that second explanation he also allows for the following, "[E]ven a rapidly growing cancer may still represent overdiagnosis if detected when it is very small or in a patient with limited life expectancy" (3).

Based on these definitions, it is clear that overdiagnosis does not depend on the aggressiveness or any other feature of the cancer but rather whether or not the cancer causes death. An indolent cancer which either does not progress or progresses slowly can be a cause of overdiagnosis, but so can an aggressive cancer which is rapidly progressing but the person dies from an unrelated cause. According to these definitions, a cancer is not a genuine cancer unless the person dies from it, if death is from another cause it represents overdiagnosis. Even more to the point, the definition of overdiagnosis is not clinically defined, not even based on pathology. Instead it is an epidemiologic definition even though it is generally viewed as a clinical finding. This is unique in medicine. Nowhere else is a clinical diagnosis defined by the cause of death regardless of the actual cause. Welch explicitly makes this point as follows (3):

- The conundrum in overdiagnosis is that clinicians can never know who is overdiagnosed at the time of cancer diagnosis. Instead, overdiagnosis can only be identified in an individual if that individual (I) is never treated and (II) goes on to die from some other cause.

In reviewing the current conceptualization of overdiagnosis, Miettinen, often described as the "father of modern epidemiology" stated the following in his book on Epidemiological Research: Terms and Concepts (4):

- Recently, various critics of screening for a cancer have adduced a very different concept of overdiagnosis: rule-in diagnosis about a latent, preclinical case which never will become patent/ overt/clinical on account of death from some other cause. This represents the epitome of malformed concept.

To demonstrate the challenge with the current formulation of overdiagnosis, consider the following. Imagine a set of identical twins both of whom have coronary artery disease and both develop lung cancer. One of the twins foregoes any treatment of his coronary artery disease and dies from a heart attack. The other twin treats his coronary artery disease by having stents placed and avoids having a heart attack but the lung cancer progresses and eventually causes death. In this example, the cancers are identical, life threatening, yet for the twin who foregoes treatment of his comorbidity the cancer represents a case of overdiagnosis whereas for the other twin it is a genuine cancer. From a clinical perspective, the diagnoses were the same, lung cancer; from an epidemiologic perspective, one of them represents overdiagnosis. Given this understanding of the actual definition of overdiagnosis, even aggressive, advanced stage small cell carcinoma could represent a case of overdiagnosis. Nevertheless, there persists this general belief that overdiagnosis can only represent an indolent form of cancer.

Overdiagnosis has been described as "a side effect of screening for early forms of cancer" (5). This relates to the idea that within the unscreened population there exists a pool of subclinical cancers that would never become apparent during the lifetime of the screening participants, but because of the screening it is discovered leading to a cancer diagnosis, whereas if the person was never screened it would remain undiscovered throughout the lifetime of the participant. When conducting a typical stop-screen trial design where there are a limited number of screening rounds and long term follow, these overdiagnosed cancers can manifest as excess cancers in the screening arm with no mortality reduction. In the absence of overdiagnosis, it is expected that within a screening trial the screening arm will initially find more cancers, after screening has stopped, then there should be "catch up" of the number of cancers in the control arm. When this catch up does not occur, overdiagnosis is the presumed cause. However, this does not necessarily imply that the screening is only finding indolent cancers, the lack of catch up could also be due to comorbidities causing deaths (e.g., cardiac deaths) in the control arm before the cancer becomes manifest. It does not have to be interpreted as the screening is only finding indolent cases.

One of the first screening trial to be interpreted as demonstrating overdiagnosis was the Mayo Lung Project (MLP) whose results were published in 1986 (6). The MLP randomized participants to chest X-ray screening plus 3-day pooled sputum, every 4 months for 6 years versus annual

chest X-ray and sputum, with randomization occurring after all participants had an initial baseline screen (prevalence). No significant mortality reduction was found in that trial, and it was estimated that 51% of the screen-detected cancers were overdiagnosed (3). In other words, all of the excess cancers in the screening arm were "overdiagnosed." While this has generally been taken to mean that those excess cancers were indolent, there were actually six times as many deaths in the MLP due to other causes rather than lung cancer with heart disease the most common among them (6,7). These frequently occurring competing causes of death likely affected the rate of overdiagnosis. This trial had a major impact on lung cancer screening policy in the United States leading to a recommendation against chest radiographic screening and deepening the concerns regarding overdiagnosis. It is important to now consider what would happen in other screening trials with different designs. When performing a screening trial the initial round of screening has the greatest frequency of slow growing cancers, this is generally referred to as length-bias sampling. In the case of the MLP that baseline round was excluded and participants were only randomized after they had undergone an initial screen. A trial that would perform the random assignment prior to that baseline round would be expected to have an even greater percentage of overdiagnosed cases due to the large number of slower growing cancers found in the baseline round. In addition, a more sensitive test, such as CT scanning would also be expected to find more cancers when they are early and more likely to be indolent and thus, a screening trial that uses CT instead of chest radiographs would also be expected to have a greater number of overdiagnosed cases. We examine the actual findings in several studies.

With the publication of the Early Lung Cancer Action Project (ELCAP) results showing that low-dose CT was superior to chest radiography in 1999 (8), interest in CT screening for lung cancer greatly increased. However, along with great enthusiasm for early diagnosis, concern for overdiagnosis and other harms intensified. So much so, that it was even postulated that CT was incapable of finding genuine cancers and only could find indolent cancers, the type that would not progress or lead to death. In his article, "Is our natural-history model of lung cancer wrong?" Bach postulated the following model based on his interpretation of the literature (8):

- Under this model, patients with early lung cancer would no longer be assumed to be patients with pre-advanced lung cancer, and patients with advanced lung cancer would not be assumed to have had detectable and curable disease in the past.
- Like the other studies of screening by chest radiography, this analysis suggests that early-stage lung cancers detected by [LDCT] screening are not precursors of advanced lung cancers.

Thus, according to the above theory, 100% of screen detected cancers using CT would be overdiagnosed and low-dose CT screening would yield an even greater extent of overdiagnosis than chest radiographic screening.

Given the above considerations regarding factors that would influence the extent of overdiagnosis it is now useful to see what actually happened when reviewing results from various trials.

In 1993, the year ELCAP began their study on CT screening, the Prostate Lung Colorectal Ovarian Trial (PLCO) began. The PLCO compared chest x-ray to usual care (no routine screening) and published its results in 2011 (9). Here the randomization occurred prior to the baseline round and therefore it would be expected that the extent of overdiagnosis would be more than the MLP, however, the rate of overdiagnosis was found to be an insignificant 6% (10). The National Lung Screening Trial (NLST) which compared annual CT screening with annual CXR screening began enrolling participants in 2002 and published its core results in 2011 (11). A follow-up report was published in 2019 to specifically evaluate for extent of overdiagnosis (12). We would have expected the extent of overdiagnosis to be the highest in the NLST, yet its results demonstrated that when cases of BAC were excluded, there was essentially 0% overdiagnosis, and even when including those cases of BAC, which are easily managed through serial observation to assess for progression, it was estimated at 3%.

Thus, we see that results from these trials are completely at odds with what would have been expected to occur based on an understanding of where indolent cancers are most likely to manifest. So spectacular is this misunderstanding of the concept of overdiagnosis that it led voices for major societies producing guidelines in the field of screening to predict that there would be virtually 100% overdiagnosis when in actuality the NLST trial showed it was virtually 0%. To further demonstrate the confusion around this topic, the Danish Lung Cancer Screening Trial (DLCT) which compared CT screening to usual care found the extent of overdiagnosis was 67%, while the ITALUNG screening trial, which had a similar design, found no evidence of overdiagnosis (13,14).

Along with the difficulty in defining overdiagnosis and estimating its extent in the screening population, there has been the idea that screening leads to harm, in the form of unnecessary treatment because the person would die of a competing cause of death. This is generally referred to as overtreatment and has been described as the main cause of harm associated with overdiagnosis. If an indolent cancer, one that would not progress can be identified then surely there would be no benefit in performing treatment, especially if the treatment had associated risks. This idea has led to the concept of watchful waiting. This is now commonplace for certain subsets of prostate cancers and there is now a class of lung cancer, those that manifest as nonsolid nodules whereby they can be safely followed on an annual basis without any harm (15,16). It is also noteworthy that the necessary evidence to make decisions about which subtypes of lung cancers could be managed in this manner did not come from randomized trials, but instead came from studying those cases in the context of clinical care.

As we have tried to describe above, the current concepts of overdiagnosis and also overtreatment are confusing and results from screening trials have rendered them essentially meaningless, at least as they are currently defined. The root cause of this confusion relates to the way it has been defined. Its current definition is an epidemiologic one, and is being used to define what should be thought of as a clinical entity. If one were to rationally define overdiagnosis, it would be based on understanding features that describe the aggressiveness of the cancer that is being diagnosed. So, for example, a nonsolid lung cancer, which behaves in an indolent manner, would represent a case of overdiagnosis if instead of being recognized as being indolent it was diagnosed as an aggressive one. Here, the concept of overtreatment then naturally flows from this because an aggressive treatment here, perhaps with surgery and chemotherapy, would not be necessary when the cancer could safely have been monitored. In this example overtreatment would result from the overdiagnosis, but it could also occur independently if the cancer was correctly diagnosed as being indolent yet still had an aggressive treatment. Along with this formulation of overdiagnosis, the concept of underdiagnosis can be adduced and with it the attendant undertreatment. Note that these proposed definitions are based on clinical features of the cancer and these concepts remain in the clinical domain not in epidemiology and therefore allow for an actual diagnosis to be made and do not rely on it being made based on the cause of death.

In conclusion, the current formulation of the concept of overdiagnosis has led to confusion and results that are in all essence uninterpretable and meaningless. This is not to say that indolent cancer is not a real phenomenon, because the evidence for this is quite clear and management plans for how to deal with them, especially in lung cancer where they can be directly identified on CT scans are now well established. The problem has arisen due to the attempt to define a clinical entity based on an epidemiologic concept. When focusing on the clinical aspects of the aggressiveness of the cancer, this provides a rational path for thinking about overdiagnosis as a clinical entity and also provide a path toward thinking about the associated research. It implies that there needs to be better understanding of methods to determine the aggressiveness of the tumors, including through the use of novel biomarkers, and issues of competing causes of death while obviously important, need to be studied independently.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editor (Witold Rzyman) for the series "Implementation of CT-based screening of lung cancer" published in *Translational Lung Cancer Research*. The article has undergone external peer review.

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tlcr-20-736). The series "Implementation of CT-based screening of lung cancer" was commissioned by the editorial office without any funding or sponsorship. Dr. Yankelevitz reports other from Accumetra, other from GRAIL, outside the submitted work. In addition, Dr. Yankelevitz is a named inventor on a number of patents and patent applications relating to the evaluation of diseases of the chest including measurement of nodules. Some of these, which are owned by Cornell Research Foundation (CRF), are non-exclusively licensed to General Electric. As an inventor of these patents, I am entitled to a share of any compensation which CRF may receive from its commercialization of these patents. He is also an equity owner in Accumetra, a privately held technology company committed to improving the science and practice of image-

based decision making. Dr Yankelevitz is also on the advisory board of GRAIL, Pfizer and AstraZeneca. Dr. Henschke is a named inventor on a number of patents and patent applications relating to the evaluation of pulmonary nodules on CT scans of the chest which are owned by Cornell Research Foundation (CRF). Since 2009, Dr. Henschke does not accept any financial benefit from these patents including royalties and any other proceeds related to the patents or patent applications owned by CRF. Dr. Henschke is the President and serve on the board of the Early Diagnosis and Treatment Research Foundation. I receive no compensation from the Foundation. The Foundation is established to provide grants for projects, conferences, and public databases for research on early diagnosis and treatment of diseases. Recipients include, I-ELCAP, among others. The funding comes from a variety of sources including philanthropic donations, grants and contracts with agencies (federal and non-federal), imaging and pharmaceutical companies relating to image processing assessments. The various sources of funding exclude any funding from tobacco companies or tobacco-related sources. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Decision Memo for Screening for Lung Cancer with Low Dose Computed Tomography (LDCT) (CAG-00439N). Available online: https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=274. Accessed June 01/2020
2. NCI Dictionary of Cancer Terms. Available online: www.cancer.gov/publications/dictionaries/cancer-terms/def/overdiagnosis. Accessed June 01/2020
3. Welch GH, Black WC. Overdiagnosis in cancer. J Natl Cancer Inst. 2010;102:605-13.
4. Miettinen OS. Epidemiological Research: Terms and Concepts. Dordrecht: Springer; 2011:20.
5. Overdiagnosis. Available online: https://en.wikipedia.org/wiki/Overdiagnosis. Accessed June 01/2020.
6. Fontana RS, Sanderson DR, Woolner LB, et al. Lung cancer screening: the Mayo program. J Occup Med 1986;28:746-50
7. Marcus PM, Bergstralh, Fagerstrom RM, et al. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. J Natl Cancer Inst 2000;92:1308-16.
8. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999;354:99-105.
9. Oken MM, Hocking WG, Kvale PA, et al. PLCO Project Team. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. JAMA 2011;306:1865-73.
10. Wender R, Fontham ET, Barrera E Jr, et al. American Cancer Society lung cancer screening guidelines. CA Cancer J Clin 2013;63:107-17.
11. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365:395-409.
12. National Lung Screening Trial Research Team. Lung Cancer Incidence and Mortality with Extended Follow-up in the National Lung Screening Trial. J Thorac Oncol 2019;14:1732-42.
13. Heleno B, Siersma V, Brodersen J. Estimation of overdiagnosis of lung cancer in low-dose computed tomography screening. JAMA Intern Med 2018;178:1420-2.
14. Paci E, Puliti D, Lopes Pegna A, et al. Mortality, survival, and incidence rates in the ITALUNG randomised lung cancer screening trial. Thorax 2017;72:825-31.
15. Yankelevitz DF, Yip R, Smith JP, et al. CT screening for lung cancer: nonsolid nodules in baseline and annual repeat rounds. Radiology 2015;277:555-64.
16. Lung CT screening reporting & data system. 2014. Available online: www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads. Accessed June 01/2020.

Cite this article as: Yankelevitz DF, Henschke CI. Overdiagnosis in lung cancer screening. Transl Lung Cancer Res 2021;10(2):1136-1140. doi: 10.21037/tlcr-20-736

