SJ Quinney College of Law, University of Utah

Utah Law Digital Commons

Utah Law Faculty Scholarship

Utah Law Scholarship

11-2019

Frontiers in Precision Medicine IV: Artificial Intelligence, Assembling Large Cohorts, and the Population Data Revolution

Adam Bress

Rich Albrechtsen

Monika Baker

Jorge L. Contreras

Zachary Fica

See next page for additional authors

Follow this and additional works at: https://dc.law.utah.edu/scholarship

Part of the Health Law and Policy Commons, and the Science and Technology Law Commons

Authors

Adam Bress, Rich Albrechtsen, Monika Baker, Jorge L. Contreras, Zachary Fica, Austin Gamblin, Chelsea Ratcliff, Bianca E. Rich, Matt A. Szaniawski, Alyssa Thorman, Chad VanSant-Webb, and Willard Dere

Frontiers in Precision Medicine IV: Artificial Intelligence, Assembling Large Cohorts, and the Population Data Revolution

Adam Bress[1], Richard Albrechtsen[2], Monika Baker[2], Jorge L. Contreras[3], Zachary Fica[2], Austin Gamblin[2], Chelsea Ratcliff[2], Bianca E. Rich[2], Matt A. Szaniawski[2], Alyssa Thorman[2], Chad VanSant-Webb[2], Willard Dere[2]

Affiliations:
[1]University of Utah Health, [2]University of Utah – School of Medicine, [3]University of Utah - S.J. Quinney College of Law, Salt Lake City, UT.

Abstract:
Large cohort studies and more recently electronic medical records (EMR) are being used to collect massive amounts of genetic information. Implementation of artificial intelligence has become increasingly necessary to interpret this data with the goal of augmenting patient care. While it is impossible to predict what the future holds, policy makers are challenged to create guiding principles and responsibly roll out these new technologies. On March 22, 2019, the University of Utah hosted its fourth annual Precision Medicine Symposium focusing on artificial intelligence, assembling large cohorts, and the population data revolution. The symposium brought together experts in medicine, science, law and ethics to discuss and debate these emerging issues. A link to the full conference recording can be found at
http://www.kaltura.com/index.php/extwidget/preview/partner_id/670542/uiconf_id/36005582/embed/auto?&flashvars%5BstreamerType%5D=auto&flashvars%5BplaylistAPI.kpl0Id%5D=0_yfqter8f

Opening Keynote Session: Assembling Large Cohorts for Precision Medicine

Genetic information continues to be collected at a rate that exceeds our current ability to interpret and translate these finding into meaningful healthcare solutions. Yet, paradoxically the continued growth of biobanks and cohort studies provides the necessary power and diversity to elucidate potential pathophysiology and pharmacologic targets of complex polygenic diseases, thus accelerating the translational processes.

Dr. Lon Cardon, a statistical geneticist and Chief Scientific Officer at BioMarin Pharmaceutical, argues that the benefits of Precision Medicine accrue with the translation of genetic variants into better therapies for patients with both rare and complex common diseases. As with new technologies, the broader field of Precision Medicine is impacted by the "hype cycle," where a newly invented technology does not initially deliver on the ambitions and expectations, which is followed by an exodus of researchers and interest. Eventually the technology confers benefits and becomes profitable. One of the main challenges that initially faced precision medicine is the lag between genetic discoveries such as the cloning of *CFTR* in cystic fibrosis in 1989 and the clinical translation into medicines almost 30 years later. Other well-known genetic disorders such as Duchenne muscular dystrophy and phenylketonuria (PKU) were also discovered decades ago, but clinical applications are only now coming into fruition. With regards to the

1

"hype cycle," precision medicine is rebounding, starting by delivering targeted therapies in oncology and more recently, for rare and common diseases.

The discovery process has been rapidly accelerating, with respect to rare monogenic diseases. What used to take years is now being discovered at a rate of three per week, largely due to new sequencing technologies and initiatives such as the Rare and Undiagnosed Disease Network.[1] In conjunction with the increase in gene discoveries by researchers, the Food and Drug Administration (FDA), has responded in kind, by increasing the number of approvals of orphan drugs for rare diseases. There was a record number of approvals in 2015 and continues to increase.[2] Even with the increase in orphan drugs, there are still plenty of new discoveries that are waiting for a clinical application. However, discoveries surrounding common complex diseases such as type II diabetes mellitus (T2DM) or coronary artery disease (CAD) are not progressing at the same rate. Instead of using family pedigrees with probands to identify genetic causes of monogenic disorders, common complex genetic disorders require cohorts with large sample sizes to identify loci of interest. Dr. Cardon introduced the Wellcome Trust Case Control Consortium (WTCCC) as an example of how a large cohort was assembled to investigate seven common complex diseases using GAWS with 2,000 cases to 3,000 controls.[3] Prior to the WTCCC there were only 10 genes identified for common complex diseases, which the WTCCC doubled to 20. In a little more than a decade later, there are now over 6,000 genes with variants identified for common complex diseases.[4] The plethora of data is overwhelming, and the field rapidly progressed from having too few to too many targets in such a short period. Specifically, with regards to inflammatory bowel disease, associations with over 200 loci have been identified. These findings could be paralytic, and the question remains: which are optimal drug targets for new therapies?

When considering the intersection of genetic variants and drug effects, Dr. Cardon highlighted that genetic effect size does not equal drug effect size. For example, in the setting of lowering the risk of cardiovascular disease, two specific genes *HMGCR* and *PCSK9* can be compared. The loss-of-function of *PCSK9* has a dramatic effect in lowering risk of CAD; by contrast, much more modest salutary effects are seen with the variant rs12916 in *HMGCR*. However, when comparing an *HMGCR* inhibitor such as rosuvastatin to the *PCSK9* inhibitor alirocumab in randomized controlled trials, the treatment effects on lowering LDL cholesterol appear to be comparable. In relating the example of *HMGCR* and *PCSK9* to the more than 6,000 loci identified for common complex diseases, the genetic effects of any one locus may seem inconsequential, but targeting those genes could have a dramatic drug effect, thus each locus could be considered as a serious target. In the drug discovery domain, drugs that have an underlying genetic basis have double the clinical success rate and is the best predictor of success. Furthermore, the success rate increases to ten-fold when considering a rare disorder, due to their large genetic effects that are monogenic in nature.[5] Until only recently, the dogma was to discover and target the "drug-able genome" or the idea that only certain genes were amenable or targetable by chemical or biological approaches. This is no longer the case with the advent of oligonucleotides, modified RNA, or gene-based therapies, as we can now replace or modify almost any gene in the genome. Gene editing has also completely transformed basic research by allowing researchers to create mutations of interest, instead of painstakingly sampling populations in order to find a mutation of interest, thereby saving a significant amount of time and money.

With the genetic information and the modalities to implement them, the third component is getting patients into large cohort studies, which is key to unlocking their full potential. Assembling all of the pieces together highlights the complexity of modern drug discovery in an ocean full of opportunities with no clear place to start. According to Dr. Cardon, one of the best

2

places to start, is to utilize rare genetic disorders that are "experiments of nature." In returning to the example of *PCSK9*, individuals with loss-of-function alleles of *PCSK9* are at lower risks of CAD with lower levels of LDL cholesterol, as compared to those with gain-of-function alleles who have a higher risk of CAD and higher levels of LDL. Thus, human genetics can mimic therapeutic dose-response curves and is highly attractive for proof of mechanism, efficacy, and safety. These experiments of nature can lower the risk of the drug discovery process and even reduce or eliminate the need for basic research, but currently there are relative few excellent examples, making each highly desirable if observed. When considering a complex phenotype such as height, one locus was identified in 2007[6] whereas a total of 712 loci had been identified by 2018[7]. However, only four loci—FGFR3, CNP (a modulator of FGFR3), NPRB (a receptor for CNP), SHOX (part of the FGFR3 signaling pathway)—have the properties of an ideal drug target based on genetic data. Their attractiveness is due to the following observation: gain-or loss-of-function variants in these four loci are associated with greatly increased or reduced height, thus it seems that this particularly FGFR3 pathway seems to contain potentially valuable drug targets.

The most effective way to identify these experiments of nature is through the use of large, diverse cohorts with a broad reach that not only include patients of the phenotype of interest, but of unrelated phenotypes as well. As a member of the advisory board, Dr. Cardon highlighted how the *All of Us* Precision Medicine initiative is one of the largest to date with a goal of enrolling one million participants, is patient oriented and driven, and will be a catalyst for positive change in medical research. *All of Us* was built upon lessons learned from other large cohorts, by emphasizing size, genetic and phenotypic diversity, and will have active data collection through tools such as the use of wearables. Two other examples of large cohorts provided by Dr. Cardon, *BiobankUK* and *Genomics England* were used in designing *All of Us*.

BiobankUK enrolled 500,000 healthy people ages 40 to 69 to study the effects of genetic predisposition and environmental exposure to the development of new disease. Data from health records and physical measures such as lab values, vital signs, and electrocardiograms are recorded a passive way, requiring little effort from participants and physicians. Since its launch in 2012, the process aims to be transparent as the data are open and available to all, with over 400 peer-reviewed papers having been published to date. To further emphasize the power of the large cohort that *BiobankUK* provides, Dr. Cardon provided the example of the discovery a glucose-lowering variant in *GLPR1* that is protective for CAD.[8] The discovery process, which took five years and over 1,500 emails to assemble data from 50 cohorts, was replicated and validated by *BiobankUK* in 90 minutes and three emails. Not only have large cohorts been transformational in identifying new gene discoveries but have also led to the creation the next big thing: polygenic risk scores, which evaluates the combined effect of multiple loci for common complex diseases. Previously, the predictive value of risk for given loci for T2DM was approximately equivalent to that of body mass index, as each locus individually has a minimal impact on disease risk. However, when considering the unique combination of loci as part of the polygenic risk score, the patients at the tails of polygenic risk score have a predictive value that is equivalent to having one pathogenic allele in a monogenic Mendelian disorder.[9]

Genomics England, a more recent example of a large cohort program, has almost 100,000 patients and focuses on rare diseases, having specifically oversampled undiagnosed neurological diseases and cognitive deficits. Investigators using the *Genomics England* cohort have the potential to identify patients—such as PKU or lysosomal storage disease-- who are either un- or misdiagnosed and for whom an existing treatment is readily available. An unexpected route to Precision Medicine identified by *Genomics England* is the identification of

3

incidental findings that will help characterize parents who are asymptomatic or probands for whom a genetic alteration was not identified. The example provided are mothers who were sequenced together with their affected child(ren), only to find that the child did not carry genetic mutation(s) for PKU, but instead were exposed to elevated levels of phenylalanine in utero from asymptomatic mothers with undiagnosed maternal PKU. Dr. Cardon cautioned that while large cohorts have been transformational in genetic research, families still have to be involved as these cases of PKU would not have been solved if the mother had not been included. The United Kingdom wants to take the next step forward by sequencing the genomes of five million newborns in the next five years.[10] In addition to the work in the United Kingdom, the International Cohorts Summit, held in March 2018, looked to create minimum criteria for large cohorts with at least 100,000 healthy participants who are available for biospecimen collection and longitudinal follow-up. Overall, 60 cohort studies that include 25 million people (with expansion to 36 million) in 24 countries were identified.[11] The availability of data from these studies offer exciting possibilities, but the challenge remains: How can we translate these data into meaningful therapies?

Dr. Cardon concluded that no single population-based collection is enough. There is always a need for more size, diversity, phenotypic breadth, and depth. One way to accomplish this is to have greater investment in non-population-based samples such as including large families, founder populations, and consanguineous pedigrees, which can clarify and amplify genetic signals. He emphasized the need for better harmonization between clinical and phenotypic data and genomic data, which can be particularly problematic when trying to combine multiple large cohorts across healthcare systems. Investing in computational infrastructure and data analysis tools are important ingredients for success. Through the utilization of genetic information, modalities to implement them, and getting patients into large cohort studies, the opportunities are endless, where using experiments of nature to study rare diseases with a populational context leads to insights into common complex disease leading to translation of new diagnostics and therapeutics.

Session 1: ELSI Issues in Assembling Large Cohorts for Precision Medicine

Researchers and clinicians alike should be aware of the ethical concerns in both the methods used to collect genetic information, but also implementing treatments based on genetic information in the absence of thorough physiological understanding or in lieu of the current standard of care. From the research perspective, this relies on transparency in cohort study recruitment methods; from the clinician perspective, this relies on integration of patient preferences and understanding of treatments. In order to achieved transparency and patient understanding, communication methods should be tailored to the audience and account of discrepancies in health care literacy.

    A.   Assembling Large, Longitudinal Cohorts: Ethical Considerations

Dr. James Tabery is an Associate Professor of Philosophy at the University of Utah. He studies the philosophy of science and applied ethics. Dr. Tabery addressed the ethical concerns in assembling large longitudinal cohorts for genetics research. The session started with a viewing of the *All of Us* Anthem[12], a promotional video put together by the NIH *All of Us* research program.[13] The goal of the program is to collect data from one million Americans and create a research database that can help answer important health questions and create more personalized care. The program has an emphasis on genetic data but recruiting material

4

emphasizes lifestyle, environmental factors, and family history. In the recruiting video, there is no specific mention of genetics, DNA, or other biomedical information.

Historically, longitudinal cohorts have come in many different forms. Though the *All of Us* program and its British counterpart the UK biobank, focus on genetics with hundreds of thousands of participants, many previous programs have instead focused on environmental and lifestyle components in smaller cohorts in the thousands. One of the first longitudinal cohorts, The Framingham Heart Study, began in 1948 and started by recruiting 5,209 adults from the town of Framingham Massachusetts[14]. The study followed these participants as well as their children, grandchildren, and spouses ever since its inception, collecting information that has led to the identification of cardiovascular risk factors. It is still going on today and is considered the gold standard of longitudinal cohorts. Other cohorts have taken a similar approach such as the Dunedin longitudinal study[15] that followed the lives of 1037 babies born 1972-1973 at Queen Mary Maternity Hospital, in Dunedin, New Zealand and National Children's Study that attempted to follow 100,00 children from before birth to after age 21 across the United States.

When starting a large longitudinal cohort many design decisions must be made. How a study is designed is directly reflective of the goals of the study. Decisions about the size, the types and amount of data collected, time frame, and types of participants included all must be made. Certain types of data sets lend themselves better to different types of research and the way in which you design a study dictates how that data will be used. Due to the scarce resources available for biomedical research, such projects must also consider ethical implications of who benefits from the research, if it is the best use of scarce biomedical research resources, and if the way you are promoting and advertising your study is in alignment with the actual goals of the study.

Dr. Tabery proposed that sufficient ethical consideration is met when there is alignment of study design with clear communication of study goals to potential participants. In medicine consent is paramount and if you are misrepresenting the goals of a research project to its participants this is a concerning breach of trust and ethics. Furthermore, when research resources are limited it is unethical to represent a study as filling a need and addressing issues that it does not. Though *All of Us* advertises itself as a lifestyle, environmental, and genetics research tool it is primarily designed for genetics research[13]. For example, the study has 3 genome sequencing centers and only a genomics working group, and prioritizes collection of DNA samples via blood but not longitudinal, accurate lifestyle or environmental data at this time. Though non-genomic facets of the study are promised later, *All of Us* appears to be primarily a genetics research project. Such is highlighted in the data collection methods, the enrollment materials, and the recruitment and advertising materials.

Longitudinal cohorts come in many different designs with many different goals, all with their own merits. The value of each type of study and its design can and should be debated to best direct limited funds. This is not possible when a project is incompletely forthright in its intent and has some misalignment between the design of this cohort and the stated recruitment goals.

   B.   Crossing the Finish Line: Translating Precision Medicine to the Clinic

Dr. Kyle Brothers, a clinician from the University of Louisville, addressed the difficulty of translating and further implementing precision medicine into the clinic. Focusing on the final steps that must be taken to bridge the gap between development of new, tailored treatments and the administration of these therapies to their particular target subpopulations.

5

Dr. Brothers likened Precision Medicine to the theories of Alan Turing, a mathematician and early pioneer in the realm of computing who is widely known for his contributions to decrypting German messages in Project Ultra during World War II. The device responsible for encrypting sensitive data was known as the Enigma cipher machine, which takes information (input) and encrypts it (output). In this analogy, the goal of Precision Medicine is to take the output (a disease state or phenotype) and determine what is required (medications, interventions, etc.) to revert it to its original message (a state of health). In classical biomedical science, Dr. Brothers argues, we are attempting to use the human body to understand mechanisms that lead to disease. This is contrasted with Turing's later career, where he attempts to determine whether, without understanding the human mind (the "Black Box"), we can contrive a model capable of simulating it. In summary, Dr. Brothers argues that the two approaches to Precision Medicine mirror the two stages of Turing's career – the first, gaining insights based on a mechanistic understanding of the model, and the second, gaining insights independent of a mechanistic understanding of the model.

This discussion led to the following question: Can non-mechanistic (so-called "Black Box") Precision Medicine insights be used in the rollout of a clinical plan? That is, can we make clinical decisions, such as employing a specific treatment, without a complete understanding of how they function? To dissect this issue, Dr. Brothers turned to the topic of Machine Learning. Companies such as Google and Amazon represent effective examples of machine learning that rely on previous user behavior to prioritize search hits or make recommendations for purchase, respectively. However, these Machine Learning models only function effectively because there is an element of human choice.

For a machine learning algorithm to be useful in medicine and employed as part of the clinical decision-making tree, Dr. Brothers argues that clinicians and patients need enough information to decide whether to follow a Precision Medicine ("Black Box") recommendation, similar to how a decision is required whether to purchase an Amazon-recommended product. In such cases, the Machine Learning algorithm needs to be transparent enough to both the provider and the patient to make a good clinical judgement.

Two approaches were outlined to deal with this gap. The first involves the use of "Black Box insights" to inform mechanistic basic science. From here, mechanisms are understood prior to implementation resulting in care based on mechanistic reasoning. Here, so-called "Big Data" is parlayed into mechanistic understanding. The second approach involves an Evidence Based Medicine approach in which non-mechanistic insights are used to identify subpopulations and precision treatments, followed by comparative effectiveness research that ultimately gives rise to a new standard of care. However, there are issues inherent in Comparative Effectiveness Research trials in Precision Medicine. These include dealing with rare subgroups for which design of clinical trials is difficult, as well as carry-over, order, and irreversibility effects for would-be enrollees who have failed a standard (or other previous) therapy. Finally, Dr. Brothers touches on the concept and challenge of equipoise as it relates to Comparative Effectiveness Research. Can standard of care be withheld in order to study precision therapies? Can precision therapies be withheld in cases where a patient is known to have a specific genotype or disease characteristic that is known to fail standard therapy? These questions further address the critical challenges of Precision Medicine, which involve its transformation from theoretical tool to authentic therapeutics or interventions with substantive clinical utility.

6

C. Communication Issues in Recruitment to Large Cohorts for Precision Medicine

Kimberly Kaphingst ScD, is a Professor in the Department of Communication and the Director of Cancer Communication Research at the Huntsman Cancer Institute (HCI) at the University of Utah. She discussed cohort recruitment from a communication perspective. She noted that *health literacy* is a useful framework for thinking about communication skills that individuals will need in order to navigate this environment. Health literacy is "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions."[16] This skillset will play a key role in people's ability to understand information they encounter during cohort recruitment and consent processes and, ultimately, make an informed decision about whether to participate.

Health literacy levels vary widely in the US population. To illustrate how this applies to precision medicine cohort recruitment, Dr. Kaphingst highlighted data from several studies. The first, a large national assessment conducted by the U.S. Department of Education, showed limited levels of health literacy among many U.S. adults.[17] Roughly 36% of U.S. adults had *basic* or *below basic* levels of health literacy, meaning they are likely to struggle with many tasks in the health care system. About half of adults (53%) had *intermediate* health literacy skills, and only 12% were considered to be *proficient*—that is, to possess the skills required to understand complex presentations of abstract or unfamiliar subjects, such as consent forms.

These data suggest that making informed decisions about whether to join a precision medicine cohort may be a challenge for many adults. However, it may be especially challenging for certain population subgroups. Dr. Kaphingst pointed out that individuals from underserved communities, such as Black/African American, Hispanic/Latino, or American Indian/Alaska Native individuals, are more likely than those from other racial and ethnic groups to have limited health literacy. Older age and lower educational attainment also predict limited health literacy.

Knowledge levels specific to precision medicine and genetics may also be relatively low in the U.S. Dr. Kaphingst described the results of a recent study, which found that among individuals with adequate health literacy, only one-third were familiar with the term "genomics", one-fourth with the term "precision medicine", and one-fifth with the term "biobank."[18] Familiarity with these terms was markedly lower for people with limited health literacy. Further, among those with lower health literacy, only half indicated familiarity with the term "DNA" and one-third with "genetic testing" and "gene." Hence, a substantial portion of the population might not understand the implications of participating in precision medicine cohort studies.

Attitudes toward precision medicine research appear to vary across population subgroups. Whereas most U.S. adults surveyed by the National Human Genome Research Institute in 2016 were favorable toward the *All of Us* research, this study found a notable education-related gap in responses.[19] Those with lower educational attainment (i.e., a high school diploma or less) were less supportive of the research and were also less willing to participate than those who had attended college.

As a final example, Dr. Kaphingst described possible racial differences in people's willingness to share biospecimens based on findings that emerged in her own research. In one study, Black men in a prostate cancer cohort were somewhat more likely than White men to consent to linkage of their biospecimens with their personal health information.[20] However, Black participants were also less likely to feel comfortable sharing their data at all, even when it would be de-identified. Notably, overall willingness to consent to the use of biospecimens with linked

7

personal health information was low for both Black and White men.

Several factors are likely to influence recruitment for *All of Us* and other large population-based cohorts, and many of these factors are connected to health literacy. Thus, Dr. Kaphingst stressed that communication approaches that facilitate understanding across the spectrum of health literacy levels will be essential to consider. A lack of understanding could make individuals less likely to participate in research cohorts—or, conversely, more likely to do so— and either outcome has considerable ethical and social implications.

Dr. Kaphingst concluded by highlighting research-based communication strategies that can support informed decisions about precision medicine cohort participation. The first is adopting a *plain language* approach when developing recruitment and consent materials; namely, using common words instead of jargon and defining technical terms. For instance, the idea of biobanks—for which participants are often asked to provide consent to unknown future use of their data—is a difficult concept for many people to grasp.

Dr. Kaphingst emphasized the importance of testing materials with target audiences and undergoing multiple rounds of revision as needed. In addition to clear language, evidence-based formats can go a long way toward facilitating understanding of complex concepts. Rather than relying on text-based materials, conveying information via audiovisual and in-person modes (e.g., healthcare providers or community liaisons) can also support understanding about research participation across literacy levels. Dr. Kaphingst noted that interactive digital tools embedded in patient portals could be an effective way to help walk people through the processes of recruitment and informed consent. Finally, developing communication materials targeted to different population groups (e.g., based on race/ethnicity, age, or health literacy level) could be an effective strategy for ensuring that cohorts include diverse groups and are representative of the U.S. population.

Session 2: The Role Of AI/Deep Learning in Genetic Analyses and Phenotyping of Large Cohorts

Large genetic databanks are collected through above mentioned cohort studies, but additional largely-untapped data forms exist within all electronic health records (EHR). Currently, various groups are testing machine learning models that utilize information available in the EHR to augment patient care. As Implementing AI in healthcare becomes a reality along with it comes interest from big-named corporations. While it is impossible to predict what the future holds policy makers are challenged to band together to create guiding principles and responsibly rollout these new technologies.

   A.   Machine-Readable Phenotype Descriptions and the Utah Genome Project.

Dr. Mark Yandell, the co-director of the USTAR Center for Genetic Discovery at the University of Utah, opened the final session with a talk that encompassed the importance and the process of readying electronic health records, or EHR, for machine learning.

He began with the important distinction that needs to be made between the terms "controlled vocabulary", which is a set of terms used to describe data, and ontology, which goes a step further and includes relationships between these terms. This extra step makes the data more useful when analyzed by a computer. One specific ontology, the Human Phenotype Ontology (HPO), describes patient phenotypes for the purpose of genetic analysis. It includes free-text

descriptions that a human can understand, as well as machine-readable phenotype descriptions that a computer can use.

Now that we have phenotypic descriptions of data that a computer can read, we can assign these phenotypes to specific genes using a program developed here at the University of Utah called Phevor (Phenotype Driven Variant Ontological Re-ranking Tool). This approach has many advantages, including the ability to detect new genes associated with a phenotype, the ability to be used on large sets of data, and the lack of requirement for a candidate gene list for each specific phenotype.

So how does all this relate to the Utah Genome Project? The University of Utah Hospitals and Clinics have a real-time database of the information captured in the EHR called the Enterprise Data Warehouse, with over 3.7 million individuals and 14 years of data. Some of the most valuable but untouched information in the EHR are in the physician's free-text medical notes for each individual patient. If we converted this free-text data into machine-readable text (like HPO terms), we could analyze this data with their corresponding genomic sequences. Practice runs have indicated that it would take 6 hours on 64 core processors to run every clinic note in the data warehouse. This would have real world implications for identifying rare disease phenotypes and potential targets for new treatment possibilities.

   B.   <u>White Coat, Black Box: Augmenting Clinical Care with AI in the Era of Deep Learning</u>

Dr. Jenna Wiens, is the Morris Wellman Assistant Professor of Computer Science and Engineering (CSE) at the University of Michigan in Ann Arbor. Dr. Wiens described current challenges and ongoing research in the use of deep and representative learning in the augmentation of clinical care. The context for Dr. Wiens research is that in today's world of clinical care, massive amounts of patient data are collected including demographics, diagnoses, medications, imaging, clinician notes and much more. Projected physician shortages, increasing clinician burnout, and lack of proper clinical tools are contributing to a growing number of medical errors, now the third leading cause of death in the U.S. There is a substantial need and opportunity for artificial intelligence (AI) to help overcome these obstacles.

Dr. Wiens and her team are currently validating AI models which construct patient risk stratification algorithms to help predict the occurrence of hospital acquired infections such as *Clostridium difficile*. These models can leverage existing electronic health record (EHR) data points such as medication lists, procedures, rooms assignments, assigned healthcare staff, lab results, and even proximity to other infected patients to predict infection occurrence. Current models show promise in identifying at-risk patients 5 days in advance of when detectable signs and symptoms may appear. Using these predictive models could guide preventative clinical interventions such as targeted cleaning or infection testing. These interventions could decrease length of stay, reduce costs, and improve patient outcomes. Similar models are being designed to predict disease trajectories of severe disease like acute respiratory distress syndrome (ARDS), graft versus host disease (GvHD), sepsis and shock, acute renal failure, and other life-threatening diseases encountered in the hospital setting. However, the development of these AI models presents significant technical challenges that can be alleviated by the use of representation learning and interpretable models.

Representation learning is an important strategy employed in the development of deep learning models. Typically, new training models rely on hand-engineered variables for representing data

9

to the machine learning-based prediction model. For example, heartbeat tracings may be represented to a model based on peak amplitude or interval periods between heart beats. The designation of data representing the tracings is critical for how the predictive model learns to distinguish between different heart beats. Choosing the wrong representation may lead to difficulty in identifying pathologic heart rhythms. These hand-designed representations can be labor-intensive and error-prone. In contrast, representation learning is the process by which the model utilizes deep neural networks to independently represent data sets without user-derived input. These techniques eliminate the need for time-intensive representation and have led to breakthroughs in fields such as image recognition and translation, with use in identifying skin lesions or compiling pathology reports. Limited implementation of these tools in healthcare is largely due to technical challenges due to small sample sizes and model interpretability, with the latter considered here.

A model is highly interpretable when the representations are hand-engineered. In this case, the underlying thought-process of the model is known and understood by researchers. However, with representation learning, the properties of the of the neural network algorithms are unknown. It is important to be able to query the model to learn what factors were considered significant and then correlate those items with clinical knowledge. Models will lack common sense, even if they are interpretable. Thus, interpretability is neither necessary nor sufficient and must integrate clinical knowledge in model design. In the case of *C. difficile* infection prediction, each significant variable identified by the prediction model is assigned a weighted binary designation depending on any known association with *C difficile* infections. This regularization penalty favors a solution that is sparse in the set of unknown variables and dense in the set of known variables. These and similar techniques can lead to models that are not only data-robust but clinically credible.

While these tools are criticized as being unknowable "black-box" models, Dr. Wiens suggests that similar tools are already common in clinical practice. For example, a physician may not know the physics behind the function of a digital thermometer, but they are used every day because they are thoroughly validated and trusted. If the model is meticulously vetted and validated, it can certainly be trusted even if the underlying mechanisms are not understood by clinicians.

Before these models are tested in expensive clinical trials it is critical that they are prospectively validated. The current model for predicting *C.difficile* infections is being prospectively validated in silent mode on daily live streams of patient data. This validation enables real time review of errors which provide crucial clinical context. These processes help shed light on things that should be included or excluded in the model. The evaluation of interventional decision-making and patient outcomes will also be a critical component of future validation.

Prospective validation of similar models will require infrastructure to integrate patient data with artificial intelligence. As they enter clinical practice, these machine learning-based prediction models should not be feared as black boxes but as useful tools developed and interpreted alongside the clinical knowledge of domain experts. Such models have the potential to reduce healthcare costs and medical errors, thereby improving clinical outcomes.

    C. <u>Precision Health, AI and Ethics</u>

Steve Mutkoski, JD is a government affairs director at Microsoft focusing on healthcare policy and technology issues. He described some of the intersections, limitations, and ethical concerns

inherent in emerging AI technologies for healthcare applications. His background as a Microsoft lawyer specializing in public policy give him insight through Microsoft's cloud computing services, Azure, which currently stores data and runs computing workloads for a number of private and public sector healthcare organizations. He stressed that with each new wave of technology comes tremendous opportunity but also a tremendous responsibility to make sure that it is utilized in ways that benefit society and avoid pitfalls that can undermine trust in new technology. He stressed that public trust is something that can be easy to lose if those in the precision medicine world aren't careful in how they develop and deploy AI systems. Silicon Valley's policy of moving fast and breaking things might not be the right approach with this emerging technology for healthcare applications. As a warning of what might happen, Mr. Mutkoski pointed to the recent loss of trust in autonomous vehicles after a highly-publicized 2018 autonomous vehicle crash that resulted in a pedestrian fatality in Arizona.

Public trust in new technologies can be undermined by overpromising and misinformation. For example, 23andMe's original panel of 250 diagnostic tests for genetic health may have promised too much. As a result, the FDA required 23andMe to stop offering health-related genetic testing until it could meet reasonable scientific standards. When 23andMe re-released its health-related testing offering (after substantial engagement with the FDA), it had far fewer diagnostic options. 23andMe's initial enthusiasm for giving the public information may have created unrealistic expectations in the science, and the FDA had to step in to prevent potential harm.

A second example is offered by an article by a CBC reporter who sent her and her twin sister's DNA to be analyzed by five different genome sequencing companies. Each company reported a slightly different genetic profile. When asked, one geneticist responded that genetic testing of this nature was as much an art as a science. Mr. Mutkoski disagrees, arguing that genetics is a science, and needs to protect itself against bad headlines like these in order to prevent a loss of trust.

Mr. Mutkoski suggested two avenues to help prevent such a loss of trust. First, acquiring patient input will better get them on board, comfortable and aware of new technologies. Patient education should be a large part of preventing adverse public opinion. Second, he encouraged active engagement with policymakers to ensure that policymakers understand the technologies we are using and the often complex healthcare related solutions we are building on that technology.

Mr. Mutkoski recently led a group in writing an article about the potential of AI and Precision Healthcare published in the Fall 2018 issue of *The SciTech Lawyer*, published by the American Bar Association's Section of Science and Technology Law.[29] This article suggests a series of principles for the responsible development of AI in the healthcare sector today: (a) Reliability and Safety, (b) Fairness, Inclusion and Bias, (c) Transparency and Accountability, (d) Privacy and Security and (e) Future Workforce Issues.

Data Security and Privacy Issues: It can be difficult to predict some of the ways with which we will use genetic data. Similar problems can emerge under different names and labels on it while new problems can seemingly emerge which make those involved look back and say "Why didn't we see this coming and prepare for it?"

One example is law enforcement's use of genetic databases as a powerful tool to solve crimes, as seen in the Golden State Killer case. While Mr. Mutkoski believes that particular use largely has society's stamp of approval, the realization that one person's genetic information can help

11

identify very distant relatives, who in no way consented to the use of this "familial" genetic data, creates some tension with expectations around the concept of notice and consent.

Another example is the *All of Us* program, an ambitious program to create a "biobank" of one million genetic donors in the United States. One concern raised with the All of Us program is what limits would be placed on third party access to donor data, either through civil or criminal processes?   Mr. Mutkoski pointed to the experience that large global cloud computing providers have had with the data they hold for customers, which in essence has create a series of "honeypots" for personal data that civil litigants and law enforcement authorities have sought to obtain for use as evidence in civil or criminal matters.

Some believe that the recently-enacted 21st Century Cures Act addresses these problems about third party access thought civil or criminal legal process. Mr. Mutkoski, however, argues that issues likely remain. As a federal statute, it is unclear that the Cures Act provisions would limit State courts from allowing discovery of such data for a variety of purposes. Other relevant legislation such as the 2008 Genetic Information Nondiscrimination Act (GINA) prevents certain discriminatory uses of genetic data (for instance, health insurance companies are precluded from using genetic data to deny coverage and prohibits an employer from using genetic information for decisions like hiring and firing) but other potentially problematic and discriminatory uses are not prohibited (long term disability insurance and life insurance are not covered by GINA).

The databases and biobanks themselves are not the problems. Rather, we as a society need to agree on a set of policies regarding the appropriate use of health data and by extension highlight as many of the "prohibited" uses as possible. It may be impossible to predict how this data will be used, as it would have been impossible for Henrietta Lacks to have been given informed consent for all of the future uses of her cells. However, Mr. Mutkoski cautions, just because we cannot predict them all, does not mean that we should not be on the lookout for potential harms.

Fairness, Inclusion and Bias: Mr. Mutkoski pointed out two potential sources of bias in the development of AI-enabled precision health solutions. First, there are some challenges with data itself, specifically much of the genomic testing and study to date has been performed on Caucasian populations. Second, often the scientists driving these studies are a largely homogeneous group of 20-40-year-old white males. As the old adage goes, the invention often reflects the inventor. Diversity and inclusion initiatives are being undertaken to combat this but will take time to change the current paradigm in research.

Safety and Efficacy: Mr. Mutkoski spoke briefly about the need for well-designed trials to validate the safety and efficacy of AI tools in healthcare. He cited several recent news articles claiming that AI is more efficient and safer than current technologies, and that AI can in some cases detect cancer more accurately than pathologists. In many cases though, a deeper inspection of the claims suggests that the AI algorithms hold promise but have not yet been validated to the level of headlines, and certainly not to the level needed to deploy the algorithms in a clinical setting. We have a well-established framework for review of traditional medical devices which diagnose or treat medical conditions, and we must clarify where these new AI-enabled algorithms must go through that existing process. Many of these new technologies will need to go through those processes. Properly designed trials, not just headlines, will be an essential part of those clearances and approvals. Yet, he added, there will be challenges with validating machine learning programs that are "unlocked", and that change based on new data they encounter after their safety and efficacy have been verified. Do the results of validating

studies stay valid when the underlying algorithm changes? It is one thing to train an algorithm on a given dataset and verify its behavior, but do we trust the "continuous learning" capabilities built into it such that we can be confident that future output from the "updated" algorithm will be valid and accurate? Our current regulatory frameworks will almost certainly need to be extended and expanded to properly cover continuous learning scenarios.

D. Translational Frameworks: Transcriptomics & Breast Cancer

Dr. Nicola Camp, a professor in the department of human genetics at the University of Utah, concluded the symposium and spoke about translational frameworks, meaning the way data is shared and communicated, in transcriptomics and breast cancer.

Dr. Camp recognized that there is currently a shift from clinical data (i.e. physical exam findings and cell counts) to molecular data (i.e. SNPs and gene profiles). Therefore, it will be important to represent data across translational areas so that we use the same language when discussing big data. The goal of Dr. Camp's work was to find a platform for research to share information and data across the translational spectrum.

The characteristics of gene expression was the starting point for developing a common platform for discussion of gene expression. Specifically, the characteristics of gene expression in breast tumors were explored. Typically breast tumors were divided into 5 distinct groups: Luminal-A, Luminal-B, HER2-enriched, Basal-like, and Normal-like (meaning the biopsy may have been of normal breast tissue). These 5 groups were differentiated based on 50 genes, and any breast cancer could be assigned to one of these groups based on the expression profile of the 50 genes [21,22].

Some companies including Mammaprint, OncotypeDX, and Prosigna have developed methods of improving on the "5-bin" classification system by adding additional genes and using statistical tools to predict prognosis and recurrence rates of cancers based on their genetic profile [23-25]. This innovation also led to the discovery that certain treatments were not as effective or necessary in certain women; for example, women with a better prognosis and lower longer-term risk could receive less toxic hormonal therapy rather than more traditional chemotherapy. Thus, complicated genetic data could be synthesized into a more understandable "risk score" that is translational across disciplines.

This led Dr. Camp to explore tumor genomics, and how we classify and communicate about breast tumors. In collaboration with Intermountain Healthcare, and utilizing the Utah Population Database, pedigrees were linked to pathology records and tumor genetics in the general population were mapped. With a population baseline, it is possible to compare a cancer's genetic profile with that of the general population. The hypothesis was that if a particular type of cancer within a pedigree were genetically different than that of a sporadic mutation in the general population, such a difference would indicate some kind of genetic susceptibility within a family [26]. However, after analysis of 15 different familial tumors, there was not any appreciable difference between the genetic profile of the familial cancers and the sporadic cancers in the general population.

This led to the question: Does one variable (the 5 groups of breast cancer) provide sufficient data to properly describe a tumor? Or does the genetic profile require additional dimensions to describe its biological variability? The comparison was drawn to Newton's Color Theory which

described color as one spectrum divided into 7 entities ranging from red to violet. In reality there are colors that don't fit onto Newton's color spectrum, just like there are breast cancers that don't quite fit into one of the 5 tumor categories. However certain colors like teal do not fit well into the spectrum. A more contemporary, and accurate, analysis of color is in a three-dimensional space with magenta, cyan, and yellow being the three dimensions. With this approach, the entire spectrum of color can be generated using a 3-value code from these three dimensions. Therefore, breast cancers might be able to be classified based on a multi-dimensional space, instead of a linear spectrum.

Using the "multi-dimensional reduction" approach, it is possible to generate a dimensional space that allows for more descriptive analysis of tumors. Two cancers that look identical in a one-dimensional space can end up very different from one another when additional variables are added; therefore, one could develop different treatment strategies and have different recurrence rates. Instead of analyzing all 50 genes as one score, analyzing different families and groups of genes as one variable would allow for more personalized, effective stratification according to risk and prediction of responsiveness to specific therapeutic regimens.

The problem with a multi-dimensional analysis is the lack of an accurate way to visually represent 5-dimensional space for breast tumors. Thus, a revolutionary new platform was developed based on the color wheel, with a 5-dimensional space represented by a wheel with space for the value of each variable, much like a Red/Green/Blue color code.
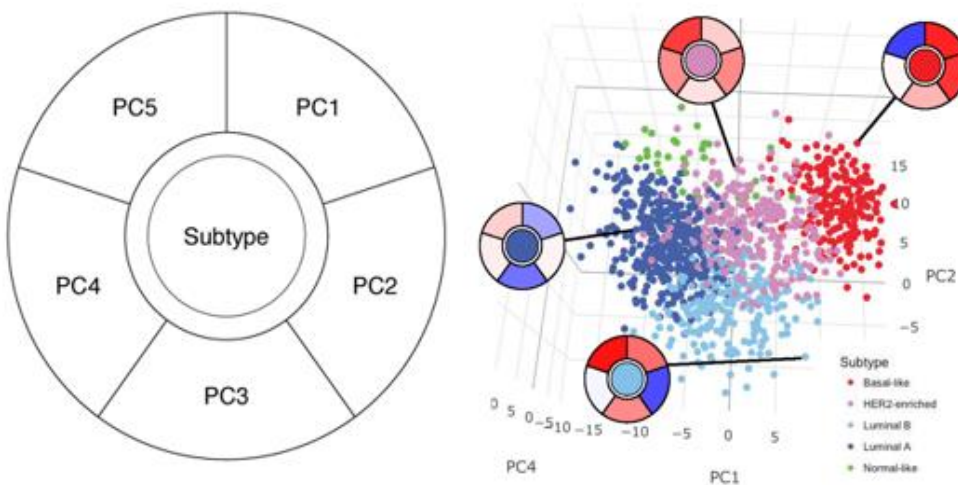

Figure 1 – Multiple quantitative dimensions of one breast cancer presented in a multidimensional space

In Figure 1, use of this 5-dimensional space allows for more accurate and informative analyses. For example, analysis of one principal component (PC), which is the linear combination of the gene expression, was most descriptive of the tumor in isolation from any other variable, and was found to very accurately explain recurrence risk. This finding was not present in a one-dimensional analysis and was confirmed with clinical research[27]. With this multi-dimensional approach applied to a completed European clinical trial (Geicam), the addition of paclitaxel (to a standard 3-drug chemotherapy regimen) was also shown to improve life expectancy by two additional years in node-positive women with low values in PC3/PC4. By contrast, paclitaxel

14

was found to have no impact on the life expectancy for the subset of women (44%) with high values of PC3/PC4[27].

The purpose of the multi-dimensional space was to find a difference between sporadic cancer in the general population and familial grouped genetic susceptibility to cancer. This analysis allowed for high-risk pedigrees to be found within families, and for their genetic profiles to be mapped. The PC variables were found to be inherited within families; by contrast, the one-dimensional characteristic classifications were variable[27].

Dr. Camp believes that we will be able to predict the trajectory of future breast cancers by applying this five-dimensional space to their gene profile, which would allow for better screening and treatment guidelines. In the future, there is the potential to correlate SNPs to the PC variables and know a woman's tumor profile before a clinically apparent tumor develops.

In summary, transcriptomics needed a different framework to have a more standardized discussion about breast tumors. Within the new quantitative framework, 5 PC variables are compared instead of one cancer subtype, resolution can be added, outcomes can be more accurately predicted, therapeutics can be targeted to maximize benefit, and pedigrees can be analyzed to understand how cancer is inherited. Big data may harbor many important discoveries, but for detection and rapid translation, the data framework matters.

<div align="center">References</div>

1. Chong JX, Buckingham KJ, Jhangiani SN, Boehm C, Sobreira N, Smith JD, Harrell TM, McMillin MJ, Wiszniewski W, Gambin T, Coban Akdemir ZH, Doheny K, Scott AF, Avramopoulos D, Chakravarti A, Hoover-Fong J, Mathews D, Witmer PD, Ling H, Hetrick K, Watkins L, Patterson KE, Reinier F, Blue E, Muzny D, Kircher M, Bilguvar K, López-Giráldez F, Sutton VR, Tabor HK, Leal SM, Gunel M, Mane S, Gibbs RA, Boerwinkle E, Hamosh A, Shendure J, Lupski JR, Lifton RP, Valle D, Nickerson DA; Centers for Mendelian Genomics, Bamshad MJ. The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *The American Journal of Human Genetics*. 2015;97(2):199-215. doi: 10.1016/j.ajhg.2015.06.009.
2. EvaluatePharma. Orphan Drug Report 2015. http://info.evaluategroup.com/rs/607-YGS-364/images/EPOD15.pdf. Published 2015. Accessed September 27, 2019.
3. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145),661-78.
4. CARDIoGRAMplusC4D Consortium. A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nature Genetics*. 2015;47(10):1121-1130. doi:10.1038/ng.3396.
5. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J, Cardon LR, Whittaker JC, Sanseau P. The support of human genetic evidence for approved drug indications. *Nature Genetics*. 2005;47(8):856-60. doi:10.1038/ng.3314.
6. Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JR, Elliott KS, Hackett R, Guiducci C, Shields B, Zeggini E, Lango H, Lyssenko V, Timpson NJ, Burtt NP, Rayner NW, Saxena R, Ardlie K, Tobias JH, Ness AR, Ring SM, Palmer CN, Morris AD, Peltonen L, Salomaa V; Diabetes Genetics Initiative; Wellcome Trust Case Control Consortium, Davey Smith G, Groop LC, Hattersley AT, McCarthy MI, Hirschhorn JN,

Frayling TM. A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nature Genetics*. 2007;39(10):1245-50.

7. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, Frayling TM, Hirschhorn J, Yang J, Visscher PM; GIANT Consortium. Meta-analysis of genome-wide association studies for height and body mass index in~ 700000 individuals of European ancestry. *Human Molecular Genetics*. 2018;27(20):3641-3649. doi:10.1093/hmg/ddy271.

8. Scott RA, Freitag DF, Li L, Chu AY, Surendran P, Young R, Grarup N, Stancáková A, Chen Y, Varga TV, Yaghootkar H, Luan J, Zhao JH, Willems SM, Wessel J, Wang S, Maruthur N, Michailidou K, Pirie A, van der Lee SJ, Gillson C, Al Olama AA, Amouyel P, Arriola L, Arveiler D, Aviles-Olmos I, Balkau B, Barricarte A, Barroso I, Garcia SB, Bis JC, Blankenberg S, Boehnke M, Boeing H, Boerwinkle E, Borecki IB, Bork-Jensen J, Bowden S, Caldas C, Caslake M; CVD50 consortium, Cupples LA, Cruchaga C, Czajkowski J, den Hoed M, Dunn JA, Earl HM, Ehret GB, Ferrannini E, Ferrieres J, Foltynie T, Ford I, Forouhi NG, Gianfagna F, Gonzalez C, Grioni S, Hiller L, Jansson JH, Jørgensen ME, Jukema JW, Kaaks R, Kee F, Kerrison ND, Key TJ, Kontto J, Kote-Jarai Z, Kraja AT, Kuulasmaa K, Kuusisto J, Linneberg A, Liu C, Marenne G, Mohlke KL, Morris AP, Muir K, Müller-Nurasyid M, Munroe PB, Navarro C, Nielsen SF, Nilsson PM, Nordestgaard BG, Packard CJ, Palli D, Panico S, Peloso GM, Perola M, Peters A, Poole CJ, Quirós JR, Rolandsson O, Sacerdote C, Salomaa V, Sánchez MJ, Sattar N, Sharp SJ, Sims R, Slimani N, Smith JA, Thompson DJ, Trompet S, Tumino R, van der A DL, van der Schouw YT, Virtamo J, Walker M, Walter K; GERAD_EC Consortium; Neurology Working Group of the Cohorts for Heart; Aging Research in Genomic Epidemiology (CHARGE); Alzheimer's Disease Genetics Consortium; Pancreatic Cancer Cohort Consortium; European Prospective Investigation into Cancer and Nutrition– Cardiovascular Disease (EPIC-CVD); EPIC-InterAct, Abraham JE, Amundadottir LT, Aponte JL, Butterworth AS, Dupuis J, Easton DF, Eeles RA, Erdmann J, Franks PW, Frayling TM, Hansen T, Howson JM, Jørgensen T, Kooner J, Laakso M, Langenberg C, McCarthy MI, Pankow JS, Pedersen O, Riboli E, Rotter JI, Saleheen D, Samani NJ, Schunkert H, Vollenweider P, O'Rahilly S; CHARGE consortium; CHD Exome+ Consortium; CARDIOGRAM Exome Consortium, Deloukas P, Danesh J, Goodarzi MO, Kathiresan S, Meigs JB, Ehm MG, Wareham NJ, Waterworth DM. A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease. Science Translational Medicine. 2016;8(341):341ra76-341ra76. doi: 10.1126/scitranslmed.aad3744.

9. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, Kathiresan S. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*. 2015;50(9):1219-1224. doi:10.1038/s41588-018-0183-z.

10. Health and Social Care-National Health Service. Matt Hancock announces ambition to map 5 million genomes. https://www.gov.uk/government/news/matt-hancock-announces-ambition-to-map-5-million-genomes. Published 2018. Accessed September 27, 2019.

11. Global Genomic Medicine Collaborative. International Cohorts Summit Meeting Summary [Internet]. https://g2mc.org/wp-content/uploads/2018/11/International-Cohorts-Summit_-March-26-27-2018-Meeting-Summary_FINAL.pdf. Published 2018. Accessed September 27, 2019.

12. Program, A. O. All of Us Anthem. https://www.youtube.com/watch?v=4AkSlb4yoCw. Published 2018. April 23. Retrieved May 15, 2019.

13. NIH-funded genome centers to accelerate precision medicine discoveries. https://allofus.nih.gov/news-events-and-media/announcements/nih-funded-genome-centers-accelerate-precision-medicine-discoveries. Published 2016. Accessed April 28, 2019.

14. About the Framingham Heart Study. https://www.framinghamheartstudy.org/fhs-about/. 2019. Accessed May 27, 2019.

15. About Us | The Dunedin Study - Dunedin Multidisciplinary Health & Development Research Unit. https://dunedinstudy.otago.ac.nz/about-us. 2019. Retrieved May 27, 2019.

16. Ratzan SC, Parker RM. Health literacy. *National library of medicine current bibliographies in medicine. Bethesda: National Institutes of Health, US Department of Health and Human Services.* 2000.

17. Kutner M, Greenberg E, Jin Y, and Paulsen C. The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy (NCES 2006–483). *U.S. Department of Education. Washington, DC: National Center for Education Statistics.* 2006.

18. Williams JR, Yeh VM, Bruce MA, Szetela C, Ukoli F, Wilkins CH, Kripalani S. Precision medicine: Familiarity, perceived health drivers, and genetic testing considerations across health literacy levels in a diverse sample. *Journal of Genetic Counseling.* 2018;*28*(1):59-69. doi:10.1007/s10897-018-0291-z.

19. Kaufman DJ, Baker R, Milner LC, Devaney S, Hudson KL. A survey of US adults' opinions about conduct of a nationwide Precision Medicine Initiative® cohort study of genes and environment. *PLoS One.* 2016;*11*(8):e0160461.

20. Kaphingst KA, Blanchard M, Milam L, Pokharel M, Elrick A, Goodman MS. Relationships between health literacy and genomics-related knowledge, self-efficacy, perceived importance, and communication in a medically underserved population. *Journal of Health Communication.* 2016;21 Suppl1:58-68. doi: 10.1080/10810730.2016.1144661.

21. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. *Nature.* 2000;406(6797):747-52

22. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lønning PE, Brown PO, Børresen-Dale AL, Botstein D. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences.* 2003;100(14):8418-23.

23. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* 2002;415(6871):530-6.

24. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *New England Journal of Medicine.* 2004;351(27):2817-26.

25. Kelly CM, Bernard PS, Krishnamurthy S, Wang B, Ebbert MT, Bastien RR, Boucher KM, Young E, Iwamoto T, Pusztai L. Agreement in risk prediction between the 21-gene recurrence score assay (Oncotype DX®) and the PAM50 breast cancer intrinsic classifier™ in early-stage estrogen receptor–positive breast cancer. *The Oncologist.* 2012;17(4):492-8. doi:10.1634/theoncologist.2012-0007.

26. Sweeney C, Bernard PS, Factor RE, Kwan ML, Habel LA, Quesenberry CP Jr, Shakespear K, Weltzien EK, Stijleman IJ, Davis CA, Ebbert MT, Castillo A, Kushi LH, Caan BJ. Intrinsic subtypes from PAM50 gene expression assay in a population-based breast cancer cohort: differences by age, race, and tumor characteristics. *Cancer Epidemiology and Prevention Biomarkers*. 2014;23(5):714-24. doi:10.1158/1055-9965.EPI-13-1023.

27. Camp NJ, Madsen MJ, Herranz J, Rodríguez-Lescure Á, Ruiz A, Martín M, Bernard PS. Re-interpretation of PAM50 gene expression as quantitative tumor dimensions shows utility for clinical trials: application to prognosis and response to paclitaxel in breast cancer. *Breast Cancer Research and Treatment*. 2019;175(1):129-139. doi:10.1007/s10549-018-05097-5.

28. Rainie L. The rise of the e-patient. Pew Research Center Internet and the American Life Project. http://www.pewinternet.org/Presentations/2009/40-The-rise-of-the-e-patient.aspx. October 7, 2009. Accessed January 11, 2012.

29. Lawry T, Mutkoski S, Leong N. Realizing the Potential for AI in Precision Health. *SciTech Lawyer*. 2018; 15(1): 23-27.