

Allying with Donors to Link Health and Medical Information with Stem Cell Lines Can Advance Disease Modeling while Enhancing Data Access

A fundamental concern in disease modeling with stem-cell-derived cells is the need to connect information about cellular and tissue dysfunction in the lab with the manifestation of pathologies in individuals. To do this precisely necessitates linking personally identifiable information about individual patients, including medical records, biosensor data, and other potentially sensitive personal health information (PHI), to stem cell lines made from their donated tissue. Linking rich phenotypic data with stem cell lines, however, comes with a risk of informational harm, should donor anonymity be breached. To address this risk, the International Stem Cell Forum's Ethics Working Party (ISCF EWP) recently made recommendations for protecting donor-specific data (Isasi et al., 2014). Although we agree with the spirit of the recommendations, we argue that an alternative approach that actively engages donors would be better able to resolve data access issues while advancing disease modeling science.

Although regulations currently exist to protect participants' privacy (OHRP, HIPAA, EU Data Protection Directive), such protections make it difficult to link identifiable PHI to biospecimens, and not all conditions are covered (Hogle, 2014). Data access policies are often institution specific and may either be too restrictive to enable sharing among scientists or too open to protect donors' privacy (Kreiner and Irion, 2013). Sensitive information about diagnoses, lifestyle, and socioeconomic or demographic information could be exposed, creating stigma and prejudice and harms to family relationships, employment, or insurability. The situation is more complex for research using induced pluripotent stem cell (iPSC) lines than embryonic stem or immortalized tumor-derived cell lines because there is likely to be a living, uniquely identifiable donor—and family members—who may potentially be harmed (e.g., the case of HeLa cells; Hudson and Collins, 2013). The ISCF EWP recommendations repre-

sent a positive step toward balancing these concerns (Isasi et al., 2014). However, ambiguities in the guidelines pose challenges for their implementation. There is no clear definition of what constitutes "approved" uses, who is "unauthorized," or what constitutes "necessary sharing." There is little capacity to audit uses of data once acquired or what kinds of sanctions might be functionally relevant, should misuses occur. Overall, the recommendations are based on assumptions of what the concerns of individuals and researchers might be, without considering what exactly "privacy" and "data security" mean in a digital, postgenomic era. Privacy may be differently understood by ethicists, scientists, and donors. Some studies suggest, for example, that donors want their information to be broadly shared (Rodriguez et al., 2013). Furthermore, security and privacy concerns evolve as donors' experiences with disease change, and there may be issues specific to each disease. The stem cell community has not yet addressed broader issues underlying the current challenges of data linkage or more fundamentally, how to address any emerging governance issues.

We argue that donors should be more than bodies from which to extract biosamples and data; rather, they could be valuable partners in crafting the best approaches to acquire sensitive yet critically important information and deal with other emerging governance issues. To enable partnerships among donors, researchers, ethicists, and clinicians, an alliance could be nucleated from existing organizations. For example, a network of researchers from NINDS (e.g., within their iPSC consortia) could identify key patient advocacy groups and clinics with significant patient pools of the diseases of interest (e.g., Huntington's, Parkinson's, or amyotrophic lateral sclerosis), and a core reprogramming facility (e.g., the Salk Institute Stem Cell Core) could serve as an operations center, hosting workshops and housing online forums. Such an alliance could

design a systematic, scalable infrastructure to more effectively deal with data access and sharing concerns. Critical choices could be made together, such as *which* data need to be protected, whether some data could be *selectively* shared to particular parties, and *how* data could be made more secure. Potential outcomes of such collaboration include the innovation of tools and procedures, such as ways to scrub identifiable raw data within remote devices prior to uploading into cloud servers, or utilizing new informatics tools to generate "synthetic data" from authentic data (Malin et al., 2011). For example, a NINDS alliance could engage Parkinson's disease patients from whom cell lines have been created to determine which biosensors and data management practices work well to furnish secure data while causing minimal intrusion into family life.

For the disease modeling field, it is particularly important to capture the actual illness experiences of patients, since *in vitro* work, biorepository data, and even the limited information contained in medical records does not provide a comprehensive picture of what is happening *in vivo* as the disease progresses. An alliance could illuminate important information about disease phenotypes that is often incompletely captured through existing infrastructures. Our model can help guide choices in the lab (Saha and Jaenisch, 2009) about what cell type to use, at what stage of development, and what interventions could rescue the diseased phenotypes. If the alliance uncovers a relationship between the onset of tremors and reported pesticide exposures, Parkinson's disease researchers could stress the donor's iPSC-derived neurons with particular pesticides such as paraquat. If donors report digestive problems or the medical record shows frequent use of feeding tubes, cystic fibrosis researchers could focus on the donor's stem-cell-derived intestinal organoids, rather than iPSC-derived lung airway epithelia cells alone, and

rescue phenotypes in the organoids with gastrointestinal enzymes used to treat symptoms. If speech dysfluencies decrease in audio recordings from donors treated with particular drugs, Fragile X Syndrome researchers could focus on the effects of the drug on neuronal subtypes involved in phonological memory that connect the cerebellar vermis to the frontal lobe.

Our model is consonant with similar contemporary interactive approaches designed to facilitate information flow among donors, researchers, clinicians, bioethicists, and policymakers (Kaye et al., 2012; Saha and Hurlbut, 2011). The “Informed Cohort” model, for example, allows participants to manage their medical data and refine phenotypic data online, contribute additional biomaterials, and communicate regarding research findings (Holm and Taylor, 2012). We recognize that alliance activities may require substantial time and effort from all participants and the technical capacity to sustain an interactive system. Yet, such work may also craft new practices that build trust among all parties, sharpen collected phenotypic

data, and enhance recruitment of highly motivated donors. Overall, our model provides a productive means by which varied types of expertise can be integrated to consider relevant research and governance questions together.

ACKNOWLEDGMENTS

We are grateful to P. Ossorio, B. Hurlbut, and the larger University of Wisconsin’s Stem Cell and Regenerative Medicine Center community for helpful discussions. We also acknowledge generous financial support from the Society in Science Foundation.

WEB RESOURCES

The URLs for data presented herein are as follows:

Boston Children’s Hospital Gene Partnership Informed Cohort Oversight Board, <http://www.genepartnership.org/about-tgp/informed-cohort/>

HeLa Genome Data Access Working Group, <http://acd.od.nih.gov/hlgda.htm>

National Institute of Neurological Disorders and Stroke (NINDS) iPSC Consortia, http://www.ninds.nih.gov/news_and_events/news_articles/NINDS_ipsc_consortia_2011.htm

OHRP: Guidance on Research Involving Coded Private Information or Biological

Specimens, <http://www.hhs.gov/ohrp/policy/cdebiol.html>

REFERENCES

- Hogle, L.F. (2014). *Regenerative Medicine Ethics: Governing Research and Knowledge Practices*. (New York, New York, USA: Springer).
- Holm, I.A., and Taylor, P.L. (2012). *Minn J Law Sci Technol* 13, 669–690.
- Hudson, K.L., and Collins, F.S. (2013). *Nature* 500, 141–142.
- Isasi, R., Andrews, P.W., Baltz, J.M., Bredenoord, A.L., Burton, P., Chiu, I.-M., Hull, S.C., Jung, J.-W., Kurtz, A., Lomax, G., et al. (2014). *Cell Stem Cell* 14, 427–430.
- Kaye, J., Curren, L., Anderson, N., Edwards, K., Fullerton, S.M., Kanellopoulou, N., Lund, D., MacArthur, D.G., Mascalzoni, D., Shepherd, J., et al. (2012). *Nat. Rev. Genet.* 13, 371–376.
- Kreiner, T., and Irion, S. (2013). *Cell Stem Cell* 12, 513–516.
- Malin, B., Loukides, G., Benitez, K., and Clayton, E.W. (2011). *Hum. Genet.* 130, 383–392.
- Rodriguez, L.L., Brooks, L.D., Greenberg, J.H., and Green, E.D. (2013). *Science* 339, 275–276.
- Saha, K., and Hurlbut, J.B. (2011). *Nature* 478, 312–313.
- Saha, K., and Jaenisch, R. (2009). *Cell Stem Cell* 5, 584–595.

Krishanu Saha^{1,2,3,*}
and Linda F. Hogle^{2,3,*}

¹Department of Biomedical Engineering,
University of Wisconsin-Madison, Madison,
WI 53715, USA

²Wisconsin Institute for Discovery,
University of Wisconsin-Madison, Madison,
WI 53715, USA

³Department of Medical History and
Bioethics, University of Wisconsin-Madison,
Madison, WI 53715, USA

*Correspondence: ksaha@wisc.edu (K.S.),
lfhogle@wisc.edu (L.F.H.)

<http://dx.doi.org/10.1016/j.stem.2014.04.012>