Policies and practices of data-intensive primary care in the precision-medicine era

Vasiliki Rahimzadeh[1], Gillian Bartlett[1]

Author details:
Vasiliki Rahimzadeh[1],
Department of Family Medicine,
McGill University,
5858 Cote des Neiges,
Suite 300, Montreal,
QC H3S-1X1

Gillian Bartlett[1]
Department of Family Medicine,
McGill University,
5858 Cote des Neiges,
Suite 300, Montreal,
QC H3S-1X1

Corresponding author:
Vasiliki Rahimzadeh[1]
Department of Family Medicine,
McGill University,
5858 Cote des Neiges,
Suite 300, Montreal,
QC H3S-1X1
E-mail: vasiliki.rahimzadeh@mail.mcgill.ca
Tel: 514-887-7030

Abstract

Since our first writing on the topic in 2014, the wheels of technological progress in next generation sequencing and precision medicine have not ceased turning. The clinical frontier is on the cusp of yet another technological revolution, one that is chiefly driven by new forms and functions of data, capable of transforming the clinic into a living, learning ecosystem of discovery and care. Primary care is, as it was in 2014, situated at the nexus of this technological progress, community health and whole person care. With greater demands for data of increasing volumes, veracities and validities to meet new practice-based needs, what then constitutes primary 'data' in primary care? In what forms does this data come, and for what/whom should it serve in the post genomic era? This article takes up the ethical, legal and social dimensions of these questions as they relate to the precision medicine movement in North America. It pays special attention to collaborative synergies in fields such as epigenetics, the data requirements needed to support them, and the future of a 'data-intensive' primary care.

Keywords:
Precision medicine, primary care, big data, data sharing, genomics

1. Introduction

This article should be read in conversation with, and as a general update to the ethical, legal and social implications (ELSI) of integrating genomic medicine into routine primary care. Since our first writing on the topic(1), it comes as no surprise that the wheels of technological progress in next generation sequencing and precision medicine have not ceased turning. Precision medicine became its own federal research program and funding priority under President Obama in 2015(2). It inspired a new generation of disruptive e-health applications, enabling patients to be active agents in their own healthcare. The precision medicine movement furthermore brought the concept of citizen science to bear on the ways in which public engagement directs the sails of future clinical research with participants as partners in this process (3). It is also fortunate that many of the ELSI considerations we discussed in 2014 received targeted policy attention, including the return of incidental findings(4,5) and long overdue updates to consent modalities for secondary use of data(6,7). The new revisions to the Common Rule were nearly 20 years in the making, and are evidence of such prioritization in ELSI policy in the United States (8).

To be sure, precision medicine encompasses related clinical approaches. These include personalized and genomic medicine, yet expands the target of personalization beyond the genome to include a focus on broader lifestyle and health behaviors, as well as how environmental exposures influence them. We adopt the National Institute of Health (NIH) definition of precision medicine for the purposes of this discussion, as an "approach to disease treatment and prevention that seeks to maximize effectiveness by taking into account individual variability in genes, environment, and lifestyle. Precision medicine seeks to redefine our understanding of disease onset and progression, treatment response, and health outcomes through the more precise measurement of molecular, environmental, and behavioral factors that contribute to health and disease. This understanding will lead to more accurate diagnoses, more rational disease prevention strategies, better treatment selection, and the development of novel therapies"(2).

Yet just as we approach new regulatory and scientific horizons, the clinical frontier of precision medicine is on the cusp of yet another technological revolution. This one is chiefly driven by new forms and functions of data, capable of transforming the clinic into a living, learning ecosystem of discovery and care (9). Defined by the then Institute of Medicine, a healthcare system that "learns" is one in which "knowledge generation is so embedded into the core of the practice of medicine that it is a natural outgrowth and product of the healthcare delivery process and leads to continual improvement in care"(10). Collection, storage and sharing of genomic data along with other health-related data is critical to this endeavor. Equally critical is the electronic medical record (EMR), which serves as the vessel for linked genomic and health-related data upon which precision medicine innovations within a learning healthcare system depends (see for example 10-12).

With greater demands for data of increasing volumes, veracities and validities, what then constitutes primary 'data' in primary care? In what forms does this data come, and for what/whom should it serve in a learning healthcare system aimed at delivering precision medicine? This article takes up the practical, ethical and legal dimensions of these questions as they relate to the precision medicine movement in the United States.

The following sections proceed first with a discussion of the type of data needed in part to elucidate novel patterns of disease (e.g. etiology, treatment, and prevention) common to the primary care encounter. These diseases may include chronic conditions and those identified in early childhood as a result of newborn and other early childhood screenings. We describe how, not unlike in 2014, the complexity of genetic and environmental etiologies of chronic disease continues to pose extensive scientific challenges for improved care approaches and drug development. In order to meet the data-intensive demands of these clinical tasks, linked phenotypic and genotypic data must be made available to researchers; and lots of it. We discuss the regulatory considerations of such sharing at a time when public fears of genetic discrimination and other forms of undue genetic exceptionalism abound.

Next, we explore the ethical-legal considerations of new forms of health outcomes data in primary care, and map them onto the "entrustable professional activities" (EPAs) proposed by the Competencies Working Group of the Inter-Society Coordinating Committee for Physician Education in Genomics(14). We discuss three of five EPAs relevant for primary care—family history, genomic testing and treatment based on genomic results—and highlight the types of health data primary care providers (PCPs) may utilize to demonstrate these professional competencies. We contend that in aggregate volumes and on a direct patient level, outcome data can contribute to more comprehensive clinical utility and validity of genomic testing, with the hope of expanding the opportunity for such testing more routinely to chronic disease. On a health systems level, outcome data can be closely monitored for quality assurance and patient safety, can inform resource allocation decisions within institutions and when granted appropriate access to bonafide researchers can identify unmet needs that future research can address.

2. Primary care and epigenetics: a shared refuge for complex data

Precision in the diagnostic tools and interpretation of molecular data is a precondition of personalized/genomic medicine. As next generation sequencing supplements typical diagnostic batteries in primary care, accurate interpretation and subsequent referral will be among the chief responsibilities for the primary care clinician. There is evidence to suggest that institutions and medical educators are responding to calls for improved genetics training for future as well as practicing clinicians, including in primary care (14). Even with advanced genomic literacy, however, discovering variants of unknown significance (VUS) will in the future be challenging for both clinical follow up and referral. The overall number of variants with unknown significance is disease-specific, however variants have perhaps been best characterized in fields such as oncology(15). The discovery of VUS can be especially resource intensive, requiring extensive specialist follow up and further testing.

Recent advances in the field of epigenetics are elucidating gaps in understanding VUS. There has been growing interest in epigenetics mainly because the "lack of identified genetic determinants that fully explain the heritability of complex traits, and the inability to pinpoint causative genetic effects in some complex diseases, suggest possible epigenetic explanations for this missing information" (16). Several authors underscore the important role primary care can play in contributing to the multi-parametric data needed for epigenetic analyses (17) e.g. family history, sociocultural and demographic data, as well as general lifestyle and

environmental information especially relating to pre-and perinatal conditions. Epigenetics-epi (above) genetics is concerned with the interactions between, and modifications of genes along with environment exposures that may influence observable phenotypes. In relying on a combination of molecular and other relevant health data, epigenetics can be seen as a disciplinary liaison bringing together primary care, public health and genetics. What's more, epigenetics and primary care embrace complex systems, of which social determinants of health and the environmental influences on genetics both exemplify. The shared appreciation for complexity (social and genetic) as empirical and philosophical foundations of understanding human health further testifies to the compatibility between primary care and epigenetics within the learning healthcare system.

3. (de) Identifiable data links and the ethical technicalities of sharing

In addition to being an individual patient archive, the electronic medical record (EMR) is increasingly becoming an informational clearinghouse for both phenotypic and genotypic patient data. The analytic potential of such linked data for monitoring patient outcomes motivated large federal investment in how to better integrate genomic and EMR data within large institutional consortia in support of a learning healthcare system (18). Such data integration augments the sophistication of patient outcome data when clinicians document how disease manifests, and therapeutic efficacies compare across patient populations with (dis)similar phenotypic, genotypic and demographic traits. On a health systems level, outcome data that draw on linked sources of genotypic and phenotypic information available in the EMR can outline gaps in therapeutic approaches, identify specific populations for whom targeted personal therapies or screenings are most effective, and highlight areas for further research where institutional performance may be poor (19).

The enthusiasm for greater integration of linked data sources has been met with equal concern for adequate data protection and security(20–22). The informational utility of a comprehensive EMR rests on the ability to link, share and analyze the patient data it stores. The realities of data de-identification and technical approaches to keeping it secure are a moving target for patients and primary care clinicians to keep appraised, never mind the ever-changing lexicon of the types of security techniques used (see **Box 1** for examples most applicable in the primary care context). Phenotypic or genotypic data alone can be enough to accurately de-identify a patient given certain parameters and specificities of the types of health data and access governing its use. Such de-identification can be affront to explicit patient consent, particularly if the patient contributed their data under presumed conditions of anonymity (23,24). The promise of absolute anonymity in the post genomic era is, however, a false one at best. While deliberate privacy breaches are rare, the likelihood of de-identification inevitably heightens when multiple data sources are linked together. Although the recent ransomware hacking of the National Health Service and other databases in the United Kingdom exemplified unauthorized access to patient data, this data was not further disclosed or shared with unauthorized third parties. Patients whose data were stolen during the cyberattack experienced few serious repercussions as a direct result. Other instances of data de-deidentification involved bioinformaticians who purpose-fully demonstrated how certain linkage patterns with genomic and other health-

Box 1—Lexicon of applicable data protection mechanisms and definitions most in the primary care context (*adopted from the Data Sharing Lexicon of the Global Alliance for Genomics and Health* (27)).

Anonymisation	The irreversible delinking of identifying information from associated data.
Coding/ Pseudonymisation	The act of replacing an identifier with a code for the purpose of avoiding direct identification of the participant, except by persons holding the key linking the code and identifier
Confidentiality	The ethical and legal obligation of an individual or organization to safeguard data or information by controlling access as authorized by law or by the data donor.
Controlled/ Restricted Access	Access to data that is subject to conditions and an approval process.
Data Linkage	The process by which records representing the same entity or individual are linked across multiple data sources.
Data Protections	The set of laws, policies and procedures that aim to minimize intrusion into people's privacy, uphold confidentiality, and penalize undue intrusions and/or breaches.
Data Security	The protection of the confidentiality, availability and integrity of data.
De-identification	The removal or alteration of any data that identifies an individual or could, foreseeably, identify an individual in the future
Disclosure	The revelation of confidential information about an individual.
Encryption	A mechanism of safeguarding stored data or information by making those data or information unreadable without access to the correct decryption method.
Identifiable/ Personal Data	Data that alone or in combination with other data may reasonably be expected to identify an individual
Information	Data that have already been interpreted, i.e. they have meaning in a specific context.
Open Access	Making data available without restriction.
Privacy	The right and freedom to control access to information about oneself.
Registered Access	A system of authentication and self-declaration prior to providing access to data.
Re-Identification	The act of associating specific data or information within a dataset with an individual.
Return of Results	Communication of research results to an individual or a designated health care provider or family member.
Secondary uses	Using data or biospecimens in a way that differs from the original purpose for which they were generated or collected.
Trusted Third Party	An individual or organization that safeguards access to information linking individuals to their data and biospecimens.

related data could unintentionally identify patients despite anonymization (25,26). Rather than restrict sharing outright, such breaches in data security should instead motivate infrastructure scientists, bioinformaticians and researchers alike to recalibrate the threshold of realistic informational risks (and benefits) of such sharing in partnership with the patients whose consent they seek. For the implications of not sharing data can be as deleterious to the future of clinical progress as unfettered sharing without informational safeguards is to respect for persons.

Strict data anonymization in primary care may be the exception rather than the rule, however. The clinician-patient relationship, for instance, relies on longitudinal disclosure and information sharing throughout the lifecourse, and a relationship that depends on data identifiability. Data privacy norms in primary care may therefore differ considerably from the genomic research context, and pose an added data protection challenge when operationalizing the integrated EMR discussed herein(28). An intermediate data security schema is hence needed, which strikes a proportionate balance between primary care and data-intensive research objectives in genomics (29). This schema would ideally preserve the identifiability of the data so as to enable clinical follow up with the patient. It would also afford a proportionate degree of security commensurate with the nature and type of data shared to prevent patient de-identifiability upon secondary data use. Coded/pseudonymized data (Box 1) with controlled or registered access (24) may be ideal in this regard.

Ensuring direct patient benefit from data collected and stored appropriately, and facilitating secondary use of this valuable data through responsible data sharing, has fueled a corollary open science movement in the genomics research community (30–33). We argue, as do others, that with appropriate ethics governance and federal resources, "it may be reasonable to consider public benefit as a goal, or even an obligation, in the collection and analysis of [patient] data"(34). Consolidated genomic and primary care EMR data could therefore be transformed into a "centralized public resource" within the learning healthcare system (34). Within this system, individual patient benefits are made possible by the contributions and analysis of data from other patients.

Although clinicians and researchers make concerted efforts to keep linked patient data strictly de-identified(35–37), this again may not always be preferable in the primary care context. It may be advantageous, for example, to update patients on emerging evidence that suggests clinical significance for the variants they harbor, but which were previously unknown to their treating physician. Some authors argue clinical follow up for VUS falls under a physician's moral duty to rescue and a duty to provide ancillary care (38), both of which require patient de-identification. PCPs may also want to follow up with patients who underwent genomic testing as children, but for whom there might be reason e.g. family history or symptomatic family members to believe they carry a gene for an adult onset disorder that was not disclosed during childhood as per existing guidelines(4,39). Testing children for later onset disorders or those without clinical action-ability during childhood is not endorsed (40–43), yet the longitudinal familial relationships fostered between PCP and patients allows for prevention and identification of known genetic conditions that may run in the family.

4. Genomic competencies and associated data needs in primary care

Advances in, and decreasing costs of sequencing generates far more genomic data than can at present be deemed clinically relevant to inform patient care. Many studies report that PCPs lack the necessary genomic literacy to use genomic information in routine patient care(44,45), however this trend may be slowly changing. Professional development, educational competencies and training programs specifically for enhancing translational knowledge around genomic medicine(46) are largely to thank for improving genomic literacy among PCPs and others(47,48). The Inter-Society Coordinating Committee for Physician Education in Genomics (ISCC) spearheaded one of these educational initiatives. Convened by the National Human Genome Research to develop a framework for genomics practice competencies, the ISCC outlined 5 "entrustable professional activities" (EPA) centered on integrating genomic information in routine clinical care. **Figures 1-3** illustrate how three of these EPAs are particularly relevant in the primary care context, and the associated data needed to realize the practice-based benefits of making genomics a more routine component of clinical care: i) family history, ii) genomic testing and iii) treating the patient using genomic results. Seven types of data are categorized based on their relevance to the specific competencies outlined above. These data vary in their identifiability, sensitivity and clinical utility and, as such, define the ethical-legal parameters determining its use. Each data category is described in turn.

4.1 *Genomic data*

Genomic data may refer to data derived from whole genome or exome sequencing either in the clinical or research context. It may also come from direct to consumer testing companies, in which results reflect known gene-disorder correlations but are not validated in clinical accredited laboratories(49,50). Genomic data may vary in its risk of de-identifiability when shared outside a research context. As a result, new questions emerge. For example, can substitute decision makers e.g. parents/guardians or other care providers have full authority over a patient's genomic data(51)? Can a parent make their child's data publicly available in a large population database? Would the permissibility change in line with whether direct therapeutic benefits were anticipated as a result of such sharing? Such questions have been the focus of recent empirical research (52,53), and a forth-coming policy delphi conducted by author VR among pan-Canadian pediatric genomic researchers.

It has been argued elsewhere that precision medicine practiced within a learning healthcare system promises to further blur what has increasingly become an indiscriminate line between research and care (54). The massive volume of genomic data needed to make sound associations between the human genome and (chronic) disease means that sharing data in the name of broader societal benefit in fact improves the chances of direct clinical benefit for an individual patient (55–57). Advances in rare genetic disease testify to this (58–61). But even greater volumes and varieties of quality genomic data are needed to witness similar successes in chronic disease due to the complexity of gene, environment and behavioural interactions (62).

4.2 Medical history data

Medical history data constitutes data typified in a standard electronic medical record. It may include results from diagnostic screens (genetic or other), resea-rch participation, imagining, in-patient admissions and related records, as well as

information on prescription medications. Medical history data relates solely to the individual patient, whereas family history data provides chronological, familial context to one's clinical profile.

4.3 Family history data

Whereas information in the medical record are individual and patient-specific, genetic information is, by nature, inherited. Thus, genomic data is one of the only types that can be common to both medical and family history records. Family history data broadly encompasses clinical, as well as non-clinical information regarding health and well-being of the family unit. While family histories can be compiled using multiple individual patient medical records, they are more often informed by data that is of narrative form. Patient narratives may also not follow linear timelines. Yet, they this narrative data is critical in that it adds anecdotal richness to health information used to care for the whole person, which includes the health of biologically related individuals across time and space.

4.4 Sociocultural data

Race, religious affiliation, sex/gender, as well as country of origin together constitute pertinent sociocultural and demographic data that can be encompassed by family history, and thus better contextualize genomic data. Genetic diseases that are disproportionately common in certain populations illustrate this mutually informative dynamic between genetics and cultures of place(63,64). Like family history data, sociocultural and demographic data can be narrative or documentary. Moreover, sociocultural data can often lend guidance and anticipatory insight into how communicative events between patient and provider might unfold, such as in disclosing incidental findings discovered during genomic testing or deciding on the trajectory of care based on these findings (see for example 46–48).

4.5 Socioeconomic and demographic data

It is undisputed that income and education are key social determinants of health, yet the extent to which they influence access to care varies in line with healthcare organization and delivery systems. Medical records often do not contain explicit details on patient income, employment or available neighbourhood resources, yet these may be inferred based on other demographic information such as address or insurer. PCPs and community health advocates contend maybe they should. PCPs often practice in the same communities in which they live, affording them richer insight the local cultures, norms and politics that (in)directly factor into healthcare access. Growing disparities in health on the basis of socioeconomic indicators are motivating the prioritization of what some authors term 'community vital signs' (68). Put simply, socioeconomic/demographic data that "convey contextual social deprivation and associated risks based on where patients live […] that could influence point-of-care decisions." Until outlined in professional standards or guidelines (see Section 4.6), coverage for genomic testing with adequate clinical follow up could introduce a new socioeconomic determinant of health in the precision medicine era(28). These effects could be accentuated in privatized healthcare jurisdictions like the United States, and in publicly funded systems where testing is recommended for prevention, but not medically necessary as yet.

4.6 Regulatory and professional guidelines data

Regulations (in particular data privacy and protection) as well as professional practice guidelines comprise two sources of data relevant to the

governance of primary care at the individual provider as well as macro policy levels. First, sensitive clinical data is collected, stored and shared under conditions set out in data privacy regulations. These conditions have needed to be more ameliorable to adaptation than ever before based on the deluge of data precision medicine both generates and draws on. As learning healthcare systems encompass both research and care activities, research ethics regulation too becomes a relevant regulatory feature of primary practice. Indeed, the Common Rule that regulates all research with human subjects in the United States only recently underwent the most significant reform in two decades in order to facilitate the evidence-based practices that a learning healthcare system affords (69,70).

Second, professional practice guidelines are important dissemination vehicles for bench-to-bedside innovations in care that precision medicine makes possible. In other words, professional practice guidelines situate research discoveries within existing practices and norms specific to the activities of a professional body. The American College of Medical Genetics (ACMG) guidelines on the return of incidental findings(71) and interpretation of sequence variants (72) are but two examples that illustrate where innovations in genomic diagnostics meet standards for professional clinical practice.

4.7 Institutional data

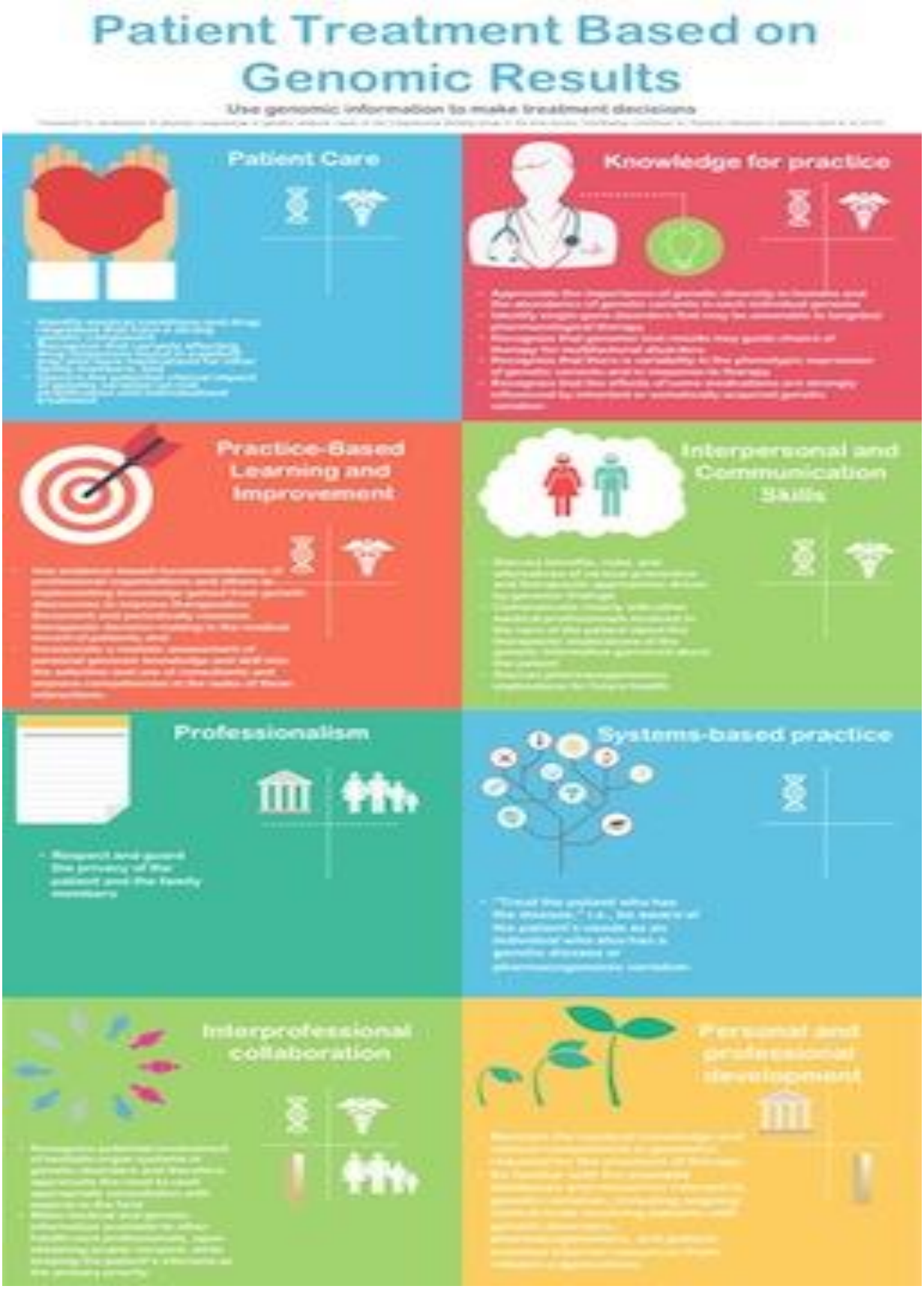
By institutional data, we refer to health provider knowledge of resources and services offered through the institution, and how PCPs access them in the best interest of their patient. This form of data further encompasses information on the patient populations the institution services, complementing the sociocultural and socioeconomic forms of data described earlier. Because primary care is the first point of access to the healthcare system, patient referral is a key professional responsibility of PCPs. Referral requires obtaining, and deploying institutional data to secure on behalf of the patient resources needed to optimize their care.

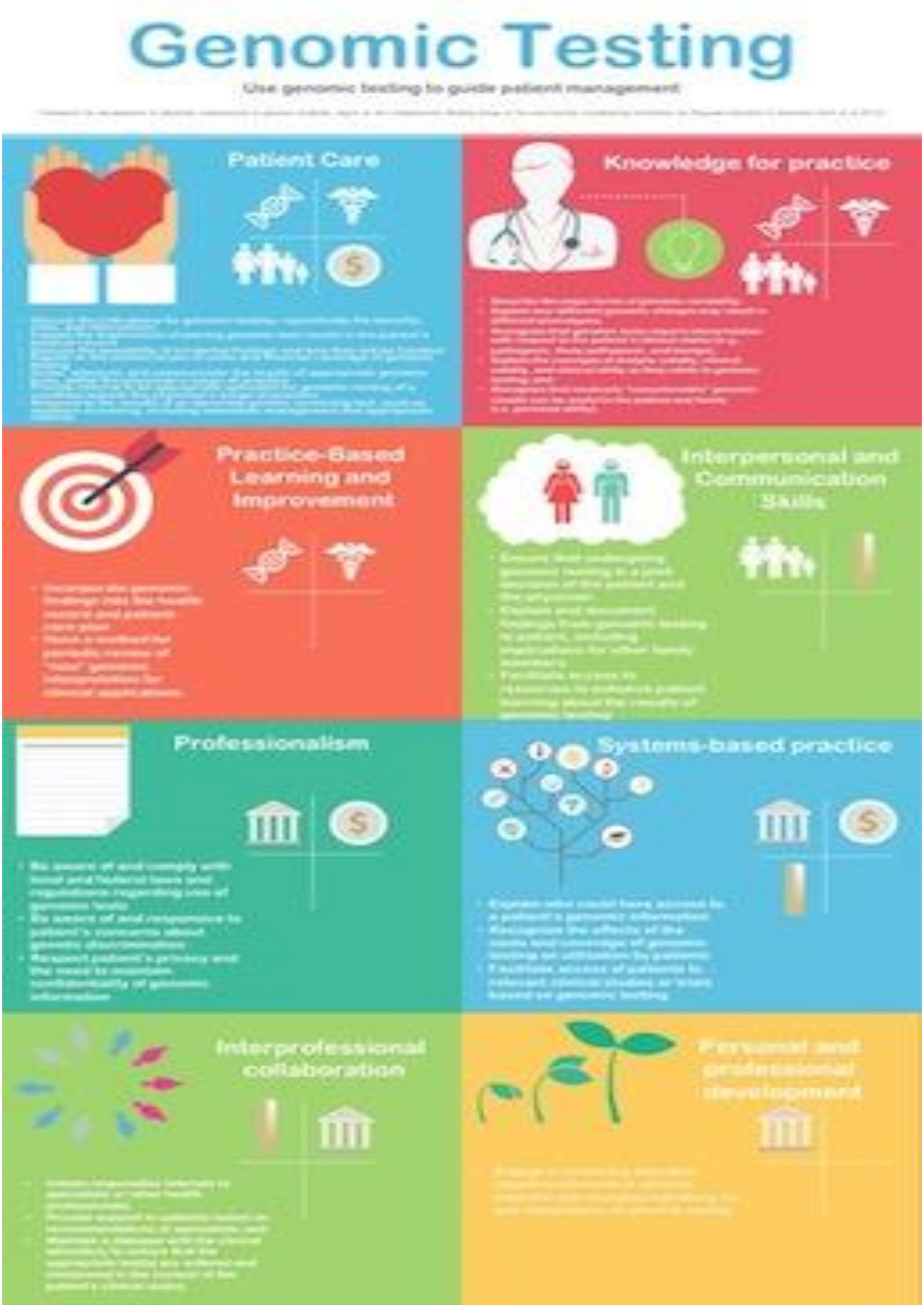
5. Conclusion

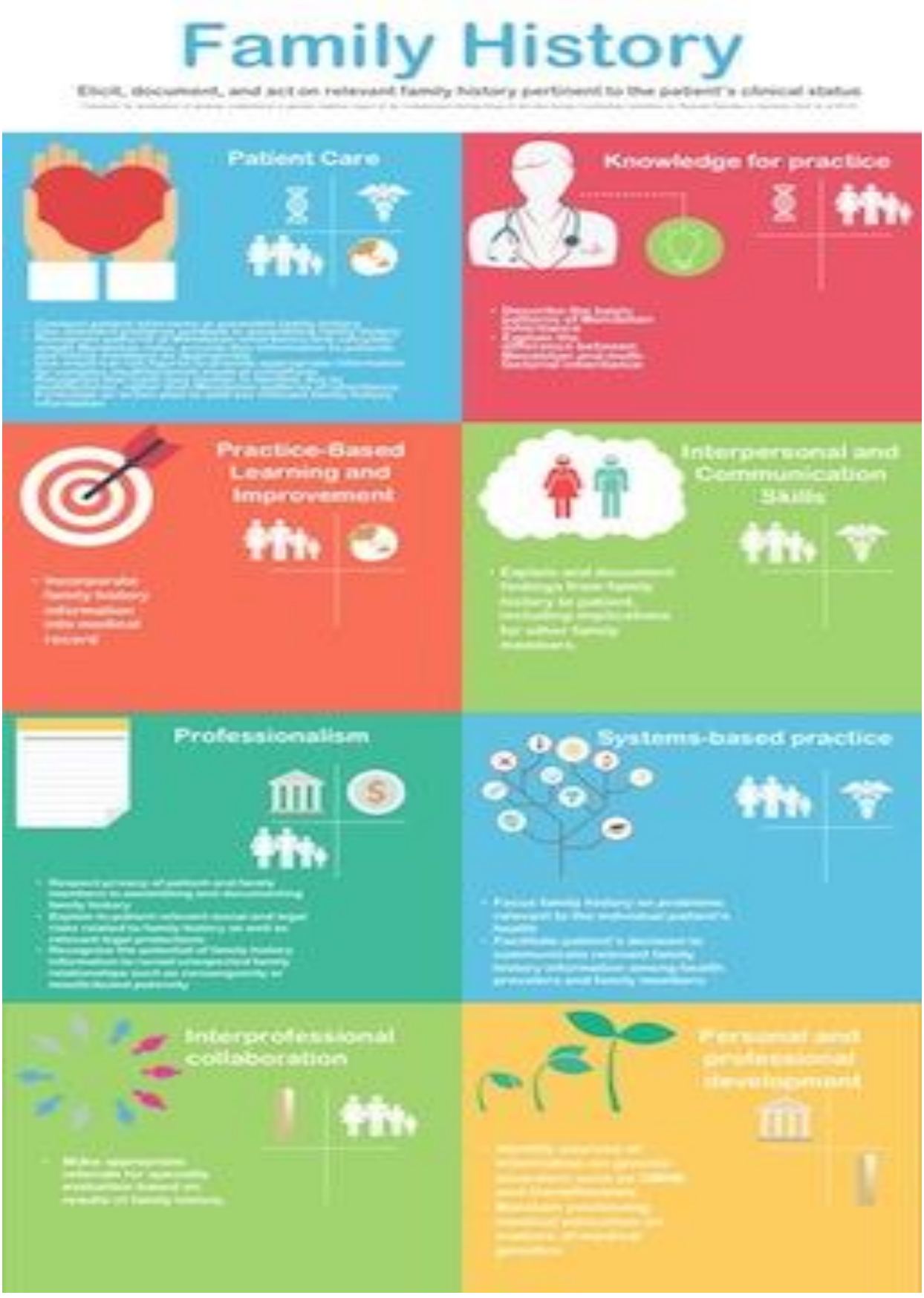
The compendium of primary care data sources, types and purposes are-like healthcare itself-multi-parametric and complex. There is growing interest in, as well as advanced computational methods to make sense of such complexity in ways that bring us closer to the reality of precision medicine that so many have anticipated. Given the diversity of information sources needed to attend to the whole person, primary care may perhaps be ahead of other clinical disciplines in its embrace of the data-intensive inputs that precision medicine requires. Genetic/genomic data should be considered an additional information source in this regard to be contextualized with the myriad data that typify a primary care patient record. We argue, as others do, the individual as well as population benefits yielded from sharing genomic and associated primary care data are not only ethically defensible, but necessary to the contemporary practice of whole person care in the post genomic era. Data in the volumes, varieties and veracities we describe herein are the building blocks upon which the successful implementation of precision medicine ultimately rests, and finds PCPs in particularly strategic positions to deliver on its promises.

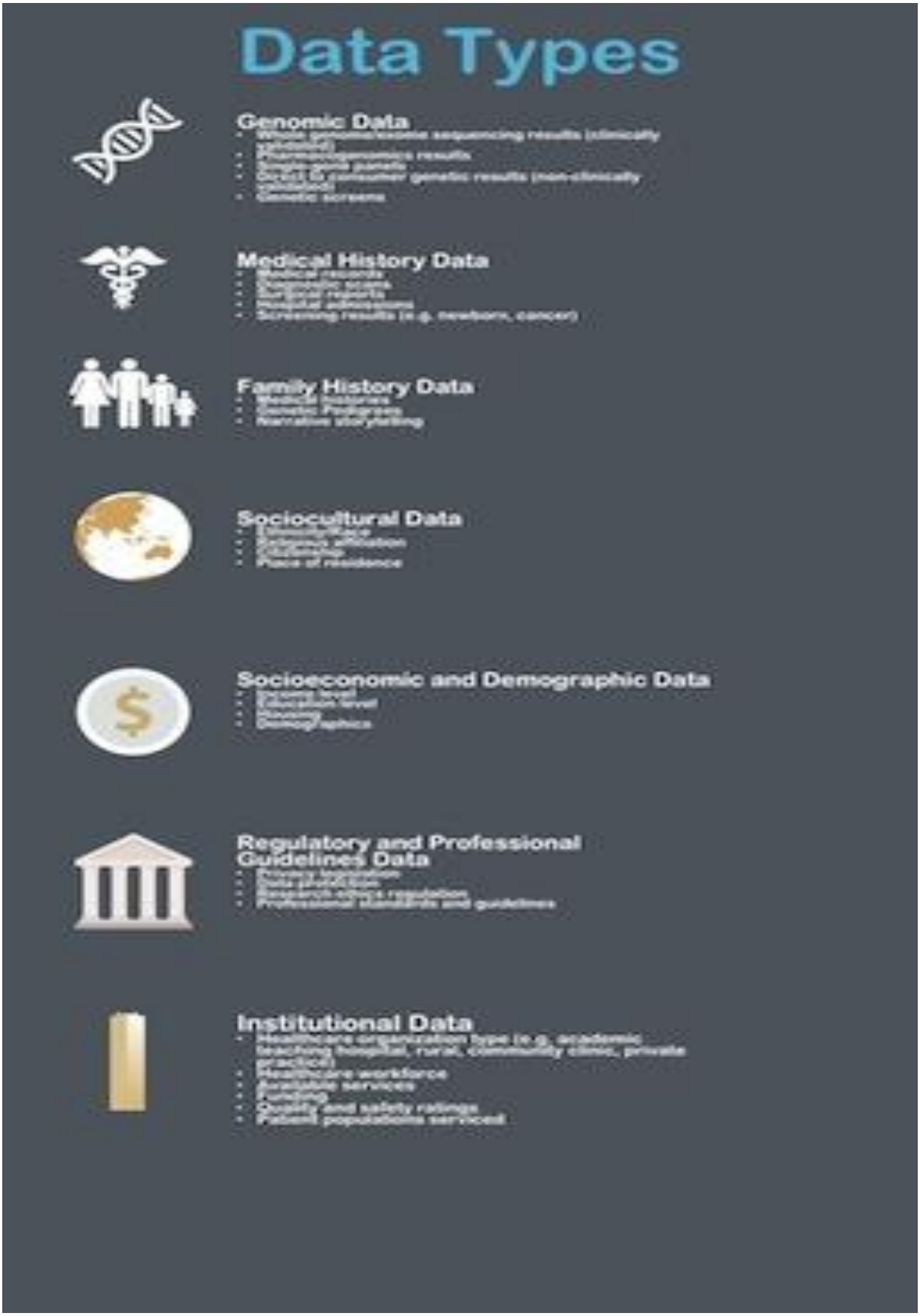
Acknowledgements

The authors wish to thank the following funders for the support of this work: Vanier Canada Graduate Scholarship CIHR# 359258 and the Precision Medicine Policy Network (Genome Canada).









References

1. Rahimzadeh V, Bartlett G. Genetics and primary care: where are we headed? J Transl Med. 2014 Jan;12(238).
2. Precision Medicine Initiative (PMI) Working Group. The precision medicine initiative cohort program – building a research foundation for 21st century medicine. Vol. Sept 17, Precision Medicine Initiative (PMI) Working Group Report to the Advisory Committee to the Director, NIH. 2015.
3. Hafen E, Kossmann D, Brand A. Health data cooperatives - Citizen empowerment. Methods Inf Med. 2014;53(2):82–6.
4. Boycott K, Hartley T, Adam S, Bernier F, Chong K, Fernandez BA, et al. The clinical application of genome-wide sequencing for monogenic diseases in Canada: Position Statement of the Canadian College of Medical Geneticists. J Med Genet. 2015 Jul 1;52(7):431–7.
5. Green RC, Goddard KAB, Jarvik GP, Amendola LM, Appelbaum PS, Berg JS, et al. Clinical Sequencing Exploratory Research Consortium: Accelerating Evidence-Based Practice of Genomic Medicine. Am J Hum Genet. 2016;98(6):1051–66.
6. Schmietow B. Ethical dimensions of dynamic consent in data-intense biomedical research-paradigm shift, or Red Herring? In: Ethics and Governance of Biomedical Research. Switzerland: Springer International Publishing Switzerland; 2016. p. 163–73.
7. Kaye J, Whitley EA, Lund D, Morrison M, Teare H, Melham K. Dynamic consent: a patient interface for twenty-first century research networks. Eur J Hum Genet. 2015;23(10):141–6.
8. Department of Homeland Security, Department of Agriculture, Department of Energy, National Aeronautics and Space Administration, Department of Commerce, Social Security Administration, Agency for International Development, Department of Housing and Urban and the D of T. Federal Policy for the Protection of Human Subjects. U.S. Department of Health and Human Services; 2017.
9. Grady C, Wendler D. Making the transition to a learning health care system. Hastings Cent Rep. 2013;43(1):32–3.
10. Institute of Medicine. Ed. Olsen LA, Aisner D, McGinnis JM E. The learning healthcare system: Workshop Summary. Roundtable on Evidence-Based Medicine. 2007.
11. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform. 2017; 201(February):1135–44.
12. Zhou SM, Fernandez-Gutierrez F, Kennedy J, Cooksey R, Atkinson M, Denaxas S, et al. Defining Disease Phenotypes in Primary Care Electronic Health Records by a Machine Learning Approach: A Case Study in Identifying Rheumatoid Arthritis. PLoS One. 2016;11(5): e0154515.
13. Forrest BCB, Margolis P, Seid M, Colletti RB. PEDSnet: How a prototype pediatric learning health system is being expanded into a national network. Health Aff. 2014; 33(7):1171–7.
14. Korf BR, Berry AB, Limson M, Marian AJ, Murray MF, O'Rourke PP, et al. Framework for development of physician competencies in genomic medicine: report of the Competencies

Working Group of the Inter-Society Coordinating Committee for Physician Education in Genomics. Genet Med. 2014;16(11):1–6.
15. Eggington J, Burbidge L, Copeland K, Roa B, Pruss D, Bowles K, et al. Current Variant of Uncertain Significance Rates in BRCA1 , BRCA2 and Lynch Syndrome (MLH1, MSH2, MSH6, PMS2, EPCAM) Testing. In 2012. p. 1–3.
16. Riddihough G, Zahn LM. What Is Epigenetics ? Science (80-). 2010;330(6004):611.
17. Wright R, Saul RA. Epigenetics and Primary Care. Pediatrics. 2013;132(Supplement 3):S216-223.
18. McCarty C, Chisholm R. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. BMC Med. 2011;
19. Larson EA, Wilke RA. Integration of Genomics in Primary Care. Am J Med. 2015;128(11):1251e1-1251e5.
20. Kaye J. The tension between data sharing and the protection of privacy in genomics research. Annu Rev Genomics Hum Genet. 2012 Jan;13:415–31.
21. Garrison NA, Sathe NA, Antommaria AHM, Holm IA, Sanderson SC, Smith ME, et al. A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States. Genet Med. 2016 Jul 19;18(7):663–71.
22. Arias JJ, Pham-Kanter G, Gonzalez R, Campbell EG. Trust, vulnerable populations, and genetic data sharing. J Law Biosci. 2015;(October):lsv044.
23. Steinsbekk KS, Kåre Myskja B, Solberg B. Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? Eur J Hum Genet. 2013;21(9):897–902.
24. Dyke SO, Dove ES, Knoppers BM. Sharing Health-Related Data: A Privacy Test? Genome Med. 2016;16024:1–6.
25. Gymrek M, McGuire A, Golan D, Halperin E. Identifying personal genomes by surname inference. Science (80-). 2013;339:321–4.
26. Shringarpure SS, Bustamante CD. Privacy Risks from Genomic Data-Sharing Beacons. Am J Hum Genet. 2015;97(5):631–46.
27. Global Alliance for Genomics and Health (GA4GH). Data Sharing Lexicon. 2016.
28. Hazin R, Brothers KB, Malin BA, Koenig BA, Sanderson SC, Rothstein MA, et al. Information Into Electronic Health Records. 2014;15(10):810–6.
29. Wright CF, Hurles ME, Firth H V. Principle of proportionality in genomic data sharing. Nat Rev Genet. 2015;17(1):1–2.
30. Walport M, Brest P. Sharing research data to improve public health. Lancet. 2011 Feb 12;377(9765):537–9.
31. Harris TL, Wyndham JM. Data Rights and Responsibilities: A Human Rights Perspective on Data Sharing. J Empir Res Hum Res Ethics. 2015;10(3):334–7.
32. Knoppers BM. Framework for responsible sharing of genomic and health-related data. Hugo J. 2014 Oct 17;8(1):3.
33. Wilbanks J, Friend SH. First, design for data sharing. Nat Biotechnol. 2016;1–3.
34. Adkins DE. Machine learning and electronic health records: A paradigm shift. Am J Psychiatry. 2017;174(2):93–4.
35. Fullerton SM, Lee SS-J. Secondary uses and the governance of de-

identified data: lessons from the human genome diversity panel. BMC Med Ethics. 2011;12(1):16.
36. Shabani M, Borry P. Challenges of web-based personal genomic data sharing. Life Sci Soc Policy. 2015;11(1):3.
37. Tucker K, Branson J, Dilleen M, Hollis S, Loughlin P, Nixon MJ, et al. Protecting patient privacy when sharing patient-level data from clinical trials. BMC Med Res Methodol. 2016 Jul 8;16(S1):77.
38. Vos S, van Delden JJM, van Diest PJ, Bredenoord AL. Moral Duties of Genomics Researchers: Why Personalized Medicine Requires a Collective Approach. Trends Genet. 2017;33(2):118–28.
39. Anderson JA, Hayeems R, Shuman C, Szego MJ, Monfared N, Bowdin S, et al. Predictive Genetic Testing for Adult-Onset Disorders in Minors: A Critical Analysis of the Arguments For and Against the 2013 ACMG guidelines. Clin Genet. 2014 Jul 21;
40. Evans JP. Return of results to the families of children in genomic sequencing: tallying risks and benefits. Genet Med. 2013 Jun;15(6):435–6.
41. Holm IA, Savage SK, Green RC, Juengst E, McGuire A, Kornetsky S, et al. Guidelines for return of research results from pediatric genomic studies: deliberations of the Boston Children's Hospital Gene Partnership Informed Cohort Oversight Board. Genet Med. 2014 Jul;16(7):547–52.
42. Knoppers BM, Deschênes M, Zawati MH, Tassé AM. Population studies: return of research results and incidental findings Policy Statement. Eur J Hum Genet. 2013 Mar;21 (3):245–7.
43. Knoppers BM, Zawati MH, Sénécal K. Return of genetic testing results in the era of whole-genome sequencing. Nat Rev Genet. 2015;16(9):553–9.
44. Carroll J, Makuwaza T, Manca D, Sopcak N, Permaul J, O'Brien MA, et al. Primary care providers ' experiences with and perceptions of personalized genomic medicine. Can Fam Physician. 2016;62(October): 626–35.
45. Houwink EJ, van Luijk SJ, Henneman L, van der Vleuten C, Jan Dinant G, Cornel MC. Genetic educational needs and the role of genetics in primary care: a focus group study with multiple perspectives. BMC Fam Pract. 2011;12(1):5.
46. Khoury M, Gwinn M, Yoon P. The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and. Genet Med. 2007;9(10):665–74.
47. Kaye C, Korf B. Genetic literacy and competency. Pediatrics. 2013;132(Suppl 3):S224-30.
48. Saul RA. Genetic and genomic literacy in pediatric primary care. Pediatrics. 2013;132(Suppl 3):S198-202.
49. Agurs-Collins T, Ferrer R, Ottenbacher A, Waters EA, O'Connell ME, Hamilton JG. Public Awareness of Direct-to-Consumer Genetic Tests: Findings from the 2013 U.S. Health Information National Trends Survey. J Cancer Educ. 2015;30(4):799–807.
50. Bartlett G, Avard D, Knoppers BM. A new twist on an old problem: primary care physicians and results from direct-to-consumer genetic testing. Per Med. 2013 Nov;10(8): 827–33.

51. Thorogood A, Deschênes St-Pierre C, Knoppers BM. Substitute Consent to Data Sharing: A Way Forward for International Dementia Research? Oxford J Law Biosci. 2016;1–26.
52. Burstein MD, Robinson JO, Hilsenbeck SG, McGuire AL, Lau CC. Pediatric data sharing in genomic research: attitudes and preferences of parents. Pediatrics. 2014 Apr;133(4): 690–7.
53. McGuire AL, Oliver JM, Slashinski MJ, Graves JL, Wang T, Kelly PA, et al. To share or not to share: A randomized trial of consent for data sharing in genome research. Genet Med. 2011 Nov;13(11):948–55.
54. Faden RR, Kass NE, Goodman SN, Pronovost P, Beauchamp TL, Ethics C. An Ethics Framework for a Learning Health Care System: Hastings Cent Rep. 2013;43(1):S16–27.
55. Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, et al. Sharing Clinical Trial Data: A Proposal From the International Committee of Medical Journal Editors. Ann Intern Med. 2016 Apr 5;164(7):505.
56. Kaye J, Heeney C, Hawkins N, de Vries J, Boddington P. Data sharing in genomics--re-shaping scientific practice. Nat Rev Genet. 2009 May;10(5):331–5.
57. ACMG Board of Directors. Laboratory and clinical genomic data sharing is crucial to improving genetic health care: a position statement of the American College of Medical Genetics and Genomics. Genet Med. 2017;
58. Dragusin R, Petcu P, Lioma C, Larsen B, J??rgensen HL, Cox IJ, et al. FindZebra: A search engine for rare diseases. Int J Med Inform. 2013;82(6):528–38.
59. Mascalzoni D, Paradiso A, Hansson M. Rare disease research: Breaking the privacy barrier. Appl Transl Genomics. 2014;3(2):23–9.
60. Darquy S, Moutel G, Lapointe A-S, D'Audiffret D, Champagnat J, Guerroui S, et al. Patient/family views on data sharing in rare diseases: study in the European LeukoTreat project. Eur J Hum Genet. 2015 Jun 17;epub ahead.
61. Brownstein CA, Holm IA, Ramoni R, Goldstein DB. Data sharing in the undiagnosed diseases network. Hum Mutat. 2015 Oct;36(10):985–8.
62. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet. 2014 May;15(5):335–46.
63. Rehm HL. Evolving health care through personal genomics. Nat Rev Genet. 2017;18(4):259–67.
64. Aksentijevich I, Torosyan Y, Samuels J, Centola M, Pras E, Chae JJ, et al. Mutation and haplotype studies of familial Mediterranean fever reveal new ancestral relationships and evidence for a high carrier frequency with reduced penetrance in the Ashkenazi Jewish population. Am J Hum. 1999;64(4):949–62.
65. Fernandez C V, Strahlendorf C, Avard D, Knoppers BM, O'Connell C, Bouffet E, et al. Attitudes of Canadian researchers toward the return to participants of incidental and targeted genomic findings obtained in a pediatric research setting. Genet Med. 2013 Jul;15(7):558–64.
66. Burke W, Matheny Antommaria AH, Bennett R, Botkin J, Clayton EW, Henderson GE, et al. Recomm-endations for returning genomic incidental findings? We need to talk! Genet Med. 2013 Nov;15(11):854–9.

67. Harris ED, Ziniel SI, Amatruda JG, Clinton CM, Savage SK, Taylor PL, et al. The beliefs, motivations, and expectations of parents who have enrolled their children in a genetic biorepository. Genet Med. 2012;14 (3):330–7.
68. Hughes LS, Phillips RL, DeVoe JE, Bazemore AW. Community Vital Signs: Taking the Pulse of the Community While Caring for Patients. J Am Board Fam Med. 2016;29(3):419–22.
69. Friesen P, Kearns L, Redman B, Caplan AL. Rethinking the Belmont Report? Am J Bioeth. 2017;17(7):15–21.
70. Pearl O 'rourke P, Healthcare P. The Final Rule: When the Rubber Meets the Road. 2017;5161(July).
71. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013 Jul;15 (7):565–74.
72. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405–23.