How to prevent overdiagnosis

Arnaud Chiolero\textsuperscript{a}, Fred Paccaud\textsuperscript{b}, Drahomir Aujesky\textsuperscript{b}, Valérie Santschi\textsuperscript{c}, Nicolas Rodondi\textsuperscript{b}

\textsuperscript{a} Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital, Switzerland
\textsuperscript{b} Department of General Internal Medicine, Bern University Hospital, Switzerland
\textsuperscript{c} La Source School of Nursing Sciences, University of Applied Sciences Western Switzerland, Lausanne, Switzerland

Summary

Overdiagnosis is the diagnosis of an abnormality that is not associated with a substantial health hazard and that patients have no benefit to be aware of. It is neither a misdiagnosis (diagnostic error), nor a false positive result (positive test in the absence of a real abnormality). It mainly results from screening, use of increasingly sensitive diagnostic tests, incidental findings on routine examinations, and widening diagnostic criteria to define a condition requiring an intervention. The blurring boundaries between risk and disease, physicians’ fear of missing a diagnosis and patients’ need for reassurance are further causes of overdiagnosis. Overdiagnosis often implies procedures to confirm or exclude the presence of the condition and is by definition associated with useless treatments and interventions, generating harm and costs without any benefit. Overdiagnosis also diverts healthcare professionals from caring about other health issues. Preventing overdiagnosis requires increasing awareness of healthcare professionals and patients about its occurrence, the avoidance of unnecessary and untargeted diagnostic tests, and the avoidance of screening without demonstrated benefits. Furthermore, accounting systematically for the harms and benefits of screening and diagnostic tests and determining risk factor thresholds based on the expected absolute risk reduction would also help prevent overdiagnosis.

Key words: diagnosis; overdiagnosis; overtreatment; cancer screening; prevention

Introduction

Traditionally, a diagnosis is based on clinical symptoms and signs, and a patient’s past medical history. The development of preventive medicine and diagnostic technologies, in a context of a predominance of chronic conditions with a long pre-clinical phase, have changed the diagnostic process, expanding the possibilities of interventions across asymptomatic individuals and blurring the boundaries between health, risk and disease [1].

An unwanted effect of these developments is overdiagnosis, defined as the diagnosis of an abnormality that is not associated with a substantial health hazard [2–5] and that patients have no benefit to be aware of [6]. Over-diagnosis is neither a misdiagnosis (diagnostic error), nor a false positive result (positive test in the absence of a real abnormality). Nevertheless, it leads to overtreatment that provides, by definition, no benefit but that can cause harm and induce costs.

Our goal is to comprehensively review the causes of overdiagnosis (table 1) and to propose interventions to prevent its occurrence (table 2).

Literature search

We sought for publications related to overdiagnosis in the Medline database via PubMed (1950 to March 2014). Search terms were “overdiagnosis” OR “overdetection”. Some 1603 potential citations were identified and screened. We selected high quality general reviews about overdiagnosis and original research studies assessing the existence and the frequency of overdiagnosis, in all domains of medicine and public health. A similar search was conducted in Google scholar to find additional publications. We also conducted a hand search of bibliographies of all relevant articles. Additionally, websites dedicated to the issue of overdiagnosis were searched for.

Causes of overdiagnosis

Overdiagnosis can directly result from screening tests, use of increasingly sensitive diagnostic tests, incidental findings on routine examinations, and widening diagnostic cri-

<table>
<thead>
<tr>
<th>Table 1: Causes of overdiagnosis [2–6].</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximal causes:</td>
<td></td>
</tr>
<tr>
<td>Screening</td>
<td></td>
</tr>
<tr>
<td>Increasing sensitivity of diagnostic tests</td>
<td></td>
</tr>
<tr>
<td>Incidental findings (following laboratory or radiological examinations, genetic tests, etc.)</td>
<td></td>
</tr>
<tr>
<td>Widening diagnosis criteria to diagnose a condition requiring an intervention</td>
<td></td>
</tr>
<tr>
<td>Distal causes:</td>
<td></td>
</tr>
<tr>
<td>Blurring boundaries between risk and disease</td>
<td></td>
</tr>
<tr>
<td>Physician's fear of missing a diagnosis and fear of litigation</td>
<td></td>
</tr>
<tr>
<td>Patient's need of reassurance</td>
<td></td>
</tr>
<tr>
<td>Financial incentives</td>
<td></td>
</tr>
</tbody>
</table>
teria to define a condition requiring an intervention (table 1) [2–5]. Indirectly, the blurring boundaries between risk and diseases, the fear of missing a diagnosis and the patients’ need for reassurance are further causes of overdiagnosis.

Screening and overdiagnosis of cancers

Screening is a major source of cancer overdiagnosis. An overdiagnosed cancer is defined as a tumour whose histology indicates a cancerous lesion but which will not evolve to become symptomatic or lethal. These tumours constitute a reservoir of potentially detectable cancers. The size of this reservoir is estimated by post mortem histological examination of tissue in people who have not suffered from or have not died of such a cancer. For example, 30% to 70% of men aged over 60 have asymptomatic prostatic cancerous lesions [2]. Screening causes a shift in tumour grade or histology toward more indolent cancers [7, 8]. Therefore, cancer screenings tend to detect malignancies that are growing more slowly on average than cancers clinically diagnosed. This is the source of the “length-time bias” (or “prognostic bias”) that must be taken into account in the evaluation of screening. In extreme cases, screening can detect cancer that would have never become symptomatic before the death of the screened individual: these cases are overdiagnosed.

Prostate cancer

Since 2012, the U.S. Preventive Services Task Force (USPSTF) recommends not to perform routine screening of prostate cancer by the determination of prostate-specific antigen (PSA) [9]. In view of the frequent side effects (urinary incontinence, erectile dysfunction) due to treatment and of unclear benefits, the balance between risks and harms has been considered unfavourable by the USPSTF [9]. A recent Cochrane review concluded that screening had no effect on mortality from prostate cancer [10]. There is strong evidence that screening for prostate cancer by PSA is an important source of overdiagnosis. Indeed, prostate cancer is heterogeneous and many tumours found by screening grow slowly and are not life threatening [9, 11] and PSA is unable to distinguish aggressive cancers requiring intervention from indolent cancers [12]. Furthermore, the large increases in the incidence of prostate cancer in US and several European countries are compatible with overdiagnosis. The estimates are uncertain but between 17% and 66% of prostate cancers detected by PSA screening would be overdiagnosed [9, 11]. The probability of overdiagnosis is higher in older men, having a relatively low Gleason score and a relatively low level of PSA at diagnosis [13]. Overdiagnosis and overtreatment are among the reasons cited by the USPSTF to recommend against PSA screening for prostate cancer [9].

Breast cancer

The USPSTF recommends routine screening for breast cancer with mammography every two years for women 50 to 74 years [14]. Randomised studies have shown that routine screening reduces mortality from breast cancer by about 20% [15]. The net benefit of screening is directly related to absolute baseline risk of breast cancer and increases with age [16]. Although to a lesser extent compared with prostate cancer, some cases of breast cancer discovered following mammography are overdiagnosed [15–18]. In a recent review of randomised trials, it was estimated that 19% of cases detected by screening every 3 years beginning at the age of 50 years were overdiagnosed [15]. However, there were large uncertainties on the frequency of overdiagnosis estimates, ranging from 0% up

<table>
<thead>
<tr>
<th>Method</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoid certain diagnostic tests or screening</td>
<td>Systematic screening for prostate cancer by PSA measurement [9]</td>
</tr>
<tr>
<td>Reduce the frequency of screening test</td>
<td>Breast cancer screening every 2 years instead of every year [18]</td>
</tr>
<tr>
<td>Targeted screening</td>
<td>Screening for breast cancer with mammography: mortality benefit vs harms of overdiagnosis [14–16]</td>
</tr>
<tr>
<td>Screen a diagnostic test or screening</td>
<td>Cardiovascular risk estimation based on age or on the basis of a clinical score (e.g., Framingham)</td>
</tr>
<tr>
<td>Anticipate the consequences of abnormalities discovered following a diagnostic test or screening</td>
<td>The relative increase in the cardiovascular risk associated with high cholesterol is similar in old people and in young people compared to same age people with low cholesterol. However, the absolute increase in the risk and the benefits of treatment are low in young people with high cholesterol</td>
</tr>
<tr>
<td>Screening with combined diagnostic and prognostic tools (biomarker, personalised medicine) [43, 45, 53]</td>
<td>Cancer vs indolent lesions of epithelial origin (IDLE) [53]</td>
</tr>
<tr>
<td>Prognosis estimation to decide whether or not to screen and treat</td>
<td></td>
</tr>
<tr>
<td>When assessing a risk factor, consider the absolute risk of disease associated with this factor and expected absolute risk reduction through intervention or treatment</td>
<td></td>
</tr>
<tr>
<td>Avoid conflicts of interests in guideline panel committees</td>
<td></td>
</tr>
<tr>
<td>Change terminology for conditions with a high probability of indolence to prevent overtreatment [53]</td>
<td></td>
</tr>
</tbody>
</table>

Swiss Medical Weekly · PDF of the online version · www.smw.ch
to 36% of invasive breast cancers diagnosed during the screening period [14].

**Overdiagnosis of other cancers**
Overdiagnosis could explain in part the increase in the incidence of melanoma [19], as well as lung [20, 21], kidney, and thyroid cancers [22, 23], which are often discovered incidentally following radiological examinations (CT scan, MRI, ultrasound) [3].

**Increasing sensitivity of diagnostic tests and incidental findings**
The use of increasingly sensitive radiological diagnostic tests is a source of overdiagnosis for several conditions, including pulmonary embolism [24, 25]. Due to the high mortality risk, pulmonary embolism is a diagnosis that physicians do not want to miss. Therefore, the increasing sensitivity of diagnostic tools, such as multi-detector CT-scans, is clinically appealing. However, this leads to an increased detection of less severe cases of pulmonary embolism (e.g., sub-segmental) for which the benefit of anticoagulation is doubtful [24, 25].

In the US, the incidence of pulmonary embolism was stable between 1993 and 1998 but increased from 62 to 112 cases per 100,000 persons between 1998 and 2006, a period characterised by the growing use of more sensitive multi-detector CT-scans. Despite the increased number of detected pulmonary embolism, mortality from this condition has decreased only slightly since 1993 [24]. A pattern of massive increase in the incidence without significant change in mortality is compatible with overdiagnosis. Indeed, there is a true reservoir of potentially detectable pulmonary embolisms [26, 27].

The increasing use of diagnostic imaging, in conjunction with an increasing sensitivity, leads to frequent incidental findings of abnormalities whose clinical hazard is uncertain but for which follow-up tests are required and treatment are, eventually, prescribed [28]. A classical example is the adrenal incidentaloma found on abdominal CT-scans. A report of 25 autopsy studies estimates the prevalence of an incidental adrenal mass to be 6% and up to 4% of abdominal CTs reveal an unexpected adrenal mass [29]. The great majority of adrenal incidentalomas are benign adenomas. However, because these incidentalomas can also be a secreting adenoma, pheochromocytoma, adrenocortical carcinoma, or metastatic carcinoma, they require additional diagnostic procedures and a potential increase of overdiagnosed and overtreated cases.

Whole-body radiological imaging is a major potential source of overdiagnosis. For example, among 1,192 persons with no particular health problem and having undergone a whole-body CT-scan, abnormalities were found in 86% of patients, with an average of 2.8 abnormalities per patient [30]. Genetic tests are also a source for incidental findings and overdiagnosis [31].

**Widening diagnosis criteria**
A more subtle form of overdiagnosis is related to the lowering of thresholds, for example of blood pressure, blood cholesterol or blood glucose, to define conditions requiring treatment [32, 33]. Lowering thresholds are justified by the continuous association between the variables of interest and the risk of diseases, such as between the level of systolic blood pressure and the risk of cardiovascular diseases. Due to this continuity in the risk, it is theoretically useful to decrease blood pressure at any level in order to decrease the risk of blood pressure related diseases.

However, the relationship between a given risk factor and the risk of disease is generally log-linear [34, 35]. Thus, a pressure reduction of 10 mm Hg is associated with a greater absolute risk reduction if the blood pressure is initially at 160 mm Hg than if it is at 140 mm Hg: the higher the blood pressure, the greater the absolute risk reduction. Below a given threshold, the absolute benefit of treatment becomes minimal and insufficient in regards to the potential harm [35]. This is a situation of overdiagnosis and overtreatment. Lowering thresholds also leads to a substantial increase in the number of cases treated for high blood pressure, high blood cholesterol, diabetes or their precursors, such as pre-hypertension or pre-diabetes [35]. Widening of diagnostic thresholds, and the resulting overdiagnosis, leads to an epidemic of diagnoses [3, 36]. To limit the risk of overdiagnosis, it is key to consider the absolute risk of disease because the higher the absolute risk, the lower the probability of overdiagnosis, and this is true at any level of risk factors. More generally, the blurring boundaries between risk and disease contribute to causing overdiagnosis [37]. Indeed, being at risk of diseases is increasingly confounded with the disease itself, and it is true in particular for chronic diseases such as diabetes or various cancers.

**Physician’s fear of missing a diagnosis and patients’ need for reassurance**
Another cause of overdiagnosis is the physician’s fear of missing a diagnosis and the patient’s need for reassurance. The fear of missing a diagnosis is at the heart of the physician ethos, and as such, is difficult to fight. The fear of litigation also exists because in some jurisdictions there is a risk of litigation when a test is not done, but not when a test is done even if it results in the identification of indolent abnormalities [38]. Financial incentives for physicians and from pharmaceutical and diagnostic industries are also causing overdiagnosis [39]. More broadly, societal “more is better” attitudes and blind beliefs in new technologies are major drivers for more testing, and hence more overdiagnosis.

**Methods to estimate the frequency of overdiagnosis**
There is no consensus on how to estimate the frequency of overdiagnosis. The excess incidence method is based on the comparisons of cancer incidence in screened and unscreened populations. Thus, if there is no overdiagnosis, the cumulative-incidence should be equal in both populations with a follow-up extended over a few years after the end of screening period; in case of overdiagnosis, the cumulative-incidence will be larger in the screened population, due to the excess cases detected (fig. 1) [40]. A stand-
ard calculation to estimate the percentage of overdiagnosed cases is:

\[
\text{Percentage of overdiagnosis} = \frac{\text{incidence in screened population} - \text{incidence in unscreened population}}{\text{incidence in unscreened population}}
\]

One difficulty is to compare screened and unscreened populations having the same underlying risk of cancer, ideally using randomised trials data. If observational data are used, it is necessary to adjust for cancer risk factors and to account for the actual degree of participation to screening, the time interval between screening tests (overdiagnosis tends to be more frequent with shorter intervals [18]), the sensitivity of the screening test, and the lead-time [41]. Another method is based on the modelling of the lead-times bias [42]. This method requires information on the incidence of cancer without screening, and on the pattern of screening dissemination [39].

Over-diagnosis due to cancer screening can also be assessed by the analyses of secular trends in cancer stages. In theory, following the introduction of screening, the incidence of cancers initially increases and, after a few years, return to the pre-screening level (fig. 1) [42, 43]. More specifically, a few years after the introduction of screening, the number of cases diagnosed at an early stage should have increased and the number of cases at an advanced stage should have, ideally, decreased, with the total number of cases unchanged [42, 43]. In case of overdiagnosis, however, once screening is implemented, the incidence remains above the pre-screening level [42, 43].

Prevention of overdiagnosis

A case of overdiagnosis would be identifiable if the person was not treated for the detected abnormality, never suffered from this abnormality, and eventually died of an unrelated cause [2]. However, once an abnormality is identified, it is almost impossible - and often unethical - not to investigate and not to treat the patient even if there is a high probability of overdiagnosis. Nevertheless, overdiagnosis can be prevented (table 2). Increasing awareness of healthcare professionals and patients about overdiagnosis and avoiding unnecessary diagnostic tests, useless routine examinations, and inefficient screening should be advocated. One major difficulty is that many screening tests are very popular [44] and even when a given screening test has been shown to be inefficient, it will often continue to be proposed and performed [16]. Recently, the American Board of Internal Medicine has launched the “Choosing Wisely” initiative (www.choosingwisely.org) with the aim of identifying useless tests with a potential of overdiagnosis [28]. The Swiss Society of Internal Medicine has launched a similar campaign, Smarter Medicine (www.smartermedicine.ch).

Targeted screening also reduces the risk of overdiagnosis. Screening with combined diagnostic and prognostic tools (e.g., using biomarkers before screening to assess the susceptibility of the persons to develop the disease and for the aggressiveness of the disease detected) could help target the persons with the greatest probability of benefits and the lowest risk of overdiagnosis [43, 45]. Unfortunately, there is no convincing pre-clinical-prognostic tool to determine the evolution of most at-risk or chronic conditions. Research devoted to personalised medicine should try to help design targeted screening strategy.

An intervention on risk factors such as elevated blood pressure is justified only by its ability to substantially reduce the associated absolute risk of disease. It is therefore essential to educate physicians and patients to frame risk in absolute terms rather than in relative terms. More generally, before doing a test, it is necessary to estimate the absolute risk associated with the condition and to do the test only if it is possible to reduce this risk substantially. Furthermore, overdiagnosis can occur with underdiagnosis in the same population. For instance, while some patients at low cardiovascular risk are treated, some high-risk patients are not treated. This may be due to the fact that the decision to treat is based on the level of the risk factor itself (e.g., cholesterol level), and not on the associated absolute cardiovascular risk [46]. Prevention of overdiagnosis should not be done at the cost of increasing under-diagnosis and under-treatment. The financial interests of diagnostic and pharmaceutical industries are also a source of overdiagnosis [39]. Recommendations for diagnostic procedures or for the definitions of conditions requiring treatment should be elaborated by experts who have no ties with these industries. A recent study showed that many panel members of guidelines proposing a widening of the definition of common conditions had financial ties with the healthcare industry [47]. Furthermore, none of these guidelines reported potentials harm of such a widening.

Once an abnormality is detected, the goal must be to prevent overtreatment. Tools like prognostic scores can help distinguishing abnormalities associated with a substantial risk and requiring treatment from indolent ones. For conditions with a high risk of overdiagnosis, a watchful waiting strategy should be considered. For example, under certain conditions, patients with a prostate cancer of low grade and low volume discovered by screening may be initially fol-
lowed without treatment [48]. However, these patients may choose to be treated quickly because of an underlying fear that the risk of cancer-related death may have been underestimated [48]. Further, physicians may fear an underestimation of risk and anticipate that they might regret not having recommended intervention [49].

Physicians should also discuss with their patients the benefits and harm associated with a diagnostic or screening test, including the risk of overdiagnosis [50]. Significance of harm (including overdiagnosis) and benefits associated with a given test depends on individuals’ value and preferences. Ideally, the goal is to reach a shared decision [51] and balanced information on the probabilities of harm and benefits, accounting for uncertainties, has to be provided to the patients [16, 18, 52]. Some patients will accept the risk of overdiagnosis, and other will not.

Finally, changing terminology for some conditions based on their degree of health risk could help prevent overtreatment. Minimal-risk, low-grade, and non-invasive breast and prostate lesions could be better labelled “indolent lesion of epithelial origin” (IDLE) rather than “cancer”, this term implying a high probability of death without treatment [45, 53]. Ductal intraepithelial neoplasia (DIN) and prostate intraepithelial neoplasia (PIN) have been proposed [45, 53]. Some indolent thyroid cancers have been named micropapillary lesions of indolent course (microPLICs) [54]. Changes in nomenclature for low risk cancers would also help patients and physicians choose active surveillance.

Conclusion

Overdiagnosis is mainly the consequence of a technology-driven medicine that aims to improve patient outcomes by detecting disease in its earliest form. Overdiagnosis is of growing relevance in population dominated by chronic conditions having long pre-clinical stage. Although early detection has been shown to be beneficial for several conditions, it also increases the probability of finding insignificant abnormalities, whose treatment is not associated with any benefit but can harm the patient. Moreover, overdiagnosis diverts healthcare professionals from caring about other health issues and generates costs. The increase in healthcare costs, the over- and undertutilisation of some care, the debates about the effectiveness of various screening, and the growing role of patients in medical decisions require concern about overdiagnosis.

References


54 Brito JP, Morris JC, Montori VM. Thyroid cancer: zealous imaging has increased detection and treatment of low risk tumours. BMJ. 2013;347:f4706.
Figure 1

Panel A) Following the implementation of cancer screening, assuming no true change in the frequency of cancer, the observed incidence will initially increase before reaching again, in the long term, the incidence observed in the pre-screening era. In case of over-diagnosis, in the long term, the incidence does not return to pre-screening level due to the excess cases detected [43]; Panel B) Comparing cumulative incidence of screened and unscreened population can demonstrate over-diagnosis due to a screening test done regularly during a given period of life, e.g., mammography for breast cancer. During the period of screening, the cumulative incidence is higher with screening than without screening. Few years after the end of screening period, the cumulative incidence in the screened population reach the cumulative incidence in the unscreened population if there is no over-diagnosis. In case of over-diagnosis, the cumulative incidence remains higher [40, 43, 45].