Digital phenotyping – Editorial

Editorial

BIG DATA & SOCIETY

Big Data & Society
July–December: 1–4

DOI: 10.1177/20539517221113775
journals.sagepub.com/home/bds
SAGE

Digital phenotyping – Editorial

Lukas Engelmann and Ger Wackers

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

There is an astonishing posthuman promise in digital phenotyping, as Beth Semel recently argued (Semel, 2022). The goal of digital phenotyping enthusiasts is no less than to bypass the human observer as a deeply flawed threshold of medical knowledge production. The second goal is then – ultimately – to rid the human body and mind of its frailty and to utilise technology for a 'world without disease' (Topol and Corr, 2019). This promissory rhetoric is not only geared towards the disruption of dated medical conventions but comes equipped with bold, revolutionary concepts. Objective knowledge, based on aggregated, automated, and sweeping data collection to deliver granular, minute, and personalised healthcare; digital phenotyping is a collection of ideas, technologies, and practices to realise a powerful and futuristic vision of a medicine far beyond human capacities. This posthuman promise might be naive and driven by an abundant positivism, but as a small movement, made up of medical researchers and digital disruptors alike, it has continuously gathered steam over the last decade. The purpose of this collection is foremost to take stock and to collect a range of critical questions for a first revision of what digital phenotyping might be and what it could potentially become.

The meaning of digital phenotyping is not as well defined as the many publications in this growing body of scholarship might suggest. Some of that vagueness has been captured in the critical literature. Birk and Samuel, in their sociological analysis, have described the term recently in more general terms as an analytical concept that presumes simply that diseases and illness are by and large 'measurable by digital devices' (Birk and Samuel, 2020). This assumes that a person's experience of any kind of suffering is always in one way or another expressed in the digital traces of their behaviour. The leg injury that might result in a different mobility pattern; measurable tremors in the thumb control of smartphones as a sign of Parkinson's; sudden lack of social interaction as a sign of depression: digital phenotypes can in theory be defined for any illness and disease and captured by any of the sensors, devices, and technologies, through which humans leave digital traces. Loi, in his ethical and philosophical exploration of the digital phenotype, assumes it in more general terms to be 'an assemblage of information in digital form, that humans produce intentionally or as a by-product of other activities, and which affects human behaviour' (Loi, 2018). Many questions remain, not least why and how this concept seeks association with genetic terminology. What does the wholesale capturing of a human's digital traces as phenotype imply? What does it mean to group a sheer endless range of symptoms within the paradigm of inheritable traits and how does this framing structure research on and with digital phenotypes?

The phrase itself was coined by the physician Sachin Jain and colleagues at Harvard in 2015 in a letter to *Nature Biotechnology*. Conceptually, they conceived of *digital phenotyping* with reference to Richard Dawkins' elaborations on the 'extended phenotype' (Jain et al., 2015; Dawkins, 1982). Not only did they see digital technologies equipped to deliver a never-before-seen mass of potentially valuable data for diagnostics and prognostics but importantly these data were produced beyond the brief and cursory encounters between patients and physicians. The full-scale exploitation of these data would enable new insight into disease expressions over a lifetime. This was not only an expansion of surveillance but would open a new paradigm of medical knowledge production: rather than just recording symptoms in a medical consultation, 'digital phenotypes redefine disease expression in terms of the lived experience of individuals, which expands our ability to classify and understand disease' (Jain et al., 2015). In a 2017 JAMA article, the American neuroscientist Thomas R. Insel conceptualized digital phenotyping into nothing less but a 'New Science of Behaviour' (Insel, 2017). Since then, the phrase has given

focus and purpose to a growing field of digital disruptors, innovative physicians, and medical researchers, who seek – once again – to overcome human fallibility with yet another form of analytical approach that seeks fresh 'insights "born from data"' (Kitchin, 2014, 2).

This themed section of *Big Data and Society* explores digital phenotyping as a data-collection instrument and rationale in the world of data-driven health research. Equipped with promises of speed and immediacy, granularity, and comprehensiveness, digital phenotyping seeks to syphon observational data about human attributes and behaviour. Notably, this practice exceeds the realm of traditional medical data, as it is geared towards *repurposing* data from individual's online behaviours and from actively and passively produced information in wearables and smartphones explicitly beyond the structured information exchanges between patients and their physicians. Information inferred from these data is to improve disease classification, risk assessment, treatment optimisation, and public health prevention. Digital phenotyping practices are supposed to marshal the computational capacity of personally owned digital devices to shorten the loop between continuous, ecologically valid data generation and intervention delivery.

Importantly, digital phenotyping practices in medicine and public health are not on the distant horizon of medical utopianism. They are happening today, transforming extant practices and health care policies and give rise to emergent ways of seeing. Simultaneously, digital phenotyping is imbued with as yet unrealised but rhetorically powerful imaginaries of a revolutionised future. Whereas much research in digital phenotyping is still in a proof of concept and validation stage, the imagined cognitive superiority of digital data infrastructures and machine learning technologies is exerting its transformative power today. Despite the unsolved onto-epistemological conundrum of the relationship between the digital doppelgänger and the elusive real-world entities they are purported to represent, digital phenotyping practices are decentring both extant diagnostic practices and individual experiences of disease.

As the contributions to this theme show, practices and tools associated with digital phenotyping are hailed to refine the classification and understanding of psychiatric, infectious, and chronic conditions. By tailoring to individual phenotypes, they hold a generalised promise of optimisation, both in terms of effectiveness and economics. The data-driven approach is supposed to overcome vague and unstructured clinical observations to offer new, highly standardised pathways towards complete symptomatology of all diseases and conditions. However, digital phenotyping has not only offered a new technopolitics of collecting health-related data, its conceptual implications and technological affordances have also begun to reshape the foundations of medical knowledge production, raising epistemological questions across medical and public health specialities.

To capture the transformation in data collection, to discuss the often-undisclosed risks and problems of digital disruption and to critically engage the technology's impact on our understanding of health and illness, the collection of papers in this special theme focusses on three dimensions of digital phenotyping. Most contributions present selected case studies on *emergent medical practices* that integrate digital phenotyping in existing frameworks of care and medical and public health research. Almost all research articles and commentary pieces investigate and highlight the particularly prominent emergence of digital phenotyping within the realms of mental health research and *psychoinformatics*. The third dimension concerns the conceptual and philosophical underpinnings that inform digital phenotyping to address the *epistemological reconfiguration* of knowledge production in medicine and public health.

Emergent medical practices

The contributions to this dimension offer new empirical material with concise case studies to illuminate how digital phenotyping changes and challenges existing modes of medical practice. *Baumgartner* introduces the position of digital phenotyping in the project of precision, or personalised medicine, to clarify how the exploitation of data is envisioned to shift health care towards preventive and proactive interventions. Engelmann's contribution looks at the spurious claims that inform research built on the automated prediction of depression on Twitter (Engelmann, 2022), while Milne et al. focus on the 'behavioural surplus' of surveillance capitalism (Milne et al., 2022; Zuboff, 2019), where the recognition of Alzheimer's disease is potentially entangled with voice-activated home assistants. Rowe's article introduces 'Opioid360', a system to determine risk factors for potential opioid addiction to guide and improve prescription practices (Rowe, 2021). Birk et al. pay attention in their commentary to the impact of digital phenotyping on the pathologization of mental distress (Birk et al., 2021), while Green and Svendsen warn of the historical biases in automated diagnostic technologies (Green and Svendsen, 2021). Hussain and Bowker provide a critical commentary on the enrolment of physicians in digitally enabled policing of prescription practices (Hussain and Bowker, 2021), and Lucivero and Halliwell conclude the special theme with a critical comment on the development of RADAR-AD, an Alzheimer's prediction system, and the Minerva Initiative, which seeks to deliver phenotypes of rare disorders (Lucivero and Hallowell, 2021).

Psychoinformatics

Within the second dimension, the authors map out the rapid uptake and analyse the outstanding popularity of digital

phenotyping in mental health research and digital psychiatry (Pickersgill, 2019). Engelmann's article draws attention to the often-shallow underpinnings of the promises of deep medicine, when the assumptions and presumptions that inform the prediction of depression of Twitter users are scrutinized (Engelmann, 2022). Milne et al. demonstrate in their paper how digital phenotyping extends and exaggerates a biological reductionism in the research environments of Alzheimer's disease (Milne et al., 2022), while Rowe shows that the analysis of social factors of addiction within a digital phenotyping context contributes to their individualised assessment (Rowe, 2021). Birk et al. show likewise how digital phenotyping advances a 'psychocentrism', attributing mental distress increasingly to innate capacities of individuals (Birk et al., 2021). Wackers traces the legacy of the new psychoinformatics to the influence, in the late 19th and early 20th century, of physicists' conceptions of fundamental measurement in the development of psychometric instruments. These instruments were deployed in a new *correlational* psychology (Wackers, 2022).

Epistemology

The third dimension assembles a series of perspectives to inspect closely what kind of conceptual allegiances and theoretical kinships are invoked through often pragmatically introduced digital phenotyping. Baumgartner asks for increased analytical attention to the use and analysis of data in precision medicine and digital phenotyping, so that the often implied 'genetic determinism' and 'data fundamentalism' are acknowledged (Baumgartner, 2021). Engelmann investigates the depth of 'deep medicine' and asks how the emergence of digital phenotyping has been coupled to the reinvigoration of the dream of total archives, risking the introduction of totalising and normative conceptions of health and disease (Engelmann, 2022). Wackers rereads Dawkins' *Extended Phenotype* and uses the notion of the genotype to explore the problematic of probabilistic inference of causality; and the fallacy of the reification of generative, causal infrastructures (genotypes) underlying phenotypes derived from mathematical abstractions purported to express the strength of association (correlation) between variables defined as phenotypic that were assumed to share a common internal cause (Wackers 2022).

Acknowledgements

Most of the papers in this special section have been originally presented in the digital phenotyping panel at the 2020 (online) meeting of the Society for Social Studies of Science. As co-editors, the authors like to thank the journal editors for their help and patience, all the authors for their contributions and the reviewers for their productive suggestions and comments. Research for this editorial has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 947872).

Declaration of conflicting interests

The author(s) 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: This work was supported by the H2020 European Research Council, (grant number TER 947872).

ORCID iDs

Lukas Engelmann https://orcid.org/0000-0002-2175-0156
Ger Wackers https://orcid.org/0000-0002-8467-6029

References

Baumgartner R (2021) Precision medicine and digital phenotyping: digital medicine's way from more data to better health. *Big Data & Society* 8(2): 20539517211066452.

Birk RH and Samuel G (2020) Can digital data diagnose mental health problems? A sociological exploration of "digital phenotyping". *Sociology of Health & Illness*. https://doi.org/10.1111/1467-9566.13175.

Birk R, Lavis A, Lucivero F, et al. (2021) For what it's worth. Unearthing the values embedded in digital phenotyping for mental health. *Big Data & Society* 8(2), 20539517211047320.

Dawkins R (1982) *The Extended Phenotype: The Gene as the Unit of Selection*. New York: Freeman.

Engelmann L (2022) Digital epidemiology, deep phenotyping and the enduring fantasy of pathological omniscience. *Big Data & Society* 9(1): 20539517211066452.

Green S and Svendsen MN (2021) Digital phenotyping and data inheritance. *Big Data & Society* 8(2): 20539517211036800.

Hussain MI and Bowker GC (2021) Phenotyping as disciplinary practice: data infrastructure and the interprofessional conflict over drug use in California. *Big Data & Society* 8(2), 20539517211031256.

Insel TR (2017) Digital phenotyping: technology for a new science of behavior. *JAMA* 318(13): 1215.

Jain SH, Powers BW, Hawkins JB, et al. (2015) The digital phenotype. *Nature Biotechnology*, May. https://doi.org/10.1038/nbt.3223.

Kitchin R (2014) Big data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): 2053951714528481.

Loi M (2018) The digital phenotype: A philosophical and ethical exploration. *Philosophy & Technology*, June. https://doi.org/10.1007/s13347-018-0319-1.

Lucivero F and Hallowell N (2021) Digital/computational phenotyping: what are the differences in the science and the ethics? *Big Data & Society* 8(2): 20539517211062884.

Milne R, Costa A and Brenman N (2022) Digital phenotyping and the (data) shadow of Alzheimer's disease. *Big Data & Society* 9(1): 20539517211070748.

Pickersgill M (2019) Digitising psychiatry? Sociotechnical expectations, performative nominalism and biomedical virtue in (digital) psychiatric praxis. *Sociology of Health & Illness* 41(S1): 16–30.

Rowe R (2021) Social determinants of health in the big data mode of population health risk calculation. *Big Data & Society* 8(2): 20539517211062880.

Semel BM (2022) Listening like a computer: attentional tensions and mechanized care in psychiatric digital phenotyping. *Science, Technology, & Human Values* 47(2): 266–290.

Topol E and Corr J (2019) Eric Topol, "Deep Medicine: How Artificial Intelligence Can Make…. New Books Network." Available at: https://newbooksnetwork.com/eric-topol-deep-medicine-how-artificial-intelligence-can-make-healthcare-human-again-basic-books-2019 (accessed 25 May 2022).

Wackers Ger (2022) Fitness trackers of the mind: digital phenotyping from biology to psychoinformatics. *Forthcoming in Big Data & Society*.

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: Barack Obama's Books of 2019*. London: Profile Books.