WHITE PAPER

The United States' Precision Medicine Initiative Cohort Program – Investigating the Ethical and Legal Aspects Surrounding Consent

January 9, 2017

Datacenter Solutions Group – Health and Life Sciences

"This document addresses many important issues not only for PMI but for all broad research initiatives."

– Christine Suver,
Head of Open Science Data Governance,
Sage Bionetworks

Emily Kirby, Anne-Marie Tasse and Bartha Knoppers

Public Population Project in Genomics and Society (P3G), Montreal, Canada

Chiara Garattini and Rick Cnossen

Intel Corporation, Health & Life Sciences

Correspondence to: chiara.garattini@intel.com

Executive Summary

The Precision Medicine Initiative ("PMI") is an ambitious project that aims to build a longitudinal cohort representative of the American population by collecting samples and data from one million participants. As currently planned, these individuals would be recruited by selected partner health-provider organizations ("HPOs") or through other types of recruitment sites where volunteers would be recruited directly. Given that the effort was launched recently, there are ongoing announcements by funding agencies presenting key collaborating institutes and partners. Since the PMI aims to recruit participants as early as 2017, addressing issues related to consent becomes crucial, including the medium used to document consent and the content and language employed in consent forms. In this context, this report aims to provide an overview of the issues related to consent and relevant to building the PMI cohort. Several recommendations are provided to guide the cohort development and ensure alignment with existing international initiatives.

More specifically, the report begins by identifying legal and ethical aspects of consent related to the PMI, including its geographic scope, covered data, approach to existing data sets, and anticipated obstacles. Options for capturing consent (referred to as the consent "medium") contemplated by the

Table of Contents

PMI Working Group Report are examined, including the use of electronic consent, dynamic consent and computable consent methods. The various options related to the core elements of the proposed PMI cohort that will be presented to participants (the consent "substrate"), are identified and categorized.

Next, existing global efforts related to consent that could be leveraged within the PMI context are examined through a review of consent elements found in eight U.S.-based biobanks and cohorts (Million Veteran Program; Kaiser Permanente Research Program on Genes, Environment and Health; Partners Healthcare Biobank, Geisinger MyCode; Marshfield Clinic Personalized Medicine Research Program; Mayo Clinic Biobank; Children's Hospital of Philadelphia) and six international initiatives (UK Biobank; 100 000 Genomes Project; Lifelines; Ontario Health Study; H3Africa; MalariaGEN). A comparative analysis of these projects reveals several points to consider, including the scope of consent, the sources of data collection and possible linkage, the possibility of recontacting participants, questions surrounding the return of results, sharing of data and samples, as well as matters of commercialization.

Furthermore, relevant international and national standards related to consent, particularly in the biobanking context, are identified, in order to provide a brief overview of their potential use within PMI. Leverageable standards/guidelines identified include process standards (related to *how* consent should be obtained), technical standards/guidelines (*what* consent should contain or express in terms of structure), and semantic standards/guidelines (the *meaning* of the consent language used).

Finally, the report concludes by summarizing a number of current limitations related to the environment PMI is evolving in, and proposing several consent-specific recommendations. Potential obstacles and uncertainties identified include the lack of a central governing entity for the PMI, an unclear privacy framework, evolving guidance on the use of next generation sequencing technologies, the matter of the return of results and the availability of counseling resources, and finally, the evolving regulatory framework. As a result, the following recommendations are set forth, for further consideration and discussion:

- Adopt a consent "medium" and mechanisms compatible with a longitudinal cohort approach;
- Ensure interoperability of the consent "substrate" with international initiatives;
- Account for health literacy when developing consent form language;
- Address the obstacles related to the issue of the return of results;
- Examine the issue of data sharing within the U.S.A. and internationally;
- Name the governing entity(ies) responsible for the Precision Medicine Initiative Cohort Program ("PMI-CP");
- Clarify the applicable privacy framework.

WHITE PAPER

Introduction: The Precision Medicine Initiative (PMI)

The Precision Medicine Initiative: A brief background

Announced by President Obama in his State of the Union Address on January 20, 2015, the Precision Medicine Initiative ("**PMI**") aims to build a cohort of one million, or more participants between 2016 and 2020, to create a resource for researchers working to understand the many factors that influence health and disease. The PMI Cohort Program ("**PMI-CP**"), also known as the All of Us research program, will collect and share participants' samples, data from a number of different sources such as electronic health records, environmental risks and devices. It is part of a series of recent efforts in the United States to build a roadmap for precision medicine[1] by collecting quality data and samples from a large number of individuals over time. Eventually, this data could serve to predict individual health.

This ambitious project also ties into significant efforts that several American agencies are undertaking to modernize elements of the legal framework applicable to research with human subjects, such as Common Rule and the Food and Drug Administration's ("**FDA**") medical device framework,[2] particularly as they relate to genomic technologies.

Multiple stakeholders and agencies are involved, some of which have already received funding and begun building PMI-CP, including the White House Office of Science and Technology Policy, the National Institutes of Health ("**NIH**"), the FDA, the Office of the National Coordinator for Health Information Technology ("**ONC**"), the National Institute of Standards and Technology and the Office of Civil Rights.

Proposed approach for building the PMI-CP

Prospective recruitment. The PMI Cohort Program Working Group Report[3] (the "**Working Group Report**"), issued on September 17, 2015, recommended that the PMI-CP cohort be built on the basis of two distinct recruitment strategies, namely (1) recruitment of participants through health provider organizations ("**HPOs**") and (2) by enrolling individuals who wish to volunteer directly (self-referred participants).

Inclusion of retrospective cohorts. Although the option of including retrospective cohorts (i.e. previously collected samples and data from existing cohorts) within the PMI-CP infrastructure, was briefly considered, this option appears to have been set-aside for now by the PMI Working Group, as adopting this approach would "limit the ability of the PMI-CP to serve as a national resource accessible by many researchers, and would hinder standardization of participant communication and engagement, sample collection, and data collection."[3] Therefore, based on the Working Group's recommendations, the prospective recruitment option appears to have been retained as a preferred method to enroll participants. Nonetheless, where possible and consented to, participants from existing cohorts and biobanks could be recontacted to recruit them into the PMI-CP.

Key announcements. At the time of preparing this report, many key questions related to setting up the biobank, database and related infrastructures remain unanswered. This, in turn, implies that recommendations formulated may evolve as new announcements are made in the upcoming months. Indeed, many important considerations which will influence recruitment, consent and the application norms and regulations will depend on key choices made by the agencies involved.

After the Working Group Report was issued several funding decisions were made by the participating agencies regarding the launch of PMI efforts (summarized in Table 1). In particular, on July 6, 2016, the NIH announced an award of $55 million to start building the cohort.[4] As part of this news release, key partners were also identified. They will play an important role in the participant recruitment infrastructure, which will be launched later in the year.

Finally, in addition to this operational infrastructure, several other actors have been partnering with governmental agencies to support background research and policy work.

TABLE 1: SUMMARY OF KEY INSTITUTIONS INVOLVED IN THE PMI-CP.

ROLE	FUNCTION	NIH-FUNDED ORGANIZATION (MORE FUNDING ANNOUNCEMENTS MAY BE MADE IN THE UPCOMING MONTHS)
Biobank	Will build the PMI-CP biobank and support the collection, analyses, storage and distribution of biospecimens.	Mayo Clinic
Data and Research Support Center	Will acquire, organize and provide secure access to the PMI-CP datasets, and provide research support and analysis tools.	Vanderbilt University Medical Center (Nashville, Tennessee), working with the Broad Institute (Cambridge, Massachusetts) and Verily Life Sciences (Mountain View, California)
Participant Technologies Center	Will enroll participants through direct enrollment and develop, test, maintain and upgrade the PMI-CP mobile applications to enroll, consent and communicate with participants.	Scripps Research Institute (San Diego, California), Vibrent Health (Fairfax, Virginia) and partners (including Sage Bionetworks, Walgreens, PatientsLikeMe)
Healthcare Provider Organizations (HPOs)	Will engage their patients and enroll participants in the PMI-CP through regional medical centers and community-based HPOs (Federally Qualified Health Centers)	Initial announcement of funded HPOs (more announcements to come in the coming months): *Regional medical centers:* • Columbia University Medical Center (New York, New York) • Northwestern University (Chicago, Illinois) • University of Arizona (Tucson, Arizona) • University of Pittsburgh (Pittsburgh, Pennsylvania) *Federally-Qualified Health Centers:* • Cherokee Health Systems (Knoxville, Tennessee) • Community Health Center, Inc. (Middletown, Connecticut) • Eau Claire Cooperative Health Center (Columbia, South Carolina) • HRHCare (Peekskill, New York) • Jackson-Hinds Comprehensive Health Center (Jackson, Mississippi) • San Ysidro Health Center (San Ysidro, California) *Veterans Affairs:* • Medical centers across the U.S. *California Precision Medicine Consortium:* • University of California San Diego, with partners Cedars-Sinai Medical Center, Los Angeles; San Diego Blood Bank; University of California, Davis; University of California Health; University of California, Irvine; University of California, San Francisco; and University of Southern California, Los Angeles *Geisinger Health System, Danville, Pennsylvania* *New England Precision Medicine Consortium:* • Partners HealthCare System and its hospitals, Massachusetts General Hospital and Brigham and Women's Hospital, with Boston University and Boston Medical Center *Trans-American Consortium for the Health Care Systems Research Network:* • Henry Ford Health System, Detroit, with partners Baylor Scott and White Research Institute, Dallas; Essentia Health, Duluth, Minnesota; Spectrum Health, Grand Rapids, Michigan; and University of Massachusetts Medical School, Worcester

Nature of participation in PMI-CP

In addition to the collection of biological specimens, the types of data to be collected from participants could include, but would not be limited to, "clinical and insurance claims data, survey and demographic data, genomic and other biospecimen-derived data, and mobile, implantable, or other equipment or device data, all of which may be electronically stored or on paper."[5]

Another important component of the PMI initiative is increased participant engagement in all stages of the project. The Working Group, for instance, recommends that the cohort be developed using a "highly interactive and proactive participation model."[3] Amongst a number of proposed approaches, this could include "a standardized consent protocol to ensure consistency in the terms and conditions that all PMI cohort participants agree to […]"[3] The Working Group also recommends that, to adequately engage participants, the consent protocol should: give participants the option to join supplementary or complementary studies outside the PMI cohort; where possible, return individual and aggregate results to each participant; allow participants to set preferences to dictate how much information they receive, and be able to change their preferences throughout their participation in PMI-CP.[3]

Finally, the Working Group recommendations set the stage for the development of more granular details regarding both consent content, the "substrate" of consent, as well as the form the consent should, that is, the "medium" for obtaining, recording and updating consent.

Context and scope of this report: PMI and the issue of consent

Building on preliminary scoping work undertaken to identify the ongoing initiatives to assemble the PMI-CP, this report aims to examine the time-sensitive matter of consent. Indeed, as it is hoped that recruitment of participants will begin as early as 2017, both the "substrate" and the "medium" of consent will need to be clearly articulated.

More specifically, based on background issues related to consent, we first identify the legal and ethical aspects of consent as it relates to the PMI initiative, the geographic scope, covered data, approach to existing data sets, as well as any anticipated obstacles (Section "PMI and consent: Overview of the Current Proposals"). Next, we canvas existing global efforts related to consent that could be leveraged within the PMI context (Section "Overview of current consent approaches in U.S.-based and international longitudinal cohorts and biobanks"). For instance, we review the consent policy and standards for fourteen biobanks and cohorts, eight of which are U.S.-based, in order to identify existing consent options in cohort studies (Appendix I). We also identify relevant international and national standards related to consent, and provide a brief overview of their potential use within PMI (Section "Contextualizing PMI in the broader national and international consent ecosystem"). Finally, we conclude by identifying potential obstacles of the PMI consent approach, as currently proposed, and suggest several key consent-specific recommendations (Section "Conclusions: Points to consider, obstacles and key recommendations"). This report was prepared with the objective of informing a future stakeholder meeting, in order to clarify certain issues raised and further debate the proposed recommendations.

To do so, we have adopted number of different methodological approaches, including for instance, ongoing monitoring of PMI-related press releases, literature review related to precision medicine (in the U.S. and internationally), general web-based searches related to the PMI initiative, as well as collection and review of consent form from select cohorts.

However, the development of the PMI-CP is rapidly evolving, and announcements regarding funding and project partners are made regularly. Therefore, we recommend ongoing monitoring of such information to ensure that information presented in this report remains relevant and accurate in the upcoming months.

We finally note that while the White House has released reports on *Privacy and Trust Principles* (November 9, 2015),[6] as well as a *Data Security Policy Principles and Framework* (May 25, 2016),[5] there is still no formal position on consent. Therefore, the timing of this report is opportune, as the time is ripe for reflection on this central pillar of the ambitious PMI-CP.

PMI and consent: Overview of the current proposals

The concept of "consent" encompasses two distinct yet related notions, and both raise ethical and legal issues.

First, given the rapidly evolving technological infrastructure surrounding biomedical research, documentation of consent is no longer seen as solely a pen-and-paper process. Indeed, there are a number of emerging mechanisms to carry out the consent process and maintain ongoing communication with participants, in an interactive and engaging way. The first part of this section addresses several approaches to obtaining, recording and updating

consent (the "medium") that have been envisaged.

Second, the content of the consent form itself also raises several important questions, especially in the context of precision medicine research that employs a longitudinal cohort approach. This aspect is examined in the second part of this section. Both of these facets of consent are addressed in light of the publicly available background work done by several PMI working groups and committees.

Approaches to obtaining consent (the "medium")

As previously noted, a central aim of the PMI initiative is to actively engage participants in all stages of the project, consent being one of the key steps. To implement this objective, much thought has revolved around what consent mechanisms would be appropriate.

Electronic consent

With the increasing availability of technologies such as smartphones and wearables, capable of capturing significant amounts of data on individuals, such devices have considerable potential in the field of health research. Indeed, increased options for data collection and analysis now expand the locus of health research beyond clinical and institutional settings.

Thus, in some cases, the use of electronic media for research purposes may also require a departure from the traditional in-person, paper-based informed consent process. Electronic consent ("**e-consent**") can be defined as any method of obtaining and recording participant consent by any electronic means, such as a computer, smart-phone or other device. While e-consent has not yet been widely adopted by biobanks, its potential to improve efficiency and even increase enrolment has been raised.[7] Furthermore, the use of an electronic format may facilitate presenting information in a variety of different formats, such as graphics, audio, video, weblinks. Use of an electronic multimodality platform may be, "more universally accessible and adaptable for the widest range of potential research participants with diverse skills and limitations"[8] (ex: disabled participants, the elderly). However, after the initial consent, use of multimodality platforms, such as apps or web portal, have, in certain cases, been linked to increased attrition rates.[9,10]

The PMI Working Group has recommended that "an electronic consent protocol should be used across all PMI cohort participants to ensure consistency, minimize organizational burdens, and maximize participant recruitment"[3] (Recommendation 4.5), and that e-consent be a core component of recruitment. However, depending on its design, e-consent options may be considered both for the initial recruitment to the PMI-CP, including both HPO-mediated recruitment and through direct volunteers, or where specific app-based data collection or research are thereafter initiated, for example for any additional modules to PMI participation. Finally, several authors note that self-administered electronic-based consent should not entirely replace the staff facilitated, in-person options, as for certain individuals or groups, a paper-based approach may be more accessible.[9,11] Therefore, staff-facilitated consent options should generally be made readily available alongside self-administered approaches, so as to reduce any potential bias.

Recent NIH funding announcements have listed several institutes that will act as "participant technology centers", developing mobile applications for enrollment of participants. This implies that some form of e-consent could be envisaged. Sage Bionetworks is named among the participant technology partners, and has been developing an e-consent tool for mobile-mediated research.[12] This tool proposes a framework that is customizable to the requirements of a given study, with eight recommended core elements of informed consent based on 45 CFR 46, subpart A (the "Common Rule").[12] Through the use of this tool, studies can design a self-paced and self-administered electronic informed consent process for app-mediated research studies.[12]

In addition to the existing Sage Bionetworks e-consent tool, other initiatives are currently examining the use of electronic media to complete the consent process. Mobile device-based health research initiatives, sometimes called "mobile Health" or "mHealth" have been gaining popularity in the research setting. However, there are no clear standards or guidelines on how to implement such device-based consent for research. Groups such as the Global Alliance for Genomics Health's ("**GA4GH**") Mobile Health Consent Task Team[13] was convened to create a statement of best practices highlighting key issues and considerations for both academic and industry researchers, designers, and coordinators as well as interested members of the general public.

Active efforts to start implementing e-consent mechanisms on a larger scale are currently underway in many different settings. For instance, in collaboration with its Institutional Review Board ("**IRB**"), Vanderbilt University has been piloting the use of e-consent forms in researcher project undertaken at the university.[14] As part of this initiative, e-consent forms can be created and implemented as part

of the REDCap (Research Electronic Data capture) platform.[15,16] These forms enable participants to access and sign consent forms (via a "wet signature" – PIN number, typed name, or finger/stylus signature, as a REDCap feature). Through this system, consent can be documented via a computer, mobile phone, or tablet. Similarly, Duke University's reports several new projects using an e-consent interface (PLATFORM study, use of Apple Research Kit)[16], including video modules and the opportunity to interact with study coordinators in case of questions. In addition, as part of the PMI, Vanderbilt has been funded to lead the Direct Volunteer pilot studies, which includes the development of an interface for obtaining consent.[17] To do so, it can presumably leverage and adapt its current e-consent framework to a direct volunteer type of recruitment.

Furthermore, the use of e-consent in a biobanking context has been tested by the Partners HealthCare Biobank,[9] and it could provide an important case study to inform a wider implementation as part of the PMI-CP. In complement to its traditional mailed and in-person enrollment strategies, the biobank launched its e-consent platform in June 2014, aiming to recruit 75,000 participants by 2018. As opposed to its previous staff-facilitated recruitment strategy, Partners HealthCare's e-consent approach consists of approaching potential participants via email to invite them to participate in the biobank. Participants are then consented via a website featuring multimedia content about the biobank. By piloting this approach, Partners HealthCare hopes to leverage this mechanism for use in the PMI-CP.[9] Preliminary results from this system suggest that establishing an e-consent system presents certain challenges such as IT (for example, participant online authentication procedures) and administrative requirements (for example, IRB approval of the approach). Use of e-consent for biobanking is hoped to facilitate certain aspects of recruitment and data collection, which could include an extended outreach, increase in the number of participants who can potentially be consented to the biobank, potential increase in participant understanding through the use of multimedia tools, as well as an increased completion rate of initial questionnaires. However, in practice, some limitations were identified in piloting this e-consent approach. For instance, the pool of participants whose identity can be electronically validated was found to be limited; the method is largely dependent on electronic tools and their availability within the Partners HealthCare system (ex: large-scale email alerts); there is a potential decrease in the diversity of participants enrolled depending on accessibility to computer technology; and a decrease in the follow-up to provide blood samples since e-consent is often asynchronous to sample collection events (ex: phlebotomy).[9] It is also important to note that overall enrolment rates using exclusively an e-consent approach were reported to be lower, as compared to approaching potential participants in-person in the hospital waiting room (for instance, 3.5% of emailed patients (n=7078 patients) consented to participate in the biobank, as compared to 51% (n= 28,930 patients) who were approached in-person, between June 2014 and January 2016).[9] Further work may be required to examine whether email outreach can be used to reach a greater number of individuals than in-person approaches, and augment recruitment numbers.[9]

Finally, use of e-consent is also reported in several citizen science and direct-to-consumer ("**DTC**") initiatives. For example, the uBiome microbiome sequencing project uses an e-consent form to interact with participants, and document consent remotely.[18] Similarly, upon registration, 23andMe participants are guided through an electronic format of the consent form.[19]

While there is a range of possible technological innovations in the realm of e-consent, the regulatory context in which it is inscribed remains in flux, and gaps remain in the literature regarding its use, effectiveness, and the potential challenges.[8,20]

Computable consent

Computable consent, or sometimes also called machine readable consent, is a proposed mechanism whereby the format in which consent is captured is understandable by a computer, thereby bridging the gap between a static document (ex: the paper-based consent form) and the data it contains (ex: restrictions on data use, participant choices, etc.).[21] It is hoped that given the amount of data generated by PMI, and possible granular consent options, there will be an opportunity to help drive the implementation of machine readable e-consent.

The Office of the National Coordinator ("**ONC**") for Health Information Technology is a key participant in the PMI initiative, and will be responsible for advancing data standards, addressing relevant privacy policies, and advancing innovation in health IT. A PMI Task Force was set up by the ONC to "identify data standards and implementation specifications that make health IT data available to participants and researchers for precision medicine."[22] The task force held a series of meetings between July 2015 and May 2016, many of which addressed the notion of consent. More specifically, within the scope of its "Interoperability Roadmap", the task

force narrowed the scope of its review of consent-related issues to examine "computable consent," as part of the broader computable privacy and interoperability agenda. Although the task force has not yet made a report publicly available, it has indicated that it intends to make connections with the GA4GH to address computable consent in the research context, and has recommended that the ONC convene a stakeholder group to address granular, dynamic computable consent. It also raises specific issues that need further investigation, including: education of participants and data providers with respect to data-related rights and permitted uses, as well as agreement on an interoperability roadmap across data sources (from consent, to EHR data from labs, hospitals and pharmacies, to data from non-provider sources such as claims databases and mobile apps.)[23]

Finally, specifically regarding consent choices, the ONC Health IT PMI Task Force suggests that three levels of rules must be computable, namely: (1) background rules such as Health Insurance Portability and Accountability Act ("**HIPAA**")-related permitted uses of health information; (2) basic choice participants make about the use and disclosure of health information; and (3) granular choices that a participant can make regarding legally sensitive conditions, which can evolve over time to enable choices about disclosure.[23] The Task Force stresses that standards need to be developed to enable such computable approaches,[23] highlighting in particular, the need to develop consent standards in this field.[24]

Dynamic consent

The notion of dynamic consent has received much attention in the academic literature, and particularly in the recently evolving context of precision medicine. It is often presented as an alternative to broad consent, as a means to more fully engage participants. Generally described as a "personalised, communication interface to enable greater participant engagement in clinical and research activities,"[25] it places the participant at the center of the decision-making process as it allows for consent interactions. Dynamic consent has also been proposed as a mechanism which could be used for the sharing of health information in electronic health records ("**EHR**"), thereby allowing participants to more readily provide or withdraw their consent over time, as well as providing information to participants about how their personal data are used[26].

Dynamic consent has been raised in the PMI Working Group report as a way to "enable participants to actively engage in an informed, voluntary and ongoing manner."[3] The concept is also invoked in several of the Working Group's recommendations and comments, including those on recontact for future studies (Recommendation 4.6), on the return of results and information, and on participant access to data and sharing with healthcare providers (Recommendation 5.20).[3]

There is no consensus as to how dynamic consent choices should be presented to the participant in the particular context of cohort studies. Use of "opt-in"/ "opt-out" statements has been suggested.[25,27] However, considering the risk of "consent fatigue" and the need to maintain standardized core data over time, the level of granularity that should be made available to participants is open to debate. While some pilot projects using dynamic consent models have been reported, there are still very few examples of large-scale, longitudinal biobanking efforts using this approach.[27] Moreover, "classical" longitudinal cohort studies using a broad consent are not static but use recontact mechanisms for both internal updates and invitations to join new studies. Indeed, broad consent cohorts employ ongoing communication efforts to avoid attrition and maintain interest while providing updates on findings. There are also challenges associated with using a dynamic consent model, including the technical capacity to interface between patient choice, which may change over time, and actual downstream data/sample use.[25] Moreover, given the ongoing interactions with the project proposed in the dynamic consent model, therapeutic misconception may become a greater risk than with a one-time consent approach followed by updates as the understanding of the unique epidemiological nature of cohort studies may be replaced by a clinical one.[28,29]

Internationally, the UK-based the EnCoRe Project piloted the use of dynamic consent in a biobanking context.[30] The dynamic consent web interface was designed to "enable biobank participants to revisit consent choices and have a more active involvement with the research of the biobank."[30] Results from this pilot study suggest that biobank participants welcome the opportunity to have a greater interaction and receive more information about the use of their samples through an online interface. However, wide-scale testing of a dynamic consent interface, including an interface to inform participants of downstream research uses, as well as, more generally, testing its efficacy against other models, remains to be undertaken.

There is therefore a need for further piloting of dynamic consent before wide-scale implementation. Indeed, this matter goes beyond the design of consent interfaces, as a proper implementation of dynamic consent options requires an ability to trace

consent choices back to data use, for example, through metadata that can be updated in real-time.

As part of its work on computable consent, the ONC PMI Task Force has looked at dynamic computable consent, and concluded that while there are existing standards, there is no clear implementation guidance and no alignment between HIPAA and the Common Rule on this matter. In addition, implementation of dynamic consent would require that an individual's health information be connected to his or her sharing choices as updated and changed by the individual.

Scope of consent (the "substrate")

Although there is currently no publicly-available information regarding a PMI template consent form, we identified several PMI Working Group recommendations that mention elements that will need to be included in such a form. The following subsection examines key consent elements mentioned in order to guide the discussion in the next section ("Overview of current consent approaches in U.S.-based and international longitudinal cohorts and biobanks"), related to a comparative analysis of national and international populational biobanking efforts.

Currently the Working Group is considering including the following in the PMI, and consent would be required for each of these elements. Participants would be required to:

Provide a certain amount of **data**. Current proposals on types of data to be collected could include the following (some of which are likely optional):

- Individual demographic and contact information
- Consent and personal preferences for participation in the project
- Self-reported measures
- Behavioral and lifestyle measures
- Sensor-based observation through phones, wearables and home-based devices
- Structured clinical data derived from EHRs
- Unstructured and specialized types of clinical data derived from EHRs
- PMI baseline health exam
- Healthcare claims data
- Research specific observations
- Biospecimen-derived laboratory data
- Geospatial and environmental data
- Other data, such as social networking information, Twitter feeds, social contacts from cell phone text and voice, and OTC medication purchases
- Other suggestions raised by the PMI Working Group:
  - Exposure data
  - Data on family relationships (explicit linkage between records)
  - Gender identity

Provide a **biospecimen** (parallel collection of biospecimen for the HPO's use is possible)

Agree to **share EHR data**, if available, although having EHR data may not be a requirement for direct volunteers

See a healthcare provider for an initial baseline health measurements

Give permission to collect **identifying information**, which is to be kept secure but that could be shared for recontact for approved studies

Give permission to use identifiers to link across disparate data sources using identifiers

Give permission to allow future **recontact**

Give permission to share specimens and data for future research uses

- Data sharing options should be provided to participants (control and access to their data).

Be allowed to **withdraw** participation at any time:

- Clarification that data already contributed to research studies or included in existing aggregate data sets cannot be withdrawn if participants withdraw from the PMI-CP.

Select options for the return of research results:

- Aggregate results: aggregate results for all studies should be made available to participants.
- Individual results: basic health information (ex: serum cholesterol, blood pressure) should be provided at enrollment.
  - Unclear strategy in the Working Group recommendations, in part due to issues related to Clinical Laboratory Improvement Amendments certification ("**CLIA**")/FDA medical device regulations.
  - Data remains in the PMI cohort unless results are transmitted back to participants' physicians, where consented to and where possible.
  - Data is not transferred into the clinical record, but participants could choose to share the data with their HPO, outside of the PMI.

Overview of current consent approaches in U.S.-based and international longitudinal cohorts and biobanks

Based on the preliminary considerations detailed in the PMI Working Group Report, approaches to consent adopted in eight U.S.-based biobanks (Million Veteran Program; Kaiser Permanente Research Program on Genes, Environment and Health; Partners Healthcare Biobank, Geisinger MyCode; Marshfield Clinic Personalized Medicine Research Program; Mayo Clinic Biobank; Children's Hospital of Philadelphia) and six international longitudinal cohorts or biobanks (UK Biobank; 100,000 Genomes Project; Lifelines; Ontario Health Study; H3Africa; MalariaGEN), were examined. Tables presenting results are provided in Appendix 1, and a comparative analysis of these findings is summarized below.

Key points to consider for the development of PMI-CP consent documents

The comparative analysis of U.S. and international longitudinal biobank and database cohorts reveals several key common elements, as well as divergences between U.S. approaches as compared with international initiatives. We summarize below several points to consider with respect to the development of consent clauses and approaches.

Scope of consent: In setting the scope of the biobank initiatives, it is important to mention the longitudinal and distinct nature of such cohort studies. Such projects generally include a broad range of environments and sources of data, and allow for its use in future research. Variations of broad consent for future use are generally employed. In addition, although the consent forms we examined did not always do so, we recommend explicitly mentioning that subsequent studies may involve genetic or genomic analyses, raising unique familial and privacy issues.[31]

Therefore, this information should be conveyed consistently, especially in a precision medicine context.

Data collection: The predominant forms of data collection in a longitudinal cohort and biobanking context are surveys/questionnaires and data collection from the medical record. For the PMI, a key consideration is ongoing access to EHR data: this information will be updated over time. As is done in many of the cohorts examined, a clear indication of the ongoing nature of access to medical records should be stated. In addition, if any linkage with other available datasets is anticipated, this should also be clearly indicated as was the case in the Ontario Health Study, LifeLines and MVP. However, should the type of records to be linked be expected to change in the future, a broader statement regarding access to other types of records should be used, following the example of the Ontario Health Study. Finally, where legally permitted, if a participants' record is to be linked to those of other participants, such as family members, this information should be mentioned in the consent form as well (see the Children's Hospital of Philadelphia Biobank).

Recontact: Broad recontact clauses are generally preferable to allow participants to choose whether to participate in additional components of the research that have new objectives, or additional collection of samples or data. Most participants in the cohorts examined were recontactable in some way, however, the broad purposes of the recontact should be made clear, as to not exclude any reason for recontact that may arise. Furthermore, recontact is another element of consent where participants could express preferences, for example, on the type of recontact they wish to receive. We also note that for the purpose of leveraging existing large U.S. cohorts to the PMI, most biobanks examined specifically allow recontact to propose other studies, with some exceptions, such as the Mayo Clinic and the Cincinnati Children's Hospital Center – Better Outcomes for Children.

Return of results: The topic of the return of research results is where there is the most divergence across the cohorts examined. This issue has also been extensively addressed in the literature.[32] The variation in approaches to the return of results we observed could be in part due to limits imposed by the jurisdictional context of each cohort. Options reviewed range from:

- No return (MVP, Marshfield Clinic, H3Africa) in largely epidemiological studies;
- Return of baseline health information at initial assessment only (UK Biobank, LifeLines);
- Limited return when results of medical significance are discovered (Kaiser Permanente, Partners' Healthcare Biobank, Mayo Clinic, Cincinnati Children's Hospital); or
- Providing choices to return testing information for potential clinical guidance (100,000 Genomes Project, Geisinger MyCode, Children's Hospital of Philadelphia).

The question of the return of results is a difficult one to address, especially in the context of genomics and precision medicine. It is also further complicated in the U.S. context, where there is regulatory debate as to whether research results from genetic testing can be returned for use in a clinical context, without certain regulatory certifications and approvals. This is evidenced by the

fact that for almost all U.S. cohorts surveyed results are only returned when they are clinically significant and validated by a CLIA-accredited laboratory. However, the Geisinger MyCode Project and Children's Hospital of Philadelphia, which offer a broader return of information, including for use for potential therapeutic optimization, were different in this respect (see Appendix I for a comparison of the types of results returned). Internationally, the UK's 100,000 Genomes Project is perhaps the most permissive, in that testing results regarding the disease examined are systematically returned to physicians, and further options are offered to participants with respect to other types of testing results. Arguably, however, the considerations related to the return of results are somewhat different when recruiting participants with a pre-existing medical condition (as is done in the 100,000 Genomes Project), where results may be relevant to the condition for which the participant is being treated (ex: rare diseases, cancer, etc.).

Data and sample sharing: Data and sample sharing are largely regulated by the scope of consent provided. However, we note that there is an increasing trend to specifically mention international data sharing in consent forms, in order to be able to enable researchers to share data globally.[33] A clear mention of a possible cross-border transfer of data is especially important given the current important changes in several regional privacy regulations (for example, the European General Data Protection Regulation).

Commercialization: Possible future commercialization of research findings is another important consideration, as transparency on this is often perceived as a matter of public trust, especially in the context of donated biological samples and data.[34] In the consent forms reviewed commercialization was addressed either with respect to access to the biobank's resources by commercial entities or by possible commercialization of research findings more generally.

Recontact: Finally, as mentioned in the PMI Working Group Report, the research participants in the cohorts examined are almost all recontactable for future studies and so could help to leverage the recruitment of participants to the PMI-CP. Given the broad range of consent options and approaches adopted by these different cohorts, retrospective use of previously collected samples and data from these cohorts to PMI-CP may be possible via their recontact mechanisms. However, the need for a core standardization set of variables argues in favor of a new consent. Therefore, based on this review, we concur with the PMI Working Group in suggesting that a new consent, specific to the PMI-CP be developed.

Contextualizing PMI in the broader national and international consent ecosystem

The PMI will be operating within the U.S. Common Rule regulatory framework. The Common Rule outlines baseline regulatory requirements for informed consent in human subjects research. Additionally, there are state regulatory requirements that will need to be addressed.

However, beyond the U.S. national regulatory framework, several international and national guidelines and directives on implementing consent should also be considered to foster alignment of PMI with other international initiatives. Indeed, several precision medicine and populational cohorts exist around the world, and alignment of PMI with other international initiatives will strengthen its outcomes and utility. We examine several of these guidelines, standards and projects, with a view to possibly leveraging efforts towards developing a common approach for consent in the PMI. While this does not constitute an exhaustive review of all available standards, guidelines, and best practices, its aims to summarize key documents with a particular emphasis on those applicable to biobanking and genomic research.

Process standards and best practices

Process standards or best practices relate to *how* consent should be approached and documented within participating organizations, such as HPOs.

International

Several general human subject research and biobanking standards include provisions regarding the process of obtaining consent.

The International Conference on Harmonization ("**ICH**") E6 *Harmonised Tripartite Guidelines: Good Clinical Practice* (1996),[35] which pertains to clinical trials and not to biobanking specifically, provides general guidance on the consent process by detailing the process for obtaining written informed consent (section 6.2). Guidance includes information on compliance with regional and local norms, revision of consent materials when appropriate, the use of non-technical language, the information process, as well as a list of core elements to include in the information provided to the participant. Furthermore, building on the consent provisions found in the ICH E6 guidelines, the 2015 draft ICH E18: *Guidelines on Genomic Sampling and Management of Genomic Data* (Draft 2015, Step 3, comments pending)[36] aims to provide guidance on genomic sample

collection to evaluate efficacy and safety of a drug for regulatory approval. With respect to consent, while referring to ICH E6 for general consent matters, E18 (draft) specifies that within a genomic context, "[i]deally, informed consent for the collection and use of genomic samples should permit broad analysis of the samples (e.g., sets of genes, transcriptome, whole genome sequencing) regardless of the timing of analysis. Additional elements might include the possibility to use the samples for assay development, disease research, or pharmacovigilance."[36]

Similarly, the Organization for Economic Co-operation and Development ("**OECD**") *OECD Guidelines on Human Biobanks and Research Databases*[37] provides some policy guidance on the development of biobanks and databases, including key recommendations and best practices on obtaining informed consent for such purposes. In particular, it emphasizes basic notions in the biobanking context, including that consent provisions should indicate the types of sample and data to be collected, the type of data anticipated to be derived from the analysis of the samples, and the health and other records to be accessed, their intended uses, as well as storage conditions and duration. The guidelines also emphasize the importance of recontact in the context of longitudinal biobanks, as well as, where permitted under local law, obtaining a consent that will permit human biological specimens and/or data to be used to address unforeseen research questions.

Furthermore, the International Society for Biological and Environmental Repositories' ("**ISBER**"), *ISBER Best Practices for Repositories: Collection, Storage, Retrieval, and Distribution of Biological Materials for Research* (2012),[38] addresses informed consent best practices in the context of biobanking (Section L.2.200). It requires that the consent process allow participants to make an informed choice about whether to provide samples to a biorepository, and where applicable, to agree to future research uses. In such cases, general information about the possible future research uses should be provided, in accordance with applicable national or local regulations and policies.

Other international guidelines have been developed specifically on consent. A notable example is the GA4GH's *Consent Policy* (2015),[39] which builds on the principles found in its *Framework for Responsible Sharing of Genomic and Health-related Data*.[40] The *Consent Policy* is based on the respect for the ability of data donors to make their own decision, as well as respect for the practices of medicine and research. It provides principled and practical guidance on consent issues related to the sharing of genomic and health-related data, including transparency, accountability, data quality and security, privacy, data protection and confidentiality, risk-benefit analysis, accessibility and dissemination of both prospective and legacy data. As such, it applies to a wide range of stakeholders involved in genomic and health-related data sharing including researchers, research participants and patient communities, publishers, research funding agencies, data protection authorities, hospitals, research ethics committees, industry, ministries of health, and public health organizations.

Finally, while the use of wearable devices and wellness applications has started to be integrated into health research, and is of interest to biobanking initiatives, there have been few international guidelines regarding their use, and in particular, on consent matters. Recently, the Future of Privacy Forum has produced *Best Practices for Consumer Wearables & Wellness Apps & Devices,*[41] which seek to address regulatory gaps in the collection, use and processing of data collected through wearable and apps. These preliminary guidelines propose that use of covered data for research purposes, including secondary use, should follow existing regulatory informed consent processes and requirements. Recognizing that data sensitivity lies on a spectrum and privacy protections ought to be calibrated according to this spectrum, the guidelines propose relevant factors to consider when assessing data sensitivity, which, in turn, can affect consent requirements. Beyond consent, the guidelines lists factor to be taken into account when considering whether a secondary research use of collected data is compatible with the primary use consent. These factors include, for instance, the existence of a link between the purpose of collection and the intended secondary use, the context in which the data have been collected, the nature and sensitivity of the data, the possible consequences of secondary use on the user, and the existence of appropriate privacy safegards.[41] While these constitute preliminary suggestions, further investigation and guidance on the use of data collected through the use of wearables and app-based platforms will be required.[41]

National (U.S.)

In addition to international guidance, and regulatory Common Rule requirements, U.S.-specific guidelines provide some insight into mechanisms detailing the process of obtaining consent for biobanking purposes.

Recently, the National Cancer Institute ("**NCI**") has adopted *Best Practices for Biospecimen Resources* (2016)[42], which include specific recommendations on the creation of bioresources. Section C.2 offers a detailed overview of the current

U.S. regulatory framework applicable to obtaining informed consent for data and sample collection. It sets forth several general recommendations on the process of obtaining consent, including the following, of relevance to the PMI context:

- NCI recommends seeking informed consent of participants who provide biospecimens and data, and preferences should be considered, for instance regarding engagement in future research and recontact. E-consent or mobile consent strategies may streamline the consent process and improve participant understanding;
- Personal, religious, and culturally held-beliefs and traditions should be respected in the use of data and specimens, and the biospecimen resource should track any relevant restrictions or instructions based on such beliefs;
- Information about the biospecimen resource governance, placement of data in the medical records, documentation pertaining to consent, as well as confidentiality and privacy protection measures should be provided to the participant;
- The consent document should disclose whether biospecimens may at some point be de-identified or subsequently used for secondary research, beyond that described in the original consent form;
- New concepts that strengthen the informed consent process, including community advisory boards, community consultation and community based participatory research projects are being piloted.

These general recommendations are complemented by more specific proposals related to key elements that should be presented in the informed consent form.

With respect to the use of e-consent, while there are currently few standards regarding its use in research, in March 2015, the FDA issued draft guidance on the *Use of Electronic Informed Consent in Clinical Investigations*[43]. In an effort to harmonize standards across different governmental agencies, the Office for Human Research Protections ("**OHRP**") is considering a final FDA-OHRP document.[8] The draft Q&A presented in the document provide clarification and recommendations on procedures that may be followed when using an e-consent protocol. These aim to help ensure the protection of the rights, safety and welfare of human subjects; the comprehension of information conveyed via e-consent; the appropriate documentation of the consent process when electronic media are used; and, the quality and integrity of data captured via e-consent.[43] It addresses a range of questions such as the presentation of information to the participant, the location of the e-consent process, interacting with the participant, assessing understanding, privacy issues, obtaining HIPAA authorizations through e-consent, as well as the use of electronic signatures to document the process.[43]

Technical standards and best practices

Technical standards are used to define *what* the consent should contain or express, that is, structural components and the reliability of transactions in terms of interoperability. These standards have been slowly emerging in relation to computable consent initiatives with a view to ensuring that the structure and syntax of consent are made interoperable.

International

Several international technical standards have addressed *what* consent should be able to capture. For instance, Health Level 7 ("**HL7**") –*Version 3 Standard: Medical Records; Data Access Consent, R1* was designed to allow health IT systems to communicate a patient's consent to the collection, access, use or disclosure of their personal health information, or to subsequently withdraw such consent. While primarily addressing the clinical considerations, such as consent to sharing and use of medical records, these guidelines may also be relevant where ongoing linkage with medical records for research purposes is proposed. The document is intended for use primarily by healthcare IT vendors, EHR systems, departmental systems and healthcare providers. Therefore, in a research context, its relevance lies in the interoperability of consent directives in the healthcare and research settings.

The HL7 Consent Tracker proposal has been presented to the HL7 Community-Based Collaborative Care ("**CBCC**") Work Group.[44] While no directive or standard has been produced by the CBCC Work Group, its aims are to "[provide] a flexible means for tracking and conveying intra- and inter-enterprise the lifecycle, provenance, privacy, and security metadata used for Consent Directive management workflows, such as obtaining patient authorizations and restrictions, and Consent Directive enforcement using security labeling and privacy protective services to comply with access control decisions."[44] If developed, this standard would attempt to build a process to convey the minimal set of metadata to manage consent directives in various workflows, including both clinical and research settings.

The International Standards Organization ("**ISO**") has developed the *ISO/TS 17975: 2015(E) Technical Specification on "Health informatics – Principles and data requirements for consent in the Collection, Use or*

Disclosure of personal health information" (2015),[45] which offers a technical framework for capturing "informational consent" for the processing of health information. Its scope is primarily processing of health information in a clinical context, but extends to capturing consent for research uses of such information. It provides an "informational consent" framework that can be specified and used in specific policy areas to inform decisions related to the communication of personal health information. It also addresses its processing and use for research purposes, and the design of paper and electronic "informational consent" forms. Minimum data requirements for consent documentation are proposed, such as what data are to be collected and used or disclosed, for what purposes, to whom, the date the consent comes into effects, the length of the agreement, if/when the decision has been renewed, if/when it has been revoked, what information about the data was presented to the subject, what data are permitted for use, what the purpose of use is, and more detailed requirements.[45] As such, the proposed ISO framework could be useful in examining how to bridge the various mechanisms used to capture consent to use and disclosure of health record information in a clinical context, with the mechanisms used to consent participants in the PMI research context.

Finally, the GA4GH's Automatable Discovery and Access Task Team has been working to create an "Automatable Discovery and Access Matrix,"[46] which provides a standardized way to represent any and all consent and other conditions of use that apply to a resource, making such information unambiguous, computer-readable and hence directly available for digital communication, searching and automation activities. It proposes a structured set of metadata (i.e., data about either data or an object annotated by data) which can facilitate the effective communication of use conditions, or provide the basis for the discovery of these conditions, or increasingly automate decisions around granting access based on these conditions. More specifically, this tool principally specifies and organizes 42 general concepts that may or may not be employed to define acceptable use and conditions of use in any one setting. Some of these concepts are hierarchical to others in the set, and all of which are grouped into three sections, namely "Permissions" (mainly relevant to consent), "Terms" (typically relating to legal/contractual matters), and Meta-Conditions (over-arching topics).[47] Ultimately, use of this tool could help in capturing consent-based restrictions (amongst others) in an automatable and computer-readable manner. The proposed tool is availableto the broader research community.[48]

Semantic standards and best practices

Semantic standards are concerned with preserving the *meaning* of consent forms, by ensuring that consistent language or codes are used in their implementation, or that language and codes are translated or interpreted correctly. Several semantic standards have been developed by various organizations both internationally and in the U.S., including for instance, template consent forms, consent clauses, and academic publications. For the purpose of this report, we have selected examples based on their generic application to biobanking, as well as their potential usefulness to a precision-medicine, data sharing context.

International

Internationally, several guidelines address semantic standards for developing consent forms.

A number of tools have been developed by the GA4GH, specifically with respect to consent to genomic data sharing. For instance, *Consent Codes*[49] proposes a structure for recording data use "categories" and "requirements" based on consent information. Such codes can be used to standardize data use groups, based on previously existing consent protocols and consent forms. Analysis of various research participant and patient consent protocols, the Consent Codes identified 19 consent-based restrictions and categorize them into "primary categories" or "secondary categories" and as "requirements," based on their occurrence and on the resulting criteria for data access.[49] Use of the proposed approach could avoid introducing unnecessary new restrictions on data use,[49] based on consent language and elements used.

In addition, the GA4GH *Consent Tools*[50] provides template consent language and clauses which can be used for different types of situations including: (i) conditions for international data sharing of legacy data, (ii) clauses specific to international data sharing, and (iii) a generic international data sharing prospective consent form. All require adaptation according to local social, cultural and legal specificities.

The Public Population Project in Genomics and Society ("**P^{3}G**") has also developed a publically available template consent form for biobanking,[51] which can be tailored to local requirements. In addition, its International Policy interoperability and data Access Clearinghouse ("**IPAC**") provides an open-access database of sample clauses,[52] including consent form clauses to assist researchers in building ELSI policy documents.

National (U.S.)

As previously mentioned, the NCI has adopted *Best Practices for Biospecimen Resources* (2016),[42] which not only presents general recommendations on the process of consent, but also specific recommendations on what consent forms for bioresources should contain, that is, the "elements of consent" (Section 2.3). Such key elements include, for instance:

- The right to refuse biospecimen donation, and no impact on treatments;
- The purpose of collection of biospecimens and data;
- The source of collection of biospecimens;
- The custodian of the biospecimen;
- The use of biospecimens (including the possibility of secondary uses);
- How data and biospecimens will be shared;
- How data and biospecimens will be stored (ex: de-identification) and how long they will be kept for;
- Who may access biospecimens and data;
- Whether there is a plan for sharing general findings;
- Whether there is a plan to return individual research results (and if so, whether they will be placed in the medical record);
- The nature of risks associated with the use of genomic technologies.[42]

Interestingly, the use of a tiered system for secondary use of biospecimens, in which several types of uses are proposed, is examined. Example of tiered consent categories are suggested to guide to simplify use of such systems and avoid them being overly burdensome to participants.[42] Furthermore, the use of a "one-time general consent" strategy for broad use of specimen and data in IRB-approved future research studies is also presented as an acceptable approach.[42]

In addition, the NIH's National Human Genome Research Institute ("**NHGRI**") provides template consent forms[53] and suggested clauses specific to genomic research[54] for several types of NIH-funded research involving genomic sampling such as biobanking, WGS/WES, and pediatric research. This collection is mainly specific to the U.S. context.

Finally, with respect to e-consent semantic tools, Sage Bionetworks' Participant-Centered Consent toolkit ("**PCC**"),[55] which is openly licensed, aims to transform the concept of consent from a signature on a legal form to a process that educates. The toolkit contains a variety of resources such as a PCC visual dictionary of icons and concept animations, e-consent workflows, design documents and templates, and FAQs that aim to create an informed consent process. In partnership with the Scripps Research Institute, Sage Bionetworks has received NIH funding for consent and governance within PMI.

Other noteworthy consent-related projects

Several other noteworthy initiatives could also be of interest as case studies in the development of certain innovative facets of the PMI proposal.

The eMERGE Network,[56] housed at Vanderbilt University, is a national network organized and funded by the NHGRI that combines DNA biorepositories with EHR systems for large scale, high-throughput genetic research in support of implementing genomic medicine. Amongst other consent-related initiatives, its Consent Task Force has produced a model consent language document[57] that provides clauses for the collection and storage of human biospecimens and data for future research, particularly those collections that have an electronic medical records component. Work on ELSI issues is ongoing, as part of this network's efforts.

More recently, the NIH has funded the LawSeq project, led by Susan M. Wolf at the University of Minnesota. LawSeq will convene a national Working Group of top legal and scientific experts to analyze current U.S. federal and state law, regulation, and guidance on translational genomics and to generate consensus recommendations. The goal of LawSeq is to build a solid legal foundation for successful integration of genomics into clinical use via online public databases, stakeholder meetings and publications.[2]

The National Patient-Centered Outcomes in Research network ("**PCORnet**") works to transform the culture of clinical research from one directed by researchers to one driven by the needs of patients and those who care for them. It has developed several resources to guide in the implementation of patient-centric tools. One of these resources is its *Informed Consent Recommendations for Specimen Acquisition and Future Use*[58] that presents different forms of consent such as tiered, dynamic, and broad.

With regards to participant engagement, the Genetic Alliance's Platform for Engaging Everyone Responsibly ("**PEER**")[59] project offers, "a customizable registry system used to collect privacy-assured health information, which enables individuals to set their own sharing, privacy and data access preferences in a granular and dynamic manner."[59] It could therefore serve as a model for the implementation of participant-centric options such as dynamic consent, sharing preferences

setting and engagement. It is currently being piloting in certain PCORnet disease groups and it has partnered with PrivateAccess to develop its consent management and access control technology.

A final noteworthy initiative is the ONC/NIH partnership, Sync4Science pilot project.[60] NIH, the ONC, and the Harvard Medical School Department of Biomedical Informatics will coordinate the S4S pilot in collaboration with six EHR vendors to test the use of open, standardized applications to give individuals the ability to easily and securely contribute their EHR data to research. It will therefore form part of the ecosystem related to consent to data sharing via participant-mediated tools.

Conclusions: Points to consider, obstacles and key recommendations

While the PMI is rapidly setting the stage for the construction of its million participant longitudinal cohort, several key issues remain to be discussed. The following analysis of potential obstacles highlights several opportunities for the development of the PMI-CP. The concluding recommendations address important consent-related issues that must be resolved for the future sustainability and international interoperability of this important endeavor.

Potential obstacles identified

First, there is still a limited understanding of the likely benefits and limits of precision medicine technologies, particularly in a research setting. While there is great hope in their potential, knowledge of such technologies is rapidly evolving. This, in turn, makes it difficult to fully foresee all future considerations in today's consent form language. Therefore, consent elements must be broad enough to account for any directions research may take in the future, and unanticipated applications, benefits and risks related to participation.

Second, as currently proposed, the PMI-CP aims to recruit participants from a number of different sources. This includes recruiting self-referred participants, current patients from HPOs, and participants from several different institutions. With such a wide range of sources, use of a uniform consent form for the entire initiative might prove difficult. Furthermore, although a single IRB of record was proposed for the PMI, local implementation of the program will depend on reliance agreements and acceptance of at a minimum, common consent elements. While this is a departure from the traditional review process under the Common Rule before proposed modifications, how this approach will be implemented in local institutions remains to be seen.

Third, several cross-disciplinary regulatory frameworks are currently being revised, which in turn may impact the consent and its content. These include, for instance, consumer protection legislation, privacy legislation, regulatory regimes related to drug and devices approvals, diagnostic testing and research governance.[20] Such legal frameworks take on even more importance in the context of governing biobanking and database activities, where consent for future unspecified research uses of tissues and data is asked of participants. The regulatory regime adopted could thereby impact the scope of what can be consented to. Furthermore, what can be done in the future may be different as regulations can change.

More particularly, we would like to emphasize the following limitations in the current regulatory framework in which the PMI is evolving:

Unclear definition of the central governing entity for the PMI

As currently proposed, the PMI appears to be a collaborative effort across different institutions and partners. Custodianship of the biobank (Mayo) currently appears distinct from database management. From an accountability perspective, it will be important to define governing entities, in order to inform potential participants of such details, and more generally, to establish the legal and regulatory framework of the study.

Applicable privacy regulations[61]

There is an ongoing debate in the legal literature regarding privacy rules applicable to data and samples collected as part of the PMI-CP. Indeed, data collected to form the PMI-CP may originate from different custodians of personal health data. However, differing privacy regulatory frameworks apply, or not, to these different custodians (ex: state-specific privacy legislation, federal privacy legislation, Common Rule, HIPAA, etc.) which could impact the type of privacy protections and measures in place with respect to PMI participation.[62,63] This uncertainty could, in turn, impact how privacy-related matters, for example, information on privacy protections, risks related to participation, access to PMI data by law enforcement authorities and participants' right to access PMI data, are presented in consent forms. Clarification of the applicable privacy frameworks is needed to develop the appropriate, common, consent language.

Evolving guidance on the use next generation sequencing (NGS) technologies

On July 8, 2016, in an attempt to clarify its position on the matter of NGS tests

and the regulation of "medical devices", the FDA issues draft guidance on (i) the *Use of Standards in FDA Regulatory Oversight of Next Generation Sequencing (NGS)-Based In Vitro Diagnostics (IVDs) Used for Diagnosing Germline Diseases*[64] and (ii) *Use of Public Human Genetic Variant Databases to Support Clinical Validity for Next Generation Sequencing (NGS)-Based In Vitro Diagnostics*.[65] While both these guidance documents are part of an attempt to help the assessment of safety and efficacy of NGS tests, they also have an indirect impact on issues such as the possible return of results, and accompanying consent elements, emanating from use of NGS technologies. In addition, CLIA regulations regarding laboratory certification also limits, in certain cases, the types of results that can be returned in a research context. Indeed, research laboratories that test human specimens but do not report patient-specific results for the diagnosis, prevention or treatment of any disease, or the assessment of the health of individual patients, do not require certification under CLIA. The interplay between FDA guidance and CLIA regulations with regards to NGS-based results therefore needs to be carefully considered ahead of making any promises with respect to returning research findings to participants, as such options, or the way in which they are implemented, may eventually be shaped by the current regulatory regime. Attesting to the importance of addressing this matter, the FDA held a public workshop in March 2016 on "Patient and Medical Professional Perspectives on the Return of Genetic Test Results,"[66] which elicited several comments and responses from a number of actors in the research and patient community.[67]

Return of results and availability of counseling resources

The context in which data and samples are collected is another important question related to the development of return of results consent language and mechanisms. In particular, as the PMI-CP intends to be as inclusive as possible and accept self-referred volunteers, careful consideration must be made regarding the return of results to participants who may not be able to seek medical follow-up or genetic counseling.

Recommendations

Adopt consent mechanisms and "medium" compatible with a longitudinal cohort approach

By their very nature, longitudinal population cohorts are complex endeavors and require robust data collection processes. Indeed, consenting to be a member of a cohort means participating in a population project with a standardized set of core measurements and questionnaires that must be consistently applied in order to achieve statistical quality and significance, which is achievable under a broad consent model.[28] In order to allow for implementation of broad consent and maintain participant engagement over time, appropriate consent media and associated communication tools should be adopted. This would facilitate respecting consent choices, maintaining an ongoing right to withdraw, allowing participants to be asked for permission for further recontact, and the return of results.

Ensure interoperability of consent "substrate" with international initiatives

To maximize the utility of the PMI-CP, both nationally and internationally, efforts should be made to ensure PMI is interoperable with other international initiatives. This was, for instance, a concern reported in the National Children's Study ("NCS") final report, indicating that: "the NCS has not been designed to coordinate with other global longitudinal environmental cohort studies. Thus it is unlikely that worldwide methodological experience and scientific progress will be fully leveraged".[69]

With respect to consent, interoperability requires several items. From a semantic standpoint, the consent forms must be explicitly designed to allow international data (and sample) sharing (see, for example, the GA4GH consent tools,[50] and the GA4GH Data Sharing Lexicon[70]). Second, language used in the consent form must be aligned with international standards, particularly with regards to privacy terminology.[71] Third, the use of core consent elements, or similar international consent standards can be facilitated by use of terminology to categorize downstream conditions of use.[49,70]

Account for health literacy when developing consent form language

A recent National Academies of Science workshop report has raised several issues regarding health literacy in a precision medicine context.[72] Empowerment of research participants is a central theme of PMI, and requires conveying adequate information at all stages of the project. Consent is one of the many facets of PMI where matters of health literacy will present a challenge. Indeed, in asking over a million participants to take part in a long-term project, undergo medical assessments and complete questionnaires, may present a challenge for retainment over the duration of the project. Accessible language and information should be presented and should account for the disparity of health literacy in the populations expected to enroll in PMI: if trust and understanding are not established at the onset, there is a risk of attrition and uneven distribution of subpopulation enrollment.[72]

Address the obstacles related to the issue of the return of results

Additionally, clarification of the regulatory framework applicable to clinical laboratory validation, and medical devices regulation of genetic testing, and how it applies in a research context, is required to properly assess options regarding the return of results. Jurisdictional and regulatory obstacles should be addressed, and clarifications provided to adequately inform participants regarding the possibility of returning individual results or incidental findings.

Examine the issue of data sharing within the U.S. and internationally

Since the PMI-CP is expected to recruit participants across a number of different institutions, issues regarding data sharing, both within the U.S., for example, sharing of EHR data from one site to another, as well as internationally, should be addressed. Will all sites agree to baseline data sharing, given that this means transferring substantial information between different U.S. and international jurisdictions with different privacy frameworks? Any anticipated limits or hurdles on data sharing should be clarified in order to adequately inform participants of the scope of possible sharing options.

Name the governing entity(ies) responsible for the PMI-CP

As currently planned the PMI-CP will be under the responsibility of several institutions and entities. Matters of custodianship of both samples and data are unclear, and still under development as PMI partnerships are formalized. This, in turn, influences information provided to participants, as well as applicable privacy, security, legal rights, and overall ethical governance. A clear governance structure should be developed prior to development of consent materials.

Clarify the applicable privacy framework

Finally, in order to develop appropriate consent language and inform participants of how their data and samples will be kept confidential, clarification of the privacy framework applicable to the PMI-CP as an entity is required. In particular, given the involvement of several custodians of personal health information across PMI, clarification of applicable privacy norms is important to adequately inform participants or risks involved and protections in place.

In conclusion, several key questions must be addressed before a complete informed consent process and accompanying forms can be devised. A clear consent policy, to accompany the existing PMI *Privacy and Trust Principles*[6] and *Data Security Policy Principles and Framework*[4] could be a first step in establishing the underpinnings of the PMI-CP development. We welcome an open discussion on the obstacles and recommendations raised in this report, as part of a broader stakeholder consultation, including both international and national experts, so as to align PMI efforts with the global biobank consent ecosystem.

Acknowledgements

We would like to thank Megan Doerr and Christine Suver at Sage Bionetworks for their detailed comments and suggestions. We would also like to thank colleagues David McLauchlan at the Centre of Genomics and Policy (McGill University) for his research assistance, and Alice Borrelli, Bhavik Kanabar and Tram Huyen at Intel for their support throughout the process.

For more information visit:

www.intel.com/health

http://www.p3g.org/

List of abbreviations

CLIA: Clinical Laboratory Improvement Amendments

CBCC: HL7 Community-Based Collaborative Care Work Group

DTC: direct-to-consumer

E-consent: electronic consent

EHR: electronic health record

ELSI: ethical, legal and social issues

FDA: United States Food and Drug Administration

GA4GH: Global Alliance for Genomics and Health

HIPAA: Health Insurance Portability and Accountability Act of 1998

HL7: Health Level 7

HPO: Health Provider Organization

ICH: International Conference on Harmonization

IRB: Institutional Review Board

ISBER: International Society for Biological and Environmental Repositories

ISO: International Standards Organization

MVP: Million Veterans Program

NIH: United States National Institutes of Health

NCI: United States National Cancer Institute

NCS: United States National Children's Study

NGS: next generation sequencing

NHGRI: United States National Genome Research Institute

NIH: United States National Institutes of Health

NPRM: Notice of Proposed Rule Making

OECD: Organization for Economic Co-operation and Development

OHRP: United States Office of Human Research Protections

ONC: United States Office of the National Coordinator for Health Information Technology

P3G: Public Population Project in Genomics and Society

PMI: Precision Medicine Initiative

PMI-CP: Precision Medicine Initiative Cohort Program (also referred to as the "All of Us" research program)

Appendix 1: Summary of consent elements from selected biobanks (U.S. and International)

TABLE 2: U.S.-BASED BIOBANKS AND COHORTS

COHORT	HIGHLIGHTS OF FINDINGS
Million Veteran Program (MVP)	The Million Veteran Program (MVP) is a national, voluntary research program funded by the Department of Veterans Affairs (VA). The goal of MVP is to partner with veterans receiving their care in the VA Healthcare System to study how genes affect health. To do this, MVP will collect blood samples and health information from one million Veteran volunteers. • Scope of consent: broad consent to future use of samples for genetic and health studies, in diseases that affect Veterans. • Duration of participation: only one sample collection, sample is kept indefinitely. Surveys could be sent not more than once a year. Ongoing collection of information from medical record for life. • Sample and data collection: blood collected once (DNA/molecular analysis), data collected from baseline survey, postal surveys, ongoing access to medical records, linkage to research records in VA research records database. • Recontact: recontact is possible for additional surveys, as well as participation in additional research. • Return of individual results: no return of individual results. • Data and sample sharing: sharing is restricted to researchers at the VA, other federal health agencies, and academic institutions within the United States. • Commercialization: no access by commercial entities mentioned, but use of samples could lead to commercial products or treatments. *Source: VA Research Consent Form (version date 09/12/2012). Available online, at:* *http://www.research.va.gov/mvp/InformedConsentForm.pdf*
Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH)	The Kaiser Permanente Research Program on Genes, Environment, and Health examines the genetic and environmental factors that influence a number of different diseases. The resource is linking together comprehensive electronic medical records, data on relevant behavioral and environmental factors, and biobank data (genetic information from saliva and blood) from 500,000 consenting health plan members. • Scope of consent: broad consent for analyses, including genetic • Duration of participation: information may be used indefinitely • Data collection: ongoing collection from medical record; surveys • Recontact: possible for other RPGEH studies • Return of individual results: no return of health or medical results. Except if result is of "substantial medical importance", RPGEH will recontact and ask participant if he/she wants to know the result. • Data and sample sharing: "Outside researchers" may be given de-identified information (including genetic information)` • Commercialization: Kaiser Permanente may participate in project with commercial companies, genetic tests may be developed *Source: FAQ and Privacy and Confidentiality information.* Available online at: https://www.dor.kaiser.org/external/DORExternal/rpgeh/faq.aspx?ekmensel=194f64c3_47_69_btnlink; and https://www.dor.kaiser.org/external/DORExternal/rpgeh/privacy.aspx?ekmensel=194f64c3_47_51_btnlink
Partners Healthcare Biobank	The Partners Healthcare Biobank aims to foster collaborations among investigators, physicians and patients seen at Partners HealthCare hospitals. This project works to develop the tools needed to expedite research and the translation of discoveries into clinical care. Samples are maintained in an institution-wide repository, which is a growing resource available to Partners investigators and research groups. They are available for distribution to Partners investigators with required approval from the Partners IRB. They are linked to clinical data stored in the Partners Research Patient Data Registry (RPDR), as well as some additional health information and survey data. To date, more than 45,000 patients have consented to participate in the Partners Biobank. • Scope of consent: plan to do many types of biological and genetic research • Duration of participation: samples and information stored indefinitely • Data collection: medical record (ongoing collection), surveys • Recontact: recontact possible to obtain additional information or to propose other studies • Return of individual results: no return of individual research results ("not the same as clinical tests"). However "if experts from the Biobank decide that results from [the participant's] samples are of high medical importance [the biobank] will attempt to contact [the participant]. In some situations, follow up testing might be needed in a

	certified clinical lab". • Privacy risks: measures in place to protect privacy, and limited risk with data sharing, however "we cannot predict how genetic information could be used in the future". • Data and sample sharing: sharing of cell lines, DNA sequence information, health information and research results shared with other central tissue or data banks such as NIH-sponsored, so that researchers from around the world can use them (sharing done via these other databases). Coded samples and health information may be shared with researchers at Partners institutions as well as researchers at non-Partners institution or with for-profit companies working with Partners researchers (limited sharing by Partners itself). • Commercialization: collaboration with for-profit companies possible for commercialization. *Source: Consent form received by email from biobank representative on August 2, 2016.*
Geisinger MyCode	The Geisinger MyCode project provides a research platform for Precision Medicine. Launched in 2007, it provides a system-wide biobanking program to link samples and EHR data for broad research use. Currently, approximately 100,000 participants are included in the program. • Scope of consent: DNA analysis for Geisinger-approved studies • Duration of participation: samples and data will be kept in biobank indefinitely. • Data collection: data collected from medical record (no explicit mention of ongoing access) • Recontact: recontact possible to collect additional information or present new studies. • Return of individual results: participants agree to be told about results Geisinger gets from studying their samples that might help them. This includes "information that could be important to [the participant's] health (and the health of [his or her] family members). Examples of types of results that could be returned are provided (ex: development of serious diseases, drug response, etc.). There is a detailed clause about the possible return of individual results and the process to return them (informing participant and healthcare provide, placing information in medical record). • Privacy: Acknowledgement of unknown risks ("We do not know every possible risk that might come up in the future") • Data and sample sharing: no explicit mention of international data sharing. List of institutions with which information can be shared (Geisinger, Universities/medical schools; Government agencies; Public agencies; Companies) • Commercialization: commercialization possibilities mentioned. Commercial companies may have access to data. *Source: Geisinger MyCode Research Consent/Authorization Form (dated 02/22/2016).* Available online at: http://www.geisinger.org/for-researchers/partnering-with-patients/includes/pdf/MyCode%20Main%20Consent%20V29.pdf
Marshfield Clinic Personalized Medicine Research Program	The main goal of this project is to apply genetic science to human health. The project will try to do this by creating a Personalized Medicine Research Database. The purpose of the database is to store questionnaire, genetic and medical record information. The objective is to enroll 40,000 individuals in the project. • Scope of consent: Broad consent for different types of research uses: Many researchers will use this database. Researchers will try to learn what genes cause common diseases and to find out which genes predict reactions to drugs. • Duration of participation: database will continue indefinitely (following the health over time). • Data collection: information generated from the sample (DNA, plasma, serum) will be linked to information from medical record (ongoing) and other available medical samples. • Recontact: recontact possible for allowing samples to be used for a specific project, to obtain additional information, update questionnaire or information, to obtain additional sample, to participate in future studies. • Return of individual results: individual research results are not returned. • Data and sample sharing: coded data and samples may be shared with trusted research partners outside of the Marshfield Clinic system. Only approved researchers and staff will have access to data (internal review and IRB review prior to access). No explicit mention of international data sharing. • Commercialization: mention that commercial products, patents or licenses may be developed as a result of the project. *Source: Consent form received by email from biobank representative on August 4, 2016.*
Mayo Clinic	The Mayo Clinic Biobank aims to provide samples for different types of research studies. Many of the studies are aimed at gaining a better understanding of how a person's genes (DNA) may influence overall health and wellness. According to its website, the Mayo Clinic Biobank has reached its goal of enrolling more than 50,000 participants and is no longer recruiting new participants. • Scope of consent: broad consent to future research on the interaction between genes and disease. • Duration of participation: indefinitely. Ongoing participation in project (samples and data kept after death). • Data collection: information collected from medical record (ongoing), questionnaire. • Recontact: recontact possible to administer additional questionnaires, collect additional blood sample.

	• Return of individual results: possible return of individual results based on decision of an internal committee. No options provided to participants. • Data and sample sharing: Biospecimen Trust Oversight Group will be in charge of review access requests. Collaboration with Mayo researchers is required. No explicit mention of international sharing. • Commercialization: commercialization possibilities mentioned. *Source Consent Form (approved July 16, 2015).* Available online at: http://www.mayo.edu/research/documents/biobank-consent-form/doc-20086428.
Children's Hospital of Philadelphia	Children who are treated at the Children's Hospital Healthcare Network and their parents may be eligible to take part in a major initiative to collect more than 100,000 blood samples. • Scope of consent: Broad consent to use samples in information for studies aiming to better understand the causes of complex pediatric disorders. The Center will also search for genes that determine which drugs are most likely to work for an individual patient in order optimize the benefits and minimize any side effects from the drugs. • Duration of participation: taking part lasts 15-30 minutes, and updating of information from medical record once a year thereafter. No further tests required. • Data collection: information collected from child's medical record (and parent medical record, where available). • Linkage with parental medical record: optional portion requesting parents to provide access to their own medical record to help determine impact on child's health. • Recontact: no systematic recontact. Recontact is an "opt-in" choice provided at the end of the consent form. • Return or results: Some individual results returned, if accepted (option at the end of consent form). More specifically, this could include information on disease risk and medication (ex: drug dosage). • Data and sample sharing: samples and data may be shared with outside laboratories (such as NIH, university research groups, and pharmaceutical companies) to store/analyze samples. • Commercialization: pharmaceutical companies mentioned as potentially receiving data or samples. Not specific mention of commercialization. *Source: FAQs and Informed Consent materials (updated on March 24, 2016).* Available online at: https://caglab.org/index.php/for-participants/ic.html
Cincinnati Children's Hospital Medical Centre – Better Outcomes for Children Project	The Cincinnati Children's Hospital "Better Outcomes for Children Project" aims to improve child health by finding new ways to prevent, identify and treat diseases and conditions. The goal is to ask all patients who registers at CCHMC in a participating clinic if they would like to donate any leftover samples for the project. • Scope of consent: Broad consent for a wide variety of conditions, including genetic analysis. • Duration of participation: Samples from leftover tissue and data from medical visits are kept indefinitely (or, for samples, until they are used up). Not explicitly mentioned, but access to information from medical visits appears to be ongoing. • Data collection: information is collected from the CCHMC medical records and research records. • Recontact: recontact is optional and possible only to return important findings. • Return or results: option to return results for "important information that would affect [the participant]/[the child's] medical care". Individual results are only returned if important information is found about a major disease that can be prevented or treated. Option to refuse recontact for the return of results. • Data and sample sharing: data and samples may be shared with doctors and scientists at CCHMC and their partners, which may be outside the medical center and may include commercial partners. • Commercialization: Inclusion of commercial partners is possible. *Source: Better Outcomes for Children Consent Form (approved on 3/10/2015).* Available online at: https://www.cincinnatichildrens.org/patients/visit/during/stay/better-outcomes-for-children

TABLE 3: INTERNATIONAL BIOBANKS AND COHORTS

COHORT	HIGHLIGHTS OF FINDINGS
UK Biobank (United Kingdom)	The UK Biobank is a national health resource, with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia. From 2006 to 2010, the UK Biobank recruited 500,000 individuals aged between 40-69 years, from across the UK. Participants are recruited via their National Health Service record and are directed to an assessment center. • Scope of consent: broad consent for a number of different future uses, including analysis of genetic data. • Duration of participation: follow-up for several years after the assessment visit, through medical record and other records that may be related to health. • Data collection: assessment visit, questionnaires, baseline measures, sample collection (blood, saliva, urine), follow-up surveys, ongoing access to health records and other records that could be related to health (occupational, residential information). • Linkage: linkage with other records related to health. • Recontact: recontact possible to follow-up questionnaires or assessment visits. • Return or results: return of baseline assessment, no return of individual research results. • Data and sample sharing: sharing permitted with approved researchers (scientific and ethics approval), and includes researchers in other countries as well as commercial companies. • Commercialization: Access by commercial companies permitted. *Source: Information Leaflet and Consent Form*. Available online at: http://www.ukbiobank.ac.uk/wp-content/uploads/2011/06/Participant_information_leaflet.pdf?phpMyAdmin=trmKQlYdjjnQIgJ%2CfAzikMhEnx6
100,000 Genomes Project (United Kingdom)	Led by Genomics England, the project aims to sequence 100,000 genomes from around 70,000 people. Participants are NHS patients with a rare disease, plus their families, and patients with cancer. • Scope of consent: broad consent for research, genomic sequencing of samples (note: recruitment limited to cancer, rare diseases, severe reaction to infection) • Duration of participation: not mentioned, but ongoing access to medical record in the future (linkage). • Data collection: samples (blood, tumour (where applicable) undergo full sequencing, data collected from medical records, linkage with other sources possible (GP, disease registries, hospital records, social care records) • Recontact: recontact is possible to collect additional information or to participate in future studies. • Return or results: several options for the return of results are presented, including: (i) non-optional return of individual results related to main conditions (i.e. required in order to take part in the project); (ii) incidental findings are optionally returned with additional screening potentially offered; (iii) optional carrier testing offered to parents (rare diseases); (iv) potential benefit of results to other family members is mentioned. • Data and sample sharing: data sharing possible with researchers and organizations approved by Genomics England, including for-profit healthcare companies, including outside the UK but only "in-house" access (no downloading of data). • Commercialization: access by commercial entities mentioned. *Source: Participant Information Sheets and Consent Forms (in particular, the Cancer Information Sheet).* Available online at: https://www.genomicsengland.co.uk/taking-part/patient-information-sheets-and-consent-forms/
LifeLines (Netherlands)	LifeLines is a multidimensional cohort study and biobank and offers a data resource to study a broad scope of (epi)genetic, biomedical, environmental and psychosocial factors in relation to healthy ageing, disease development, and general well-being. The cohort study started in 2006 and has an estimated size of 165.000 individuals to be followed-up for 30 years. • Scope of permitted research: broad consent to genetic epidemiological studies across a broad range of multifactorial diseases. • Duration of participation: participation over an extended period of time, measurements every 4-5 years (health assessments and questionnaires). • Data collection: sample collection (blood, urine), data collected from medical records, regular assessments, regular questionnaires, linkage with other databases (central bureau of statistics, national registration of causes of death) • Recontact: recontact possible for additional questionnaires and health measures. • Contact of family members: contact of other family members to invite them to participate in LifeLines is allowed. • Return or results: only return of initial health assessment returned to the GP (blood pressure, blood values, lung function, ECG, cholesterol, blood glucose and interview results). No return from genetic analysis mentioned. • Data and sample sharing: Not mentioned in the consent form. However, Data Access Policy indicates that researchers worldwide can apply for access to data and samples.

	• Commercialization: Not mentioned. *Source: Data Access Policy, Appendix 3: Consent Form (translated, dated June 25, 2015).* Available online at: https://lifelines.nl/upload/file/lifelines+data+access+policy_%5B1%5D.pdf
Ontario Health Study (part of the Canadian Partnership for Tomorrow Project)	The Ontario Health Study is part of the Canadian Partnership for Tomorrow Project, which is recruiting and following the health of 300,000 Canadians in five cohort studies across the country. The aim of the particular OHS study is to investigate risk factors that cause diseases like cancer, diabetes, heart disease, asthma and Alzheimer's. Consent and questionnaires are completed online. • Scope of permitted research: broad consent for research into chronic diseases. • Duration of participation: information collected will be kept at least until 2059. Participation is long-term (lifetime). • Data collection: optional blood sample (not required for participation), data from questionnaires, linkage of various records including health insurance claims database, Ontario cancer registry, other databases that may be developed in the future, access to administrative medical records or medical databases and personal medical records. • Recontact: recontact possible for collecting additional information or measurements (questionnaires), for future studies or to provide additional samples. • Return of research results: participants may receive tailored health information on their homepage based on information provided in questionnaires. No other results. • Data and sample sharing: questionnaire answers can be shared with researchers around the world for approved health-research projects. • Commercialization: commercial entities may have access to data, commercialization possible. *Source: Internal document (transcript of the OHS online consent form)*
H3Africa (International/Africa)	The H3Africa biorepository is a pan-African effort to enable African researchers to carry out large-scale studies on African populations. H3Africa aims to use of state-of-the art genomic technologies, in combination with clinical and environmental analyses, with the aim of understanding the interaction of genes and the environment in health and disease. H3Africa will include several GWAS studies (from 1,000 individuals affected by a particular disorder and another 1,000 or so as a healthy control group for comparison, up to 10,000 or more individuals). • Scope of permitted research: broad consent for future use of samples in genetic studies, for a number of different diseases. • Duration of participation: no mention on the length of time of conservation of samples or data. • Data collection: blood sample collection (and genetic analysis), collection of health information (not indicated from where information will be obtained) • Recontact: recontact possible regarding certain uses of samples. • Return or results: in general, no return of individual research results [but can be other options implemented by member projects] • Data and sample sharing: data and sample sharing possible with researchers across the world but prior rights of African researchers. • Commercialization: possible development of commercial products mentioned. *Source: H3Africa Guidelines for Informed Consent Form (August 2013).* Available online at: http://www.health.uct.ac.za/sites/default/files/image_tool/images/116/H3Africa%20Guidelines%20on%20Informed%20Consent%20August%202013.pdf
MalariaGEN (International)	MalariaGEN is a scientific network that connects researchers and clinicians in malaria-endemic countries with DNA sequencing technologies and genomic research. Through a number of multi-centre projects, the network provides a framework for generating, integrating and sharing genetic and genomic data, and for investigating key questions about malaria biology and epidemiology. A template consent form for participating projects was analysed. Individual projects are responsible for creating their own consent forms. • Scope of permitted research: consent to study samples and genetic information for research on malaria. • Duration of participation: participation involves collection of a sample. Sample is kept until all tests are done (no specific end-date). • Data collection: no mention of access to health records or other health data collection. • Recontact: not mentioned. • Return of research results: not mentioned. • Data and sample sharing: material shared with researchers in the UK who are working on the project. • Commercialization: not mentioned. *Source: MalariaGEN Informed Consent Template*. Available online at: https://www.malariagen.net/sites/default/files/content/page/MalariaGEN_Informed_Consent_Template.pdf

References

1. Committee on A Framework for Developing a New Taxonomy of Disease, Board on Life Sciences Division on Earth and Life Studies & National Research Council. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. *Towar. Precis. Med. Build. a Knowl. Netw. Biomed. Res. a New Taxon. Dis.* (2011). doi:10.17226/13284
2. Consortium on Law and Values in the Environment Health and the Life Sciences. LawSeq: Mapping & Shaping the Law of Genomics. at https://consortium.umn.edu/research/lawseq-mapping-shaping-law-genomics
3. Precision Medicine Initiative (PMI) Working Group. The precision medicine initiative cohort program – building a research foundation for 21st century medicine. *Precis. Med. Initiat. Work. Gr. Rep. to Advis. Comm. to Dir. NIH* **Sept 17,** 1–108 (2015).
4. NIH awards $55 million to build million-person precision medicine study. at https://www.nih.gov/news-events/news-releases/nih-awards-55-million-build-million-person-precision-medicine-study
5. The White House. Precision Medicine Initiative: Data Security Policy Principles and Framework. 1–10 at https://www.whitehouse.gov/sites/whitehouse.gov/files/documents/PMI_Security_Principles_Framework_v2.pdf
6. The White House. Precision Medicine Initiative: Privacy and Trust Principles. 1–4 at https://www.whitehouse.gov/sites/default/files/microsites/finalpmiprivacyandtrustprinciples.pdf
7. Simon, C. M., Klein, D. W. & Schartz, H. A. Traditional and electronic informed consent for biobanking: a survey of U.S. biobanks. *Biopreserv. Biobank.* **12,** 423–9 (2014).
8. Simon, C. M., Klein, D. W. & Schartz, H. A. Digitizing Consent: The Food and Drug Administration's Draft Guidance on Electronic Informed Consent. *IRB Ethics Hum. Res.* **38,**
9. Boutin, N. T. *et al.* Implementation of Electronic Consent at a Biobank: An Opportunity for Precision Medicine Research. *J. Pers. Med.* **6,** (2016).
10. Becker, S. *et al.* mHealth 2.0: Experiences, Possibilities, and Perspectives. *JMIR mHealth uHealth* **2,** e24 (2014).
11. Kirby, E., Zawati, M. H. & Knoppers, B. M. Electronic Consent to Health Research in Canada. *Can. Bar Rev.* **91,** 419–435 (2012).
12. Doerr, M., Suver, C. & Wilbanks, J. Developing a Transparent, Participant-Navigated Electronic Informed Consent for Mobile-Mediated Research. at http://ssrn.com/abstract=2769129
13. Global Alliance for Genomics and Health. Mobile Health Consent. at http://genomicsandhealth.org/working-groups/our-work/mobile-health-consent
14. Vanderbilt University. e-Consent (FAQs). (2015). at https://www4.vanderbilt.edu/irb/2015/03/e-consent-faqs/
15. Research Electronic Data Capture (REDCap). at https://projectredcap.org/
16. Lawrence, C. & Power, J. eConsent Guidance and Use Cases. (2016). at http://nchica.org/wp-content/uploads/2016/06/Lawrence-Power.pdf
17. Howser, J. VUMC to Lead pilot program for Precision Medicine Initiative Cohort Program. (2016). at https://news.vanderbilt.edu/2016/02/25/vumc-to-lead-pilot-program-for-precision-medicine-initiative-cohort-program/
18. uBiome. uBiome Support. at https://app.ubiome.com/support
19. 23andMe. 23andMe Consent Document. at https://www.23andme.com/en-ca/about/consent/
20. Nicol, D. *et al.* Precision medicine: drowning in a regulatory soup? *J. Law Biosci.* (2016). doi:10.1093/jlb/lsw018
21. Hendler, J. A Primer on Machine Readability for Online Documents and Data. *Data.gov* at https://www.data.gov/developers/blog/primer-machine-readability-online-documents-and-data
22. HealthIT.gov. Precision Medicine Task Force. at https://www.healthit.gov/FACAS/health-it-standards-committee/hitsc-workgroups/precision-medicine-task-force
23. HealthIT.gov. Standards: Precision Medicine Task Force (Wednesday, May 11, 2016 meeting). (2016). at https://www.healthit.gov/FACAS/calendar/2016/05/11/standards-precision-medicine-task-force
24. Health IT Standards Committee. HIT Standards Committee Precision Medicine Task Force Final Transcript August 5, 2015. (2015). at https://www.healthit.gov/FACAS/sites/faca/files/PMTF_Transcript_Final_2015-08-05.pdf
25. Kaye, J. *et al.* Dynamic consent: a patient interface for twenty-first century research networks. *Eur J Hum Genet* **23,** 141–146 (2015).
26. Williams, H. *et al.* Dynamic Consent: A Possible Solution to Improve Patient Confidence and Trust in How Electronic Patient Records Are Used in Medical Research. *JMIR Med. Informatics* **3,** e3 (2015).
27. Thiel, D. B. *et al.* Testing an Online, Dynamic Consent Portal for Large Population Biobank Research. *Public Health Genomics* **18,** 26–39 (2015).
28. Knoppers, B. M. Precision medicine: a matter of regulation or collaboration? *J. Law Biosci.* (2016). doi:10.1093/jlb/lsw049
29. Steinsbekk, K. S., Kåre Myskja, B. & Solberg, B. Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? *Eur. J. Hum. Genet.* **21,** 897–902 (2013).
30. Teare, H. J. A., Morrison, M., Whitley, E. A. & Kaye, J. Towards 'Engagement 2.0': Insights from a study of dynamic consent with biobank participants. *Digit. Heal.* **1,** (2015).
31. National Institutes of Health (NIH) - National Human Genome Research Institute. Informed Consent for Genomic Research. (2016). at https://www.genome.gov/27026588/informed-consent-for-genomics-research/
32. Knoppers, B. M., Zawati, M. H. & Senecal, K. Return of genetic testing results in the era of whole-genome sequencing. *Nat Rev Genet* **16,** 553–559 (2015).
33. Burton, P. R. *et al.* Size matters: just how big is BIG?: Quantifying realistic sample size requirements for human genome epidemiology. *Int. J. Epidemiol.* **38,** 263–273 (2009).
34. Caulfield, T. *et al.* A review of the key issues associated with the commercialization of biobanks. *J. Law Biosci.* **1** , 94–110 (2014).
35. ICH Harmonised Tripartite Guideline. ICH harmonized tripartite guideline: Guideline for Good Clinical Practice. **47,** 45–50 (1996).
36. ICH. ICH E18: Guidelines on Genomic Sampling and Management of Genomic Data (Draft). *2015*
37. Organization for Economic Co-operation and Development. OECD Guidelines on Human Biobanks and Research Databases. (2009). at https://www.oecd.org/sti/biotech/44054609.pdf
38. International Society for Biological and Environmental Repositories. 2012 Best Practices for Repositories Collection, Storage, Retrieval, and Distribution of Biological Materials for Research International Society for Biological and Environmental Repositories. *Biopreserv. Biobank.* **10,** 79–161 (2012).
39. Global Alliance for Genomics and Health. Consent Policy. (2015). at http://genomicsandhealth.org/work-products-demonstration-projects/consent-policy
40. Global Alliance for Genomics and Health. Framework for Responsible Sharing of Genomic and Health-related Data. (2015). at https://genomicsandhealth.org/files/public/Consent Policy %28Final - 27 May 2015%29.pdf
41. Future of Privacy Forum. *Best Practices for Consumer Wearables & Wellness Apps & Devices*. (2016). at https://fpf.org/wp-content/uploads/2016/08/FPF-Best-Practices-for-Wearables-and-Wellness-Apps-and-Devices-Final.pdf
42. National Cancer Institute. *NCI Best Practices for Biospecimen Resources*. (2016). at https://biospecimens.cancer.gov/bestpractices/2016-NCIBestPractices.pdf
43. U.S. Department of Health and Human Services *et al.* Use of Electronic Informed Consent in Clinical Investigations - Questions and Answers - Draft Guidance. *2015* at http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm436811.pdf
44. Health Level Seven (HL7). Consent Tracker FHIR Resource Proposal. at http://wiki.hl7.org/index.php?title=Consent_Tracker_FHIR_Resource_Proposal
45. International Standards Organisation. ISO/TS 17975:2015 - Health informatics -- Principles and data requirements for consent in the Collection, Use or Disclosure of personal health information. (2015). at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=61186
46. Global Alliance for Genomics and Health. Automatable Discovery and Access Task Team. at https://genomicsandhealth.org/working-groups/our-work/automatable-discovery-and-access
47. Global Alliance for Genomics and Health ADA Task Team. Automatable and Discovery Access Matrix Guidance Document. (2016). at https://genomicsandhealth.org/files/public/ADAM_GuidanceDocument_02Sept2016_Draft.pdf
48. Global Alliance for Genomics and Health. Automatable Discovery and Access Matrix. (2016). at http://genomicsandhealth.org/work-products-demonstration-projects/automatable-discovery-and-access-matrix
49. Dyke, S. O. M. *et al.* Consent Codes: Upholding Standard Data Use Conditions. *PLoS Genet* **12,** e1005772 (2016).
50. Global Alliance for Genomics and Health. Consent Tools. at https://genomicsandhealth.org/files/public/GA4GH-Consent Tools-FINAL.pdf
51. Public Population Project in Genomics and Society (P3G). Generic Information Pamphlet and Consent Form. at http://www.p3g.org/system/files/biobank_toolkit_documents/P3G Generic Info Pamphlet and Consent Form for Biobanks_0.pdf
52. Public Population Project in Genomics and Society (P3G). P3G-IPAC Generic Clauses / Agreements Database. at http://www.p3g.org/resources/ipac
53. National Institutes of Health (NIH) - National Human Genome Research Institute. Informed Consent for Genomic Research. at https://www.genome.gov/27559023/informed-consent-sample-consent-forms/
54. National Institutes of Health (NIH) - National Human Genome Research Institute. Special Considerations for Genome Research. *June 16, 2016* at https://www.genome.gov/27559024/informed-consent-special-considerations-for-genome-research/
55. Sage Bionetworks. Participant-Centered Consent Toolkit. at http://sagebase.org/platforms/governance/participant-centered-consent-toolkit/
56. Electronic Medical Records and Genomics (eMerge Network). at https://emerge.mc.vanderbilt.edu/
57. The Electronic Medical Records and Genomics (eMERGE), Workgroup Consent & Community Consultation & Force Informed Consent Task. Model Consent Language. (2009). at http://www.ninds.nih.gov/research/clinical_research/application_process/eMerge_model_language.pdf
58. PCORnet. Informed Consent Recommendations for Specimen Acquisition and Future Use. (2015). at http://www.pcornet.org/wp-

content/uploads/2015/06/BiorepositoriesTF_InformedConsentGuidance_22June2015.pdf
59. Genetic Alliance. Platform for Engaging Everyone Responsibly (PEER). at http://www.geneticalliance.org/programs/biotrust/peer/
60. Sync4Science. Sync4Science. at http://www.sync4science.org/
61. Rothstein, M. A. Some lingerings concerns about the precision medicine initiative. *J. Law, Med. Ethics* **43,** (2016).
62. American Bar Association - Health Law Section. *Re: Precision Medicine Initiative: Proposed Privacy and Trust Principles*. at http://www.americanbar.org/content/dam/aba/administrative/healthlaw/aba_pmi_comments_15.authcheckdam.pdf
63. World Privacy Forum. The Precision Medicine Initiative and Privacy: Will Any Legal Protections Apply? at http://www.worldprivacyforum.org/wp-content/uploads/2016/05/WPF_PrecisionMedicineInitiative_May2016_fs.pdf
64. Federal Drug Administration. Use of Standards in FDA Regulatory Oversight of Next Generation Sequencing (NGS)-Based In Vitro Diagnostics (IVDs) Used for Diagnosing Germline Diseases - Draft Guidance. at http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM509838.pdf
65. Federal Drug Administration. Use of Public Human Genetic Variant Databases to Support Clinical Validity for Next Generation Sequencing (NGS)-Based In Vitro Diagnostics. at http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM509838.pdf
66. U.S. Food and Drug Administration. Public Workshop - Patient and Medical Professional Perspectives on the Return of Genetic Test Results, March 2, 2016. at http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm478841.htm
67. Regulations.gov. Patient and Medical Professional Perspectives on the Return of Genetic Test Results - Docket Summary. at https://www.regulations.gov/docket?D=FDA-2015-N-4809
68. Department of Health and Human Services. NPRM for federal policy for the protection of human subjects. (2015). at https://www.federalregister.gov/articles/2015/09/08/2015-21756/federal-policy-for-the-protection-of-human-subjects
69. National Institutes of Health Advisory Committee to the Director - National Children's Study (NCS) Working Group. Final Report - December 12, 2014. (2014). at http://acd.od.nih.gov/reports/ncs_wg_final_report.pdf
70. Global Alliance for Genomics and Health. Data Sharing Lexicon. (2016). at https://genomicsandhealth.org/work-products-demonstration-projects/work-products/data-sharing-lexicon
71. Wallace, S. E. What Does Anonymization Mean? DataSHIELD and the Need for Consensus on Anonymization Terminology. *Biopreserv. Biobank.* **14,** 224–230 (2016).
72. Joe Alper (Rapporteur); Roundtable on Health Literacy; Board on Population Health and Public Health Practice; Health and Medicine Division; National Academies of Sciences Engineering and Medicine. *Relevance of Health Literacy to Precision Medicine: Proceedings of a Workshop*. (2016).

 Printed in USA Please Recycle 335624-001US