Back to basics: overdiagnosis is about unwarranted diagnosis

Bjørn Hofmann [1,2]

Correspondence address:

Bjørn Hofmann, Institute for the Health Sciences at the Norwegian University of Science and

Technology in Gjøvik, Norway, (e-mail: bjoern.hofmann@ntnu.no)

Affiliations:

1. Institute for the Health Sciences at the Norwegian University of Science and Technology in Gjøvik,

   Norway, bjoern.hofmann@ntnu.no

2. Centre of Medical Ethics, Institute for Health and Society, Faculty of Medicine, University of Oslo,

   PO Box 1130, Blindern, N-0318 Oslo, Norway (e-mail: b.m.hofmann@medisin.uio.no)

Back to basics: overdiagnosis is about unwarranted diagnosis

Abstract:

Epidemiological studies of overdiagnosis are challenged by unclear definitions and absence of unified measures. This spurs great controversies. Etymologically, overdiagnosis means too much diagnosis and stems from the inability to discern what is important from what is not. Accordingly, in order to grasp, measure, and handle overdiagnosis we should revive medicine's original goal and re-connect diagnosis to what matters to professionals and patients, i.e., knowledge and suffering respectively. This will make overdiagnosis easier to define and measure, and eventually less difficult to reduce.

Introduction

Overdiagnosis is frequently defined as the diagnosis of a condition "that would otherwise not go on to cause symptoms or death."(1, p.605) While there is reasonable agreement that overdiagnosis is negative and should be avoided and that it is distinct from false positive test results, there is no agreement on how to define (2-5) or to measure it (6, 7). Some of the issues of the debates are listed in Table 1.

[Table 1]

The objective of this study is to analyze overdiagnosis in terms of its etymological elements: *over* and *diagnosis*. The analysis will show that overdiagnosis is *unwarrantedly identifying something as a disease due to prognostic uncertainty*. In particular, two conceptual aspects will be identified: development uncertainty and progression uncertainty. Additionally, three alternative approaches to handling these challenges will be discussed. The goal is to make overdiagnosis easier to comprehend, measure, and eventually to prevent. The analysis will start with the etymology of overdiagnosis.

Overdiagnosis = too much diagnosis

Literally, overdiagnosis means too much diagnosis. However, most of the controversy has been about what we mean by "over" or "too much" and much less attention has been payed to the key term "diagnosis." The etymology of diagnosis in general, and the analysis of the concept and process of *diagnosis* in particular, may bring us closer to an answer to some of the crucial questions.

Diagnosis stems from Latin and ancient Greek *diagignőskein* meaning to discern (from *dia-* meaning apart and *gignoskein* meaning to learn). Hence, etymologically diagnosis means discerning by means of learning or gaining new knowledge. A diagnostic test may provide information that makes it possible to discern what disease the patient has. Accordingly, we tend to say that a person is "diagnosed with a disease." The diagnosis gives the disease a label.

Diagnosis may have other functions than discerning disease, for example it may be informative with respect to prognosis and be action guiding with respect to selecting a treatment. There is also a difference between the verb, "to diagnose," the activity, "diagnostics," and the result, "the diagnosis." As we will see, the process of finding indicators and of "diagnosing" is very similar. However, in order to address the challenges with overdiagnosis, a differentiation between them is suggested.

Etymologically, overdiagnosis means *unwarrantedly giving a person the label of a disease*. To clarify when this happens, we can investigate the diagnostic process and start with the case where the disease label is warranted, i.e., in the case of manifest disease and suffering.

Manifest disease and suffering

"Human suffering" involves unpleasant sensory or emotional experience (8) and refers to states of severe human distress that threaten human agency (9). Suffering is related to pain as pain can result in or be part of suffering, but pain is not a necessary condition for suffering (10, 11). Moreover, suffering can be a bodily, a mental, a social, and an existential or spiritual experience (12).

Manifest disease is tangible characteristics of such suffering. Typical cases are (visual) lesions, tumors, and malformations. For example, a person may be suffering great pain, social stigma, and existential angst due to a visible tumor in his jaw. Identifying suffering and diagnosing manifest disease is widely considered to be warranted, as it calls for measures to help people. Hence, correctly diagnosing a suffering person with manifest disease cannot result in overdiagnosis.

Overdiagnosis resulting from prognostic uncertainty

Overdiagnosis does not occur due to lack of test accuracy either, as reduced accuracy results in *false diagnosis*, but not *unwarranted diagnosis*. Unwarranted diagnosis, on the other hand, occurs when there exists information that an identified entity can result in manifest disease and suffering, but

where we do not know whether it will do so. The reason for this is that it is uncertain how the condition that we correctly identify will develop. Hence, overdiagnosis is a result of *prognostic uncertainty*, i.e., we do not know whether what we have (correctly) identified will develop into manifest disease and suffering.

This type of uncertainty occurs when the indicator, such as the risk factor, predictor, precursor, or a combination of these, are only loosely connected to manifest disease and suffering. You do not know whether what you have found by a positive test will develop into manifest disease and suffering even though you could verify that the test result is correct. Because we want to catch disease early, and because perfect predictions do not (yet) exist, we catch too much, i.e., *too much diagnosis*.

However, we face another problem as well. Even when we have used a so-called "gold standard test" to give a person a disease label, we still do not know whether the person will experience manifest disease and suffering. For example a person with localized prostate cancer (Gleason ≤6) only has a chance of about 18% over 20 or more years that the tumor will progress to manifest disease or suffering (13). The reason for this is that diseases may progress differently and in ways that our diagnostic ("gold standard") tests are not able to differentiate. This latter problem occurs because what we define as disease is not tightly connected to what people experience, i.e., manifest disease and suffering. There is not a 1-to-1 mapping of what professionals find and what matters to people.

Hence, overdiagnosis is *unwarrantedly identifying something as a disease due to prognostic uncertainty*. We do not know whether what we find will develop to manifest disease or suffering. Hence, the information provided by the test does not make it possible to discern what matters to people from what does not.

Accordingly, prognostic uncertainty comes in two kinds. One results from our lack of knowledge of whether indicators will develop into manifest disease and suffering. This kind of uncertainty may be called *development uncertainty*. The other, which may be called *progression uncertainty*, stems from

the fact what we do not know whether what we call "disease" according to a "gold standard test" will develop into manifest disease and suffering.

Hence, overdiagnosis occurs because you do not know whether what you have identified by a test will develop to manifest disease and suffering or not. This *prognostic uncertainty* has two sources:

1. **Development uncertainty**: You have identified indicators (by a test) that you do not know will develop into manifest disease and suffering.

2. **Progression uncertainty**: You have identified a condition that is labelled "disease" (according to a "gold standard test") but you do not know whether it will progress into manifest disease and symptoms.

Figure 1 indicates the relationship between overdiagnosis and prognostic uncertainty.

Figure 1   Overdiagnosis resulting from prognostic uncertainty with two sources
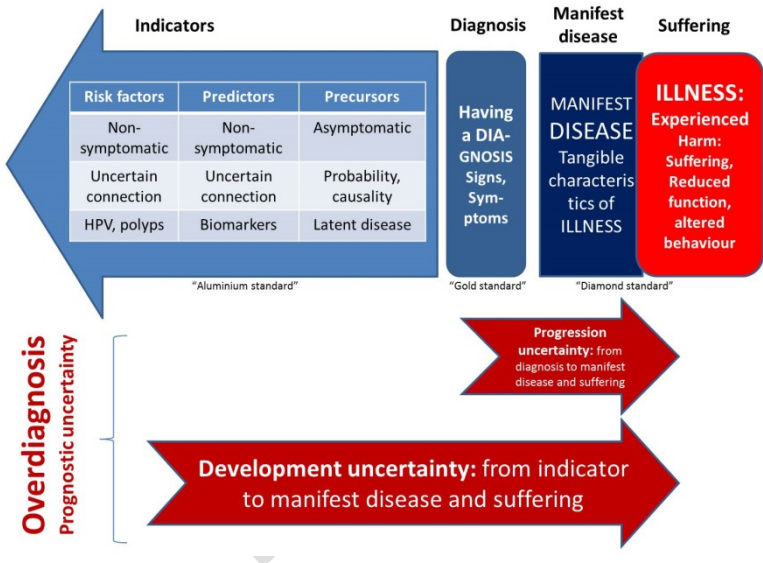
The impure gold standard of tests

New tests are normally tested against "gold standard tests," which previously were mostly pathology tests but ever more often are imaging or biomarker tests. However, a disease identified with a correct positive gold standard test does not necessarily progress to manifest disease and suffering,

because the tests are not perfect. Even less, do we know whether what is identified with an indicator will develop to manifest disease and suffering?

Accordingly, the ultimate criterion for a diagnostic test is whether it is able to identify manifest disease or suffering. We may call this "the diamond standard." If we need a name for the indicator tests, we may thus call it "the aluminium standard" of tests. The core problem with overdiagnosis, therefore, is the detachment of what we find from what matters to people. We lack "diamond standard tests."

Diagnosis and disease

As we have seen, traditionally and etymologically diagnosis means discerning by means of new knowledge. However, over time "diagnosis of a disease" has come to mean "identification of a specific disease." Hence, diagnosis has become synonymous with disease. When you have a diagnosis, you have a disease. However, due to *progression uncertainty* the connection between diagnosis and disease has become supple and given room for diagnostic errors, such as overdiagnosis. Overdiagnosis occurs because of two mechanisms:

1. We call something a disease even though we do not know whether it will progress to manifest disease and suffering.
2. We identify risk factors, predictors, or precursors of disease and connect these to specific diseases even though we do not know whether they will develop to manifest disease and suffering.

Hence, we do not have sufficient knowledge and cannot differentiate what we identify with our tests from disease. Thus, literally speaking, it is not a diagnosis. Therefore, "over-diagnosis" appears to be a misnomer. This leaves us with (at least) three alternatives:

1. Restore the original meaning of diagnosis, and use it only where we can differentiate cases of manifest disease and suffering from those who are not.

2. Decouple diagnosis from disease, where having a diagnosis does not mean that you have or will not get manifest disease and suffering.

3. Decouple disease from suffering. Stick to the connection between diagnosis and disease, but decouple them from harm and suffering.

1. Back to basics: reconnecting diagnosis to disease

If we require that diagnosis has to be based on knowledge that makes it possible to differentiate cases related to manifest disease and suffering from those who do not, we have solved the conundrum of overdiagnosis. Then every case of diagnosis will be correctly connected to manifest disease and suffering and there will be no over-diagnosis.

Unfortunately, this approach will make much of what we diagnose today falling outside the new and stricter definition of diagnosis. Actually, much of what we call diagnosis today will be overdiagnosis. This does not mean that the activity of looking for indicators (risk factors, predictors, or precursors) has to be abandoned. It only means that it does not qualify as finding or identifying (manifest) disease. Differentiating and identifying indicators does then not fall under the concept of diagnosis. Even more, identifying something that does not progress to manifest disease and suffering cannot rightly be called disease, even if we do so today (according to "gold standard tests"). This makes prognostic uncertainty dominated by development uncertainty.

This alternative may make it easier for people to understand that an identified indicator does not necessarily develop into disease. A person does not have to be anxious that they will experience manifest disease and suffering, and a physician does not have to be afraid s(he) will be litigated if (s)he does not treat the person. This approach also sticks to the intuitive understanding that a diagnosis refers to a disease. "Having an indicator" is very different from "having a diagnosis", and

we may need another word for identifying indicators. (An "indicator search" may give an "indicator finding" or an "indico" or "inveniendo" from latin).

Now the (scientific) task becomes to investigate the relationship between an identified indicator and manifest disease and suffering (comparing to "the diamond standard"). It does not solve the challenge with finding and informing about the chance of experiencing manifest disease and suffering (when you have found the indicator). However, it makes it clear that what you find is not disease. You may then "indicate disease" instead of "diagnose disease" and do "indicator search" instead of "diagnostics." See Figure 2.

Figure 2   Three alternative relationships between indicators, diagnosis, and disease
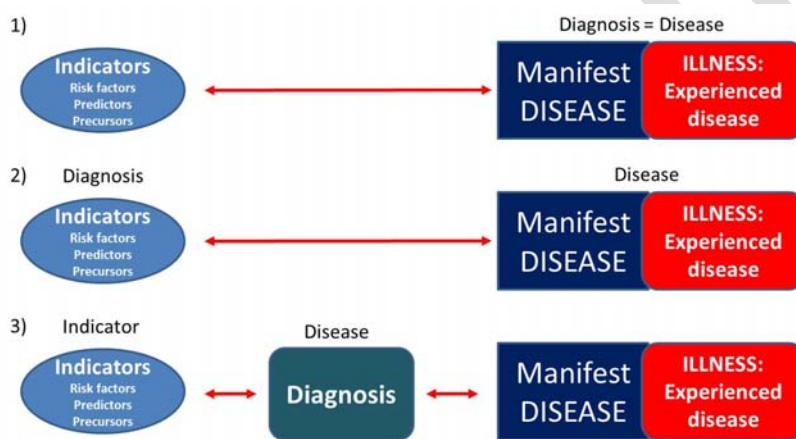
2. Detaching diagnosis from manifest disease

The other alternative is to decouple diagnosis from disease. Having a diagnosis does not mean that you will experience manifest disease and suffer. It only means that you have an increased chance of experiencing it. This is more or less where we are today, where it is uncertain whether you will experience manifest disease or suffering when you have a diagnosis.

The term overdiagnosis then tries to capture those cases of diagnosis that do not connect to disease. However, in this approach diagnosis has been decoupled from its original meaning – and disconnected from what people experience, i.e., manifest disease and suffering. Accordingly, informing people that they have a diagnosis, but not a disease, or even a high risk of getting manifest

9

disease and suffering, is challenging, as people's conception of diagnosis is so closely connected to their conception of disease. Moreover, this alternative conflates development uncertainty and progression uncertainty as both indicators and diagnosed conditions may develop into manifest disease.

3. Detaching disease and diagnosis from harm and suffering

A third alternative is to keep a close connection between diagnosis and disease but to decouple them from manifest disease and suffering. Accordingly, a person can have a diagnosis and a disease, but this condition does not need to be experienced as harmful. This alternative underscores the distinction between development uncertainty and progression uncertainty and is held (explicitly or implicitly) by scholars that define disease and overdiagnosis referring to "non-harmful disease" (14-16) or "inconsequential disease"(17).

While this approach can provide a precise definition of overdiagnosis, it faces the challenge of explaining that diagnosis and disease does not mean experienced harm or suffering. It goes beyond where having a diagnosis means that you have a disease (as in alternative 2). It also claims that having a diagnosis and a disease does not mean that you experience manifest disease and suffering. This may appear paradoxical as you can have disease without any dis-*ease*.

There are of course other alternatives. One can abandon the conception of overdiagnosis altogether and talk about erroneous diagnosis or diagnostic error. Moreover, one can desert the focus on diagnosis and look at the outcomes of various health care practices. However, as I have tried to illustrate, the concept of overdiagnosis appears to be relevant to highlight *unwarranted diagnosis due to prognostic uncertainty*.

What does it solve?

Which approach one chooses depends on what one wants to obtain. They address the various challenges with overdiagnosis, listed at the beginning. For example all three approaches dismiss

competing mortalities. If a person dies from a car crash instead of from manifest disease, this does not make the person overdiagnosed. In the first approach a person having a diagnosis has manifest disease and is suffering, so a car crash is irrelevant. In the second approach, the person is overdiagnosed because what is found (indicators or conditions) are not connected to the disease. This is independent of a subsequent car crash as well. According to the third approach, something is defined as overdiagnosis in terms of its (low) chance of developing to something harmful. What matters is the progression uncertainty. Therefore the car crash is not influencing this case (directly) either.

Moreover, overdiagnosis does not depend on treatment outcomes. While diagnostic efficacy can be assessed in terms of therapeutic efficacy and patient outcome efficacy (18) overdiagnosis is about diagnosis and not about treatment. This counters some conceptions of overdiagnosis (2).

The approaches differ, however, when it comes to whether you can have a correct diagnosis without having the disease and whether it is possible to give a general definition of disease. See Table 2.

[Table 2]

Approach 1 and 2 both differentiate medicalization from overdiagnosis. Medicalization is to make certain phenomena, such as sorrow, count as manifest disease and suffering. This is different from identifying indicators and making diagnosis of conditions that may or may not develop into disease (due to prognostic uncertainty). Approach 3, on the other hand makes medicalization part of overdiagnosis. Hence, it is much more inclusive.

Where should we go?

With reference to its etymology overdiagnosis is giving too many persons a disease label because the information we gain by testing is insufficient to discern who will have manifest disease and suffering

in the future and who will not. Hence, *overdiagnosis is prospectively unwarranted diagnosis stemming from prognostic uncertainty*.

If we want to bring medicine and its activity closer to what matters to people we should restrict the use of diagnosis to its original meaning. This *purist approach* could significantly reduce overdiagnosis in the future, but make much of what we call diagnosis today overdiagnosis. We should still identify indicators, but acknowledge that they are by no means diagnosis of manifest disease and suffering. Indicators do not discern disease, but only indicate a potential development of disease. Finding an indicator (or having an "indico") does not give you qualified knowledge about your present or future health. Conveying and accepting uncertainty is crucial for good and trusted care (19, 20).

Alternatively, we have to accept that disease and diagnosis are two different things. Disease would then not be the reference for diagnosis. This would reduce the problem of overdiagnosis, as all true positive diagnosis would be correct by definition. However, detaching diagnosis from disease would make diagnoses of less value to people for understanding whether they will develop and experience manifest disease and suffering. Diagnoses may mainly be useful in professional heuristics for identifying risk factors, predictors, and precursors. Informing about diagnosis then has to be detached from information of disease, due to progression uncertainty. Having a diagnosis does not need to be(come) harmful for you.

The most liberal approach is to detach indicators from diagnosis and diagnosis from disease. This gives overdiagnosis two sources: development uncertainty and progression uncertainty. It also allows for what has been called "maldetection overdiagnosis"(3) or "overdetection"(4). However, it allows for disease without *dis-ease* and detaches medicine from its original objective and subject matter: the suffering human being.

No doubt, it is difficult to change professional and public conceptions of disease and diagnosis. However, in order for these concepts not to be subject to heterogeneous drivers, such as technology (21), we need to take control. The meaning of these concepts will certainly change in the future, as

they have done previously. However, it is crucial that we govern this change so it helps instead of harming people. In general it is essential that we avoid detection without prospective benefit.

Conclusion

Overdiagnosis means too much diagnosis, and etymologically diagnosis means discerning based on knowledge. More specifically, overdiagnosis is a result of prognostic uncertainty: we do not know whether what we have identified will become manifest disease and suffering. Labelling indicators, such as risk factors, predictors, and potential precursors as "diagnosis" unwarrantedly connects such indicators to disease. Overdiagnosis underscores the difficulty of differentiating between what is important and not. One practical reason for this is that new indicators and increased precision detaches diagnosis from manifest disease resulting in inflation in diagnosis. Thus, overdiagnosis is a result of (conceptually) removing diagnosis from disease. As such, diagnosis deflects medicine from its original goal: to help people who are suffering by identifying what is bothering them and by treating it to alleviate their ailment. Accordingly, we must revive medicine's original goal and re-connect diagnosis to what matters to professionals (knowledge) and to patients (avoid suffering). Only then can we measure and handle overdiagnosis.

References

1.   Welch HG, Black WC. Overdiagnosis in cancer. Journal of the National Cancer Institute. 2010;102(9):605-13.
2.   Carter SM, Degeling C, Doust J, Barratt A. A definition and ethical evaluation of overdiagnosis. Journal of medical ethics. 2016;42:722-4.
3.   Rogers WA, Mintzker Y. Getting clearer on overdiagnosis. Journal of evaluation in clinical practice. 2016;22(4):580-7.
4.   Brodersen J, Schwartz LM, Heneghan C, O'Sullivan JW, Aronson JK, Woloshin S. Overdiagnosis: what it is and what it isn't. Evidence-based medicine. 2018;23(1):1-3.
5.   Davies L, Petitti DB, Martin L, Woo M, Lin JS. Defining, Estimating, and Communicating Overdiagnosis in Cancer Screening. Annals of internal medicine. 2018;169(1):36-43.

6.   de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiologic reviews. 2011;33(1):111-21.

7.   Etzioni R, Gulati R, Mallinger L, Mandelblatt J. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. Annals of internal medicine. 2013;158(11):831-8.

8.   DeGrazia D. What is suffering and what sorts of beings can suffer. In: Green R, Palpant N, editors. Suffering and Bioethics. New York: Oxford University Press; 2014. p. 134-54.

9.   Cassel EJ. The nature of suffering and the goals of medicine. The New England journal of medicine. 1982;306(11):639-45.

10.   Fordyce WE. Pain and suffering: a reappraisal. American Psychologist. 1988;43(4):276.

11.   Clarke BM. On suffering: pathways to healing & health. Hanover, NH: Dartmouth College Press; 2011.

12.   Hofmann B. Suffering: Harm to Bodies, Minds, and Persons. In: Schramme T, Edwards S, editors. Handbook of the Philosophy of Medicine. Dordrecht: Springer Netherlands; 2017. p. 129-45.

13.   Nickel B, Moynihan R, Barratt A, Brito JP, McCaffery K. Renaming low risk conditions labelled as cancer. BMJ. 2018;362:k3322.

14.   Rogers WA, Mintzker Y. Getting clearer on overdiagnosis. Journal of evaluation in clinical practice. 2016.

15.   Rogers WA, Mintzker Y. Casting the net too wide on overdiagnosis: benefits, burdens and non-harmful disease. Journal of medical ethics. 2016(42):717-9.

16.   Rogers W, Walker M. Précising definitions as a way to combat overdiagnosis. Journal of evaluation in clinical practice. 2018;24(5):1019-25.

17.   Hersch J, Barratt A, Jansen J, Irwig L, McGeechan K, Jacklyn G, et al. Use of a decision aid including information on overdetection to support informed choice about breast cancer screening: a randomised controlled trial. Lancet. 2015(385):1642–52.

18.   Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Medical decision making. 1991;11(2):88-94.

19.   Hatch S. Uncertainty in medicine. BMJ (Clinical research ed). 2017;357.

20.   Macdonald H. Navigating uncertainty. BMJ (Clinical research ed). 2017;357.

21.   Hofmann B. Diagnosing overdiagnosis: conceptual challenges and suggested solutions. European Journal of Epidemiology. 2014;29(9):599-604.

Tables

Table 1   Topics of disagreement and examples of challenges with defining overdiagnosis

Challenge / Topic of Disagreement	Explanation / Example
Competing mortalities	Is a person who dies from a car crash before he dies from the identified disease overdiagnosed?
Including treatment outcomes	Does the extension of overdiagnosis depend on treatment outcome?
Non-harmful disease	Can you have a correct diagnosis without having the disease?
Medicalization versus overdiagnosis	Is medicalization part of overdiagnosis?
Unity	Is it possible to define a general or unifying concept of overdiagnosis, or do we have to rely on specific definitions for specific purposes?
Symptoms	Can persons with symptoms be overdiagnosed or does overdiagnosis only apply to person without symptoms?

Table 2   How challenges are addressed by the three approaches to overdiagnosis

Question	Approach 1	Approach 2	Approach 3
Is a person who dies from a car crash before he dies from the identified disease overdiagnosed?	NO	NO	NO
Does the extension of overdiagnosis depend on treatment outcome?	NO	NO	NO
Can you have a correct diagnosis without having the disease?	NO	YES	YES
Is it possible to define a general or unifying concept of overdiagnosis, or do we have to rely on specific definitions for specific purposes?	YES	NO	NO
Is medicalization part of overdiagnosis?	NO	NO	YES
Can persons with symptoms be overdiagnosed?	YES	YES	YES