BLOCKCHAIN™ IN HEALTHCARE TODAY

A Pragmatic Solution to a Major Interoperability Problem: Using Blockchain for the Nationwide Patient Index

Michael L. Gagnon[1] and Grant Stephen[2]

Affiliations: [1]Senior executive and Executive Director of HealtHIE Nevada, Qualis HealthInsight, [2]CEO of bPrescient

Keywords: Blockchain; Master Patient Index; MPI; Distributed Ledger; Patient Identity; HIE; Health Information Exchange.

Corresponding Author: Grant Stephen, gstephen@bprescient.com

Section: Cases/Pilots/Methodologies

Associating the health-related records and transactions of patients with their numerous "identities" as they interact with different healthcare providers, payers, pharmacy benefit managers and other entities is an expensive and complex problem. With many years of experience addressing this issue in different healthcare systems and Health Information Exchanges (HIEs), it is apparent that there is now a compelling and relatively straightforward technical solution for this problem. Presented here is a broadly feasible and technically compelling argument for a blockchain-based approach to addressing these issues. At the same time, challenges ahead and potential strategies to address them are discussed.

Associating the health-related medical records and patient transactions with their numerous "identities" as they interact with different healthcare providers, payers, pharmacy benefit managers (PBMs), and other entities is a complicated problem.[1,2] Failure to efficiently make such associations impacts care and is responsible for considerable wasted time and resources across the U.S.

A blockchain-based approach could be used to index patient identities and locations of their health records in Health Information Exchanges (HIE) and other networks of patient data across the country. Accordingly, the objects for this article are listed here.

1. Set out the various problems of patient identity and their costs.
2. Explore the traditional approaches to tackling the problem, and their limitations.

3. Convey a conceptual design that utilizes blockchain to address the problem of patient identity on a national level.
4. Explore the weaknesses in the concept that will need to be tackled along with strategies for doing so.
5. Lay the basis for a broader conversation to evolve the concept.

Blockchain's characteristics as a decentralized, highly resilient, and secure log of transactions make its architecture a strong fit for the problem of maintaining an accurate association between patients and their records. Blockchain offers a high integrity mechanism for locating data and monitoring precisely how it changes over time.

THE PROBLEM OF PATIENT IDENTITY

How to collect patient data and make it accessible to any organization that performs a service for that patient is a long-standing challenge. In the early 2000's, the conceptual architecture of a master patient index (MPI) coupled with a record locator service (RLS) started being considered for HIEs. By the mid-2000's regional health information organizations (or RHIOs as HIEs were originally called) started springing up as regional or state-wide data aggregators. Almost all of them used some form of the original MPI/RLS model.

In some of the early models "edge servers" were put into place at each participating organization.[3] Under this system, an edge server was a server appliance located within the security perimeter of each HIE-participating organization. The edge server staged the data that each organization shared with the HIE. The record locator service maintained an index of the information but did not house clinical data. The edge server model quickly fell out of vogue as staging data in this special server required significant resources, did not improve security, and did not scale well to small organizations like provider offices.

Next, HIE's started aggregating data, indexing patients through a master patient index, and storing the data in individual repositories for each organization. The MPI effectively became the RLS, and HIEs could share data with treating providers and other participating organizations.

CONNECTING NETWORKS

However, this did not address the issue of how HIEs would communicate with each other, which was increasingly important as individuals relocated around the country, visited multiple providers perhaps in different states, or simply changed insurance. The result was they became disconnected from their health information that was stored in their previous health system EHR or HIE.

The idea of a "network of networks" was long discussed,[4] but it never gained much traction due to implementation costs and lack of an organizational entity that was ready to create this full "open" network. Instead the concept of point-to-point network connections was promoted by organizations like the eHealth Exchange (Sequoia Project) using the "Integrating the Healthcare Enterprise" (IHE) concepts of patient demographic queries (XCPD) and cross community access (XCA) to retrieve documents.[5] While this has been a good first step, it has significant limitations in terms of identifying locations where all patient data exist and scaling the number of connections required for connecting to these sources.

LIMITATIONS OF CURRENT CONNECTING METHODOLOGIES

One thing that seems certain is that a completely federated set of patient data repositories which depend on a brokered broadcast query (XCA) to look up patients using demographics (XCPD) will not scale. This is made clear by looking at the basic math regarding the number of connections required to integrate a fully meshed network of HIEs or other clinical data repositories.[6]

Assume we level out at 200 aggregated clinical data repositories that could be HIE's, integrated delivery networks (IDNs), and other networks such as CommonWell Health Alliance. To connect all 200 networks would require (n*n-1)/2 connections or (200*199)/2=19,900 connections. Each HIE would have 199 connections to query. Having federated repositories (HIEs) of data is not the main problem. The idea that we can do this without some type of index is the core issue.

The SHIEC project

Once such effort to create a network of HIEs is the existing SHIEC Patient Centered Data Home (PCDH) project. PCDH routes data using HL7 v2 messages to other HIEs when you receive care away from your "home" HIE. The basic idea is that if you become ill or are injured away from home and visit an emergency department or hospital, the "away" HIE sends an Admission Discharge Transfer (ADT) message to the "home" HIE based on the zip code of your home address. The "home" HIE responds with a care summary.

Once your visit is complete at the "away" HIE, it creates a visit summary (either a v2 or v3 message) and returns that to the "home" HIE so your primary care provider has these clinical data for future follow-up. This is a straightforward approach that can be implemented by almost all HIEs quickly. It has already shown promise in the Western Region and other regions of the PCDH. In essence, it eliminates the issue of having to query all HIE's by having a predetermined zip code index.

Using this zip code index within the PCDH is better than XCPD/XCA but it could be improved. Zip codes as the "marker" for a true patient data index leaves "holes" in all data for a patient. Consider the millions of patients who summer in the north and winter in the south, or those who recently moved: Most of their clinical data do not exist in their current "home" HIE. The preferred solution is to develop a *nationwide* patient index. However, having a single MPI in the cloud also suffers from issues such as scalability, data quality and accessibility.

A NATIONAL NETWORK OF PATIENT IDENTITY BROKERS

To address the problems identified above, we propose the following concept: A National Network of Patient Identity Brokers with a blockchain-based record locator.

Implement a limited number (~6) of Patient Identity Brokers (PIBs) nationwide. Each regional HIE, IDN or other "network" would connect to one or two PIBs for performance and redundancy and send all their patient demographics via ADT messages. Each PIB would have a Master Patient Index. Current patient identity matching logic has its issues and limitations, but it is far better than demographic queries, and it will improve over time.

Instead of storing the index in the patient identity brokers, the PIB's can manage the required business logic while the index itself would be stored in a single permissioned patient identity blockchain. The blockchain would not contain any Protected Health Information (PHI) but would be the index to all the locations where the patient's clinical data exist (nationwide RLS).

While the index would be in the blockchain, clinical data would remain off-chain. The PIB would be the broker to the patient identity blockchain and would perform specific business functions such as patient matching, ADT message processing, patient private key management, managing relationships of patients to healthcare organizations, and managing the blockchain index. In our core use case described below the PIB would process ADT messages from participating organizations which would “register” this organization as having a relationship with the patient, update patient demographics, and grant access to the patient identity blockchain by managing public and private keys.[7] Any organization (HIE, IDN, or other network) with proper authorization (private keys) could get access to the index. Healthcare organizations could request access (keys) by sending an ADT message to the PIB, which would return the patient’s private key to the organization. The organization would then have access to the index.

In our model the PIBs manage the public and private keys that control access to the patient identity blockchain. A future consideration would allow patients to control their own private keys to the location of their medical records. A user facing application could be developed to allow a patient to manage their consent for certain providers to view their data. While the design of such an application is outside the scope of this paper, the underlying architecture of PIBs coupled with a patient identity blockchain would likely be very conducive to an application of this type. One can imagine how access to sensitive data such as substance abuse data, which is governed by 42 CFR Part 2, could be managed and controlled in this model with the aforementioned consent application.

When a local HIE or healthcare organization wanted to perform a query, it would perform an *indexed broadcast query* using a regional patient identifier which would greatly improve identification of all patient data sources.

The PIB could perform other services. For example, the PIB could implement the standardization of transactions and the normalization of the data required to implement the blockchain’s smart contracts. The smart contracts would effectively be things like the data use requirements and authorization for access. After performing a query, the PIB could consolidate the responses into a single patient summary (C-CDA) to return to the requesting organization.

HIPAA compliance would be managed in our model in much the same way that it is managed in the patient-centered data home or Sequoia eHealth Exchange. Participants sign a participation agreement, which would include a business associate agreement and describe the rules of participation and the data use standards. Access to the data locations would be managed by the PIBs as described above but consent to access the actual medical records would still be managed by the source organizations and subject to their patient consent and data use requirements. The proposed architecture of the Patient Identity Blockchain is displayed in Figure 1.

LIMITATIONS OF CURRENT MATCHING METHODS

There are many issues with our current deterministic and probabilistic matching methods in MPI's. The PIBs will need to employ the best matching algorithms and methods available, including using data sources in an appropriate way (e.g., phone numbers, credit reports, previous addresses).

CORE USE CASE

The core use case for loading patients into the blockchain and using the index for querying patients is as shown in Table 1.

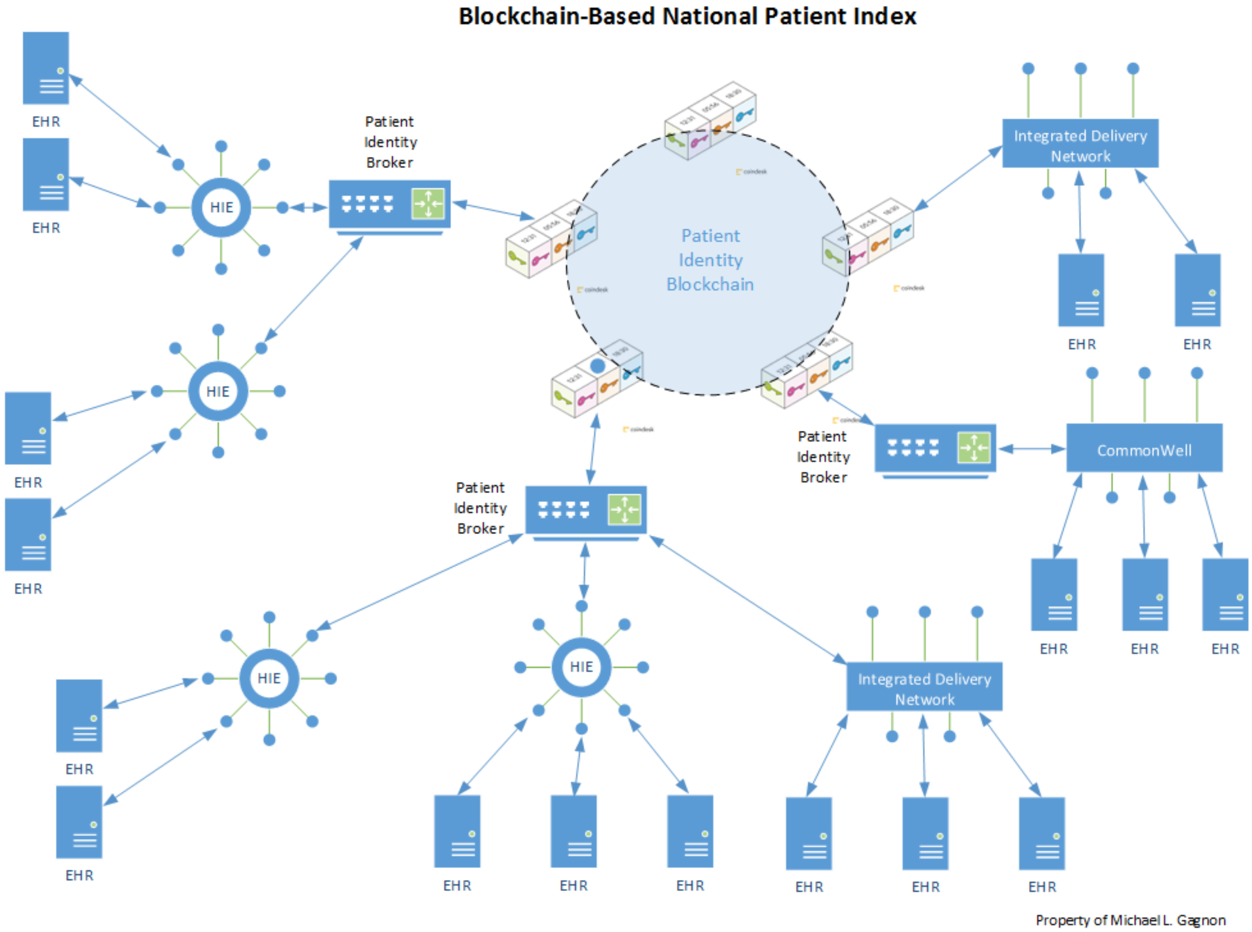

Figure 1. Proposed architecture of the Patient Identity Blockchain (EHR: Electronic Health Records)

There will be false negatives (duplicates); and each PIB might employ biometrics to improve matching. Matching solutions that use facial recognition and iris scans to improve patient matching already exist.[8] Regional MPIs at HIEs will need to start "scoring" demographics from source organizations to ensure that patient data sent to the PIB and then put into the blockchain meet the smart contracts' data standards. This part of the solution clearly requires continuous quality improvement.

The recently released Office of the National Coordinator (ONC), Trusted Exchange Framework Common Agreement (TEFCA) could help in this regard with a set of required patient demographics and the framework for participation. The framework agreements could propose permitted uses and other policies, which could utilize chaincode and become blockchain smart contracts. Perhaps this could be the purview of the Recognized Coordinating Entity (RCE) as proposed by TEFCA.

ADVANTAGES OF THE BLOCKCHAIN APPROACH TO MPI

The advantages to using blockchain as the architecture for solving the patient indexing problem at the HIE level are listed in Table 2.

Table 1. Core use case for loading patients into the blockchain and using the index for querying patients

Event	Activity
Healthcare event	1. Patient encounter at a connected provider organization.
System activation	2. That organization sends an ADT message with patient demographics to a broker; which includes a master patient index. a. It also submits a “consent” flag indicating whether this patient’s data can be queried at this time. Subsequent messages may turn consent on or off. 3. The master patient index uses deterministic and probabilistic logic to determine if the patient already exists in the index. The patient is scored as a match, a probable match, or a non-match. 4. If the patient is a non-match, then a new patient private key and universal ID are created. If the patient is a match, these already exist.
Reply	5. A message is returned to the original provider organization with the patient’s private key and a universal identifier.
Ensuring privacy	6. The originating organization can store the patient’s private key and universal ID in their local MPI. 7. The originating organization then passes a URL or IP address to the broker, which is the node acting as the responding gateway for queries from other participants. 8. The broker then performs two functions: a. Add an index to the blockchain for this organization’s responding gateway for that patient, in effect identifying that this organization has data on this patient. b. Update the patient’s best demographics “golden record” with new information. The broker MPI determines whether the new data should supersede what is already stored. 9. The originating organization opens the chain to identify all locations where a patient’s data may be located (or use the broker). 10. Using an existing HL7 Standard XCA query the originating organization queries these locations and retrieves clinical documents. 11. Depending on the capabilities of the originating organization’s systems, the broker may consolidate steps 9 and 10 into a single clinical document to be returned to the originating organization.
Audit trail	12. When a query is performed, and data exchanged, the blockchain stores the audit trail of what data went to what organization.

ADT: Admission Discharge Transfer; HL7: Health Level Seven International; MPI: Master Patient Index; XCA: Cross-Community Access

ISSUES TO BE ADDRESSED

There are issues with using blockchain, specifically when it comes to performance and maturity. Performance of a large scale blockchain, which must be used to retrieve clinical data, often in real time, can be a significant issue. Requiring each node to process each transaction may make it resilient to cyberattacks, but it also limits transaction processing speeds. In addition, the chain nature of a blockchain requires that each transaction be serialized, which can slow updates. In this regard, the design of the index can definitely improve the performance of the system.

Table 2. Advantages to using blockchain as the architecture for solving the patient indexing problem at the HIE level

Advantage	Comment
1. **Reliability**	Most importantly, it will directly impact our ability to care for patients.
2. **Cost effectiveness**	Currently, large amounts of money and resources are wasted due to limitations of alternative approaches.
3. **Scalability**	Efficient scalability is a crucial advantage of the blockchain approach, especially on a regional or national level.
4. **Responsiveness**	An appropriately designed blockchain approach will provide a fast mechanism to associate patients with their up-to-date records.
5. **Simplicity**	Blockchain avoids many of the governance and access issues that plague other solutions to these problems.

We can envision two different chains for each patient, one for the transactions performed against the index such as patient registration events or queries, and a second chain which would be an index of the locations of patient data. The first chain could become very large as ADT registration events are quite verbose and queries would become routine. The second chain of locations where a patient's data are stored, would probably remain quite small as long as HIEs, integrated delivery networks or other integration networks perform the initial integration. For the vast majority of patients, this number would probably be ten or less.

TECHNICAL OPTIONS

As for the two most prevalent blockchain platforms, Hyperledger and Ethereum are both relatively immature which can lead to unforeseen issues with deployment and software bugs.[9] Over time both of these platforms will continue to improve and make enhancements such as Ethereum's concept of "sharding" which requires a far smaller number of nodes to process each transaction.[10] Then there are entirely new concepts for encrypted ledgers such as IOTA's Tangle. The IOTA Tangle was originally proposed as a solution for connecting the Internet of Things (IoT) and uses a more interesting underlying data structure called a *directed graph.*[11] Instead of a very simple chain, which is effectively a secure linked list, the directed graph only requires each transaction entering the tangle to approve two previous transactions. Any unapproved transaction is called a "tip" and the more transactions that approve any given transaction the more confidence the system has with this transaction.

Ultimately the directed graph structure or Ethereum's sharding could be more effective as the model for our patient transactions, but a traditional blockchain would be sufficient for at least during the first phase of the patient index. This is because, as mentioned above, the total size of any patient's chain of all the locations where their medical information exists will be modest (perhaps <10) especially when HIEs play the role of initially aggregating the information. At this size a traditional blockchain should perform adequately.

CONCLUSION

This paper presents a high level conceptual design for a blockchain approach to the patient indexing problem and an outline of the clear advantages it has over other approaches. However, developing a nationwide patient index is not a simple task. It will involve overcoming many design challenges.

At this point, we are limiting our concept of an index of patient medical data locations to the U.S., but we envision this expanding to other countries over time. The models for the necessary US data use agreements are already in place through the SHIEC patient centered data home and CareEquality. However, to implement this approach internationally will be more complex.

The goal is not to make MPIs free from issues but, instead, to recognize that patient ID queries based on demographics will not scale and that, in the interests of patient care and cost control, a nationwide patient index is required. We should not make the perfect be the enemy of the good with regards to the Patient Identity Brokers.

The fundamental point is that, through collaborations with other disciplines and stakeholders, blockchain offers the opportunity to finally ensure that a complete record of a patient's clinical data is truly available to the patient and clinician regardless of where the patient received care.

Acknowledgments: The authors thank Andrew Hinkle for his technical advice.

Contributors

To fulfil all of the criteria for authorship, every author of the manuscript has made substantial contributions to **ALL** of the work and participated sufficiently in the work to take public responsibility.

Funding statement: The authors confirm that no third-party funding supported or influenced the development of this paper.

Conflict of Interest: The authors confirm that there are no known conflicts of interest associated with this publication.

REFERENCES

1. Wang X, Ling J. Multiple valued logic approach for matching patient records in multiple databases. J Biomed Inform. 2012 Apr;45(2):224-30. doi: 10.1016/j.jbi.2011.10.009. Epub 2011 Nov 10.
2. Posnack S. ONC Patient Matching Challenge. *healthit.gov.* [Online] May 1, 2017. URL: https://www.healthit.gov/buzz-blog/interoperability/demystifying-patient-matching-algorithms/.
3. Shapiro, Jason S. and Kuperman, Gilad. Health Information Exchange. [book auth.] Ken Ong. *Medical Informatics, An Executive Primer.* Chicago: HIMSS, 2011, pp. 147-160.
4. *Information Exchange: 'Lex Parsimoniae'.* Overhage, J. Mark. 2007, *Health Affairs* 26, no. 5, pp. 595-597.

5. The Sequoia Project. About the Sequoia Project. 2018. URL: https://sequoiaproject.org/about-us/
6. Network Topologies | Computer Networks. *Geeks for Geeks.* [Online] Disqus, 2017. https://www.geeksforgeeks.org/network-topologies-computer-networks/.
7. Adams C, Lloyd S. *Understanding PKI, Concepts Standards and Deployment Considerations.* Boston : Addison Wesley, 2003. 0-672-32391-5. URL: https://www.pearson.com/us/higher-education/program/Adams-Understanding-PKI-Concepts-Standards-and-Deployment-Considerations-2nd-Edition/PGM216259.html
8. Trader J. *Iris Recognition vs. Retina Scanning—What are the Differences?* M2SYS Blog On Biometric Technology 2012 URL: http://www.m2sys.com/blog/biometric-hardware/iris-recognition-vs-retina-scanning-what-are-the-differences/.
9. Ray J. Sharding FAQs. *ethereum/wiki.* [Online] Github, June 15, 2018. https://github.com/ethereum/wiki/wiki/Sharding-FAQs.
10. Singh N. Ethereum or Hyperledge Fabric. Quillhash. 2018. URL: https://medium.com/quillhash/ethereum-or-hyperledger-fabric-259f3c9b8da6
11. Popov, Serguei. *The Tangle.* Berlin, Germany : IOTA Foundation, 2018.