European Journal of Epidemiology (2020) 35:1115–1121
https://doi.org/10.1007/s10654-020-00700-w

ESSAY

When will individuals meet their personalized probabilities? A philosophical note on risk prediction

Olaf M. Dekkers[1] · **Jesse M. Mulder**[2]

Received: 13 October 2020 / Accepted: 13 November 2020 / Published online: 28 November 2020

Abstract
Risk prediction is one of the central goals of medicine. However, ultimate prediction–perfectly predicting whether individuals will actually get a disease–is still out of reach for virtually all conditions. One crucial assumption of ultimate personalized prediction is that individual risks in the relevant sense exist. In the present paper we argue that perfect prediction at the individual level will fail–and we will do so by providing pragmatic, epistemic, conceptual, and ontological arguments.

Keywords Prediction · Individual risk · Reference class · Philosophy · Ontology

Joe

Meet Joe. Joe, a 60-year-old male, with hypertension, overweight and hypercholesterolemia, wants to know his cardiovascular risk. Based on the Framingham risk score, you calculate his predicted 10-year risk for myocardial infarction to be 10%. Joe, not completely satisfied, asks whether a more precise estimation would be possible, maybe even a truly 'personalized risk estimation'. You decide for a more personalized approach, a risk prediction model including measured circulating proteins [1], and provide Joe with a 5.8% 5-year predicted risk. Joe tries to reformulate this risk: 'Suppose I have seventeen identical copies, one of us will get a myocardial infarction in the next 5 years.' But realizing that his predicted risk, although more personalized now, seems still a population average, Joe further challenges the risk prediction: 'Now I know I'm one of those seventeen identical copies. But why is it not possible to tell me which one of those I am? Why can't you provide me with a *truly* individualized risk?' Joe challenges you to come up with a risk score of either 0 or 1 (for the given time-frame). (Mind that Joe explicitly adopts a frequentist's approach to risk. Although such an approach is often displayed in epidemiology [2], other approaches to risks are possible [3]).

Risk classification for different conditions is an old enterprise; and the Framingham study for instance, with its developed risk scores, has clearly contributed to cardiovascular risk management [4]. However, Joe is right in challenging the risk prediction, for the ultimate prediction–perfectly predicting whether individuals will actually get a disease–is still out of reach for virtually all conditions (ignoring prediction over exceedingly short time frames, the only plausible exceptions are monogenetically caused diseases). The optimists may hasten to add: 'not yet, but in the future, we will be able to have perfect predictions', an optimism echoed in the promise of *personalized medicine*, which takes into account individual variability for example in genes, environment, and lifestyle for each person [2].

In the present paper, we define *ultimate personalized prediction* (or *ultimate prediction* for short) as the prediction that assigns individuals a 0 or 1 true estimated risk (within a well-defined time-interval). In this paper we will argue philosophically that ultimate personalized disease prediction will fail; the concepts developed will apply likewise to predictions of disease occurrence and treatment effects. For the sake of the argument, we will assume that variables used for prediction are perfectly measured, and that prediction tools are modelled both optimally and validly. We also assume that for perfect prediction, whether at the individual or group level, all independent causal pathways of a condition should be included in the prediction model (for a technical

✉ Olaf M. Dekkers
o.m.dekkers@lumc.nl

[1] Department Clinical Epidemiology LUMC Leiden (OMD), Leiden University Medical Center, Albinusdreef 2, 2333 Leiden, ZA, The Netherlands

[2] Department of Philosophy and Religious Studies, Utrecht University (JMM), Utrecht, The Netherlands

exploration of causal thinking in risk prediction see [5], for a conceptual exploration see [6]).

One crucial assumption of ultimate personalized prediction is that individual risks in the relevant sense exist, i.e. that at some level it is inevitable that someone will (not) get a disease. Challenging this ultimate prediction can go in two directions: one can challenge the possibility of *knowing* this individual risk (a so-called epistemic position), or one can challenge the *existence* of a true individual risk (an ontological position). Clearly, if such individual risk does not exist, it can also not be known, whereas the opposite is not necessarily true, as, if such risk exists, we may still not be able to determine the risk for an individual.

On the need for complete information

Imperfect prediction is often interpreted epistemically, i.e., attributed to our limited state of knowledge: we must have failed to measure or take into account certain factors that are relevant for the prediction. It could then be argued that more detailed knowledge about disease causes will improve prediction, and finally enable ultimate prediction.

But what is the scope of this claim? Suppose, to take a rather extreme example, that a big asteroid crashes into the Earth, eliminating all life on Earth, two weeks after we gave Joe his estimated 5-year risk of 5.8% for a myocardial infarction. Had we taken the approaching asteroid into account, we would of course have had to change our calculation for Joe–after all, his chance for a myocardial infarction within the two remaining weeks of his life would have been close to zero. Now, although 'relevant for the prediction' in a very broad sense, this is not the sort of information that prediction should be based on. Prediction in medicine should be based on information that lies within the purview of medicine. What falls outside of that purview must simply be treated as background assumptions–when assessing his cardiovascular risk, we assume that this background will, for Joe's lifetime, remain roughly stable (no asteroids, no war). And that in turn implies that, even if an ultimate prediction for Joe's case is possible (i.e., one which assigns a cardiovascular risk of either 1 or 0), this result will be conditional on that large and largely unspecifiable background assumption. This is a major qualification that we need for our notion of ultimate prediction to make sense, even though this assumption of stable background conditions remains mostly implicit.

We should note that, without this qualification, the idea of perfect prediction boils down to the metaphysical thesis of determinism: that, given the precise state of the entire universe right now, all future states of the universe are already predetermined, so that there is only one possible future course of events. And, given quantum mechanics, we do have good scientific reason to think that this grand metaphysical thesis is false. (A more detailed philosophical discussion on determinism falls outside the scope of this paper; see [7, 8]).

Disease outcomes are predicted on the basis of a set of measured variables X1-Xn at time T = 0. For example, in prediction of cardiovascular diseases, this usually comprises classical risk factors such as cholesterol, smoking, blood pressure, diabetes, currently combined with data from genomics, metabolomics and lipidomics. However, ultimate prediction will only be successful if the set of variables is complete, in the qualified sense we have just introduced: that all medically relevant and independent risk predictors are included. For, if medically relevant factors outside the variables included make a difference regarding cardiovascular risk, ultimate prediction will fail. Now, clearly, cardiovascular risk is determined over the course of life by many determinants which are not constant over time. Think of cortisol, a stress hormone associated with cardiovascular risk. Cortisol levels change per day, per hour and even per minute. Arguably, ultimate cardiovascular risk prediction requires continuous measurements of cortisol over life, up till T = 0. The same holds for variables like blood pressure, or dietary patterns. As most variables are non-constant over time, the only exception probably being genes, this requirement of lifetime measurements applies to all risk factors X1-Xn up to T = 0. We call this the *Variable-Lifetime matrix*. A completely filled-out variable-lifetime matrix at T = 0 may now be thought sufficient (though probably not necessary) for ultimate prediction in our qualified sense. It may be, of course, that for certain variables the information included in the matrix displays a simple recurring pattern (say), in which case including the recurring pattern is sufficient. However, this makes no substantial difference to our picture, also as at least some variables do not display strong recurring patterns (diet, perceived stress).

Obviously, such a complete *Variable-Lifetime matrix* is highly unlikely to be realized, which is a first argument why ultimate prediction will not materialize. This may strike one as a mere epistemic point reflecting our limited knowledge state, or even just a pragmatic point concerning what is practically feasible in the foreseeable future, and not a fundamental argument. In response, one may point out that measurement *itself* is in fact an intervention which has real (ontological) biological as well as psychological consequences, so that increasing measurement intervals and scope will increasingly change the one that is being measured in ways that may make a difference [9] for the prospective risk assessments–and frequent and invasive measurement may even make a difference in what is being measured itself. For example, cortisol levels are likely influenced by (the knowledge of) measuring it, even if the actual effects of measurement may be very small. It would be rather awkward, of course, if Joe learned that his ultimate personalized risk

for myocardial infarction was 1, but that this was partly so because he was being measured all the time (perhaps his risk would have been 0 if personalized medicine had simply left him alone). Then again, future techniques may allow for instant data accumulation over time in humans, in which case one could treat the measurement itself not as an eventful intervention, but simply as a constant factor that is there anyway (a very smart smartwatch). We will not further discuss this topic here; our aim was just to highlight the conceptual point that such measurement is not just subject to merely pragmatic constraints but involves serious ontological and epistemic intricacies as well.

The fundament and limitation of prediction: comparing like with like

Generally speaking, the current concept of risk prediction is based on the course of a reference population to predict disease risks in new patients [2]. For example, the Framingham Risk score was originally developed in a cohort of 5,209 men and women [4]. These risk scores are now used for persons like Joe to predict their cardiovascular risk. The basic assumption is thus that the reference population is similar to people like Joe: identical clinical characteristics implies identical predicted risk (the converse is obviously not necessarily true as a similar predicted risk may be due to different underlying risk patterns).

Assigning a person to a reference population to determine someone's risk invites two problems. The first is that someone can be assigned to different reference populations based on different sets of predictors. Each individual has innumerable characteristics and can therefore potentially be assigned to innumerable reference populations; this is known as the reference class problem [10]. Joe, for example, was assigned to a reference population based on classical risk factors combined with measured proteins, but could also have been assigned to a reference class based more on psychological (perceived stress) and lifestyle factors (diet, sleep pattern, yoga). Interestingly, these different reference populations can lead to different risk estimations for the same person and the same outcome [11].

The second problem related to reference populations becomes clear when we consider the number of potential variables needed for ultimate prediction; for some diseases, like obesity or diabetes type 2, merely listing all the genetic loci associated with the condition already yields quite a large number of variables [12]. However, with only a modest set of 20 dichotomous variables the number of potential reference classes exceeds one million. In fact, for many multifactorial diseases, considering more than 20 variables to predict their occurrence is not overdone. This means that for most reference populations no predicted risks have been established empirically, and that in turn implies that the ideal of personalized medicine, viz., taking into account all variables involved in disease risk, is bound to fail, given that the large number of predictors 'strikes the curse of dimensionality' [13]. One could again argue that this dimensionality problem is epistemic and may be solved by increasing computing power; however, it is difficult to see how we can increase the knowledge base for all these reference classes, as this knowledge is based on available data and available populations [14]. And these are limited. (Although it is admittedly always possible to generate more (relevant) data, there are surely both practical and ethical limits to (our possibilities of) increasing the available populations.) And so we may safely conclude that this problem can, theoretically, only be solved if we will be able to predict from biology and not just from population statistics.

Does biology behave deterministically? There may be a role for randomness

Back to Joe with his 5.8% predicted risk. The fundamental question is the ontological one whether indeed a true 1-or-0 risk exists (for Joe). The existence of an individual 1-or-0 risk would mean that if we would have 1,000 copies of Joe under identical conditions and we let time start at $T=0$, all copies either will get the disease or will not get the disease. If a 1-or-0 risk at the individual level does not exist, then a population average is the closest we can get, even though the points discussed before would apply. Indeed, it has been claimed by several authors that such individual risks do not exist [11]. This question touches the basis of disease biology. If individual 1-or-0 risks do not exist, the combination of all causal factors to predict risks at $T=0$ does not fully determine this risk.

For several organs, it is an intriguing phenomenon that if cancer occurs in one of a pair of organs, there is little elevation of cancer risk on the contralateral organ [15]. This is for example the case for testis, breast and adrenal cancer. This may at first glance be a surprise, as major risk factors (diet, smoking, genes, circulating carcinogens) presumably behave largely identical for contralateral organs. Why then such risk-difference between contralateral organs, even though some local differences may be involved? This consideration has given rise to the idea that chance may play an ontological role [16]. If so, then, even if we assume perfect knowledge and a complete variable-lifetime matrix, and treat our target individual as a closed system (on which more below), we will not be able to predict the event to occur–not because we missed something, but simply because there is nothing more to determine. Then, we have to conclude that what happens is ultimately partly a matter of mere chance. A second explanation would be that, although contralateral

organs are very similar, there are still minor differences which finally translate into very divergent future states, a feature described as 'sensitive dependence on initial conditions (SDIC)' [17]; in such explanatory context, chance is not necessarily involved.

It is actually very hard to see how one could ever determine empirically whether chance plays an ontological role in medicine. It may also be that for some, but not for all diseases, chance plays a role [18]. If, however, chance is involved in disease development, it follows straightforwardly that we will never be able to give Joe his ultimate individual 1-or-0 risk–simply because it does not exist. Reconceiving the individual risk, we could then say that it was, say, 0.8–where the fact that it lies somewhere between 0 and 1 is then understood to reflect an ontological reality rather than epistemic limitations [3].

Here, we come in the vicinity of the vexed question what assigning an objective (ontological) probability to a single event really comes to. Very briefly, such assignment may be taken to rest on a distribution over a certain range of events, as in classical frequentist theories, or it can be genuinely single (as in, e.g., 'single-case' propensity theory; see [19, 20]. See [21] for in-depth discussion.

The set of variables used to predict is not a closed system

Let us assume that everything in Joe is measured perfectly and let us also assume that we have perfect data on several persons with a risk-profile identical to Joe's (and let us assume that the reference class problem is solved and that chance plays no role). Will we be able to predict occurrence of a cardiovascular event in Joe with certainty, on the assumption that all things remain equal outside of the medically relevant sphere? Only if all future events and behaviour that affects cardiovascular risk can be perfectly predicted based on all measurements at $T=0$.

But suppose Joe likes a drink, resulting in an alcoholic hepatitis, only a couple of weeks after he received his calculated cardiovascular risk. Joe decides to become a total abstainer and succeeds in doing so. This will probably make a difference to his cardiovascular risk. If the previously calculated risk must now be updated, it is shown to not have been ultimate after all. However, although not yet realistic, it is conceivable that behaviour prediction is possible when everything is measured (genes affecting behaviour, Joe's responses to lifestyle events, etc.). If so, then the calculated cardiovascular risk will have already taken into account the hepatitis and Joe's subsequent behavioural response, so that no update of that risk needs to be undertaken. If this is possible, the system to predict Joe's risk may be called 'closed' at $T=0$.

This suggests the following picture: setting aside the unlikely arrival of an apocalyptic asteroid and similar external disruptions of the background against which Joe lives his life, we can treat him as a *closed system*; and then, the variable-lifetime matrix will be all we need to enable perfect prediction.

However, this picture rests on an ontological assumption that must be rejected, for there is no such thing as a closed system in medicine. Here are some examples that illustrate why; we will further develop the underlying philosophical point in the next section. Suppose that, the same day you provided Joe with his 5.8% predicted risk, he stumbles while descending the stairs, and is unfortunate enough to break his leg in the ensuing fall. His lifestyle changes, becoming more sedentary for a few months. This will increase his cardiovascular risk, and though it is obviously an accident, it doesn't look like we can assign it to a sudden change in the background assumption, as we could in the case of the apocalyptic asteroid. Similar examples can be easily thought of, such as a change in lifestyle because a good friend got a myocardial infarction, or Joe's moving to Southern Italy and adopting a Mediterranean diet after having fallen in love with an Italian woman, or changing environmental conditions because the city finally decides to shut down the air-polluting coal plant, etc. Clearly, all of these circumstances will affect cardiovascular risk but could not have been predicted based on all variables known at $T=0$. Perhaps some of the examples (changing environmental conditions) can be dismissed as they seemingly violate the assumption of background stability. But the more we do so, the more we lose Joe, in all his concreteness, entirely from sight, so that any 'ultimate' prediction that then in the end results will have little value, if at all, to Joe.

So even when we assume that the background against which Joe is living his life remains roughly constant, we cannot conceive of Joe as a closed system. Future factors not included in, or predicted by, the model, may influence risk, and because they were not included, ultimate prediction will fail. Even more, conveying to Joe what his cardiovascular risk is may be considered an intervention in Joe's life, and as such may encourage Joe to change his behaviour. Joe may then decide, after being confronted with an ultimate risk prediction of 1, to eat more salad and do yoga every day. And paradoxically, the very idea of an ultimate prediction appears to preclude precisely such adjustments to one's life based on the prediction: being ultimate, it must have already included whatever behavioural adaptations one is going to implement upon learning about it. This argument develops the argument displayed above, that measurement itself can affect human biology and psychology, and therefore risk.

It is a hallmark of biology that organisms are open systems, so that the processes in which living beings engage are interrupted or affected all the time [22, 23], and as we

have argued, this can occur by factors that were not included in the initial set of predictors. Cleary, the longer the time and the number of steps involved, the more difficult disease prediction will be as more events can affect the disease process. As with the weather, predicting the rain likelihood for tomorrow will often work, though predicting the exact temperature in Amsterdam three years hence is impossible [24].

Doesn't this, again, merely reflect our limited (medical or meteorological) knowledge state? That suggestion takes us back to our above discussion of the unqualified idea of ultimate prediction, where we concluded that that idea ultimately collapses into a mere statement of determinism, on which such ultimate prediction indeed seems possible (in theory), but in that case has to be based not on medically relevant factors but rather on the current state of the entire universe.

Biological processes

We have seen that there are practical, epistemic, conceptual, and ontological limits to the idea of ultimate prediction. The most central, philosophically speaking, is tied to our point that it is impossible to treat Joe as a 'closed system' in the sense that disease occurrence does not follow deterministically from all variables known at $T=0$ (not even if we assume stable background conditions). In fact, this is what medicine is about, as interventions aim to modify a disease process in such ways that the predicted course will differ from what would have occurred without the intervention. We will now attempt to sketch a more substantial philosophical underpinning for that observation.

To start, note that most of the variables used to predict Joe's health capture the *state* in which Joe's body, or a certain part of it, is. Examples are his cortisol level, blood pressure, or genetic make-up. The question is, now, how we should read such variables. According to one view, which found influential expression in the philosophy of the so-called 'new mechanism' [25], biological entities are conceived as 'mechanisms', and the entities or 'things' of which a mechanism is composed are fundamental. And biology is indeed full of descriptions of mechanisms, such as those of spindle formation and cancer genesis.

The question is whether such a mechanistic view can fully account for what happens in human biology. Think of a mechanical clock; it can be stopped, so that its parts, the 'things' that comprise the mechanism, do not engage in any clock-related processes for a while. Then, the clock can be wound up again, and the mechanism runs as it used to. By contrast, we cannot similarly take out Joe's cortisol, store it in the fridge, put it back after 24 h and hope Joe will run again. More generally, stopping all the biological processes within an organism typically results in the death of that organism. Also, the cortisol that we measure today, is not the same as the cortisol we measured last week, as the cortisol is newly formed constantly. In the biological case the 'things' are by far not as stable and rigid as in the case of mechanisms. The crucial point, then, is that the components of biological mechanisms *are* typically dependent on the life-processes of their host organ or organism. In short, then, processes in biology are as fundamental as things [7, 23]. This can be thought of in a similar way to physics, where what happens is determined by the interplay between things and laws [17].

Now back to the measured variables. What we measure is the state in which (some part of) the organism is. Our measurements thus relate to the various biological entities or things at the time of measurement. On the mechanistic picture, then, these measurements indeed capture fundamentally what goes on. But if processes are also fundamental in human biology, measuring entities may not capture the full picture. The tension between entities and processes can be seen when considering 'aging'. Aging is clearly a process, but attempts to completely reduce aging to variables (actual age) or things (telomere length) have failed to fully capture the aging process.

This invites the question: how does one measure processes? Well, one way to do so is, of course, to track the changes in (the relevant part of) an organism, and this might be done by successive measurements of the kind we have been considering. The variable-lifetime matrix introduced earlier in fact precisely does that. Suppose Joe has a grandson whose height is now 104 cm, one year later 109 cm. This growth is the result of a (growth) process, even if we have not directly measured the process itself. And when predicting growth, for example in Joe's grandson, we generally combine information on actual height and the parents' final height with information from growth curves. Such growth curves, based on combined growth data from many children, can be said to represent the growth process. Not including such growth curves when predicting height, will make the predicting task quite difficult and likely more inaccurate. However, there are not many instances in medicine where processes are included in the prediction model–not even via a proxy, like the growth curves. Still, standard textbooks in physiology and medicine are full of references to biological processes, such as clot formation, cell-growth, aging or metabolism. (We will not discuss the semantic-philosophical issue whether 'force', 'power' or 'process' is a better term describing the dynamic part of biology; also, for our argument, it is not extremely important that a full definition of process seems lacking [23, 26]).

If, now, dynamic processes and not only things are ontologically fundamental yet not incorporated in our prediction models, this may be another reason why ultimate prediction will fail. Think of cholesterol as included in Joe's prediction.

Although relevant for the development of atherosclerosis, a single value may not well capture the dynamic process of atherosclerosis formation in Joe. The same applies to actual age, which may not well capture Joe's aging process; in fact, medicine agrees that aging is a relevant process, without good variables representing this process. Second, several processes coincide, and the same material entity may induce different effects depending on which processes are active [22]. Similarly, genes unfold their effects depending on spatial, mechanical and electromechanical interactions with other molecules [13]. And silencing processes should be taken into account as well, which are typically not states or things. This means that, not considering processes and their interactions, we will fail to estimate the dynamics of disease development.

Now, again, one may argue that this is an epistemic or merely practical issue, as we currently lack tools to observe and measure the dynamics of disease processes directly. But there is more. The point is that processes are *ontologically* different from things and the states they are in. Processes can be interrupted and influenced; processes combine in non-trivial ways–in a word, processes are by their very nature 'open'. And this, we argue, is the philosophical reason why one cannot meaningfully treat Joe as a closed system. Accordingly, this is how health is generally considered, as a system that is open to interventions and changes in lifestyle.

Conclusions

We have argued that perfect prediction at the individual level will, with some rare exceptions, fail–and we have done so by providing pragmatic, epistemic, conceptual, and ontological considerations. And indeed, large scale genomic data, for instance, only add little to the prediction of complex conditions like obesity [27] or cardiovascular disease [28], nor have they brought prediction at the individual level closer [29]. It is also fundamentally unlikely that adding to the complexity of the data by combining different -omics, even when combined with machine learning techniques to master these data, will enable perfect prediction [30]. This does not mean that predicted population averages are not improving as a result of a more personalized approach; and such population averages may indeed be helpful to inform policy makers. And arguably, personalized medicine can be seen as an attempt to bring down a person's medical situation to the smallest epidemiological unit for which population based evidence exists (be it treatment or prediction) [31]. Why then so much emphasis on the impossibility of ultimate prediction? Isn't any scientific effort worthwhile that improves prediction in for example cardiovascular medicine [32], even if ultimate prediction is impossible?

We think our exploration of the limits of ultimate prediction is important as it highlights fundamental hurdles to disease prediction. Furthermore, we have highlighted fundamental questions relating to the philosophy of biology, in particular the question of chance and whether the ontological relationship between things and processes is different in biology as opposed to the non-living world. And such reflections might help us to arrive at a deeper understanding of the reality that our predictions aim to latch onto, and thereby of the limits of prediction.

Acknowledgements We thank Ewout Steyerberg, George Davey-Smith, Doranne Thomassen and Jan Vandenbroucke for constructive comments on a prefinal version of the manuscript.

Author contributions Both authors contributed to the study conception and writing.

Funding JM was funded by NWO (VENI) Grant number 275–20-068.

Code availability Not applicable.

Complaince with ethical standards

Conflict of interest No conflict of interest

Availability of data and material Not applicable.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication The authors both consent with publication.

References

1. Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, et al. Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. JAMA. 2016;315(23):2532–41.
2. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. BMJ (Clinical research ed). 2018;363:k4245.
3. Goodman SN. Probability at the bedside: the knowing of chances or the chances of knowing? Ann Intern Med. 1999;130(7):604–6.
4. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. Lancet (Lond, Engl). 2014;383(9921):999–1008.
5. Piccininni M, Konigorski S, Rohmann JL, Kurth T. Directed acyclic graphs and causal thinking in clinical risk prediction modeling. BMC Med Res Methodol. 2020;20(1):179.
6. Ickerman BA, Hernán MA. Counterfactual prediction is not only for causal inference. Eur J epidemiol. 2005;35(7):615–7.
7. Mulder JM. A Vital Challenge to Materialism. Philosophy. 2016;91(2):153–82.
8. Earman J. A Primer on Determinism: D. Reidel; D. Reidel Publishing Company; Dordrecht 1986.

9. Cartwright N. Hunting Causes and Using Them: Approaches in Philosophy and Economics. Cambridge University Press, Cambridge 2009.
10. Dahabreh IJ, Hayward R, Kent DM. Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. Int J Epidemiol. 2016;45(6):2184–93.
11. Stern RH. Individual risk. J clin hypertens (Greenwich, Conn). 2012;14(4):261–4.
12. Park KS. The search for genetic risk factors of type 2 diabetes mellitus. Diabetes metab J. 2011;35(1):12–22.
13. Personalized GA, Medicine P. In: The Routledge Companion to Philosophy of Medicine. Jeremy R Simon, Harold Kincaid: Edited ByMiriam Solomon; 2016.
14. Saracci R. Epidemiology in wonderland: Big Data and precision medicine. Eur J Epidemiol. 2018;33(3):245–57.
15. Smith GD. Post-modern epidemiology: when methods meet matter. Am J Epidemiol. 2019;188(8):1410–9.
16. Smith GD. Epidemiology, epigenetics and the "Gloomy Prospect": embracing randomness in population health research and practice. Int J Epidemiol. 2011;40(3):537–62.
17. Hoefer C. Causal Determinism. In: Zalta EN, editor. The Stanford Encyclopedia of Philosophy (Spring 2016 Edition). https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/
18. Coggon DI, Martyn CN. Time and chance: the stochastic nature of disease causation. Lancet (Lond, Engl). 2005;365(9468):1434–7.
19. Giere RN. Objective Single-Case Probabilities and the Foundations of Statistics. Studies in Logic and the Foundations of Mathematics. 74: Elsevier; 1973. p. 467–83.
20. Gillies D. Varieties of propensity. Br J Philosophy Sci. 2000;51(4):807–35.
21. Hájek A. Interpretations of probability. In: Zalta EN, editor. Stanford Encyclopedia of Philosophy (Fall 2019 Edition); 2019. https://plato.stanford.edu/archives/fall2019/entries/probability-interpret/
22. Dupré JA, Nicholson DJ. A Manifesto for a Processual Philosophy of Biology. In Nicholson DJ, Dupre JA, editors. Everything Flows: Towards a Processual Philosophy of Biology. Oxford University Press, USA; 2018.
23. Dupré J. Living Causes. Aristotelian Society Supplementary Volume 2013;87(1):19–37.
24. Vandenbroucke JP. Cerebral sinus thrombosis and oral contraceptives there are limits to predictability. BMJ (Clin res ed). 1998;317(7157):483–4.
25. Machamer P, Darden L, Craver CF. Thinking about mechanisms. Philos Sci. 2000;67(1):1–25.
26. Mulder JM. Varieties of Power. Axiomathes. 2020. https://doi.org/10.1007/s10516-020-09481-z
27. Loos RJF, Janssens A. Predicting polygenic obesity using genetic information. Cell Metab. 2017;25(3):535–43.
28. Morris RW, Cooper JA, Shah T, Wong A, Drenos F, Engmann J, et al. Marginal role for 53 common genetic variants in cardiovascular disease prediction. Heart. 2016;102(20):1640–7.
29. Joyner MJ, Paneth N. Promises, promises, and precision medicine. J Clin Investig. 2019;129(3):946–8.
30. Wilkinson J, Arnold KF, Murray EJ, van Smeden M, Carr K, Sippy R, et al. Time to reality check the promises of machine learning-powered precision medicine. The Lancet Digital Health. https://doi.org/10.1016/S2589-7500(20)30200-4
31. Smith GD, Egger M. Incommunicable knowledge? Interpreting and applying the results of clinical trials and meta-analyses. J Clin Epidemiol. 1998;51(4):289–95.
32. Knowles JW, Ashley EA. Cardiovascular disease: The rise of the genetic risk score. PLoS Med. 2018;15(3):e1002546-e.