# Blockchain and Health IT: Algorithms, Privacy, and Data

WHITE PAPER

August 8, 2016

Prepared for:

Office of the National Coordinator for Health Information Technology
U.S. Department of Health and Human Services

Prepared by:

Allison Ackerman Shrier, Anne Chang, Nadia Diakun-thibault, Luca Forni,
Fernando Landa, Jerry Mayo, Raul van Riezen
Project PharmOrchard$^{\text{TM}}$ of MIT's Experimental Learning "MIT Fintech: Future Commerce"
&
Thomas Hardjono MIT Connection Science

Correspondence:
Allison Ackerman Shrier
aackerman26@gmail.com
617-971-6915

# Summary

The President's **Precision Medicine Initiative (PMI)** is "enabling a new era of clinical care through research, technology, and policies that empower patients, researchers, and providers to work together toward the development of individualized care".[1] Its commitment to privacy and security in the setting of responsible data sharing and transparency is articulated in the "Privacy and Trust Principles"[1] and the "Data Security Policy Principles and Framework"[2], developed by an interagency working groups including the Office of the National Coordinator for Health Information Technology in conjunction with multiple stakeholders.

In this paper, we review the threats to the security, confidentiality, integrity, and availability of PMI data. PMI organizations can mitigate these challenges through a new system architecture in development at MIT -- the OPAL/Enigma project[3] -- which creates a peer-to-peer network that enables parties to jointly store and analyze data with complete privacy, based on highly optimized version of multi-party computation with a secret-sharing. An auditable, tamper-proof distributed ledger (a permissioned blockchain) records and controls access through smart contracts and digital identities. We conclude with an initial use case of OPAL/Enigma that could empower precision medicine clinical trials and research.

MIT's OPAL/Enigma challenges traditional data security paradigms. Centralized databases cannot assure security and data integrity, regardless de-identification and controlled access requirements. Safe, vetted queries that are distributed to private, encrypted databases assure that organizations and participants can share health care data with cryptographic guarantees of privacy with various stakeholders, assuring momentum for a new era of medical research and practice.

# Contents

## Introduction

In an online world where users expect instant information and seamless flow of data, stakeholders expect new technologies to be absorbed by society and industries as soon as it is available. However, this has not been the case of the healthcare industry, despite technologic advances and economic relevance of technological innovation.
The Office of the National Coordinator for Health Information Technology (ONC)[4] acknowledged that healthcare stakeholders should assist in creating a new infrastructure for the industry and its users, whether government agencies, drug developers, doctors or patients.

The foundation of a new healthcare IT system lays in, among other measures, the creation of a platform that allows interoperability; the adoption of Electronic Health Records (EHR), creation of standards and formats that can be widely adopted by the stakeholders; safe storage of all the data that is collected by all healthcare agents with the highest regard to participant privacy and secure and efficient exchange of health information between the parties respecting the privacy of all those whose information is involved.

In this paper, we posit the adoption of OPAL/Enigma[5,6], an encrypted platform that is able to create a secure environment for the storage and analysis of healthcare information by using blockchain technology, as an effective solution to address the privacy and security concerns of the stakeholders. OPAL/ENIGMA is also a potential tool to resolve infrastructural matters, such as time and cost related to the analysis, storage and manipulation of health information. A framework for precision medicine trials, development of more comparative trials, cheaper development of drugs and assignment of more effective treatments to patients are also potential future benefits.

## Privacy, security, and trust in healthcare IT

### Overview of guiding principles

Interoperability not only means having the ability to exchange information, but also being able to use the exchanged information.   In this regard, in addition to real-time, secure data sharing, there should also be mechanisms that ensure data provenance (identifying the original source of the information), data verification/accuracy and that the respective consent has been duly obtained. Failure in obtaining and communicating patient consent could give cause to liabilities related to breach of electronic protected health information. The concerns on being able to comply with different state regulations specially in connection with patient privacy and rules on health information exchange, has led to debate on how data sharing should be handled, to what purposes it should be used and who should have access to it.

The ONC has a 10-year vision for an "interoperable health IT ecosystem", in which health information flows seamlessly and securely to the right people, at the right place, at the right time.[4,7]  This interconnected "learning health system" will enable the rapid dissemination of new knowledge to support the use of best evidence in the care of all patients, lower healthcare costs, improve population health, empower patients, and drive innovation.   A guiding principle is to protect privacy and security in all aspects of interoperability, with strong and effective safeguards as well as greater transparency, in order to establish and maintain public trust.

Current ONC standards and goals rely heavily on encryption technology for secure data transmission between organizations in a healthcare IT ecosystem. In an increasingly interconnected healthcare IT system, security is difficult to ensure for all organizations in the larger healthcare IT ecosystem that will have access to PMI data. Perimeter defenses such as hardware and software intrusion detection systems fail to secure vital health IT infrastructure.[8] The failure rate is high and the root cause is "the human factor", our reliance on human maintenance and intervention. A "learning health system" must leverages lessons learned about cybersecurity to answer new questions.

In order to unlock the potential value of data sharing, we need better solutions to manage data security. Centralized IT systems offers advantages in terms of efficiency, however, frequent data breaches, lack of transparency, and loss of data integrity have led to the adoption of networks which distribute authority among many trusted actors, so that significant security compromises would require consensus. In addition to this distributed consensus mechanisms, blockchain is a distributed ledger that provides a immutable and auditable record of actions (and actors). Technologies based on blockchain may provide more optimal solutions for inherently safe health IT ecosystems.

We posit that the ONC should support blockchain-based architectures that meet the needs of future health IT ecosystems can lead to improved clinical guidelines and practices.

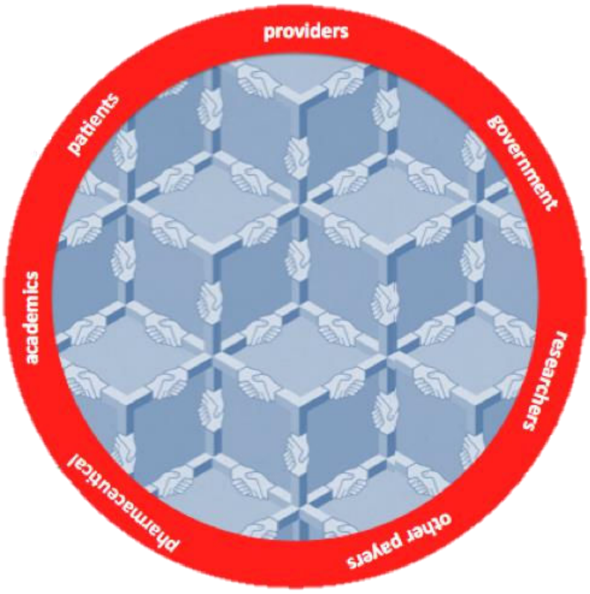

Figure Blockchain will interconnect stakeholders[7]

### Tension between data sharing and privacy

Two competing needs must be satisficed: the need for data-sharing between stakeholder (government public sector, industry private sector, healthcare providers, researchers, and patients) and the need for privacy and confidentiality. These create an unsustainable ecosystem if based on current legacy protocols and legacy hardware/software that do not guarantee data integrity. The first requirement of an optimal solution is a permissioned distributed ledger, which automatically tracks every change to the data, so it is auditable, trustworthy, and transparent (the blockchain).[8,9] On top of the blockchain, is a layer that provides secure, privacy-preserving processing. In order to ensure that entities are engaging in data sharing that is appropriately permissioned, without revealing private or confidential information, there must be a mechanism for automation of data processing that verifies permissioned identities, will ultimately bind the parties in a data usage agreement that is compliant. Finally, data must remain encrypted, during storage and processing in personal data stores. We posit that MIT OPAL/Enigma may provide an architecture with these attributes.

## OPAL/Enigma challenges the security paradigm

### Overview of OPAL/Enigma

Resolving the tension between data sharing and privacy, in order to unlock the potential value of that data, is a challenge for healthcare IT.[3,10–13] Although centralized databases for processing may be conceptually simpler, distributed networks with privacy maximizing algorithms can maximize security for data processing, in compliance with HIPAA, confidentiality and other regulatory and ethical requirements without the limited scalability of current distributed processing platforms.

OPAL/Enigma[3,6] envisions distributed data repository architecture on a **peer-to-peer (P2P)** network where data is encrypted at its repository so that raw data is never released. Data remains secure during storage and analysis, because data can be queried, but only by queries that are permissioned by digital identity credentials for specific data operations defined by legally binding smart contracts. An unalterable and auditable record of patterns of communications between data and operators, including credentials and data operations is recorded, creating a distributed cryptographic ledger, or permissioned blockchain.

At the level of the **data repository node**, data is encrypted by **Enigma** and remain encrypted in storage and during computation, which counters internal data theft. The data repository owner has control over granularity of answers to queries and therefore privacy. The **OPAL** algorithm essentially "moves the algorithm to the data" by using distributed query processing to ship queries and sub-queries to data repositories, where computation occurs, so that each data repository returns only de-identified aggregated results. The queries can be permissioned by smart contracts, or **Query Smart Contract (QSC),** that legally bind the **Querier** (person or organization), the data repository and other entities in a data usage agreement. QSC are the mechanism for operating within existing trust frameworks and respecting the core values, responsible strategies, and legal principles for participating entities. The blockchain provides assurance that QSC are honored with a tamper-proof and auditable history of identity and operations of data access.

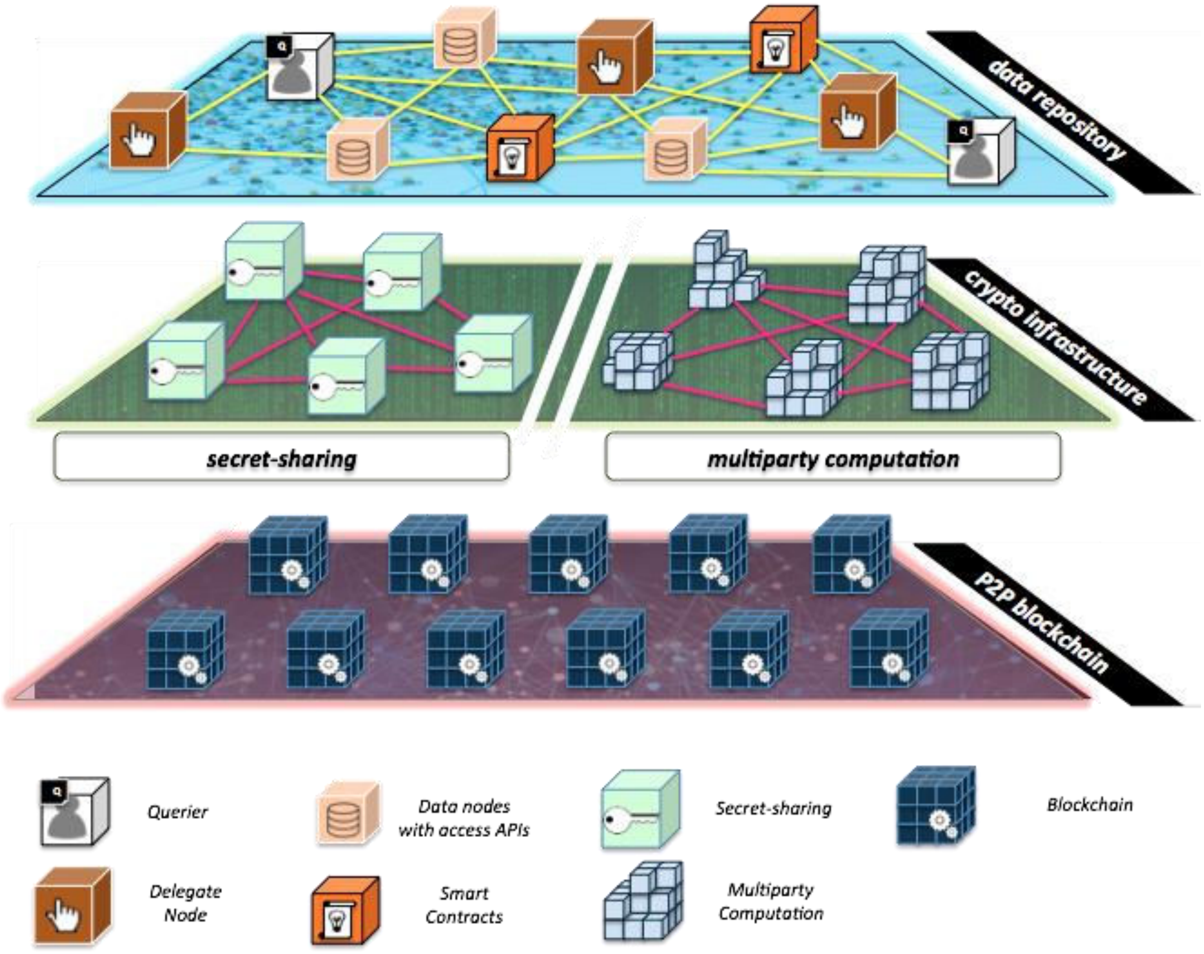

Figure Layers of OPAL/Enigma **On top of the P2P network, or blockchain, Enigma uses two cryptographic constructs simultaneously (secret-sharing and MPC), so that OPAL can direct data exchange with QCS.**

## Enigma provides cryptography

The MIT Enigma system is part of a larger initiative at MIT called **OPAL-EAST (Open Algorithms for Equity, Accountability, Security, and Transparency**).[12,13] MIT Enigma employs two core cryptographic constructs simultaneously atop a Peer-to-Peer (P2P network of nodes). These are **secret-sharing** (a la Shamir's Linear Secret Sharing Scheme (LSSS))[14] and **multiparty computation (MPC)**.[15–17] Although secret sharing and MPC are topics of research for the past two decades, the innovation that MIT Enigma brings is the notion of employing these constructions on a P2P network of nodes (such as the blockchain) while providing "Proof-of-MPC" (like proof of work) that a node has correctly performed some computation.

In secret-sharing schemes, a given data item is "split" into a number of ciphertext pieces (called "shares") that are then stored separately. When the data item needs to be reconstituted or reconstructed, a minimum or "threshold" number of shares need to be obtained and merged together again in a reverse cryptographic computation. For example, in Naval parlance this is akin to needing 2 out of 3 keys in order to perform some crucial task (e.g. activate the missile). Some secret sharing schemes possess the feature that some primitive arithmetic operations can be performed on shares (shares "added" to shares) yielding a result without the need to fully reconstitute the data items first. In effect, this feature allows operations to be performed on encrypted data (similar to homomorphic encryption schemes).

The MIT Enigma system proposes to use a P2P network of nodes to randomly store the relevant shares belonging to data items. In effect, the data owner no longer needs to keep a centralized database of data-items (e.g. health data) and instead would transform each data item into shares and disperse these on the P2P network of node. Only the data owner would know the locations of the shares, and can fetch these from the nodes as needed. Since each of these shares appear as garbled ciphertext to the nodes, the nodes are oblivious to their meaning or significance. A node in the P2P network would be remunerated for storage costs and the store/fetch operations.

The second cryptographic construct employed in MIT Enigma is multiparty computation (MPC). The study of MPC schemes seeks to address the problem of a group of entities needing to share some common output (e.g. result of computation) whilst maintaining as secret their individual data items. For example, a group of patients may wish to collaboratively compute their average blood pressure information among them, but without each patient sharing actual raw data about their blood pressure information.

The MIT Enigma system combines the use of MPC schemes with secret-sharing schemes, effectively allowing some computations to be performed using the shares that are distributed on the P2P. The combination of these 3 computing paradigms (secret-sharing, MPC and P2P nodes) opens new possibilities in addressing the current urgent issues around data privacy and the growing liabilities on the part of organizations who store or work on large amounts of data.

## OPAL provides query smart contracts

An important concern in the broad areas of PMI is the need to maintain the privacy of individuals (e.g. patients) while being able to perform computations (e.g. statistical analysis) on data pertaining to those individuals. In many studies (e.g. clinical trials) it is often more relevant to obtain aggregate answers instead of narrow answers pertaining to a small fraction of the study participants.

One paradigm shift being championed by the MIT OPAL/Enigma community is that of using "pre-fabricated" queries (e.g. SQL queries) that have been analyzed by experts and have been vetted to be "safe" from the perspective of privacy-preservation.[18,19] The term "Open Algorithm" (OPAL) here implies that the vetted queries ("algorithms") are made open by publishing them, allowing other experts to review them and allowing other researchers to make use of them in their own context of study.

The next step in the Open Algorithms paradigm is the use of smart contracts to capture these safe algorithms in the form of executable queries residing in a legally binding digital contract. A query smart contract will require that querier authorization requirements be encoded within the contract to be recorded on the blockchain. A query smart contract that has been vetted to be safe and digitally signed by expert can be stored on nodes of the blockchain, so the entities seeking to query can find prefabricated safe QSCs that match the intended application.

### Network

The P2P network contains **data repository nodes** where the data is entered, encrypted, stored, and analyzed. In general, standardized interoperability will be a feature to facilitate data exchange later in development- as noted later in the discussion of the use case many electronic health records are legacy systems, which do not provide interoperability with standardized APIs, and alternative methods will be used in anticipation of this standardization. The node may be a partition on a server with limited processing power. More processing power could be effectively loaned by **delegate nodes** on the P2P network, which work on a query or subquery by locating the relevant data repositories, sending the appropriate subquery to each data repository, and receiving individual answers and collating the results received from these data repositories for reporting to the querier. Delegate nodes assure that all the conditions of the QSC have been fulfilled (e.g. QSC has valid signature; identity of querier is established; authorization to access APIs at data repositories has been obtained; payment terms has been agreed if there is a remuneration scheme, etc.). A hierarchy of delegate nodes may be involved in the completion of a given query originating from the querier entity. **QSC nodes** will maintain vetted QSCs and function as a point of origination of the query for the querier. All QSC nodes will maintain distributed copies of the ledger; more nodes could maintain the ledger but further development is needed to determine the appropriate balance of redundancy to ensure consensus as a way of security-proofing the network vs. efficiency, and the balance between security vs. efficiency for methods establishing proof-of-work. As OPAL/Enigma is currently in development, but likely limitations will be computational expense and interoperability, as there is a lack of standardized APIs in the healthcare IT relative to other industries, which underscores the importance of mission of ONC.

## OPAL/Enigma applied to precision medicine

### Precision Medicine

Permissioned distributed cryptographic ledgers focused on privacy, like OPAL/Enigma, are creating novel tools to empower innovation in healthcare through overcoming the data sharing challenges like privacy.[10,12,13] For example, a goal of President Obama's **Precision Medicine Initiative (PMI)** is to set "the foundation for a new way of doing research that fosters open, responsible data sharing with the highest regard to participant privacy." (https://www.whitehouse.gov/precision-medicine) We will discuss a specific use case, with implications for not only precision medicine but also other federal healthcare research efforts.

Precision medicine is an innovative approach that provides a holistic understanding of a patient's health, disease and condition, and a means to choose treatments that would be most effective for an individual. Translating initial successes to a larger scale will require a coordinated and sustained national effort, and PMI will guide these efforts through **PMI Cohort Program** and **PMI for Oncology**, which are focusing activities on engaging one million Americans to volunteer their health data for clinical research purposes and expanding precision medicine clinical trials and laboratory models in order to develop a

National Cancer Knowledge System - a comprehensive database that would be useful to researchers as well as clinicians actively treating patients.

For many conditions, today's best preclinical models do not recapitulate the individual nuances for patients, and so precision medicine has spurred the development of innovative clinical trial design. The novel trial designs have evolved not only to study low prevalence subsets in a resource-effective fashion, but also to understand the nuances of the translational scientific issues that complicate drug development. Many of these precision medicine trials, several of which are under the auspices of the PMI, use varieties of biomarker-driven adaptive platform trials. Platform trials simultaneously test multiple therapies, and so they can yield comparative evidence between therapies.[20] Adaptive trials prospectively plan dynamic modification of one or more specified aspects of the study design and hypotheses during the trial, based on the statistical analysis of data from subjects in the study. Modifications may include dosage, sample size, and choice of drug or combination therapy, patient selection criteria, and biomarkers. Adaptive platform trials often use Bayesian framework, with data borrowing techniques between treatment arms, in order to evolve and learn the best fit between a therapy and a group of individuals with given biomarker.[21–27] If used correctly, with designs to reduce the false positive rate, then adaptive platform clinical trials are cost and time efficient and produce superior comparative evidence.

## Use Case

Our goal is to develop an infrastructure for improving precision medicine clinical trials and creating methods for the development of national knowledge systems in cancer and other therapeutic areas. Executing adaptive platform and other precision medicine trials at scale will require the development of infrastructure which allows data abstraction from electronic health records, secure data handling with regards to patient confidentiality, HIPAA and other regulatory/compliance, and data exchange for statistical analysis across pharmaceutical industry competitors. In brief, it needs an interoperable IT infrastructure that can performs computationally complex analysis, while preserving the privacy of patients and the confidential intellectual property of pharmaceutical competitors, within a highly regulated trust network. MIT OPAL/Enigma is an ideal solution for a clinical trial blockchain as its MPC algorithms enable data borrowing and statistical analysis without a trusted 3rd party across a distributed network with security provided by high level encryption of distributed data, digital identification proofs and smart contracts that ensure compliance with regulatory and pharmaceutical sponsor wishes.

The pilot of the clinical trial blockchain would create the IT infrastructure for an adaptive platform trial run by a consortium; the design, including the clinical sites and the therapies, would be predetermined by that entity. A startup to enter into a Business Associate agreement would have to be formed, termed PharmOrchard$^{TM}$ given the compliance requirements for clinical trials, and smart contracts (self executable computer protocols based on legally binding contracts) would be written specifically to run on node networks, and to be enforced immediately once the pre-determined conditions are met.

The new entity would utilize the OPAL/Enigma infrastructure in order to write an algorithm which sends queries and sub-queries to the data repository nodes at the clinical sites, controlled by legally-binding, blockchain secured smart contracts which determine the permissions and the content of the queries, in compliance with all regulatory guidance on standard trial protocol as well as the principles outlined by “Precision Medicine Initiative: Privacy and Trust Principles”. Queries would investigate if patients were eligible for the trial given the inclusion criteria, were eligible for a given treatment arm based on biomarker profiles, were tolerant of treatment by predetermined CTCAE and patient-reported parameters, and were responding by standardized and patient-reported outcome measures. Statistical analysis using standard Bayesian methods would determine, as biomarker-correlated response and safety data becomes available, if more or fewer patients should be allocated to different treatment arms as treatments prove more likely to succeed towards regulatory approval or fail. This analysis is distributed and performed on the encrypted data in situ at the data repository nodes.

In the first phase of this implementation, the data repository nodes at each of the clinical sites (clinical sites function as the data owner) would require a local solution to abstract the clinical data in real time from electronic health records (EHR) - ideally a local interoperability solution as demonstrated by TRANSCEND for the I-SPY family of adaptive trials at UCSF[28–30] although some clinical sites may have a more traditional data entry. Data repositories could also reside at the pharmaceutical companies sponsoring the therapies, as there is a minimum size for the P2P network, although it is the QSC nodes that enable appropriately permissioned smart contract queries. The QSC nodes would exist at PharmOrchard$^{TM}$ and consortium governance headquarters at minimum for vetting.

The NIH is working with National Human Genome Research Institute, in parallel with other institutions, in order to develop standardized consent forms that anticipate indefinite biobanking, future testing and reinterpretation of historic testing for clinical trial patients, and PMI-WG [31] are developing “a standardized consent protocol to ensure consistency in the terms and conditions that all PMI cohort participants agree to”.

In the first phase of the pilot, the data could queried securely by others through permissioned queries; the data remains always encrypted at the clinical center/data repository node which generated it. This creates a new paradigm for broadening transparency and data sharing of clinical trial data, well beyond calls by clinicaltrials.gov, ICHME or others for databases of individual patient data that are either highly abstracted/limited or secured insufficiently with standard de-identification techniques. For example, the FDA, NCI, the companies, or another highly permissioned querier like the clinical trial sites themselves could have access to this unique and rich knowledge system. More importantly, there is the underlying tamper-proof distributed cryptographic ledger of every query, including the querier, so that regulatory agencies could perform an audit of the data access to ensure it has been handled responsibly at any time.

### Implications of OPAL/Enigma applied to precision medicine

The largest adaptive platform trials anticipate on the order of a thousand patients, which is an ideal size for a pilot implementation of MIT OPAL/Enigma during its development phase. In addition, taking advantage of the well-developed clinical trial infrastructure with defined consent, patient recruitment, trial management, and data handling protocols will allow rapid deployment. Even a small scale implementation has the potential to influence the industry's outlook on adaptive platform trials, which heretofore have required governance and oversight from government and academic consortia because of the need for a trusted 3rd party, and thus been unable to demonstrate financial sustainability for industry. Drug development involves multiple stakeholders including pharma, patients, providers, regulators, public health advocates, and payers, and as the emerging discussion of healthcare costs draws attention to major beneficiaries - namely pharma - the unmet needs of other stakeholders will inevitably change the socio-cultural, technological, economic, political, legal, environmental and ethical landscape of clinical trials. Adaptive platform trials have the potential to not only deliver comparative evidence to enable precision medicine, but provide faster and optimized access to therapies for patients, treatment guidance and informational tools for providers, financial sustainability to industry, and better comparative evidence for regulators, public health experts, and payers. Further iterations could involve trial designs including SMART protocols, a type of adaptive platform trial that enable switching therapies in the setting of drug resistance, another area of interest of the PMI.

At a later phase in development of this Enigma-based implementation will enable participant empowerment through innovative and responsible access to information, inspired by "Precision Medicine Initiative: Privacy and Trust Principles". Participants could be permissioned on the blockchain to access their own data or aggregated research data, and thus form a pilot for the PMI-Cohort programs or PCORI that will need IT solutions that in their own words "Pioneer a new model of research that engages clinical trial participants, responsible data sharing, and privacy protection."

We note that broader implementation of the MIT OPAL/Enigma architecture could:

- Encourage public-private partnerships
- Ensure data access
- Allow for open data
- Promote "Big data" applications in health IT

## Recommendation

Using secure MPC on P2P network, which allows process of distributed encrypted data in response to appropriately permissioned queries controlled by smart contracts, OPAL/Enigma has the potential revolutionize responsible data sharing and privacy protection. Given the extraordinary need for security in the healthcare IT sector, and the broad potential for innovation critical to national needs, we recommend the following:

- R&D for an OPAL/Enigma-based implementation in precision medicine clinical trial, designed by an entity which can enact a Business Associate agreement with an NCI supported consortium seeking an IT infrastructure to support the "Precision Medicine Initiative: Privacy and Trust Principles
- Discussion with NIST, ONC, and other standard-setting and regulatory bodies on the critical interoperability question
- Further discussion of implications to PMI- Cohort Program
- R&D for open source Enigma, OPAL, and related projects at MIT

## References

1. The White House. Precision Medicine Initiative : Privacy and Trust Principles. 1–4 (2015).
2. *THE WHITE HOUSE Precision Medicine Initiative: Data Security Policy Principles and Framework*. (2016).
3. Zyskind, G., Nathan, O. & Pentland, A. Enigma: Decentralized Computation Platform with Guaranteed Privacy. (2015). at <http://arxiv.org/abs/1506.03471>
4. ONC. Connecting Health and Care for the Nation A Shared Nationwide. 1–25 (2014). at <http://www.healthit.gov/sites/default/files/nationwide-interoperability-roadmap-draft-version-1.0.pdf>
5. Zyskind, G., Nathan, O. & Pentland, A. Enigma: Decentralized Computation Platform with Guaranteed Privacy. *Proc. IEEE Symp. Secur. Priv. Work.* 180–194. (2015). at <http://arxiv.org/abs/1506.03471\nhttp://enigma.media.mit.edu/\nhttp://enigma.media.mit.edu/enigma_full.pdf\nhttp://www.arxiv.org/pdf/1506.03471.pdf>
6. MIT Internet Trust Consortium. Project Enigma. at <http://www.mit-trust.org/projects/>
7. Page, D. A Strategist's Guide To Blockchain. (2016). at <http://www.strategy-business.com/article/A-Strategists-Guide-to-Blockchain?gko=0d586>
8. Shrier, D., Wu, W. & Pentland, A. Blockchain & Infrastructure ( Identity , Data Security ). 1–19 (2016).
9. Shrier, D., Sharma, D. & Pentland, A. Blockchain & Financial Services : The Fifth Horizon of Networked Innovation WHITE PAPER EXCERPT. 1–10 (2016).
10. Pentland, A., Reid, T. & Heibeck, T. Revolutionizing medicine and Public Health. *World Innov. Summit Heal.* 40 (2013).
11. World Economic Forum. *Personal data : The emergence of a new asset class*. *Forum American Bar Association* (2011). at <http://www.weforum.org/reports/personal-data-emergence-new-asset-class>
12. Pentland, A. Reality Mining of Mobile Communications: Toward A New Deal On Data. *Soc. Comput. Behav. Model.* 1 (2009). doi:10.1007/978-1-4419-0056-2_1
13. de Montjoye, Y.-A., Wang, S. & Pentland, A. On the Trusted Use of Large-Scale Personal Data. *IEEE Data Eng. Bull.* **35,** 5 – 8 (2012).
14. Shamir, A. 'How to Share a Secret',. *Commun. Assoc. Comput. Mach.* **22,** 612–3. (1979).
15. Yao, A. C. 'Protocols for Secure Computations.' *Proc. Twenty-Third Annu. Assoc. Comput. Mach. Symp. Theory Comput.* 160–164. (1982).
16. Chaum, D., Crépeau, C. & Damgård, I. Multiparty Unconditionally Secure Protocols. *Proc. Twent. Annu. ACM Symp. Theory Comput.* 11–19 (1988). doi:10.1007/3-540-48184-2_43
17. Abbe, E. A., Khandani, A. E. & Lo, A. W. Privacy-preserving methods for sharing financial risk exposures. in *American Economic Review* **102,** 65–70 (2012).
18. Hazard, J. & Hardjono, T. CommonAccord: Towards a Foundation for Smart Contracts in Future Blockchains, Blockchains and the Web: W3C Workshop on Distributed Ledgers on the Web, June 2016. in
19. Hardjono, T., Greenwood, D. & Pentland, A. Towards a Trustworthy Digital Infrastructure for Core Identities and Personal Data Stores, Proc ID360 Conference on Identity, University Texas at Austin, May 2013. in
20. Berry, S. M., Connor, J. T. & Lewis, R. J. The platform trial: an efficient strategy for evaluating multiple treatments. *JAMA* **313,** 1619–20 (2015).
21. Berry, D. A. How to take clinical research to the next level. *Fortune Insider* (2015). at <http://fortune.com/author/donald-berry/>
22. Berry, S. M., Broglio, K. R., Groshen, S. & Berry, D. A. Bayesian hierarchical modeling of patient subpopulations: Efficient designs of Phase II oncology clinical trials. *Clin. Trials* **10,** 720–734 (2013).
23. Berry, D. A. Emerging innovations in clinical trial design. *Clin. Pharmacol. Ther.* **99,** 82–91 (2016).
24. Berry, D. A. Bayesian clinical trials. *Nat. Rev. Drug Discov.* **5,** 27–36 (2006).
25. McClellan, M., LaVange, L., Pazdur, R., Sridhara, R., Kosorok, MR., Berry, DA., Mehta, C. Pioneering Statistical Approaches to Accelerate Drug Development through Adaptive Trial Designs. (2016). at <https://custom.cvent.com/00BF8C9066844371A697530EA9BB54B7/files/875adda52b2440afa33004ed4689f664.pdf>
26. Berry, D. a. The Brave New World of clinical cancer research: Adaptive biomarker-driven trials integrating clinical practice with clinical research. *Mol. Oncol.* **9,** 951–959 (2015).
27. Bates, S. E. *et al.* Advancing Clinical Trials to Streamline Drug Development. *Clin. Cancer Res.* **21,** 4527–35 (2015).

28. Esserman, L. J. & Woodcock, J. Accelerating identification and regulatory approval of investigational cancer drugs. *JAMA* **306,** 2608–9 (2011).

29. Esserman, L. The I-SPY Master Trials : A Model for Accelerating the Pace of Getting the Right Drugs to the Right Patients. at <find url if needed>

30. Esserman, L. J. A Model for Accelerating Identification and Regulatory Approval of Effective Investigational Agents. *Cureus* **4,** (2012).

31. Precision Medicine Initiative (PMI) Working Group. *The precision medicine initiative cohort program – building a research foundation for 21st century medicine*. *Precision Medicine Initiative (PMI) Working Group Report to the Advisory Committee to the Director, NIH* **Sept 17,** (2015).