Genetic information: How much to share, how much to keep private in age of Big Data

To what extent should personal genetic information be protected?

One big concern revolves around how much information should be supplied to health insurance companies. If you carry a BRCA gene for breast and ovarian cancer, which makes it likely but not certain that you will develop cancer if you are female, you certainly don't want this to increase the premium that you pay. And if you choose an expensive procedure, such as prophylactic mastectomy and reconstruction because you carry the gene, you want the assurance that the insurance company would cover the cost.

Taking another perspective, some people might not wish to know if they are genetically prone to a rare disease for which there's no effective treatment, particularly if having the gene for the disease does not necessarily mean that the person will ever develop it. Many men, for example, carry the BRCA gene, which only slightly elevates their chance of male breast cancer, but they can also pass that mutation on to their daughters. A person who may carry that gene might not want other people to know while others might want to share that information. And those who may be impacted, from insurance companies to relegates, might also want to know who carries it.

The issue of genetic privacy can get even more confusing if we consider mental health. In a statement published in 2013, the International Society of Psychiatric Genetics (ISPG) points to a "history of successful use of genetic tests for several neuropsychiatric disorders", but cautions that this applies mostly to conditions affecting brain development in early life. Thus, genetic testing is very reliable for screening infants for conditions like Huntington disease, fragile X syndrome, phenylketonuria and Down syndrome. These are medical conditions with easily recognized physical effects on various body systems, but that also include mental health symptoms.

In contrast, while genome wide association studies (GWAS) have attempted to reveal genetic patterns for mood, psychotic and substance abuse disorders that typically appear in young adulthood, the genetics of these conditions turns out to be extremely complex. The complexity is not only because of non-biological factors (i.e. *nurture* as opposed to *nature*), but also due to multiple genes interacting in a plethora of combinations. Therefore, we do not seem to be approaching an era when something as simple as a blood test can reliably identify individuals who are likely to be a danger to others, or to themselves.

In addition to highlighting the complexity of psychiatric genetics, the ISPG statement emphasizes the importance of informed consent. Anyone subject to genetic testing must agree to it. That's very straightforward decision to make when a person, or his or her infant, is being screened for an illness to facilitate early intervention. In such cases, the chances seem low that the genetic information would be used for anything but to help the individual who is screened. But what about using the information for to help others, or the population as a whole?

This question will come up increasingly, for our society, and for those deciding whether to give consent for genetic testing, as we move further into the age of computational biology. Sometimes called biological

data mining, computational biology takes data that have already been obtained in the field or laboratory for various reasons and utilizes them to answer big questions.

For instance, a recent study based at MIT and Harvard pooled numerous medical genetic data from many individuals and analyzed the data in relation to the prevalence of blood cancers (leukemias and lymphomas). The analysis revealed an association between certain genetic patterns, particularly mutations of three genes, and both categories of blood cancers. The association is so strong that blood tests can now be developed using the revealed genetic patterns to identify individuals with lymphoma and leukemia in *premalignant* stages. Premalignant means that a cancer has not advanced, or has not spread, enough to cause symptoms, and is so well isolated that it can be removed from the body with relatively simple treatments. In other words, premalignant disease is completely curable if you know it exists, and thus tests for the newly discovered genetic pattern constitute preventive screening.

Genetics and computational biology are evolving so quickly that it's impossible to predict the science and medical questions that they'll be used to research, let alone the answers that they'll reveal, in the years to come. But what does seem to be predictable is one question that patients, human subjects, and policy makers will be asking over and over as biological science continues to develop: to what extent should genetic information be kept private?

David Warmflash is an astrobiologist, physician, and science writer. Follow @CosmicEvolution to read what he is saying on Twitter.