

FHIRChain: Applying Blockchain to Securely and Scalably Share Clinical Data

Peng Zhang^a, Jules White^a, Douglas C. Schmidt^a, Gunther Lenz^b,
S. Trent Rosenbloom^c

^a*Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, Tennessee, USA*

^b*Varian Medical Systems, Palo Alto, California, USA*

^c*Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA*

Abstract

Secure and scalable data sharing is essential for collaborative clinical decision making in telemedicine. Conventional clinical data efforts are often siloed, however, which creates barriers to efficient information exchange and impedes effective treatment decision making for remote patients. This paper provides four contributions to the study of applying blockchain technology to clinical data sharing in the context of technical requirements defined in the "Shared Nationwide Interoperability Roadmap" from the *Office of the National Coordinator for Health Information Technology* (ONC). First, we analyze these requirements and their implications for blockchain-based systems. Second, we present FHIRChain, which is a blockchain-based architecture designed to meet ONC requirements by encapsulating HL7's *Fast Healthcare Interoperability Resources* (FHIR) standard for shared clinical data. Third, we demonstrate a FHIRChain-based decentralized app using digital health identities to authenticate participants in a collaborative decision making case study. Finally, we highlight lessons learned from our case study.

Keywords: Blockchain, Interoperability, Digital Health Identity, Smart Contract, Health Data Sharing

Email addresses: peng.zhang@vanderbilt.edu (Peng Zhang), jules.white@vanderbilt.edu (Jules White), d.schmidt@vanderbilt.edu (Douglas C. Schmidt), gunther.lenz@varian.com (Gunther Lenz), trent.rosenbloom@vanderbilt.edu (S. Trent Rosenbloom)

1. Introduction

The importance of data sharing in collaborative decision making. Secure and scalable data sharing is essential to provide effective collaborative treatment and care decisions for patients, especially in telemedicine [1] where patients are remotely diagnosed and treated. Data sharing helps improve diagnostic accuracy [2] by gathering confirmations or recommendations from a group of medical experts, as well as preventing inadequacies [3] and errors in treatment plan and medication [4, 5]. Likewise, aggregated intelligence and insights [6, 7, 8] helps clinicians understand patient needs and in turn apply more effective remote treatments.

For example, groups of physicians with different specialties in cancer care form tumor boards that meet regularly to discuss cancer cases, share knowledge, and determine effective cancer treatment and care plans for patients [9]. Regional virtual tumor boards are also being implemented via telemedicine [10, 11] for institutions that lack inter-specialty cancer care due to limited oncology expertise and resources [12].

Data sharing barriers to collaborative decision making. In practice, many barriers exist in the technical infrastructure of telemedicine and other health IT systems that impede the secure and scalable data sharing across institutions, limiting support for collaborative clinical decision making. Examples of such barriers include the following:

- **Security and privacy concerns.** Despite the need for data sharing, concerns remain regarding protection of patient identity and confidentiality in telemedicine [13]. Virtual medical treatment may increase the risk of health data breaches through electronic storage and transmission without a highly secure infrastructures in place breaches, which can result in severe financial and legal consequences [14]. Likewise, without requiring face-to-face interactions between providers and patients, medical identity theft may occur more frequently [13].
- **Lack of trust relationships between entities.** Trust relationships between care providers and/or institutions [15] are an important precondition to digital communications [16] and data sharing in the absence of a custodian over shared data. Telemedicine clinics are typically networked with larger hospitals [17], but may not establish communications with private or smaller practices.

- **Scalability concerns.** Large-scale datasets may be hard to share electronically due to limitations in bandwidth, restrictive firewall settings, etc, such as in rural areas [18]. Lack of scalability can also impact overall system response time and transaction speed [19].
- **Lack of interoperable data standards to ensure data understandability.** Without the adoption of interoperable data standards (such as HL7’s *Fast Healthcare Interoperability Resources* (FHIR)[20] standard for shared clinical data), clinical data can vary in formats and structures that are hard to interpret and integrate into other systems [21].

What is needed, therefore, is a standards-based architecture that can integrate with existing telemedicine systems to enable secure and scalable clinical data sharing for improving collaborative decision support.

Research focus and contributions → Evaluating blockchain-based secure and scalable sharing of clinical data to support collaborative clinical decision making. In recent years, blockchain technologies have been increasingly touted [22, 23, 24] as a technical infrastructure to support clinical data sharing that promotes care coordination. A key property of blockchains is their support for “trustless disintermediation,” which enables multiple parties who do not fully trust each other to exchange digital assets (such as the Bitcoin cryptocurrency [25]), while still protecting their real identities from each other. This paper focuses on a key research question related to addressing the barriers described above: *is it feasible to use blockchain technologies to securely and scalably share healthcare data for improving collaborative clinical decision support?*

This paper provides the following contributions to assessing the feasibility of blockchain technologies in clinical data sharing to improve collaborative decision support:

- We describe key technical requirements defined by the *Office of the National Coordinator for Health Information Technology* (ONC), who defined a “Shared Nationwide Interoperability Roadmap” [26] for creating an interoperable health IT system. These technical requirements include verifiable identity and authentication of all participants, ubiquitous and secure data store and exchange, permission to access data source, consistent data formats, and maintaining system modularity.

- We present the design and implementation of a blockchain-based architecture called FHIRChain that meets the ONC technical requirements for sharing clinical data between distributed providers. FHIRChain uses HL7’s FHIR data elements (which have uniquely identifying tags) in conjunction with a token-based design to exchange data resources in a decentralized and verifiable manner without actually moving the data.
- We demonstrate a FHIRChain-based *decentralized app* (DApp) that uses digital health identities to easily authenticate participants and manage data access authorizations in a case study of clinical data sharing. This DApp enables users to share specific and structured pieces of information (rather than an entire document), thereby increasing the readability of data and flexibility of sharing.
- We highlight key lessons learned from our case study and evaluate how our FHIRChain-based DApp can be further extended to support other technical requirements for improving advanced healthcare interoperability issues, such as coordinating various stakeholders (*e.g.*, insurance companies) across the industry and providing patients with easier (and more secure) access to their own medical records.
- We also explore the data exchange issues that blockchains do not yet solve, such as semantic interoperability and healthcare malpractice and unethical use of data, which remain as future research problems in this space.

Paper Organization. The remainder of this paper is organized as follows: Section 2 provides an overview of blockchain technologies and the Ethereum platform, which is an open-source blockchain implementation that supports programmability via “smart contracts;” Section 3 analyzes ONC’s key technical requirements for sharing clinical data and their implications for blockchain-based designs; Section 4 describes our blockchain-based architecture, called FHIRChain, designed to meet ONC requirements, and motivates why specific architectural decisions were made; Section 5 highlights the results of our case study that applied a FHIRChain-based DApp to provide collaborative clinical decision support; Section 6 compares our work with related blockchain research in the healthcare domain; and Section 7 presents concluding remarks and summarizes our lessons learned and future work on extending the FHIRChain architecture.

2. Overview of Blockchain

The most popular blockchain (*i.e.*, the Bitcoin blockchain [25]) is a public distributed ledger designed to support financial transactions, such as Bitcoin and other cryptocurrencies. A public blockchain operates in a peer-to-peer fashion with all transactions distributed to each network node (called a “miner”) for verification and admittance to the blockchain. These miners validate available transactions and group them into blocks, as shown in Figure 1. Miners then compete in solving a computationally expensive cryptographic puzzle to append their block to the blockchain sequence.

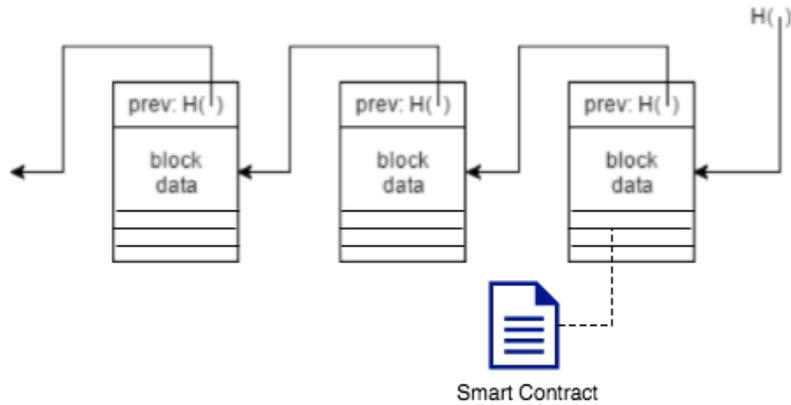


Figure 1: The Blockchain Structure: a Continuously Growing and Immutable List of Ordered and Validated Transactions

The Bitcoin blockchain uses the proof-of-work process described above to achieve consensus by

- incentivizing miners to contribute powerful hardware and electricity to the network with small amounts of cryptocurrency as rewards and
- discouraging rogue actors against malicious control of the system.

After a block is added to the blockchain, its transaction history is secured from tampering via cryptography.

While the Bitcoin blockchain is widely deployed, other blockchain technologies have emerged recently. In particular, the Ethereum blockchain [27] provides a more generalized framework via “smart contracts” [28] that allow programs to run on the blockchain and store/retrieve information. Smart

contracts, shown in Figure 1, enable code to execute autonomously when certain conditions are met. They can also store information as internal state variables and define custom functions to manipulate or update this state. Operations in smart contracts are published as transactions and thus occur in a globally sequential order. These operations are deterministic and verifiable by miners in the Ethereum blockchain to ensure their validity.

The mechanisms described above make a blockchain decentralized and immutable, thereby removing the need for a trusted central authority. These properties make blockchain technologies attractive to health IT researchers and practitioners as a promising solution to improve clinical communications, while protecting the privacy of healthcare participants. The remainder of this paper examines the question of whether blockchains can be leveraged to securely and scalably share clinical data that enables collaborative decision support.

3. Technical Requirements for Blockchain-Based Clinical Data Sharing

The ONC’s “Shared Nationwide Interoperability Roadmap” defines technical requirements and guiding principles for creating interoperable health IT systems [26]. After analyzing these requirements, we posit it is feasible to craft a blockchain architecture that meets its key technical requirements. Significant challenges must be overcome, however, when attempting to utilize blockchains to share clinical data for providing collaborative decision support.

This section first analyzes five key technical requirements that are fundamental to clinical data sharing systems and then discusses the implications of these requirements on blockchain-based architectures. Sections 4 and 5 then describe how we developed and applied our blockchain-based architecture, called FHIRChain, to create a decentralized app that meets the ONC requirements in the context of collaborative clinical decision making.

3.1. Requirement 1: Verifying Identity and Authenticating All Participants
ONC requirement summary. The ONC requirements state that an identity ecosystem should be employed to minimize identity theft and provide redress in case of medical identity fraud, while complying with individual privacy regulations. Providers, hospitals, and their health IT systems should be easily identity-proofed and authenticated when exchanging electronic health

information. Healthcare systems today, however, lack consistently applied methods and criteria for both identity proofing and authentication across organizations, *e.g.*, different network service providers have different policies or requirements and may not acknowledge the methods applied by other network service providers.

One of the most popular—and least complex—approaches to exchange data is through direct secure messaging [26]. The Direct Project (Direct) was launched to create a standard way for participants to send authenticated, encrypted health information directly to known, trusted recipients over the Internet [29]. Providers who use an EHR system without Direct integration, however, cannot benefit from the direct messaging capability.

Implications for blockchain-based system design. For a blockchain-based system storing identification information (such as personal email) directly on-chain is problematic [30]. In particular, a property of conventional public blockchains (such as the Bitcoin blockchain) is information openness, *i.e.*, all data and associated modification records are publicly and immutably recorded [25]. To meet this requirement, therefore, a blockchain-based system should support identifiability and authentication of users while encapsulating sensitive user information. Section 4.2.1 shows how we address the identifiability and authorization requirement in FHIRChain via digital health identities based on public key cryptography [31].

3.2. Requirement 2: Storing and Exchanging Data Securely

ONC requirement summary. The ONC requirements state that information should be shared in a secure and private manner and not altered in an unauthorized or unintended way, while still making the information available when needed by those authorized to access it. Data encryption is recommended as a means to send data over networks (*i.e.*, data in motion) and when it is stored (*i.e.*, data at rest). The procedures by which encryption keys are generated, distributed, stored, rotated, and revoked must be secure and tightly controlled.

Implications for blockchain-based system design. There has been recent interest [32, 33] in using blockchain as decentralized storage for encrypted health data. As discussed in Section 2, however, the open and transparent nature of blockchain raises privacy concerns when attempting to integrate blockchain into the health IT domain. Although sensitive data can be encrypted, flaws

in encryption algorithms or software implementations may expose the data contents in the future. To ensure long-term data security, therefore, a data storage design should be “simple” to minimize software bugs [34], *e.g.*, by not storing sensitive data (encrypted or not) on-chain, while still allow the data to flow from one user to another [35].

Another implication of storing data on the blockchain is system scalability. Each blockchain transaction (such as storing data in a smart contract and modifying the data) is associated with a small fee paid to the miner for validating and then logging the transaction to the blockchain, as described in Section 2. If the entire clinical data ecosystem is stored on the blockchain, however, then each time new data is available or existing data is modified, the data residing on the blockchain must be updated via a smart contract data operation, leading to significant long-term operation costs. Section 4.2.2 shows how we address this requirement in FHIRChain with a hybrid on-chain/off-chain store and exchange via data reference pointers.

3.3. Requirement 3: Ensuring Permissioned Access to Data Sources

ONC requirement summary. The ONC advocates computable privacy that represents and communicates the permission to share and use identifiable health information. New technological advances should enable individuals to document their permissions electronically, which are then honored appropriately in circumstances where they are required. Modern web-based apps manage access via OAuth [36]. OAuth is an open standard for access delegation that enables Internet users to grant websites or applications access to their information on other websites without disclosing their passwords.

Implications for blockchain-based system design. Unfortunately, since smart contract operations only occur in the blockchain space to ensure deterministic outcomes, services (such as OAuth) that exist off the blockchain cannot be used. Determining if alternative approaches exist to provide data access permission given this constraint is key to assessing the feasibility of blockchain-based designs in healthcare. Section 4.2.3 shows how we address this requirement in FHIRChain via a token-based permission model.

3.4. Requirement 4: Applying Consistent Data Formats

ONC requirement summary. The ONC requirements state that to satisfy interoperability needs, health IT systems should be implemented with an intentional movement and bias toward a standard identified by ONCs most recently finalized *Interoperability Standards Advisory* [37]. The data exchanged

should be structured, standardized, and discrete (granular [38]) information. Likewise, standards should use metadata where possible to allow human users to communicate this context along with pieces of structured data.

Implications for blockchain-based system design. To provide collaborative clinical decision support, health IT systems must present shared data to clinicians in a structured and readable format [39]. This requirement implies that these IT systems should enforce existing, commonly accepted clinical data standard(s) instead of introducing new data exchange formats. Section 4.2.4 shows how we address this requirement in FHIRChain by enforcing FHIR standards.

3.5. Requirement 5: Maintaining Modularity

ONC requirement summary. The ONC requirements state that since medicine and technology will inevitably change over time, scalable health IT systems should also preserve the abilities to evolve by maintaining modularity. When divided into independent and connected components, modular systems become more resilient to change. In particular, modularity enhances flexibility, which in turn enables innovation and adoption of new, more efficient approaches over time without overhauling entire systems.

Implications for blockchain-based system design. Modularity requires a careful system design to avoid information locking due to the immutability of smart contracts. Every change to a smart contract code creates a new contract instance on the blockchain, nullifying previous versions and their data. To minimize dependencies and the need to upgrade, therefore, smart contracts should be loosely coupled with other components in the system. Section 4.2.5 shows how we address this requirement in FHIRChain by applying the *model-view-controller* (MVC) pattern [40].

4. FHIRChain: a Blockchain-Based Architecture for Clinical Data Sharing

This section first presents an overview of FHIRChain, which is a blockchain-based architecture we developed to meet the ONC requirements for secure and scalable sharing of clinical data described in Section 3. We then explain why specific architectural decisions were made to address each requirement.

4.1. FHIRChain Overview

Figure 2 shows the FHIRChain architecture we devised to address key ONC requirements. This architecture provides a general data sharing solution applicable to a wide range of health IT systems and also serves as the basis for our decentralized app (DApp) in Section 5, which customizes FHIRChain to support collaborative clinical decision making using a case study of a telemedicine tumor board. The dashed ellipse in Figure 2 represents

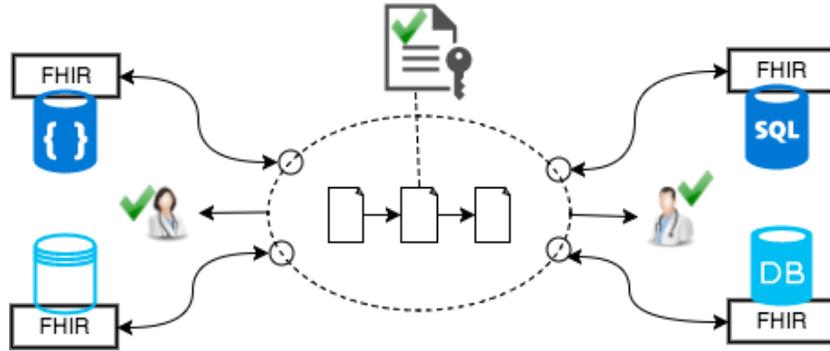


Figure 2: Architectural Components in FHIRChain

a blockchain component that mediates data sharing between collaborating medical professionals (represented by providers with green check marks). Clinical data to share are represented by heterogeneous database symbols and are normalized via FHIR standards to create a common format that enhances the readability of shared data. Secure database connectors (represented as small circles) connect siloed data sources to the blockchain by exposing their reference pointers in smart contracts (represented by linked documents) on the blockchain.

Smart contracts are also used to store an immutable timestamped transaction log (represented as a keyed file symbol zoomed in from the blockchain) of all interactions related to data source sharing and accessing via reference pointers. These logs include specific sharing information regarding what access has been granted to which user by whom and who has accessed which resource. To ensure the validity of shared data, FHIRChain can be configured to only approve participation from certified clinicians and healthcare organizations.

4.2. FHIRChain Architectural Decisions that Address Key ONC Technical Requirements

Below we explain why we made specific architectural decisions to address each ONC requirement presented in Section 3.

4.2.1. Addressing Requirement 1: Verifying Identity and Authenticating All Participants

Context. Blockchains like Ethereum and Bitcoin provide pseudo-anonymous personal accounts (*i.e.*, public addresses composed of random hash values) for users to transact cryptocurrencies. These native identities, however, do not address healthcare needs for identifiability or authentication.

Problem. By design, public blockchains are globally accessible to anyone with Internet access and allow each participant to hold any number of blockchain accounts to minimize traceability of account holders. However, this ONC requirement specifies that all U.S. healthcare participants should be identifiable, implying the need for an entirely separate, traceable user base from the native identities. A key problem is thus how to properly define identities for healthcare users participating in clinical data sharing while protecting sensitive personal information on the blockchain.

Design choice → use of a digital health identity. Inspired by the success of secure shell (SSH) [41] and blockchain address generation mechanism, FHIRChain employs public key cryptography [31] to manage identities in the framework. In public key cryptography, a pair of mathematically related public and private keys is used to create digital signatures and encrypt data. It is computationally infeasible to obtain the private key based on the public key. Public keys can thus be shared freely, allowing users to encrypt content and verify digital signatures. Likewise, private keys are kept secret to ensure only the owners of the private keys can decrypt content and create digital signatures.

FHIRChain generates a cryptographic public/private key pair (also used for encryption, as described in Section 4.2.3) for each data sharing provider *e.g.*, in-house care providers and remote clinicians in telemedicine practice. The public key is associated with each user’s digital health identity. These public identities can be recorded in the blockchain for tamper-proofing, and users holding the corresponding private keys can authenticate to use the system.

FHIRChain’s design applies a smart contract to maintain health users’ identifiability without exposing personal information on the blockchain. It also replaces the need for a traditional username/password authentication scheme with the use of a public/private cryptographic key pair for authentication. In a general clinical setting, these digital health identities—private keys—would be hard to manage for patients. FHIRChain, however, only creates these identities for clinicians to facilitate data sharing, which consequently enables more effective collaborative decision making for patients.

4.2.2. Addressing Requirement 2: Storing and Exchanging Data Securely

Context. A key capability offered by blockchains is their support for trustless transactions between parties who lack trust relationships established between them. Bitcoin is the most common example of this trustless exchange via its native cryptocurrency. Blockchains are peer-to-peer by nature and thus contribute to the ubiquitousness of digital assets being transacted.

Problem. Health data as digital assets are much more complex and harder to share *en masse*. There are also privacy and security concerns associated with its storage in an “open” peer-to-peer system (*i.e.*, public blockchains), such as encryption algorithms applied to protect data being decryptable in the future [35]. A key problem is thus how to design a blockchain-based health IT system so that it balances the need for ubiquitous store and exchange and the concerns regarding privacy of the data and scalability of the system.

Design choice → hybrid on-chain/off-chain data store and exchange via reference pointers. Instead of storing encrypted health data in the blockchain, an alternative option is to store and exchange encrypted metadata for acquiring the protected data (*i.e.*, a reference pointer to a data source) that can also be combined with an expiration for short-term sharing. Exchanging reference pointers allows providers to maintain their data ownership and choose to share data at will.

As shown in Figure 2, FHIRChain attaches a secure connector to each database. Each connector generates appropriate reference pointers that grant access to the data. These reference pointers are digital health assets that can be transacted ubiquitously with reduced risks of exposing the data.

A benefit of not directly exchanging shared data *en masse* is scalability. As discussed in Section 3.2, each transaction or operation on the blockchain (*e.g.*, querying a smart contract state variable value or updating it) is associated with a small fee paid to the miner for verification and then inclusion

to the blockchain. Transacting these lightweight reference pointer data is more efficient, both in terms of time and cost in production health IT systems because small changes to data generally do not require modifications to reference pointers.

4.2.3. Addressing Requirement 3: Permission to Access Data Sources

Context. Data references can be stored on the blockchain for ubiquitous access via a smart contract. However, access rights must be granted only to authorized providers for viewing the data. As discussed in Section 3.3, OAuth is a popular platform for communicating permissions in web-based apps that are not blockchain-based.

Problem. Smart contracts cannot directly use external services like OAuth since they do not produce deterministic outcomes that can be verified by blockchain miners. A key problem is thus how to design a mechanism that balances the need of permission authorization for clinical data and blockchain requirements for deterministic outcomes.

Design choice → token-based permission model. To overcome the limitation with public blockchains, FHIRChain protects the shared content via a secure cryptographic mechanism called “sign then encrypt” [42]. This design uses the users’ digital health identities to encrypt content so that only users holding the correct digital identity private keys are able to decrypt the content. FHIRChain also generates a new pair of signing keys for each participant and registers the public portion of signing keys alongside users’ digital identities.

To concretely demonstrate this workflow, Figure 3 provides an example of using FHIRChain to create and retrieve an access token. Suppose provider *Alice* would like to initiate sharing of her patient’s data, denoted as D_{Alice} (with a reference pointer, denoted as RP_{Alice}) with another provider *Bob*. FHIRChain creates a digital signature on the shared content RP_{Alice} , with *Alice*’s private signing key SKS_{Alice} for tamper-proof as a first step. With *Bob*’s public encryption key, PK_{Bob} , FHIRChain encrypts the signed RPS_{Alice} to obtain an encrypted token $EncRPS_{Alice}$, and then stores $EncRPS_{Alice}$ in a smart contract for ubiquitous access.

When *Bob* wants to obtain the content *Alice* sent, he must use his corresponding private encryption key SK_{Bob} to decipher the real content of $EncRPS_{Alice}$. *Bob* also verifies that this content was indeed provided by *Alice* with her public signing key PKS_{Alice} . This authentication process is

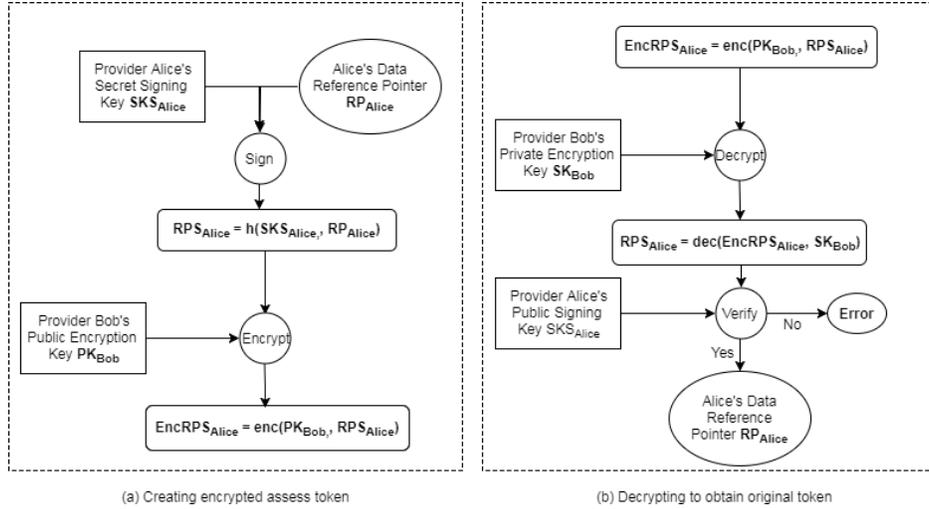


Figure 3: Example of the Creation and Retrieval of an Access Token Using FHIRChain.

automated by the DApp server component interfacing the smart contract, as discussed in Section 4.2.5.

The purpose of digital signing is to ensure that a resource is indeed shared by the sender and is not tampered with. Likewise, encryption is used to protect the information against unauthorized access and spoofing. Access to a resource can be approved or revoked at any time via a state update in the smart contract where all permissions are logged.

4.2.4. Addressing Requirement 4: Consistent Data Formats

Context. Clinical research data can exist in various formats and structures, which when shared with other providers from different organizations, may or may not be meaningful.

Problem. Blockchain-based health IT systems should facilitate data sharing while adhering to some existing standard(s) for representing the clinical data. A key problem is thus how to design a blockchain-based architecture to enforce the application of existing clinical data standard(s).

Design choice → enforcing FHIR standards. HL7's FHIR standards use a popular form of data structure, JSON [43], for exchanging clinical information. JSON is more compact and readable compared to XML used by other data formatting standards, which enables more efficient transmission

of JSON-encoded data. It is also compatible with many software libraries and packages. As more health IT systems upgrade their data exchange protocols to comply to FHIR standards, FHIRChain enforces the use of FHIR to shared clinical data by validating whether the generated reference pointers follow the FHIR API standards [20].

4.2.5. Addressing Requirement 5: Maintaining Modularity

Context. Health IT system updates and/or upgrades are necessary to adopt more efficient, secure, or prevalent technology as it advances.

Problem. If functions in a smart contract have too many dependencies on the rest of a health IT system, then each upgrade to the system must deploy a new contract, which requires restoring data from previous versions to prevent loss. A key problem is thus how to design a modular data sharing system that minimizes the need to create new versions of existing contracts when the system is upgraded, such as for improving the user interface design.

Design choice → *applying the model-view-controller (MVC) pattern.* The MVC pattern [40] separates a system into three components: (1) the *model*, which manages the behavior and data of a system and responds to requests for information about its state and instructions to change state, (2) the *view*, which manages the display of information, and (3) the *controller*, which interprets user inputs into appropriate messages to pass onto the *view* or *model*.

The FHIRChain architecture applies the MVC pattern to separate concerns with individually testable modules as follows: (1) a model in the form of an immutable *blockchain component* is used to store data via smart contracts; (2) a view provides a front-end *user interface* that accepts user inputs and presents data; (3) a controller is a *server* component with control logic that facilitates interactions with data between the *user interface* and *blockchain component*, such as queries, updates, encrypting and decrypting contents; and (4) a controller-invoked *data connector* service is used to validate the implementation of FHIR standards and create reference pointers for the data sources upon requests from the server.

The workflow for updating data access is shown in Figure 4 by the following steps 1-4:

1. A user first authenticates through the user interface (UI), and when successfully authenticated, data access permission request can be input to the system;

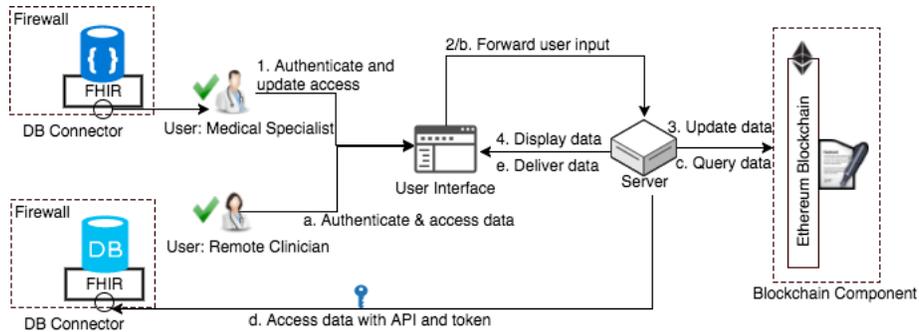


Figure 4: Composition and Structure of the FHIRChain Architecture with Modular Components.

2. The UI forwards user’s request to the server;
3. The server logs permissioned or revoked access in the blockchain component (BC); and
4. The server updates UI with proper response to notify the user.

Likewise, the workflow for accessing a data source is outlined in the following steps a-e:

- a) The user first authenticates via the UI, and when successfully authenticated data access request can be input to the system;
- b) UI forwards users request to the server;
- c) The server queries BC for current user’s access token(s);
- d) When permission is valid, the server decodes the access token(s) with correct keys supplied by user and uses the decrypted reference pointer to obtain actual data from the DB connector to the proper database;
- e) When data has been retrieved from the data source via DB connector, the server updates UI to display data in a readable format.

FHIRChain stores all relevant information in smart contracts, decoupling data store from the rest of the system. This decoupling enables future upgrades to all other components without losing access to—or locking out—existing users or their permission information.

5. Case Study: Applying FHIRChain to Create a Telemedicine DApp

This section first describes the structure and functionality of a *decentralized app* (DApp) that customizes the FHIRChain architecture described in

Section 4 to support collaborative clinical decision making via a telemedicine tumor board case study. We then summarize the benefits and limitations of our DApp case study.

5.1. Overview of the FHIRChain DApp Case Study

The DApp has an intuitive user interfacing portal that facilitates the sharing and viewing of patient cancer data for a telemedicine tumor board to collaboratively create treatment plan for cancer patients. In addition, the DApp implements a notification service [44] to alert collaborative tumor board members when new data access is available for review. Our DApp customizes the FHIRChain architecture in a private Ethereum testnet to address the various ONC requirements described in Section 3, as discussed below.

Verifying identity and authenticating participants with digital identities, as discussed in Section 4.2.1. Our DApp contains a *Registry* smart contract that maintains the digital health identities of providers who registered with our app. The registry maps provider email (or phone numbers) from a public provider directory to both their public encryption (used as digital identity) and signing keys, which are generated automatically at user registration time. Figure 5 demonstrates the user registration and authentication workflow.

Storing and exchanging data securely with FHIR-based reference pointers, as discussed in Sections 4.2.2 and 4.2.4. Our DApp defines two cancer patient databases and referencing paths to patient data entries using the open-source HapiFHIR [45] public test server. Validation of the FHIR implementation is performed via regular expression parsing of the paths against the FHIR APIs [20].

Permissioning data access with token-based exchange, as discussed in Section 4.2.3. Our DApp also contains an *Access* smart contract that logs all user interactions and requests on the portal, *e.g.*, what resource is shared or no longer shared with which provider by whom and when. These access logs are structured as a mapping between user digital health identities (public encryption keys) and authorizations to custom-named access tokens (represented as a nested object associated with a *true/false* boolean value indicating if an access token access is granted for a provider). In case of an access revocation, authorization is set to *false* and the associated token is set to an empty value. The workflow of this process is shown in Figure 6.

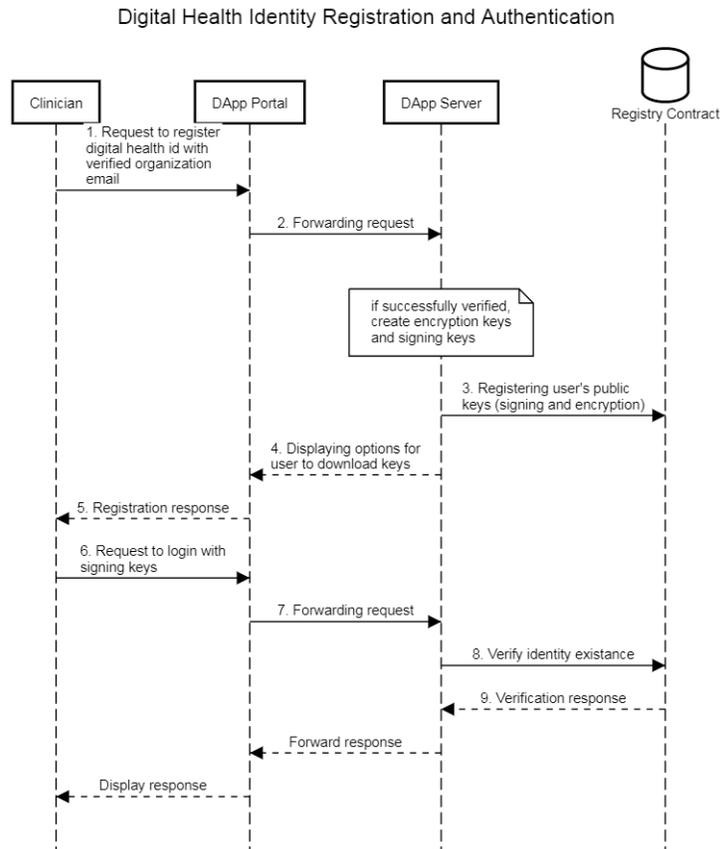


Figure 5: Workflow of the User Registration and Authentication Process in the FHIRChain DApp.

Maintaining modularity with the MVC pattern, as discussed in Section 4.2.5. The *view* component is a user interfacing portal that accepts provider user input, including registration and authentication credentials (corresponding keys) and data access information (tumor board member email to query, a reference pointer to securely access data, and approval/revocation of access). Figure 7 is a screenshot of our DApp, presenting the following features (1) display recent sharing events related to the user, (2) display reference pointer APIs created by logged in user and available actions, and (3) display all references shared with logged in user and the option to view data.

The portal then forwards the user requests along with data input to the

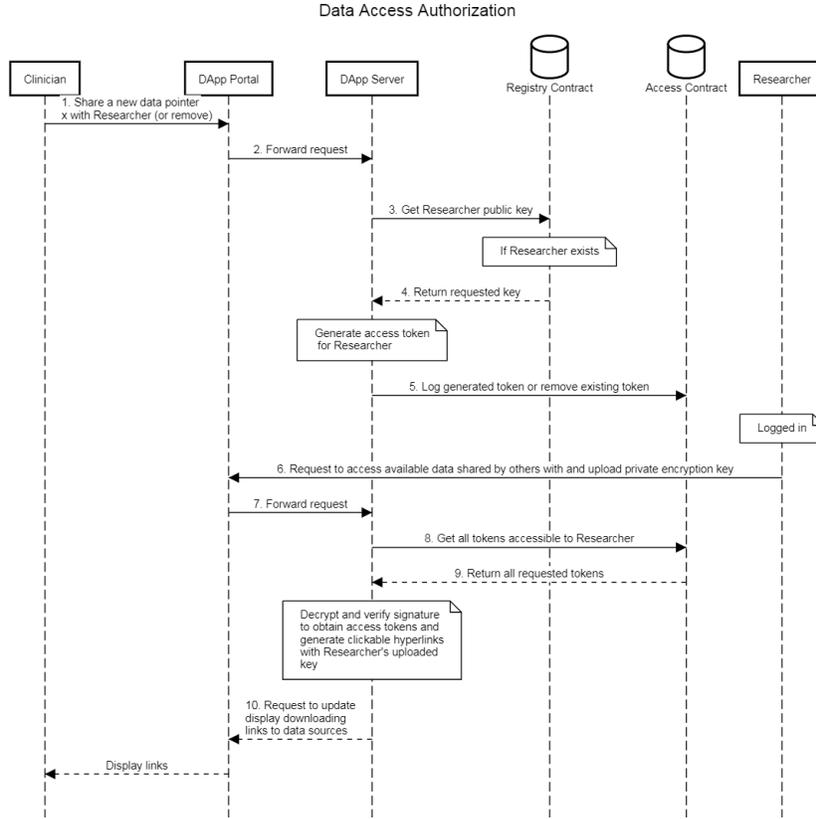


Figure 6: Workflow of Access Authorization in the FHIRChain DApp.

server component, where all the complex logic is encapsulated.

The *DApp server* is responsible for all functions and control logic, including verifying provider user email account, generating cryptographic keys, token creation via signing and encryption, token retrieval via decryption and signature verification, forwarding requests and delegating tasks between the *portal* and *blockchain*. The *blockchain* component is an independent *model* component containing two smart contracts for ubiquitous storing and persisting event logs of data access.

5.2. Benefits of Our FHIRChain DApp Case Study

To enhance modularity, we applied the "separation of concerns" principle [46] to decompose our DApp into independent components. FHIRChain employs a peer-to-peer API exchange protocol that references data pointers

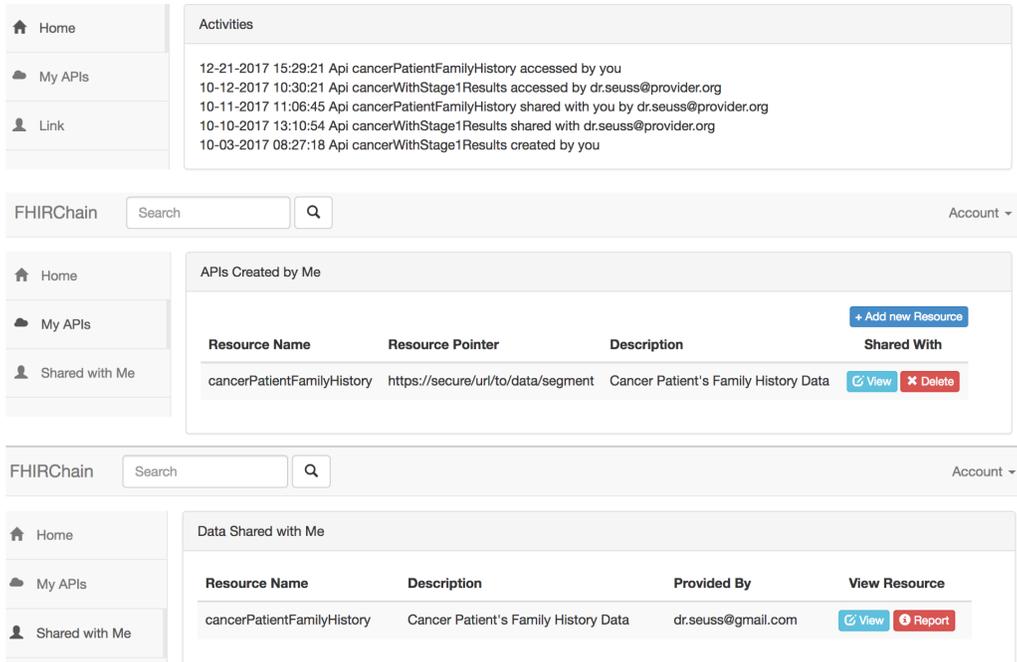


Figure 7: Screenshot of Our FHIRChain-based DApp User Interface.

stored in a smart contract on the blockchain. This design provides two key benefits:

- Exchanged information becomes lightweight and increases scalability since system performance remains the same regardless of the original size of the data,
- Data is not transmitted electronically across institutional boundaries, thereby reducing the risk of data being compromised.

Moreover, permissions to access a data source can be given or revoked at will by providers across various institutions regardless of their trust relationships. All data sharing and access activities are logged in a transparent history for auditability. In addition, exchanged data follows standard FHIR formats and thus can be displayed to providers with readability.

The adoption of public key cryptography provides “trust” to the participants in the following two ways:

- **Identifiability and authentication.** Given the computation power today, it is infeasible to impersonate a user without knowing their pri-

vate key, and the only way a user can be authenticated to use our service is to provide the correct private key paired with their public key registered on the blockchain. On the other hand, it is trivial to create a new public/private key pair in case of a user’s private key being lost or stolen. This “digital identity” approach has been successfully adopted in Estonias government and healthcare infrastructure [47].

- **Permission authorization.** With public key encryption securing their data reference pointers, users can trust that none other than the intended data recipient can view what they have shared. FHIRChain never shares the reference pointer with any user. Instead, RP is used to display the data content when it is decrypted with an authorized user’s private key. In addition, users can approve or revoke data access at any time, and the request takes effect immediately.

5.3. *Limitations of Our FHIRChain DApp Case Study*

Our FHIRChain DApp was designed with several assumptions in mind, so it incurs the following limitations:

- **Does not address semantic interoperability.** FHIRChain cannot address data exchange challenges related to semantic interoperability that are not yet fully captured by the FHIR standards. To provide semantics to clinical data, therefore, manual inspection and mapping of predefined ontologies from medical and health data experts are required, which remain the focus of future research in this space.
- **Cannot control clinical malpractice.** The intended users of FHIRChain are clinicians interested in collaboratively providing clinical decision support for remote patients. Our current design trusts that the data being exchanged using our DApp is not abused, misused, or unethically redistributed by users. In future work we will explore options to minimize these risks, such as tracking data credibility using cryptographic hashing or zero knowledge proofs [48] (ability to demonstrate the truth of a statement without revealing additional information beyond what its trying to prove [49]) along with each reference pointer. Naturally, clinical malpractice may still occur (just as in any other health IT system) since we cannot fully control these behaviors.
- **DApp deployment costs.** Unlike existing public blockchain, such as Ethereum, our DApp is developed using a private testnet that imposes

no interaction costs (*e.g.*, transaction fees). Our DApp would thus not be free of charge if deployed on a public blockchain. However, the convenience provided by a public blockchain may justify the cost of usage versus the costs of licensing, running, and maintaining a private clinical data exchange infrastructure.

To overcome these limitations in future work, we will deploy our DApp in a permissioned consortium blockchain platform with trusted parties to ensure consensus through a variation of proof-of-work that incentivizes mining with cryptocurrency rewards. For instance, [50] proposes to use aggregated data as mining rewards in their system, while MultiChain [51] enforces a round-robin mining protocol in their blockchain. With the ability to replace monetary incentives to maintain consensus on the blockchain, the cost to use this blockchain-based service will be lower in the long run, although the initial deployment may still be expensive.

6. Related Work

Due to the growing interest in using blockchain for health IT systems, related work has explored various blockchain-based design considerations and prototypes. This section summarizes this related work and compares it with our research on FHIRChain and DApps that provide collaborative clinical decision support for remote patients.

6.1. Conceptual Design Considerations

Krawiec et al [52] present several existing pain points in current health information exchange systems and the corresponding opportunities provided by blockchain technologies. They also discuss how blockchain can be leveraged in the health IT systems so that patients, health providers, and/or health organisations can collaborate. Nichol et al [53] present an analysis that assembles concepts in blockchain-related technologies and speculates on how blockchain can be used to solve common interoperability problems facing healthcare.

A white paper from IBM [54] takes a broader approach by highlighting the challenges in the healthcare industry and providing concrete use cases to showcase potential applications of blockchain technologies. More recently, Zhang et al provide design recommendations for creating blockchain-based healthcare systems with a case study [44] and propose assessment metrics

for evaluating such systems [35], which include a subset of the technical requirements defined in the ONC roadmap.

6.2. Blockchain Prototype Designs

Ekblaw et al [50] create a decentralized record management platform that enables patients to access their medical history across multiple providers. This platform uses a permissioned medical blockchain to manage authentication, data sharing, and other security properties. Their blockchain design integrates with existing provider data storage to enable interoperability by curating a representation of patient medical records. Medical researchers are incentivised to contribute to mining of the blockchain for collecting aggregated metadata as mining rewards.

Dubovitskaya et al [55] also propose a permissioned blockchain framework on managing and sharing medical records for cancer patient care. Their design uses a membership service to authenticate registered users using a username/password scheme. Patient identity is created using a combination of personally identifying information (including social security number, date of birth, names, and zip code) and encrypted for security. Medical data files are uploaded to a secure cloud server, with their access managed by the blockchain logic.

6.3. Differentiating Our Research on FHIRChain from Related Work

Several factors differentiate our research on FHIRChain from the related work described above. For example, we presented a blockchain-based framework called FHIRChain whose architectural choices were explicitly designed to meet key technical requirements defined by the ONC interoperability roadmap. Likewise, our FHIRChain-based DApp demonstrates the use of digital health identities that do not directly encode private information and can be replaced for lost or stolen identities, even in a blockchain system. In addition, FHIRChain provides a token-based access exchange mechanism that conforms with FHIR standards. Finally, we leverage public key cryptography to simplify secure authentication and permission authorizations.

7. Concluding Remarks

This paper described the FHIRChain prototype designed to provide collaborative clinical decision support for remote patients using blockchain technology and the FHIR protocol. Complemented by the adoption of public key

cryptography, the design addresses five key requirements provided by the ONC interoperability roadmap, including user identifiability and authentication, secure data exchange, permissioned data access, consistent data formats, as well as system modularity.

The following are the key lessons we learned thus far from designing and implementing our DApp based on FHIRChain:

- **FHIRChain can provide trustless, decentralized storage.** FHIRChain alleviates proprietary vendor-lock found in conventional health IT systems by leveraging its blockchain component as a decentralized storage. It enables sharing of clinical data without established trusts, while providing clinicians with secure and scalable collaborative care decision support.
- **FHIRChain facilitates data exchange without moving data.** The FHIR standards provide resource APIs to reference specific pieces of structured data. By adopting FHIR and combining it with blockchain technologies, FHIRChain creates lightweight reference pointers to siloed databases and exchange these pointers instead of actual data. For telemedicine clinics or clinics in rural areas in particular, this approach can overcome network limitations by enabling large-scale data sharing without electronic data transmission, in addition to reduce risks of compromised data.
- **Public key cryptography can be effective for managing digital health identity in data sharing.** FHIRChain creates public keys as digital health identities associated with each collaborating care entity (provider or organization administrator). The benefits to this strategy include: (1) *easy authentication* since a clinician only needs to provide their private key associated with their identity, (2) *integrity* since by signing the exchanged reference pointers FHIRChain can easily verify that it was provided by the signed provider and has not been modified, and (3) *remedy to lost or stolen keys* since a new key can be created easily to replace the old key and associate with the same user. However, there is a drawback to using digital identities for patients in a general clinical setting. Managing these identities—private keys—is challenging because unlike conventional passwords, private keys are hard to remember and require technical training for patients to manage their own keys. Nevertheless, there are approaches for managing private keys

for larger populations, such as using key wallets [56, 25] or embedding private keys to physical medical ID cards [57].

In summary, our FHIRChain-based DApp demonstrates the potential of blockchain to foster effective healthcare data sharing while maintaining the security of original data sources. The design of FHIRChain can be further extended to address other healthcare interoperability issues, such as coordinating other stakeholders (*e.g.*, insurance companies) across the industry and providing patients with easier (and secure) access to their own medical records.

Acknowledgements

The work presented in this paper was sponsored in part by funding from Varian Medical Systems.

References

- [1] M. Berman, A. Fenaughty, Technology and managed care: patient benefits of telemedicine in a rural health care network, *Health economics* 14 (2005) 559–573.
- [2] C. Castaneda, K. Nalley, C. Mannion, P. Bhattacharyya, P. Blake, A. Pecora, A. Goy, K. S. Suh, Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine, *Journal of clinical bioinformatics* 5 (2015) 4.
- [3] H. Singh, T. D. Giardina, A. N. Meyer, S. N. Forjuoh, M. D. Reis, E. J. Thomas, Types and origins of diagnostic errors in primary care settings, *JAMA internal medicine* 173 (2013) 418–425.
- [4] R. Kaushal, K. G. Shojania, D. W. Bates, Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review, *Archives of internal medicine* 163 (2003) 1409–1416.
- [5] G. D. Schiff, O. Hasan, S. Kim, R. Abrams, K. Cosby, B. L. Lambert, A. S. Elstein, S. Hasler, M. L. Kabongo, N. Krosnjar, et al., Diagnostic error in medicine: analysis of 583 physician-reported errors, *Archives of internal medicine* 169 (2009) 1881–1887.

- [6] D. B. Taichman, J. Backus, C. Baethge, H. Bauchner, P. W. De Leeuw, J. M. Drazen, J. Fletcher, F. A. Frizelle, T. Groves, A. Haileamlak, et al., Sharing clinical trial data: A proposal from the international committee of medical journal editors, *Annals of internal medicine* 164 (2016) 505–506.
- [7] E. Warren, Strengthening research through data sharing, *New England Journal of Medicine* 375 (2016) 401–403.
- [8] N. Geifman, J. Bollyky, S. Bhattacharya, A. J. Butte, Opening clinical trial data: are the voluntary data-sharing portals enough?, *BMC medicine* 13 (2015) 280.
- [9] G. E. Gross, The role of the tumor board in a community hospital, *CA: a cancer journal for clinicians* 37 (1987) 88–92.
- [10] J. Rieke, H. Bartelink, Telemedicine and its impact on cancer management, *European Journal of Cancer* 36 (2000) 826–833.
- [11] C. L. Marshall, N. J. Petersen, A. D. Naik, N. V. Velde, A. Artinyan, D. Albo, D. H. Berger, D. A. Anaya, Implementation of a regional virtual tumor board: a prospective study evaluating feasibility and provider acceptance, *Telemedicine and e-Health* 20 (2014) 705–711.
- [12] L. Levit, A. P. Smith, E. J. Benz Jr, B. Ferrell, Ensuring quality cancer care through the oncology workforce, *Journal of Oncology Practice* 6 (2010) 7–11.
- [13] M. Terry, Medical identity theft and telemedicine security, *Telemedicine and e-Health* 15 (2009) 1–5.
- [14] A. S. Downey, S. Olson, et al., Sharing clinical research data: workshop summary, National Academies Press, 2013.
- [15] G. Hripcsak, M. Bloomrosen, P. FlatelyBrennan, C. G. Chute, J. Cimino, D. E. Detmer, M. Edmunds, P. J. Embi, M. M. Goldstein, W. E. Hammond, et al., Health data use, stewardship, and governance: ongoing gaps and challenges: a report from amia’s 2012 health policy meeting, *Journal of the American Medical Informatics Association* 21 (2014) 204–211.
- [16] G. Hartvigsen, M. A. Johansen, P. Hasvold, J. G. Bellika, E. Arsand, E. Arild, D. Gammon, S. Pettersen, S. Pedersen, et al., Challenges in telemedicine and ehealth: lessons learned from 20 years with telemedicine in tromso, *Studies in health technology and informatics* 129 (2007) 82.

- [17] M. Maheu, P. Whitten, A. Allen, E-Health, Telehealth, and Telemedicine: a guide to startup and success, John Wiley & Sons, 2002.
- [18] R. LaRose, S. Stover, J. L. Gregg, J. Straubhaar, The impact of rural broadband development: Lessons from a natural field experiment, *Government Information Quarterly* 28 (2011) 91–100.
- [19] A. B. Bondi, Characteristics of scalability and their impact on performance, in: *Proceedings of the 2nd international workshop on Software and performance*, ACM, pp. 195–203.
- [20] D. Bender, K. Sartipi, HI7 fhir: An agile and restful approach to healthcare information exchange, in: *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, IEEE, pp. 326–331.
- [21] R. L. Richesson, J. Krischer, Data standards in clinical research: gaps, overlaps, challenges and future directions, *Journal of the American Medical Informatics Association* 14 (2007) 687–696.
- [22] R. Das, Does blockchain have a place in healthcare, *Forbes*. <https://www.forbes.com/sites/reenitadas/2017/05/08/does-blockchain-have-a-place-in-healthcare> (2017).
- [23] M. Mettler, Blockchain technology in healthcare: The revolution starts here, in: *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on*, IEEE, pp. 1–3.
- [24] A. Azaria, A. Ekblaw, T. Vieira, A. Lippman, Medrec: Using blockchain for medical data access and permission management, in: *Open and Big Data (OBD), International Conference on*, IEEE, pp. 25–30.
- [25] S. Nakamoto, *Bitcoin: A peer-to-peer electronic cash system*, 2008.
- [26] K. DeSalvo, E. Galvez, Connecting health and care for the nation: a shared nationwide interoperability roadmapversion 1.0, *Health IT Buzz* (2015).
- [27] V. Buterin, et al., *Ethereum white paper*, 2013.
- [28] D. Johnston, S. O. Yilmaz, J. Kandah, N. Bentenitis, F. Hashemi, R. Gross, S. Wilkinson, S. Mason, The general theory of decentralized applications, *dapps*, GitHub, June 9 (2014).
- [29] Direct project, Available at <https://www.healthit.gov/policy-researchers-implementers/direct-project>, ????

- [30] G. Greenspan, Blockchains vs centralized databases, Available at <https://www.multichain.com/blog/2016/03/blockchains-vs-centralized-databases>, Accessed 2017-12-31.
- [31] A. J. Menezes, P. C. Van Oorschot, S. A. Vanstone, Handbook of applied cryptography, CRC press, 1996.
- [32] A. Al Omar, M. S. Rahman, A. Basu, S. Kiyomoto, Medibchain: A blockchain based privacy preserving platform for healthcare data, in: International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage, Springer, pp. 534–543.
- [33] X. Yue, H. Wang, D. Jin, M. Li, W. Jiang, Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control, Journal of medical systems 40 (2016) 218.
- [34] R. Shea, Simple contracts are better contracts: What we can learn from the meltdown of the dao, Available at <https://medium.com/@ryanshea/simple-contracts-are-better-contracts-what-we-can-learn-from-the-dao-6293214bad3a>, Accessed 2017-12-31.
- [35] P. Zhang, M. A. Walker, J. White, D. C. Schmidt, G. Lenz, Metrics for assessing blockchain-based healthcare decentralized apps, in: 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), pp. 1–4.
- [36] D. Hardt, The oauth 2.0 authorization framework (2012).
- [37] Introduction to the isa, Available at <https://www.healthit.gov/isa/>, Accessed 2017-12-31.
- [38] M. Kim Futrell, Structured data (2013).
- [39] K. Kawamoto, T. Hongsermeier, A. Wright, J. Lewis, D. S. Bell, B. Middleton, Key principles for a national clinical decision support knowledge sharing framework: synthesis of insights from leading subject matter experts, Journal of the American Medical Informatics Association 20 (2013) 199–207.
- [40] A. Leff, J. T. Rayfield, Web-application development using the model/view/controller design pattern, in: Enterprise Distributed Object Computing Conference, 2001. EDOC'01. Proceedings. Fifth IEEE International, IEEE, pp. 118–127.

- [41] T. Ylonen, C. Lonvick, The secure shell (ssh) protocol architecture (2006).
- [42] H. Krawczyk, The order of encryption and authentication for protecting communications (or: How secure is ssl?), in: *Advances in CryptologyCRYPTO 2001*, Springer, pp. 310–331.
- [43] D. Crockford, The application/json media type for javascript object notation (json) (2006).
- [44] P. Zhang, J. White, D. C. Schmidt, G. Lenz, Applying software patterns to address interoperability in blockchain-based healthcare apps, *arXiv preprint arXiv:1706.03700* (2017).
- [45] Hapi-fhir, Available at <http://fhirtest.uhn.ca/>, Accessed 2017-12-31.
- [46] H. Ossher, P. Tarr, Using multidimensional separation of concerns to (re) shape evolving software, *Communications of the ACM* 44 (2001) 43–50.
- [47] R. M. Alvarez, T. E. Hall, A. H. Trechsel, Internet voting in comparative perspective: the case of estonia, *PS: Political Science & Politics* 42 (2009) 497–505.
- [48] C. Rackoff, D. R. Simon, Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack, in: *Annual International Cryptology Conference*, Springer, pp. 433–444.
- [49] G. Greenspan, Understanding zero knowledge blockchains, Available at <https://www.multichain.com/blog/2016/11/understanding-zero-knowledge-blockchains/>, Accessed 2017-12-31.
- [50] A. Ekblaw, A. Azaria, J. D. Halamka, A. Lippman, A case study for blockchain in healthcare: medrec prototype for electronic health records and medical research data, in: *Proceedings of IEEE Open & Big Data Conference*.
- [51] G. Greenspan, Multichain private blockchain white paper, Available at <https://www.multichain.com/download/MultiChain-White-Paper.pdf>, 2015.
- [52] R. Krawiec, D. Housman, M. White, M. Filipova, F. Quarre, D. Barr, A. Nesbitt, K. Fedosova, J. Killmeyer, A. Israel, et al., Blockchain: Opportunities for health care, in: *Proc. NIST Workshop Blockchain Healthcare*, pp. 1–16.
- [53] J. B. Peter B. Nichol, Co-creation of trust for healthcare: The cryptocitizen. framework for interoperability with blockchain (2016).

- [54] I. G. B. S. P. S. Team, Blockchain: The chain of trust and its potential to transform healthcare our point of view (2016).
- [55] A. Dubovitskaya, Z. Xu, S. Ryu, M. Schumacher, F. Wang, Secure and trustable electronic medical records sharing using blockchain, arXiv preprint arXiv:1709.06528 (2017).
- [56] S. Even, O. Goldreich, Y. Yacobi, Electronic wallet, in: Advances in Cryptology, Springer, pp. 383–386.
- [57] G. Anthes, Estonia: a model for e-government, Communications of the ACM 58 (2015) 18–20.