Published in final edited form as:

Urol Oncol. 2014 February ; 32(2): 198–201. doi:10.1016/j.urolonc.2013.09.016.

“Use it or lose it” as an alternative approach to protect genetic privacy in personalized medicine

Jennifer K. Wagner, J.D., Ph.D.,

University of Pennsylvania, Center for the Integration of Genetic Healthcare Technologies (CIGHT), 1112 Penn Tower, 399 S. 34th St., Philadelphia, PA 19104, 215-614-0689 (phone), 215-615-5315 (fax)

Jessica T. Mozersky, Ph.D., and

University of Pennsylvania, Center for the Integration of Genetic Healthcare Technologies (CIGHT), 1112 Penn Tower, 399 S. 34th St., Philadelphia, PA 19104, 215-614-0209 (phone), 215-615-5315 (fax)

Reed E. Pyeritz, M.D., Ph.D

University of Pennsylvania, Center for the Integration of Genetic Healthcare Technologies (CIGHT), 11-133 Smilow Center for Translational Research, 3400 Civic Center Blvd., PA 19104, 215-614-0933 (phone), 215-573-8606 (fax)

Jennifer K. Wagner: jennifer.wagner@uphs.upenn.edu; Jessica T. Mozersky: jessica.mozersky@uphs.upenn.edu; Reed E. Pyeritz: reed.pyeritz@uphs.upenn.edu

Scholars have spent considerable time discussing the challenges and importance of protecting privacy in the context of genomic medicine [see, e.g., 1]. There is a great deal of research and clinical potential to be gained from storing massive amounts of genotypic data but this must be weighed against the possible risks to individual privacy, a quandary some have described as “privacy versus the gold-mine” [2]. DNA is a unique identifier and has the potential to reveal medical (and non-medical) information about individuals and by consequence their family members. For instance, genomic data could reveal undesired pre-symptomatic health information, non-paternity, or even be used to frame a suspect at a crime scene [3]. Access to, or disclosure of, this information, whether authorized or not, could lead to various forms of misuse and discrimination (e.g. employment, insurance, and financial). The advent of whole genome sequencing has compounded privacy concerns because it could reveal entirely unanticipated information particularly as our ability to (re)interpret genomic sequence data is continually improving [1; see also 4-6].

In addition, genomic information could have consequences beyond the individual from whom it was derived, including both lineal relatives (e.g., children and grandchildren) and collateral relatives (e.g., siblings, cousins, nieces, and nephews) who may be unaware that the individual is undergoing sequencing. Intra-familial privacy issues are complex and vary from one family to another; however, in the clinical context physicians typically give precedence to individual patient privacy and autonomy over the interests of a patient’s

© 2013 Elsevier Inc. All rights reserved.

Correspondence to: Reed E. Pyeritz, reed.pyeritz@uphs.upenn.edu.

Conflicts of Interest:

The authors have no conflicts of interest to disclose.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

relatives. For example, one patient who is found to carry a mutation in the BRCA1 or 2 genes, which increases the risk of breast and/or ovarian cancer, may not want this information revealed to family members for a variety of reasons whereas another patient receiving the same may choose to share this with all of his/her family members, only specific family members he/she worries may be at heightened risk, or even complete strangers to advance biomedical research [see, e.g., 7]. A patient's care might be improved if the health care provider were able to ascertain genetic information of the relatives as well as that limited information on family medical history known to and shared by a specific patient; however, the idea of linking genomic records between family members or creating joint accounts containing certain genetic data to which multiple patients' records have access has not yet garnered strong support.

A frequent if not ubiquitous assumption in these discussions has been that the integration of whole genome sequencing (WGS)¹ in health care would consist of a one-time sequencing of the patient's genome and subsequent storage of the WGS data in the patient's electronic health record (EHR) [e.g. 8-10]. Those WGS data could then be interpreted and reinterpreted by health care providers over the patient's lifetime (or collaboratively "managed" by patients and clinicians [11]) as knowledge about the clinical relevance of genomic variants increases. One of the biggest clinical challenges to interpretation of an individual's genome currently is the uncertain significance of many DNA variants, although this will likely improve over time. This ongoing need for reinterpretation of an individual's genome has raised additional clinical concerns such as who is responsible for re-contacting patients when and if new information emerges, and how patients can be expected to give informed consent when the potential implications of WGS are unknown currently [12]. With such access to WGS data alongside scientific and technological capabilities for health care providers to mine those WGS data (sometimes for information that exceeds the scope of patient expectations or immediate health concerns), policy discussions have focused on risk management of "the incidentalome" [e.g., 13-14]. Discovering multiple abnormal incidental findings, including the dreaded variants of unknown significance (VUS), could place many undue burdens on clinicians and patients alike.

Discussions that center on incorporating all genomic data into the EHR have the effect of medicalizing the genome [see also 16] by assuming all genomic information is relevant for determining medical risk when some portions of the genome have no known medical relevance whatsoever, despite being useful for non-healthcare purposes (such as ancestral or forensic information). Here, in contemplating patient privacy in personalized medicine, we question the medicalization of genomes in a broad sense, not only questioning the restriction of an individual's access to genomic information by requiring such data to be obtained only through a health care provider, but also questioning the *a priori* medical relevance of all genomic sequence data. While genomic data can, and do, provide clinically relevant information, the entire genomic sequence will not be relevant or necessary in most contexts. Each locus in the genome has its own evolutionary story and an anthropological (not just medical) genetics perspective is necessary. Some specific loci in the genome may be medically relevant for some individuals, in some contexts, and during some stages of development while irrelevant for other individuals, in other contexts, or during other stages of development. Assessing the medical relevance of genomic data is limited by our present understanding of normal human genomic variation, reporting biases of positive results in the literature, and the under-representation of genomic research involving individuals from racial and ethnic minorities. Thus, the determination of medical relevance of genomic data is

¹In this article we refer to whole genome sequencing (WGS) although we acknowledge that whole exome sequencing (WES) is currently often being used. We anticipate that as costs continue to drop whole genome sequencing may become the norm and replace WES entirely.

appropriately a locus-by-locus, patient-by-patient, visit-by-visit, case-by-case decision. We further question the assumption that WGS data should be incorporated into the EHR without deliberate consideration given to privacy and standard of care problems that may accompany the “hoarding” of genomic data in clinical systems. We emphasize that our suggestion relates to the storage of clinical data as opposed to research data, where storage of the entire genome may indeed be necessary.

Proposing a “Use it or Lose it” Strategy for WGS Data in Health Care

The cost to generate a human WGS has fallen rapidly in the last few years [17], and the coveted \$1000 genome may soon be within arm’s reach. Clinical applications of WGS are no longer cost-prohibitive [18]. Moreover, WGS can be done in roughly a day, with the speed of sequencing now 50,000 times faster than it was in 2000 [19]. WGS costs are descending to levels comparable to diagnostic genetic tests [see, e.g., 20]. Clinical uptake of WGS is expected to intensify as the cost of generating a patient’s WGS falls, while the speed of generating and analyzing the WGS increases. Yet privacy concerns are unresolved: that WGS data are themselves identifiers and can be associated with publicly available information relatively easily is gaining broader recognition [21]. At what point does the cost of storage of WGS data locally or on the cloud –and the associated costs (economical, logistical, structural, and otherwise) of properly maintaining data security and protecting the privacy of health information in compliance with the HIPAA/HITECH/GINA Final Rules [22; see also 23-24] –itself become an avoidable act of “raiding the medical commons” by straining already limited health care resources [25]?

Quite simply, the mass default storage of WGS data generated in clinical settings may expose patients, and their family members, to unnecessary privacy risks. Various solutions have been proposed ranging from closed or controlled-access databases [e.g. 26] to entirely publicly accessible databases that provide no privacy protection (such as the Personal Genome Project) [2]. An alternative approach may be to treat WGS as any other diagnostic test, not ordered once per lifetime of a patient (as is often envisioned today) but, rather, ordered as needed for specific medical purposes and to guide clinical intervention. Upon generation of the WGS data ordered for a specific purpose, those raw data from the WGS used as a basis for the medical decisions and course of treatment could be incorporated into the EHR along with the record of the diagnosis and course of treatment. The rest of the WGS data would be discarded. This proposed “use it or lose it” policy has several advantages that could help mitigate the currently contentious issues regarding “privacy versus the gold-mine” [2]. Our “use it or lose it” strategy would limit the genomic privacy risks of the patient by retaining only the genomic information necessary for the patient’s current management. Moreover, such a policy would reduce the ‘incidentalome’ and associated findings such as variants of unknown significance or unwanted knowledge about pre-symptomatic disease or reproductive risk. This would also avoid the potential duty to re-contact patients about an unlimited number of conditions with which their genomic data may become associated in the future. In trying to grapple with these issues, the American College of Medical Genetics and Genomics (ACMG) recently published policy recommendations that 57 particular variants, genes and conditions be reported as incidental findings, regardless of the purpose for which clinical sequencing was initially carried out [14]. This ACMG recommendation has generated controversy, leading some to suggest that such a policy violates patient autonomy by denying individuals the right to opt out of receiving such information [27]. In contrast, our suggested “use it or lose it” policy would facilitate patient autonomy by increasing the transparency for the purpose of WGS in the clinical context and by providing multiple opportunities for a patient to express current preferences for WGS information within each specific context in which WGS data are being considered (conceptually in line with “self-guided management” of sequencing results [11] and in

recognition of intra-individual genomic variability [see, e.g., 28] making it appropriate not merely to *reinterpret* but *resequence* an individual's genome multiple times over a lifetime). This policy would also clarify the standard of care for the physician and institution generating, using, and storing the WGS data. Transparency and clarification of the standard of care would ensure patient and physician expectations are aligned before the WGS is performed, thereby increasing accountability and also potentially reducing legal liabilities.

Context is everything

We propose that privacy interests are conceptualized not only as patient interests in protecting or blocking access to information but also as interests in “contextual integrity” [29]. In other words, we advocate facilitating the appropriate flow of the right information, to the right people, at the right time, and for the right purposes. A “use it or lose it” policy toward WGS of patients may provide an important foundation to establish the appropriate scope of standard of care for genomic analysis and interpretation in clinical settings. Such a policy not only would document/memorialize the quality of WGS methods used for generating that patient's genomic data but also would document/memorialize the temporally-specific biomedical understanding of genomic variants at the time the WGS was interpreted for that patient. In so doing, the policy would highlight the problem-specific motivation for considering the WGS data and may help to explain any selection and confirmation biases that occur. We argue that professional judgment is “indispensible” in genomic medicine [30], thus it seems a policy promoting the transparency of those subjective decisions about what portions of the genome were relevant for that patient under those circumstances, what methods were taken to interpret those selected portions of the genome given current biomedical understanding, and what recommendations or course of treatment is appropriate based on that interpretation would be advantageous and enable accountability for those decisions. Furthermore, the policy would avoid the medicalization of the patient's genome *a priori* and allow the portions of the genome not relied upon for the present medical decisions be kept out of the EHR. While opponents to return of WGS data to research participants often invoke the “therapeutic misconception” [see, e.g., 31], scholars rarely acknowledge that the *a priori* medicalization of the genome implicitly lends credence to misguided notions of genetic determinism that patients and the general public already have. To continue to medicalize the genome indiscriminately not only exposes patients to privacy risks but also exposes health care providers to liability risks widening under “loss of chance” doctrines (i.e., legal theories, often invoked in failure to diagnose or intervene early, that permit individuals to recover for medical malpractice not only for an unfavorable outcome but also for reduced odds of a more favorable outcome) [see, e.g. 32] for health outcomes with even minor genetic components.

Someday never comes

An expected criticism of our proposed policy is the idea of “waste not, want not,” the idea that the entirety of a patient's WGS data should be saved for a rainy day just in case it may find a future medical use. Indeed, it is anticipated that some patients might feel safer and indicate preferences for storing their entire WGS data; however, storage in this manner might in the long-term endanger patient welfare and expose family members, to whom the data relates probabilistically, to unknown or undesired privacy risks. WGS technology has been continually improving in terms of genome coverage and reliability of variant calling. Why should subsequent patient care be limited by the quality of the sequencing technology at the time the patient's genome was first sequenced? The legal standard of care for the practice of medicine is by design flexible to accommodate emerging medical tools and techniques. Genomic medicine will be no different. With this in mind, it seems likely to us that when future instances would arise during which analysis of the patient's genomic

information is perceived as medically relevant, a health care provider would likely order a new WGS even if the patient has already had one done (assuming, of course, that it is affordable and feasible to do so). If such a prediction is true, for what benefit has the storage of the WGS data served to offset the privacy risks associated with that storage? Discarding WGS information could also help to avoid costly technical challenges such as the need for adequate infrastructure and computer storage space that such data will inevitably require.

Concluding remarks

Our proposed “use it or lose it” policy is not without limitations, particularly under today’s current WGS capabilities and constraints. For example, saving only the WGS data relevant for the present medical problem means that other relevant data, some of which may have clinical utility, would be unavailable for emergency uses (e.g. pharmacogenomics purposes for which there may not be time to re-sequence prior to treatment delivery) and for monitoring. Additionally, while WGS may no longer be cost prohibitive, patients and insurers may not be willing to pay for the generation of WGS data that is perceived as unchanging during the patient’s lifetime (notwithstanding exceptions due to somatic mutations, cancers, etc.). Yet as we have outlined in this paper, “hoarding” genomic data is not necessarily cost effective and may come at the expense of patient privacy. This proposal of a “use it or lose it” policy is mentioned not for uncritical adoption but, rather, to encourage deliberate consideration of the pros and cons of storing genomic information for future use in clinical contexts (not research contexts). The assumption that long-term genomic data storage is appropriate should, in our opinion, be questioned. It may be that the clinical hoarding of genomic data not only places patient privacy at risk but also future patient care.

Acknowledgments

Preparation of this article was supported in part by Grant No. P50HG004487-05 and Grant No. K99HG006446-02 from the National Human Genome Research Institute (NHGRI). The content of this article is solely the authors’ responsibility and may not represent the official views of the NHGRI or the University of Pennsylvania. The authors thank Dan Vorhaus, Misha Angrist, Mark DeLong, and Barbara Bernhardt for feedback during the early conceptualization of this article.

References

1. President’s Commission for the Study of Bioethical Issues. [May 10, 2013] Privacy and Progress in Whole Genome Sequencing. Oct. 2012 Available at http://bioethics.gov/sites/default/files/PrivacyProgress508_1.pdf
2. Belmont J, McGuire AL. The futility of genomic counseling: essential role of electronic health records. *Genom Med*. 2009; 1(5):48.10.1186/gm48
3. Brenner S. Be prepared for the big genome leak. *Nature*. 2013; 498:139. [PubMed: 23765454]
4. McGuire AL, et al. Confidentiality, privacy, and security of genetic and genomic test information in electronic health records: points to consider. *Genet Med*. 2008 Jul; 10(7):495–9. doi:10.1097/GIM.0b013e31817a8aaa. [PubMed: 18580687]
5. Lunshof JE, Chadwick R, Vorhaus DB, Church GM. From genetic privacy to open consent. *Nat Rev Genet*. 2008; 5:406–411. [PubMed: 18379574]
6. Kaye J. The tension between data sharing and the protection of privacy in genomics research. *Annu Rev Genomics Hum Genet*. 2012; 13:415–31.10.1146/annurev-genom-082410-101454 [PubMed: 22404490]
7. Kuehn BM. Groups Experiment with Digital Tools for Patient Consent. *JAMA*. 2013; 310(7):678–680.10.1001/jama.2013.194643 [PubMed: 23989528]
8. Ginsburg GS. Realizing the opportunities of genomics in health care. *JAMA*. 309(14):1463–64.10.1001/jama.2013.1465 [PubMed: 23571581]

9. Manolio TA, Chisholm RL, Ozenberger B, et al. Implementing genomic medicine in the clinic: the future is here. *Genet Med.* 2013; 15(4):258–267.10.1038/gim.2012.157 [PubMed: 23306799]
10. Chute CG, Kohane IS. Genomic medicine, health information technology, and patient care. *JAMA.* 2013; 309(14):1467–8.10.1001/jama.2013.1414 [PubMed: 23571583]
11. Yu J-H, Jamal SM, Tabor HK, Bamshad MJ. Self-guided management of exome and whole-genome sequencing results: changing the results return model. *Genet Med.* 2013; 15(9):684–690.10.1038/gim.2013.35 [PubMed: 23619276]
12. Pyeritz RE. The coming explosion in genetic testing – is there a duty to recontact? *N Engl J Med.* 2011; 365(15):1367–9.10.1056/NEJMp1107564 [PubMed: 21995382]
13. Kohane IS, Masys DR, Altman RB. The Incidentalome: A Threat to Genomic Medicine. *JAMA.* 2006; 296(2):212–215.10.1001/jama.296.2.212 [PubMed: 16835427]
14. Green, RC.; Berg, JS.; Grody, WW.; Kalia, SS.; Korf, BR.; Martin, CL., et al. [May 23, 2013] ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing. Available at http://www.acmg.net/docs/ACMG_Releases_Highly-Anticipated_Recommendations_on_Incidental_Findings_in_Clinical_Exome_and_Genome_Sequencing.pdf
15. van El CG, Cornel MC, Borry P, Hastings RJ, Fellmann F, Hodgson SV, et al. Whole-genome sequencing in health care: Recommendations of the European Society of Human Genetics. *Euro J Hum Genet.* 2013; 21:580–584.10.1038/ejhg.2013.46
16. Terry, S.; Cook-Deegan, R. [May 23, 2013] Your Genome Belongs to You. Health Affairs Blog. Jun 8, 2012 Available at <http://healthaffairs.org/blog/2012/06/08/your-genome-belongs-to-you/>
17. Wetterstrand, KA. [May 23, 2013] DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). Available at: www.genome.gov/sequencingcosts
18. Biesecker LG, Burkner W, Kohane I, Plon SE, Zimmern R. Next-Generation sequencing in the clinic: are we ready? *Nat Rev Genet.* 2012; 13(11):818–24.10.1038/nrg3357 [PubMed: 23076269]
19. Venter JC. Multiple personal genomes await. *Nature.* 2010; 464:676–677.10.1038/464676a [PubMed: 20360717]
20. Ball MB, Thakuria JV, Zaranek AW, et al. A public resource facilitating clinical use of genomes. *PNAS.* 2012; 109(30):11920–11927.10.1073/pnas.1201904109 [PubMed: 22797899]
21. Hayden EC. The Genome Hacker. *Nature.* 2013; 497:173–174.
22. DHHS. Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules Under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act; Other Modifications to the HIPAA Rules. Final Rule Fed Reg. Jan 25; 2013 78(17):5566–5702.
23. Solove, D. [August 26, 2013] The HIPAA-HITECH Regulation, the Cloud, and Beyond. Jan 22, 2013 Available at www.safegov.org/2013/1/22/the-hipaa-hitech-regulation,-the-cloud,-and-beyond
24. Pate, S. [August 26, 2013] HIPAA Compliance on the Cloud. Aug 20, 2013 Computerworld. Available at <http://blogs.computerworld.com/cloud-security/22647/can-you-be-hipaa-compliant-cloud>
25. McGuire AL, Burke W. An unwelcome side effect of direct-to-consumer-personal genome testing: raiding the medical commons. 2008; 300(22):2669–71.10.1001/jama.2008.803
26. Greenbaum D, Sboner A, Mu XJ, Gerstein M. Genomics and Privacy: Implications of the New Reality of Closed Data for the Field. *PLoS Computational Biology.* 2011; 7(12):e1002278.10.1371/journal.pcbi.1002278 [PubMed: 22144881]
27. Wolf AM, Annas GJ, Elias S. Patient Autonomy and Incidental Findings in Clinical Genomics. *Science.* 2013; 340:1049–1050. [PubMed: 23686341]
28. Zimmer, C. DNA Double Take. *NY Times:* Sep 17, 2013 Available at www.nytimes.com/2013/09/17/science/dna-double-take.html?pagewanted=all&_r=0 [September 18, 2013]
29. Nissenbaum, HF. Privacy in context: technology, policy, and the integrity of social life. Stanford, CA: Stanford Law Books; 2010.
30. McGuire AL, McCullough LB, Evans JP. The indispensable role of professional judgment in genomic medicine. *JAMA.* 2013; 309(14):1465–6.10.1001/jama.2013.1438 [PubMed: 23571582]

31. Angrist M. You never call, you never write: why return of 'omic' results to research participants is both a good idea and a moral imperative. *Per Med.* 2011; 8(6):651–657.10.2217/pme.11.62 [PubMed: 22199990]
32. Weigand TA. Lost Chances, Felt Necessities, and the Tale of Two Cities. *Suffolk Univ Law Rev.* 2010; 43:327–392. 373.