

## ANALYSIS

## How precision medicine and screening with big data could increase overdiagnosis

Precision medicine based on big data promises to revolutionise disease prevention but increases the challenge of determining which abnormalities will be clinically important, argue **Henrik Vogt and colleagues**

Henrik Vogt *postdoctoral fellow*<sup>1 2 3</sup>, Sara Green *assistant professor*<sup>4 5</sup>, Claus Thorn Ekstrøm *professor*<sup>6</sup>, John Brodersen *professor*<sup>7 8</sup>

<sup>1</sup>Centre for Medical Ethics, University of Oslo, Oslo, Norway; <sup>2</sup>Hybrid Technology Hub, University of Oslo, Oslo, Norway; <sup>3</sup>General Practice Research Unit, Norwegian University of Science and Technology, Trondheim, Norway; <sup>4</sup>Department of Science Education, University of Copenhagen, Copenhagen, Denmark; <sup>5</sup>Centre for Medical Science and Technology Studies, University of Copenhagen, Denmark; <sup>6</sup>Biostatistics, Department of Public Health, University of Copenhagen, Denmark; <sup>7</sup>Centre of Research and Education in General Practice, University of Copenhagen, Denmark; <sup>8</sup>Primary Health Care Research Unit, Region Zealand, Denmark

Since the Human Genome Project in the 1990s, there has been discussion of how precision medicine (or personalised medicine) might prevent morbidity and mortality by diagnosing disease early or finding risk factors in apparently healthy people.<sup>1</sup> At first, the idea was mostly tied to genome analysis. However, during the past two decades many other big data and machine learning technologies have emerged to measure and analyse other factors.<sup>2-7</sup>

“Big data” can mean many things.<sup>8</sup> We use it to describe newer, data intensive technologies that might enable precision medicine.<sup>4</sup> These include “omic” technologies such as genomics (DNA), proteomics (protein), transcriptomics (RNA), and metabolomics (metabolites).<sup>7</sup> It also includes data from imaging, electronic health records, social media, new biosensors,<sup>30</sup> and self tracking technologies such as the Apple smartwatch that can detect arrhythmias.<sup>9</sup> In future such sensors may be able to pick up signals from blood, sweat, or environmental exposures.<sup>3</sup>

Proponents of screening based on these new technologies say they enable unprecedented monitoring of the human body.<sup>2-7</sup> However, we believe that their plans, which are still largely theoretical, come with a high risk of overmedicalisation<sup>10-13</sup> and overdiagnosis.<sup>14 15</sup>

### What is different about big data screening?

Current screening typically involves measuring one variable or a few variables, either once or at long intervals. Big data screening differs from traditional screening because it can measure many different types of variable (from the molecular to the social). It can include many variables at the same time,

and measurement can be repetitive or continual, showing bodily changes over time.

This may create a highly detailed description of each person's genotype and phenotype (biological traits).<sup>3</sup> Large datasets from various sources can be used to screen individuals for many diseases and risk factors at once (box 1).<sup>5-7 9 16-18</sup>

#### Box 1: Three early big data screening studies

##### Pioneer 100 Wellness Project (P100)<sup>5</sup>

This project collected data from 108 apparently healthy people aged 21-89 over nine months, including whole genome sequences, results of 218 clinical tests, and assays of up to 643 metabolites, 262 proteins, and 4616 microbes, as well as sleep and activity measurements from wearable devices.

The study found vitamin D deficiency in 95 people, high mercury levels in 81, diabetes or prediabetes (HbA1c  $\geq 5.7\%$ ) in 52, cardiovascular risk (dyslipidaemia) in 73, haemochromatis in 1 as well as several abnormalities associated with inflammation and nutrition.<sup>5</sup>

##### Precision medicine screening study<sup>6</sup>

This study used clinical tests, whole genome sequencing, whole body magnetic resonance imaging, dual energy x ray absorptiometry, metabolomics, a new test for prediabetes (Quantose IR), echocardiography, electrocardiography, and cardiac rhythm monitoring in 209 people.

It identified diabetes or risk of diabetes in 118 people, atherosclerotic disease or risk in 29, non-alcoholic fatty liver disease in 28, early stage neoplasia in 4, atrial fibrillation in 2, other medically significant arrhythmias in 2, third degree heart block in 1, primary biliary cholangitis in 1 and xanthinuria in 1.<sup>6</sup>

##### Longitudinal big data approach for precision health<sup>7</sup>

This study had a particular focus on diabetes and followed 109 people for a median of 2.8 years. It used quarterly clinical measures; analysis of the genome, immunome, transcriptome, proteome, metabolome, and microbiome; and monitoring with wearable devices. Among its findings were diabetes or prediabetes in 59 people, cardiovascular disease in 9, cardiac arrhythmias in 2, cardiopathy (genetic) in 1, obstructive sleep apnoea in 1, Lyme disease in 1, lymphoma in 1, myeloma in 1, monoclonal gammopathy of undetermined significance in 1, and 15 genetic mutations thought to require action.<sup>7</sup>

## What does big data screening mean for overdiagnosis?

Overdiagnosis is common in any medical screening and generates waste and harm.<sup>14 15</sup> It occurs when we diagnose asymptomatic abnormalities that would not cause a problem. This might be because they disappear spontaneously, do not progress, or progress so slowly that people die from other causes.<sup>14 15</sup>

The main problem for big data screening is that monitoring many features of the body with highly sensitive technologies is bound to detect many abnormalities but without the ability to tell which, if any, will become clinically manifest. As a result, more people may be labelled with more harmless conditions.

The risk of overdiagnosis in precision medicine screening is illustrated by the three studies described in [box 1](#). The Pioneer 100 study measured numerous variables repeatedly over nine months in 108 people who had no known health problems. All of the participants had multiple abnormalities detected ([box 1](#)).<sup>5</sup> Similarly, a precision medicine screening study diagnosed abnormalities in 164 of 209 participants,<sup>6</sup> and a study presenting “a longitudinal big data approach for precision health” discovered “more than 67 clinically actionable health conditions in 109 people and identified multiple molecular pathways associated with metabolic, cardiovascular, and oncological pathophysiology.”<sup>7</sup>

Such sweeping findings likely include substantial overdiagnosis<sup>9-12</sup> for reasons we set out below.

*Screening multiple factors, multiple times for multiple problems*—Big data can observe “everything” that goes on in a human body over time.<sup>2-7</sup> However, if “everything” about the body is registered, then everything that medicine defines as abnormal will also be detected. Only some of these abnormalities will become symptomatic. The rest will represent overdiagnosis. The total risk of an individual becoming overdiagnosed increases with the number of abnormalities screened for.

*A culture of better, more and earlier*—The vision of precision medicine screening is driven by technological innovation and associated with known cultural drivers of overdiagnosis, such as the beliefs in healthcare that “more is better,” that “new is better,” and that “prevention is better than cure.”<sup>2-7 19</sup> Early projects and visionary publications are also focused on diagnosing abnormalities early,<sup>3-5 7</sup> which drives overdiagnosis.<sup>14 15 19</sup>

*Screening in lower risk groups*—Cheap genome testing and monitoring technologies enable more widespread screening. Thus more abnormalities in larger populations with low risk may be picked up, increasing overdiagnosis.<sup>14 15 19</sup>

*Expanded disease definitions*—Overdiagnosis increases with wider definitions of abnormality (low diagnostic thresholds) as more people are defined with more problems that demand medical attention but that may not develop symptoms. Pioneering studies screened for low risk conditions such as prediabetes.<sup>5-7</sup>

*More sensitive technology*—Unlike false positive results, which depend on the specificity, overdiagnosis is increased by the sensitivity of screening tests. New screening technologies such as fine grained molecular tests may be more likely than older tools to pick up more and milder abnormalities. The natural course of previously undetected, mild cases of atrial fibrillation picked up by a smart watch, for example, have not been studied properly and may not carry the same risk as previously detected cases.<sup>20</sup> But our bias may be to treat these new cases like we

treat previously researched conditions in the high risk end of the disease spectrum (spectrum bias).<sup>21</sup>

## Challenge for precision medicine

If everything “abnormal” can be detected, precision medicine screening faces a huge challenge. It must predict which of the abnormalities will cause symptoms or death. We are currently not able to predict which abnormalities will produce clinical symptoms, but longitudinal research, better algorithms, machine learning, and improved causal understanding may improve future understanding. This understanding can inform definitions of abnormality and guide decisions about what should be screened for.

Early empirical results suggest that algorithms combining multiple variables may improve risk predictions.<sup>16-18</sup> For example, one study showed that polygenic risk scores could identify a substantially larger percentage of a population at greater than threefold risk for five common diseases than previous DNA analysis.<sup>18</sup> However, when polygenic risk scores were applied to five isolated diagnoses, 20% of the population was found to be at “high risk” for at least one of them.<sup>18</sup> If such algorithms are applied to many diagnoses, everybody is likely to be labelled at high risk of something.<sup>14</sup>

Such risk scores could also be used to redefine low risk conditions as something harmless, which may reduce overdiagnosis. Our view of conditions that have previously been regarded as high risk may also need to be readjusted as something that is relatively harmless and not to be diagnosed.

There are some specific challenges. A biological challenge is that the complex systems of the body make it difficult to predict the course of disease in individuals.<sup>22-24</sup> A related statistical problem is called the bias-variance dilemma ([fig 1](#)). When creating predictive algorithms, it is tempting to include more variables to increase the accuracy and the ability of the model to reflect the real life complexity. However, because it is hard to reliably estimate many variables at once, the greater accuracy (reduced bias) gained becomes outweighed by an increase in the variance or noise (loss of precision).<sup>25</sup>

## Harnessing the benefits of precision medicine

Big data provides a unique opportunity for increasing our knowledge of disease development. But the creation of beneficial precision medicine will require patience, tolerance, and wisdom. The problem of overdiagnosis in big data screening and precision medicine requires more attention from professionals, funding agencies, ethical committees, and policy makers.

We need to rethink risk. Our intolerance of risk and disease increases the risk of overdiagnosis.<sup>14 15 19</sup> Big data precision medicine may seem to offer the prospect of eliminating uncertainty, but the benefits may be overshadowed by waste and harm resulting from identifying conditions that do not cause real problems. We have to accept that there will always be some degree of risk, morbidity, and mortality. The prospect of being able to register “everything” that goes on in the body will clarify just how much medicine defines as abnormal, and highlight the need to discuss what should be regarded as important. We need large prognostic studies with decade long follow-up that investigate the natural course of different abnormalities and quantify overdiagnosis.

Shared decision making is often seen as the way to reduce overdiagnosis as it includes a discussion of what is important for individual patients as well as the potential benefits and risks of screening. However, technological advances may mean that empowered consumers will bypass the medical establishment and screen themselves without full understanding of the consequences, potentially driving the new preventive medicine.<sup>26</sup> Companies such as Facebook, Apple, Amazon, and Google now both use monitoring data and include consumers in studies to perform big data health research.<sup>27</sup> The medical profession and regulatory bodies may lose their ability to regulate screening, and it is unclear if citizen, consumer, and patient involvement will increase or decrease overdiagnosis. In a context where people are convinced that extensive self monitoring is useful, leaving medical decisions to individuals may create as much harm as benefit.<sup>28</sup>

One pragmatic approach to minimise the risks of big data screening is for it to be guided by traditional screening principles, such as those developed by Wilson and Jungner for the World Health Organization.<sup>29</sup> Targeting only serious conditions for which there are effective treatments, for example, would limit the number of abnormalities that are screened for and thus reduce overdiagnosis. Everyone working to advance the use of precision medicine for prevention must be alert to the possible harms of overdiagnosis.

#### Key messages

Precision medicine based on big data promises to improve disease prevention but entails a massive, new form of screening

The wide scope of big data screening also risks increased detection of abnormalities that will never be clinically relevant

We need a clearer understanding of the natural course of multiple markers and the value of repeated measurement of markers

Overdiagnosis should be a concern for anyone working to advance preventive precision medicine

Competing interests: We have read and understood BMJ policy on declaration of interests and have no relevant interests to declare.

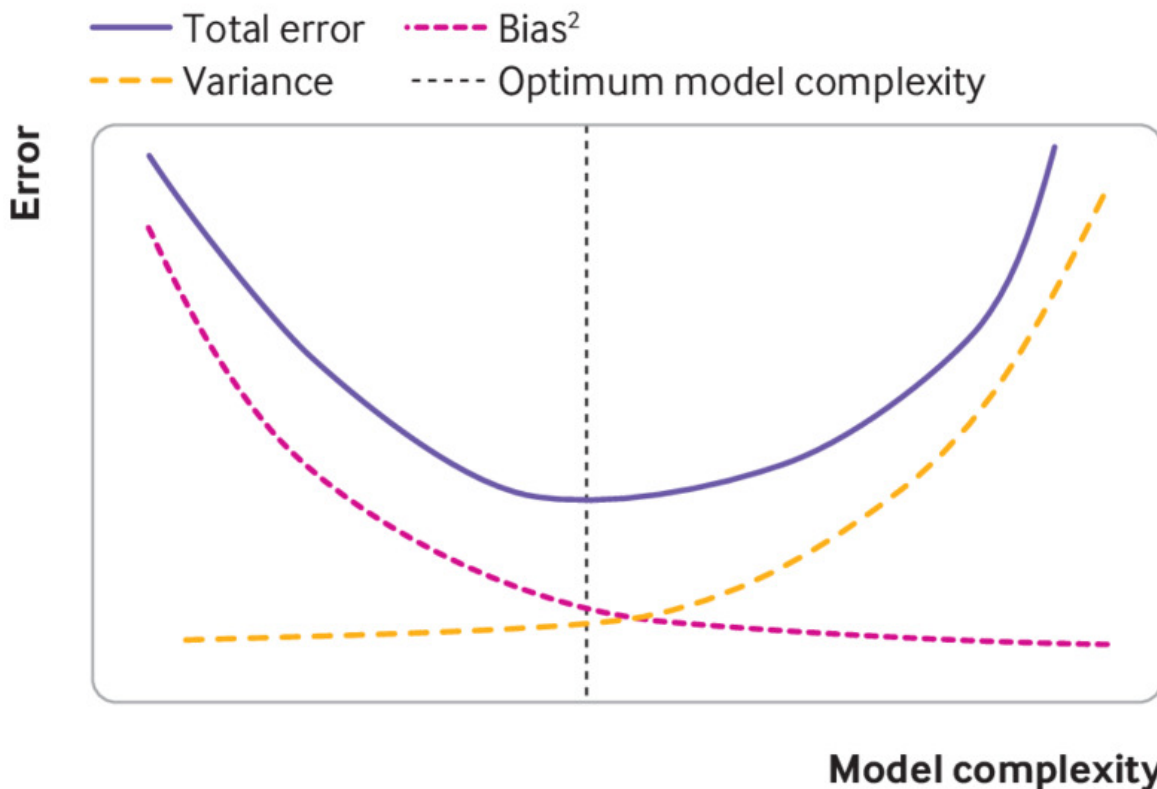
Provenance and peer review: Commissioned; externally peer reviewed.

- Shendure J, Findlay GM, Snyder MW. Genomic medicine—progress, pitfalls, and promise. *Cell* 2019;177:45-57. 10.1016/j.cell.2019.02.003 30901547
- National Institutes of Health. All of Us research program, National Institutes of Health Core Protocol v1.7. 2018 Mar 28.
- Torkamani A, Andersen KG, Steinhubl SR, Topol EJ. High-definition medicine. *Cell* 2017;170:828-43. 10.1016/j.cell.2017.08.007 28841416

- Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. *BMC Med Genomics* 2015;8:33. 10.1186/s12920-015-0108-y 26112054
- Price ND, Magis AT, Earls JC, et al. A wellness study of 108 individuals using personal, dense, dynamic data clouds. *Nat Biotechnol* 2017;35:747-56. 10.1038/nbt.3870 28714965
- Perkins BA, Caskey CT, Brar P, et al. Precision medicine screening using whole-genome sequencing and advanced imaging to identify disease risk in adults. *Proc Natl Acad Sci U S A* 2018;115:3686-91. 10.1073/pnas.1706096114 29555771
- Schüssler-Florenza Rose SM, Contrepois K, Moneghetti KJ, et al. A longitudinal big data approach for precision health. *Nat Med* 2019;25:792-804. 10.1038/s41591-019-0414-6 31068711
- Few S. *Big data, big dupe: a little book about a big bunch of nonsense*. Analytics Press, 2018: 96.
- Koshy AN, Sajeev JK, Nerlekar N, et al. Smart watches for heart rate assessment in atrial arrhythmias. *Int J Cardiol* 2018;266:124-7. 10.1016/j.ijcard.2018.02.073 29887428
- Diamandis EP. The hundred person wellness project and Google's baseline study: medical revolution or unnecessary and potentially harmful over-testing? *BMC Med* 2015;13:5. 10.1186/s12916-014-0239-6 25575898
- Hofmann B, Welch HG. New diagnostic tests: more harm than good. *BMJ* 2017;358:j3314. 10.1136/bmj.j3314 28720607
- Vogt H, Hofmann B, Getz L. The new holism: P4 systems medicine and the medicalization of health and life itself. *Med Health Care Philos* 2016;19:307-23. 10.1007/s11019-016-9683-8 26821201
- Vogt H, Green S, Brodersen J. Precision medicine in the clouds. *Nat Biotechnol* 2018;36:678-80. 10.1038/nbt.4210 30080826
- Welch HG, Schwartz L, Woloshin S. *Overdiagnosed: making people sick in the pursuit of health*. Beacon Press, 2011.
- Brodersen J, Schwartz LM, Heneghan C, O'Sullivan JW, Aronson JK, Woloshin S. Overdiagnosis: what it is and what it isn't. *BMJ Evid Based Med* 2018;23:1-3. 10.1136/ebmed-2017-110886 29367314
- Mega JL, Sirtziel NO, Smith JG, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet* 2015;385:2264-71. 10.1016/S0140-6736(14)61730-X 25748612
- Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One* 2017;12:e0174944. 10.1371/journal.pone.0174944 28376093
- Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50:1219-24. 10.1038/s41588-018-0183-z 30104762
- Pathirana T, Clark J, Moynihan R. Mapping the drivers of overdiagnosis to potential solutions. *BMJ* 2017;358:j3879. 10.1136/bmj.j3879 28814436
- Waldo AL, Camm AJ. Atrial fibrillation: atrial high-rate events (AHRES): look and you will find—then what? *Circulation* 2017;136:1795-7. 10.1161/CIRCULATIONAHA.117.030705 28851730
- Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;299:926-30. 10.1056/NEJM197810262991705 692598
- Firth WJ. Chaos—predicting the unpredictable. *BMJ* 1991;303:1565-8. 10.1136/bmj.303.6817.1565 1773173
- Nebert DW, Zhang G. Personalized medicine: temper expectations. *Science* 2012;337:910, author reply 910-1. 10.1126/science.337.6097.910-a 22923560
- Smith GD. Epidemiology, epigenetics and the 'Gloomy Prospect': embracing randomness in population health research and practice. *Int J Epidemiol* 2011;40:537-62. 10.1093/ije/dyr117 21807641
- Fortmann-Roe S. Understanding the bias-variance tradeoff. 2012. <http://scott.fortmann-roe.com/docs/BiasVariance.html>
- Topol E. *The patient will see you now*. Basic Books, 2015.
- Sharon T. The Googleization of health research: from disruptive innovation to disruptive ethics. *Per Med* 2016;13:563-74. 10.2217/pme-2016-0057 29754546
- Johansson M, Jørgensen KJ, Getz L, Moynihan R. "Informed choice" in a time of too much medicine—no panacea for ethical difficulties. *BMJ* 2016;353:i2230. 10.1136/bmj.i2230 27160274
- Wilson JMG, Jungner G. *Principles and practice of screening for disease*. WHO, 1968.
- Xu S, Jayaraman A, Rogers JA. Skin sensors are the future of health care. *Nature* 2019;571:319-21. 10.1038/d41586-019-02143-0 31316200

Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to <http://group.bmj.com/group/rights-licensing/permissions>

### Figure



**Fig 1** The bias-variance dilemma—as more variables are included in algorithms (increasing model complexity), the accuracy increases, but beyond a certain threshold the precision decreases (reduced bias, increasing variance). This results in increasing total error and poorer overall predictive values (adapted from Scott Fortmann-Roe)<sup>25</sup>

BMJ: first published as 10.1136/bmj.l5270 on 13 September 2019. Downloaded from <http://www.bmj.com/> on 14 September 2019 at Helsebiblioteket gir deg tilgang til BMJ. Protected by copyright.