Blockchains and electronic health records

Ben Yuan, Wendy Lin, and Colin McDonnell

Table of Contents

1 Introduction

1.1 The Blockchain

Since the creation of Bitcoin in 2009 and its continued relatively wide adoption, considerable interest has developed in the consensus mechanisms underpinning the cryptocurrency. Bitcoin's success stems in large part from the robustness of these mechanisms, which provide a means to achieve decentralized, trustless currency issuance, transaction validation, and transaction settlement - removing the implicit centralization requirement for these tasks.

Bitcoin's consensus model centers around the "blockchain", a data structure and set of algorithms designed specifically for achieving Byzantine fault-tolerant consensus around the state of a global transaction ledger. The key principles of the blockchain data structure as used in Bitcoin may be summarized thus:

- Transactions are bundled into blocks. For a block to be valid, all its constituent transactions must also be valid according to the global 'start' state.
- Blocks have parent blocks. The global 'start' state corresponding to any given block may be reconstructed by replaying all of its ancestor blocks in normal chronological order. For a block to be valid, all of its parent blocks must also be valid.
- Blocks also carry certain data used to prove that a certain amount of computation power was expended in its creation. For a block to be valid, its proof-of-work data must be valid according to the scheme being used.
- Consensus among "correct" participants requires that they eventually all converge on the same history. Bitcoin participants take as the 'most recent' block some valid block for which the total estimated work of the block and its ancestors is greatest. As long as blocks are always being added by "correct" participants to the block they believe is "most recent", and the "correct" participants outnumber the "incorrect" participants in terms of computing power, the "correct" participants do tend to converge to the same global state.
- "Correct" participants are given an incentive to continue creating blocks. In Bitcoin, this incentive takes the form of currency issuance; a successful block creator can issue itself currency according to agreed-upon rules.

The result is a durable transaction ledger, secured by consensus among multiple parties, that does not obligatorily rely on trust in any single party to function; no single party can alter or remove any portion of the "canonical" transaction record without performing a very large amount of work. A transaction ledger that is globally accessible, easy to verify, and difficult to modify provides evident benefits when used as the underpinning of a digital currency: it allows anyone to verify that a given unit of currency being spent has not already been spent in the past, and prevents past transactions from being arbitrarily retracted. However, a tamper-resistant ledger of this form can be used for purposes other than currency, wherever a requirement for censorship-resistant, repudiation-resistant data publication exists. We examine the relative

potential applicability of this particular aspect of blockchain technology, in comparison to alternative solutions, with respect to electronic medical records.

1.2 Current State of Electronic Medical Records in the US

Electronic medical records (EMRs) today are fragmented across myriad hospitals, private practices, labs, pharmacies, and, increasingly, private companies collecting data from wearable devices. This fragmentation will only increase as more frequent job changes, greater mobility, and the rise of specialty care drive more changes in insurance plans, greater reliance on multiple healthcare providers, and the need to access healthcare services from a higher number of outlets.

It is well acknowledged since the 1990s that reducing this fragmentation by increasing the ease with which EMRs are accessed and transferred across organizations will improve our healthcare system. However, attempts to implement solutions have run into barriers, as addressed in Vest and Gamm's paper titled "Health information exchange: persistent challenges and new strategies"[1] [7], including:

- healthcare providers' hesitation to share what they perceive to be proprietary data
- patient concerns about security and privacy
- lack of strong political will from regulators
- historically costly technological solutions, whose costs often fall to healthcare providers but whose benefits often accrue to patients, payers (e.g. insurance companies), and the healthcare system as a whole [5]

Vest and Gamm's model reduces the space of healthcare stakeholders to payers, providers, patients, and governmental entities. We found it difficult to hypothesize about how a given stakeholder would react to a proposed change or incentive structure without modeling the landscape in greater detail. Below is a diagram representing the heterogeneity of the stakeholder categories in Vest and Gamm's model, as well as the space of interactions between these parties and, where relevant, the middlemen that mediate such interaction.

It's within this complex system of incentives currently in flux that we will attempt to derive an optimal data structure for EMRs that will address key problems of the status quo. In evaluating these structures, we will assess both costs and benefits to the primary stakeholders of the healthcare system -- patients, (healthcare) providers, payers (including Medicare / Medicaid and private insurance), and regulators. For costs, we will assess both financial costs as well as the mental cost associated with behavior change relative to the status quo. For benefits, we have identified the key goals each stakeholder has for an EMR system and will assess how well different data structures address these goals.

[1] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2995716/

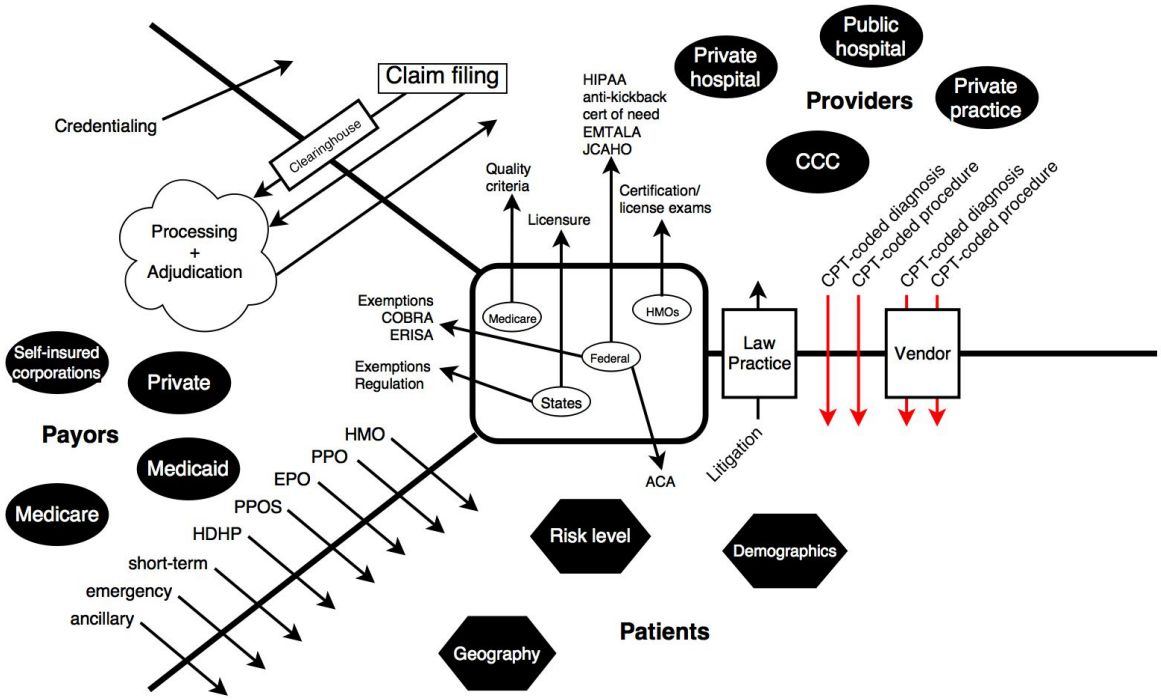

Figure 1: A map of stakeholders and interactions in the healthcare ecosystem.

2 Data Structure Enumeration

2.1 Relevant Qualities of Data Structures

We try to define the space of healthcare data management solutions by identifying the relevant properties of interest and considering each possible combination of these properties.

When trying to enumerate important or relevant properties of data management solutions for a complex industry like healthcare it's important to understand the specific needs of the stakeholders involved. In health care, three factors are particularly important: data lineage/integrity, data security, and interoperability. We consider each of these three briefly.

1. Ensuring *data lineage* and *data integrity* is a big one. If you're handling research data or test results whose integrity can directly influence people's health, then you want to know that that data hasn't been tampered with since creation. As such, it's in the public interest for anyone to verify that research data is secured by a solid chain of custody from birth. Similarly, if you're handling medical records you want to know that those records were generated by a credible source. Unfortunately "a doctor in another state" may or may not fall into that category.
2. Ensuring *data security* is another obvious concern -- we need to ensure that records can't be retrieved by people who are unauthorized to view them. On the other side of the

coin, all parties who are authorized to view them should be able to view them in a hassle free way, including emergency medical personnel, the PCP of a patient, the patient herself, and anyone else the patient wishes to bring into the loop. Not only that, but the process of adding or subtracting access permissions should be painless and instantaneous.

3. Ensuring *data mobility, integration, and interoperability* is the final piece of the puzzle. What's the point in having legal access to a medical record if you have to fly across the country to exercise that right? Medical records should be capable of movement between providers with a minimal amount of friction. The current process is a far cry from this -- it frequently involves phone calls, paperwork, and FedEx. Even digital records are frequently incompatible between the electronics systems of different providers.

But how do we convert these specifications into a small set of features or properties that we can use to enumerate categories of data structures? We settled on three properties that collectively encompass the above specifications: change tracking, decentralization, and proof-of-work. Change tracking refers to the ability to see the state of the system and its contained data at an arbitrary point in the past, as opposed to merely the most recent version of the data. This of course achieves the specification of data lineage and integrity mentioned above. Decentralization refers to a distribution of control of the actual servers and devices storing the data among many discrete autonomous entities, which helps guarantee data security and requires some degree of interoperability. Proof of work refers to the puzzle-solving executed by Bitcoin miners that enables them to collaboratively mine blocks and achieve consensus surrounding a version of history. This property also helps maintain data security and integrity while necessarily involving an agreed-upon interoperable block structure.

2.2 Candidate Data Structures

We look at all possible permutations of these three properties below and describe a feasible and reasonable data structure that falls within that category. In general, we assume a well-implemented system based on sound technology and only critique each proposal based on its inherent properties and propensity to demonstrate various vices and virtues.

The most significant bit represents proof-of-work. The middle bit represents decentralization. The least significant bit represents change tracking.

000	a traditional, centralized database administered by one entity, likely a provider or a governmental organization
001	a traditional, centralized database with change tracking
010	a distributed peer-to-peer encrypted database, perhaps employing distributed hash tables with many redundant copies of data

011 a distributed version control system such as Git

100 not considered (as proof-of-work does not make much sense in isolation)

101 a private blockchain with all nodes controlled by a single entity with proof-of-work required to implement a change

110 a distributed database without change tracking, shared among many stakeholders, and requiring proof of work to implement a change

111 a blockchain with proof-of-work, of which we consider two major variants:

- federated blockchains with a shared but controlled ownership of mining nodes among a set of shareholders, including the government, providers, payers, and vendors
- pure, public blockchains such as the Bitcoin blockchain, with no centralized or federated control on mining power

We evaluate all of these options in turn for viability and situational aptitude.

3 Data Structure Analysis

3.1 Scorecard

We created a scorecard -- displayed in Table A -- for evaluating specific proposals regarding the administration of EHR data management. It is derived from the model described in Vest and Gamm's paper as well as an analysis of the most important issues to each stakeholder. We use this scorecard to analyze each of the solutions enumerated above.

		Patients	Providers	Payers	Regulators
Costs	Mental / Behavioral				
	Financial				
Benefits		Patient has access to own complete record?	Easy and fast to access records?	Costs are lower, i.e. fewer needless procedures?	Ease of enforcing regulation, i.e. care is auditable?
		Patient controls privacy of record?	Easy and fast to modify records?	Costs are lower, i.e. disease prevention, compliance?	Quality of care is improved?

	Security of patient record assured?	Quality of care is improved, i.e. fewer preventable mistakes?	Makes fraud more difficult?	Easy to monitor public health, epidemics, health trends?
	Patients can discern insights from their data, i.e. it is sensible?	Increased safety from malpractice cases?	Does not jeopardize customers or make it easier to switch?	
		Does not jeopardize customers or make it easier to switch?		

Table A: A scorecard used to evaluate EHR proposals.

3.2 Analysis of Candidate Solutions

3.2.1 Data Structures without Change Tracking (000, 010, 100, 110)

Ensuring accurate and complete provenance of records is an important goal in healthcare. When a patient receives a copy of his or her own health record, or when a doctor or payer receives a health record from a distant office, the recipient would like to ensure that the record is complete and correct. Patients and providers want assurance that no important medical history facts have been unknowingly altered or wrongly introduced; payers need accurate information on procedures and treatments performed.

Any system that manages electronic health records should provide some mechanism by which changes to a given record may be tracked and verified, at least by anyone with the capability to read the record. An auditor - who may be a patient, or a doctor, or a payer, or a regulator - should be able to determine when a particular value for a particular attribute was created, as well as what values were present before, subject to any useful and reasonable privacy restrictions a patient may wish to place on this information.

With a good change tracking mechanism, data recipients can be assured that the data they are receiving is the product of a sensible record-keeping process - and if the observed change history is in great conflict with the previously observed history, or otherwise indicates behavior outside reasonable expectations, then the data recipient is justified in demanding an explanation. Any system without robust change tracking cannot provide this crucial property, since it becomes much more difficult for a data recipient to ascertain the legitimacy of any data received - especially when that data does not conform to expectations.

We thus do not consider systems without auditable change tracking in our discussion of electronic medical records.

3.2.2 Traditional, Centralized Data Structure (001)

A fully centralized model assumes that clients rely on the word and work of a *single authority* for the world state. This single authority performs all authentication, authorization, data processing, and data storage. An example is the US Social Security Administration, which is the single authority on one of the key identifiers and means to access government benefits.

In healthcare, the most credible and powerful central authority is probably the Centers for Medicare and Medicaid (CMS). As the largest payer, the CMS sets the ground rules for how and which health procedures get reimbursed, which then reverberates across the industry.

The trouble is that healthcare is not a contained system like that of social security and many parties must frequently read and write to the EMR database. Thus, the natural evolution of using a traditional, centralized data structure has lead to today's world where many entities maintain their own world states based on the limited information they have from the data they have access to. There is no common world state across these organizations, and patients and providers must do the legwork to reconcile and unify these world states in instances of patient mobility or collaborative delivery of care.

Even if government musters up the substantial political force of will to centralize all data under the CMS, this change will likely require a multi-billion dollar government project, using the much simpler Healthcare.gov's $500+ million cost as a benchmark. In this realization of a fully centralized system, regulators will have to bear all of the financial and mental cost, requiring relatively little behavior change or financial contribution from other stakeholders (though one can argue that ultimately the patients as tax-paying citizens bear the financial burden). A benevolent, enlightened, and sophisticated government would then be able to help each stakeholder realize his objectives. More realistically, centralized databases lead us to where we are today, whereby we generate no additional mental or financial costs, but must accept its failure to address all of the aforementioned goals of an optimal EMR system.

To protect the sensitivity of this data across multiple parties, the government introduced HIPAA, the Health Insurance Portability and Accountability Act. Any organization that deals with protected health information must ensure that all required physical, network, and process security measures are in place and must abide by privacy rules that aim to involve patient sign-off and sharing the bare minimum of data to achieve goals. Without the ability to achieve these goals through other means, HIPAA uses severe civil and criminal fines to penalize bad actors after the fact.

A completed scorecard can be found in Appendix B.

3.2.3 Distributed Database with Change Tracking (011)

We now consider a distributed data management solution that tracks all changes to the EHR over time. We are assuming a “best-case scenario” with respect to the technology. It should be entirely possible to implement a secure distributed versioning system that allows fine-grained permissioning of both read and write access. The record should be secure and private. It should ideally be possible for various organizations to receive statistics relating to medical records without accessing the raw data itself.

One can imagine a distributed system of servers that track a large amount of data over time, and is capable of rolling back to any previous state of the system. Changes to the record are represented as sets of additions and subtractions from the previous state of the system. The system can be made tamper resistant with chained hash pointers and signature dependencies to prevent any clandestine modification of the record, in a similar manner to the method by which some software version control systems ensure change lineage and integrity.

The parallel to version control systems may also provide a useful metaphor for understanding the user experience of the providers. One can have a “master branch” of the patient’s record that the provider “checks out” when providing care to the patient. In the process of diagnosing and treating the patient, the provider can augment its local branch of the record, then “merge” the changes into the master branch when a conclusion or diagnosis has been reached. This lets the provider run appropriate follow-up tests before publishing potentially misleading or mistaken test results to a patient’s record. There can be policies in place regarding the frequency of merging record changes; for instance, it might be necessary to publish changes before prescribing a pharmaceutical or after certain types of tests.

One important question affects the performance of any proposal involving a decentralized version-control approach to EHR management: who physically controls the data? Some options include the state government, the federal government, the patients themselves, a non-governmental trustless network of servers (similar to the Bitcoin network, minus proof of work), providers (perhaps that meet a given size according to some metric), insurance companies, or any combination of the above. Any of these options may be viable. Historically, the burden for hosting similar data has fallen to medical providers and to the government. An incentive scheme that somehow motivated insurance companies or patients to put forth time and monetary resources to host data would vastly improve the chances of a solution taking off.

If the requirement for always-online access to the most recent data by any authorized party may be dropped, then even simpler solutions for data storage may be possible, while retaining the system’s tamper resistance and auditability. One can imagine the patient carrying a write-only memory device, perhaps in the guise of an insurance card or similar, to which signed changes to the patient’s record are written at each provider visit. Taking the data “offline” in this way does seem to hinder payer and regulator access, as the “canonical copy” would exist on a device that

spends most of its time disconnected from the world. One should of course ensure that the storage device used is of a commonly readable type and uses an easily readable data encoding, so that patients retain the ability to easily read their own records.

A completed scorecard can be found in Appendix C.

3.2.4 Private Blockchains (101)

Private blockchains are a bad idea in general. They eliminate the benefits of a decentralized network capable of trustless transactions and robust consensus. A blockchain whose entire mining pool is controlled by a single entity degenerates to a traditional centralized system with a bit of cryptographic auditability sprinkled on top. What's more, this auditability can be achieved through other means besides the mechanisms used in blockchains. To quote Vitalik Buterin, "there is no reason to believe that the optimal format of such authentication provision should consist of a series of hash-linked data packets containing Merkle tree roots; generalized zero knowledge proof technology provides a much broader array of exciting possibilities about the kinds of cryptographic assurances that applications can provide their users." [1]

However, this proposal may not be entirely without merit. There are vanilla implementations of blockchains that are presumably easier for a hospital to get running than implementing secure zero-knowledge proof protocol. Additionally, if the data format of these private blockchains is somehow standardized early in the process, then the adoption of private blockchains by the administering stakeholder may improve data portability and interoperability. However, this standard would likely have to be mandated by a state or local government, which would be better off simply mandating a standard data format for conventional records.

A completed scorecard can be found in Appendix D.

3.2.5 Partially Open Blockchains (111)

A federated blockchain consist of several parties that jointly create the world state and attempts to replace Bitcoin's distributed network of voluntary miners with proprietary computers belonging to approved users that process transactions. Operations on this world state may affect multiple parties simultaneously, and a federated blockchain would force the network to share responsibility over each other's databases.

In the case of healthcare, such a group would likely include regulators, providers, and payers. Federations are likely to be organized by systems of care, most likely identified by geography, such as community or state. Patients are assumed to stay within these systems of care that cross organizations. Most likely, a federated blockchain will be applied on top of an existing health information exchange community as a way to further reduce costs and help the community reach financial sustainability.

Current sharing of data across these systems of care exist with the combination of centralized data structures within each individual organization and HIPAA-compliant data transmission. With the blockchain, organizations can come together and jointly create a public (to the federation) truth that each organization can modify with requisite proof-of-work. Because miners are distributed across organizations, each organization checks every other organization's database modifications.

Such a system is more secure than the status quo, as organizations are able to aggregate computing power to secure the blockchain. Also, data is redundant across organizations, avoiding single points of failure. Fewer parties means it's quicker to modify the system's code and revert actions; nodes are well-connected and manual intervention can quickly fix a lot of faults and enable faster confirmation times. Actions on the chain are also cheaper, requiring less "wasted" proof-of-work. It is simple to rewrite the rules, especially regarding read permissions, which is a useful property in the context of EMRs. However, many of these goals can also be achieved without proof-of-work, which is an expensive form of security.

Additionally, removing patients from the picture removes one of the key beneficiaries of a more liquid EMR system and ends up leaving mostly antagonistic parties at the table. Given a solution like this requires system-wide buy-in, it's unlikely that without patients at the table pushing for this solution, other parties will reach consensus in adopting a federated blockchain.

A completed scorecard can be found in Appendix E.

3.2.6 Public, Open Blockchains (111)

Considering the potential issues with deploying and ensuring acceptance of more closed systems, we turn our attention to the possibility of building an electronic medical record system using a public, open blockchain as a trust anchor. It is of course not desired to place medical records directly on such a blockchain, as any information committed to an open blockchain is naturally globally visible; this property would immediately introduce serious privacy concerns for the patients described by the records being kept. Additionally, if a public blockchain like Bitcoin is to be used, the restrictions on data storage for the host blockchain must be respected; Bitcoin itself only permits 80 bytes of user-chosen data to be added to the blockchain in a given transaction [2], so full medical records could not easily be stored directly even if privacy were a non-issue.

However, if we permit a secondary data storage mechanism, e.g. a distributed hash table with open participation and custom access control mechanisms, then we may have the tools needed to build a sensible privacy-respecting electronic medical record system. The Enigma protocol [9] describes a privacy-respecting programmable substrate, using secret sharing and secure multi-party computation to achieve Turing-complete computation over private data, and using an open blockchain to perform identity management, access control, and auditing. Enigma is constructed such that the off-chain network permits and incentivizes open participation, allowing

anyone to participate in keeping the Enigma system running, and such that the public blockchain stores audit records of off-chain activity, allowing anyone to verify that the off-chain network is operating correctly without being able to discover private data.

Since Enigma is highly programmable, we can construct our EMR system essentially however we wish. It need not follow particularly complicated rules; as an example, by default, a given patient's medical record can be readable only by the patient, and only upon request by a provider and/or payer (signalled by e.g. a smartphone application) does the record become accessible and writable as necessary by the respective parties. Such an architecture, instantiated correctly, gives the patient access and control over their own complete medical record without imposing the singular burden of storing or transmitting it, while allowing all parties to participate in and verify correct operation of the network.

If patients are willing in practice to disclose the necessary data, then being able to compute over complete medical records brings advantages to providers, payers, and regulators. Providers and payers can assess the medical need of any given procedure in the context of a patient's entire medical history, potentially enabling otherwise unavailable insights and potentially reducing the incidence of medically unnecessary work. Regulators, without needing to handle the actual records, are still able to compute trends over medical records in aggregate, potentially giving them the tools needed to discover public health trends essentially as they happen.

The largest issue with any such system, supposing the pieces work as advertised, is the issue of key management. Identity in any Enigma-based system is tied to private keys; should these keys be lost or otherwise compromised, control of the corresponding identity is lost. This can be especially problematic if a patient loses control of the key owning a corresponding medical record, as direct control of the medical record is lost. This situation is not entirely impossible to recover from; for instance, the key itself may be distributed by a secret-sharing mechanism to multiple partially trustworthy parties, and a key recovery mechanism derived from that, or the entire medical record may be retrieved (if disclosed in such form in the recent past) from the patient's last visited provider and reissued under a new key. Any system using private keys as identities must consider the key recovery issue, but it is especially important in the case of medical record management - as loss of an entire medical record would be problematic for the corresponding patient, in the context of future and ongoing medical care.

Because in this architecture the patient has final control over what data gets disclosed to whom, providers still face the prospect of having to disclose data they may perceive as proprietary to parties that may be potential competitors. As with any personal health record system, providers must be convinced that the net benefits of contributing comprehensive information to a patient's record outweigh the net benefits of concealing information perceived proprietary.

A completed scorecard can be found in Appendix F.

4 Recommendation, Observations, and Conclusion

Given our analysis, we believe blockchain-based technology is a viable choice for EMR management. Notably, the lack of any single entity that everyone trusts to run a centralized system indicates that a decentralized one might be favorable, and the minimization of required trust relationships seems like a good fit for such an environment. The weaknesses that blockchain technology currently presents, such as lack of high-volume processing and difficulty handling private data, can increasingly be addressed with advancements like Enigma and Bitcoin-NG, and we believe the high amount of developer attention on the blockchain will continue to resolve other weaknesses that emerge.

However, while blockchains are *a* good choice for this application, consideration must still be given to alternatives that achieve the same goal of enabling complete, auditable patient-owned personal health records. Arguably, alternatives built on decentralized change-tracking databases can be comparably effective at enabling completeness, auditability, and data control if well-engineered. To settle the question decisively, it may become necessary to instantiate more comprehensive system designs and conduct more detailed cost-benefit analyses with the more complete designs in hand.

The bigger question is how to get the healthcare system to adopt any new EMR management system, given the complex and often competing interests involved in the ecosystem. The benefits to patients of having ownership of a comprehensive health record are appealing, but patients are historically also the most disempowered of the stakeholder groups. Achieving a successful implementation of patient-controlled EMRs requires a compelling enough message to mobilize patients and patient advocacy groups to jumpstart the initiative. Previous instances have shown that regulators rarely have urgency in initiatives without sufficient citizen attention, another reason for mobilizing patients [3].

Then, given a change like this requires system-wide support, it makes sense to implement this first in a contained health community, ideally one where

- there are many small providers who are on fragmented EMR systems but don't have the ability to provide all healthcare services alone. This solution can allow them to become part of a physically distributed, full service provider. This argument becomes more compelling as healthcare providers continue to specialize and fragment, as indicated by the rise of minute clinics and Uber-style doctors on demand. Additional key arguments include cost savings versus the current method of exchanging information (fax, security, high-cost HIPAA-compliant technology solutions) and the inevitable march of healthcare payments towards a pay-for-performance basis, which requires coordinated care
- there is one or two payers, to reduce the number of parties from whom we must get buy-in
- there is a forward-looking regulator who will provide support if the solution gains traction

Some good candidates for initial communities are ones that have established healthcare information exchanges, such as the Utah Health Information Network.

For analysis regarding other potential applications of the blockchain, we suggest a similar initial approach of mapping out the stakeholders in the space and identifying their key behavioral / financial costs and motivators. A lack of trust across participating parties in each other or one central entity is a good initial indication that a blockchain may be a workable solution. One should consider whether the constraints of the application may permit a distributed database solution without coupling to a proof-of-work requirement. Then, given that any new solution of these forms often requires system-wide change, one should assess how different parties in the system are impacted and which parties are likely to provide the primary thrust for adoption.

5 Bibliography

[1] Buterin, Vitalik. "On Public and Private Blockchains." *Ethereum Blog*. N.p., 07 Aug. 2015. Web. 16 Dec. 2015.

[2] "Change the Default Maximum OP_RETURN Size to 80 Bytes by Flavien · Pull Request #5286 · Bitcoin/bitcoin." *GitHub*. N.p., n.d. Web. 16 Dec. 2015.

[3] Cordina, Jenny, Rohit Kumar, and Christa Moss. "Debunking Common Myths about Healthcare Consumerism." McKinsey and Co., Dec. 2015. Web. 16 Dec. 2015.

[4] "Health Care Fraud and Abuse." (2009): n. pag. Center for Medicare and Medicaid Services. Web.

[5] "Health Information Privacy." *HHS.gov*. US Dept of Health and Human Services, n.d. Web. 16 Dec. 2015.

[6] Kessler, Glenn. "How Much Did HealthCare.gov Cost?" *Washington Post*. The Washington Post, 24 Oct. 2013. Web. 16 Dec. 2015.

[7] Vest, J. R., and L. D. Gamm. "Health Information Exchange: Persistent Challenges and New Strategies." *Journal of the American Medical Informatics Association* 17.3 (2010): 288-94. Web.

[8] "What Is HIE (Health Information Exchange)?" HealthIT.gov, n.d. Web.

[9] Zyskind, Guy, Oz Nathan, and Alex Pentland. "Enigma: Decentralized Computation Platform with Guaranteed Privacy." (n.d.): n. pag. Web.

6 Appendices

6.1 Appendix A Blank Scorecard

		Patients	Providers	Payers	Regulators
Costs	Mental / Behavioral				
	Financial				
Benefits		Patient has access to own complete record?	Easy and fast to access records?	Costs are lower, i.e. fewer needless procedures?	Ease of enforcing regulation, i.e. care is auditable?
		Patient controls privacy of record?	Easy and fast to modify records?	Costs are lower, i.e. disease prevention, compliance?	Quality of care is improved?
		Security of patient record assured?	Quality of care is improved, i.e. fewer preventable mistakes?	Makes fraud more difficult?	Easy to monitor public health, epidemics, health trends?
		Patients can discern insights from their data, i.e. it is sensible?	Increased safety from malpractice cases?	Does not jeopardize customers or make it easier to switch?	
			Does not jeopardize customers or make it easier to switch?		

6.2 Appendix B Scorecard for traditional centralized database

		Patients	Providers	Payers	Regulators
Cost	**Mental / Behavioral**	**None**	**None**	**None**	**High** - must take on burden of setting up and maintaining centralized database
	Financial	**None**, potentially additional taxes	**None**	**None**	**High** - potentially multi-billion dollar initiative
Benefits		Patient has access to own complete record?	Easy and fast to access records?	Costs are lower, i.e. fewer needless procedures?	Ease of enforcing regulation, i.e. care is auditable?
		High - centralized database creates one source of complete records, provided government grants patients access	**High** - centralized database creates one source of complete records	**High** - centralized database creates one source of complete records, so can decrease likelihood of redundancy or actions based off incomplete data	**High** - controls database, so can audit at any point
		Patient controls privacy of record?	Easy and fast to modify records?	Costs are lower, i.e. disease prevention, compliance?	Quality of care is improved?
		Low to none - unless there are policy changes, the regulator controls the record	**Medium** - depends on how regulator implements rules of database	**Low** - no additional ability to influence patient compliance or prevention unless this is an additional feature built into database	**High** - centralized database creates one source of complete records, so easier to coordinate care and data across healthcare-related parties
		Security of patient record assured?	Quality of care is improved, i.e. fewer preventable mistakes?	Makes fraud more difficult?	Easy to monitor public health, epidemics, health trends?
		Low - one database to hack into with all EMR data? yes, please!	**High** - centralized database creates one source of complete records, so can decrease likelihood of redundancy or actions based off incomplete data	**Low** - requires central authority to police against fraud, thus requiring a lot of resources	**High** - controls database, so can audit or run analyses at any point
		Patients can discern insights from their data, i.e. it is sensible?	Increased safety from malpractice cases?	Does not jeopardize customers or make it easier to switch?	
		Unclear - depends on if regulator builds tools for others to develop analysis on top of data	**Medium** - likely to help prevent *unfounded* malpractice cases	**Low** - data portability makes it easier for patients to switch to another provider.. Can limit this through agreed-upon rules.	
			Does not jeopardize customers or make it easier to switch?		
			Low - data portability does make it easier for patients to have lower cost to switch to another party. Can limit this through agreed-upon rules		

6.3 Appendix C Scorecard for distributed database with history tracking

		Patients	Providers	Payers	Regulators
Costs	Mental / Behavioral	None	High, providers must digitize all records in standard format and cede ownership	High, payers must digitize all records in standard format and cede ownership	Low, makes their lives much easier
	Financial	None	Medium short term: there will be transitioning costs Low long term: easier to access records, distributed IT maintenance costs	Medium short term: there will be transitioning costs Low long term: easier to access records, distributed IT maintenance costs	Low, potentially cost-saving
Benefits		Patient has access to own complete record?	Easy and fast to access records?	Costs are lower, i.e. fewer needless procedures?	Ease of enforcing regulation, i.e. care is auditable?
		High, record is coherent and unified	High, no more chasing down records	High, previous test results etc are all available to providers	High
		Patient controls privacy of record?	Easy and fast to modify records?	Costs are lower, i.e. disease prevention, compliance?	Quality of care is improved?
		High, assuming good implementation	High, assuming good implementation	Medium	High
		Security of patient record assured?	Quality of care is improved, i.e. fewer preventable mistakes?	Makes fraud more difficult?	Easy to monitor public health, epidemics, health trends?
		High, assuming good implementation	High, this should follow naturally from easy access to medical records	Medium, there is more oversight to changes to medical record but many fraud types are still possible	High
		Patients can discern insights from their data, i.e. it is sensible?	Increased safety from malpractice cases?	Does not jeopardize customers or make it easier to switch?	
		High, this should follow naturally from having access to a complete record	High, full auditability	Low, this would lower barriers to switching insurance providers	
			Does not jeopardize customers or make it easier to switch?		
			Medium, this system would likely make it easier for patients to change providers		

6.4 Appendix D Scorecard for private blockchains

		Patients	Providers	Payers	Regulators
Costs	**Mental / Behavioral**	None	High (assuming providers administer a blockchain)	High (assuming payers administer a blockchain)	Low (will have to retrain to accommodate new data management system)
	Financial	None	Medium (requires initial costs to set up, but private blockchains can reach near optimal efficiency...no expensive PoW)	Medium (requires initial costs to set up, but private blockchains can reach near optimal efficiency...no expensive PoW)	Low (will have to retrain to accommodate new data management system)
Benefits		Patient has access to own complete record?	Easy and fast to access records?	Costs are lower, i.e. fewer needless procedures?	Ease of enforcing regulation, i.e. care is auditable?
		No change from status quo: the record is still fragmented across organizations and may still be difficult to synthesize.	Depends on degree and quality of communication between administering entities. Likely low.	Low, there is no guaranteed improvement in care.	Medium, blockchains likely enforce a minimum quality standard on data management, and have built in change tracking
		Patient controls privacy of record?	Easy and fast to modify records?	Costs are lower, i.e. disease prevention, compliance?	Quality of care is improved?
		Medium: this should be possible assuming well-implemented private blockchains, but patient ust trust administering entities	High is blockchain run locally. Low if blockchain run by govt or payer.	Low, there is no guaranteed improvement in care.	Low, there is no guaranteed improvement in care
		Security of patient record assured?	Quality of care is improved, i.e. fewer preventable mistakes?	Makes fraud more difficult?	Easy to monitor public health, epidemics, health trends?
		Yes, the form factor of a blockchain necessitates better security than many conventional systems	Low, records are not much more accessible than they are now.	Low, institutions have complete control over their internal reporting system.	Low, only to the extent that switching to blockchains improves data portability
		Patients can discern insights from their data, i.e. it is sensible?	Increased safety from malpractice cases?	Does not jeopardize customers or make it easier to switch?	
		No change from status quo.	High, the form factor of a blockchain necessitated greater auditability than most current systems	Low, only to the extent that switching to blockchains improves data portability	
			Does not jeopardize customers or make it easier to switch?		
			Low, these blockchains would be managed by individual providers, just as EHR databases are now.		

6.5 Appendix E Scorecard for partially-open blockchains

		Patients	Providers	Payers	Regulators
Costs	**Mental / Behavioral**	**None**	**High** - change current data transfer behaviors and create roles for miners	**High** - change current data transfer behaviors and create roles for miners	**High** - change current data transfer behaviors and create roles for miners
	Financial	**None**	**Medium** - cost of miners and mining, offset by cheap data transfer	**Medium** - cost of miners and mining	**Medium** - cost of miners and mining
Benefits		Patient has access to own complete record?	Easy and fast to access records?	Costs are lower, i.e. fewer needless procedures?	Ease of enforcing regulation, i.e. care is auditable?
		Medium to High - as long as patient stays within federation and federation allows	**High** - as long as the record is within the federation, this should improve dramatically	**High** - as long as patient is within federation, coordinated care should be easier and reduce needless procedures	**Medium** - must audit all federations, but the act of auditing is much easier as organizations are clustered
		Patient controls privacy of record?	Easy and fast to modify records?	Costs are lower, i.e. disease prevention, compliance?	Quality of care is improved?
		Low - no changes in privacy controls within federation	**Medium** - within the federation, this is true bureaucratically but difficult computationally, as it requires proof-of-work	**Low** - no additional ability to influence patient compliance or prevention	**High** - within the federation, it will be much easier to coordinate care and data across organizations
		Security of patient record assured?	Quality of care is improved, i.e. fewer preventable mistakes?	Makes fraud more difficult?	Easy to monitor public health, epidemics, health trends?
		Medium - any attempts by members to take improper action can be identified by rest of group and corrected	**High** - as long as the patient remains within the federation, it will be much easier to coordinate care and data across organizations	**Medium** - any attempt by members of the federation to commit fraud can be quickly identified by rest of federation and corrected in absence of collusion	**Medium** - instead of tracking across even more fragmented database, only need to track down the data across different federations
		Patients can discern insights from their data, i.e. it is sensible?	Increased safety from malpractice cases?	Does not jeopardize customers or make it easier to switch?	
		Low - no changes, though data organized across a federation system might help patient to see holistic picture of health upon accessing own data	**Medium** - likely to help prevent *unfounded* malpractice cases	**Low** - data portability does make it easier for patients to have lower cost to switch to another party in the federation. Can limit this through federation agreed-upon rules	
			Does not jeopardize customers or make it easier to switch?		
			Low - data portability does make it easier for patients to switch to another party in the federation. Can limit this through federation agreed-upon rules		

6.6 Appendix F Scorecard for open, public blockchains

We first describe this hypothetical solution in more detail. We are considering an EMR management system using the Enigma system as a backend. This system is implemented as such:

- Enigma uses a public blockchain to store proofs of correct execution, and an off-chain network for accomplishing secure distributed data storage and multiparty computation.
- A patient maintains “ownership” authority over his or her own medical record.
- In a simpler system (treating this blockchain-based system as a “database node”), the patient can choose to disclose certain elements of their record to a provider or payer on demand, and the updated record ends up in the provider / payer system.
- In a more comprehensive system (implementing the entire system in Enigma), the patient may also be able to choose to revoke visibility on certain elements of their data. *This is hard to ensure without regulatory support and careful auditing, and may not actually be practical.*
- Enigma charges fees for computation and storage.

		Patients	**Providers**	**Payers**	**Regulators**
Costs	**Mental / Behavioral**	**Medium / High** - Have to do key management somehow to maintain access to records. Recovery from key compromise is hard.	**Medium / High** - Recovery from key compromise is hard. Key theft is still a vector for dangerous data compromise (though detection may be easier).	**Medium / High** - Key theft is still a vector for dangerous data compromise, so best practices must be followed.	**Medium / High** - Key theft is still a vector for dangerous data compromise, so best practices must be followed.
	Financial	**Medium** - If patients own their data, they may be liable for the storage costs. Additionally, increased costs for provider and payers may be passed down through higher insurance and medical bills.	**Medium** - Providers may be liable for costs related to access of patient data. Some costs may be recouped by participation in the Enigma network on the compute side.	**Medium** - Payers may be liable for costs related to access of patient data. Some costs may be recouped by participation in the Enigma network on the compute side.	**Medium** - Regulators may be liable for costs related to computation over patient data. Some costs may be recouped by participation in the Enigma network on the compute side.
Benefits		Patient has access to own complete record?	Easy and fast to access records?	Costs are lower, i.e. fewer needless procedures?	Ease of enforcing regulation, i.e. care is auditable?
		High - Ensurable by design, with appropriate client support.	**Not known**; dependent on the performance of the Enigma off-chain network. Historically, computation on encrypted data has been slow in practice.	**High** - As long as patient agrees to disclose relevant records to providers, and providers are willing to trust records.	**High** - As long as the necessary process can be encoded in an Enigma smart contract. Trivially doable in the simple model if data revocation is not a concern.

Benefits	Patient controls privacy of record?	Easy and fast to modify records?	Costs are lower, i.e. disease prevention, compliance?	Quality of care is improved?
	High - Enigma enables fine-grained access and disclosure control to be implemented in a smart contract.	**Not known**; dependent on the performance of the Enigma off-chain network. Historically, computation on encrypted data has been slow in practice.	**High** - As long as patient agrees to disclose relevant records to providers, and providers are willing to trust records.	**Potentially high** - to the degree to which complete medical records are actually provided and contribute to quality of care.
	Security of patient record assured?	Quality of care is improved, i.e. fewer preventable mistakes?	Makes fraud more difficult?	Easy to monitor public health, epidemics, health trends?
	High - By design if the EMR system is constructed properly.	**Possibly high** - as long as providers are willing to provide useful data in the shared medical record, providers can consider care in context.	**High** - With appropriate system design, veracity of claims can be audited against medical history, if patient agrees to disclose the necessary data.	**Potentially high** - Dependent on whether patients are willing to share this data in practice. Enigma contract language allows aggregate computation without giving regulators access to raw data, but this must be understood by patients.
	Patients can discern insights from their data, i.e. it is sensible?	Increased safety from malpractice cases?	Does not jeopardize customers or make it easier to switch?	
	Possibly high - Dependent on any client software to present patient record in a sensible way.	**Not known**. Allows proving that certain data was used and that certain procedures were followed, but does not eliminate human factors entirely.	High, it is likely that an open universal data system will decrease barriers to switching insurance companies	
		Does not jeopardize customers or make it easier to switch?		
		High, it is likely that an open universal data system will decrease barriers to switching providers		