A definition and ethical evaluation of overdiagnosis: response to commentaries

Carter, S. M., J. Doust, C. Degeling and A. Barratt (2016)

It is a privilege to have respected colleagues engage with our ideas: such exchanges are critical to moving this new field forward. First, some paradigmatic issues.

1. Rogers and Mintzker[1] are correct to say we reject an objectivist conception of disease. However on p. 5, the authors imply that we think professional communities knowingly reject realism in developing disease definitions. This is not the case: rather we suspect that greater reflective insight into the constructedness of definitions in the professions may help counter the institutionalisation of overdiagnoses.

2. Hofmann[2] proposes that when professionals agree that diagnoses are correct (the first part of our definition), they are guided by knowledge about consequences (the second part of our definition): in his view this makes the second part of our definition redundant (p. 1 and throughout). We agree that, ideally, diagnoses would track consequences: unfortunately they often do not. Moynihan and colleagues, for example, have shown that expert panels rarely consider the consequences (harms particularly) of the disease definitions—or redefinitions—that they propose.[3] Diagnoses, as argued, are negotiated social and political achievements determined on multiple grounds beyond outcomes, thus we reaffirm a distinction between correctness and consequences.

3. Hofmann, similarly, rejects our observation that sometimes (note, not always) the distinction between overdiagnosis and false positives is blurred. In an ideal world this would be a bright line, but again, in practice, it is not. As we observed, a gold standard does not always exist, and for some conditions there are competing gold standards (e.g. Polycystic Ovary Syndrome: PCOS).[4] In such cases the distinction will, unfortunately, be unclear.

4. On the distinction between medicalisation and overdiagnosis raised by Morrison,[5] we have found earlier work by Hofmann extremely useful.[6]

Overdiagnosis as harmless disease

Rogers and Mintzker's central concern (p. 2) is that we have widened the definition of overdiagnosis 'to include any conditions in which there is a negative balance of benefits to harms, irrespective of the nature of the disease in question'.[2] We agree. The authors contend that this is too broad; we respectfully counter that their definition is too narrow, for several reasons.

First, overdiagnosis does not only apply to diagnosis of disease, but also increased risk (as Morrison notes[5]), conditions, and disorders.

Diagnosis of harmless disease is one type of overdiagnosis. The most obvious example is diagnosis of indolent or regressive (non-progressive) cancers. These would never become symptomatic so if detected through screening they are always overdiagnosed.[7] Epidemiologists refer to the bias introduced by such overdetection as length time bias; natural history models of the frequency of overdiagnosis often neglect it.[7-10]

However this is only a subset of overdiagnosis.

Some overdiagnosis is of harmful conditions, but where the harms likely to arise from the diagnosis, and subsequent treatment, exceed the harms produced by the condition itself. Examples are numerous, and include Attention Deficit Hyperactivity Disorder (ADHD), genomic variation, low testosterone, mild depression, PCOS, mild blood pressure, and mildly elevated cholesterol. The symptoms that prompt a diagnosis of ADHD, for example, are not harmless, involving educational and behavioural challenges that limit children's future interests. However, it seems likely that more harm than good is being done by diagnosing and medicating large numbers of children, particularly those with mild symptoms.[11] The problem is that this diagnosis, currently considered correct, is producing an unfavourable balance of benefits to harms, not that ADHD is a harmless disease.

Another subset of overdiagnosis is overdiagnosis due to the interaction between competing mortality and rate of progression of the diagnosed disease.[8] This is well-described in cancer epidemiology, for example, having been incorporated into at least six established models of breast cancer natural history.[12] It occurs when a progressive (i.e. harmful) disease (such as cancer) is correctly diagnosed, but the person dies first of another cause. For example, a woman is diagnosed with early breast cancer via mammography, receives treatment, but dies from a heart attack before the breast cancer would—if undetected—have become symptomatic. Rogers and Mintzker[1] (p. 4) and Hoffman[2] (p. 1) assert that diagnosis and overdiagnosis should be linked to prognosis. In a perfect world this would be so, but it is: a) often not the case; and b) sometimes not possible because of overdiagnosis due to competing mortality, which is acknowledged as an inherent risk of early detection. Other examples include the diagnosis of elevated cholesterol or blood pressure in a patient with end-stage dementia.

Note that the commonly-observed discrepancy between increasing incidence and unchanged mortality in overdiagnosed conditions (Rogers and Mintzker[1] p. 6) is not evidence that the disease is harmless, but that the diagnosis does not effectively decrease mortality.

In our view, our definition captures the many types of overdiagnosis, while a definition limited to diagnosis of harmless disease does not.

Implications of a broad definition of overdiagnosis based on the balance between benefits and harms

Rogers and Minztker express concern about three main implications of our wider definition.[1] We agree that these implications follow from our definition, but find them unproblematic.

1) The status of diagnoses (as overdiagnoses versus beneficial diagnoses) may change over time. We see this as inevitable, due to the constantly evolving nature of medical knowledge and practice.

2) Sometimes overdiagnosis arises from features of the healthcare system rather than biological features of the disease. We accept this and believe it to be appropriate. For example, in very low income countries without functioning healthcare systems, a cancer diagnosis may provide nothing beyond a label, as basic care including cancer treatment, pain relief, psychological support, or palliation are unavailable. In this case, in our view, a cancer diagnosis would be an overdiagnosis. It may be a form of tragic overdiagnosis that had not previously occurred to us: the practitioner diagnoses correctly, and aims to serve her patients' best interests, but unavoidably the diagnoses provide no benefit (and may harm by adding psychological suffering). If that health system changed so that cancer patients routinely received needed care, overdiagnosis arising from lack of access to care would stop. This highlights the structural nature of overdiagnosis and its solutions.

3) On our definition, the effectiveness of treatment partly determines whether overdiagnosis occurs. Thus when routine treatments for certain diseases were more harmful than the diseases themselves, those diseases, on our definition, were always overdiagnosed. (This may well take in much of the history of medicine, including the example of starvation therapy for type 1 diabetes.) If treatments become more benign, the balance would shift. Rogers and Mintzker ask (p. 4): if treatment for early breast cancer was easy, effective and cheap, would mammographic screening still produce overdiagnosis? We think it would not. Invasive breast cancer can cause suffering and fatality if left untreated; current treatment for early breast cancer produces psychological, physical and financial harms and opportunity costs. Overdiagnosis is currently thought to be occurring in populations of screened women based on a potentially unfavourable or marginal balance between the number of women spared early breast cancer death versus the number of women harmed by treatment.[13] Imagine, instead, that women could treat early breast cancer by swallowing one affordable, highly effective pill that had no side effects. The balance between breast cancer harms and treatment harms would radically shift, and—we contend—overdiagnosis

would no longer occur. Conversely, Rogers and Mintzker suggest that under our definition, Huntington's Disease (a genetic disease which usually becomes symptomatic only in midlife) would always be overdiagnosed, because the disease is incurable. We disagree. Although it cannot be cured, in a well-functioning health system, a Huntington's Disease diagnosis can be on-balance beneficial (and thus not overdiagnosis). Diagnosis can provide an explanation of symptoms, reassurance and understanding, access to reproductive autonomy, and services and treatments that would make life better with, than without, a diagnosis. To summarise, we are convinced that identifying overdiagnosis depends on recognising the balance of all of the broad medical, social and psychological benefits and harms that will arise from correct diagnosis and consequent treatment in a particular context, compared with the consequences of not diagnosing.

Evaluating and balancing benefits and harms

Both Hofmann[2] and Rogers and Mintzker[1] express concern regarding the complexity of evaluating and balancing benefits and harms. We agree that it is unfortunate that this task is so difficult, but can't see any way of avoiding it. Our section on judging benefits and harms (p. 3-5) approaches the problem, and argues that decisions will likely combine experiential, technical, and deliberative approaches, avoiding over-reliance on a single perspective. Hofmann[2] expressed concern regarding community juries (p. 2) and subjective consumer judgements (p. 2): we reaffirm that these should be part, but only part, of the answer to this difficult problem (we also need, for example, more transparent and systematic expert processes for defining diagnoses, and better empirical evidence about harms). Rogers and Mintzker[14] (p. 6) suggest that calculating benefits and harms in populations is impossible because of variation in individuals' values. This precise problem is a constant challenge to Evidence Based Medicine (EBM)[15]; in our view this has prompted continuous incremental refinement of EBM rather than its abandonment, and the same seems possible here. The line between beneficial diagnosis and overdiagnosis will never be absolute, and at both a population and an individual level, as Morrison notes,[5] there will always be uncertainty in this judgement. Judgement is unavoidable nonetheless.

A population-level definition

Rogers and Mintzker[1] and Hofmann[2] each raise questions about our decision to define overdiagnosis at a population, rather than an individual, level. As we discussed, overdiagnosis is done to, and affects individuals (and their families), but can, currently, only be estimated (quantified, or 'statistically inferred' as Hofmann notes on p. 2) by examining population-level data. In well-financed healthcare systems, most diagnosed individuals will receive something like standard care; it is thus impossible to know what would have happened had the condition been left undetected and untreated (the non-identifiability problem). With future knowledge we may be able to accurately identify non-progressive, preclinical disease and thus be able to quantify overdiagnosis due to 'harmless' disease at an individual level, at the time of diagnosis. If achieved (and this new knowledge was incorporated into diagnostic standards) this could reduce the problem of overtreatment and potentially of overdiagnosis. Overdiagnosis due to competing mortality, however, will never be knowable at an individual level, absent circumstances in which no intervention is offered to people with progressive clinical disease. So

at least some overdiagnosis will only ever be statistically inferred in populations, at least for the foreseeable future. And, as we noted, overdiagnosis entails corporate harms via the re-orientation of health systems, which only make sense at a population level.

We do not, as Hofmann suggests, wish to prevent individuals from receiving probabilistic risk information from population-based studies (p.2). Evidence suggests that well-designed decision aids (based on population-based evidence) can improve individuals' knowledge, reduce decisional conflict, and increase accuracy of risk perception.[16] That such aids cannot remove the non-identifiability problem is not, in our view, a reason to abandon risk communication altogether.

Relevant professional communities

Both Hofmann[2] and Rogers and Mintzker[1] raise questions about the 'relevant professional community' that might make a judgement about the correctness of a diagnostic label. We agree that a lot turns on 'relevance', and that this needs further elaboration. Our rough working conception is that a relevant professional community has: 1) critical mass; 2) shared standards for diagnosis (whether formally developed or instantiated through practice); and 3) ability to apply such diagnoses in practice (regulated or not). This deliberately elastic conception extends beyond formal structures of western medical authority (e.g. it could include various complementary and alternative medical practitioners, and thus make it possible for them to engage in overdiagnosis); however it also deliberately excludes idiosyncratic outliers (mavericks not practicing according to any standard).

Relevance is unavoidably arguable. On p. 2-3, for example, Hofmann asserts that there is no relevant professional community that supports diagnoses of Low T. Contra this, others see the Low T campaign as a textbook exercise in building off-label prescription;[17] and evidence suggests that this greatly increased prescribing rates.[18] This could not have occurred without a professional community willing to adopt 'Low T' as a correct diagnosis. We emphasise that in our definition, the relevant professional community need only accept that the diagnosis is *correct.* As discussed above, we do not assume that this shared view of correctness contains any considered agreement about benefits, harms or prognosis. This, we believe, obviates the concerns Hofmann expresses regarding misdirected and tragic diagnosis on p.3. However we agree with Hofmann, and Rogers and Mintzker, that further elaboration of the 'relevant professional community' component of our definition is needed.

The normative typology

As acknowledged in the commentaries, the three types are ideal: we do not suggest that they will ever be perfectly expressed, and the degree to which instances of overdiagnosis have tragic, misdirected or predatory features will always be arguable. Hofmann is concerned (p. 3) about reliance on un-knowable intentions, a well-worked normative problem. We note that, as is common in public health ethics (e.g. [19-21]) our analysis relies not on intentions but on goals. It is also not uncommon, through legal discovery and other means, for evidence of relatively predatory collective or institutional goals to become available, particularly in respect of commercial actors.

We expect that it will be some time before these matters are settled, and look forward to continuing to debate them with our colleagues.

References

1. Rogers W, Mintzker Y. Casting the net too wide on overdiagnosis: benefits, burdens and non-harmful disease. J Med Ethics. 2016 (This issue).

2. Hofmann B. Defining and evaluating overdiagnosis. J Med Ethics. 2016 (This issue).

3. Moynihan RN, Cooke GPE, Doust JA, Bero L, Hill S, Glasziou PP. Expanding Disease Definitions in Guidelines and Expert Panel Ties to Industry: A Cross-sectional Study of Common Conditions in the United States. PLoS Med. 2013;10(8):e1001500.

4. Dumesic DA, Oberfield SE, Stener-Victorin E, Marshall JC, Laven JS, Legro RS. Scientific Statement on the Diagnostic Criteria, Epidemiology, Pathophysiology, and Molecular Genetics of Polycystic Ovary Syndrome. Endocr Rev. 2015;36(5):487-525.

5. Morrison M. Overdiagnosis, medicalisation and social justice. J Med Ethics. 2016 (This issue).

6. Hofmann B. Medicalization and overdiagnosis: different but alike. Medicine, Health Care and Philosophy. 2016:1-12.

7. Etzioni R, Gulati R. Oversimplifying Overdiagnosis. J Gen Intern Med. 2014;29(9):1218-20.

8. Welch HG, Black WC. Overdiagnosis in Cancer. J Natl Cancer Inst. 2010;102(9):605-13.

9. Baker SG, Prorok PC, Kramer BS. Lead Time and Overdiagnosis. J Natl Cancer Inst. 2014;106(12).

10. Zahl PH, Jorgensen KJ, Gotzsche PC. Overestimated lead times in cancer screening has led to substantial underestimation of overdiagnosis. Br J Cancer. 2013;109(7):2014-9.

11. Thomas R, Mitchell GK, Batstra L. Attention-deficit/hyperactivity disorder: are we helping or harming?2013 2013-11-05 23:31:21.

12. Mandelblatt J, Cronin K, Koning HD, Miglioretti DL, Schechter C, Stout N, et al. Collaborative Modeling of U.S. Breast Cancer Screening Strategies. Rockville, MD: Agency for Healthcare Research and Quality US Department of Health and Human Services, 2015.

13. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. The Lancet. 2012;380(9855):1778-86.

14. Rogers WA, Mintzker Y. Getting clearer on overdiagnosis. J Eval Clin Pract. 2016.

15. Greenhalgh T, Snow R, Ryan S, Rees S, Salisbury H. Six 'biases' against patients and carers in evidence-based medicine. BMC Med. 2015;13(1):1-11.

16. Stacey D, Legare F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2014(1).

17. Schwartz LM, Woloshin S. Low "T" as in "template": how to sell disease. JAMA Intern Med. 2013;173(15):1460-2.

18. Baillargeon J, Urban RJ, Ottenbacher KJ, Pierson KS, Goodwin JS. Trends in androgen prescribing in the United States, 2001 to 2011. JAMA Intern Med. 2013;173(15):1465-6.

19. Kass NE. An ethics framework for public health. Am J Public Health. 2001;91(11):1776-82.

20. Ortmann LW, Barrett DH, Saenz C, Bernheim RG, Dawson A, Valentine JA, et al. Public Health Ethics: Global Cases, Practice, and Context. In: Barrett DH, Ortmann LH, Dawson A, Saenz C, Reis A, Bolan G, editors. Public Health Ethics: Cases Spanning the Globe. Switzerland: Springer 2016. p. 3-35.

21. Juth N, Munthe C. The Ethics of Screening in Health Care and Medicine: Serving Society or Serving the Patient? Springer; 2012.