

DEREK ARMSTRONG



Geneticists should offer data to participants

Sarah Nelson was refused access to her own genome data. How long before volunteers who face this attitude turn away from science?

As a human-genetics researcher, I analyse the DNA of thousands of anonymous strangers. Earlier this year, I got to experience the other side of a consent form — and was left disappointed. When another research group asked me to donate my own genetic material for their whole-genome sequencing project, I asked in exchange for access to my raw data — to explore, to play with and just to have on file. Not surprisingly, my request was refused: the status quo for biomedical and genetic studies is not to return individual-level data to participants.

I still joined the study, but the irony is not lost on me that my personal data will be available to thousands of scientists (including me) through restricted-access databases. As awareness and usefulness of this information increases, I fear that potential volunteers who are refused access to their genetic data will become less willing to donate them to science. The genetics-research community must therefore update its stance on returning personal data.

Granted, there are well-founded reasons why studies don't typically return the data. Researchers rarely recontact participants, and doing so could draw resources and attention away from the primary project goals. Research is not medicine, and returning data can create the misleading impression that researchers are offering health care. Historically, there has also been little reason to return genetic data, because volunteers couldn't access tools to receive, store or understand them.

But beyond the lab, more people now want and expect access to all kinds of personal data, a trend that shows no signs of slowing. Health data are no exception, as evidenced by the flood of wellness and 'mobile health' apps that are now coming to market. Future generations will take for granted that our personal computing devices are vehicles for almost unlimited 'quantified self' and self-tracking activities.

Several online platforms exist to help people to explore their genetic data, developed by for-profit companies, academic groups or by self-taught citizen scientists. Since launching in 2011, the site openSNP has drawn more than 4,000 users, half of whom have uploaded genetic data. The DNA.LAND platform has attracted more than 32,000 contributors since its release last October. Launched in 2008, and therefore one of the earliest third-party interpretation tools, Promethease reports performing hundreds of analyses daily. Other tools, such as GEDMatch and Genome Mate Pro, attract thousands of users who are eager to incorporate genetic analyses into their genealogical research.

Many scientists are suspicious and occasionally derisive of consumer or 'recreational' genomics. Although these products have their flaws, they underscore what citizens can and want to do with their genetic data. For the genetics-research community to maintain its good relationship with volunteers, it must take these activities more seriously. Many current large-scale genetics-research studies rely on legacy collections, and have not had to navigate the new 'participant-as-owner'

culture. But legacy studies cannot fuel future research indefinitely.

The genetics-research community needs to develop an anticipatory infrastructure to return raw data to interested participants. The format for returning genotype data would probably vary according to the nature of the study — depending, for instance, on whether next-generation sequencing or microarray genotyping was performed. Funding opportunities should include resources for participant data return. Institutional review boards need to be able to review the mechanisms that studies propose for returning genetic data. Research groups should develop and adopt informed-consent procedures so participants can make decisions about acquiring their raw data, including the limitations of self-directed interpretation and analysis.

Making data-return practicable might require building technical systems such as secure web interfaces. However, we already have secure and robust methods to share data within the scientific community, so perhaps the necessary change in culture is a bigger hurdle.

Some research initiatives are already experimenting. In early September, the New York Genome Center in New York City released Seeq, a smartphone app and research platform through which individuals can pay a modest fee (around US\$50) to receive their whole-genome sequence and some interpreted reports (such as ancestry composition and microbiome profiles). In turn, the researchers amass genomes for their own research projects.

Most of the available consumer-genomics tests look at a very small portion of the genome, but newer and more powerful exome and whole-genome sequencing options offer more detail at a rapidly decreasing cost. MyGene2, a web portal created by researchers at the University of Washington in Seattle, enables the sharing of genetic-sequencing and medical data across families, physicians and researchers, with the goal of tackling rare genetic diseases. Genomics tools such as this flourish when people are stewards of their own genetic data.

This era of big data begs big questions, including who should own health and research data. Legally, research participants may not own their biological specimens, or the data extracted from them, once these have been donated to scientific studies. But as researchers, don't we have an obligation to respect the individual autonomy of participants seeking their raw data? Asking for access is not the same as asking for ownership or control, just for a reasonable reciprocity. Let's have the conversation. After all, if potential participants can obtain their genetic data from a growing number of commercial companies, they might turn their backs on traditional research studies altogether. ■

Sarah Nelson is a researcher and PhD student at the Institute for Public Health Genetics at the University of Washington in Seattle. e-mail: sarahcn@uw.edu

**GENOMICS TOOLS
FLOURISH
WHEN PEOPLE ARE
STEWARDS OF THEIR
OWN GENETIC
DATA.**