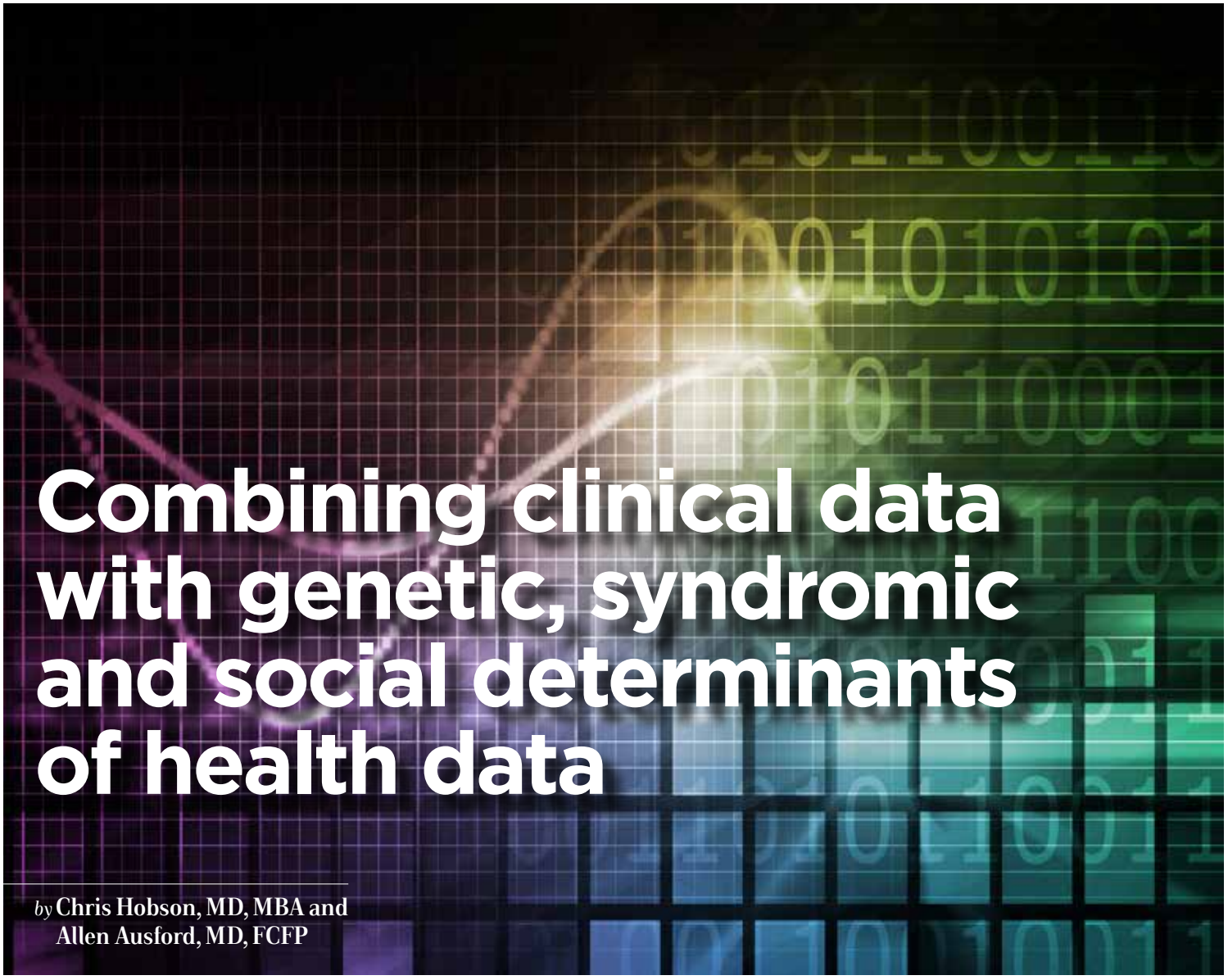

Combining clinical data with genetic, syndromic and social determinants of health data

by Chris Hobson, MD, MBA and Allen Ausford, MD, FCFP

According to the Precision Medicine Initiative, precision medicine is "an emerging approach for disease treatment and prevention that considers individual variability in genes, environment, and lifestyle for each person."[1] Precision medicine's stated goal is to treat individual patients with significantly greater precision and accuracy by analyzing data from large, diverse populations.

Today, the technical systems aggregating wide ranging data are the large health information exchanges (HIEs). A well-populated HIE contains broader information than a local electronic medical record (EMR) does and thereby enables improved approaches to precision medicine. HIEs have traditionally included clinical and behavioral/lifestyle data (e.g. problem list, past clinical and social history, medications, allergies, lab and radiology results, clinical documents, immunizations, etc.) stored in a central clinical data repository.

For the purposes of this discussion, we view integrated data as yielding consolidated information that is usable by clinicians and can be analyzed in context; it is not simply a series of separate data silos, as may be found in some electronic health records (EHRs) and HIEs. Adopting a standards-based approach that unifies both genetic data and clinical information systems is crucial to making sense of genetic testing results in a complete clinical context.

The existence of so much data presents novel challenges and opportunities. Challenges of this highly data-dependent enterprise include how to capture and store non-traditional data;

how to retrieve and display it in a way that makes sense to clinicians (and not overwhelm them); how to validate patient generated data; and how to use the data to identify patients who are most deserving of attention.

The opportunity is that by having a more accurate and precise understanding of patient populations, we can intervene in their care in ways that are more likely to be successful, providing the best possible care and patient experience at the lowest possible price.[2]

This paper highlights concepts that are important to better understand how varying types of data may be combined to facilitate and continue the progress of precision medicine.

Data Management

While HIEs have improved access to vast amounts of clinical data, we note issues and challenges remain that may hinder an organization's ability to manage data and practice precision medicine:

- Inaccurate or incomplete medication data: Medication data is critical to the safe care of patients with any errors having potentially catastrophic impacts. Consequently, medication data needs to be as comprehensive as possible, handled with due care and subject to extensive configuration testing before and during deployments.
- Privacy regulations: The privacy environment is constantly changing and evolving. Privacy rules and consent requirements must be respected, though they can have a restricting effect on the availability of important clinical data. Once members of the public understand the purpose of collecting and researching their genetic data, it is unusual for patients to object.
- Unstructured text: Text-based documents contain much valuable content though require technology such as natural language processing to make them digitally useable.
- Incomplete fields in the EMR: Data from the EMR communities are not always available or integrated.

Exciting recent developments have added significantly to the nature, volume, and value of additional data available to assist in managing patients as well as patient populations. In addition to traditional clinical data, we can now paint a picture with traditional clinical data, PH syndromic data, genomic data, social determinants of health, environmental and patient-generated data (see **Figure 1**). We consider these options in turn below:

Public Health Syndromic Data

Public health syndromic surveillance using inpatient and ambulatory clinical care electronic health records systems and other external data, like geographic or demographic data, is a relatively new practice, though it has been commonly used for predicting flu epidemics. A current example of leveraging this type of data for precision medicine can be seen in the recent North American measles outbreaks as public health officials look to identify vulnerable, non-immunized, populations.

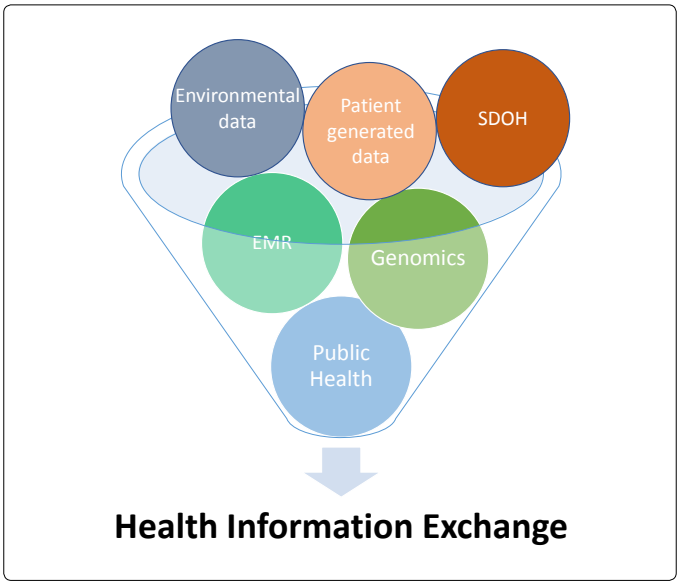

Figure 1: An illustration of recent data sets that can now be used to assist in managing patients and patient populations.

A simple illustration of this outbreak can be seen in a search on Google Trends "measles (disease classification)" for the period from July 1, 2016 to July 1, 2019; three peaks are evident for the periods Sept 2016, April 2018, and the most prominent flare up during February to April 2019 (see https://trends.google.com/trends/explore?date=2016-07-01%202019-07-01&q=%2Fm%2F0g2gb and **Figure 2**).

Another example of this type of data in action is that many organizations now screen patients during admission to emergency departments and hospitals for fever, cough, rash and travel to endemic areas to quickly isolate possible carriers of high-risk diseases.

Environmental Data

In the broad picture, environmental issues are also responsible for more morbidity and mortality than is commonly considered in patient care situations today. Clinicians should have more detail on those factors available at the point of care to better address them. Examples of environmental issues where the data is not typically available to clinicians today include lead exposure (i.e. Flint, Michigan), airborne exposures like asbestos, and radon found in homes. Exposure risks may be pinpointed to a zip code, and when integrated at the patient level in a clinical record, enabling further recommended actions including screening for the relevant disease and advising the patient on how to reduce or mitigate their risk.

Genomic Data

Although progress in the area of genomic data is extremely rapid, we should remember that as it stands today, some of the best genomic data can be found in the clinical record in the simple structured family history section. The family history is a standard part of most clinical assessments and the information frequently changes the clinical management of patients. For example, a family history of breast cancer in multiple first-degree relatives should trigger discussion about possible BRCA testing; a family history of colon cancer or adenoma should trigger a screening colonoscopy instead of fecal immuno-chemical test (FIT) stool testing; and a family history of heart attack in a young family member should alter the Framingham risk score for lipid screening.

Turning to actual genomic data, it is important to point out that it is only available for some and not all patients, and our understanding of the correlations with phenotypic disease impacts are still largely in the area of research. This will continue to be a major focus for academic research and health systems need to stay up to date with the latest developments.

For example, Geisinger Health System on the U.S. east coast is amassing a genomic database that already has 190,000 people signed up for genetic analysis, and the goal to give this

opportunity to every Geisinger patient.[3] As we develop and extend such databases, and pair genomic data with phenotypic clinical data, the correlation between genetic sequences and specific diseases will become both more accurate and more precise. The work of the Human Phenotype Ontology[4] will be critical in the evolution of HIEs and their ability to handle both clinical and genomic data in one record.

While detailed genomics data integration is often touted as the next medical disruptor[5], it is still relatively early on in its evolution. As more third-party genetic test offerings appear, patients are taking it upon themselves to do self-testing and sharing the results via patient portals. Improved integration of genomic data with HIEs over time will help guide clinician decision making, support secondary uses, and significantly improve patient care.

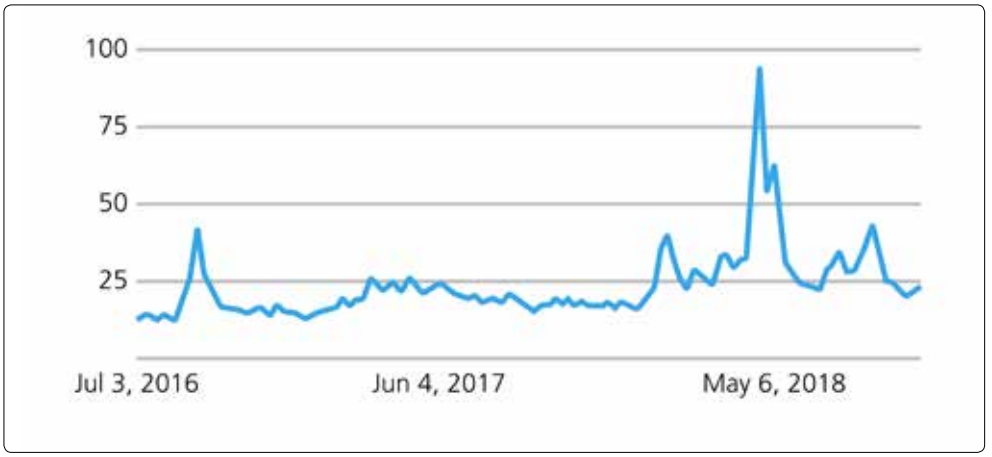

Figure 2: A Google Trends plot of "measles (disease classification)" for the period from July 1, 2016 to July 1, 2019: Three peaks corresponding to outbreaks are evident for the periods Sept 2016, April 2018, and the most prominent, a flare up during February to April 2019.

Social Determinants of Health (SDOH)

SDOH contribute to as much as 40 percent of an individual's health outcomes[6], so an understanding of social factors provides invaluable insight into improving a patient's health status and for predicting negative outcomes such as high-risk pregnancy, early hospital readmissions, poor compliance with prescribed medications, and susceptibility to chronic disease. Consider for instance that social isolation can increase the risk of heart disease by 32 percent[7] and lower levels of education are associated with higher rates of smoking and shorter life expectancy. Many primary care visits are linked to the effects of stress arising from social factors such as financial worries, family responsibilities, marital relationships, and so on. A new area with potential to add value for clinicians is the availability of additional information such as utility bills, traffic tickets, mortgage certificates, judgements and liens, address changes, and educational attainment.

Incorporating data on SDOH that address unmet social, financial, or economics needs of patient populations significantly improves clinician and care team understanding of the individual. In turn, addressing those needs should enable provider organizations to improve the overall health of a patient to a greater extent than a single-minded focus on the purely clinical issues.

Traditionally, behavioral issues are included under the umbrella of a "social history" and include items such as alcohol use, smoking, illicit drugs, diet, exercise, and high-risk sexual behavior. Such items are also powerful pointers to potential disease and should be included in the individual patient record and in the analysis of a population's health outcomes.

'Overall, we recommend that PGHD be separated from the rest of the medical record. Once the data has been reviewed by a clinician, it may be imported into the relevant sections of the record and used like other data for purposes such as to trigger clinical decision support rules. The data needs to be marked as "source = patient" and provenance of the data maintained'

Patient-Generated Health Data (PGHD)

Recently, patients have been able to monitor and submit significant health-related data in their electronic health record through a variety of devices and platforms. Those enabling technologies include patient portals, personal health records, mobile and Internet-of-things (IOT)-based devices.

A major challenge for clinicians using patient-generated data is how providers can validate and use the information submitted – that is, to what extent can clinicians trust data entered in the patient's record. For example, patient self-generated data can easily contain large volumes of normal data, which can quickly overload the practitioner without improving the quality of care. On the other hand, an example of patient-generated data that appears to have great value is the collection of family history data via online questionnaires. A recent study found that patients felt empowered through sharing their family history because they anticipated future predictive value from having the information integrated into their electronic health records. However, respondents raised concerns about validity (e.g. not reflecting the complexity of family history) and privacy (e.g. possible influence on insurance and employment).

Overall, we recommend that PGHD be separated from the rest of the medical record. Once the data has been reviewed by a clinician, it may be imported into the relevant sections of the record and used like other data for purposes such as to trigger clinical decision support rules. The data needs to be marked as "source = patient" and provenance of the data maintained.

Another important category of patient-generated data is the "patient preferred goals." Any health system must be ultimately judged by its ability to deliver outcomes that match the patient's preferences. In certain situations, the patient's goals do not align with those of their physicians or clinical teams, giving rise to a multitude of problems including poor compliance. Patients have a legitimate right to express their preferences and to expect that

they will be respected by their clinical team. Patient preferred goals should be integrated into the care plan and considered when determining any major interventions.

Assessing the Value of Diverse Data

As we look at new data types, decisions need to be made regarding the value of that data for clinical and population health purposes. If data does not add value to patient care or to population health, then it should not be included. A risk of providing too much information at the individual patient level may cause clinician overload to the point where they cannot process large volumes of data.

Ideally, we need tools to translate data into insights and actions that clinicians can undertake. We need to consider factors such as:

- What is the "signal-to-noise ratio" for any specific new data?
- Does the data add value from the perspective of managing the patient or managing a patient population?
- Does the data retain its meaning once exchanged? Can it be integrated safely and correctly into the existing comprehensive clinical record?
- Can the data be exchanged from source systems to an HIE or population health registry?
- Can it inadvertently lead to alert fatigue due to multiple clinical workflow interruptions?

Using a benefits realization model can minimize data overload. Ideally, we should select a benefit first and then work backwards to identify which small subset of indicators, outputs, and outcomes help to achieve the benefit thereby eliminating the need to review large amounts of extraneous data. Machine learning is also going to be key to making sense of large amounts of diverse data by finding the relevant indicators for each individual patient and for a population.

Public health registries are another source of relevant data. For example, we commonly ensure all patients over the age of 65 receive pneumovax and we commonly miss giving it to eligible younger patients with certain comorbidities like COPD. Linking those on specific chronic disease registries with their immunization history can alert us to an important gap in their care.

In addition, alert fatigue is increasingly a concern for clinicians. If we enable a mixture of high value and low value alerts, the usual response by clinicians is to ignore most alerts. Reducing interruptive alerts to a small number of individual clinician-determined high value items is key.

Looking Ahead

Clinicians typically have to make decisions in very time-limited settings. In order to reach an accurate diagnosis and treatment plan, it is important that all relevant high-value data, including new types of data, are available and considered. Completeness is important, as clinical decisions are better made on the basis of holistic information.

Today, a wide range of new data is available and can be aggregated, presented to the clinician, and analyzed for population-based purposes. All the data items listed in this paper have the potential to be leveraged for predictive modeling as well as directly in clinical care.

An ongoing collaborative effort by all stakeholders – patient, provider, and payer – is required to achieve precision medicine's promise of highly effective care personalized for a patient. JoPM

Dr. Chris Hobson, MD, MBA is the Chief Medical Officer with Orion Health. He is responsible for the clinical direction of Orion Health products and solutions, and ensures the protection of patient safety and privacy in everything Orion Health does. He is a primary care practitioner and internist certified in healthcare informatics. He has extensive international clinical, technical, and policy-level expertise, drawing from his background successfully developing and implementing health information technology globally.

Dr. Allen Ausford, MD, FCFP is a Clinical Professor with the Department of Family Medicine at the University of Alberta and is actively involved in teaching medical students and residents. He chairs many clinical IT committees and served as CMIO for the health region. His involvement is at the regional, provincial, national, and international levels in EMR and EHR applications design, adaptation, integration, interoperability, change management, clinical decision support, patient portal consulting, and chronic disease management. He is currently a member of the Orion Health Clinical Advisory Team (EHR) and co-chair of the Alberta Connect Care (CIS) Clinical Decision Support Committee.

References

1. Precision Medicine Initiative https://obamawhitehouse.archives.gov/precision-medicine
2. Institute for Health Improvement – The IHI Triple Aim Initiative http://www.ihi.org/Engage/Initiatives/TripleAim/Pages/default.aspx accessed Wed Jun 6 2019
3. MyCode Community Health Initiative / "What is my Code?" https://www.geisinger.org/precision-health/mycode accessed July 10 2019
4. The Human Phenotype Ontology https://hpo.jax.org/app/ accessed June 5 2019
5. Ginsburg, G. S., & Phillips, K. A. (2018). Precision Medicine: From Science to Value. Health affairs (Project Hope), 37(5), 694–701. doi:10.1377/hlthaff.2017.1624
6. Minnesota Department of Health "Social Determinants of Health" https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-of-health/objectives accessed Jun 5, 2019
7. Schroeder, Steven A We Can Do Better – Improving the Health of the American People, N Engl J Med 2007; 357:1221-1228 DOI: 10.1056/NEJMsa073350