GENETIC TESTING AND MOLECULAR BIOMARKERS
Volume 17, Number 3, 2013

Pp. 175–177
DOI: 10.1089/gtmb.2012.1542

PERSPECTIVE

The Haystack Is Made of Needles

Sharon F. Terry,[1] Robert Shelton,[2] Greg Biggers,[1,3] Dixie Baker,[4] and Kelly Edwards[5]

Developing genetic tests that have clinical utility and validated biomarkers presents many challenges. Much has been written about these challenges for the development of genetic test evidence (Khoury *et al*., 2010; Horn and Terry, 2012) and biomarker validation (Lesko and Atkinson, 2001; Surh, 2009). One consistent thread through these challenges is the lack of well-characterized cohorts. This is already generally true for common conditions (<5% of patients with cancer participate in clinical trials [Michaels, 2012]), and it seems likely this problem will only worsen as cancers are stratified by genomic tests. Finding a robust cohort has always been difficult for rare diseases, simply by virtue of their rarity (Griggs *et al*., 2009; Gliklich and Leavy, 2011). Finding individuals to create these cohorts can feel like finding a needle in a haystack. And yet, this way of framing the problem unnecessarily boxes us in. As a solution, we are designing and pursuing research innovations that turn that framing on its head. Notice what happens to the possibilities when we consider that we have before us not a needle in a haystack but a haystack composed of needles!

Instead of trying to enroll participants in a clinical trial by seeking only those few rare individuals with a particular variation or individuals affected by a rare disease (the needles in a haystack), we need to engage all individuals (because the haystack is made of needles). This simple concept has been discounted as too expensive, producing excessively noisy and dirty data, and being fundamentally impractical in an age when researchers are burdened by national and institutional regulations that make it difficult to engage anyone but a well-defined cohort formed around a hypothesis. As more people are becoming aware of this hidden potential, we are seeing a shift from hypothesis-constrained cohort identification toward the creation of hypothesis-generating registries and repositories (Ritchie *et al*., 2010; Manolio *et al*., 2012). Although this is a welcome shift, it generally originates in institutions seeking robust resources in the form of registries to support their own research, or those that are bravely encouraging open data sharing (Norman *et al*., 2011). As a result of this limited scope, each has limited value for creating meaningful cohorts for advancing widespread research objectives. We speculate that given the right tools, it is possible to create a novel type of registry that overcomes these limitations and identifies more robust cohorts for both common and rare diseases. This shift in capability requires looking at biomedical research from the perspective of the people who will benefit most: each of us. This work starts from the premise that a carefully maintained haystack made of needles will be of use to researchers and appealing to participants.

Genetic Alliance has a long history of engaging the public with comprehensive and accessible health information and in building capacity in disease advocacy organizations to create platforms to better engage their respective populations. Two major results of these endeavors are Disease InfoSearch (www.diseaseinfosearch.org) and the Genetic Alliance Registry and BioBank (GARB; www.biobank.org). Since their creation, Genetic Alliance has experienced some of the challenges described earlier. As increasingly powerful and less costly technologies are developed for genomic testing, data analysis, and exchange, Genetic Alliance has also been mindful of the ethical issues inherent in data sharing; it led the coalition that moved the Genetic Information Nondiscrimination Act of 2008 (GINA) through the US Congress over 12 and a half years. GARB was created in 2003, the only layperson-owned and -managed cross-disease registry and biorepository platform in the world, but unfortunately it has not yet reached critical mass. The many registries created by disease advocacy organizations are highly trusted, but they lack the cross-disease features needed to support a systems approach to diseases; ease of discovering cohorts; and appropriate, individualized access to clinical trials for their participants.

We observed a common characteristic in all these databases of personally identified information. In these registries—whether owned by disease advocacy organizations, universities, government agencies, or industry (and indeed, even in the case of GARB)—the individuals who elect to participate hand over control of their data and samples to a proxy. Although that proxy is often considered a trusted entity, the fact is that the individual who contributed data is no longer a participant in the registry. Data context is lost because the experience of the individual and the community in which she or he lives are not available. More often than not, the opportunity for engaging individuals in their own health, and moving translation forward toward effects on actual health outcomes, is neglected in favor of more conventional research goals.

[1]Genetic Alliance, Washington, District of Columbia.
[2]Private Access, Inc., Irvine, California.
[3]Genomera, Inc., Palo Alto, California.
[4]Martin, Blanck & Associates, Redondo Beach, California.
[5]UW Center for Genomics and Healthcare Equality, University of Washington, Seattle, Washington.

We need an innovative solution built on the simple technologies that have emerged for citizen engagement in other sciences and in a variety of industries. Individuals need direct involvement with the "science" in the form of making decisions about who can use their data, for what purpose, and when. They need to see and experience a stronger connection between their health, the health of their loved ones, and advancing biomedical research (Terry and Terry, 2011).

To accomplish this, individuals need the ability to determine their own privacy and sharing preferences, which will vary widely among individuals in different circumstances and over time. Personal preferences and context are critically important and call for privacy controls that are more granular and enable privacy settings that may change as circumstances and available information change. New technology, called privacy layer by private access, makes this possible (Shelton, 2011). Unique and equally dynamic governance structures are being tested (O'Doherty *et al.*, 2011). Assembly of these resources and expertise has led to the creation of the Registry for All Diseases (Reg4ALL; http://www.Reg4ALL.org), which recently won the Sanofi Collaborate Activate Innovate Challenge (http://www.collaborateactivate.com/).

Reg4ALL uses a tiered approach to participation in research that is best explained through a highway metaphor. Entering the highway is facilitated by Disease InfoSearch. This technology will guide users from one of more than 13,000 disease listings (which provide information and support resources for persons affected by these conditions) and onto an "on-ramp." The on-ramp is made accessible according to the privacy settings created by each participant using Private Access, and from the on-ramp the participant proceeds to the Reg4ALL highway. Each lane represents a higher level of participant involvement and disclosure, as the participants experience the opportunity to provide more sensitive data and to make them available (or not) to researchers. All data that traverse the highway are protected by strong security policies and technology.

The first lane consists of general questions in a gamified survey, meaning that users will automatically be able to see the results of each question as they answer it and, in doing so, appreciate how their answers compare with those of others. As users progress to the second engagement lane, they are asked to provide information on the common data elements, published by the National Institutes of Health as important for all diseases, enabling the measurement of the same things, the same way, across all diseases. Moving to the third engagement lane, users are asked to respond to many in-depth questions specific to a disease or conditions affecting the respondents.

Researchers will access Reg4ALL using RecruitSource, a software application that enables researchers to search the database and determine whether a cohort is available for their studies. Any researchers who wish to engage must have institutional review board approval for their study before they can query whether the registry contains any data of potential interest. And even then, the only data that these researchers may access are determined by each participant's personal privacy settings. This may be accomplished through a participant's advance authorization or in response to a participant's being notified of interest by that particular researcher and given an opportunity to consent in advance to the information being provided to them. Reg4ALL is overseen by an external Biotrust Ethics Team composed of leading international experts in research governance, dynamic consent, data sharing, and participatory research practices.

Reg4ALL was not created overnight. It is more than15 years in the making and employs enormous private and public experiences and investment, catalyzed by the Sanofi Collaborate/Activate Innovation Challenge award. Crowdsourced, participant-controlled shared data hold the promise for fundamentally transforming authentic investment and participation in biomedical research and helping to accelerate the development of diagnostics, biomarkers, and treatments. A revolution is upon us—let's seize the moment!

Acknowledgments

The authors are deeply indebted to an enormous number of people and institutions. This is a good thing because it indicates the collaborative partnerships that are necessary to build a true participant-centric system. First, Sanofi and the collaborators in the Sanofi Challenge contributed immensely to this project: Suzanne Vernon of CFIDS Association of America, Marin Aurand of National Psoriasis Foundation, and Ginny Mason of the Inflammatory Breast Cancer Research Foundation. For most of Genetic Alliance's years of work on this, Michele Lloyd-Puryear in her former position as chief of Genetics Services Branch, Maternal and Child Health Bureau, Health Services Resources Service Administration, was both morally and materially supportive (Disease InfoSearch was built in part with Health Resources and Services Administration funding). We are immensely grateful to the generous staff of the National Center for Biotechnology Information of the National Library of Medicine, National Institutes of Health. Steve Groft, director of the Office of Rare Disease Research, National Institutes of Health, offered technical assistance and cheers us from the sidelines. Project Miso has offered its expertise for years on this project and is a wonderful creative and technology partner in our web presentation. Traitwise's gaming technology offers us an excellent platform for consumer engagement. The Biotrust Ethics Team, led by Kelly Edwards with the UW Center for Genomics and Healthcare Equality, guides us with their worldclass expertise: Nick Anderson, Greg Biggers, Leila Jamal, Jane Kaye, Kieran O'Doherty, Suzanne Vernon, and David Winickoff. We celebrate each of the individuals who engage in our participant-driven studies who have shown us this is more than just a theory. Our respective boards and staff members (especially Alyson Krokosky, Lauren Youngborg, James O'Leary, Lisa Wise, Natasha Bonhomme, Scott McDaniel, Susan Barba, and Barbara Wilkinson), many of whom worked nights and weekends often without compensation because they are passionate about serving those who need us most, brought us here.

References

Gliklich R, Leavy M (2011) Patient registries and rare diseases. Appl Clin Trials 20, 1–5.

Griggs RC, Batshaw M, Dunkle M, (2009) Clinical research for rare disease: opportunities, challenges, and solutions. Mol Genet Metab 96, 20–26.

Horn EJ, Terry SF (2012) Precision medicine: generating real-world evidence for companion diagnostics. Genet Test Mol Biomark 16, 75–76.

Khoury MJ, Coates RJ, Evans JP (2010) Evidence-based classification of recommendations on use of genomic tests in clinical practice: dealing with insufficient evidence. Genet Med 12, 680–683.

Lesko LJ, Atkinson AJ Jr (2001) Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. Annu Rev Pharmacol Toxicol 41, 347–366.

Manolio TA, Weis BK, Cowie CC, et al. (2012) New models for large prospective studies: is there a better way? Am J Epidemiol 175, 859–866.

Michaels M, Weiss ES, Guidry JA, et al. (2012), The promise of community-based advocacy and education efforts for increasing cancer clinical trials accrual. J Cancer Educ 27, 67–74.

Norman TC, Bountra C, Edwards AM, Yamamoto KR, Friend SH (2011) Leveraging crowdsourcing to facilitate the discovery of new medicines. Sci Translat Med 3, 88mr1.

O'Doherty KC, Burgess MM, Edwards K (2011) From consent to institutions: designing adaptive governance for genomic biobanks. Soc Sci Med 73, 367–374.

Ritchie MD, Denny JC, Crawford DC, et al. (2010) Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. Am J Human Genet 86, 560–772.

Shelton RH (2011) Electronic consent channels: preserving patient privacy without handcuffing researchers. Sci Transl Med 3, 69cm4.

Surh LC, Lesko LJ, Hobbs S, et al. (2009) Fit-for-purpose pharmacogenomic biomarkers in drug developme nt: a project team case study with 'what-ifs'. Pharmacogenomics 10, 137–147.

Terry SF, Terry, PF (2011) Power to the people: participant ownership of clinical trial data. Sci Transl Med 3, 69cm3.

Address correspondence to:
Sharon F. Terry, MA
President & CEO
Genetic Alliance
4301 Connecticut Avenue, NW
Suite 404
Washington, DC 20008

E-mail: sterry@geneticalliance.org