Big-data studies need to be part of policy discussion

To the Editor — The new definition of clinical trials by the National Institutes of Health (NIH) has raised manifold comments by scientists working in the field of basic behavioural and brain research both in the United States and internationally. Clearly, the aim of increasing accountability and transparency of clinical research is not controversial, and measures need to be identified to support that goal. However, the new definition changes the scope of prototypical clinical trials from the concept of 'medical strategy, treatment or device' and 'medical approaches' to 'health-related biomedical or behavioural outcomes', that is, from 'healthcare' to 'health related'. It might be argued that the new definition leads to an inflationary use of the term for studies addressing almost any change in the physiological range of normal behaviour or body reaction. This could be interpreted as pathologization (or medicalization) of normal processes and life cycles with all its philosophical, ethical and sociological dimensions and societal implications.

Moreover, a definition emphasizing health-related and behavioural outcomes as a result of intervention could eventually include big-data approaches. An increasing number of studies, including non-medical research, collect and analyse biomedical and behavioural data. Deep learning represents an increasingly powerful tool to extract detailed information regarding health status or personality traits from seemingly non-health-related data, such as motion profiles or buying behaviour. The concrete context hereby defines whether a study is health related or not.

Many of such studies and their results are not made public[1]. In addition, a large portion of data is being collected, processed and analysed by Internet providers, where workflows are not necessarily transparent. Big-data studies may not only analyse data, but also manipulate subjects' behaviour and/or biomedical outcomes.

For example, a recent study used search engines with masked biases, which left people unaware of being manipulated but changed voting behaviour and intentions[2]. The authors concluded that search engine companies, which are currently unregulated, could affect the outcomes of close elections worldwide[2].

The study is exemplary, because the approach can easily be transferred to many other scenarios, including health topics. A study evaluating massive emotional contagion through social networks[3] raised the question of when ethical approval is necessary. It's unclear where such studies belong under the new definition and how they can be captured, especially when research is funded by the private sector.

In the context of big data, the broad definition of clinical trials may not resolve a problem, but postpone it to another level considering new ways of intervention at the level of behaviour and health in the era of big data. It might be useful to distinguish between different categories — clinical studies in the traditional meaning with their strict regulations, and health-related and behavioural studies, addressing basic neuroscience, behavioural and sociological questions with their specific constraints, including, for example, institutional review board approval, transparency and documentation. A broader discussion seems to be necessary, on a national and an international level, to elucidate the ethical and legal regulations necessary to handle health-related big-data studies, to make sure that scientific progress is not impeded, freedom and privacy are being respected, and transparency and accountability are ensured. ❐

Katrin Amunts[1,2]

[1]*Cécile and Oskar Vogt Institute for Brain Research, University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.*
[2]*Institute for Neuroscience and Medicine, INM-1, Research Centre Jülich, Jülich, Germany.*
e-mail: k.amunts@fz-juelich.de

Published online: 22 January 2018
https://doi.org/10.1038/s41562-018-0292-9

References

1. Wilbanks, J. T. & Topol, E. J. *Nature* **535**, 345–348 (2016).
2. Epstein, R. & Robertson, R. E. *Proc. Natl. Acad. Sci. USA* **112**, E4512–E4521 (2015).
3. Kramer, A. D. I., Guillory, J. E. & Hancock, J. T. *Proc. Natl. Acad. Sci. USA* **111**, 8788–8790 (2014).

Competing interests

K.A. is deputy chair of the German Ethics Council, which in November 2017 published a statement on big data and health. She is also scientific director of the European Human Brain Project, which connects neuroscience and computational approaches on a large scale. K.A. declares present funding by the NIH. She does not and has never held an advisory role for the NIH, and she is not and has not been involved in any planning of clinical trials policy or implementation.