Enforcing Human Subject Regulations using Blockchain and Smart Contracts

Article · March 2018

DOI: 10.30953/bhty.v1.10

9 authors, including:

Olivia Choudhury
IBM Research
26 PUBLICATIONS **41** CITATIONS
SEE PROFILE

Amar Das
IBM
149 PUBLICATIONS **2,435** CITATIONS
SEE PROFILE

BLOCKCHAIN™ IN HEALTHCARE TODAY

Enforcing Human Subject Regulations using Blockchain and Smart Contracts

Olivia Choudhury,[1] Hillol Sarker,[2] Nolan Rudolph,[3] Morgan Foreman,[4] Nicholas Fay,[5] Murtaza Dhuliawala,[6] Issa Sylla,[7] Noor Fairoza,[8] Amar K. Das[9]

Authors

[1]Olivia Choudhury, MS, PhD, postdoctoral researcher, IBM Research, Cambridge, Massachusetts, USA. [2]Hillol Sarker, MS, PhD, postdoctoral researcher at IBM Research, Cambridge, Massachusetts, USA. [3]Nolan Rudolph, BS, research engineer, IBM Research, Cambridge, Massachusetts, USA. [4]Morgan Foreman, BSPsy, healthcare data scientist, IBM Research, Cambridge, Massachusetts, USA. [5]Nicholas Fay, MCS, healthcare data scientist, IBM Research, Cambridge, Massachusetts, USA. [6]Murtaza Dhuliawala, MCS, research engineer, IBM Research, Cambridge, Massachusetts, USA. [7]Issa Sylla, BA, research engineer, IBM Research, Cambridge, Massachusetts, USA. [8]Noor Fairoza, MS, Dev Ops Engineer, IBM Research, Cambridge, Massachusetts, USA. [9]Amar Das, MD, PhD, Director of Learning Health Systems, IBM Research, Cambridge, Massachusetts, USA

Corresponding Author

Olivia Choudhury, PhD, Olivia.Choudhury1@ibm.com

Keywords: Blockchain, Clinical Trial, "Common Rule", Data Security and Privacy, Distributed Ledger, Healthcare and Medical Research, Human Subject Regulations, Hyperledger Fabric, Protection of Human Subjects, Smart Contracts

Section: Use Cases/Pilots/Methodologies

Recent changes to the Common Rule, which govern Institutional Review Boards (IRB), require implementing new policies to strengthen research protocols involving human subjects. A major challenge in implementing such policies is an inability to automatically and consistently meet these ethical rules while securing sensitive information collected during the study. In this paper, we propose a novel framework, based on blockchain technology, to enforce IRB regulations on data collection. We demonstrate how to design smart contracts and a ledger to meet the requirements of an IRB protocol, including subject recruitment, informed consent management, secondary data sharing, monitoring risks, and generating automated assessments for continuous review. Furthermore, we show how we can employ the immutable transaction log in the blockchain to embed security in research activities by detecting malicious activities and robustly tracking subject involvement. We evaluate our approach by assessing its ability to enforce IRB guidelines in different types of human subjects studies, including a genomic study, a drug trial, and a wearable sensor monitoring study.

Keywords: Blockchain, Clinical Trial, "Common Rule", Data Security and Privacy, Distributed Ledger, Healthcare and Medical Research, Human Subject Regulations, Hyperledger Fabric, Protection of Human Subjects, Smart Contracts

The Federal Policy for the Protection of Human Subjects in Research (the "Common Rule") has recently undergone regulation revisions (the "Final Rule"). The Common Rule is the subpart A of the Department of Health and Human Services (DHHS) regulations, 45 CFR part 46, and outlines the basic provisions for Institutional Review Boards (IRBs), informed consent, and Assurances of Compliance. The Final Rule revision is expected to be effective July 2018 and has been the source of much ethical debate around consent management.

IRBs are established under the Common Rule to review and approve of research that is not directly conducted by a federal department. An IRB protects the rights and welfare of human research subjects recruited to participate in a research activity conducted at its affiliated institution. The board has the authority to approve, require modifications to, or disapprove of a research protocol based on the federal regulations and local policies at their institution. An IRB must ensure that the research protocol details the implementation of adequate informed consent and study procedures so as not to jeopardize the rights, safety, or wellbeing of the human subjects.[1]

Obtaining informed consent is one of the most sensitive and complex ethical issues in clinical research.[2] No entity may involve a person as a subject in research without obtaining the legally effective informed consent of the subject or the subject's legally authorized representative.[3] While there are many different types of consent, broad consent has been used by the research community for many years to collect, store, and use subjects' data and samples for unspecified future research.

The Final Rule, designed to address broader types of research, creates new regulations for establishing a framework of broad consent as a substitute for traditional informed consent. Prior to the Final Rule, there were only two alternatives for using identifiable data or biospecimens in a research study for which researchers had not secured study-specific consent: (1) obtaining an IRB waiver of consent or (2) removing personal identifiers. The final rule creates new exemption categories for the storage, maintenance, and research of data and biospecimens involving identifiable information under which broad consent is a condition for the exception. Exempt research in the new categories is required to undergo limited IRB review to ensure adequate privacy safeguards are in place for identifiable private information and identifiable biospecimens. To enforce broad consent, the healthcare institution has to maintain a tracking system of biospecimens approved for future research. The Final Rule also creates a provision that multi-site research will use a single IRB for the part of the research conducted within the United States, effective in 2020. An individual institution from this group, however, may still conduct an additional internal IRB review not limited to the standard regulatory guidelines. Finally, the Final Rule removes the requirement for ongoing research studies that received an expedited review to conduct a continuing review. This is also the case for studies that have completed interventions and are solely analyzing data or continuing observational follow up. [4]

A major challenge of enforcing the Common Rule and the Final Rule regulations is that, once a protocol is approved by the IRB, determining whether the protocol procedures are being violated is difficult. Institutions can easily become overwhelmed in the pursuit of compliance with the DHHS regulations, often due to manual record keeping.[5] Although federal regulators expect institutions to adopt better data management strategies, institutions continue to struggle on this front, leading to corrective actions becoming necessary. When subjects

withdraw consent from a study, their data is still allowed to be used by the researcher. For data that can be used for multiple studies, especially biological samples, it is difficult to guarantee that it will not be used for research that conflicts with the subject's values. Further, to prepare for continuing or final reviews of the research, investigators spend a significant amount of time compiling the data within a strict deadline. To mitigate these challenges and ensure proper enforcement of the guidelines and regulations set forth by the IRB, we propose a novel approach that leverages the distributed, shared ledger technology called blockchain.

Although blockchain gained prominence in the financial domain through the popular cryptocurrency Bitcoin,[6] it has been deemed as a promising solution for several applications in healthcare and medical research.[7-10] Many healthcare applications have proven the benefits of this technology in building a secure platform for managing and analyzing sensitive healthcare data. MedRec[11,12] is a decentralized record management system to manage electronic medical records using Ethereum (an open-source, public, blockchain-based distributed computing platform).

The Orange Consent Management Service[13] offers a consent management system for eHealth based on Hyperledger Fabric (a permissioned blockchain infrastructure). In this paper, we demonstrate the use of private blockchain in designing a system that enforces the requirements of IRB protocols, as defined in the Common Rule and the Final Rule. The system incorporates the functionalities and regulatory constraints in the smart contract, stores consent and sensitive information on the ledger, and monitors risks and generates results for continuous review using the immutable transaction log. By being an integral part of how data in a research study are collected, stored, managed, and analyzed, it insures that the regulations are consistently enforced across all the entities involved, including investigators, subjects, and research organizations.

HUMAN SUBJECT REGULATIONS

A. IRB Protocol Review

After submission to an IRB a research protocol goes through one of three types of review: (1) exempt, (2) expedited, or (3) full board. The type of review is determined by the level of risk based on certain categories defined in 45 CFR 46.101(b) and 45 CFR 46.110. Exempt review can occur for protocols that involve anonymous or publicly-available data, including surveys, retrospective chart reviews, or analysis of specimens without subject identifiers. Research that falls under exempt review still requires registration with the IRB. An expedited review can be used by an IRB when the research protocol involves no more than minimal risk to a human subject. Examples of expedited research protocols include studies collecting samples of DNA, voice recordings, or specimens with subject identifiers. All other research that does not fall into the categories of exempt and expedited review is subject to a full board review.

The IRB conducts a continuing review of research protocols that underwent full board reviews at intervals deemed appropriate to the degree of risk, but not less than once per year. The continuing review examines any changes or negative instances that occurred in the research, including withdrawals, adverse events, and unanticipated problems. If the research has been completed, the continuing review will become the final report. If the research needs to extend past the approval period of the IRB, the protocol must be resubmitted and approved for renewal.

B. IRB Protocol Requirements

The criteria for IRB approval of research includes the following seven requirements, according to 45 CFR 46.111(a):

1. Risks to subjects are minimized by using study procedures which are consistent with sound research design and which do not unnecessarily expose subjects to risk.
2. Risks to subjects are reasonable with respect to anticipated benefits, based only on the risks and benefits resulting

from the research, not collateral therapies.
3. Selection of subjects for the study is equitable, given the purpose of the research and the setting in which it will be conducted.
4. Informed consent from the subject, per CFR 46.116(a), must include the purpose of the research, expected duration of the subject's participation, and the details of the procedures to be followed. The basic elements also cover the benefits and foreseeable risks to the subjects, disclosure of appropriate alternative procedures or courses of treatment, and a description on how confidentiality of records identifying the subject will be maintained. CFR 46.116(b) describes additional elements in informed consent such as unforeseeable risks, anticipated circumstances under which the subject's participation may be terminated by the investigator, consequences of a subject's decision to withdraw from the research, procedures for orderly termination of participation by the subject, and approximate number of subjects involved in the study.
5. Informed consent must be documented using a written consent form approved by the IRB and signed by the subject or the subject's legally authorized representative. Under a broad consent model, this consent then carries over to the secondary usage of the data as it is de-identified and requires no further consent for secondary usage.
6. When appropriate, the research plan must make provisions for monitoring the data collected to ensure the safety of subjects.
7. When appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of data.
8. In addition, CFR 46.111(b) mandates safeguards are included for subjects that are likely to be vulnerable to coercion or undue influence, such as children, prisoners, pregnant women, mentally disabled persons, or economically or educationally disadvantaged persons, to protect the rights and welfare of these subjects.

C. Informed Consent

The process of informed consent is a dynamic and ongoing process. Even though consent for the research is given, subjects are not obligated to continue the research until the completion of the study period or study activities. Once consent is withdrawn, researchers are obligated to no longer contact the subject about collecting more data. The researchers are, however, still able to use the data collected up to the point of withdrawal. Consent may also be required again if any major changes happen to the subject's ability or willingness to take part in research or if any new major information has become known during the conduct of the study.[14]

When consent is needed, there are multiple types that can be used for research studies. The most customary form of consent is specific consent, which is only applied to one research study. Biobank research may involve subjects participating in multiple research questions or studies. Obtaining specific consent from each time new research questions arise can be challenging. As a result, broad and blanket consent are used. Blanket consent allows research without any restrictions to data, while broad consent allows for a wide range of future studies, subject to specified restrictions. A major issue with blanket consent is that the subject's data may be used for studies that conflict with the individual's values. Meta-consent goes a step further and allows individuals to express preferences for the consent they want to give for certain types of research to ensure the research aligns with their values. Dynamic consent facilitates the consent process with two-way ongoing communication with researchers and subjects through on-line platforms.[15]

BLOCKCHAIN TECHNOLOGY

A. Public versus Private Blockchain

When Bitcoin's[6] rising popularity showcased there could be trust in a network despite the fact that no particular node could be trusted, it paved the way for many further implementations and uses of open, public blockchains. For any public or permissionless blockchain network to function, some consensus model must be implemented such that potentially malicious or faulty actions of a compromised user will be negated by the participation of the remaining users. Bitcoin is successful due to its use of a Proof of Work (PoW) consensus model.[16] While PoW has been relied upon for creating a trustworthy system to ensure malicious users cannot interfere or tamper with the network for personal gain, it comes at a cost of high energy usage and slow transaction rates. This is due to the fact that "miners" are tasked with solving a computationally intensive puzzle in order to add a block of transactions to the chain, and they are rewarded in cryptographic tokens upon success. Other proposed methods of instilling trust in a network include Proof of Stake (PoS)[17] and Proof of Elapsed Time (PoET)[18] Proof of Stake eliminates the energy usage and transaction rate sacrifices of PoW, but still requires the network to incorporate a cryptographic token. Participation in such networks comes at a cost, in tokens, to the user. Although PoET gets away from the energy or token cost of PoW and PoS, it still cannot achieve immediate finality of transactions, as a fork in the blockchain can temporarily exist before being solved algorithmically. Its reliance on the literal passing of time also means it cannot realize the high transaction rate possible with other models.

The above models solve trust in a permissionless network. However, private blockchains operate under much different circumstances.[19] These are based on permissioned networks, which restrict who can join the network, read the ledger, propose transactions, and participate in consensus. Therefore, permission to participate in the network can be limited only to known, trusted entities. If a particular use case is suitable for a closed network rather than a public one, advantageous simplifications can be made. Although a cryptographic token can be implemented if such a need exists, it is not required to incentivize mining (PoW) or to prove one has a financial stake in the network (PoS). Transactions in a permissioned network can be considered immediately final, as the possibility of having to resolve a fork is nonexistent. Due to immediate finality as well as expedited consensus, their transaction rates are higher than any existing public blockchain model can produce.

In addition to increased transaction speed and finality, permissioned networks also benefit from improved privacy. The authors of MedRec[11,12] pursued a blockchain framework built on top of Ethereum,[20] a PoW-based, permissionless network, to develop a solution for electronic medical record management. While it offered many improvements over traditional systems, an acknowledgement was made that frequency-based analysis of even encrypted data transactions on the public chain could provide insights on network activity to unwanted third parties. This issue is solved by the inherent design of permissioned networks, where only approved nodes could view the underlying activities.

B. Implementations of Private Blockchain

Within the domain of private or permissioned blockchain frameworks, there are several options to consider. Many of the first pursuits in this area were created to become solutions in the financial sector. These networks, including Quorum,[21] Ripple,[22] and Chain.[23] are open-source and very promising in their own right, but were not created with healthcare-specific considerations in mind. Many of them, such as Quorum, fork the popular permissionless blockchain network Ethereum or other cryptocurrency-focused designs, and add permissioning and other functionalities as needed. In contrast, Hyperledger Fabric [24] was designed from the ground up to enable permissioned, secure use of distributed ledger technology.[25] It allows for modular inclusion of different consensus models and membership service providers, as well as the creation of private channels within a network that only

specified participants operate on. This flexibility in design and privacy makes it a promising solution for applications in medical and healthcare research. In prototyping a system to store patient's consent for sharing sensitive health data with health practitioners, researchers adopted Hyperledger Fabric to build a platform for secure and efficient management of sensitive data.[13]

C. Hyperledger Fabric

Hyperledger Fabric, introduced by The Linux Foundation, is one of the primary private blockchain frameworks currently available.[24] It is based on a permissioned network comprising only interested stakeholders as participants. This restricts anyone from joining the network, updating the ledger, or initiating transactions. A network typically consists of multiple nodes, a smart contract implementing the business logic, and a ledger maintaining transaction log and its state as a key-value store. Nodes are logical entities running on a physical server that can be maintained by participants. They can be categorized into client, peer, and orderer nodes. Client nodes invoke transactions and are connected to both peers and orderers. Peer nodes maintain the ledger and receive state updates in the form of blocks. They can also act as endorsers for verifying and validating a requested transaction. Unlike public blockchains, Hyperledger Fabric employs an endorsement policy that defines the necessary conditions for a valid transaction. A transaction is approved only when it acquires endorsement signatures from designated endorsers, as defined in the policy. Orderers support communication between clients and peers. When a client invokes a transaction request, the message is broadcasted to all peers. On receiving signature from all the endorsers, the orderer broadcasts a message to all peers to update their copy of ledger. This is further illustrated in Figure 1.

Hyperledger Fabric provides an additional layer of security by creating private channels between members of the network. Each channel maintains a ledger that can only be accessed by the members of that channel. Since it does not rely on the compute-intensive Proof of Work protocol to attain consensus, the overhead of transactions is significantly reduced. This helps the system to be scalable when increasing the workload or adding users.[26] Since the immutable transaction log records all transaction requests, the system can also track unauthorized or malicious transactions initiated by nodes. We further describe the different components of Hyperledger Fabric

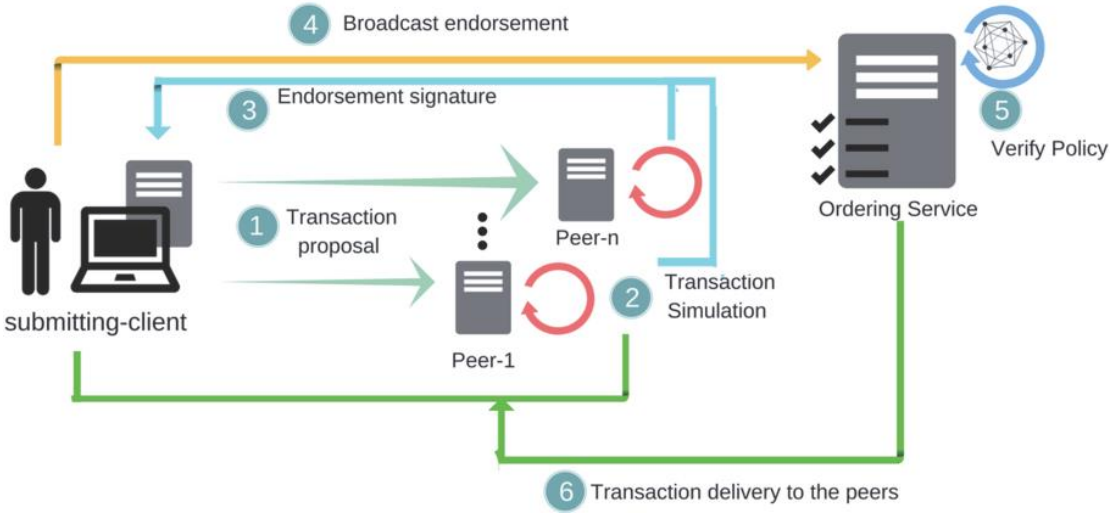

Figure 1: Outline of transaction flow in Hyperledger Fabric. It depicts a use case where: (1) A client node sends a transaction proposal to the endorsing nodes (for this example, all the peer nodes act as

endorsers). (2) Endorsers simulate the transaction and generate endorsement signature. (3) Client node collects endorsement signatures. (4) Client node sends the signatures to the ordering service. (5) Ordering service verifies signatures and broadcasts a message to the peer nodes to update their ledger.

1. *Smart Contract.* Smart contract, referred to as chaincode in Hyperledger Fabric, is the business logic written in a machine readable and executable language. The features offered by the smart contract are agreed upon by relevant parties and define the functionalities afforded to the members of the network. It enables fine-grained access checks to verify the authenticity of proposed transactions. Unlike a regular database, the ledger can only be accessed or updated through the functionalities defined in the smart contract. Once the smart contract is installed and instantiated on a network, it can only be updated upon mutual agreement between the parties.
2. *Transaction.* Transactions encompass the act of invoking the functionalities outlined in the smart contract. Users can initiate transactions to read or write to the ledger with the help of application program interface (API) calls. A request is sent as a transaction proposal to endorsing peers for simulation. The results of the simulation are collected and sent to the orderer node for verification. Since all the peers in a network maintain an identical copy of the ledger, simulations should produce the same results. Once the results are verified, the orderer broadcasts a message to all peers to update their ledger. Transactions are considered successful upon distribution by the orderer, indicating the proposal was successfully simulated and accepted upon validation. Although an unsuccessful transaction cannot update the ledger, the request is logged in the history, in support of an auditable system. Figure 1 depicts the scenario of transaction flow in Hyperledger Fabric.
3. *Endorsement Policy.* The endorsement policy informs the committing peers on the validity of a proposed transaction. The decision is based on the contents of the collected simulation results, endorsing peer signatures, and authorization certificates. The policy defines a list of endorsing peers and number of endorsements required to validate a proposed transaction. Once all the criteria are met, the results of the simulated transaction are propagated to the relevant peers, thus updating their ledgers and data to ensure consistency. Hyperledger Fabric further allows more granular access control that requires involvement at the level of participant, rather than the high-level organization. Based on the use case, the system can be designed to meet both the requirements.

System Design

To address known challenges with IRB protocol implementation and ongoing enforcement, we designed a blockchain-based system in which IRB protocols are integrated authoritatively as steps in the research process. In our system, all transactions pertaining to consent management, data collection, and data sharing are executed through a smart contract's programming logic. Interactions with the smart contract make the system secure, efficient, and auditable, thereby ensuring reliable enforcement of IRB guidelines.

A. Consent Management

Our design includes a simple interface for subjects to interact with the system. This consent may also include granular access rights for data. They can specify the users, duration, and type of data they intend to share. Once consent is obtained from the subject, the details are securely stored on the ledger. Since consent is typically collected and maintained by the principal investigator or coordinator of the study, they can have direct access to this information on the ledger. All other entities must request access to the data collected during the study through an access-control server, which communicates with a consent server to validate the authenticity of the request. If a subject withdraws from a study or revokes permission to share data, the corresponding information on the

ledge is updated through the smart contract. The updated consent information gets reflected upon subsequent verification by the consent server. In Figure 2, we demonstrate our design of informed consent management using blockchain technology.

B Data Collection

After receiving consent, the principle investigator (PI) or research coordinator collects data from subjects. Methods of data collection are tailored to the research objectives of the specific study in question, dictated by the study protocol, and enforced by the smart contract. This may involve a set of multimodal measurements, or, in the case of longitudinal research, repeated measurements. Personal information that can potentially identify a subject are retrieved and stored separately. Since blockchain provides a secure platform, we use the ledger to store protected health information (PHI) attributes and consent information. This ensures that only authorized entities in the network have access to data. Due to the distributed nature of the ledger, data is replicated on multiple nodes in a channel. This can cause an overhead when storing large volumes of data on the ledger, often experienced in genomic studies. To address this challenge, we store the high-volume data collected during the study in a database. Access to this database is restricted by an access-control server. Depending on the type of consent, identifiable information from the data is removed prior to storing it in the database or repository.

In the case of verbally administered interviews, the identity of subjects may be revealed by their response to a questionnaire. Our proposed system generates a unique key for each subject and stores it on the ledger. Before transferring data to the database, each record is purged by replacing the real identity with this key, making the data anonymized for downstream analysis. This is also relevant to clinical trial of drugs, where subjects are typically divided into a control group and a treatment group. Once the study coordinator records outcome measures and cases of adverse effects, the identifiable information can be saved on the ledger and the demographics and outcome measures can be stored in the database. Focus group is another conventional method of qualitative health research, where the study coordinator records audio or video of subjects. Such recordings pose a risk of identifying the individuals from their speech traits. The study coordinator or a speech-to-text module should annotate recordings and replace any mention of a specific person's name by his or her unique, anonymous key. In some studies, certain types of data may raise an additional risk of privacy. For instance, in genomic studies, a set of specific genes can be used to uniquely identify a subject. For such cases, we use an intermediate step of de-identification to obfuscate identifiable traits before storing the data. Pre-processing may require one-step or two-step de-identification. Generalization, suppression, randomization, and sub-sampling are some of the widely-used techniques of de-identification.[27]

C. Data Sharing

Once data are collected and stored in a database, a third-party research organization may request access to these data through the access-control server. For blanket consent, the least restrictive type of consent, the access-control server does not restrict data access. For all other consent types, it conveys the request to the consent server, which verifies the access rights stored on the ledger for authentication. For valid requests, the access-control server queries the database to retrieve data of consented subjects, thereby sharing the data only with intended research organizations while maintaining a strict consent system. If a subject withdraws consent for data sharing, the updated access rights on the ledger will restrict further data access. An auditor requesting statistics of the collected data will follow a similar procedure in order to have the data shared with them. The method of secondary data sharing implemented by our blockchain-based solution is depicted in Figure 2.

Hyperledger Fabric offers an additional layer of data security through endorsement policy, which authenticates a transaction request. Such a policy

defines the set of endorsing nodes that must simulate a requested transaction to return a response and a signature for proof. For transactions that involve sharing data with third-parties, the nodes corresponding to research subjects and investigators can act as endorsers. Even if a third-party acquires necessary access rights for secondary data sharing, any discrepancy in the simulated results will immediately halt data access. The rules can be tuned to conform to the requirements of different types of transactions. The endorsement policy coupled with granular access control empower the data owners to restrict unauthorized data sharing.

EVALUATION

In this section, we describe how our proposed blockchain-based system enforces the major requirements of IRB protocols. These mechanisms are evaluated for coverage and completeness against the specific IRB protocols from three different research studies: a genomic study,[28] a drug trial,[29] and a wearable sensor monitoring study,[30] Our scheme is shown to be sufficient for enforcing the IRB guidelines of all three studies.

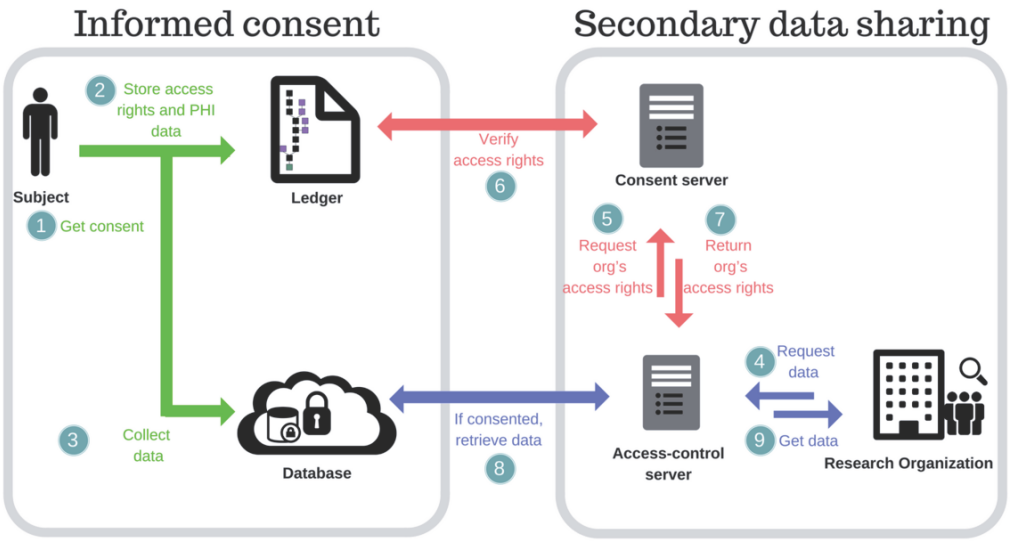

Figure 2: Design of our proposed blockchain-based system for enforcing IRB protocol. The workflow involves: (1) Subjects consenting to participate in the study. (2) Storing access rights for data sharing and PHI attributes of subjects on the ledger. (3) Storing data collected from subjects during the study in a database. (4) A third-party or research organization requesting to access data through access-control server. (5) Access-control server requesting consent server to verify access rights of the organization. (6) Consent server using smart contract to retrieve access rights stored on the ledger. (7) Consent server responding to access-control server with the access rights. (8) For a valid data request, access-control server fetching data from database. (9) Access-control server returning the data to research organization.

A. Time and Place of Study

Protocols often dictate when and where a study can take place. We enforce these requirements through the smart contract. Before conducting a study, we programmatically set limits based on the approved time window for different phases of the study. Since each subject in the study corresponds with an entry on the network's ledger, the number of subjects is controlled by setting a fixed limit when the study-specific contract is instantiated. Likewise, to enforce geographic limits we set additional checks in the smart contract to accept or reject subject enrollment based on their location. For example, a subject or the representative, given that they are providing reliable information, may enter a zip code which is checked against a set of pre-defined accepted zip code values. This

verification will dictate whether enrollment can successfully progress.

B. Subject Selection

To satisfy regulations on subject selection, as described in the third requirement of clause CFR 46.111(a), inclusion and exclusion criteria for participation in the study can be checked in the smart contract prior to enrollment. The smart contract verifies specific login information against the ledger to determine if the attributes of that individual satisfy enrollment criteria for a given study. This provides an immutable provenance of the subjects and the data collected for a research study. Duration of recruitment is also defined in the smart contract. This given interval is set upon registration and made transparent to give the users a deeper view of what they are involved in. As one of the additional elements of informed consent in the fourth requirement of CFR 46.111(a), the number of subjects is restricted by setting an upper bound within the enrollment functionality of the smart contract. At the time of registration, each subject is presented with the terms and conditions of participation in the study. As required by the fifth requirement of CFR 46.111(a), a copy of this consent is also stored in the blockchain network. This transaction is recorded for future need, with no possibility of being tampered or altered without consent. There is a set of criteria agreed upon and implemented during initial enrollment that specify the ability to withdraw from a given study. For CFR 46.116(b), the withdrawal criteria are also enforced through the smart contract, where a subject is informed of his or her eligibility to withdraw and any criteria that must be met to do so. The transactions can track if certain incentives or benefits offered to a subject were discontinued as a consequence of withdrawal from the study.

C. Informed Consent Management

As per CFR 46.116(a), the IRB protocol requires the investigator to define the duration of a subject's involvement. The smart contract is programmed such that all the functionalities implementing a subject's involvement, such as enrollment and informed consent, are operative within a specified time interval. To ensure the fifth requirement of CFR 46.111(a) are met, the consent granted by subjects to participate in the study and share their data are stored on the ledger. This consent is verified by the consent server for an organization requesting data access. Figure 2 illustrates the underlying method of informed consent implemented by our system. The digital record keeping through the smart contract and ledger offers an additional security measure. Based on the type of consent defined in the protocol, subjects provide necessary information pertaining to it. For example, if the subjects grant specific consent, their data can only be used for that research study. Accessing their data for any other study is restricted by the system. If a research organization not enlisted as an intended user of the data tries to access it, it will be blocked by the access-control server and stored as an unauthorized transaction request in the transaction log. This is further explained in the secondary data sharing part of Figure 2. In dynamic consent, subjects may also revoke consent to partially or completely share their data. Our proposed system supports this by allowing or limiting data access accordingly. The consent further indicates if the subject has agreed to the investigator contacting them for secondary research or changes in the current study. The access-control server, consent server, and the smart contract embedded in our proposed system enforce the requirements for all consent types.

D. Secondary Data Sharing

De-identification or removal of personal information is often required prior to collecting, storing, sharing, or analyzing data. As mentioned in the seventh requirement of CFR 46.111(a), investigators must declare the underlying method of de-identification, storing protected health information (PHI), and duration of storage. Since a private blockchain network restricts unauthorized access to ledger, all the sensitive information is securely saved on the ledger, which can only be accessed by the investigator. The ledger also stores the mapping

between code and identifiable information. In addition, subjective assessments, such as HIPAA (Insurance Portability and Accountability Act of 1996) compliance, taken by investigators, PI, or study coordinators are recorded here. For more restrictive type of consent, the coded, de-identified data collected from subjects can be stored in a database. Certain data formats may require intermediate steps to further obfuscate sensitive information before making it available on the database. For example, in genomic studies, a set of genes that may uniquely identify a subject can undergo a two-step de-identification process before sharing it for public usage. In the case of more generic consent types, such as broad or blanket consent, the data may be stored in a repository without de-identification.

For the most restrictive consent scenario, when a third-party research organization requests data through the access-control server, they communicate with the consent server to retrieve corresponding access rights previously consented by the subjects. The smart contract verifies if the request was initiated by an intended organization for a valid data type and duration. If all the criteria are met, the consent server responds to the access-control server with the list of consenting subjects and the data type. In addition to this, the endorsers' signatures collected during the phase of transaction endorsement provide a further means of authenticity. An approval from both granular access rights and endorsement policy allows the organization to fetch data from the database. Hence, our blockchain-based system guarantees secure storage of sensitive information and authorized access of permitted data, as required by the IRB guidelines.

E. Safety Measures

Collateral benefits are a natural side effect of research studies, which are often beneficial to the recipient. All test results that are generated from a specific study are recorded with provenance and maintained for each subject. If any underlying condition that is found as a side effect of the study or the regiment of the study provides positive outcomes to the patient, it is recorded using the underlying technology.

Risk assessment and adverse effects are major concerns for ensuring safety during and after a research study as they can often be surprising, and sometimes, detrimental. Each transaction is timestamped in the ledger which can record any incident of risk or adverse effect that occur within the timeline of a study, thus addressing the requirements of the sixth requirement of CFR 46.111(a). These data can later be referenced as sparsely labeled, time series data for analysis to be performed on top of the timeline of each trial conducted. This final form of reporting in conjunction with continuing review ensure safety measures are enforced, and any negative effects are handled immediately. Any self-reported adverse effects by the subjects can also automatically trigger an alert to the study coordinators to allow for just-in-time intervention and close monitoring. As required by CFR 46.111(b), additional safeguards for vulnerable subjects can be recorded and tracked during the course of the study. Such measures may also incorporate a validation or endorsement by the subjects to ensure they received the desired treatment or safeguard measures.

F. Continuing Review

Principal investigators of on-exempt studies, that did not undergo expedited review, must compose and send continuing reviews to the IRB board at predefined intervals. We compose functionalities in the smart contract to query and organize data required for the review, as well as generate automated reports. While defining the network's ledger, we include all the pertinent data fields that can be potentially used for later submission. Since transactions that alter the ledger are tracked in an immutable and sequential order, our system captures reliable timelines of subjects' detailed involvement that can be submitted for review. For example, the IRB must be informed of any subject withdrawal, including the reason.[31] When composing the continuing review, the smart contract defines ledger queries which reveal all

subjects (ledger entries) who have withdrawn consent and provide details surrounding their withdrawal.

As an additional consideration, the institution conducting project overview can be added as a network node so that it can access the shared ledger to directly make queries and propose changes. If the IRB chooses to enforce suspension or termination due to conclusions reached from the continuing review, or a missed or improper submission, the smart contract can be written such that the IRB can revoke all subjects' participation consent on the ledger. While traditionally, IRB intervention is recognized at an administrative level, this proposed method offers an accompanying technical enforcement to provide further trust such that the IRB's decision is respected.

DISCUSSION

The recent amendments to The Federal Policy for the Protection of Human Subjects or the Common Rule have sparked controversy around the status quo of security and privacy measures. The revisions defined in the Final Rule necessitate more comprehensible and transparent informed consent management, revised concept of identifiable biospecimens, consent for research involving biospecimens and identifiable data, and considerable measures for privacy safeguards. Moreover, it is often difficult to track and ensure that the rules and regulations approved by the IRB are imposed throughout the study. To alleviate these challenges, we propose a novel data management framework based on blockchain technology that implements and enforces the requirements of human subject regulations, as outlined in the IRB protocol. We provide an overview of blockchain technology, distinguishing between public and private blockchain frameworks to elucidate our adoption of a private blockchain implementation in designing the system. We describe how we leverage smart contracts to enforce the requirements of IRB guidelines, the ledger to securely store sensitive data and prevent unauthorized access, and the transaction history to generate statistics and track activities.

Although blockchain technology is a promising solution for healthcare applications, its adoption in this community is still at a nascent stage. To realize its full potential, we must continue exploring different implementations of this technique and their applicability in the domain of medical and healthcare research. The effectiveness of our blockchain-based solution largely depends on all the entities of a research study adopting and embracing this new technology. Hyperledger Fabric allows updating the smart contract upon successful endorsement from network participants to accommodate changes required by the IRB during a study. However, any inconsistency in the data previously stored on the ledger or database must be resolved. Setting up a private blockchain network requires a comprehensive knowledge of the underlying technique, which can be a challenging task.

For future work, we intend to assess the scalability of our system when increasing the number of nodes and workload. We would also like to extend the system for active involvement of the regulatory board to automate and log communication with them. Although our proposed system currently becomes effective after the approval of IRB protocols, it may help to also include the IRB review process as a part of the system. Finally, if permitted by the protocol, we plan to include additional functionalities in the smart contract that allow sharing study results with subjects and their healthcare providers.

Acknowledgement

We would like to thank Daniel Gruen at IBM for valuable comments and suggestions. We are also thankful to Woong A. Yoon, John Paul Filippone, and Farhan Arshad at IBM for providing technical assistance for the system described in this paper.

Funding Statement

There was no public or private funding provided in the creation of this work.

Conflict of Interest

The authors are employees of IBM, an organization with financial interest in the subject matter discussed in the article. No other potential conflicts are reported.

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.

References

1. hhs.gov. Approval of Research with Conditions: OHRP Guidance (2010) [Internet]. HHS.gov. 2016. Available from: https://www.hhs.gov/ohrp/regulations-and-policy/guidance/guidance-on-irb-approval-of-research-with-conditions-2010/index.html#section-a
2. Nijhawan LP, Janodia MD, Muddukrishna BS, Bhat KM, Bairy KL, Udupa N, et al. Informed consent: Issues and challenges. *J Adv Pharm Technol Res.* 2013;4(3):134–40.
3. Gupta U. Informed consent in clinical research: Revisiting few concepts and areas. *Perspect Clin Res.* 2013;
4. 82 FR 7149. Department of Health and Human Services. *Code Fed Regul*. 2018 Jan;7149–7274.
5. Portier W, Dunne C. Current Challenges and Opportunities in Clinical Research Compliance. *Ochsner J*. 2006;6(1):21–4.
6. Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008;9.
7. Mettler M. Blockchain technology in healthcare: The revolution starts here. In: 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom). IEEE; 2016. p. 1–3.
8. Yue X, Wang H, Jin D, Li M, Jiang W. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J Med Syst. 2016;40(10):218.
9. Irving G, Holden J. How blockchain-timestamped protocols could improve the trustworthiness of medical science. F1000Research. 2016;5.
10. Nugent T, Upton D, Cimpoesu M. Improving data transparency in clinical trials using blockchain smart contracts. F1000Research. 2016;5.
11. Ekblaw A, Azaria A, Halamka JD, Lippman A. 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. 2016.
12. Azaria A, Ekblaw A, Vieira T, Lippman A. Medrec: Using blockchain for medical data access and permission management. In: Open and Big Data (OBD), International Conference on. 2016. p. 25–30.
13. Genestier P, Zouarhi S, Limeux P, Excoffier D, Prola A, Sandon S, et al. Blockchain for Consent Management in the eHealth Environment: A Nugget for Privacy and Security Challenges. *J Int Soc Telemed eHealth*. 2017;5:21–4.
14. Title 45 Part 46. Department of Health and Human Services. Code Fed Regul. 2009 Jul;
15. Budin-Ljøsne I, Teare HJA, Kaye J, Beck S, Bentzen HB, Caenazzo L, et al. Dynamic Consent: A potential solution to some of the challenges of modern biomedical research. *BMC Med Ethics.* 2017;18(1).
16. Dwork C, Naor M. Pricing via Processing or Combatting Junk Mail. In: *Advances in Cryptology — CRYPTO' 92.* 1992. p. 139–47.
17. Vasin P. BlackCoin's Proof-of-Stake Protocol v2. Self-published. 2014;
18. Proof of Elapsed Time (PoET) [Internet]. 2017. Available from: https://intelledger.github.io/
19. Baliga A. Compliance Oversight Procedures for Evaluating Institutions. 2017.
20. Buterin V. A next-generation smart contract and decentralized application platform. *Etherum*. 2014;(January):1–36.
21. Quorum [Internet]. Available from: https://www.jpmorgan.com/global/Quorum
22. Ripple [Internet]. Available from: https://ripple.com/
23. Chain [Internet]. Available from: https://chain.com
24. Hyperledger Fabric. 2017.
25. Vukolić M. Rethinking Permissioned Blockchains. Proc ACM Work Blockchain, Cryptocurrencies Contract - BCC '17. 2017;3–7.

26. Li W, Sforzin A, Fedorov S, Karame GO. Towards scalable and private industrial blockchains. In: Proceedings of the ACM Workshop on Blockchain, Cryptocurrencies and Contracts. 2017. p. 9–14.
27. Lo B. Sharing clinical trial data: maximizing benefits, minimizing risk. *JAMA.* 2015;313(8):793–4.
28. Dressler LG. Disclosure of research results from cancer genomic studies: state of the science. *Clin Cancer Res.* 2009;15(13):4270–6.
29. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med.* 2013;158(3):200–7.
30. Silva de Lima AL, Hahn T, de Vries NM, Cohen E, Bataille L, Little MA, et al. Large-Scale wearable sensor deployment in Parkinson's patients: The Parkinson@Home Study Protocol. *JMIR Res Protoc.* 2016;5(3):e172.
31. Guidance for IRBs, clinical investigators, and sponsors, US Department of Health and Human Services and Food and Drug Administration and others. IRB Contin Rev after Clin Investig Approv Silver Spring, MD US food Drug Adm. 2012

Supplementary data: None