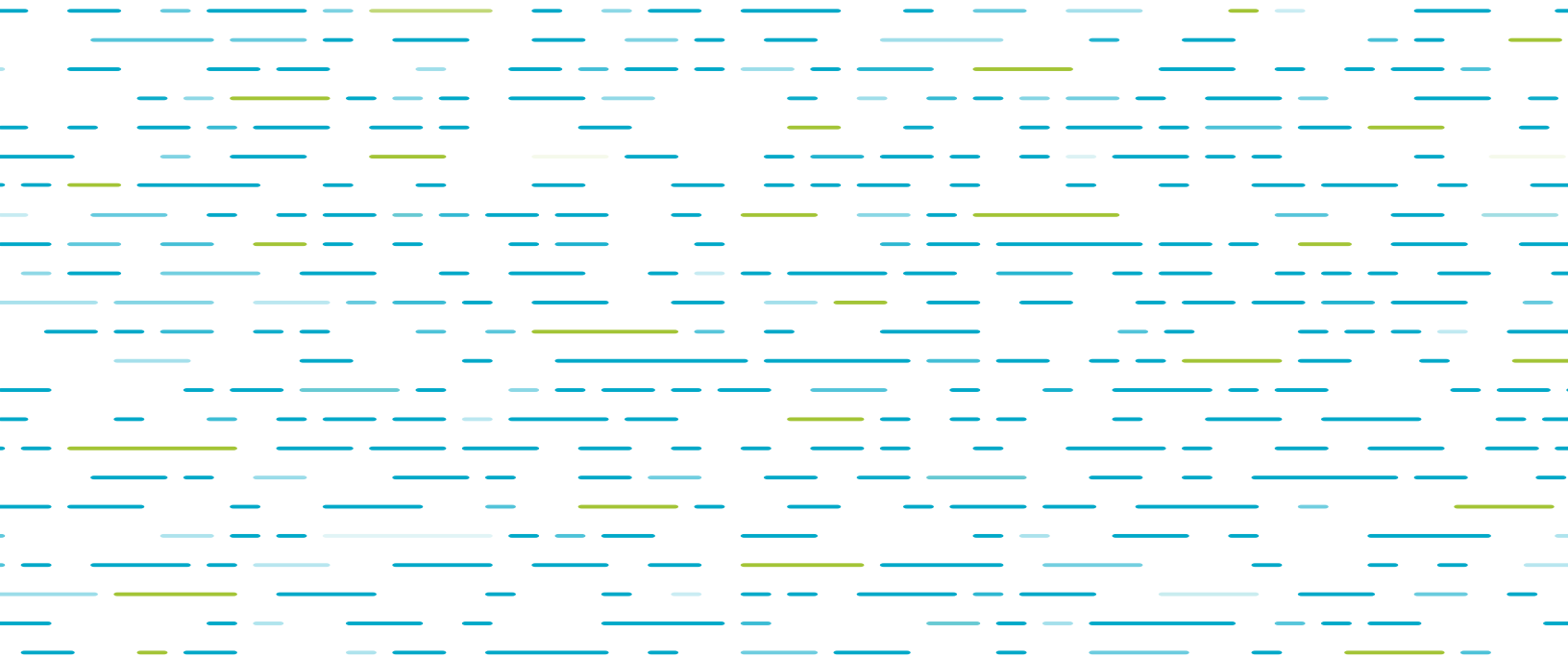
USING HEALTH DATA TO ADVANCE PERSONALIZED MEDICINE

Challenges and a Path Forward



PMC

This paper was developed by Innovation Horizons for the Personalized Medicine Coalition. The issues and opportunities discussed herein were identified by members of PMC's Health Data Working Group.

Executive Summary

The devastating Covid-19 pandemic has refocused the public's attention on the need for faster health care innovation to meet new challenges, better data for public health monitoring, and the urgency of finding treatments that work well for diverse patient populations.

It also highlighted the need for public engagement and transparency of information. The experience has taught us the critical importance of good, timely data from multiple sources to guide decisions regarding key interventions, timely investments in capacity, and evidence to determine what works well for different patient cohorts. We have realized that the same virus can cause highly differing outcomes in different people, and that therapies often don't work equally well for everyone.

The Personalized Medicine Coalition (PMC) defines personalized medicine as an evolving field in which physicians use diagnostic tests to determine which medical treatments will work best for each patient or use medical interventions to alter molecular mechanisms that impact health. By combining diagnostic tests as well as an individual's medical history, circumstances, and values, health care providers can develop targeted treatment and prevention plans with their patients. Personalized medicine can improve treatment outcomes across a range of diseases and, by targeting treatments to those who will benefit, add to the efficiency and effectiveness of health care. This vision rests on the promise of sharing and aggregating data.

Designed to inform Congressional discussions on the future of data usage in the United States, *Using Health Data to Advance Personalized Medicine: Challenges and a Path Forward* discusses challenges and opportunities related to the access to and use of new sources of high-quality data that can reveal the unique characteristics of individuals and provide strategic insights to manage health and disease conditions. This document is intended as a synopsis of key issues and policy priorities that need to be addressed to enable the maximal use of clinical data to advance personalized medicine, ensure the integrity of data, and protect vital data resources in ways that do not create barriers to research and health care applications. Executing a plan to balance these priorities will go a long way toward catalyzing efforts to accelerate innovation and bring novel and more effective personalized medicine solutions to patients.

In recent decades, biomedical research has been transformed by the ability to analyze massive amounts of data to better understand prevention, diagnosis, and treatment opportunities in health care. The genomic revolution over the past two decades has also provided an important new window on the unique biological attributes

Thus, a critical challenge of our time is to embrace the enormous opportunity for using diverse sources of data to improve health while protecting the privacy of individuals.

of an individual, and their impact on health. The use of big data complements "hypothesis-driven" research such as randomized clinical trials. Meanwhile, the health care delivery system has undergone a structural transformation through the widespread adoption of digital technologies and information sharing. And health care payment reforms have created incentives for providers to deliver "high-value" care, raising the stakes for delivering coordinated, efficient, and effective treatment identified through research and development of care delivery tools.

Access to and appropriate use of data can capitalize on the convergence of scientific and technical capabilities with simultaneous demands for improved health care outcomes. Multiple data sources contribute to a broader and deeper understanding of health and the interventions we use to preserve it. In addition to "patient" data created within health care systems and leveraged through electronic health records (EHRs), health data are now widely attainable from "consumer" sources such as health apps, wearable devices, home health monitoring devices, patient reports, and information about the social determinants of health.[1] Properly analyzed and integrated into health care decision-making, these novel data sources can augment and complement the results generated through time-tested research methods.

Thus, a critical challenge of our time is to embrace the enormous opportunity for using diverse sources of data to improve health while protecting the privacy of individuals. Starting in 1996, prior to the widespread use of digital health records, the U.S. Congress enacted a regulatory structure for enabling essential uses of data for health care delivery, administration, and research while protecting privacy of patients. In recent years, technology companies have begun harnessing data and applying sophisticated analytics to make inroads into health care delivery and product development. Under these circumstances, protections for individual privacy may not apply. Concerns that personal data could be misused or used to re-identify patients threaten to erode the trust of those who share personal data, creating hurdles for future health care research and innovation. Building on our nation's existing and effective health care privacy rules and regulations, the federal government has an opportunity to revisit the regulatory framework for health data privacy.

A fundamental tension arising from this digital transformation is ensuring the flow of data and

1. Throughout this paper we differentiate *patient* data as information created through health care encounters with the delivery system and subject to *HIPAA* as distinct from *consumer* data, which is health and health-related information generated outside of the health care system and generally not addressed by *HIPAA*. See also Figure 2.

information that enables health care innovation while protecting the privacy of individual health information regardless of its source. Ongoing technical progress is needed to facilitate the interconnection of varied data sources such as EHRs and "real-world" data collected from patient and consumer sources. Rigorous validation methods are now available to ensure integrity of data and valid, equitable analysis. New analytic techniques of machine learning and artificial intelligence applied to these interconnected data sources will enhance our understanding of individual treatment opportunities as well as opportunities for health system improvement. The vision of a "learning health system" that continuously gathers, analyzes, and integrates information at the point of care to assist clinicians and patients in making evidence-based and personalized treatment decisions is attainable within the decade.

Future policies must ensure that science prevails in establishing sound programs that use high-integrity data in well-governed systems built on a foundation of consumer privacy and control. Guideposts, rules, and best practices governing data are needed to ensure consumer protections and the agency of individuals to control their own information, while not throttling access to data for important treatment and system improvements. Future standards must ensure a high level of privacy for all health data to maintain trust in the research and health care delivery process. By cultivating synergy between patient care, data, and innovation, we have the opportunity to realize the potential of personalized medicine as a mainstay of health care.

This white paper addresses data-related opportunities and barriers directly impacting the delivery of personalized medicine in the next three years. The paper identifies three top priorities for collaboration across government, academia, and the health care industry:

1. **STREAMLINE** the development and implementation of coordinated standards for collection and interoperability of EHR data, ensuring that the data are portable, transferable, and fit for appropriate secondary research use to guide personalized medicine.
2. **ENSURE** that new legislative efforts at the federal or state level recognize the careful balance between respecting and protecting the interests of patients, empowering patients with access to their own data, honoring patient choice, fortifying privacy and security protections, and supporting the public interest when addressing use of datasets for research and development to improve quality of care and patient outcomes.
3. **ACCELERATE** development and adoption of practices and standards for datasets created from real-world data to enable maximal use by researchers, clinicians, patients, regulators, and other stakeholders that yields insights for improved decisions, equity, and outcomes in health care.

This paper examines key issues and highlights policy considerations relating to data access and use practices that maximize opportunities to develop personalized treatments necessary for improved health equity and outcomes. The paper describes the importance of data to personalized medicine and identifies key elements of a national approach that ensures data can be used to inform better health care decisions and health outcomes.

Despite our progress on uses of data, new strategies and policies are needed to achieve the promise of personalized medicine. At its conclusion, this white paper discusses strategic approaches — the path forward — that will improve access to and protections for data, enabling the analytic insights to fuel delivery of innovative, personalized treatments for those in need.

Introduction

Bridging the gap between the burgeoning engine of biomedical research and medical need is essential to accelerating the pace of new therapies.

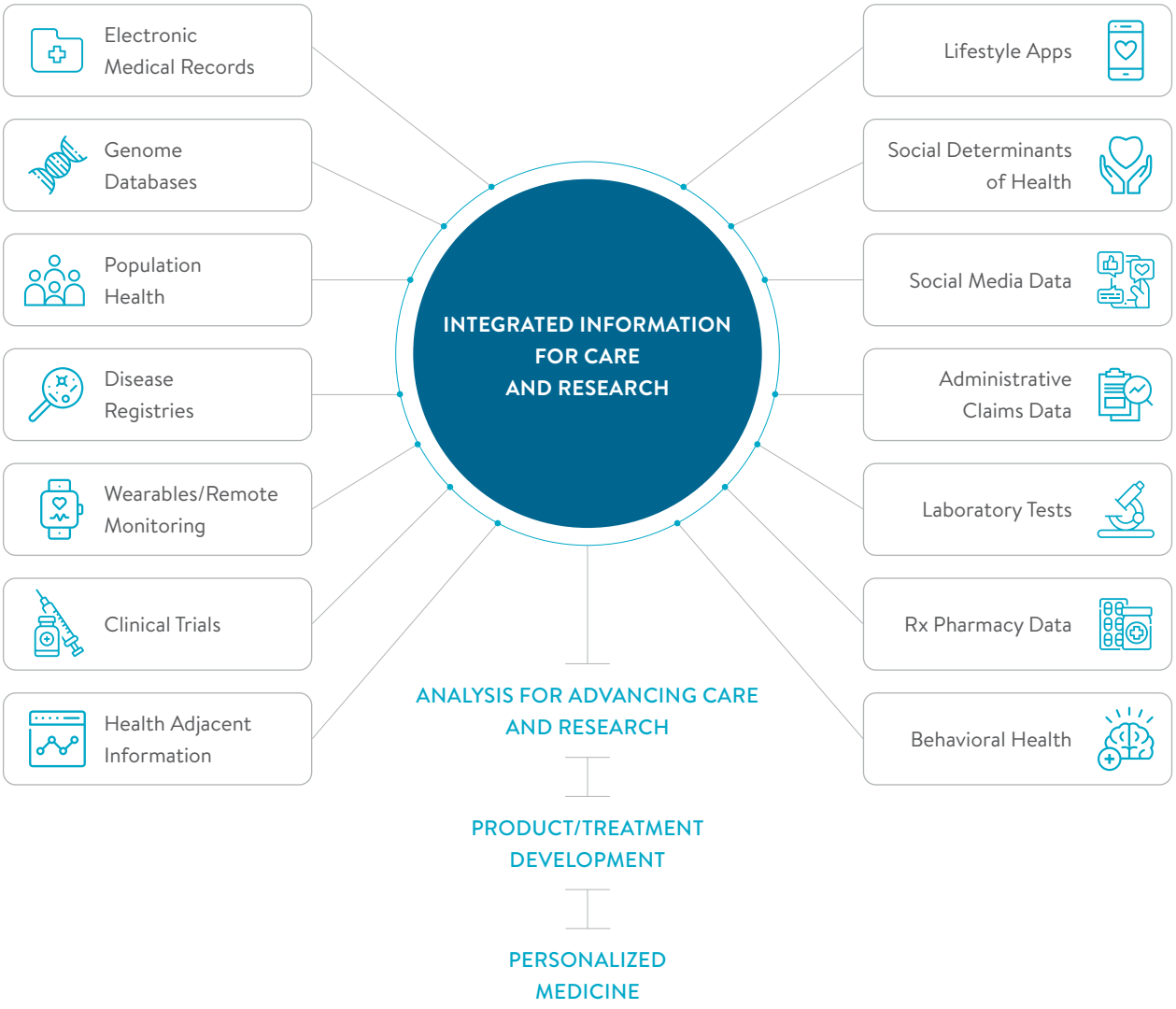
Two key elements of infrastructure needed to implement personalized medicine are innovative information technology and expanded access to new data sources. The future of personalized medicine in the American health care system depends on our ability to analyze data to guide clinicians and patients to individualized approaches that maintain wellness and overcome deadly diseases. Improved use of data and information will underpin other transformations throughout the health care system, adding value in health care services and driving improvements in efficiency, effectiveness, and equity of care.

The nation's experience responding to the Covid-19 pandemic illustrates both the promise and peril of using big data to improve health. While many members of the public rapidly authorized applications to track their health symptoms or to alert them of possible viral exposure, the Covid-19 example also illustrates how data can be used to understand patterns in disease and risks to certain populations. The pandemic experience also highlights gaps in the nation's data infrastructure and the need to ensure transparency and privacy of consumer information. See the "Lessons Learned" box below for details.

New sources of data on health and health care present tremendous opportunities to generate knowledge and improve practice. Research studies and clinical trials that have relied on clinical data from EHRs, health care claims, pharmacies, and laboratories are integrating new forms of data such as genomic information. Data from new sources such as remote health monitoring sensors can be captured with advances in telecommunications and digitization of information. Patient-reported outcomes and social determinants of health data will increasingly be used to help researchers understand with greater precision individual differences in patients' risk of disease, adverse event experiences, and response to therapies. Figure 1 illustrates some of the rapidly expanding new types of data that contribute insights to health care research and delivery. This digital information can inform researchers and clinicians about patients' risks, health conditions, and responses to therapy.

The power unleashed by the use of data to characterize disease and molecular pathways has given rise to new industries, platforms, and networks to support researchers in translating data into precise diagnostics and therapeutics. As we look to the future, breakthrough advances such as gene modification strategies that represent the cutting-edge of science and medicine will rely upon robust data resources to carefully calibrate individualized therapeutic application. The integration of machine learning (ML), powerful computer-guided data collection and analytics engines, and artificial intelligence (AI) into algorithm-aided medical practices and decision-support present new opportunities to recognize and act on improvement opportunities in health care.

FIGURE 1: Sources of Data Contributing to a Comprehensive Understanding of Health



Delivery of personalized medicine requires effective analysis of data from a wide array of sources and the engineering of learning systems that integrate results of those data into clinical practice, thereby creating a continuous loop of knowledge generation that allows providers and patients to apply this knowledge more effectively toward clinical decision-making. Experts have deemed this approach to have a "fly-wheel" effect whereby the system initially overcomes inertia and inefficiency to create momentum for knowledge development and accelerate its application in health care.

Enhanced Data Policy Will Improve Health Care Delivery and Outcomes

- ☑ Improved development of models that predict disease in diverse populations
- ☑ Enhanced understanding of complex diseases
- ☑ Better application of effective interventions and diagnostics for all communities
- ☑ Avoidance of treatments that do not work
- ☑ Accelerated opportunities for patients to have access to and benefit from emerging, promising medical discoveries
- ☑ Lower data management burden for clinicians
- ☑ Better tools to support clinician and patient treatment decisions
- ☑ Increased access to personalized treatments
- ☑ Consistent delivery of high value

Delivering on the Promise of Personalized Medicine

A key objective of personalized medicine is to frame medical decisions based on highly individualized patient characteristics such as health care history, genomic and molecular data, imaging, physiological and behavioral biomarkers, and patient preferences. The translation of innovative molecular and data analytic strategies into practices that can benefit all patients is a centerpiece of the "big data" era in health care.

To deliver the promise of personalized medicine, it is essential to simultaneously translate information into insights in order to help clinicians make the right decisions at the point of care. In practical terms, this requires greater alignment and integration of research with clinical practice, with a bi-directional flow of information. Some of the advantages of scaled personalized medicine will be better medication effectiveness, reduction of adverse events, lower costs through optimized therapy, enhanced diagnostic capability through the use of molecular information, improved disease management with the help of wearable sensors and mobile health applications, and smarter design of clinical trials.

Effective use of data enables "personalization at scale" — that is, delivering personalized medicine at all levels of health status to all individuals. Ultimately, we may be driving towards a "learning health system" that provides a supportive environment for personalized medicine by continuously integrating information, particularly about the understanding of genomic data, and making it available to clinicians at the point of care through decision-support tools.

Need for Action

While the genome revolution has resulted in increased understanding of disease stratification and targeted molecular therapies, these advances have not yet had a universal impact of improved clinical outcomes and personalized medicine. The full promise of personalized medicine goes unfulfilled due, in part, to gaps in access to data and usability of information critical to decision-making in patient care. The interpretation of a genetic test result, for example, may have tremendous value or none at all depending on the context of its application and whether the new information is integrated to present more complete knowledge of the patient's biological state, environmental influences, and social determinants of health. All signals from the accelerating progress of research suggest that the realization of that vision may not be far off.

Lessons Learned About Data From the Covid-19 Public Health Emergency

The Covid-19 public health emergency yielded important lessons relevant to the future of personalized medicine. Importantly, the public witnessed first hand how genetic variability influenced individual responses to the disease and response to treatments.

Responses also reflected diverse uses of data for research, treatment and population health. For example:

- Uses of non-health care generated data — consumer health data — derived from sources such as telecommunications (cellphones and sensors), retail commerce, and transportation were crucial for public health strategic planning.
- Innovative payment programs designed to incentivize use of digital health care services such as video visits were rapidly and safely deployed, and widely embraced by providers and patients alike.
- Regulatory agencies quickly modified reporting requirements and used enforcement discretion for privacy regulations to accommodate the public health emergency without compromising patient privacy and individual protections.
- Data from software applications embedded in mobile devices and telehealth platforms were analyzed for population-level information on the use of public testing and vaccination and for travel/physical activity patterns during the pandemic.
- Government, institutional leaders, and the public received a practical crash course in population health statistics; very rapidly the use of data dashboards and heat maps were as commonplace in our daily media dialogue as weather and traffic reports.

Overall, the features of transparency, resiliency, and interoperability of data across disparate information systems were key to our national Covid-19 response strategy and hallmarks of success. The demand for timely and accurate data brought about new networks formed by data-sharing agreements among health care institutions and private sector companies with unprecedented levels of collaboration.

Despite these advances in the uses of data, the Covid experience also revealed shortcomings in public health data information systems and wide disparities in institutional digital resources across the country. It also revealed distrust in data and evidence that must be addressed through greater transparency, education and public engagement. There is no doubt from this experience that robust, reliable, and secure data systems are a crucial element to the future of personalized medicine and vital to our future national security.

The rapid pace of technology innovation has revealed new challenges in our regulatory framework, technical infrastructure for data sharing, and emerging issues in consumer privacy protections. Work is needed to optimize the use of data to help clinicians deliver better care and help patients experience better health outcomes. During this period of transformation, it is also critical to ensure that data are available to researchers and to properly vetted product developers. This will require the adoption of technology solutions along with legal privacy protections for patients to ensure the innovation needed for the full potential of personalized medicine is realized. Without the right balance, data use will be throttled and the pace of biomedical innovation and advancement slowed.

Current Policy Framework for Data Integration

Many federal laws have been enacted to overcome barriers to safe, effective, and efficient access to and use of data. These initiatives have been supported by diverse stakeholders seeking to advance personalized medicine, including patients/consumers, policymakers, health care leaders, and researchers. Critical legislation and related updates impacting data use include the *Health Insurance Portability and Accountability Act* of 1996 (*HIPAA*), which enacted privacy protections for the class of data called protected health information (PHI) generated by health care organizations. *HIPAA* also promoted the adoption of EHRs. Later, the *Genetic Information Non-discrimination Act* (*GINA*), the *Health Information Technology for Economic and Clinical Health* (*HITECH*) *Act*, and the landmark *21st Century Cures Act* of 2016 (*Cures Act*) established guideposts for the use of data and incentivized the adoption of digital technology, among other provisions. Subsequent rules and updates have refined many aspects of each of these pieces of legislation and clarified enforcement of the laws.

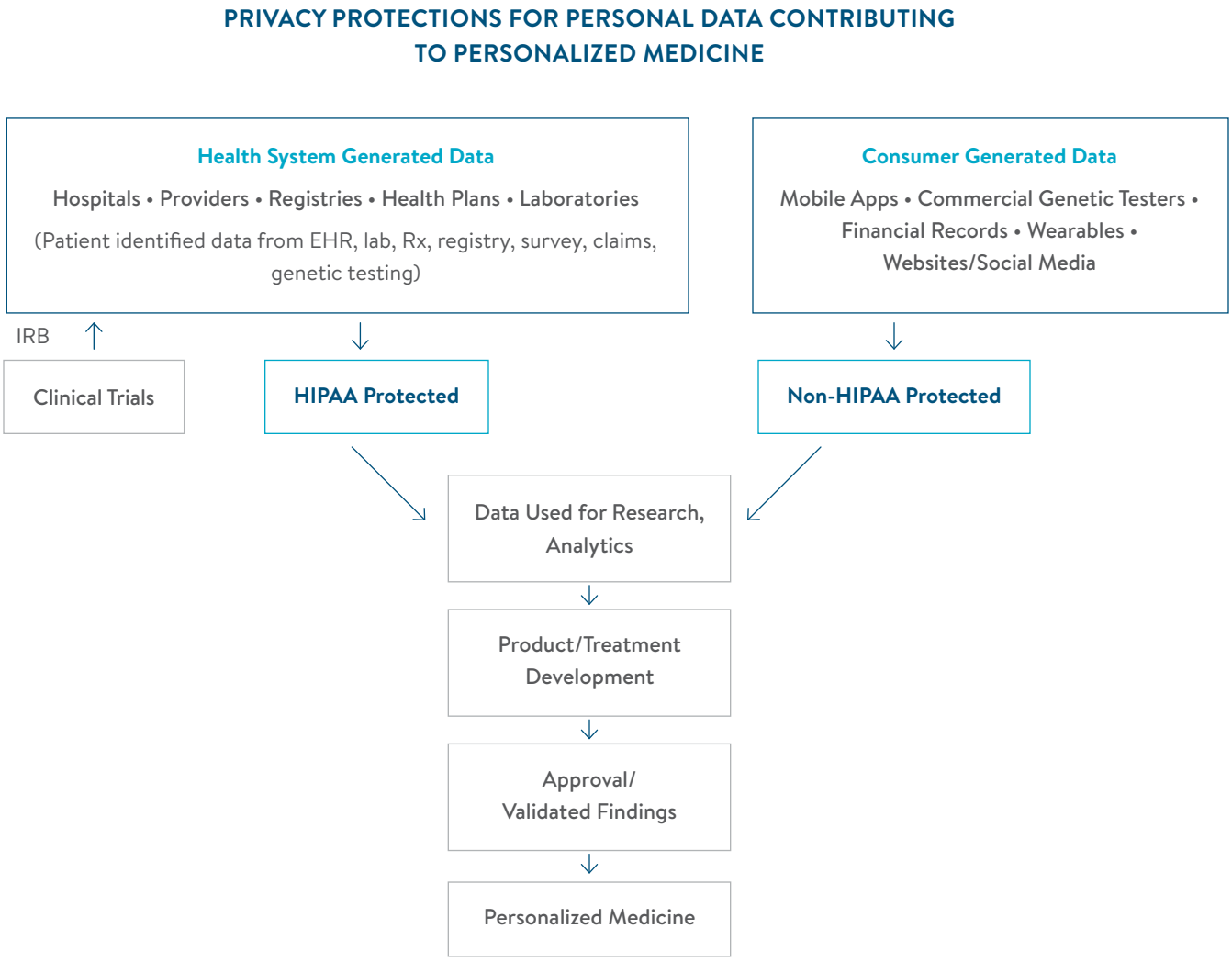
The regulatory structure that has evolved in recent years regarding differing types of health data is complex. Other legislation, for example, the *Medicare Access and CHIP Reauthorization Act* of 2015 (*MACRA*), has impacted information management requirements by tying the adoption of specific practices to health care payments, affecting the use of personalized medicine laboratory tests, prescription coverage, and clinical services. Appendix 1 highlights key features of governing legislation. Appendix 2 lists many of the federal agencies tasked with protecting patients and improving data management. The federal government has played an important role in developing data standards and leveraging payments to increase standardized adoption. Agencies such as the Centers for Medicare and Medicaid Services (CMS), the Food and Drug Administration (FDA), and the National Institutes of Health (NIH) have influential roles in driving data management and research practices by leveraging payment, market approval, oversight, and research funding. The Federal Trade Commission (FTC) has a role in enforcing privacy by prohibiting "unfair or deceptive" practices. FTC authority extends to some mobile app developers, social media sites and technology companies. The FTC does not have specific protections for privacy of health data but does require organizations to abide by their own privacy and security policies relating to health data.

New technologies and data management strategies have created additional regulatory needs. Regulations enacted or under consideration by states and internationally also influence the need

for U.S. action. For example, the internet era and emergence of large technology companies have resulted in the capture of personal information outside the scope of *HIPAA*'s protections. While this consumer-generated data has value to health care research, it also creates new privacy exposure for patients. Figure 2 illustrates sources of health care data and the regulatory divide that currently exists between *HIPAA* and non-*HIPAA* covered data.

State regulations have emerged to fill many gaps in data use and privacy protections, both leading the way and creating compliance challenges. For example, the *California Consumer Privacy Act* of 2018 (*CCPA*), one of the most far-reaching state models, enacts data privacy rights and focuses on consumer control of personal information. While *CCPA* addresses the need for consumer protection, 50 such state-led initiatives could create hurdles to multi-state, multi-center research.

FIGURE 2: General Overview of Federal Oversight Protections for Health Data Sources Supporting Personalized Medicine



Advancing Use of Data to Accelerate Patient-Centric Personalized Medicine

Three key issues will impact how quickly the U.S. can fulfill the promise of data to contribute the research insights necessary to deliver personalized medicine at scale. These key topics — how we use EHR data, how we protect patient privacy, and how we integrate "real-world data" — are examined here.

Adding to the Value of EHR Data

Electronic health records are the digital backbone of the nation's health system. Data from individual patients' clinical EHRs contribute to information underlying personalized medicine. An ideal EHR ecosystem would capture relevant information in a structured format to permit advanced analysis, offer decision-support and coordination capability for clinicians, and enable patients to understand and act on the treatment options best for them. Importantly, increased access to and use of EHR data could complement and augment findings of randomized clinical trials, contributing to more timely and robust research.

EHRs are a decentralized system of commercial and customized applications that capture real-time patient care information to support health care delivery. EHR systems offer the potential to improve health care quality and safety, consistency and standardization of data, and point-of-care access to patients' information. EHRs enable patients to interact with their providers and contribute to sharing of data across health systems, public agencies, and researchers, as allowed by privacy regulations.

Although accessible to patients, EHRs are held by health care providers, and the data therein are subject to *HIPAA* privacy protections. Clinicians enter patient data and are accountable for acting on information within the EHR. A dynamic tension lies between the interest in more comprehensive and multi-source data within the EHR and the imperative to manage the documentation burden on clinicians; while data add value for treatment decisions and research, data entry and management pose a significant burden on clinicians.

A significant barrier to robust use of EHR information is lack of interoperability across different data systems and types. The federal *Cures Act* spurred significant progress in interoperability, and many technical innovations have emerged from the private sector to enable transmission of and access to information. But interoperability challenges still often inhibit the sharing of information with other providers and systems to coordinate effective care, thus impeding the aggregation of data for research. Some barriers to maximizing the value of EHRs are driven by technical challenges, such as the use of "unstructured" data (e.g., data that are not conforming to terminology or information technology coding

standards that often provide valuable observations to understanding health care and environmental conditions); some are due to competition among EHR vendors, health care institutions, and others; and some are the unanticipated result of regulatory compliance. These barriers have been addressed through a variety of legislative initiatives, requirements, and voluntary data standards, but the changes have not yet yielded a system of usable, unobstructed data to support better patient care and research.

From the perspective of patients, EHRs are a primary source of diagnostic, genomic, therapeutic, and outcomes data. Patients have an interest in accessing information about their own care, and importantly, sharing health records with primary and specialty providers participating in their care. Although there are a few digital platforms that permit individuals to share their health records and patients legally have a right to access their records, such tools are vastly underutilized. Furthermore, health-related data from patients' digital devices created outside of the health care system are rarely integrated with EHRs for a variety of reasons. Health care organizations accountable for the EHR are often reluctant to integrate data with commercial products not subject to *HIPAA* protections; may have concerns about managing the volume of information generated outside the health system; or may have technical difficulties integrating data from outside the institutional EHR system.

On the provider side, while many EHRs have been configured to offer decision-support, there is much more that could be offered to lessen the clinician cognitive burden in health care decision-making to enable personalized medicine practices. EHRs have the potential to offer more data synthesis, interpretive insights and information about treatment options that would enable clinicians to make sense of the massive trove of clinical, genomic, and diagnostic information available at the point of care. Now and for the foreseeable future, EHRs are the digital nexus by which patients and providers could collaborate in a secured and reliable way on making personalized treatment decisions.

The clinical data captured in EHRs contributes to important research to understand health care delivery and improve treatment outcomes. EHR data are an important source of insight on health and health care and could be particularly valuable in understanding and addressing health disparities. But the value of EHR data could be

"Health care is a team sport, and various players have to coordinate to provide optimal care to a patient. Secure and consistent patient identity and consent management across players is key to improving patient safety and capturing full value from the billions of dollars invested in electronic health record systems."

KRIS JOSHI, PH.D.

Executive Vice President, President, Network Solutions, Change Healthcare; Co-Chair, PMC Health Data Working Group

vastly improved by increasing standardization of data and reducing the gaps in information. For example, standard terminology and data elements are needed to enable coupling of genomic information with other data to predict a patient's likely response to interventions. Such technical advances would need to be accompanied by appropriate protections for patients, also discussed in this paper.

To enable the full potential of using EHR data for research, improvements are also needed to reduce gaps and inconsistencies in data elements (while minimizing the burden on clinicians). Technical improvements could enable researchers to maximize the use of AI and ML to identify trends, connections, and factors influencing outcomes. This research would enable clinicians to better target, compare, and personalize interventions. When applied effectively to EHR data, AI and ML solutions offer powerful insights to guide the selection of patients who will benefit from specific interventions, thus improving the efficiency, effectiveness, and equity of care. There is also a need for alignment of institutional and federal/state policies enabling researchers to acquire data and integrate it with biological information and clinical care records.

There are signs of progress in the prioritization of interoperability of IT systems and data standardization to enhance access and usability of data. For example, the Department of Health and Human Services (HHS) Office of the National Coordinator for Health Information (ONC) recently published the final rule mandated in the *Cures Act* that addresses interoperability and information blocking, among an expanse of other issues that support personalized medicine. The rule directs the health care industry to adopt standardized application programming interfaces (APIs) to allow individuals to securely and easily access structured digital information that connects EHRs to each other and smartphone applications. The rule standardizes the Fast Healthcare Interoperability Resources (FHIR) data model and the Substitutable Medical Applications, Reusable Technologies (SMART) on FHIR standard, which specifies how consumer and medical apps can support research and care delivery.

The private technology sector has made great strides in enhancing interoperability and data sharing within the existing regulatory structure, often by developing platforms and products to connect essential information. To enhance and accelerate progress, EHRs will need ongoing upgrades in data science and engineering to service care needs with higher degrees of complexity and dimensionality of the data. Going forward, expanded domains — including genomic and consumer generated data — must be integrated with patient records to address the complex issues driving poor health outcomes. Updated oversight and processes for ensuring privacy of these data along with modernized processes for patient consent will ensure that data and analytics can be coupled to drive health care innovation. Teasing out the full value of data captured within EHR systems will be part of the solution to breaking down barriers and enabling the benefits of equitable, personalized medicine to reach all Americans.

Protecting Patient Privacy While Encouraging Innovation

Personalized medicine relies on highly individualized information about people — such as genetic profiles — coupled with targeted therapeutic solutions delivered at the right time and place. AI and ML offer the capability to meaningfully analyze enormous datasets for patterns that

"The data needed to advance personalized medicine have been siloed for too long. To optimize patient care and develop more tailored treatment options, we need policies that facilitate the use of multimodal data at scale."

LAUREN SILVIS

Senior Vice President of External Affairs, Tempus; Co-Chair, PMC Health Data Working Group

predict patient responses to personalized therapeutic regimens and validate clinical research. The widespread use of large datasets requires the highest level of privacy protections for individual health information. It is both a policy and technical challenge to find the right balance between information accessibility and privacy.

One barrier to the use of patient information for research (rather than individual patient care) is ensuring proper protections of patient health data. As shown in Figure 2, identifiable "patient" data are protected by *HIPAA*, but "consumer" data generated outside the health care system, often containing equally sensitive information, are not protected as PHI. Some types of individual data, such as mental health, substance use, and genomic information have been subject to explicit privacy regulation. Violations of *HIPAA* and other privacy laws are subject to enforcement, including fines. Institutions seeking interoperability of data across different health care systems and with consumer devices must navigate the uneven privacy protections applied to these data sources.

Current approaches to privacy protections range from obtaining consent for use of health care data, de-identifying data used for research, requiring "opt in" as a condition of using the service (common among consumer apps), and adoption of advanced technology such as blockchain or digital ledger technology systems to assure secure transmission of data across settings. In scientific and health care settings, "data governance boards" may be charged with approving use of data to protect patient privacy as well as enhancing the quality of data and data sharing. These boards may also approve use of resources (including funding and technical support personnel) to ensure privacy of data. The private sector has created innovations such as "trust networks" that facilitate consumer consent for data use based on the individuals' consent parameters.

For the massive amount of data about health conditions not subject to *HIPAA* regulations, there is little clearly enforceable protection of patient privacy. Recently, headline-generating articles illustrate consumer confusion and concern about how their information is used by apps and technology firms, and lack of clarity on what information is *HIPAA*-protected versus non-protected. Use of personal data by large technology and commercial apps has come as a surprise to many consumers.

"Unlocking the insights of real-world data — which chronicles the experiences of many more patients than the few who participate in clinical trials — has tremendous potential to answer important scientific questions. The more we learn about each patient, particularly in oncology, the closer we get to truly personalized medicine."

MICHAEL VASCONCELLES, M.D.

Chief Medical Officer, Flatiron Health; Co-Chair, PMC Health Data Working Group

To maintain trust and confidence in our system of protections, more transparency is needed to inform patients/consumers on how their personal data — from any source — may be used. The flow of data across organizations conducting research, developing clinical-support tools, and delivering care is highly complex. Organizations in the data transfer chain are often accessing data multiple steps removed from the patient. Stakeholders may need to engage in a larger discussion about how to educate the public properly to provide sufficient consent for the myriad of important uses. Consumers should be involved in the development of data sharing approaches and oversight. Private sector organizations may have an important role in facilitating dialogue and developing consensus on strategies to effectively navigate data access, transparency, and patient authorization for data use.

On the technical side, additional investment is needed to develop the sophisticated identification and authentication systems necessary to facilitate both transparency and authorization of data use. These systems should empower consumers with greater control over their health information without establishing cumbersome consent procedures that could hinder research use. A priority will be to implement effective privacy protections for consumers regardless of the origin of data, while promoting the development of technical approaches to empower consumers such as streamlined interoperability, digital authorization, and individual authentication.

Over time, duplicative or conflicting state and federal regulations could hinder aggregation of data, multi-state research and even global sharing of information. Downstream, a fragmented approach to privacy protection could limit patients' access to clinical trials and the essential therapies emerging from multi-state, multi-center research programs.

Protecting individual patient privacy while making data available for research and health care innovation is a high priority in personalized medicine. This requires scalable data privacy and consent requirements based on the type of data and intended use. It will be important to adopt a regulatory framework that protects privacy with requirements aligned across federal agencies and consistent across states while fostering innovation.

Safeguards for consumer patient health data should be enforceable regardless of where the data originate. There will need to be substantial analysis conducted to ensure that appropriate protections are developed that do not impede research and health care applications. Policy strategies adopted in the 2008 *GINA* legislation that were crucial to patient protections may offer a useful playbook for achieving the consensus needed for progress to happen elsewhere. The need for *GINA* protections engaged stakeholders in careful analysis and adopted stakeholder agreements to reach a national standard that protects individuals from loss of insurance due to information available from genetic test results or pre-existing conditions. *GINA* offered the protections necessary to enable genomic information to become more widely available for research and personalized medicine treatments.

Both *HIPAA* and *GINA* initially spurred concerns among stakeholders that patient protections would impede progress and innovation. However, both laws have stood the test of time; through careful consideration of research needs and a series of amendments and modifications over the years, they are now the trusted framework for regulation and enforcement and are an important foundation for personalized medicine. Now is the time to create a new path forward for ensuring the safe and appropriate future use of new data sources without creating barriers to patient access and research.

Strengthening Adoption of Real-World Data and Evidence

Real-world data (RWD) are the information generated outside of controlled clinical trials that contribute to our understanding of health and health care interventions. Real-world evidence (RWE) is based in the analysis of RWD collected from a variety of sources, including EHRs, claims/billing, disease registries, wearables, and patient-generated data. These data maximize the data already routinely captured in clinical care. RWD/E can fill information gaps and provide a rich longitudinal source of information on patient characteristics and patient experiences with medical products. For example, RWD/E can help regulators and researchers understand the natural history of a disease (as was done during the Covid-19 pandemic), aid in comparative effectiveness research, enhance drug labeling, and serve as a complement to clinical trial data where the study populations do not mirror the potential population impacted by the treatment.

The challenge ahead lies in optimizing the use of RWD by developing effective processes for producing, acquiring, organizing, and analyzing information to create meaningful insights and to manage potential errors in data. There are risks to using RWD in lieu of more controlled clinical trial data. For example, in real-world datasets, it is difficult to tell if there has been accidental inclusion of extraneous data or if information is missing. ML and AI technology "learn" by analyzing data from human systems and existing treatment patterns,

RWD Use Case — Muscular Dystrophy

The value of RWD was evidenced in a clinical trial resulting in the 2016 FDA approval of the drug Eteplirsen, a treatment for Duchenne muscular dystrophy. Data mining was used to create an RWD control group, which would otherwise have been impossible due to the rarity of the condition. The RWD control group eliminated the need for giving a placebo, often deemed unethical when there is the potential for a useful treatment. This approval was considered a strong use case demonstrating the value of RWD in the study of rare disease.

meaning an algorithm resulting from AI could perpetuate or exacerbate an inequity present in the underlying data. Transforming RWD into RWE for research and/or regulatory purposes requires a multistep, rigorous approach. The development of scientific, operational, and technical best practices to ensure the integrity of data and analysis is needed to inform subsequent guidance regarding principles such as traceability to source, data quality, validation, transparency, and reproducibility.

Some key benefits of utilizing RWD to influence personalized medical decision-making include:

- Improving clinical trial data, promoting diversity, and validating results in larger populations;
- Identifying and promoting opportunities for effective evidence-based treatment decisions and understanding comparative effectiveness of different treatments; and
- Improving health equity by mitigating selection bias, expanding participation in research, and helping to identify unmet clinical needs in diverse populations.

AI and ML have dramatically increased the capability of researchers to organize and analyze data points. Properly analyzed, RWD can provide insights necessary to target medical interventions and disease prevention for individuals and populations. The mining of large datasets through AI and ML can provide the RWD needed to identify possible treatments for specific cancers or rare diseases. It can also drive development and adoption of preventive therapies based on analysis of patients' molecular, clinical, behavioral, and environmental information.

FDA has taken a leading role in helping to guide the use of RWD/E in scientific studies and defining how RWD are assessed in the regulatory approval process. In 2018, FDA released the *Framework for FDA's Real-World Evidence Program*, outlining how RWD/E are used in FDA regulatory decision-making. NIH has also recognized the value of RWD/E and has taken action to ensure the integrity of RWD/E in research. NIH developed "FAIR" principles on using data (including RWD) for research. FAIR principles are applied to NIH research programs to ensure data are "Findable, Accessible, Interoperable, and Reusable." In 2020, NIH released community-developed "TRUST" principles to promote the adoption of "Transparency, Responsibility, User focus, Sustainability, and Technology" in data repositories. These data standards will be important to ensure integrity of RWD as it becomes more widely used.

Adoption of RWD to enhance research will be facilitated through implementation of widely recognized data governance arrangements. Data governance creates a system of standards and policies that provide a framework for ethical, valid, and high-integrity use of data. Private sector data governance organizations complement regulatory agencies. As non-governmental organizations responding to federal requirements, they have the potential to solve problems and develop standards quickly to help optimize data sharing and respond nimbly to market innovations.

The work to standardize and make EHRs interoperable, discussed earlier in this paper, will make an important contribution to the availability of RWD/E from EHRs and amplify the research value of genetic information. RWD will greatly enhance insights available from EHRs to better understand clinical and operational improvement opportunities. RWD research will analyze data on real-world practices to contribute to a more effective, efficient health system. Application of RWD will help lead to better, more personalized treatments with a greater probability of effectiveness. Use of RWD may also increase efficiencies in health care workflow and operations.

Properly managed, RWD offers other research and development opportunities that could improve existing methods for identifying and measuring the impact of current and future

health care interventions. For example, in real-world datasets, researchers could sample longitudinal data periodically to monitor the impact of a new treatment and evaluate outcomes over time. This use of RWD could enable researchers to monitor safety and effectiveness of pharmaceuticals in the marketplace and capture data relevant to other treatment indications. RWD/E could expand and refine indications for therapeutic drugs, molecular medicines, and medical devices, accelerating the approval process for essential therapies and medications.

Another use of RWD could be to help to reduce inequities in health care. RWD analytic methods can ensure that data are available for under-represented populations such as the elderly, racial and ethnic minority groups, and people with poorer health status, filling the gaps stemming from under-enrollment in clinical trials. For example, RWD could be derived from searchable databases to identify individuals with a targeted rare condition, enhancing recruitment into clinical trials. Ultimately, the capability to study the impact of an intervention in diverse populations should enable more equitable targeting of therapies to those who are most likely to benefit — adding overall value to health care. To capture the full value of RWD, it will be important to implement safeguards to prevent unintended bias in the analytic process. Implementation of RWD data standards will ensure that the experience and outcomes of underrepresented groups are examined to ensure equity in machine analysis, treatment algorithms, and health outcomes.

RWD Use Case — Improved Treatment for Breast Cancer

The new era of targeted drugs for personalized medical care has required new methods for testing treatment options. Incorporating RWD with randomized controlled trials (RCTs) allows modifications to the trial and/or statistical procedures of the trial after initiation, without undermining RCT validity and integrity. The use of RWD allows for studies to validate the generalizability of clinical findings in heterogenous populations.

For example, a March 2021 study published in Breast Cancer Research used the Flatiron Health Analytical Database to augment RCT data. This database includes structured and unstructured EHRs from over 280 clinics and represents 2.4 million patients currently being treated for cancer in the United States. Mortality data were acquired from multiple data sources and benchmarked against the National Death Index. The study examined breast cancer survival for patients with HR+, HER2- metastatic breast cancer taking palbociclib in combination with letrozole, compared to those taking letrozole alone. The study found improved survival for patients taking the combination regimen.

Only three percent of patients with breast cancer participate in clinical trials, and many patients seeking to participate cannot meet the RCT inclusion criteria. In this study, RWD complemented data from RCTs and helped make the findings generalizable to a more diverse patient population. Use of RWD contributed to the development of important clinical management evidence to help providers make treatment choices that optimize the benefit of oncology products.

Emerging Solutions

Issues of data access, privacy, and analysis are top priorities across health care delivery and health care policy. In this section, we identify some emerging solutions to address challenges in using health data to advance personalized medicine and recommend concrete areas of focus for the next three years.

Regulatory Updates

In the health sector, the federal government has enacted regulations to promote access to data while balancing the privacy needs of individuals. FDA's *RWE Framework*, referenced above, and a separate *Data Modernization Action Plan* have laid out pathways for assuring the quality and integrity of data used to assess health care interventions.

The *Cures Act* put many pieces in place necessary for an integrated data ecosystem. Of note, Section 4006 of the *Cures Act* directs ONC and the Office of Civil Rights (OCR) to jointly promote patient access to health information in a form convenient for the patient, in a reasonable manner, without burdening the health care provider involved. A proposed *Cures Act 2.0* update was released in June 2021 to further this work. Among other provisions, the *Cures Act 2.0* would require FDA to increase the use of RWE in treatment development and to expand diversity in patient populations in clinical trials. Also, in 2021 the HHS OCR released a proposed *HIPAA* rule that updates federal requirements for patient consent, enables multiple digital communication options, and increases the opportunities for patients to voluntarily contribute EHR data. The proposed rule takes steps towards maximizing the value of *HIPAA* and extending *HIPAA*-like protections to other sources of data. It also includes interoperability requirements.

There has also been progress in developing a more coordinated approach to privacy protections for health data generated within and outside the health care system. The White House Office of Science and Technology Policy (OSTP) is now under Cabinet-level Executive Office leadership. OSTP is situated to coordinate the levers of data policy across agencies and couple that with budget authority to effect broad changes in research and development. ONC, charged with implementing the *Cures Act* and developing data standards, is positioned to work with states to coordinate consent and privacy policy and to promote interoperability. ONC is also positioned to foster data governance systems to drive these objectives.

International Approach to Data Privacy

On the international side, the European Union (E.U.)'s *General Data Protection Regulation (GDPR)*, which took effect in 2018, was enacted to comprehensively protect personal data. Unlike the U.S. approach, which considers health data and other data separately, *GDPR* applies protections to all data processed in the E.U., extending such protection to individuals regardless of where they are at any given time — and thus extends a far reach. It defines personal data and applies privacy and security protections with fines for breaches, thereby satisfying a

desire, at least in Europe, to adopt umbrella data principles across diverse entities. The regulation addresses protections for individuals including legal bases for data use (including consent), strict enforcement authority, and requirements for transparency of data use policies. While recognizing the important privacy protections offered by *GDPR*, some, notably in the United States, have expressed concern that it creates hurdles to accessing data essential to research breakthroughs.

Recent changes in the E.U. regarding transfers of data to the U.S. might push the U.S. to adopt changes to federal privacy laws. The U.S. approach will need to strike a careful balance of offering privacy protections while fostering appropriate use of data for research innovations.

Innovative Partnerships

Innovative cross-sector partnerships are pushing the envelope to link data analysis and research findings with products and services that foster improvements in health care quality, efficiency, and outcomes. Many breakthroughs — including those in personalized medicine — rely on innovations in data integration coupled with analytic tools such as AI and ML. FDA has awarded grants to several health care institutions in collaboration with analytic partners to validate RWD against other evidence and identify methodologies for combining real-world datasets. In other examples, private sector organizations such as Carequality have organized stakeholders to agree on interoperability standards that enable information exchange across data sharing networks.

Large technology firms are seeking partnerships with health care systems to apply learning algorithms to health data — including medical, genomic, and financial data. These partnerships have much to offer in bringing cutting-edge data integration techniques, analytics, and privacy protections through innovations such as blockchain to health care research. They also illustrate the importance of updating regulatory approaches to ensure that privacy protections extend to data from any source.

Improvements in Transparency, Equity and Representation

Analysis of large datasets can contribute to understanding disparities in health care and health outcomes or to perpetuating them. Researchers and clinicians are aware that

Innovative cross-sector partnerships are pushing the envelope to link data analysis and research findings with products and services that foster improvements in health care quality, efficiency, and outcomes.

Data integration and multi-center research will be enhanced through advances in the U.S. technical infrastructure.

missing data, missing representation, and bias in care patterns impact recognition and treatment of disease in racial, ethnic, and gender minority populations. These threats to the integrity of data that could result in inequitable treatment recommendations and treatments are now being explicitly addressed in the data management and research processes. Health information exchanges and RWD are actively seeking to capture elements of social determinants of health to better identify and reduce inequities in health and health outcomes.

As an example of research that embodies representation and transparency, the NIH-led *All of Us* precision medicine initiative plans to enroll one million people in a 10-year program to gather and mine health information — including genomic data — to identify health risks, improve treatment precision, and improve clinical patterns of care. The program adheres to FAIR principles and has written principles for privacy that favor representation of participants and patient control of information. Program policies and data use are fully transparent to participants, including information on which data is being shared with researchers. International and private sector initiatives are emerging to promote platforms and strategies to reduce bias in data access, management, and analysis.

Collaborative Governance

Alongside regulated requirements for standardized data, a variety of important private sector governance approaches contribute to data integration. For example, Health Level 7 (HL7), the entity leading standards development for implementing interoperable FHIR API, is developing multiple technical approaches to ensure privacy and security of patient information. The standards are built to track consumer consent for use of data. HL7 has also created tools for auditing internal compliance and building security into coding to ensure that queries are executed according to information management policy.

Other organizations, such as the American National Standards Institute (ANSI) and the eHealth Initiative (eHI), have developed collaborative, influential models for standard setting and could have a role in education, assessment, and oversight as data protocols are standardized in a common framework. There are also models from other industries for widespread voluntary adoption of standards; security standards for credit card processing are one such example. On the consumer side, there may be a role for collaborative governance approaches that enable consumers to manage consent through a third-party intermediary, based on consumers'

expressed consent criteria. Health data trusts or data custodian organizations have been proposed as fiduciary organizations that help consumers seamlessly manage consent for use of their data.

Better Analytics to Improve Clinical Decisions

Insights from ML and AI are generated through analysis of massive amounts of digital data contributed from clinical data, disease registries, genomic and other molecular data resources, and RWD. Although the use of ML and AI to mine EHR data is still relatively limited, the tools are rapidly gaining utility in clinical laboratories and discovery sciences. AI and ML are maturing in their capacity to detect patterns and connections relevant to health by mining and creating associations between structured and unstructured data.

Data integration and multi-center research will be enhanced through advances in the U.S. technical infrastructure. Data networking and capacity to connect with consumers will be expanded through the growing availability of 5G broadband wireless networking technology. 5G may broaden the geographic reach of wireless communications to underserved communities, boosting the power of applications used by personal devices and enhancing telehealth engagement among health professionals. Innovations in information security such as blockchain and "ledger" systems that enable users to secure and authenticate their identity will also be important to facilitating data sharing among researchers and providers.

Greater Support for Clinician Decision-Making

Learning health systems (LHSs) are organizations that have adopted an integrated approach to translating data more effectively into practice. The LHS aligns science, informatics, and incentives to apply knowledge more quickly to care improvements. The National Academy of Medicine defines the LHS as a system in which "science, informatics, incentives, and culture are aligned for continuous improvement and innovation, with best practices seamlessly embedded in the delivery process, [with] patients and families active participants in all elements, and new knowledge captured as an integral by-product of the delivery experience."

The LHS harnesses the power of EHRs, applies digital tools that engage patients and clinicians to ask meaningful questions, and ensures that the knowledge that is produced is quickly integrated into care to improve health. In an ideal LHS, results and outcomes of clinical decisions inform best practices, implementation science, and new research directions. Many U.S. health systems, including the Regenstrief Institute, Mayo Clinic, and the University of California Health System, are moving towards the LHS model of creating a loop to continuously use data to understand outcomes and support clinicians in making decisions underlying effective, patient-centered personalized medicine.

The Path Forward

Recommendations for Near-Term Policy Priorities

This white paper describes the promise of a vibrant system of protected data exchange among researchers, health care providers, and patients to advance the effectiveness and efficiency of health care. Going forward, effective data policies and digital technology solutions will establish an infrastructure that translates basic and translational science information into the applied evidence, treatments, and innovations that underlie personalized medicine. Given the pace of innovation, it is essential to invest in this infrastructure now. We must implement the technology to deliver more clinical data and leverage powerful machine learning systems to translate analytic insights into reliable evidence-based treatment solutions for individual patients while protecting individual privacy.

Priorities for Enhancing the Use and Value of EHR Data

Three-Year Objective: *Streamline the development and implementation of coordinated standards for collection and interoperability of EHR data, ensuring that the data are portable, transferable, and fit for appropriate secondary research use to guide personalized medicine.*

PMC supports the concepts and patient-centered priorities that are the cornerstone of the proposed *Cures Act 2.0*. The original *21st Century Cures Act* provided a framework to address information blocking and interoperability among EHR systems. More work is needed on this legal framework through the *Cures Act 2.0* to achieve the objectives outlined in this paper. Other key features needed to address the future of data include expanding CMS authorities for coverage of innovative new health information technologies and increasing the authority of FDA to expand collection and use of RWD.

PMC supports policies that expand ONC's U.S. Core Data for Interoperability (USCDI) initiative to align data elements and other efforts to accelerate the standardization of digital exchanges across our nation's IT systems. It is essential that public and private initiatives to facilitate standardized information exchanges be supported with the goal of improving health care, population health, and research. It is particularly important to develop standardized genomic data elements to ensure that genetic information is an integral component of research inquiries.

PMC encourages NIH to deploy resources for research and training to create an initiative in support of AI and ML in clinical trials. Furthermore, as PMC recognizes the important issue of bias in datasets and the harm that misrepresented data embedded in algorithms can have, we recommend research initiatives be undertaken to develop clear guidance about avoidance of bias, and advocate for transparency of reference populations in the deployment of AI and ML tools. Such work complements PMC's overarching initiative to address health inequities among historically underserved populations, *Addressing Disparities and Improving Health Equity in Research Advancing Personalized Medicine*.

PMC encourages federal agencies, health care delivery systems, community public health

agencies, and telecommunications companies to work together to enable EHR interoperability and promote 5G broadband communications to bring personalized medicine practices to the places patients choose.

Priorities for Protecting Patient Privacy While Encouraging Innovation

Three-Year Objective: *Ensure that new legislative efforts at the federal or state level recognize the careful balance between respecting and protecting the interests of patients, empowering patients with access to their own data, honoring patient choice, fortifying privacy and security protections, and supporting the public interest when addressing use of datasets for research and development to improve quality of care and patient outcomes.*

Building on our national framework of information protections, PMC advocates for the development of ethical and practical integrated data systems that will improve access to personalized medicine through legislative and regulatory actions as well as through private sector collaboration and standard setting. Recognizing that a fragmented approach to state regulations of consumer information will create gridlock for innovation, PMC calls for Congress to develop policies that ensure privacy of consumer health information across all data collection sources. Further, we encourage close monitoring of the progress of the newly introduced *Endless Frontier Act* and encourage the consideration of civil and criminal penalties for indiscriminate uses of consumer data including intentional attempts to re-identify individuals from de-identified datasets.

PMC calls upon the HHS ONC and OCR to work with FTC, Congress, and the White House to create a new regulatory and enforcement framework for protecting non-*HIPAA* covered health data sources. These solutions should promote transparency of consumer data use and enable technology solutions to guide patient authorization/consent for uses of their data. To facilitate the coordination of policy development and actions, PMC recommends the development of a federal clearinghouse on state, federal, and international consumer health privacy protections. Further, we encourage HHS and other health-related federal agencies to expand policy research in the ethical, legal, and social aspects of the use of non-clinical data for health care research and health care practice improvements.

PMC sees great value in public and private research endeavors aimed at the development of automated “smart” systems enabling consumers to authorize or use their own data for research, care coordination and health improvement. PMC supports a research agenda

PMC advocates for the development of ethical and practical integrated data systems that will improve access to personalized medicine through legislative and regulatory actions as well as through private sector collaboration and standard setting.

to address the need for improved patient consenting processes and technical solutions for encryption and trusted data management intermediaries. This includes research in information system security programs that will enable the development of mature systems for universal credentialing and digital user-authentication as well as digital systems to assess privacy risks in digital environments on an ongoing basis.

Priorities for Strengthening Adoption of RWD in Research and Practice

Three-Year Objective: *Accelerate development and adoption of practices and standards for datasets created from real-world data to enable maximal use by researchers, clinicians, patients, regulators, and other stakeholders that yields insights for improved decisions, equity, and outcomes in health care.*

PMC recognizes that a trusted and validated system of RWD is needed to support its integration as a mainstay of research and discovery. PMC sees a need to engage patients in understanding and contributing to real-world datasets. RWD can be leveraged best to establish a secured and standards-based infrastructure for integration of clinical and non-clinical data resources such as social determinants of health data and remote patient monitoring. Research into the applications of RWD in clinical trials will be needed to validate algorithms to decrease the risk of known biases. Further, RWD applications will benefit from data sharing governance principles and best practices to ensure maximum use and benefit are achieved for those whom the data represents to avoid exploitation and misuse.

PMC also supports the expansion of translational research programs and data standards development that will expedite the integration of RWD with clinical and biological data that ultimately support individualized approaches to therapy, including valid means of addressing the uses of deidentified datasets.

PMC recommends that clinical tools be the target of advanced research and development in technology solutions by applying human-centered design to promote more effective use of data. PMC recognizes the burdens of electronic documentation and information management on health care professionals. The integration of RWD in the health care setting provides opportunities for addressing these challenges. Developing learning health systems should be a priority need for the future to enhance the use of genomic and other -omics data in personalized clinical applications.

PMC encourages payers of health care services to consider the value of research and data management to enhance treatment options and patient decision-making. Payment incentives can be used to enhance use of decision-support tools that will accelerate adoption of data-driven personalized medicine and improve the quality of care.

PMC encourages FDA and other regulatory agencies to help advance the field of data science in RWD by establishing testing environments and programs to encourage innovation and safety in automated learning systems. The recent development of an FDA Digital Health Center of Excellence should prioritize the regulatory science needed to advance the underpinnings for automated technologies to aid in the decision-making processes to guide clinicians and enable personalized medicine. The Center will expand the paradigms for integrating remote patient monitoring, AI and ML, and learning health care systems.

Conclusion

Personalized medicine is the future of patient-centered care. Advances in data management policies and data science will create an environment that fosters inquiry, insight, and the development of more personalized therapies for individuals.

Integrated data management and decision-support have the potential to transform the medical system by making data available at individual and systems levels, improving both individual and population health.

Critical issues will be:

1. Maximizing the value of data from EHRs to enhance research and improve the efficiency of health care, including burden reduction for clinicians;
2. Protecting the privacy of patient information used for clinical care and research; and
3. Designing innovative ways to turn RWD into the evidence and treatment recommendations that will improve personalized care.

Through existing regulation and private sector partnerships, our nation has made important strides towards interoperability and better analytic use of data. AI and ML offer the possibility of continued progress in achieving the broad health care benefits that can be realized from personalized medicine. Optimal use of these analytic tools requires new guideposts and safeguards to achieve the objectives of delivering personalized, equitable treatments. Moving forward, it is essential to build the technical solutions and policy constraints to ensure integrity of data and the protection of individual privacy. Technical advances must be coupled with clinical tools and a learning environment to support adoption of best practices in the real world.

Our nation's strategic choices over the next several years will have a great impact on population health and individual treatment outcomes. Closing the gaps between clinical practice and research through better use of data will advance medical progress, accelerating the delivery of personalized care to improve health outcomes, expediting evidence for new or emerging treatments, and reducing the use of ineffective treatments. In an age in which there is a consumer expectation of personalization, the health care system needs to deliver data-driven solutions leading to better outcomes, faster.

FURTHER READING

- Abul-Husn NS, Kenny EE. Personalized Medicine and the Power of Electronic Health Records. *Cell*. 2019;177(1):58-69. doi:10.1016/j.cell.2019.02.039
- Agarwala V, Khozin S, Singal G, et al. Real-World Evidence in Support of Precision Medicine: Clinico-Genomic Cancer Data As A Case Study. *Health Aff* (Millwood). 2018;37(5):765-772. doi:10.1377/hlthaff.2017.1579
- Chen XS and Patel MS. *Harvard Business Review*. Digital Health Tools Offer New Opportunities for Personalized Care. https://hbr.org/2020/11/digital-health-tools-offer-new-opportunities-for-personalized-care
- Embi PJ, Richesson R, Tenenbaum J, et al. Reimagining the research-practice relationship: policy recommendations for informatics-enabled evidence-generation across the US health system. *JAMIA Open*. 2019;2(1):2-9. Published 2019 Jan 16. doi:10.1093/jamiaopen/ooy056
- Everson J, Patel V, Adler-Milstein J. Information blocking remains prevalent at the start of 21st Century Cures Act: results from a survey of health information exchange organizations. *J Am Med Inform Assoc*. 2021; Mar 18;28(4):727-732. doi: 10.1093/jamia/ocaa323
- Gettinger A, Zayas-Caban T. HITECH to 21st century cures: clinician burden and evolving health IT policy. *J Am Med Inform Assoc*. 2021;Apr 23;28(5):1022-1025. doi: 10.1093/jamia/ocaa330
- Kramer MR and Pfitzer MW. The Ecosystem of Shared Value. *Harvard Business Review*. https://hbr.org/2016/10/theecosystem-of-shared-value
- Majumder MA, Guerrini CJ, Bollinger JM, Cook-Deegan R, McGuire AL. Sharing data under the 21st Century Cures Act. *Genet Med*. 2017;19(12):1289-1294. doi:10.1038/gim.2017.59
- Mandl KD, Kohane IS. Data Citizenship under the 21st Century Cures Act. *N Engl J Med*. 2020;382(19):1781-1783. doi:10.1056/NEJMp1917640
- McGinnis JM, Fineberg HV, Dzau VJ. Advancing the Learning Health System. *N Engl J Med*. 2021;385(1):1-5. doi:10.1056/NEJMp2103872
- McGraw D, Mandl KD. Privacy protections to encourage use of health-relevant digital data in a learning health system. *NPJ Digit Med*. 2021;4(1):2. Published 2021 Jan 4. doi:10.1038/s41746-020-00362-8
- Miksad RA, Abernethy AP. Harnessing the Power of Real-World Evidence (RWE): A Checklist to Ensure Regulatory-Grade Data Quality. *Clin Pharmacol Ther*. 2018;103(2):202-205. doi:10.1002/cpt.946
- Morris AH, Stagg B, Lanspa M, et al. Enabling a learning healthcare system with automated computer protocols that produce replicable and personalized clinician actions. *J Am Med Inform Assoc*. 2021;28(6):1330-1344. doi:10.1093/jamia/ocaa294
- Office of the National Coordinator for Health Information Technology. 2020-2025 *Federal Health IT Strategic Plan*. Available at https://www.healthit.gov/topic/2020-2025-federal-health-it-strategic-plan
- Smith M, Saunders R, Stuckhardt L, McGinnis JM, Committee on the Learning Health Care System in America; Institute of Medicine, eds. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington (DC): National Academies Press (US); May 10, 2013.

APPENDIX 1

US Legislation Governing Access to and Protection of Health Data

The Health Insurance Portability and Accountability Act (HIPAA)

HIPAA, enacted in 1996, established national standards to protect sensitive patient health information (protected health information, or PHI) and defined requirements for patient consent to share information. It set the stage for information interoperability by requiring the government to establish code sets, unique identifiers, and operating rules for electronic transactions. HIPAA addresses security standards and enforcement actions for breaches in covered information systems to protect against nonconsensual use of data.

Genetic Information Non-discrimination Act (GINA)

The 2008 GINA prohibits discrimination based on genetic information with respect to health insurance and employment. Subsequent related regulations prohibit discrimination in access to health care based on predisposition to inherited diseases.

The Affordable Care Act (ACA)

The 2010 ACA broadened insurance coverage and promoted the use of electronic transactions for payment and information sharing.

Health Information Technology for Economic and Clinical Health (HITECH) Act

The 2009 HITECH Act offered financial incentives for provider organizations to adopt electronic health information technology and enacted privacy and security requirements for the transmission of health information, with penalties for non-compliance. The law included federal investments in infrastructure, payment incentives, and regulatory penalties to hospitals and providers that did not comply.

The Medicare Access and CHIP Reauthorization Act (MACRA)

MACRA in 2015 built on earlier legislation requiring adoption of electronic health information systems and created a plan to achieve interoperability. It also adopted a payment system that required quality reporting and action towards interoperability for Medicare providers.

21st Century Cures Act (Cures Act)

The Cures Act of 2016 accelerated personalized medicine by addressing bottlenecks in the access to and use of data. Major revisions in policy were implemented to overcome “information blocking” and improve interoperability.

APPENDIX 2

US Federal Agencies and Offices Regulating Health Data

Overview of Agency Regulatory Oversight Responsibilities

Department of Health and Human Services (HHS) Office of the National Coordinator (ONC) is charged with implementing the Cures Act, enhancing interoperability, and developing privacy and security standards. ONC oversees the development of technical standards for information transmission called the USCDI and puts out advisories on ONC New Data Element and Class (ONDEC) standards for information integration. ONC's 2020-2025 Federal Health IT Strategic Plan specifically addresses the need to ensure that consumers both understand how their data may be used and how they benefit from it.

HHS Office of Civil Rights enforces HIPAA rules, which apply to individual health information captured within the health care system.

Food and Drug Administration (FDA) reviews scientific evidence on the effectiveness of drugs and devices and approves these therapies for public use. It establishes criteria on how scientific evidence will be considered in the approval process, including RWD. FDA regulates information management of medical devices, many of which relay data on device performance and/or clinical indicators. Devices may include mobile health apps, though not all apps are linked to medical devices and EHRs.

Federal Trade Commission (FTC) regulates notification of health data breaches and some aspects of privacy and data security relating to commercial products (such as mobile apps and social media).

Commerce Department's National Institute of Standards and Technology (NIST) is charged with developing overall frameworks for interoperability and risk management of information systems, including security protocols.

The Substance Abuse and Mental Health Services Administration (SAMHSA) develops rules pertaining to the privacy and sharing of patient mental health and substance use information.

HHS Centers for Medicare and Medicaid Services (CMS) administers many provisions of MACRA and has driven adoption of and payment for health information technology while protecting patient privacy. CMS exerts leverage through both rulemaking and the marketplace, as the payer for Medicare services.

The Federal Communications Commission (FCC), although not typically considered a health care regulator, has pushed to expand access to 5G communications technology. This impacts health and health data by expanding access to telehealth and bandwidth for remote monitoring and wearable monitoring devices.

White House Office of Science and Technology Policy (OSTP) leads the executive branch to shape the direction of national science and technology policy, including policy on the role of information management. OSTP also recommends strategies for ensuring data security and integrity in research. OSTP recently created a legislatively mandated National Artificial Intelligence Initiative Office. In 2021 the OSTP director was designated a Cabinet-level official, signaling interest in coordinating a national data strategy government-wide. This suggests stronger cross-government opportunities for embracing priority actions to coordinate the national data enterprise.

ABOUT US

The Personalized Medicine Coalition (PMC), representing innovators, scientists, patients, providers and payers, promotes the understanding and adoption of personalized medicine concepts, services and products to benefit patients and health systems.

PMC
PERSONALIZED
MEDICINE COALITION

1710 Rhode Island Ave., NW
Suite 700
Washington, DC 20036

@permedcoalition