



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Digital epidemiology, deep phenotyping and the enduring fantasy of pathological omniscience

Citation for published version:

Engelmann, L 2022, 'Digital epidemiology, deep phenotyping and the enduring fantasy of pathological omniscience', *Big Data and Society*. <https://doi.org/10.1177/20539517211066451>

Digital Object Identifier (DOI):

[10.1177/20539517211066451](https://doi.org/10.1177/20539517211066451)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Big Data and Society

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Digital epidemiology, deep phenotyping and the enduring fantasy of pathological omniscience

Big Data & Society
January-June: 1–13
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517211066451
journals.sagepub.com/home/bds



Lukas Engelmann 

Abstract

Epidemiology is a field torn between practices of surveillance and methods of analysis. Since the onset of COVID-19, epidemiological expertise has been mostly identified with the first, as dashboards of case and mortality rates took centre stage. However, since its establishment as an academic field in the early 20th century, epidemiology's methods have always impacted on how diseases are classified, how knowledge is collected, and what kind of knowledge was considered worth keeping and analysing. Recent advances in digital epidemiology, this article argues, are not just a quantitative expansion of epidemiology's scope, but a qualitative extension of its analytical traditions. Digital epidemiology is enabled by deep and digital phenotyping, the large-scale re-purposing of any data scraped from the digital exhaust of human behaviour and social interaction. This technological innovation is in need of critical examination, as it poses a significant epistemic shift to the production of pathological knowledge. This article offers a critical revision of the key literature in this budding field to underline the extent to which digital epidemiology is envisioned to redefine the classification and understanding of disease from the ground up. Utilising analytical tools from science and technology studies, the article demonstrates the disruptive expectations built into this expansion of epidemiological surveillance. Given the sweeping claims and the radical visions articulated in the field, the article develops a tentative critique of what I call a fantasy of pathological omniscience; a vision of how data-driven engineering seeks to capture and resolve illness in the world, past, present and future.

Keywords

Digital epidemiology, deep phenotyping, digital health, epistemology, genetics

This article is a part of special theme on Digital Phenotyping. To see a full list of all articles in this special theme, please click here: <https://journals.sagepub.com/page/bds/collections/digitalphenotyping>

Introduction

Digital epidemiologists design projects that utilise data from active and passive monitoring devices (such as smartwatches and phones) and from any potential trace that people might leave through their online actions and interactions to infer, collect, and survey pathological information. Moving far beyond social media tracking or the aggregation of data from health monitoring devices, this shiny, newish 'silicon valley' vision for epidemiology seeks to circumvent the stubborn, costly and what entrepreneurs consider inefficient medical apparatus. Since the rise and fall of Google flu trends (2008–2014), this growing field seems to offer an innovative expansion of epidemiology's data grasp, with impressive use cases across a wide range of infectious diseases, chronic conditions and mental health disorders (Park et al., 2018).

The successful operation of this new enterprise of data-driven disease surveillance relies on a technological and theoretical innovation: deep or, more recently, digital phenotyping. Both terms have come to describe a set of tools used to unify diverse and often non-medical data into coherent pictures of diseases. Deep phenotyping has emerged in the context of precision medicine in the early 2010s, mostly as a theoretical concept to expand data collection over the

Corresponding author:

Lukas Engelmann, STIS, The University of Edinburgh Institute for Studies of Science Technology and Innovation, Edinburgh, EH1 1LZ, UK.
Email: lukas.engelmann@ed.ac.uk



lifetime and beyond clinical encounters between patients and physicians (Robinson, 2012). Digital phenotyping, as the contributions to this special section demonstrate, has multiple origins in the context of the proliferation of big data in medicine. Utilising the truth claims associated with the radical empiricism of big data (boyd and Crawford, 2012) and as ‘guided knowledge discovery techniques’ (Kitchin, 2014: 6), the concept advances scientific approaches to the increasing variety and messiness of big data. As a form of ‘intensified data sourcing’ digital phenotyping is geared towards interoperability and re-purposing of data produced outside medical institutions, often in the name of personalised medicine (Hoeyer, 2019; Leonelli and Tempini, 2020; Prainsack, 2017). Epistemologically, the approach gives little concern to data collection practices – as opposed to the clinical trial or experiments – but as an abductive method it seeks to make ‘logical sense given what is already known about such data production’ (Kitchin, 2014: 6). As part of an ‘instrumentalist discourse’ on big data in health care (Stevens et al., 2018) phenotyping is brandished as innovative technology to develop new hypotheses and to deliver novel granularity to medical knowledge in place of outdated narratives, conventional knowledge and evidence practices of the past (Timmermans and Berg, 2010).

The conceptual and semantic overlap between deep and digital phenotyping is not a coincidence. When deep phenotyping became framed as an explicitly digital technology, it continued and exaggerated metaphors of depth, increasingly associated with a language of artificial intelligence and deep learning. However, as I argue below, this association of depth as a reservoir for causal inference remains vague and uncertain across the literature and is perhaps best understood as part of an exaggerated rhetoric of expectations, promises and visions.

This article offers a survey of recent literature promoting and advancing the research agenda of digital epidemiology. I ask how this technology is expected to transform the research methods in epidemiology and evaluate its epistemological implications. Utilising perspectives from science and technology studies (STS), particularly from critical data studies and the sociology of expectations (Borup et al., 2006), I analyse key scientific literature concerned with digital phenotyping since 2010 to understand how the technology is framed as promising and visionary practice, invested with the capacity to lead the transformation of medical knowledge production. I emphasise two critical aspects of this budding research framework. First, I argue that the implied proposition of a new classification infrastructure for medical knowledge production is, to a significant extent, not as utopian as it seems, but rather geared towards the resolution of shortcomings of the human genome project. Deep phenotyping shares much of its visionary claims with the promises of genomics (Fortun, 2008) and should foremost be perceived as a symptom of

what Jennifer Reardon calls the ‘postgenomic condition’ (Reardon, 2017). Unredeemed medical visions of the human genome project are sought to be patched up through a radical transformation of the observation and classification of pathological information. Second, I demonstrate that the vision of a digital epidemiology powered by deep phenotyping is problematically linked to fantasies of an overly general and overly precise medicine. Too often, this vision seems to fall prey to the naive idea of a totally confirmed world, in which a fantasy of pathological omniscience prevails. Fuelled by the technological feasibility to repurpose all data, the fantasy of a universally applicable standardization for medical diagnostics returns once again. Despite decades of growing scholarship from sociology, the history of science, from STS and data studies, ‘images of universal policy and encyclopaedic knowledge’ (Bowker and Star, 1999: 158) continue to invoke a ‘rhetoric of completeness’ (Jardine and Drage, 2018), in which the historical contingency and political thrust of classification schemes remain neglected.

To unpack this powerful fantasy, the first section delivers an outline of the expectations invested in digital epidemiology as a disruptive and innovative data-driven research enterprise that seeks to distance itself from conventional, traditional and cumbersome methods of measuring, surveying and classifying disease. Reliance on deep phenotyping furthers this vision of an epistemological transformation, rather than a digital expansion of surveillance and measurement, I argue in the second section. The third and final section will revisit the central claim of depth to critically discuss the scope and visions of digital epidemiology.

Reinventing epidemiology

To utilise data sources for surveillance which previously were not considered useful, and to transfer these into systems from which the dynamics of a disease can be inferred, has been a staple of epidemiological research since the early 20th century (Amelang and Bauer, 2019; Bauer, 2008; Morabia, 2004). Furthermore, as historians of statistics have long argued, standardisation pressure deriving from quantification always impacted on how clinical medicine and laboratory practice was carried out. To count disease, diagnostic criteria had to be unified and to account for an incidence rate, reliable practices of recording and reporting needed to be established. The history of data collection in the name of epidemiological research has never been one of merely scraping and collecting what physicians had noted down, but it is a history of establishing, negotiating and continuously adapting data standardisation (Desrosieres, 1998; Matthews, 1995).

The obvious point of comparison – although it is rarely made by the proponents of digital epidemiology and deep phenotyping – is the history of the International Classification of Diseases (currently ICD-11). As Susan

Leigh Star and Geoff Bowker argued, the ICD emerged as an important infrastructure for epidemiological and medical research and provides a ‘functioning means of coordinating information and work highly distributed over space and time’ (Bowker and Star, 1999: 139). However, its history also highlights a permanent tension between attempts of universal standardisation and the local circumstances in which these standards fail to be valuable and meaningful. But importantly, Star and Bowker emphasise that although classification schemes like the ICD might be pragmatic in nature – adaptable to an ever-changing field of medical knowledge – they still exert remarkable power over those working with classifiers in their daily practice (Bowker and Star, 1999: 69).

What sets digital epidemiology and its deep infrastructures apart from the ICD is the reliance on an ever-increasing realm of data sources. Capitalising on one significant feature of contemporary big data, it embraces the flow of data across sites, domains and across ethical and legal boundaries. To infer a diagnosis of Parkinson’s disease from the movement patterns of a mouse connected to a computer, the collected data needs to be repurposed to be interoperable in and outside of the realm of medicine (Geiger and Gross, 2021). With digital epidemiology, a classification scheme is proposed that can supposedly infer valuable diagnostics and deep causal inferences from an open-ended realm of data, which has not been collected for, nor designed for the purpose of medical diagnostics.

In a recent commentary, Milne and Costa (2020) argued that COVID-19 has offered multiple opportunities across different domains to realise such digital health futures. They cite enthusiastic CEOs, point to the regulatory flexibility offered by governments, and discuss the ubiquitous language of technological futurism with which digital resolutions are proposed to overcome the pandemic’s challenges. The COVID-19 pandemic is truly the first pandemic emerging within a ‘datafied society’ (Di Salvo and Milan, 2020), and a rhetoric of technological, data-driven fixing has left an imprint on almost every aspect of the international public health response. ‘The openings for digital health created by coronavirus’, Milne and Costa argue, ‘become propitious and productive moments for novel, data-driven diagnostic practices of the future’ (Milne and Costa, 2020). Or, to use the terms from Isin and Ruppert in their recent contribution, the pandemic has mobilised data practices and has made a new form of power based on sensor technologies ‘visible and articulable’, which had been gestating long before (Isin and Ruppert, 2020).

With COVID-19 emerged a vision of real-time epidemiology, or ‘nowcasting’ (Engelmann, 2020; Preis and Moat, 2014). Making visible how diseases are distributed through populations was no longer a historical science, reliant on data from the past, but one that was apparently able to deliver data for policy without any delay. The term ‘nowcasting’ was originally used in meteorology and had found wide application in economics (BañBura et al.,

2011). Equipped with a range of data on various indicators from the very recent past, nowcasting promises predictions of the present. Not to be mistaken with real-time surveillance, it is a methodological framework, where large-scale data collection is combined with domain-adapted modelling to develop, for example, representations of the current state of the climate, the current health of the economy or the present of an ongoing epidemic. Overcoming the common delays in case reporting and in the transnational standardisation of morbidity and mortality figures, nowcasting offers an estimate on current numbers on the basis of recent counts, smoothed to incorporate the specific dynamics of the disease (McGough et al., 2019).

In the first few months of the emerging pandemic, *Wired Magazine* ran a story on AI epidemiologists, that had supposedly determined the significance of the Wuhan outbreak ahead of national and international institutions, such as the WHO and the Centers for Disease Control (CDC) in the US. BlueDot, a Canadian AI company, had developed a system correlating novel outbreaks with commercial air travel booking data to estimate levels of transnational risk (Bogoch et al., 2020). Further analysis of travel data, this time localised to Chinese travel statistics offered, according to Moritz Kraemer et al., a ‘precise record of the spread of SARS-Cov-2 among the cities of China at the start of 2020’ (Kraemer et al., 2020). Social media data, particularly Twitter, was harnessed at times to predict global spread patterns of the disease (Bisanzio et al., 2020). Existing digital epidemiology infrastructures were adapted, as in the case of FluNearYou (<http://flunearyou.org>), to prompt patients to report on their COVID-19 symptoms for CovidNearYou (<http://covidnearyou.org>) and data from over 4.6 million members had been fed into the COVID Symptom Study to understand the distribution of COVID-19 symptoms across populations in the US and the UK (Sudre et al., 2021).

Without a doubt, epidemiology has been essential to addressing the pandemic’s multiple challenges, and according to an editorial in *Nature*, epidemiology has newly proliferated as a science over the course of 2020 (Editors, 2021). However, while dashboards might have been a new instrument to nowcast the pandemic’s everyday dynamics, they commonly relied on data collected, cleared and sorted in traditional ways through national and international institutions and health care providers. Neither the visibility of COVID-19, nor a substantial section of interventions mounted against the virus, have been informed by or were embedded within innovative systems that utilise repurposed data from the global information-exhaust of the internet. Rather, the pandemic offered a field of experimentation and calibration to bolster expectations invested in digital epidemiology. With COVID-19 understood as a ‘common environmental factor’ (Jagesar et al., 2021), the crisis offered opportunities to test and adjust digital research methods in outbreak detection or to measure the mental health fallout for

example (see for example Beukenhorst et al., 2021; Hsu et al., 2020; Mohr et al., 2020; Montag et al., 2020). To this field, COVID-19 has been an opportune moment to validate digital epidemiology as a trusted and reliable instrument of knowledge production.

Digital epidemiology – sometimes called infodemiology or computational epidemiology – originally emerged publicly with the spectacular demonstration of Google flu (2008–2014), a prototype study that successfully tracked the geographical distribution of flu through Google search trends. With the introduction of filtering and light touch modelling, Google’s engineers could arrive at the same geographic patterns of flu distribution as the CDC. However, the CDC’s numbers derived from doctors’ diagnoses, corroborated in laboratories and national clearing centres and were available with a standardised three-week delay. Google could provide flu tracking almost instantly and at a much lower cost. Nonetheless, Google Flu closed down as it lacked in accuracy what it offered in efficiency (Lazer et al., 2014). Despite the failure – or perhaps due to the interest it sparked – Google’s experiment encouraged the gestation of a veritable research landscape.

Digital epidemiology offers a radical expansion of the methods as well as the objects of epidemiological research. Tinkering with digital traces enables researchers to study new dimensions of infectious diseases, chronic conditions as well as behavioural and social phenomena, such as ‘anti-vaccination sentiment’ (Salathé et al., 2012). Since Marcel Salathé’s proposition, many claim to have gathered accurate inferences about a disease from what populations do online. Circumventing traditional gold standards of clinical diagnosis and laboratory testing, digital epidemiologists have turned to ‘discovering foodborne illness in online restaurant reviews’ (Effland et al., 2018), fighting dengue fever through mobile apps (Ahmed, 2013), mapping vaccine coverage for measles (Althaus and Salathé, 2015), identifying adverse effects of HIV drug treatment (Adrover et al., 2014, 2015), and detecting Parkinson’s disease through mouse movements (White et al., 2018) among many others (Park et al., 2018). But perhaps most striking – and concerning – is the rapid growth of a digital epidemiology of mental health: depression (De Choudhury et al., 2013), PTSD (Coppersmith et al., 2015) and schizophrenia (Hswen et al., 2018) have each become a prominent focus for automated digital diagnosis (see also Wacker, this issue).

Medical records, laboratory test results and administrative details have been the backbone of epidemiological studies, often complemented with auxiliary data such as demographics, zip codes and occupation information. The rise of digital data resources has changed the conditions for epidemiological research in at least two major ways. First, with the rise of artificial intelligence in medical research and new data-infrastructures in biological knowledge production, it has become easier to aggregate and use data from a wide range of sources and areas in medicine. Within the digital

ecosystem of healthcare provision, the entirety of significant information to capture, survey, diagnose and to intervene in the spread of disease is thought to be readily available in digital form. Electronic healthcare records, digital infrastructures in hospitals as well as digital governance offer opportunities to unleash visions of automated reporting systems, efficient surveillance methods and seamless healthcare operations fuelled by digitisation (Garrety et al., 2014; McFall, 2015; Lupton, 2017; Topol, 2019a).

Second, and perhaps more importantly, digital epidemiologists commonly seek to bypass the information delivered through healthcare systems (Brownstein et al., 2009). The traditional process, whereby data travels from the patient experiencing a symptom to a doctor’s visit, from initial clinical diagnosis to confirmative laboratory analysis, and from sentinel laboratories to national and international clearing houses, is no longer perceived as exclusive data infrastructure. Where such institutional pathways had been understood to deliver robust data quality, to offer protection of confidentiality and to adhere to democratic regulations, they are increasingly perceived as cumbersome, pedestrian and inefficient. Instead, epidemiologists have taken an interest in new types of data. Such data might have been voluntarily generated by patients who bypass health institutions (Prainsack, 2017), who share their queries while consulting online resources or contribute with self-reported data to participant-led-research projects (e.g. PatientsLikeMe). However, many of the studies listed above rely on data, which was never produced with the intent to contribute to medical or epidemiological research.

This data can be syphoned from search terms, as pioneered by Google flu trends, but also from a range of other sources in the global data deluge. Expressions of users on social media, the geographical spread of access logs on specific Wikipedia pages, gathering of movement patterns through GPS trackers, rhythms of smartphone usage, patterns of interactions with smart input and output devices, data from games and loyalty cards or even economic indicators from banking apps constitute some of the important sites for passive medical data collection. These methods are accompanied, often complemented, by active data production. Fitness trackers, smartwatches and digital glucose monitors are just some of the devices with which vital information is aggregated with potential use for epidemiological analysis. Together, these growing data streams appear to offer astonishing opportunities to infer pathological states. Significantly, they also create prospects of diagnosing and measuring the spread of disease without the knowing participation of the data producer or the data owner (Vayena et al., 2015; Mittelstadt et al., 2018; Prainsack, 2019).

Digital epidemiology seeks to capitalise on growing availability and interoperability of medical data, while also offering a conceptual framework to infer medical meaning and casual statements from the growing variability and characteristic inconsistency of big data. Consistent with

a shift in attention from “‘controlled epidemiological calculus’ to routinized practices of ‘epidemic indexing’”, as Bauer argues, this digital reinvention of epidemiology unleashes an appetite for unlimited data sources, that raises ever-more pertinent questions about the subject matter of epidemiology as much as it raises a question about a research practice seemingly out of bounds (Bauer, 2019). Deep phenotyping lies at the heart of this digital innovation of epidemiological reasoning, both as a technical fix to navigate the data deluge as well as a conceptual framework to reshape the classification of disease.

Rediscovering disease

What are the visions invested in phenotyping a disease? Phenotyping remains an odd choice of terminology to describe the novel set of research methods, mining tools and scraping technologies, which are supposed to radically redefine how diseases are observed and how their appearance is classified. With the phenotype traditionally conceived as the physical expression of one or more genes, the term usually refers to the sum of observable characteristics of an organism that are produced by the interactions between genotype and environment. Phenotyping would, within the same line of genetic thinking, be understood as a process of determining, describing and analysing the expressed traits of an organism’s genetic information (Loi, 2018). The researchers, who have since 2015 turned to the language of digital phenotyping to describe the project of newly structured observations of the expression of disease, use this terminology slightly ambiguously, if not metaphorically. With reference to Reardon, all of the practices and visions assembled under the umbrella of digital phenotyping certainly share a commitment to genetics. But intriguingly, this is not a project directly concerned with the technological capacities of genetics to identify traits and patterns in sequenced DNA data, but one that seeks ‘to render the human genome meaningful’ (Reardon, 2017: 14). Faced with a growing data deluge and within a global process of eroding trust in dominant medical institutions and standards, proponents of digital phenotyping seek to replace what they perceive as flawed human observation with structured data, while capitalising on the digital infrastructures with which meaning of life is negotiated in what Reardon has framed as the ‘postgenomic condition’ (Reardon, 2017).

In a 2015 commentary piece, the digital phenotype was defined with reference to Richard Dawkins’ proposition of an ‘extended phenotype’ (Jain et al., 2015). A phenotype, so the authors claim, should not be limited to biological processes. With reference to Dawkins, it should be acknowledged that humans, like animals, modify their environments. As humans spend their lives increasingly online, the authors ask if ‘aspects of our interface with technology [can] be somehow diagnostic and/or prognostic for

certain conditions’ (Jain et al., 2015). While the integration of digital technologies and the repurposing of user data are not novel in biomedical research, the commentary seeks to establish the digital phenotype within the budding landscape of precision medicine. Importantly, the exploitation of the digital phenotype does, so the authors argue, extend far beyond the purpose of ‘surveillance and early detection’. The appropriate analysis and careful integration of data from social media, forums, wearable technologies and mobile devices could ‘fundamentally alter our notion of the manifestation of disease’ to offer a ‘more comprehensive and nuanced view of the experience of illness’ and to continuously measure ‘manifestations of biological disease’ (Jain et al., 2015: 462).

The psychiatrists and biostatisticians John Torous and Jukka-Pekka Onnela have developed a slightly different approach. Onnela established the Beiwe research platform for digital phenotyping at the Harvard T.H. Chan School of Public Health in 2013. They focused on the impact of novel sensor technology and on the pervasive distribution of affordable and expandable measuring systems in smartphones. The impact of the smartphone and its potential to radically transform biomedical and clinical insight is in their view only comparable with the drastic transformations associated historically with microscopes in the medical sciences (Onnela and Rauch, 2016; Torous et al., 2016). Emphasising the epistemological significance of innovation in measurements and sensor technology, they suggest that digital phenotyping captures for the first time ‘moment-by-moment quantification of the individual-level human phenotype in-situ using data from smartphones and other personal digital devices’ (Torous et al., 2016). Their emphasis lies apparently on the re-evaluation of human phenotypes through the precise measurement of ‘various types of social and behavioural data that capture the subjects’ lived experiences and their interactions with people and places’ (Torous et al., 2016). Their platform is based on a smartphone architecture to collect spatial trajectories via GPS, physical movement patterns via accelerometer and audio samples via the device’s microphones. To Onnela and his colleagues, digital phenotyping did not conceptually derive from Dawkins’ extended phenotype and associated assumptions about humans’ co-configured behaviour in digital worlds. Instead, their approach to digital phenotyping is firmly located within the enterprise of precision medicine, seeking to advance the practice of *deep* phenotyping into digital ecologies.

With uncertain origins, the concept of deep phenotyping assumed popularity in the early 2010s to question the quality and granularity of the vast majority of described clinical pictures and to throw into question if physicians’ observations should continue to provide the baseline for disease classification. Already in 2008, the biochemist R.P. Tracy argued in a systematic review that genetics’ reliance on the current views of disease and the fixation on

clinical outcomes diminishes chances for establishing genotype-phenotype relations. The observable traits of diseases visible to clinicians and measurable in laboratories represent, so Tracy writes, only a small fraction of those that might actually be associated with an underlying genetic cause. Enhancing the results of genome-wide association studies to identify relevant correlations required a drastic expansion in the scope of observable traits of a disease. To do so, phenotyping would have to move substantially beyond the observations collected in the clinical encounter, so as to grasp, structure and analyse the ‘underlying pathophysiology across an individual’s life history’ (Tracy, 2008).

In 2012, *Human Mutation* published a special issue on deep phenotyping to gather current research and to further consolidate the practice’s position at the heart of precision medicine. In the opening piece, human geneticist Peter Robinson asked if the research community was ‘ready for a human phenome project?’ His paper, in line with the entire issue, proposed ways to overcome the ‘sloppy or imprecise ways’ with which physicians have so far described the appearances of a disease. Deep phenotyping instead offers ‘precise and comprehensive analysis of phenotypic abnormalities’ (Robinson, 2012: 770). Robinson hoped to better classify subpopulations at risk of rare diseases, to achieve clearer stratification of groups of patients who share a common biological basis for disease and to develop entirely new pathways of capturing patients’ responses to treatments. With a constantly growing body of genetic data, the systematic observation of expressions of what he calls the ‘morbid anatomy of the human genome’ (Robinson, 2012: 778) needs to keep up.

However, for Robinson and his fellow contributors to the *Human Mutation* special issue, the precise contours of the depth of their phenotyping remained vague. They developed an expanded and newly structured clinical assessment to refine descriptors and to expand diagnostic criteria. Not only the diagnosis of breast cancer might be important to understand the underlying mechanism of mutations like BRCA1 or BRCA2, but it might be of immense value to capture, for example, the variable treatment response in the individual’s phenotype. Further, deep phenotyping would include lateral standardisation of different clinical datasets to increase interoperability and perhaps the design of a globally standardised human phenotype ontology, (Robinson and Mundlos, 2010) which was supposed to ‘alter the way many fields of medicine are practiced’ (Robinson, 2012: 779).

‘Precision medicine requires an understanding of the precise relationship between gene and phenotype, and the stratification of diseases into subtypes according to their underlying biological mechanisms’, Cathryn Delude argues in *Nature* in 2015 (Delude, 2015). Within clinical medicine, so she states, these relationships cannot be established due to the persisting variance in the definition of

disease phenotypes. To wring clinical value from genetic data, it will be required to carry out an ‘exhaustive examination of the discrete components of a phenotype that goes beyond what is typically recorded in medical charts’ (Delude, 2015). Diabetes is a common example cited in these debates, as its diagnosis is often considered incomplete and ignorant of the many clinical subtypes with which it presents. To capture genetic variations that might sensibly contribute to the diagnostic, treatment and prevention of diabetes, so Delude explains, the clinical picture will first have to be resolved into hundreds of subtypes, which in turn require contextualisation with detailed patient data and environmental variables.

Deep phenotyping thus belongs clearly to the concepts that have emerged in the decades since the completion of the human genome project, to make sense of and to attribute meaning to a constantly growing archive of data pertinent to human life. If one of the principal promises of genetic research was to find gene variations directly responsible for the development of specific diseases (Fortun, 2008), success has thus far been sparse. As many authors and more recently Reardon pointed out, the Common Disease–Common Variant (CD-CV) hypothesis might have had a lot of purchase in the ramping up of the hype around genetic research, ‘yet, today few such variants have been found’ (Reardon, 2017: 1). Many biologists and geneticists went on to attribute the lack of immediate success not to false hypotheses or fundamental misconceptions about the role of genetic variations in the development of specific diseases, but to the persistence of dated descriptions, vague classifications and limited observations in the analysis and classification of diseases. To deliver on the promise of precision and personalised medicine and to match phenotypes to genes, the quality of observable traits needed to be improved.

In her book, Reardon offers an instructive reckoning with the impact of the human genome project. While the project has yielded to vast data sets and while its computational development offered an attractive technological model for follow-on research, it was the knowledge produced that raised more questions than answers. With the biological information of life accessible, ‘rendering it meaningful in the postgenomic era proved anything but easy’ (Reardon, 2017: 39). Researchers scrambled to develop new tools with which to salvage the initial therapeutic enthusiasm around the human genome (Landecker, 2011). Catalogues of rare variants, as for example in the HapMap project, were established with the help of more complex and more technologically intensive sequencing technologies, while others began to invest in private, market-driven endeavours to improve statistical power in the search for significant variants in companies like 23andme (Reardon, 2017: 123). Deep phenotyping, while lacking in Reardon’s book, should be understood within the same realm: a technologically driven attempt to

salvage the value and meaning of the information deluge and data infrastructures provided by the human genome project. It was yet another approach, driven by the perspective of engineering, to clear up the noise in the signal between genotype and phenotype.

The medical tech-enthusiast Eric Topol's vision for the future of 'deep medicine' implies a role for artificial intelligence to scrape the growing data deluge from medical institutions as well as novel data sources, while physicians might win back time to engage in empathic relations with their patients. While he introduces deep learning with an eye on the caveats and ethical issues – 'deep liabilities' – his description of digitally enhanced deep phenotyping is relentlessly enthusiastic. Medical specialism, limited time commitments and the brief encounters between patients and doctors are rendered remnants of what Topol calls a 'shallow medicine' (Topol, 2019b: 33). Doctors working on psychiatric wards are unlikely to offer expertise on diabetes and the pharmacologist might have only limited understanding of the genetics, which might undermine the efficiency of a prescribed drug. But more importantly, Topol laments, doctors encounter patients only in momentary settings, relying on vague descriptions, brief observations and dated classifications to arrive at diagnostics based on isolated and limited data. Deep phenotyping instead offers not only depth through multiple layering of information gathered from a range of sources, specialities and tests, but it also expands the clinical encounter over time, offering 'long' data, by 'covering as much of our lives as we can, because many metrics of potential interest are dynamic, constantly changing over time' (Topol, 2019b: 16). The expansiveness of the data collection is essential to the project of deep phenotyping in Topol's vision. It encapsulates an approach characterised by limitless depth, whereby scraping of potentially clinically valuable information is 'spanning as many types of data as you can imagine' (Topol, 2019b: 16), offering novel metrics that are collected passively, unobtrusively and permanently (Topol, 2019b: 174).

In Topol's view, vexing conditions like depression could be finally measured appropriately and their classification would no longer rely on the patient's account or on stubborn diagnostic inventories. Instead, a smartphone would be able to collect data from speech patterns, would capture shifting intonations of voice, map the reaction times from keyboard users, gather movement patterns, while tracking social media use and applying visual recognition features to selfies to search for distinctive facial attributes indicative for what would become an entirely new classification of depression (Topol, 2019b: 173). Precisely where animal models failed to replicate the specific circumstances of human psychiatric conditions such as schizophrenia, deep phenotyping is supposed to link clinical features to individual genomes for the first time. In psychiatry, the hope is to overcome what is often perceived as particularly

vague descriptions of the phenotypes of many conditions. With deep phenotyping, the neuroscientist Steven Hyman seeks to describe phenotypes as newly 'mechanistic' ones, which would enable for the first time the development of robust causal chains between biology and psychology (quoted in Delude, 2015). However, here as well as in the realm of physiological disorders, the success of deep phenotyping has not been materialised. As of 2019, as the contributions to a recent issue in *ST & HV* report, 'not a single genetic or biological marker has been identified so far that is specific to any of the major psychiatric disorders' (Rüppel and Voigt, 2019: 570).

According to the growing circle of proponents of digital phenotyping, the answer to the lack of success remains more data. This is the point where Onnela and Rauch invoke the all-empowering metaphor of the 'smartphone-as-microscope' to encourage digital phenotyping as nothing less than a revolutionary transformation in psychiatric data collection (Onnela and Rauch, 2016). Where much of *deep* phenotyping had been overly reliant on medical information, concerned with making data interoperable and multi-layered, *digital* phenotyping yields to a wealth of untapped resources that will expand the depth and length of clinically viable data collection. The explicit aim is to discover patterns, characteristics and signs, where the deep phenotyping of biological specimens, genetic data and available clinical signs has failed to offer meaningful linkage. Scraping information from infinite data is to resolve the lack of linkage of causal chains in psychiatry. In this vision, digital phenotyping has assumed utopian proportions, imagined as a technological innovation towards omniscient observation, with which a complete archive of human ailment might be abducted from an infinite wealth of data.

To seize on the potential of these technologies for epidemiological knowledge production, some developers dare to dream big. The MDs Abnoui, Rumsfeld and Krumholz proposed in a viewpoint contribution to the *Journal of the American Medical Association* a digital reconsideration of the social determinants of health in the digital age. Comparable to the expectations raised around deep phenotyping in mental health, they complain of a lack of matching progress when it comes to the 'nurture components' of disease. Social determinants, they concur, are notoriously complex to measure, as they entail 'networks and behaviour that are best revealed by what actually occurs in life' (Abnoui et al., 2018: E1). Neither patient reporting in standardised surveys, nor the sociological study of patient groups, ever manages to grasp the entirety of the data points that might elucidate drivers of disease within social and cultural contexts. However, the increasing prevalence of online social networks offers to these physicians a new window into the opaque world of social determinants. The data made available through social media offers 'measurable, actionable insights' to understand disease, but more importantly, would transform the

understanding of social determinants at large. New social biomarkers, such as ‘timing, frequency, content and patterns of posts and degree of integration with online communities’ would supposedly provide an entirely new framework to expand understanding of health and disease ‘in a truly novel way’. This fantasy is all the more chilling when the authors’ affiliation with Facebook Inc. and IBM Watson Health underlines their proposal to utilise these novel measurements to rewrite the ‘source code for nurture’ (Abnoui et al., 2018).

Discussion: The depth of epidemiological reasoning

Thomas McKeown lamented in 1983 that despite revolutionary advances in medical technologies and a never-before-seen breadth of medical knowledge, countless diseases still elude successful treatment and prevention. To develop better health strategies, the physician, epidemiologist, and historian suggested that the problem does not lie with shortcomings of the laboratory or the clinic, but rather with the way in which determinants of health had been collected and classified. Too much attention had been given to hereditary conditions, which in his words were ‘determined at fertilisation’, while the large majority of diseases was in fact ‘not so determined and manifested only in an appropriate environment’ (McKeown, 1983: 595). Classifying and understanding the influence of poverty and environmental hazards would allow public action where it was needed, while hereditary information might perhaps be only of secondary value. McKeown assumed himself to witness nothing less, but a Kuhnian change of paradigm, with the old one already gone, while the shape of the new way of seeing disease was yet to be defined.

A similar, complimentary situation appears to present itself today, almost 40 years later. Still, the mismatch remains between expectations invested in technological progress and overcoming the global burden of disease. The hopes attached to sequencing of the human genome have largely not been fulfilled. Nonetheless, personalised and precision medicine continue to accrue value as visions of the future of medicine. In stark contrast to McKeown, the vision of data-driven classification appears to have shifted to the rationale of inheritance and predetermined conditions. Importantly, this classification does not only affect all of those conditions that might be allocated directly to genetic markers via a ‘smartphone-as-microscope’ (Onnela and Rauch, 2016). As Topol and the MDs at Facebook and IBM Watson explain, even environmental and social determinants are to be captured within the model and rationale of a ‘source code of nurture’ (Abnoui et al., 2018). Digital phenotyping, it appears, is asked to provide the causal depth for newly stabilised disease

classifications, offering the appearance of a dubious aetiology, while epidemiology provides the breadth and length of data required to arrive at meaningful comparisons.

The project and vision of a digital epidemiology runs, as I discuss here, on the assumptions that all there is to know about the underlying causes of diseases as well as anything that could be potentially understood to be a classifiable indicator of a disease, can be captured in structured and stratified data. Underneath the surface awaits a depth of pathological knowledge, which – perhaps counterintuitively – reveals itself through epidemiological data scraping along the surface of populations, through the aggregation of population-level comparisons and through the large-scale association of data mined from the captured interactions of humans with digital technology. This is what I call a problematic assumption of pathological omniscience that empowers this vision of a digital, all-seeing, all-digesting epidemiology.

For this discussion section, I will refer to one exemplary case of a digital phenotype in development. This case concerns a research team at Harvard’s Computational Epidemiology group. Here, the team investigates the possibility of inferring a digital phenotype of schizophrenia using social media data (Hswen et al., 2018). The case offers an interesting example, as it demonstrates the assumptions invested in digital phenotyping while it also allows for a critical discussion of associated data practices. For their paper, they use data provided by Twitter to analyse expressions of depression and anxiety among users who self-identified as schizophrenic. Overall, their findings disclose that there are consistently higher levels of expressions of depression and anxiety among these users, compared to a sample of the general population on Twitter.

There are three important concerns, with which such studies ought to grapple with and which currently remain a lacuna in the relevant literature. First, a classification system based on these premises posits a sharp departure from the underlying principles of widely accepted – although far from ideal – systems like the DSM and the ICD. Second, the exaggerated claim of depth implies the capacity of these systems to discover hidden, so far unseen symptom-syndrome relations, which supposedly remain invisible on the surface of ‘shallow’ medicine. Third, like many similar projects in the digital health field, this disruptive innovation envisions its own utility in an uncertain future. This digital future scoping falls usually short to address urgent, obvious and largely uncontroversial drivers of illness, which are associated with the social structures of poverty, marginalization and inequality in the present.

Unlike in the ICD’s nomenclature, which represents the efforts of generations of physicians and statisticians to align observations with classifications, the aim of deep phenotyping is to deliver precise relations between genetic information and (extended) phenotypical expression. This endeavour of structuring and matching pathological information departs from the dynamic scheme with which diseases

were principally collected as the observed experience of disease. Or, as Bowker and Star put it, ‘the nomenclature of diseases and of causes of death established for the needs of statistical organization constitutes a sort of contract between the two organizations who are charged with statistical works – that is to say, the service who makes the observations and that which produces statistics with the help of these data’ (Bowker and Star, 1999: 146). Departing from this contract, the question of observation is entirely taken out of the physicians’ hands and handed over to passive monitoring of what can possibly be re-purposed into a vital sign. With medical classification historically perceived as a shared benefit between medical practice and medical research, one might ask, who precisely the benefactors of deep phenotyping would be and if physicians or epidemiologists are willing to ever accept a rule-based classification system, whose design and inputs appear to lie beyond their control.

Hswen and colleagues argue carefully in their paper that their point of departure is not a clinical diagnosis of schizophrenia but that their exploratory study relies on the information that Twitter users had put in the public sphere. Such users, the authors find in this study, ‘may express elevated symptoms of depression and anxiety in their online posts’ (Hswen et al., 2018). As the author’s stress, this finding confirms and corroborates a well-known symptom-syndrome relation, as established in clinical literature. Here, this confirmatory result thus appears as a starting point, as an establishment of a baseline of sorts to demonstrate the usability of digital epidemiology in the ‘understanding of schizophrenia by informing a digital phenotype’ (Hswen et al., 2018). This research strategy, aimed to refine the digital phenotype on the basis of people who have made their schizophrenic diagnosis public knowledge, implies a paradoxical relation to existing classification systems. While a clinical classification, such as that could be found in the ICD, is not the explicit starting point, it remains the implicit reference in the construction of a digital phenotype. Somewhat surprisingly, this endeavour falls short of radically reinventing symptom-syndrome relations but seems to solidify a historically contingent category of mental illness as generalizable information.

The history of the ICD is one structured by extensive negotiations about structures of classification and the categories in the ICD have always been a reflection of historical circumstances (Bowker and Star, 1999). Phenotyping, not just those conditions that might or might not offer a causal relation to genetic information, but to apply a genetic blueprint to re-classify the details of all diseases seeks to replace a historically grown dynamic and pragmatic system with a rigid scheme. Deploying and establishing such a scheme risks implying the persistence of a world, where each symptom can be revealed as the outcome of a discrete and generalizable causal chain, indifferent to and unaffected by the passing of time as much as by the conditions of place.

The projects and approaches assembled around the production of deep phenotypes share a curious fixation on the metaphor of depth. While such depth is everywhere implied, it appears to be almost impossible to find consistent indicators of how exactly this depth is imagined and conceptualized across this research landscape. Investigating depth suggests an excavation of something underneath the surface, a hidden and opaque reality that is perhaps shrouded by layers of theories, assumptions and beliefs that are associated with ‘shallow medicine’. Topol assumes depth where the doctor’s gaze extends into thickness, covering multiple layers of specialties and including the entirety of collectible indicators across the full life of an individual. Delude understands depth as granularity and increasing individuality in the data collected, but importantly, she also emphasises that such depth emerges out of ‘sophisticated algorithms’ which ‘integrate the resulting wealth of data with other kinds of information’ (Delude, 2015: S15). Depth in knowledge, it appears, is the result of accumulated observations, correlated across different domains and equipped with statistical power, before returned to the individual, repurposed and repackaged as precision or personalised medical expertise (Bauer, 2008; Hoeyer et al., 2019).

Returning to the example of the digital phenotype of schizophrenia, it is quite obvious that depth derives here from the power of statistical inference. Data from self-identified users with schizophrenia have been compared with data from the general population of Twitter users resulting in *P* values mostly below 0.006 (Hswen et al., 2018). The depth of the accuracy of elevated levels of expressions of anxiety and depression clearly does not derive from the excavation of complex genotype-phenotype relations, nor from the speculation about different underlying causes. Here, depth directly results from the length of the collected sample.

The second pillar of deep phenotyping can be found, as Topol and others have argued, in deep learning. The depth in deep learning, Taylor Arnold and Lauren Tilton explain, derives partly from the capacity of models to be ‘knowledgeable’ and accurately predict semantic variables in unstructured texts or images. ‘In other words, their ability to build off of existing knowledge to predict new knowledge’ (Arnold and Tilton, 2020: 310). Significantly, with the application of deep learning, disease is configured as a ‘deep problem’ and thus framed as a problem of communication and representation. In other words, a deep learning approach would assume that the true shape of a disease category like schizophrenia is already present among circulating information – in conventional diagnostics as well as in the wealth of big data – but requires the detour of multi-layered learning to achieve its desired depth and truth.

The example here does not rely on deep learning, but nonetheless utilises simple statistics to infer diagnostic criteria. However, one may ask to what extent schizophrenia constitutes a ‘deep problem’ to the study’s authors. A

deep problem is characterised by its already-present representation in existing information, but it is assumed that its nature might remain shrouded due to human error, cultural variations or false categorization. Positing schizophrenia as a deep problem, however, assumes the valance of precisely the diagnostic criteria that deep phenotyping is supposed to problematize. In other words, to research the deep phenotype of schizophrenia with a deep learning platform might yield to yet unknown aspects of experiences of those assumed to suffer from schizophrenia. However, due to the inherent confirmation bias of multi-layered systems, only those results already confirming central tenets of the already established nomenclature of schizophrenia are considered valuable. In effect, deep phenotyping risks reinstating historically controversial diagnostic criteria as elements of a new, data-driven generalised classification scheme.

The presence of such circular arguments – schizophrenia is characterised by anxiety and depression, therefore the presence of anxiety and depression is indicative of schizophrenia – leads to a third dimension, which Hoeyer has recently unpacked for personalised medicine initiatives by the Danish state. Through the metaphor of depth, population data becomes routinely enrolled in the medical approach to the individual patient in personalised medicine. Hoeyer describes for the case of Denmark a dynamic of ‘future accountability’ which configures the value of collected and ‘intensified’ population data as ‘promissory data’ (Hoeyer, 2019: 532). Its value for the individual patient within a personalised medicine framework can only be realised at a future stage. This requires on the one hand a buy-in from stakeholders into the future stakes of large-scale data-collections – like Danish public health authorities in Hoeyer’s example – but ongoing and resource intensive data-collection also serves as a vehicle to ‘avoid action’ (Hoeyer, 2019: 533) in the present. It is striking how the vast majority of contributions to a digital epidemiology driven by deep phenotyping fall within the same realm – including the schizophrenia study discussed here. While pointing to a future of data-intensive deep medicine with potential use for public health as well as personalised medical care, their research and their methodologies continue to withdraw from simpler, perhaps shallower problems, such as how to tackle the implications of poverty, inequality or underfunded health systems as towering and overwhelming determinants of health in the present.

Conclusion: Towards pathological omniscience?

Epidemiology, I have argued, has historically grappled with the division of observation and analysis, while identifying often as a mere set of inconspicuous methods. However, epidemiology is a powerful research practice, in which the foundations of medical knowledge are negotiated, adapted and transformed in the engagement with large

data sets. How disease is counted, what kind of disease is assumed countable and how causes for disease are accounted for; these were and are fundamental questions in epidemiological research and have extended impact on the ways in which disease was and is classified and perceived elsewhere. This epistemic dimension of epidemiological reasoning, however, is not new and cannot be attributed to the field’s digital transformation. The novel depth of knowledge in digital epidemiology, however, emerges with a puzzling array of naïve disruptions aimed at ‘shallow medicine’. Depth, I have argued, is not just a metaphor for granularity, fine-grained detail or improved accuracy. Depth is on the one hand envisioned as the result of an assumed invariance of causal chains, offered by the data-model of genotype-phenotype relations and fixed in the engineering rhetoric of signal and noise. On the other hand, this depth is imagined to come into being with a radical expansion of epidemiological gathering, scraping and mining data, fuelled by the erratic assumption that all data traces of human behaviour can and should potentially be re-purposed into medical information. With these imaginations, disease does indeed become framed as a deep problem, always attached to a discrete cause and always measurable as a non-ambiguous entity within the data traces of human behaviour, however, shrouded by misrepresentation and noise. Schizophrenia, as the example discussion disclosed, is no longer assumed a variable category of classification with dramatic historical inconsistency, but the disease is reimagined as a deep problem. The solution of a deep problem lies, however, not in the discovery of a deeper truth – whatever that may be – but in the recalculation of existing knowledge to predict new knowledge. Examples, such as this schizophrenia study show that deep phenotyping can thus run the risk of reifying and cementing contingent and variable categories into apparently persistent classifiers. In the resulting absence of any causal depth, it remains then an open but important question, if the depth of deep phenotyping is anything more than a shallow rhetorical device in the promissory language of future medicine.

This pathological omniscience is short-sighted and remains – perhaps willfully – ignorant of a wide area of medical and epidemiological thinking. To think of disease as merely statistical entities does away with the complex and intricate relations between disease observation, classification and treatment. To imply a pervasive presence of genotype-phenotype relations in the determination of disease continues to dramatically over-estimate the significance and indeed the social burden of pre-determined conditions among human ailments. Most significantly, this new discovery of the depth of disease is rolled out with dramatic inconsistency and relies implicitly on precisely those categories and classification schemes, it purports to disrupt and to dissolve. In effect, a digital epidemiology powered by deep phenotyping runs the risk to reinstate pragmatic

and historically contingent systems of classifications as a universal and all-powerful source code.

Acknowledgements

I like to thank all participants across various workshops and conferences, who have helped sharpening the arguments presented in this paper. I am as ever grateful to Kate Womersley for her impeccable and insightful review, to my colleagues at STIS for productive feedback on early drafts and like to extend my gratitude to the reviewers for their excellent comments and suggestions.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Research for this article has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 947872).

ORCID iD

Lukas Engelmann  <https://orcid.org/0000-0002-2175-0156>

References

- Abnoui F, Rumsfeld JS and Krumholz HM (2018) Social determinants of health in the digital age: Determining the source code for nurture. *Journal of the American Medical Association* 321(3): 247–248. <https://doi.org/10.1001/jama.2018.19763>.
- Adrover C, Bodnar T, Huang Z, et al. (2015) Identifying adverse effects of HIV drug treatment and associated sentiments using Twitter. *JMIR Public Health and Surveillance* 1(2): E7.
- Adrover C, Bodnar T and Salathé M (2014) Targeting HIV-related medication side effects and sentiment using Twitter data. *ArXiv Preprint ArXiv:1404.3610*.
- Ahmed B (2013) How smartphones became vital tools against dengue in Pakistan. NPR.Org. 2013. <https://www.npr.org/sections/health-shots/2013/09/16/223051694/how-smartphones-became-vital-tool-against-dengue-in-pakistan>.
- Althaus CL and Salathé M (2015) Measles vaccination coverage and cases among vaccinated persons. *Emerging Infectious Diseases* 21(8): 1480.
- Amelang K and Bauer S (2019) Following the algorithm: How epidemiological risk-scores do accountability. *Social Studies of Science* 49(4): 476–502.
- Arnold T and Tilton L (2020) Depth in deep learning: Knowledgeable, layered, impenetrable. In: *Deep Mediations: Thinking Space in Cinema and Digital Cultures*. Minneapolis: University of Minnesota Press, pp. 309–328. <https://doi.org/10.5749/j.ctv1gn3t58>.
- BañBura M, Giannone D and Reichlin L (2011) Nowcasting. The Oxford Handbook of economic forecasting. 8 July 2011. <https://doi.org/10.1093/oxfordhb/9780195398649.013.0008>.
- Bauer S (2008) Mining data, gathering variables and recombining information: The flexible architecture of epidemiological studies. *Studies in History and Philosophy of Biological & Biomedical Science* 39(4): 415–428.
- Bauer S (2019) Indexing, coding, scoring: The engine room of epidemiology and its routinized techno-digestions. *Somatechnics* 9(2–3): 223–243.
- Beukenhorst AL, Collins E, Burke KM, et al. (2021) Smartphone data during the COVID-19 pandemic can quantify behavioral changes in people with ALS. *Muscle & Nerve* 63(2): 258–262.
- Bisanzio D, Kraemer MUG, Bogoch II, et al. (2020) Use of Twitter social media activity as a proxy for human mobility to predict the spatiotemporal spread of COVID-19 at global scale. *Geospatial Health* 15(1): 19–24. <https://doi.org/10.4081/gh.2020.882>.
- Bogoch II, Watts A, Thomas-Bachli A, et al. (2020) Pneumonia of unknown aetiology in Wuhan, China: Potential for international spread via commercial air travel. *Journal of Travel Medicine* 27(2): 1–3. <https://doi.org/10.1093/jtm/taaa008>.
- Borup M, Brown N, Konrad K, et al. (2006) The sociology of expectations in science and technology. *Technology Analysis & Strategic Management* 18(3–4): 285–298.
- Bowker GC and Star SL (1999) *Sorting Things out. Classification and Its Consequences. Inside Technology*. Cambridge, MA: MIT Press.
- boyd d and Crawford K (2012) Critical questions for big data. *Information, Communication & Society* 15(5): 662–679.
- Brownstein JS, Freifeld CC and Madoff LC (2009) Digital disease detection — harnessing the web for public health surveillance. *New England Journal of Medicine* 360(21): 2153–2157.
- Coppersmith G, Dredze M, Harman C, et al. (2015) CLPsych 2015 shared task: Depression and PTSD on Twitter. In: *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. Denver, Colorado: Association for Computational Linguistics, pp.31–39. <https://doi.org/10.3115/v1/W15-1204>.
- De Choudhury M, Counts S and Horvitz E (2013) Social Media As a measurement tool of depression in populations. *Proceedings of the 5th Annual ACM Web Science Conference*. WebSci'13. New York, NY, USA: ACM, pp.47–56. <https://doi.org/10.1145/2464464.2464480>.
- Delude CM (2015) Deep phenotyping: The details of disease. *Nature* 527(November): S14–S15.
- Desrosieres A (1998) *The Politics of Large Numbers. A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
- Di Salvo P and Milan S (2020) The four invisible enemies in the first pandemic of the “Datafied Society”. SSRN Scholarly Paper ID 3622931. Rochester, NY: Social Science Research Network. <https://doi.org/10.2139/ssrn.3622931>.
- Editors (2021) How epidemiology has shaped the COVID pandemic. *Nature* 589(7843): 491–492. <https://doi.org/10.1038/d41586-021-00183-z>.
- Effland T, Lawson A, Balter S, et al. (2018) Discovering food-borne illness in online restaurant reviews. *Journal of the American Medical Informatics Association* 25(12): 1586–1592. <https://doi.org/10.1093/jamia/ocx093>.
- Engelmann L (2020) #COVID19: The spectacle of real-time surveillance. *Somatosphere* (blog). 6 March 2020. <http://somatosphere.net/forumpost/covid19-spectacle-surveillance/>.
- Fortun Michael. 2008. *Promising Genomics: Iceland and DeCODE Genetics in a World of Speculation*. Berkeley, CA: Univ of California Press.

- Garrety K, McLoughlin I, Wilson R, et al. (2014) National electronic health records and the digital disruption of moral orders. *Social Science & Medicine* 101: 70–77.
- Geiger S and Gross N (2021) A tidal wave of inevitable data? Assetization in the consumer genomics testing industry. *Business & Society* 60(3): 614–649.
- Hoeyer K (2019) Data as promise: Reconfiguring Danish public health through personalized medicine. *Social Studies of Science* 49(4): 531–555.
- Hoeyer K, Bauer S and Pickersgill M (2019) Datafication and accountability in public health: Introduction to a special issue. *Social Studies of Science* 49(4): 459–475.
- Hsu M, Ahern DK and Suzuki J (2020) Digital phenotyping to enhance substance use treatment during the COVID-19 pandemic. *JMIR Mental Health* 7(10): e21814.
- Hswen Y, Naslund JA, Brownstein JS, et al. (2018) Online communication about depression and anxiety among Twitter users with schizophrenia: Preliminary findings to inform a digital phenotype using social media. *Psychiatric Quarterly* 89: 569–580. <https://doi.org/10.1007/s1126-017-9559-y>.
- Isin E and Ruppert E (2020) The birth of sensory power: How a pandemic made it visible? *Big Data & Society* 7(2): 2053951720969208.
- Jagesar RR, Roozen MC, van der Heijden I, et al. (2021) Digital phenotyping and the COVID-19 pandemic: Capturing behavioral change in patients with psychiatric disorders. *European Neuropsychopharmacology* 42: 115–120.
- Jain SH, Powers BW, Hawkins JB, et al. (2015) ‘The digital phenotype’. *Nature Biotechnology* 33: 462–463. <https://doi.org/10.1038/nbt.3223>.
- Jardine B and Drage M (2018) The total archive: Data, subjectivity, universality. *History of the Human Sciences* 31(5): 3–22.
- Kitchin R (2014) Big data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): 2053951714528481.
- Kraemer MUG, Yang C-H, Gutierrez B, et al. (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science (New York, N.Y.)* 368(6490): 493–497.
- Landecker H (2011) Food as exposure: Nutritional epigenetics and the new metabolism. *BioSocieties* 6(2): 167–194.
- Lazer D, Kennedy R, King G, et al. (2014) The parable of Google flu: Traps in big data analysis. *Science (New York, N.Y.)* 343(6176): 1203–1205.
- Leonelli S and Tempini N, eds (2020) *Data Journeys in the Sciences*. X: Springer International Publishing. <https://doi.org/10.1007/978-3-030-37177-7>.
- Loi M (2018) The digital phenotype: A philosophical and ethical exploration. *Philosophy & Technology* 32: 155–171. <https://doi.org/10.1007/s13347-018-0319-1>.
- Lupton D (2017) *Digital Health: Critical and Cross-Disciplinary Perspectives*. New York: Routledge.
- Matthews JR (1995) *Quantification and the Quest for Medical Certainty*. Princeton, N.J.: Princeton University Press.
- McFall L (2015) Is digital disruption the end of health insurance? Some thoughts on the devising of risk. *Economic Sociology the European Electronic Newsletter* 17(1): 32–44.
- McGough SF, Johansson MA, Lipsitch M, et al. (2019) ‘Nowcasting by Bayesian smoothing: A flexible, generalizable model for real-time epidemic tracking’. *BioRxiv*, June, 663823. <https://doi.org/10.1101/663823>.
- McKeown T (1983) A basis for health strategies. A classification of disease. *British Medical Journal (Clinical Research Ed.)* 287(6392): 594–596.
- Milne R and Costa A (2020) Disruption and dislocation in post-COVID futures for digital health. *Big Data & Society* 7(2): 2053951720949567.
- Mittelstadt B, Benzler J, Engelmann L, et al. (2018) Is there a duty to participate in digital epidemiology? *Life Sciences, Society and Policy* 14(May): 9.
- Mohr DC, Shilton K and Hotopf M (2020) Digital phenotyping, behavioral sensing, or personal sensing: Names and transparency in the digital age. *NPJ Digital Medicine* 3(1): –2.
- Montag C, Dagum P and Elhai JD (2020) On the need for digital phenotyping to obtain insights into mental states in the COVID-19 pandemic. *Digital Psychology* 1(2): 40–42.
- Morabia A (2004) Epidemiology: An epistemological perspective. In: Morabia A (ed) *A History of Epidemiologic Methods and Concepts*, 3–125. Basel; Boston: Birkhauser Verlag.
- Onnela J-P and Rauch SL (2016) Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* 41(7): 1691–1696.
- Park H-A, Jung H, On J, et al. (2018) Digital epidemiology: Use of digital data collected for non-epidemiological purposes in epidemiological studies. *Healthcare Informatics Research* 24(4): 53.
- Prainsack B (2017) *Personalized Medicine: Empowered Patients in the 21st Century?* New York: NYU Press.
- Prainsack B (2019) Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data & Society* 6(1): 2053951719829773.
- Preis T and Moat HS (2014) Adaptive nowcasting of influenza outbreaks using Google searches. *Royal Society Open Science* 1(2): 140095.
- Reardon J (2017) *The Postgenomic Condition: Ethics, Justice, and Knowledge After the Genome*. Chicago: University of Chicago Press.
- Robinson PN (2012) Deep phenotyping for precision medicine. *Human Mutation* 33(5): 777–780.
- Robinson PN and Mundlos S (2010) The human phenotype ontology. *Clinical Genetics* 77(6): 525–534.
- Rüppel J and Voigt TH (2019) The death of the clinic? Emerging biotechnologies and the reconfiguration of mental health. *Science, Technology, & Human Values* 44(4): 567–580.
- Salathé M, Bengtsson L, Bodnar TJ, et al. (2012) Digital epidemiology. *PLoS Computational Biology* 8(7): e1002616.
- Stevens M, Wehrens R and de Bont A (2018) Conceptualizations of big data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society* 5(2): 2053951718816727.
- Sudre CH, Lee KA, Lochlainn MN, et al. (2021) Symptom clusters in COVID-19: A potential clinical prediction tool from the COVID symptom study app. *Science Advances* 7(12): eabd4177.
- Timmermans S and Berg M (2010) *The Gold Standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care*. Philadelphia, PA: Temple University Press.
- Topol E (2019a) The Topol review – preparing the healthcare workforce to deliver the digital future. *An Independent Report on Behalf of the Secretary of State for Health and Social Care February 2019*. London: NHS England. <https://topol.hee.nhs.uk/>.
- Topol E (2019b) *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. London: Hachette UK.

- Torous J, Kiang MV, Lorme J, et al. (2016) New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research. *JMIR Mental Health* 3(2): 16.
- Tracy RP (2008) “Deep phenotyping”: Characterizing populations in the era of genomics and systems biology. *Current Opinion in Lipidology* 19(2): 151–157.
- Vayena E, Salathé M, Madoff LC, et al. (2015) Ethical challenges of big data in public health. *PLoS Computational Biology* 11(2): e1003904.
- White RW, Murali Doraiswamy P and Horvitz E (2018) Detecting neurodegenerative disorders from web search signals. *Npj Digital Medicine* 1(1): –4.