

# Private UK health data donated for medical research shared with insurance companies

Publication Date: 2023-11-12

Author: Shanti Das

Section: Technology

Tags: Data protection, The Observer, Medical research, Health, Privacy, Insurance industry, Big data, Biotechnology industry, news

Article URL: <https://www.theguardian.com/technology/2023/nov/12/private-uk-health-data-donated-medical-research-shared-insurance-companies>



Sensitive health information donated for medical research by half a million UK citizens has been shared with insurance companies despite a pledge that it would not be. An Observer investigation has found that UK Biobank opened up its vast biomedical database to insurance sector firms several times between 2020 and 2023. The data was provided to insurance consultancy and tech firms for projects to create digital tools that help insurers predict a person's risk of getting a chronic disease. The findings have raised concerns among geneticists, data privacy experts and campaigners over vetting and ethical checks at Biobank. Set up in 2006 to help researchers investigating diseases, the database contains millions of blood, saliva and urine samples, collected regularly from about 500,000 adult volunteers – along with medical records, scans, wearable device data and lifestyle information. Approved researchers around the world can pay £3,000 to £9,000 to access records ranging from medical history and lifestyle information to whole genome sequencing data. The resulting research has yielded major medical discoveries and led to Biobank being considered a “jewel in the crown” of British science. Biobank said it strictly guarded access to its data, only allowing access by bona fide researchers for health-related projects in the public interest. It said this included researchers of all stripes, whether employed by academic, charitable or commercial organisations – including insurance companies – and that “information about data sharing was clearly set out to participants at the point of recruitment and the initial assessment”. But evidence gathered by the Observer suggests Biobank did not explicitly tell participants it would share data with insurance companies – and made several public commitments not to do so. When the project was announced, in 2002, Biobank promised that data would not be given to insurance companies after concerns were raised that it could be used in a discriminatory way, such as by the exclusion of people with a particular genetic makeup from insurance. In an FAQ section on the Biobank website, participants were told: “Insurance companies will not be allowed access to any individual results nor will they be allowed access to anonymised data.” The statement remained online until February 2006, during which time the Biobank project was subject to public scrutiny and discussed in parliament. The promise was also reiterated in several public statements by backers of Biobank, who said safeguards would be built in to ensure that “no insurance company or police force or employer will have access”. This weekend, Biobank said the pledge – made repeatedly over four years – no longer applied. It said the commitment had been made before recruitment formally began in 2007 and that when Biobank volunteers enrolled they were given revised information. This included leaflets and consent forms that contained a provision that anonymised Biobank data could be shared with private firms for “health-

related” research, but did not explicitly mention insurance firms or correct the previous assurances. Biobank also said commitments that “insurance companies ... will not be given any individual’s information, samples or test results” – repeated in leaflets over a 17-year period – meant to refer to identifiable information, such as that which is linked to a person’s name, rather than to other data about Biobank participants. The exact nature of the data shared with the insurance industry is not clear because Biobank does not routinely publish this and has declined so far to say. Summaries of the projects published online suggest it included de-identified, participant-level data on diseases, lifestyle and biomarkers. One company granted access, ReMark International, is a “global insurance consultancy” that underwrites a million policies a year and lists clients including Legal & General and MetLife. In its application to Biobank, approved in December 2022, the company said it needed data to develop an algorithm to predict diseases and death, using hospital records and smartwatch data to examine the relationship between lifestyle, mental health and biomarkers. Another firm given Biobank data, Lydia.ai, is a Canadian “insurtech” firm that wants to give people “personalised and predictive health scores”. The company says insurers work with it to “leverage new sources of data to make risk predictions”. It was granted access to Biobank data in January for a project linking health records to lifestyle data to “predict chronic diseases”. Club Vita, a “longevity data analytics company for pension funds & their advisors, insurers, reinsurers and asset managers” – whose clients include 400 pension funds and 25 insurers – was also granted access. Its project sought to assess data on morbidity outcomes using a range of risk factors such as gender, diseases, treatment, location and lifestyle. Prof Yves Moreau, a genetics and AI expert who has worked on projects using data from UK Biobank, said the data-sharing appeared to be a “serious and disturbing breach of trust”. He said the idea that Biobank’s public commitments could be “silently superseded” by leaflets was “weak”, and questioned whether participants understood that data could be shared with insurance firms. “The data looks very mundane – a bunch of measurements. But there are really major impacts,” he said. Prof Sandra Wachter, an expert in technology and regulation at the Oxford Internet Institute, said the cases risked eroding the trust of volunteers who “donated their data for a good cause”. She said the development of insurance products to “predict if someone will get sick” raised serious ethical concerns. Sam Smith, coordinator of medConfidential, which campaigns for the privacy of health data, said people gave data to Biobank to “help cure diseases”, not so it could be used by the insurance industry. He said: “Biobank must tell every participant what data was shared with insurance companies and why.” Biobank said it rejected any suggestion that data had ever been shared for uses that volunteers had not consented to, and said it was wrong to suggest that prior promises – which pre-dated formal enrolment at Biobank – should still apply. It added that researchers worked for “all manner of companies”, and that provided they passed its “stringent access protocols”, they could conduct research using Biobank data. Research by insurance companies into how lifestyle behaviours can improve health or help identify health risks was “consistent with being health-related and in the public interest”, it said. It added that it had consulted independent ethicists “at length” about commercial data sharing, and that “complex” applications were referred to an expert committee. Prof Naomi Allen, chief scientist at UK Biobank, said: “Our careful processes have been followed in all these cases. De-identified health data has been shared because these are bona fide researchers working on health-related research, including looking at what impacts human health and longevity – and that is what our participants signed up to help with.” There is no suggestion that Biobank data has ever been used by insurers to make direct decisions about individual policies. No physical biological samples were shared. As well as insurance sector firms, Biobank data has also been given to other companies that are not directly health-related, including pension funds and investment firms, project records show. In another case that has raised questions for Biobank, a California company whose website is covered in spelling mistakes was granted access to data. Flying Troika LLC’s website says it is a “pure research lab” offering “deep larning” solutions in sectors including insurance, pharma, manufacturing and retail. It says it has teams in 13 cities, including “Maimi” and Edinburgh. The company is understood to have sought genetic data, MRI scans and other information in April 2021, to develop a “novel AI model” that can predict ageing processes. Prof David Leslie, director of ethics and responsible innovation at the Alan Turing Institute, said: “Making explicit ... just how each of these projects counted as data being used for medical projects in the public interest would seem essential for maintaining public trust.” The Information Commissioner’s Office, the UK’s data privacy watchdog, is considering the matter. It said: “People have the right to expect that organisations will handle their information securely and that it will only be used for the purpose they are told or agree to. Organisations must provide clear, accurate and comprehensive information ... especially where sensitive personal information is involved.”