

Are genetic databases and medical confidentiality compatible?

Personalized medicine, the hoped-for use of the information in our genes to inform our medical care, may end up helping people live longer, healthier lives. Or it may not—the jury is still out. But one thing is certain: As our unique genomic data enter our medical records, researchers will be tempted to use that invaluable resource. The results may be good for science but bad for patients' privacy.

Advances in data science and information technology are eroding old assumptions—and undermining researchers' promises—about the anonymity of DNA specimens and genetic data. Databases of identified DNA sequences are proliferating in law enforcement, government, and commercial direct-to-consumer genetic testing enterprises, especially in genetic genealogy. That growth is increasing the likelihood that anyone with access to such nonanonymous "reference" databases could use them to re-identify the person who provided a "de-identified" gene sequence. People with access could include amateur genetic genealogists but also hackers.

Similarly, information about a person's health conditions or demographic characteristics can be used for re-identification. How many 6-foot-2-inch-tall 62-year-old white men are there in a given state with white hair, an artificial left hip, type A positive blood, and a prescription for warfarin?

Read full, original article: [Every Patient a Subject](#)