phg foundation
making science work for health

My healthy future

overdiagnosis

UNIVERSITY OF CAMBRIDGE

Contents

1 Introduction

The PHG Foundation programme *My healthy future* imagines a future health system where new technologies enable individuals to acquire and act on a wide range of information about themselves, their health and their risk of disease. When ill, these technologies will allow them to more precisely understand the nature of their disease, its physical, psychological and social impact and the best means of treatment, management or mitigation. Such a future is built on the biological and technological developments that enable more effective and targeted personalised medicine (sometimes called precision, stratified or P4 medicine).

In *My healthy future* we are focusing on personalised prevention, in which preventive interventions are offered according to individual risk and are personally tailored for optimum benefit. Personalised prevention complements population-based prevention programmes. It aligns with the Government's recent initiative on 'predictive prevention' as set out in the recent report *Prevention is better than cure*[1], which "envisions a world where everyone can understand their own risks", and all are offered "precise and targeted health advice - specifically designed for their demographic and their location; their lifestyle and their circumstance; and their health needs and goals".

We are imagining a future approximately twenty years from now, when individuals have access to regular personal monitoring through an abundance of devices and sensors, and will customarily access various forms of health monitoring and testing. Although some testing may be provided through health systems such as the NHS, in our preparatory work we have also identified technologies that provide information in new ways, many from commercial providers. These include, for example: continuous heart-monitoring through personal, wearable devices; examination of biomarkers through sensors in clothing; imaging-based diagnostics provided via artificial intelligence for diseases such as cancer; analysis of big data in health information systems to alert individuals about possible risk; and genomic testing via commercial providers.

To promote personalised prevention these technologies are designed to measure indicators of early disease or disease risk. However, many would caution that benefits might be offset as users are made anxious and undertake actions and interventions that might not eventually be beneficial to them. The phenomena through which this occurs are often gathered together under the general theme of 'overdiagnosis' or 'too much medicine'.

In this document, we are concerned with the many ways in which harm arises from activities grouped under the theme. We have chosen to use the term overdiagnosis as the overarching label. Sections include:

- Definitions and related concepts
- Relation between risk and disease
- Harms of overdiagnosis
- Origins and drivers of overdiagnosis
- The role of new technologies in overdiagnosis
- Detection and measurement of overdiagnosis
- A discussion from the literature of ways in which overdiagnosis may be counteracted or mitigated

We conclude with a set of questions that stem from the apparent inevitability that the introduction of new technologies for personalised prevention will increase overdiagnosis. We also seek ways of mitigating the harms that may arise.

2 Definitions and related concepts

2.1 Concepts of diagnosis

There is much debate surrounding what is meant by the 'diagnosis' of disease. For the purposes of this report, we regard it as a form of labeling, a shorthand used by physicians to classify 'abnormal' biology so that they can better manage their patients and determine with greater accuracy their prognosis.

This so called 'nominalist' approach is to be distinguished from the 'essentialist' perspective wherein diagnosis is seen as discovery - the act of discovering 'disease', itself conceptualised as a real entity pre-categorised within the body. The 'nominalist' approach admits of an (abnormal) state of biology that is the reason for a patient's signs and symptoms; what is denied is that the categorisation of that abnormal biology as different diseases is real, rather than purely a label applied to those different states of biology.

Philosophers of language such as Frege[2] drew a distinction between the reference and the sense in the use of a word. The reference is the object meant or indicated which has an underlying truth value; the sense is the thought that this expresses. In the case of disease, the reference would be the abnormal biology (which has an underlying reality) and the sense is the label (or 'disease' name) we attach to it.

If we consider 'diagnosis' as the application of a label to a supposed underlying abnormal biology, it could be suggested that we can wrongly label but we cannot 'over' or 'under' label. However, we can get it wrong in a variety of ways.

Below are three ways in which a misdiagnosis may result:

- Labeling the condition as Y when it is in fact X

  This is misdiagnosis in its simple and unconditional form. We misinterpret the symptoms and signs and label a patient's breathlessness as heart failure when it is in fact pneumonia or vice versa. We have incorrectly applied the set of rules taught to physicians to the patient's symptoms and signs.

- Applying labels that are too broad to a range of conditions that should be managed differently and have different prognoses

  This is best exemplified by prostate cancer. The diagnosis is correct but under that label there are different states of abnormal biology. It is our lack of knowledge that prevents us from giving more precise labels. Historically, the word 'dropsy' was used but we now know that the symptoms of what was called 'dropsy' can be caused by many different states of abnormal biology and we use the term no more.

- Labeling the patient as having a disease X when their underlying biology is in reality normal; or as not having disease when the patient has the underlying abnormal biology that one would normally label as disease X

> Essentially this is a misclassification error resulting from failure, through our use of tests, to determine whether a patient has a disease or not. This may arise in one of two ways. First, because we are imprecise in defining the boundary between the normal and the abnormal, for example the boundary between hypertension and normal blood pressure. Second, because of the imperfect nature of diagnostic tests themselves. Test sensitivity of 90% means that the test will only correctly identify nine out of ten of the people with disease; test specificity of 90% means that one out of every ten people who do not have the disease are incorrectly identified as having the disease. The test itself is the cause of the misclassification.

The purpose of presenting this theoretical stance is to shift the focus of the conversation to the harm that can be done to the patient. If we mislabel we have the potential to harm.

2.2 Diagnosis, overdiagnosis and the literature

Most definitions of overdiagnosis include a 'correct' diagnosis of a condition, usually with the offer of an intervention and with the subsequent finding of lack of benefit. Authors differ in the way in which they describe lack of benefit – this may be that 'the disease or condition never causes symptoms' (for example an indolent cancer that never progresses to a clinical presentation) whilst others describe this as having no 'net benefit' for the person.

A simplified definition considers 'diagnosis' to be a label which we attach to a finding of abnormal biology. A 'misdiagnosis' is when we apply a diagnostic label that leads to clinical management where there is no evidence of clinical benefit. This could include 'watchful waiting'. One of the problems is that, while a technology may enable us to make a diagnosisthere may not be a known pathway for treating, managing or preventing the condition .

The literature also introduces a number of related concepts, which bring nuances to the discussion[3] – for example:

- **'Overtreatment' -** where treatments are unnecessary and do not improve the condition or may cause harm
- **'Medicalisation' -** where experiences and human problems are interpreted as medical problems without net benefit
- **Expanded definitions of disease -** where new conditions are created and new diseases may be promoted to the public, either intentionally or unintentionally, to encourage use of health services, especially tests and medicines.

In the UK, overdiagnosis and related concepts are often discussed under the term 'too much medicine', with the underlying consensus that 'too much medicine' is problematic for individuals, society and health systems. Commentators have gone on to discuss the harms that can be caused, the origin and drivers, including the role of new technologies, methods for detection and measurement, and ways in which this phenomenon could be counteracted or mitigated. These are relevant for our consideration of the introduction of new technologies to personalise the prediction and prevention of disease, in particular to the ways in which possible harms may be mitigated.

3 Risk and personalised prevention

The risk of disease is the probability or likelihood of its occurrence. In epidemiology, risk of disease is measured by the occurrence of new cases of disease in a defined population over a defined period.

The WHO defines risk factors as any attribute, characteristic or exposure of an individual that increases the likelihood of developing a disease or injury. As well as demographic factors such as age, sex and ethnicity, risk factors include underlying genetic endowment, lifestyle and environmental exposures.

As we move into an era of personalised prevention, more tools are available to assess individual risk in order to suggest or provide risk-reducing interventions. Risk assessments might include the standard epidemiological risk factors augmented by a range of biomarkers providing information on physiological or clinical characteristics, alongside an increasing use of information derived from the analysis (automated or otherwise) of large clinical datasets.

Risk assessment tools may vary in their level of sophistication. To be useful, all must enable individuals to be placed into groups according to risk (stratified) and then to be offered different interventions or no intervention, according to the likely benefit.

3.1 Risk and harm

With respect to the general area of overdiagnosis two main problems arise:

1. The very act of assessing risk may cause harm by raising anxiety in the individual about a disease that may never affect them, even when they turn out to be 'low' or 'normal' risk. Of those who are told that they are high risk, by definition (as risk is a probability) some will not get the disease but all will have increased anxiety and many will undergo investigation, treatments or make lifestyle changes as a result of the information.
2. Problems arise because of the tendency for risk to become conflated with, or treated as, the disease itself. This may be related to the ways in which we understand the underlying biology and pathology and how we define particular diseases, but it is also related to what is understood or conveyed to the patient/individual/subject and the subsequent actions that are recommended, which may tend to medicalise the situation or may themselves cause harm.

   Examples include:

   - **Risk as measured on physiological spectrum -** hypertension is defined in terms of position on a spectrum of physiological measurement. In hypertension, the purpose of identification and treatment of elevated blood pressure is broadly to prevent future problems including heart disease and stroke. For this example, whether or not an individual is said to have the condition 'high blood pressure', is primarily definitional and in practice determined by the pronouncement of learned societies. Arguably the presence of high blood pressure is indicative of an underlying problem, for example a high resistance to blood flow caused by low elasticity of small arteries and arterioles and poor functioning of the renin-angiotensin system of the kidney. This could constitute disease in itself rather than just being a risk factor. In addition, the individual may be recommended to take antihypertensive medication, which also leads them to becoming a patient and the potential for side effects.

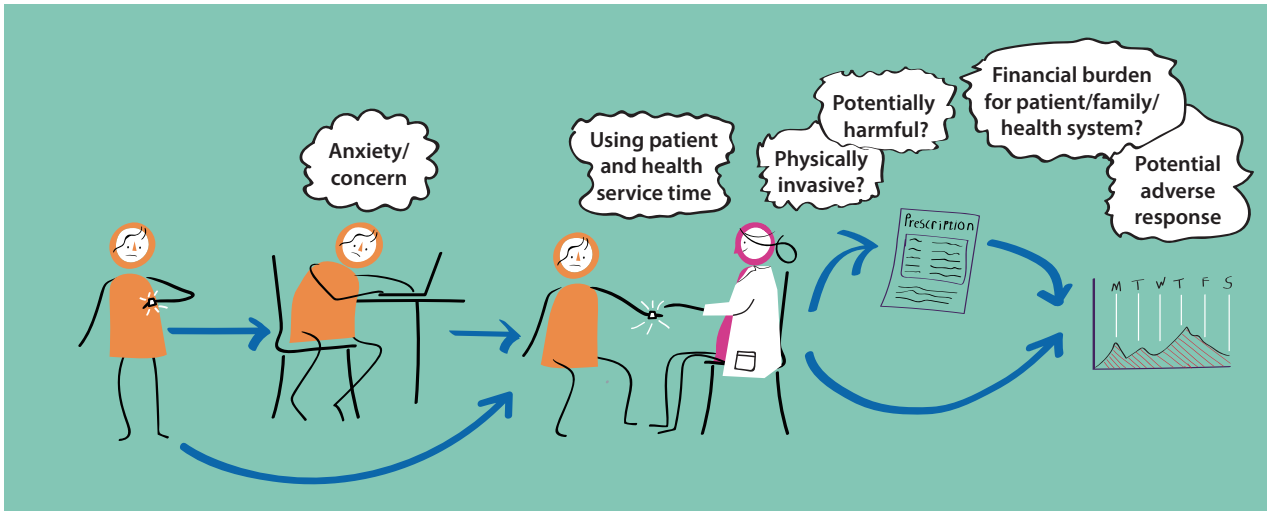
- **Risk as genetic variation -** if an individual has a *BRCA1* mutation, they have abnormal biology, but most would agree they have increased risk rather than disease. However, they may be recommended to take preventive medication or even to undergo surgery to reduce their risk of breast cancer – which inevitably leads to them becoming a patient and experiencing adverse effects.
- **Risk as adverse lifestyle factor -** the individual who is a smoker, or who has a poor diet and is obese, does not have a disease although they may be at increased risk. Interventions that alert them to risk and the provision of lifestyle advice and support may have the effect of medicalising the individual.

4 Harms of overdiagnosis

A harm can be defined as 'a wrongful setback to a person's interests'[4]. An interest in one's own health is known as a 'welfare interest', a requisite for wellbeing, and when blocked or damaged this can cause a person serious harm. Individuals have interests in avoiding unnecessary medical interventions, maintaining mental and emotional health etc. However, shared public interests also carry moral weight and have to be balanced against those of the individual. Currently, as a community we accept some risk of harm due to the overall benefit gained in identifying and treating more people with disease.

The balance of benefit and harm should be considered further as it is central to debate surrounding overdiagnosis. This requires an analysis of what types of benefits and harms matter, how they should be weighed, their magnitude and relative importance, and whether they should be measured in individuals or society.

The perspective from which the nature of harms and their balance with benefits are assessed is also of importance. Individuals will differ in how they assess these, being influenced by their underlying values. In turn, these values will be determined by innate personality, upbringing, and cultural and religious beliefs. Harms to the patient may be physical, psychological or social.



Physical harms may arise directly from the investigations that might follow identification of an abnormality or risk, from the treatment of a condition that may not have caused any harm, and from the investigation or treatment of other conditions that might also be revealed.

Psychological harms are harder to measure, but may begin at the point of first interaction with the test provider, when anxiety is raised about the potential condition and the patient has to engage with health systems for discussions, testing, waiting for results and dealing with positive or negative findings.

The psychological effects of diagnosis and treatment of the disease which is 'found' in the overdiagnosed (or wrongly diagnosed) patient are similar to those for a correct diagnosis. Social effects arise from being 'labelled' with a disease, but also include the effects of needing to take time out of work, family or other activities for diagnosis and management and future perceptions about their own health and wellbeing.

Overdiagnosis and overtreatment also have major effects on the healthcare system by diverting resources from effective to less effective care and overall to a less effective healthcare system that may also lose the trust of consumers.

However, given our focus on personalised prevention and the desire to focus on the needs and values of an individual, we should also keep in mind that individuals may differ in the way in which they rate benefits and harms, and this may in turn differ from the perspectives of healthcare providers. For example, in the case of a test result that gives knowledge of a potential future disease that is untreatable and where the future trajectory cannot be modified, the health system may decide that there are no benefits and that the harms are major. However, the patient may still value that knowledge for its own sake and also to enable future planning.

Patients may argue that they simply have the 'right to know' information about themselves. Individuals will differ in the value that they place on such knowledge and some may prefer not to know. In these circumstances, there is a question as to whether these patients are overdiagnosed, and whether the health system should seek to avoid this.

5 Origins and drivers of overdiagnosis

In the context of personalised prevention, problems under the theme of overdiagnosis arise directly from the expansion of attempts to measure risk or early disease. In particular, when risk assessment or early disease detection is offered to asymptomatic individuals or when individuals consult with symptoms and an abnormality is found, that may or may not be related to the presenting symptoms, this may prompt the physician to undertake further investigations. Such tests may be useful but could just as easily result in harm.

As noted in section 2, some level of overdiagnosis (or misdiagnosis) is intrinsic to every diagnostic test. None is perfect and the imperfection is defined by the *sensitivity* of the test i.e. its ability to correctly identify that the disease is present in an individua and by its *specificity* i.e. where normal individuals are wrongly diagnosed as having the disease.

Several factors work independently, and together, to drive overdiagnosis. They can be grouped into:

- Those that operate at the level of policy
- Those arising from the introduction of new technologies
- Those related to current lack of knowledge about diseases and their natural histories
- A set of issues related to professional practice, citizen knowledge and wider societal issues[5].

Belief systems around the phrase 'prevention is better than cure' have been widely used in policies that seek to redress the balance of healthcare. For those policies aimed at the structural determinants of health or at populations, there are few downsides for individuals. However, when aimed at individuals there is potential for harm arising in the general area of overdiagnosis.

Although in the UK the National Screening Programme follows strict criteria and pays great attention to minimising false positives and overdiagnosis, this still occurs to some extent. A focus on identifying and trying to influence individual lifestyles or precursors of disease, such as might take place within the Healthcheck programme will inevitably identify (and may harm) some individuals who will not progress to disease. In addition, instilling principles around the importance of individual prevention makes it difficult for people to make a reasoned decision not to take up screening offers, and, if early disease or increased risk is found, not to believe that their life has been saved by the intervention. This belief might be in spite of the fact that the intervention may have resulted in side effects that the early disease may never have caused harm, or that the disease would have taken the same natural course whether or not the intervention had taken place.

Overdiagnosis is not intrinsic to new technologies, however the introduction or expansion in use of technologies may drive overdiagnosis; this is discussed in more detail in section 6. Factors discussed in later sections are particular to the introduction of new, or the expansion of current, technologies; these are:

- Higher sensitivity of testing
- Increased resolution of tests
- Greater availability of testing
- Lack of accompanying improvement in treatments

Over and above the new technologies themselves is the fact that their development is often driven by commercial interests. In their attempts to deliver profits for shareholders, companies will aim to expand the applications of diagnostics to patients or individuals with lower prior risk of disease, thus inevitably increasing the overdiagnosis rates.

Abnormalities revealed by these new technologies can also lead to overdiagnosis because knowledge of the natural history or the response to treatment has not kept up. Very often we are unable to distinguish cases where the abnormality will progress to harmful disease and those where it will not. When deciding on further investigations and treatments there is a tendency to err on the side of caution. In individual interactions this may arise as a result of defensive medicine, or simply from the influence of over-enthusiastic promotion of new diagnostics.

On the patients' side it is suggested that patients rely too much on tests for reassurance and often have high expectations that the doctor will do something – favouring action over inaction. However, concern not to miss disease leads to sensitivity being favoured over specificity – underdiagnosis is minimised at the expense of overdiagnosis. This may particularly happen in conditions where abnormalities underlying disease exist as part of a spectrum of normal measurements such as in hypertension.

Within this new environment of personalised prevention the individual at the centre will be faced with decisions about what information to seek and how to act on any findings. Most consumers, however, will be poorly positioned to consider options and alternatives and will be likely to rely on information from a provider (sometimes also a seller).

6 New technologies and overdiagnosis

In this section we consider some general features of new technologies that may lead to overdiagnosis and provide some more specific examples.

6.1 General drivers of overdiagnosis

The primary aim of the workshop is to explore the impact that the introduction or expansion of health-related technologies used for personalised prevention might have on the phenomenon of overdiagnosis in the next 20 years. Health-related technologies come in many forms, from biosensors and other increasingly sensitive monitoring tools to intelligent algorithms and the application of genomics for personalising prevention and treatment.

Advances in genomic medicine could be considered to present a special case in considerations surrounding the use of tests and the potential for harm. Whole genome sequencing in particular could provide an individual with a vast amount of highly personal information applicable in different ways across an individual's lifetime, for which we currently have limited evidence surrounding actionable health-related elements. Genetic tests bring distinct issues to the discussion of overdiagnosis through their ability to provide information about potential for diseases that may occur many years into the future and that is relevant both to the person tested and to family members.

Some examples of technology areas and characteristics of new technologies that could lead to overdiagnosis in the future are given below for consideration:

- Additional 'omics tests including: transcriptomics, epigenomics, nutrigenomics, and microbiome analysis
- Liquid biopsy and circulating tumour DNA (ctDNA) tests - testing presence or type of cancer biomarkers
- Mobile biosensors including wearables, handheld devices, and implantable sensors
- Artificial intelligence and machine learning for examination of potentially diagnostic information e.g. for tumour detection in radiology imaging
- Portable diagnostics which facilitate one or more of: remote serial monitoring, point of care testing, at-home testing, and on-demand testing
- Big data and the internet of things – increased collection and analysis of personal and clinical data and increased connectivity between different data sources

Below we consider some examples of ways in which new technologies could have an impact on the prevalence of overdiagnosis.

Higher resolution may reveal new subtleties or detail in results which had previously not been apparent, leading to uncertainty about appropriate action on behalf of the patient and clinician. This is relevant across a range of new diagnostic technologies including, for example, 'omics' testing, digital radiology imaging, liquid biopsy. Although these should eventually be somewhat resolved by reevaluation, until this occurs there is a tendency towards action rather than inaction on a cautionary basis.

Greater availability and access to testing, wider use of testing, or ability to test more frequently could lead to increased cases of overdiagnosis through increased distance from associated medical advice or widening the criteria for testing to include people at lower prior risk. This may be especially true in cases where technologies provide information without sufficient context, for example information coming from mobile sensors and digital technologies such as wearables, implantable sensors and compact mobile monitors.

'**Mission creep**' in which there is broader application of tests to conditions for which they were not originally developed. This is not exclusive to new technologies, but can occur with any technology if left unchecked.

Improved monitoring without accompanying improvements in treatment/action: If treatment or health management options do not keep pace with advances in diagnostics, individuals could more readily be given labels or diagnoses that do not benefit them. This is relevant across all diagnostic technologies.

Improved identification of those truly at risk: One way in which technologies such as genomic testing might be able to mitigate some of the potential harms of overdiagnosis is through improved identification of those at risk and, equally, exclusion from further testing of those at lower risk. Higher precision or additional alternative diagnostics may result in clinicians being more confident in their decisions to exclude patients from potentially harmful treatments or actions.

6.2 Genetic and genomic testing

Genetic testing is currently undertaken to diagnose disease or to predict serious disease with the opportunity to offer a specific intervention. Where prediction is offered to healthy individuals this is largely in the context of a clinical encounter – for example, testing of relatives where a family member is diagnosed with a rare single-gene disorder (for example a hereditary cancer syndrome such as hereditary breast or ovarian cancer), or antenatal testing of a fetus where parents are known to be affected.

Opportunities for screening, where genetic testing is offered to a population of healthy people are more limited, but may include the offer of preconception carrier testing to certain ethnic groups such as those of Mediterranean origin who may carry sickle cell or thalassaemia, or the universal offer of Down syndrome screening during pregnancy. The potential for overdiagnosis has, hitherto, been managed (although not obliterated) by: strict consideration of screening criteria and formal evaluation of genetic tests that will be offered; and strict criteria for testing, target disease and evidence of clinical utility.

However, as testing has become quicker, cheaper and more readily available, and tests include more extensive analysis of the genome, there is a widespread urge to use this for further prevention – most notably by offering testing and the return of secondary findings beyond those relevant to the immediate clinical question. The potential of offering whole genome sequencing to entire populations (e.g. of newborns) is also being discussed.

Why such expansion might be considered problematic is that the positive predictive value of a test is related to the prevalence of the disease in the population being tested. So the use of genetic testing in family members where others have the disease (i.e. the prevalence in that family is higher), or in racial groups with a high prevalence of a particular disease, is more likely to be justified (the positive predictive value will be higher) than in general populations where there is no reason to suspect that the disease may be more prevalent.

Whenever genetic or genomic testing is offered to individuals or populations at low risk for the condition there is a tendency for overdiagnosis to occur. In addition, in genetics and genomics overdiagnosis might stem from the particular complexities of the conditions and their tests in the following ways:

- Symptom-free people with an associated variant are defined as 'having the disease'. The patient becomes 'labelled' by possession of a particular genetic variant even though not clinically unwell.
- There is an uncertainty in most genetic conditions about whether and when the person who possesses the disease-associated variant will become ill (penetrance) and the clinical features of the condition that will become manifest (expressivity).
- Disease is defined in the form of cellular or molecular processes rather than clinical findings, and this then leads to preventive 'treatments' even though the patient is not ill.
- A variant in the relevant gene is found that is associated with milder level of dysfunction and it is not clear how these patients should be treated.
- A variant is found and reported back to the patient even though its relationship to the disease in question is uncertain (variants of uncertain significance). This 'interpretation gap' arises from the sheer volume of variants that need be tested, the rarity of some inherited diseases in the population and the lack of evidence on the association of variants with rare disease in individuals without a family history.

6.3 Mammography

A mammogram is an image of breast tissue generated using low-dose x-rays to visualise the tissue for the detection of breast cancer. Digital mammography produces a digital rather than film mammogram, and allows for manipulation of the image for easier examination by the clinician. Tomosynthesis uses the same low dose x-rays alongside computer-based reconstructions to generate a 3D representation of the breast tissue.

Mammography is a not a particularly new technology, but breast cancer screening has seen changes that indicate something of its future trajectory – wide-scale digitisation in several countries, and the development of 3D imaging (tomosynthesis) and AI for screening. These increase test sensitivity and the ability to look for early stage cancers, and requires accompanying reevaluation of diagnostic thresholds.

The initiation of national breast cancer screening programmes using mammography was based on evidence from large-scale randomised trials conducted in the mid to late twentieth century, which showed that screening helped to prevent deaths from breast cancer. After implementation, the incidence of breast cancer in the populations eligible for screening rose between 2 and 10% in each country[6] in which it was established[6], without a subsequent comparable drop in the incidence of advanced cancer or cancer deaths. This indicates that, whilst more early stage cancers were being detected, these were not necessarily cancers that would have progressed to an advanced stage, become symptomatic or have caused early death.

6.4 Stratified screening continued

Whereas there is much concern for the potential of overdiagnosis through new technologies offered as early disease screening, it has been shown that some technologies, such as genomics, could reduce the overdiagnosis that occurs through standard mammographic screening[7]. The principle is to use risk assessment tools to assess individual risk and to offer a tailored mammography according to level of risk.

In breast cancer, risk assessments may include upwards of 100 genetic variants, combined with data on additional risk factors, such as age, sex, weight, reproductive history, alcohol intake etc. Those with a low risk score may be advised on a later-in-life start to screening, or no screening, whereas moderate risk women may be offered any of: an earlier screening start or finish for screening; more frequent screening; more sensitive screening modality (e.g. CT); or personalised lifestyle programmes. High risk women may, in addition be offered chemoprevention or surgery. Much of the purpose behind such 'stratified screening' is to reduce the risk of harm for overdiagnosis: women with less ability to benefit, because they are less likely to be diagnosed with cancer, are not exposed to the potential harm of screening.

6.5 Liquid biopsy for cancer

Liquid biopsy for cancer is a diagnostic technique that involves taking a small sample of blood to look for circulating tumour DNA (ctDNA), which may indicate the presence of cancer. Liquid biopsy is a promising technique that provides a significantly less invasive alternative to conventional tissue biopsy, and may, in the future, offer a more holistic view of cancer status.

Liquid biopsy for cancer detection has been fairly well studied in advanced stage cancers, and there is some enthusiasm for its use as a highly sensitive and convenient new diagnostic. However, the application of liquid biopsy to early stage cancer in a bid to improve lead time is less well evaluated. There are still several questions and some uncertainty around the biological mechanisms underlying liquid biopsy for cancer: what the mechanisms of ctDNA release actually are, how frequently each occurs, and how well maintained these markers are in the blood or urine.

Owing to its perceived sensitivity and ease of use, in the future we may well see clinical liquid biopsy techniques expanded to include tests for circulating RNA or epigenetic markers, some of which are currently being developed. But the sensitivity of these tests alongside 'mission creep' into detecting earlier stage tumours suggests that focus should also be placed on developing understanding of the significance of these markers in terms of natural history, developing equally nuanced responses to any diagnosis (e.g. beyond watchful waiting vs radiotherapy), and trying to understand what the likely level of overdiagnosis might be.

6.6 Molecular diagnostics

Molecular diagnostics encompass a range of techniques that utilise genetic information or the products thereof (e.g. DNA, RNA, protein) - often referred to as 'biomarkers' - to provide diagnostic information or risk scores for one or more conditions. The presence, absence or altered expression of these markers could provide information about likely risk of disease development.

Biomarkers in mental health and neurodegenerative conditions - Decades of research have implicated hundreds or biomarkers in mental health disorders including depression and dementia. Putative markers in depression are associated with many different functions, with some of the more promising including markers associated with inflammatory responses and metabolic processes. Some of the major concerns arising from the use of molecular diagnostics for conditions such as depression and dementia are the heterogeneity of these conditions and the populations to which they might be applied, and ill-defined 'normal' molecular states, and the difficulty of applying precise definitions in the context of daily clinical practice.

Mental health disorders are frequently subject to debate around diagnosis; especially concerning the benefits/disadvantages of balancing the use of guidelines, clinical context and subjectivity to obtain a conclusion most useful for the patient. Overdiagnosis of mental health disorders is discussed widely with extremes of opinion. It is suggested that (amongst other conditions) depression is being overdiagnosed in the US due to several factors, ranging from 'well-intentioned enthusiasm' as suggested by Treadwell and McCartney (2016), to the need to have a medical diagnosis in order for individuals to be able to claim on their medical insurance and over-activity attributed to financial incentives to sell solutions to patients. The consequences of diagnosing an individual with a progressive disease such as dementia, for which there are limited preventative options and no curative therapies, should not be underestimated[8].

Gene expression panels in cancer - RNA-based gene expression panels examine levels of RNA (normally messenger RNA) and could be used to help clinicians provide prognostic information following a diagnosis and help them and their patients make decisions about whether to undergo treatment and therapy selection.

OncotypeDX is one example of such a test. It is a panel used for assessing cancer recurrence risk and potentially informing treatment decisions in patients with specific types of cancer. A recent trial sought to assess the suitability of the OncotypeDX test for determining the likely benefit of chemotherapy in these patients. Results from the trial suggest that many patients with early stage cancers who might have been offered chemotherapy using conventional prognosis tools would be unlikely to benefit from the treatment, and would (with agreement) be excluded from chemotherapy if the new test were implemented.

There are still questions about the true benefits of the test over current prognosis tools, however as these technologies advance, they could help patients avoid undergoing treatment routes with harmful side effects when there is likely to be little benefit to them.

7 Detection and measurement of overdiagnosis

The need to estimate the likelihood and scale of overdiagnosis has mainly been addressed in the context of new screening programmes, and most of the discussions have arisen from studies in the clinical area of cancer. Being population programmes, rates of likely overdiagnosis are described at a population level. This information is useful to policy makers and healthcare providers, but may also help individuals decide whether to take up a screening test.

Recent work undertaken for the US Preventive Services Task Force provided guidance on defining, estimating and communicating overdiagnosis in cancer screening. In it, the authors advocate defining overdiagnosis as 'the detection of a (histologically confirmed) cancer through screening that would not otherwise have been diagnosed in a person's lifetime had the screening not been done'. They describe two main methods for estimating levels of overdiagnosis, but note that these depend on the precise metrics, study types and methods; and are affected by the disease prevalence in the population, risk of death from other causes and the balance of benefits and harms of detection and treatment for precursor lesions.

The main method for estimation is the **excess incidence approach** in which the number of overdiagnosed cancer cases equals the difference between the number of cases observed in a screened population and the number expected without screening. Data may be derived through a number of different methods, which each introduce strengths, weakness and potential biases. The methods include:

- **Follow up of randomised controlled trials** that compare clinical diagnostic and screening approaches. This is thought to be the best method for minimising biases. It provides a direct answer, but requires substantial time and resources and may have limited validity when applied to other situations.
- **Ecological and cohort studies** which make use of data routinely collected by screening programmes or national registries to compare populations invited for screening versus those not invited. These provide a direct answer and a 'real-world' view of overdiagnosis and can be used to monitor potential overdiagnosis over time. However, there is potential for confounding related to diagnosis, treatment and health status between populations and there is a requirement for investment in population registries for full and accurate information about potential confounders.
- **Modeling** can take account of areas of uncertainty by providing assumptions and can evaluate multiple screening solutions. Depending upon the extent of validation, it can be less time-consuming, but results depend entirely on these assumptions.
- **Pathological and imaging studies** that use pathological findings (such as the size or stage of tumours) informed by an understanding of disease progression and clinical manifestations. Although one of the simplest methods, this involves many assumptions and offers only a rough estimate. However, these assumptions may be refined over time as more biomarkers help to distinguish subsets of disease with different progression. In general the method can be useful for monitoring over time, but requires confidence that all diagnosed cases are ascertained.

The lead-time approach is based on an estimate of the time by which screening advances the diagnosis, i.e. the time between detection by screening and the time the cancer would have been detected clinically. For progressive cancers overdiagnosis has occurred if death from other causes occurs within this time.

General approaches to estimation of overdiagnosis have methodological limitations and sources of bias that are discussed in detail in the literature[9]. For the purposes of this document, we can conclude that methods of estimation are inexact, time-consuming and resource intensive. Where undertaken this has largely been in the context of large population screening programmes. This means that the new tests offered within a personalised prevention context - to individuals, outside population programmes and often outwith the health system (for example by commercial providers or arising from citizen generated data) - are completely untested in terms of their propensity for overdiagnosis.

8 Ways to counteract or mitigate overdiagnosis - discussions from the literature

The substantial movements on overdiagnosis in the UK and around the world make many suggestions and proposals for how the problem could be addressed[10]. Mostly these are derived as a counterbalance to the many forces driving it. Whilst many are aimed at advances in technology, including how and under what circumstances they are implemented, others look at the pervading culture in society, incentives and priorities within health systems, the role of professionals, the commercial sector, and the knowledge, attitudes and influence of patients and the public. The following range of solutions is rehearsed in the literature:

Culture - it is important to challenge the notion both with the public and the wide range of professionals involved in prevention, from policy, to research, technology development and implementation, that prevention is always better than cure, new is better than old and that more is better. Key to this building more skepticism about the real benefits and likely harms arising from attempts to implement personalised prevention through risk assessment and earlier diagnosis.

Health systems - should be reformed so that practitioners are able to work in a more person-centred way, finding solutions that are right for the individual rather than being driven by incentives that may prioritise quantity over quality. New measures of overdiagnosis should be developed and introduced. It is also important to keep a close watch on disease terminology and definitions; overdiagnosis should be one of the key considerations when these change.

Professions - need to tackle and reform the factors that underlie defensive medicine and help professionals and patients to overcome the fear of uncertainty, for example by incorporating 'watchful waiting' into management options . This will require critical thinking, education and training as part of professional development.

For industry and the commercial sector - commentators recommend that the strong role that these play in developing and marketing new technologies should be kept in check by regulation of the advertising and promotion of new tests to public and professionals and by paying attention to conflicts of interest such as industry funding of patient and advocacy groups, which may seek to generate demand for diagnosis.

Patients and public - should also be informed and educated about the possible harms of testing, particularly of screening, as well as the benefits. Opportunities should be sought within the consultation to discuss why tests may not be needed and may be unhelpful. This should be done in such a way that the patient feels that their own circumstances and needs have been well understood and they have been able to take an informed autonomous decision.

Technology - commentators advocate caution with respect to implementation – that we should not be 'overly enthusiastic about new technology and adopt perspectives that grasp our ambivalence to it, both controlling it and feeling controlled'. Although being optimistic and wishing to stimulate innovation, it is suggested that 'only those technologies that have shown real benefit should be put into routine practice' [11]. Diagnostics should be assessed as carefully as new treatments, patients should be informed about risks and benefits and should be involved in their development, assessment, implementation and use.

9 Conclusions and questions

Despite aiming to benefit people found to be at risk of disease or with early disease, it seems possible, even probable, that a major push towards personalised prevention will increase the harms caused by anxiety, investigation and unnecessary treatments in many individuals.

In principle new technologies such as genomics or various other biomarkers increase the information that should underpin personalised prevention, but, at present, understanding often lags behind the technology. We are limited in our ability to interpret the newly identified sub-categories of disease or risk strata in a way that leads to evidence based management and optimisation of the benefit/harm ratio. The rapid advance of technology and the relatively cumbersome nature of gathering evidence (for example through randomised controlled trials), particularly where this relates to the balance between benefits and harms, also beg the question whether it will be possible to generate the necessary evidence and whether current methods must be adapted.

There are many opportunities to identify people at risk or with early disease within the public health and healthcare systems, for example, through population screening programmes, in the course of a clinical encounter or in response to individual concern. We have identified ways in which new technologies might increase these opportunities. Given the pervading background culture (described in section 5) there will be many drivers for this that may be rather difficult to counteract. Very rarely are the new technologies used to identify individuals in whom interventions ought not to be offered (as was the case described for mammography in section 6.4).

Going forward, we would like to suggest ways in which technologies may be researched, developed and introduced with the aim of reducing harms as far as possible. This may require consideration of:

- Research prioritisation, design and translation – important to include a wide set of perspectives
- The use of screening tests and other methods to identify people at risk
- Health system and clinical interactions
- Disease definitions
- Professional infrastructure (e.g. disease definitions)
- Interactions with the commercial sector
- Consideration of citizen generated data

Questions for the workshop

1. Is it inevitable that some level of harm will be caused by the introduction of new technologies aimed at personalised prevention?
2. How can these harms be reduced and benefits enhanced?

References and sources

1 UK Government policy paper (2018): Prevention is better than cure: our vision to help you live well or longer; https://www.gov.uk/government/publications/prevention-is-better-than-cure-our-vision-to-help-you-live-well-for-longer

2 Frege, G. 'On Sense and Reference' in Translations from the Philosophical Writings of Gottlob Frege, edited by P Geach and M Black. Oxford: Basil Blackwell, 1960

3 Carter SM, Rogers W, Heath I, et al. The challenge of overdiagnosis begins with its definition BMJ 2015; 350 :h869

4 Feinberg, J. The Moral Limits of the Criminal Law Volume 1: Harm to Others, Oxford University Press, 1987

5 Pathirana T, Clark J, Moynihan R. Mapping the drivers of overdiagnosis to potential solutions BMJ 2017; 358 :j3879

6 Barratt A. Overdiagnosis in mammography screening: a 45 year journey from shadowy idea to acknowledged reality BMJ 2015; 350 :h867

7 Pashayan N, Morris S, Gilbert FJ et al Cost-effectiveness and Benefit-to-Harm Ratio of Risk-Stratified Screening for Breast CancerA Life-Table Model. JAMA Oncol. 2018;4(11):1504–1510. doi:10.1001/jamaoncol.2018.1901

8 Brunet M. Targets for dementia diagnoses will lead to overdiagnosis; BMJ 2014; 348:g2224

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

10 Kale MS, Korenstein D. Overdiagnosis in primary care: framing the problem and finding solutions BMJ 2018; 362 :k2820

11 Hofmann BM. Too much technology BMJ 2015; 350 :h705

Additional sources

Brodersen J, Kramer BS, Macdonald H, Schwartz et al. Focusing on overdiagnosis as a driver of too much medicine BMJ 2018; 362 :k3494

Carter SM, Degeling C, Doust J, et al. A definition and ethical evaluation of overdiagnosis Journal of Medical Ethics 2016;42:705-714.

Davies L, Petitti DB, Martin L. Estimating, and Communicating Overdiagnosis in Cancer Screening Ann Intern Med. 2018;169(1):36-43. DOI: 10.7326/M18-0694

Mayor S. New way to define disease is needed to reduce overdiagnosis, experts agree BMJ 2016; 354 :i5082

Moynihan R, Henry D, Moons KGM (2014) Using Evidence to Combat Overdiagnosis and Overtreatment: Evaluating Treatments, Tests, and Disease Definitions in the Time of Too Much. PLoS Med 11(7): e1001655. https://doi.org/10.1371/journal.pmed.1001655

Treadwell J, McCartney M. Overdiagnosis and overtreatment: generalists — it's time for a grassroots revolution. Br J Gen Pract 2016; 66 (644): 116-117. DOI: 10.3399/bjgp16X683881

PHG Foundation
2 Worts Causeway
Cambridge
CB1 8RN
+44 (0) 1223 761900

www.phgfoundation.org

Contact: intelligence@phgfoundation.org

PHG Foundation is a health policy think tank with a special focus on how genomics and other emerging health technologies can provide more effective, personalised healthcare