Online risk numbers – helpful, meaningless or simply wrong? – Reflections on online risk calculators

John-Arne Skolbekken

Department of Public Health and Nursing

Norwegian University of Science and Technology

7491 Trondheim

Norway

Phone: + 47 900 12 203

Email: john-arne.skolbekken@ntnu.no

Abstract

Among the instruments offered to citizens via digital media are risk calculators, aiming at identifying individuals at high risk of various diseases. These calculators present us with both epistemological and socioethical challenges. Tracking the history of individual risk models, this paper provide an analysis looking into their content, construction, use and functions. Epistemologically the notion of risk factor epidemiology frames an approach to public health that goes through the identification of high-risk individuals, providing a way of making public health doable without involving social, cultural and economical factors in the risk assessments. Instead, ethnicity is included in many calculators, serving as boundary objects that enables epidemiologist to avoid addressing its inherent epistemological challenges. Through this notion of individual risk, a discourse is created that provides us with the narrative of the empowered vulnerable global citizen, which is given room to look after her/his risky self, whilst ignoring the structural and political factors influencing it. In doing so, flawed calculator construction provides ample risk of getting the wrong number.

Key words: risk calculators, internet, epidemiology, public health, epistemology, ethics

Introduction

"Know your risk" is a frequently communicated public health message in many places around the world, serving as a prime example of "The New Public Health" (Petersen and Lupton 1996). This is a message with mixed connotations, as it simultaneously reminds us of our empowered position to take action to protect and promote our health, whilst also reminding us of our vulnerability as mortals. More importantly, this situation is presented as one where a strong focus on individual responsibility for health trumps calls for collective action, reflecting a neoliberal understanding of health (Dubriwny 2013). Within this frame of thinking a number of risk calculators have been developed to enable individuals to fulfil their obligations and achieve possible benefits. Put simply, these instruments estimate what is represented as an individual's risk of future disease and dying based on information about the individual's past and present. The aim of this paper is to scrutinize such instruments and the ideology behind them, through an analysis covering a number of calculators that offer individuals insights into their personal risk for coronary heart disease (CHD), type 2 diabetes (T2D), breast cancer and osteoporosis.

The risk scores included in this paper are examples of biomedicalization (Clarke et al 2010), involving an elaborated risk surveillance at the individual level, based on applications of information technology, with the aim of preventing disease, and a potential for changing our identities. Representing diagnosis of asymptomatic conditions, this knowledge may be included in a person's narrative identity in a harmful way, potentially leading to overdiagnosis and overtreatment (Walker and Rogers 2017). As we shall see, efforts to calculate and communicate an individual's health risk is a practice that has developed over the last couple of centuries. The arrival of the internet presents an opportunity to calculate risk for more individuals than ever

before, not only through the health care system, but in principle wherever and whenever people are online, including the opportunity provided by smartphone applications. Much of what is presented here is relevant for these calculators also when not available on the internet, but the magnitude of potential benefits and harms escalate with their increased availability.

Risk scores for cardiovascular disease are part of all electronic patient record systems in general practice in the UK (Noble et al. 2011). To calculate their patients' risk scores have become part of the GP's duties according to recent guidelines. Under the trademark of Your Disease Risk, characterised as "the source on prevention" The Siteman Cancer Center offers risk calculators for a dozen forms of cancer as well as well as five "other key diseases" (https://siteman.wustl.edu/prevention/ydr/). Reviewing the availability of risk models and scores for diabetes, Noble et al. (2011) found no less than 145 such models. Opportunities for risk calculation are thus abundant, but as with many other developments, there is reason to take a step back for some critical reflection as the potential for both benefit and harm has increased accordingly.

Inspiration for this paper comes from Bouk's (2015:15) observation that insurance companies sold "competing epistemological, social and ethical positions." A major aim here is thus to illustrate how the instruments that offer people knowledge about their personal health risks are the outcome of choices made by the professionals behind them, reminding the reader that it could have been different. Over the years the questions "Are we doing the right things?" and "Are we doing the things right?" have proved fruitful for critical reflection, and they have also provided

guidance in the work on this paper. As indicated by the title of the paper many of these reflections circle around the meaning of risk numbers. Acknowledging that they have no inherent meaning, but are given meaning in the course of human reflection and interaction, a recurring question is – what meaning do they really have epistemologically?

To perform the analysis presented here, concepts like blackboxing and boundary objects are borrowed from the science and technology literature. Blackboxing refers to the processes where scientific work is so successful that the work behind it is no longer visible (Latour 1999). This is what happens when something becomes a fact that is no longer questioned. In the medical literature risk scores and calculators are presented as instruments developed without shedding much light on the choices that have been made when doing so. Part of the analysis will thus seek to demonstrate that these scores do not represent a taken for granted reality, but are the outcome of choices no longer visible, as they have been blackboxed.

In the construction of individual risk there are tensions that need to be handled, like that between its universal and local applications. To describe how this can be achieved, Star and Griesemer (1989) introduced the concept of boundary objects, which they define as objects that are "both adaptable to different viewpoints and robust enough to maintain identity across them" (P. 387).

Overall the paper is based on a constructionist position, taking a critical stance on the knowledge taken-for-granted (Burr 2015). Accordingly, the analysis takes as its starting point that the notion of individual risk is one framed by and contributing to a particular public health discourse.

The studied tools for risk calculation are approached through Prior's (2003) different frames for doing document analysis. Accordingly, online calculators are treated as forms of documents although their present formats differ from the paper form habitually connected with documents. In line with the outlined constructionist position the studied documents are recognised as made for a purpose, serving as agents in a social setting, framed by and contributing to a specific medical risk discourse and the construction of primarily high-risk individuals. A major reason for choosing this approach is that it offers insights into the construction, use and functions of documents, in addition to the content analysis preferred by many researchers. By applying all four approaches, an ambition behind this paper is also to hopefully make others aware of a research area worthy of closer inspection by applying one approach on a single instrument at a time.

To give the reader an introduction to the basic characteristics of risk calculators, the analysis starts with a description of the content of selected calculators. This is followed by a presentation of the construction of calculators, focussing in particular on the relation between individuals and populations, including the construction of ethnicity. This part is perhaps the most crucial part of the analysis, as it opens black boxes and questions the scientific soundness of the calculators. Once constructed the instruments have to be used, and their use tells us about whether they are applied as expected by their constructors or whether new and ingenious ways of using them are invented by their users. Finally, the analysis looks at what functions these risk calculators have. As indicated initially, the calculators have been constructed to benefit humans, but there are reason to believe that they may also serve other functions. In presenting the analysis, it is not my

ambition to give a review covering all available instruments, as their abundance makes that a task far beyond the scope of this paper. Rather, the analysis is based on some critical cases that provide illustrations of the epistemological and socioethical issues in question.

A brief history of individual risk

To understand the present situation, a presentation of its historical background is useful. A brief history of the development of the notion of individual risk is therefore provided below, followed by a brief introduction to social epidemiology and risk factor epidemiology, respectively.

Risk calculation in medicine is based on epidemiological and actuarial data, measuring the risk of populations. It has replaced earlier understandings of causality in medicine through the acceptance of a non-deterministic model of disease known as the multifactorial model. The invention of the risk factor, serving as a proxy for causality, paired with the attribution of lifestyle as the outcome of individual choice, are imperative contributions to the development of the notion of individual risk. As illustrated by Rothstein's (2003) historical account, this development involves several innovations that are beyond the present analysis.

In the introduction it was noted that the concept of individual risk can be seen as a product of and an instrument for the goals of a neoliberal health policy. In medical practice, however, the ideological frame may be less noticeable as the notion of individual risk can be seen as serving the practical needs of life insurance companies and preventive medicine, respectively. One origin

of individual risk can be traced to the rise of the life insurance industry in America, as a construction meeting the demand for knowledge about potential individual policy customers (Bouk 2015). In this setting individual risk became a commodity, one that regulated eligibility for and the price of life insurance. It was thus not part of a health policy aiming at risk reduction in an effort to prevent mortality and morbidity, but a financial effort to guard insurance companies against economic loss resulting from selling insurance policies to high risk individuals (Rothstein 2003). This notion of individual risk is not, however, the one most frequently pointed to in descriptions of the origins of risk scores in modern medicine.

Texts depicting the development of modern risk scores point to the Framingham study as fundamental for the development of individual risk models (Holmberg and Parascandola 2010, Aronowitz 2015). The aim of this model is to identify those individuals with the highest risk of coronary heart disease, in order to identify preventive measures. According to (Aronowitz 2015) this approach was not part of the original design of the study, but came as an outcome of a process transforming it from a project based on public health achieved through collective effort to preventive medicine based on an individualised strategy in the setting of clinical practice. On the way to accomplish this, important steps were taken in the form of limiting the study to quantifiable, easily measured, clinically apparent factors. This subsequently paved the way for the construction of the notion of risk factors that could be easily identified by clinical practitioners.

Despite not being the obvious choice when the Framingham study was designed, the individual focus was presented as self-evident in the early days of publications from the study:

> *It is axiomatic in the control of disease that the first step is to identify those at high and those at low risk.*
>
> Kinch, Doyle and Hilleboe (1963:438)

This dichotomy was illustrated through the profiles of Harry Coronary and Norman (Norm) Normal. Despite their differences, the authors concluded that it was still early days for a precise identification of "a specific coronary-prone individual." If their cartoon characters were anything to judge by, Harry and Norm were white, male and middle class, reflecting the Framingham cohort. As with the life insurance risk policies, the original risk calculations were made for these groups. A common feature of the early presentation of individual risk was its primary focus on the risk of white men, both in the Framingham study and in the introduction of life insurance, where women and Afro-Americans were excluded, despite their wish to be able to purchase life insurance (Bouk 2015). As we shall see, the idea of risk scores for cardiovascular diseases has since spread to a number of other diseases, and the idea of individual risk has been carried over to other domains of medicine. In the years to follow the focus on high risk gained considerable attention over low risk, perhaps reflecting the clinical origin of this approach.

A third way into the history of individual risk scores is through the co-construction of drugs and diseases, in particular the development of drugs regulating blood pressure, blood cholesterol and blood sugar, followed by the creation of risk-factors come diseases like hypertension, hypercholesterolemia and T2D (Greene 2007). This development led to what is described as the

converged experience of risk and chronic disease (Aronowitz 2015), an experience that also is traceable to the construction of surveillance medicine wherein we are all defined as potentially sick (Armstrong 1995).

Accompanying this development has been the construction of diagnostic instruments in the form of risk scores, which have contributed to the co-construction of risk and disease, aiming to identify asymptomatic individuals in need of pharmaceutical intervention. In recent decades it has been noted that risk calculators have been developed as a tool accompanying most blockbuster-drugs (Ebeling 2011), as part of an expanding strategy enabling doctors to identify those individuals seen as being in need of chemoprevention provided by pharmaceutical products proven to have risk reducing effects in randomised controlled trials.

Sick individuals and sick populations

Although commonly presented as self-evident these days, the individual lifestyle approach to public health is not without its alternatives. Calculating an individual's risk of getting a serious and often potentially fatal disease is part of a high risk strategy where the aims of public health are presented as achieved through intervention in the lives of those individuals identified as having the most substantial risk.

The major alternative to this strategy is the mass strategy, which typically addresses whole populations, trying to improve public health by moving the population mean, acknowledging that

knowledge about which individuals that will benefit from interventions is not available. In contrast to the high risk approach, this strategy includes modifying the risk of those individuals at low risk, based on what is known as the prevention paradox (Rose 1985), where the benefit for the individual remains unknown whilst being harvested at the population level. These strategies are not necessarily seen as opposites, but as complementary strategies, addressing different parts of the population.

The differences between them is illustrated through many of the articles on the rising threat of T2D globally. Commonly articles advocating the development of a risk calculator take as their starting point an epidemic growth of this condition in recent and coming decades. In the frame of the individual lifestyle theory this epidemic is understood as the sum of individual lifestyle choices, which can be solved by means of the risk calculator and subsequent individual lifestyle change. The knowledge base behind the high risk approach is risk factor epidemiology, which strives to identify individual factors through what has been described as a "black box paradigm" (Susser 1998). This represents a reductionist position for disease causality, which can be labelled as "causality by proxy", where many of the complex social and cultural processes behind disease developments are left unaddressed whilst the focus is put on targeting one risk factor at the time.

An alternative approach to this epidemic would be to look beyond individual choices to find the social, structural and cultural factors behind them, through social epidemiology (Krieger 2011). In the event of the identification of sugar consumption as an important factor behind the T2D

epidemic, social epidemiology may point in the direction of strategies for lowering sugar consumption at the population level, including such measures as sugar embargos.

Risk calculators – a simple content analysis

This section brings a description of the content of three risk calculators, for the purpose of illustration – FINDRISC (T2D) https://qxmd.com/calculate/calculator_236/findrisc-diabetes-risk-calculator , the Diabetes UK calculator https://riskscore.diabetes.org.uk/start , and FRAX (osteoporosis) https://www.sheffield.ac.uk/FRAX/ . Closer analyses are obviously possible, but outside the scope of this paper. The aim here, is to give the unfamiliar reader a brief introduction to some important features of these instruments.

Input to the calculators come in the form of a questionnaire, covering what is considered to be the relevant risk factors contributing to an individual's risk of becoming diseased. Common background factors in all calculators are sex and age. For osteoporosis and T2D body characteristics like height and weight are also included, as well as waist circumference for the latter. In addition, factors like family history, health behaviour, relevant medical history and medication use are registered.

A feature common to FRAX and the Diabetes UK calculators is that they also ask for further demographic information, in the form of nationality and/or ethnicity (FRAX) and ethnicity

(Diabetes UK). FINDRISC contains no such questions, but has Finnish nationality as its implicit background category.

Both FINDRISC and FRAX were originally made for use by professionals, but are freely available on the internet for all to use, whereas the Diabetes UK calculator is presented as a free for all instrument. An important feature making the calculators useable by ordinary citizens is that they do not require any laboratory testing or other measurements that can not be made by an adult person. To make the registration simple to accomplish, many of the questions are in a yes/no format.

Feedback is provided instantly, in the form of numbers. In FINDRISC the score comes as a number between 0 and 26, where a score below 15 represents low to moderate risk, a score from 15 to 20 high risk, and a score above 20 very high risk of diabetes 2. Diabetes UK have four risk categories – low, increased, moderate, and high, framed in the colours green, yellow, orange and red. The scores range from 0-47 and the range for each category is given in the coloured frame. FRAX provides its numbers in the format the 10-year risk of suffering a hip fracture and other major osteoporotic fractures. Such numbers are also provided by the diabetes calculators, but only as secondary information. These scores are presented as absolute risk scores, giving the percentage of the population that is likely to be affected over the stated timespan given the combination of risk factors for this particular individual.

Diabetes UK also provides an explanation for the score, including the advice of focussing on changeable factors. Furthermore, it contains answers to the question "what to do now?", where people with moderate risk are asked for improvements, whilst those with a high risk are advised to seek professional help. People with low risk are asked to spread the word about the calculator, basically communicating that we all have a job to do in the prevention of T2D.

Disclaimers feature on both diabetes calculators, warning that they are not diagnostic tools. Such disclaimers are most likely placed there for legal reasons, for the protection of the web-owners rather than people's health. It is also worth noting that no warnings are made about the epistemic uncertainty encapsulated into the risk score, as it represents that of a number of individuals with the same characteristics and not necessarily the individual using the calculator. As will be seen from the next section, there are others things related to the construction of risk calculators that can be worthy of a disclaimer.

A critical look at the construction of risk scores

People are frequently reminded that information acquired from the internet varies in quality, requiring a cautious approach. If seen as a health service of sorts, it is not unreasonable to expect webpages offering risk calculations to provide relevant information for people to make informed choices when having their risk calculated. Such background data is not necessarily provided, like in the case of T2D risk calculator applications for smart phones (Fijacko, Brzan and Stiglic 2015). Whether the proposed users of risk calculators care to run background checks is doubtful,

but as indicated there are good reasons for them to do so, as there are important epistemological issues involved.

Two major issues in the construction of risk scores are 'what risk factors to include in the score?' and 'what populations should be included in the data sources?' The choices made strongly reflects the current status of epidemiology, often involving what Shim (2014) calls the "usual suspects", risk factors that are easily measured and standardised, eliminating relevant economic and cultural factors from the knowledge upon which the calculators are built. Furthermore, cohort studies are presented as providing the ideal data for risk scores, but many scores are based on whatever epidemiological data are available, collected for purposes other than the construction of such scores (Noble et al. 2011).

A major issue in these constructions concerns what populations the different instruments are capable of making predictions about, exemplifying what Epstein (2007) calls the politics of difference in medical research. Apart from the universal construction of subpopulations based on age and sex, the categorisation of individuals is frequently attached to nationality, ethnicity and/or race, as well as a notable absence of categories for income status and educational status.

In the USA epidemiological data are constructed for subpopulations based on notions of race and ethnicity, basically White (aka Caucasian), Afro-American, Hispanic and Asian, although there is considerable variation and disparity in this categorisation (Megyesi, Hunt and Brody 2011). Ethnicity is also included in the Diabetes UK calculator, reflecting its more recent history of

immigration from former colonies, whereas in the Australian diabetes calculator https://www.diabetesaustralia.com.au/risk-calculator the white majority population is contrasted by a mesh of aboriginal subpopulations and more recent immigrant populations.

In Europe calculators are typically constructed at the national level, such as the Finnish (Lindström and Tuomilehto 2003), German (Mühlenbruch et al. 2014), Danish (Glümer et al. 2004) and Dutch (Damman et al. 2017) diabetes calculators. In its original format the Finnish score was simply called the Diabetes Risk Score, without any indication of being anything less than a universal score (Lindström and Tuomilehto 2003).

These demographic characterisations reflect historical, political and cultural processes in their respective countries. In modern epidemiology they are serving as boundary objects (Shim 2014), providing the black box of ethnicity as a tool making the job of constructing risk scores doable without having to grapple with challenging epistemological issues.

Instead of taking on this challenge, epidemiologists have concentrated on the technical challenge of epidemiological calibration. This relates to the question of whether calculators are usable for making predictions in populations outside the one it was originally constructed for. In the US context such questions were raised in connection with both the Framingham (heart disease) and the GAIL (breast cancer) risk scores, as the narrow population base for these calculators led to questions about their suitability outside their original white populations. In the case of breast cancer risk, calibration of risk scores for Asian and Pacific Islander American Women,

characterised as successful, have been reported (Matsuno et al., 2011). Despite reported improvements, the American Heart Association's instrument for calculating risk of cardiovascular disease is not yet able to offer scores for Asian and Hispanic subpopulations based on epidemiologic studies of these populations (Kullo et al., 2014), leading to concerns over possible overestimation of risk in diverse populations (Rana et al. 2016). This development serves to illustrate that the development of these risk scores is "work in progress", but it clearly also demonstrates "the assumed normativity of whiteness" (Shim 2014) in American epidemiology.

Outside the United States there has also been efforts to see if the Framingham model could be used in other populations. These efforts has led to the acknowledgement that the risk scores can not be generalised to populations in China (Liu et al. 2004), France (Vergaud et al. 2008) and Switzerland (Marques-Vidal et al. 2009), claiming that the use of the original Framingham calculator in these populations would lead to overestimation of risk, with potential overtreatment as a likely outcome. Such concerns addresses questions of doing the things right, but omits the question of whether they are the right things.

These issues become even more pertinent when we ask if this approach can be successfully adopted in low income countries. Looking at the literature, we find ample examples of calibration of risk scores developed in high income countries to be used on population in low and middle income countries.

T2D is presented as a very potent health problem globally, affecting low- and middle-income countries in Sub-Saharan Africa (Omec et al. 2016), the Middle-East (Janghorbani, Adineh and Amini 2013) and Asia (Ku and Kegels 2013) in particular. As laboratory testing for diabetes is costly, FINDRISC is considered to be a valid and inexpensive tool, making it a popular instrument for importation in a strategy for identifying those individuals most in need of laboratory testing. Positive conclusions about its usability has been drawn in Botswana, despite modest effectiveness (Omec et al. 2016), perhaps reflecting that better alternatives are not affordable.

A common exercise in making calibration work is to adjust the cut-off point, resulting in a person identified as a high risk person with a score of 7 or higher in the Philippines, 13 or higher in Iran and 17 in Botswana, all based on local calibrations of the FINDRISC calculator. Furthermore, lack of statistical power has not deterred researchers from concluding that FINDRISC will be a useful instrument in the Philippines (Ku and Kegels 2013). To reach this conclusion they also excluded questions of exercise and diet from their modified version of the calculator, in effect omitting lifestyle from a calculator originally designed to address just that. With lifestyle out of the causal equation, it remains to be seen whether it still remains part of the solution in this context.

Pragmatic calibration manoeuvres are not restricted to low income countries, however. Similar exercises have been performed in New Zealand, leading to the conclusion that FINDRISC is effective in identifying prediabetes and T2D among overweight New Zealanders in general, but

in particular among overweight Maoris (Silvestre et al., 2017). In the construction of the Australian Type 2 Diabetes Risk Assessment Tool (AUSDRISK), creative calibration manoeuvring was applied to fit the risk of various Aboriginal populations into the equation, whilst occupation and education were disregarded despite showing significant relationships in the data (Chen et al., 2010).

Overall, epidemiological calibration seems to be an endless source of epistemic puzzles. One study concludes that FINDRISC is a robust prediction instrument in a Middle Eastern population in Iran (Janghorbani, Adineh and Amini 2013). Another claims that being part of a Middle Eastern population from Iraq should be considered to represent an independent risk indicator in Sweden (Bennet et al. 2014), based on speculation around metabolic differences.

The examples given here may reflect scattered, individual efforts to solve the problems of countries around the world. Perhaps a more unified, global effort would present a way around these problems? Probably the most ambitious development of a risk calculator so far is FRAX, developed to provide a global calculator calibrated for every country in the world (Kanis et al. 2011). Its global base is a set of clinical risk factors presented as valid for all of humans, accompanied by a national calculator calibrated for each individual country. As a work in progress, the number of national calculators have been growing steadily since its introduction in 2008, and is at the time of writing spanning 63 countries.

The extensive calibration effort behind FRAX is presented as a necessity arising from the large variations in fracture risk globally. Such variations also exist between regions and ethnic groups in different countries, but the task force developing FRAX prefer national data, believing them to be of higher quality than the regional ones. Despite the expressed preference for national data, FRAX also comprises national calculators for ethnic groups in Singapore and the USA. In doing so, FRAX is also touched by some of the problems experienced by other calculators, as the original US fracture data was based on a predominantly white population, reflecting a lack of similar data for other US subpopulations (Ettinger et al . 2010). In contrast to the situation for T2D, no ethnic subpopulations have been created in the FRAX calculators for the UK, Australia and New Zealand.

Lack of available fracture data has led the FRAX developers to create what they call surrogate models, where countries lacking data are encouraged to use data from countries believed to be similar to themselves (Kanis et al. 2011). Furthermore, in the total absence of data countries are encouraged to base their models on the fracture pattern of Sweden. Not because Swedish bones are believed to have universal qualities, but because Sweden is believed to have high quality data on fracture epidemiology (Watts, Ettinger & LeBoff 2009).

To conclude this section, it seems fair to claim that the construction of many risk calculators present more problems than solutions at present. It remains open to debate whether the problems can be solved within the frame of the individual risk or whether other approaches to public health

are required. We will return to this towards the end of the paper, but before doing so we move to the use of risk scores.

Some forms of use of risk calculators

When it comes to the use of online calculators, reports on ordinary citizens' use of them is limited in the scientific literature. One reason for this may be that such use is not as frequent as their proponents hope. The owners of the websites hosting these instruments obviously have the means to monitor their use, but this knowledge is mainly outside the public domain.

Visitors to the FRAX website are, however, informed continuously about the number of fracture risks assessed by the instrument. At the time of writing the counter shows that close to 21.5 million assessments have been made since June 2011, meaning that roughly 260 000 fracture risk assessments are made monthly or 3 million annually. These numbers demonstrate an increase compared to earlier reported numbers (Kanis et al., 2014). In 2014 the published frequencies also showed that fracture risk assessments had been made in 173 countries, but that the use of the calculators varied substantially among these nations. Its most frequent use is in North America, Europe and Oseania, whereas numbers from Africa and South East Asia are very low. The potential for offering cheap fracture risk assessments online in low-income countries, has thus not been fulfilled so far.

A similar situation is observed with regard to risk assessment tools for T2D, where a major conclusion from a review of the implementation of these instruments is that they are not widely used (Dhippayom, Chaiyakunapruk and Krass 2014). This is explained as the result of reluctance on part of both healthcare practitioners and the public when it comes to implementation, due to such factors as negative attitudes, impracticalities, economical reasons, lack of perceived severity of the health issue, and concerns about the knowledge to be gained.

The potential for use of online calculators by ordinary people is growing with the increasing number of calculators that are not requiring laboratory tests or other costly, high tech equipment to produce a risk score. In the area of T2D, at least 65 risk scores classified as non-invasive are accessible (Dhippayom, Chaiyakunapruk and Krass 2014). A similar feature is available in FRAX, which now provides fracture risk assessments without the use of bone mineral density measurements (Kanis et al. 2015). Availability of online instruments is no guarantee for their use, however.

An insight into this comes from the initial phase of the German Diabetes Risk calculator (Holmberg, Harttig, Schulze and Boeing 2011). Its website was mainly approached through direct hits rather than by means of search engines, meaning that communication about the calculator's existence through other media was paramount for its use. This indicates that online risk calculators do not hitherto serve an unmet need among ordinary citizens, but that such a need will have to be created if their use is to grow.

Claims have been made about low health literacy as a barrier against sense making of risk scores in low-income countries (Noble et al. 2011), making these tools unfit for importation. Such claims serves as a sort of victim blaming, ignoring the influence of more structural problems. The public's failure to understand risk scores the way professional actors expect them to do, is not a challenge reserved for some countries, as demonstrated by Damman et al.'s (2017) examination of lay people's understanding of the Dutch diabetes calculator. Among their findings was that a majority of experimental users missed the important take home message that people with elevated risk should make an appointment with their doctor. Furthermore, the challenge of sense making in these contexts is amply illuminated in the rich literature on risk communication (See Adelswärd and Sachs (1996) for an illustrative example).

Although disclaimers tell online users that a risk assessment is not a medical diagnosis, it is reasonable to believe that they are used as tools for self-diagnosis, a practice described as making people vulnerable to disease mongers (Jutel 2011), actors "trying to convince essentially well people that they are sick" (Payer 1992:5). Among these actors are the pharmaceutical industry, who not only produce pharmaceuticals, but also provide online risk calculators as part of the package promoting the sales of their bestselling, risk-reducing products (Ebeling 2011). This is most evident in countries where direct to consumer (DTC) marketing of drugs are allowed, including online risk assessments as an important component of such campaigns.

Although not developed directly by the industry, but with unrestricted research grants from a number of companies (Kanis et al. 2012), the explicit goal of using FRAX is to identify high-risk

individuals for the sole purpose of providing treatment in the form of pharmaceuticals. A similar development is found in relation to breast cancer, where the early version of the GAIL calculator sought to identify high-risk women as candidates for mammography, whereas the second version identifies them as candidates for chemoprevention in the form of Tamoxifen (Fosket 2004).

In the processes of pharmaceuticalisation risk scores are used by means of a continuous expansion of the number of people defined as being at risk, serving as tools of the medicalisation of normality, leading to the majority of the adult population being included in the intervention groups for risk conditions (Skolbekken 2008). This work also involves the development of guidelines, like the 120 guidelines that now have incorporated FRAX scores (Kanis et al. 2016). When the guidelines show that a risk threshold has been reached, the solution more often than not comes in the form of "drugs for life" (Dumit 2012), aiming to help the individual to gain control over a risk situation that resembles a chronic disease.

The availability of opportunities for online risk assessment has furthermore been described as part of a development that is changing the patient doctor relationship. In this scenario a combination of online risk calculators, treatment guidelines and online provision of pharmaceuticals is credited with the potential of making the primary care physician redundant (McKinley & Marceau 2008).

Others have shown that it may not be as straightforward to eliminate doctors from this equation, as many patients still trust their doctors and feel a need to consult them when faced with

challenging health decisions. For instance, the notion of patients as consumers did not come through very clearly when low dose statins were available as over the counter drugs in Britain (Will and Weiner 2015). Rather, a strong resistance was demonstrated by the public, leading to the eventual withdrawal of these drugs from the market.

Overall, despite offering independent use of risk calculators to the public, risk assessment is still mainly used as an integrated part of clinical medicine and public health practice (Noble et al., 2011), including community pharmacies (Dhippayom et al. 2012). Their use in these settings are obviously done with the best of intensions, but can they also have other effects? This is the pertinent question we turn to when looking at the function of online risk calculators.

Intended and unintended functions of risk assessments

As noted from the beginning of this paper, the intended function of risk scores is to serve as a tool for the protection and promotion of both individual and public health, by identifying high-risk individuals in need of medical intervention, whether it be in the form of lifestyle changes or pharmaceutical chemoprevention.

More specifically, in the ambitious words of Holmberg, Harttig, Schulze and Boeing (2011:106) the intended function of using risk scores comprises the tall order:

> "to correct subjective risk perception, to increase rational decision-making in hypothetical scenarios of clinical decision-making, to ensure adherence to recommended screening behaviors, and to identify those who may profit from health interventions."

When it comes to scores for diabetes risk most of them are constructed on the belief that they will have a function in disease prevention, but it turns out that only a handful of studies have been performed to check whether these beliefs are knowledge based (Noble et al. 2011). Similar concerns have been aired about the use of risk assessments in strategies for preventing CHD (Kullo et al. 2014). These are only a few examples that contribute to the much larger discussion around the successes and failures of this approach to be found in the literature, a discussion that is outside the scope of this paper. It is interesting to observe, however, that Theresa Marteau (2018), who for a long time was engaged in research on how to present risk information to change behaviour, has come to the conclusion that individual risk information does not achieve this, and that the environment has a much stronger impact. Although put in psychological terms of conscious and non-conscious thinking, it is possible to interpret this as an acknowledgement that a public health model favouring structural interventions is better than an individualised one.

Despite this acknowledgement, a major concern about the present situation is that a vital function of risk calculators is not only to contribute to what Dubriwny (2013) calls "the vulnerable empowered woman", but that they also contribute to the depiction of "the empowered vulnerable global citizen". This may not come so much through the direct use of such calculators by ordinary citizens, as from the use of these instruments as an integrated part of clinical practice.

An important function of the calculators thus comes as a contribution to the preservation of a status quo, disguising fundamental health inequalities as public health issues are individualised instead of politicised. In a global context this is demonstrated by the analysis of thresholds for interventions against osteoporotic fractures (Kanis et al. 2016), where individuals with the same risk scores have very different chances of being helped. When presented as simple and factual information, this can be seen as a missed opportunity for fulfilling equal rights for health, but rather demonstrating how health for all is outside the scope of this framework.

In a way, the neoliberal framework of which risk calculators are a component takes as its starting point that "every man is an island", omitting variables that put issues as work, education, housing and income; i.e. social class into the equation. The effect of this in people's everyday life is vividly illustrated in Shim's (2014) analysis of the many social processes leading to inequalities in cardiovascular disease, which are effectively going under the radar of risk factor epidemiology. Rather than focusing on health disparities as the outcome of social inequality, individualised risk assessment approaches outline them as related to ethnic diversity, including notions of race in the US (Pollock 2012).

Whereas ethnicity is a concept with multiple, and at best uncertain meaning from the perspective of social science (Banks 1996), it is simply another box to be ticked off without further ado in medical research. Whether ethnicity is understood as representing biology for some and culture for others, does not really matter, as ethnicity serves as a boundary object that enables medical researchers to avoid the epistemological tensions that a conceptual clarification would bring

about. For instance, when users of the Diabetes UK calculator are asked about their ethnicity they are given the options of South Asian, Black, Chinese, Mixed, White and None, but when the outcome is communicated ethnicity has become a dichotomised variable consisting of white and other, where white is scored as 0 and other as 6. These scores are furthermore communicated as unchangeable facts, pointing in the direction of ethnicity as representing a biological destiny of sorts. The "normativity of whiteness" (Shim 2014) may thus not be an exclusively American invention.

Risk calculators including boxes for ethnicity can thus be seen as reifying notions of inequality as the outcome of unchangeable differences rather than as disparities that can be seen as both unfair and amendable, worthy of political attention and public health intervention. It is thus not unproblematic to say that yes will be the undisputed answer to whether we are doing things right, and even less so about doing the right things.

Concluding remark

The internet enables the calculation of more risk numbers than ever before, by substantial parts of the global population. So far there is an underused capacity for risk calculation, as the majority of the lay public has not felt the urge to know their risk numbers. Paradoxically this may be a good thing, as whatever meaning is attached to the numbers produced, there is a real risk of getting the wrong number, for both epistemological and socioethical reasons.

Literature

Adelswärd V and Sachs L (1996) The meaning of 6.8: Numeracy and normality in health information talks. Social Science & Medicine 43 (8): 1178-1187.

Armstrong D (1995) The rise of surveillance medicine. Sociology of Health & Illness 17(3): 393-404.

Aronowitz R (2015) Risky medicine. Our quest to cure fear and uncertainty. Chicago: The University of Chicago Press.

Banks M (1996) Ethnicity: Anthropological constructions. London: Routledge.

Bennet L, Groop L, Lindblad U, Agardh CD and Franks PW (2014) Ethnicity is an independent risk indicator when estimating diabetes risk with FINDRISC scores: A cross sectional study comparing immigrants from the Middle East and native Swedes. Primary Care Diabetes 8: 231-238.

Bouk D. (2015) How our days became numbered. Risk and the rise of the statistical individual. Chicago: The University of Chicago Press.

Burr V (2015) Social constructionism. 3rd edition. London: Routledge.

Chen L, Magliano DJ, Balkau B, Colagiuri S, Zimmet PZ, Tonkin A; et al. (2010) AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. Medical Journal of Australia 192 (4): 197-202.

Clarke AE, Mamo L, Fosket JR, Fishman JR, and Shim JK (eds) (2010) Biomedicalization. Technoscience, health and illness in the US. Durham: Duke University Press.

Damman OC, Bogaerts NMM, van den Haak MJ and Timmermans DRM (2017) How lay people understand and make sense of personalised disease risk information. Health Expectations 20: 973-983.

Dhippayom T, Fuangchan A, Tunpichart S and Chaiyakunapruk, N (2012) Opportunistic screening and health promotion for type 2 diabetes: an expanding public health role for the community pharmacist. Journal of Public Health 33 (2): 262-269.

Dhippayom T, Chaiyakunapruk, N and Krass I (2014) How diabetes risk assessment tools are implemented in practice: A systematic review. Diabetes Research and Clinical Practice 104: 329-342.

Dubriwny TN (2013) The vulnerable empowered woman. Feminism, postfeminism and women's health. New Brunswick: Rutgers University Press.

Dumit J (2012) Drugs for life. How pharmaceutical companies define our health. Durham: Duke University Press.

Ebeling M (2011) 'Get with the program!' Pharmaceutical marketing, symptom checklists and self-diagnosis. Social Science & Medicine 73: 825-832.

Epstein S (2007) Inclusion. The politics of difference in medical research. Chicago: The University of Chicago Press.

Ettinger B, Black DM, Dawson-Hughes, Pressmann AR and Melton III LJ (2010) Updated fracture incidence rates for the US version of FRAX. Osteoporosis International 21: 25-33.

Fijacko N, Brzan PP and Stiglic G (2015) Mobile applications for type 2 diabetes risk estimation: a systematic review. J Med Syst 39: 124.

Fosket J (2004) Constructing “High-risk women”: The development and standardization of a breast cancer risk assessment tool. Science Technology and Human Values 29(3) 291-313.

Glümer C, Carstensen B, Sandbæk A, Lauritzen T, Jørgensen T and Borch-Johnsen K (2004) A Danish diabetes risk score for targeted screening. Diabetes Care 27 (3):727-733.

Greene JA (2007) Prescribing by numbers. Drugs and the definition of disease. Baltimore: Johns Hopkins University Press.

Holmberg C, Harttig U, Schulze S and Boeing H (2011) The potential of the internet for health communication: The use of an interactive on-line tool for diabetes risk prediction. Patient Education and Counseling 83: 106-112.

Holmberg C and Parascandola M (2010) Individualised risk estimation and the nature of prevention. *Health Risk & Society* 12(5): 441-452.

Janghorbani M, Adineh H and Amini M (2013) Finnish Diabetes Risk Score to predict type 2 diabetes in the Isfahan diabetes prevention study. Diabetes Research and Clinical Practice 102: 202-209.

Jutel AG (2011) Putting a name to it. Diagnosis in contemporary society. Baltimore: The Johns Hopkins University Press.

Kanis JA, Hans D, Cooper C, Baim S, Bilezikian JP, Binkley N, et al. (2011) Interpretation and use of FRAX in clinical practice. Osteoporosis International 22: 2395-2411.

Kanis JA, Harvey NC, Cooper C, Johansson H, Oden A, McCloskey EV et al. (2016) A systematic review of intervention thresholds based on FRAX. Archives of Osteoporosis 11: 25.

Kanis JA, Harvey NC, Johansson H, Oden A, Leslie WD and McCloskey EV (2015) FRAX and fracture prediction without bone mineral density. Climacteric 18 (Suppl 2): 2-9.

Kanis JA, Johansson H, Oden A, Cooper C, McCloskey EV and Epidemiology and Quality of Life Working Group of IOF (2014) Worldwide uptake of FRAX. Archives of Osteoporosis 9: 166.

Kanis JA, McCloskey E, Johansson H, Oden A. and Leslie WD (2012). FRAX ® with and without bone mineral density. Calcified Tissue International 90: 1-13.

Kinch SH, Doyle JT and Hilleboe HE (1963) Risk factors in ischemic heart disease. American Journal of Public Health 53(3): 438-442.

Krieger N (2011) Epidemiology and the people's health. Theory and context. Oxford: Oxford University Press.

Ku GMV and Kegels G (2013) The performance of the Finnish Diabetes Risk Score and a simplified Finnish Diabetes Score in a community-based cross-sectional screening of undiagnosed type 2 diabetes in the Philippines. Primary Care Diabetes 7: 249-259.

Kullo I, Trejo-Gutierrez F, Lopez-Jiminez F, Thomas RJ, Allison TG, Muvagh SK, et al. (2014) A perspective on the new American College of Cardiology/American Heart Association guidelines for cardiovascular risk assessment. Mayo Clinic Proceedings 89 (9): 1244-1256.

Latour B (1999) Pandora's hope. Essays on the reality of science studies. Cambridge, Massachusetts: Harvard University Press.

Lindström J and Tuomilehto J (2003) The Diabetes Risk Score. A practical tool to predict type 2 diabetes risk. Diabetes Care 26 (3), 725-731

Liu J, Hong Y, D'Agostino RB, Wu Zhaosu, Wang W, Sun, J, Wilson PWF, Kannel WB and Zhao D (2004) Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-provincial Cohort study. JAMA 291 (21): 2591-2599.

Marques-Vidal P, Rodondi N, Bochud M, Chiolero A, Pecoud A, Hayoz D, Paccaud F, Mooser V, Firmann M, Waeber G and Vollenweider P (2009) Predicitive accuracy of original and recalibrated Framingham risk score in the Swiss population. International Journal of Cardiology 133: 346-353.

Marteau T (2018) Changing minds about changing behaviour. Lancet 391: 116-117.

Matsuno RK, Costantino JP, Ziegler RG, Anderson GL, Li H, Pee D and Gail M (2011) Projecting individualized absolute invasive breast cancer risk in Asian and Pacific Islander American Women. Journal of the National Cancer Institute 103: 951-961.

McKinley J & Marceau L (2008) When there is no doctor: Reasons for the disappearance of primary care physicians in the US during the early 21st century. Social Science & Medicine 67: 1481-1491.

Megyesi MS, Hunt LM and Brody H (2011) A critical review of racial/ethnic variables in osteoporosis and bone density research. Osteoporosis International 22: 1669-1679.

Mühlenbruch K, Joost H-G, Boeing H and Schulze MB (2014) Risk prediction for type 2 diabetes in the German population with the updated German Diabetes Risk Score (GDRS) Ernahrungs Umschau 61 (6): 90-93.

Noble D, Mathur R, Dent T, Meads C and Greenhalgh T (2011) Risk models and scores for type 2 diabetes: systematic review. BMJ 343:d7163.

Omech B, Mwita JC, Tshikuka, JG, Tsima B, Nkomazna O and Amone-P'Olak K (2016) Validity of the Finnish Diabetes Score for detecting undiagnosed Type 2 diabetes among general medical outpatients in Botswana. Journal of Diabetes Research

Payer L (1992) Disease Mongers. How doctors, drug companies are making you feel sick. New York: John Wiley & Sons.

Petersen A and Lupton D (1996) The new public health: Health and self in the age of risk. London: Sage Publications.

Pollock A (2012) Medicating race. Heart disease and durable preoccupations with difference. Durham: Duke University Press.

Prior L (2003) Using documents in social research. London: Sage Publications.

Rana JS, Tabada GH, Solomon MD, Lo JC, Jaffe MG, Sung SH, Ballantyne CM and Go AS (2016) Accuracy of the atherosclerotic cardiovascular risk equation in a large contemporary, multiethnic population. Journal of the American College of Cardiology 67(18): 2118-2130.

Rose G (1985) Sick individuals and sick populations. International Journal of Epidemiology 14: 32-38.

Rothstein WG (2003) Public health and the risk factor. A history of an uneven medical revolution. Rochester: Rochester University Press.

Shim JK (2014) Heart-sick. The politics of risk, inequality, and heart disease. New York: New York University Press.

Silvestre MP, Jiang Y, Volkova K, Chisholm H, Lee W and Poppitt SD (2017) Evaluating FINDRISC as a screening tool for type 2 diabetes among overweights in the PREVIEW:NZ cohort. Primary Care Diabetes 11: 561-569.

Skolbekken J-A (2008) Unlimited medicalization? Risk and the pathologization of normality. In A. Petersen and I. Wilkinson (Eds) Health, Risk and Vulnerability (pp. 16-29). London: Routledge.

Star SL and Griesemer JR (1989) Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of vertebrate zoology, 1907-1939. Social Studies of Science 19:387-420.

Susser M (1998) Does risk factor epidemiology put epidemiology at risk? Peering into the future. Journal of Epidemiology and Community Health 52: 608-611.

Vergnaud AC, Bertrais S, Galan P, Hercberg S and Czernichow S (2008) Ten-year risk prediction in French men using the Framingham coronary score: Results from the national SU.VI.MAX cohort. Preventive medicine 47: 61-65.

Walker MJ and Rogers WA (2017) Diagnosis, narrative identity, and asymptomatic disease. Theor Med Bioethic 38: 307-321.

Watts NB, Ettinger B and LeBoff MS (2009) FRAX facts. Journal of Bone and Mineral Research 24(6): 975-979.

Will CM and Weiner K (2015) The drugs don't sell: DIY heart health and over the counter statin experience. Social Science & Medicine 131: 280-288.

Author biography

John-Arne Skolbekken is professor at the Department of Public Health and Nursing, the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. His research work covers medical risk discourse, medical screening, and the empirical ethics of biobank research.