A Hybrid Blockchain Design for Patient Recruitments and Persistent Monitoring for Clinical Trials

Yan Zhuang, Informatics Institute

Lincoln Sheets, MD, PhD, Clinical Informatics, School of Medicine

Andrea Atkins, RN, BSN, Clinical Research Center, University of Missouri Healthcare

Chi-Ren Shyu, PhD, Informatics Institute, Electrical Engineering and Computer Science Department

University of Missouri – Columbia, MO, 65211, USA

Zonyin Shae, PhD, Center for FinTech and Blockchain Research

Jeffrey Tsai, PhD, President

Asia University, Taichung, Taiwan

Chun Hsu, PhD

The Ministry of Health Clinical Trial Center of Excellence

China Medical University, Taichung, Taiwan

Abstract

Blockchain is a distributed ledger technology originally applied in the financial sector. This technology ensures the integrity of transactions without third party validation [1]. Its functions of decentralized transaction validation, data provenance, data sharing, and data integration are perfect fits for the needs of clinical trials [2].

Investigating the current workflow of clinical trial operations and conducting design thinking processes with clinicians, trial managers, informaticians, and blockchain professionals, we propose a hybrid blockchain model to tackle known issues. A public blockchain approach is used for *clinical trial recruitment* and a private blockchain approach for *persistent monitoring*. Our model addresses several challenges in the current process and provides potential solutions.

Section I: Public Blockchain for Clinical Trial Recruitment

Background: There are two major clinical trial recruitment tasks that can be handled by blockchain technology: (1) efficient *sponsor recruitment* of *patients* [3] to reach out to the right population of patients in a timely fashion, and (2) user-friendly *patient awareness of new trials* to allow automatic profile matching with trial's inclusion and exclusion criteria. Moreover, the existing public blockchain technology can increase trust between trial sponsors, trial sites (such as clinics, hospitals, and providers), and patients. This allows all parties to tackle issues related to patient safety and trial quality [4].

Challenge 1: Broadcast clinical trial general information to all clinical sites and target qualified patients without going through traditional digital media channels.

The public blockchain could broadcast information to specific recruiting physicians through the "smart contract" feature of the Ethereum product [5]. Blockchain is designed to create peer-to-peer transactions but smart contracts could broadcast a single transaction to all registered peers in the smart contract.

Challenge 2: *Validate the authenticity of sponsors, clinical trials, and real patients.*

Smart contracts could automatically validate the transactions within the blockchain to ensure the authenticity of the sponsors and clinical trials, as well as the qualification of real patients.

Reference Model:

An Ethereum public blockchain could be used to recruit patients. The benefit of using a public blockchain for recruitment is that every user could join the blockchain with minimal effort. Sponsors and all trial sites would be required to have accounts with the blockchain. These accounts would be registered in the clinical trial smart contract. After the sponsors receive the Investigational New Drug (IND) approval and determine the clinical trial criteria [6], the recruitment information would be sent to the smart contract. The smart contract would then validate the identity of the sender and check whether the trial has IND approval. The smart contract would broadcast the inclusion and exclusion criteria to all potential clinical sites.

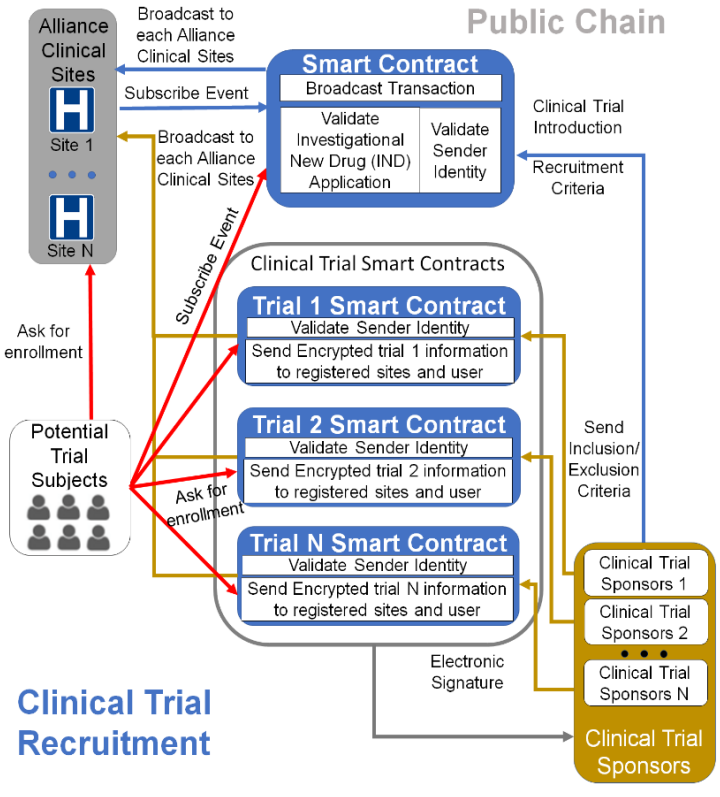

Figure 1 Public blockchain setting for clinical trial recruitment

The clinical trial blockchain will have a unique smart contract for a new trial announcement to allow sponsors to promote a new trial, and for patients as well as clinical sites to be aware of the new trial. Sponsors can create a new transaction with a brief recruitment information and the unique smart contract address. This transaction will be broadcasted to all clinical trials. If

trial sites are interested in participating in the clinical trial, they will confirm their participation to the unique smart contract by referring to this transaction ID. The sponsors would send the encrypted inclusion and exclusion criteria to clinical sites for matching qualified patients in a follow-up smart contract communication.

This model would allow sponsors to broadcast recruitment opportunities to all participating clinical trial sites without going through the traditional digital media channels. By participating in this process, trial sites would receive all clinical trial opportunities information sent by all participating sponsors. The model would ensure that every trial has an IND approval or sponsor verification approval from the Food and Drug Administration (FDA). All parties could fully trust this process.

Sponsors can only send general information through the public chain. Since a public blockchain is a distributed ledger, all details can be viewed by all users in the chain. Clinical trial sites retrieve dynamically the sponsor's smart contract address and confidential recruitment details would need to be communicated in the sponsor's smart contract.

Section II: Private Blockchain for Persistent Monitoring

Background: The current clinical trial system has multiple challenges, including recruitment and information gathering, which result in the infamous "imprecision medicine" [7] problem. Data accuracy is a challenging issue because that the FDA only receives aggregated reports from trial sponsors. Data falsification and human entry errors could happen between trial sites and trial sponsors or between trial sponsors and the FDA. Under the current system, it is difficult to perform timely corrections if there are anomalies and difficult to persistently monitor trials.

Currently, trial subjects can only communicate with clinical sites. Subject conditions are collected by diary or questionnaire. If the system features real-time collection and analysis of diary logs, it would allow sponsors and the FDA to continuously monitor relevant data.

Because the process is complex, the FDA does not have the manpower to audit every clinical trial from each clinical with certain frequencies (daily or even weekly). Instead, the FDA only receives information from sponsors with aggregated outcomes. A robust system that can ensure data accuracy with analytic capability will allow the FDA to play a more proactive role during the trial process.

Challenge 1: Reduce data integrity threats from human error and falsification.

Because patients are distributed among different health providers, data heterogeneity is to be expected. Legacy electronic data capture (EDC) platforms needs manual entry of the data. These processes are prone to have data accuracy issue [8]. Blockchain could ensure data provenance and create immutable audit trails. Each transaction's source is recorded inside the transaction. Smart contracts would only accept transactions sent directly from clinical sites and other transactions would be declined. This would ensure the accuracy of data. If data were falsified outside a smart contract during a transition, every subsequent user of the blockchain could detect that the data was not sent directly from a clinical site. The transaction would not be accepted since there would be a high chance of falsification.

Challenge 2: Collect data from multiple healthcare providers and individuals in real time.

Public blockchain creates new blocks in around 12 seconds [9], and the settings of private blockchain could be changed to make this time shorter. This near-instantaneous responsiveness would allow participants to cooperate in data transactions with no interruption of human response time.

Blockchain uses a distributed ledger technology. Every user inside a private blockchain could fully audit the data transitions. Because patients would have permission to participate in private blockchain, an application program interface (API) could be used to send their diaries or questionnaires to any other participant through their smart phones, such as health care providers and the FDA.

Blockchain is not able to connect databases outside the blockchain system. However, they can communicate with remote procedure call (RPC) servers connected to different clinical sites' electronic medical records. Code on each RPC server could engage with the smart contract within the blockchain. After receiving a request from a smart contract, the RPC server would query different databases and push the required data back to the smart contract. The whole process would be automatically performed after each request.

Challenge 3: Integrate data analytics with blockchain.

Blockchain is not able to install any third-party software. However, machine learning tools could be installed on RPC servers and could communicate with the blockchain. Like query data query processes, all the tools would run automatically after a smart contract sends the request and the RPC server would push all results back to the smart contract.

Challenge 4: Interact seamlessly with current legacy platforms.

Electronic data capture (EDC) platforms provide automated solution for data collection, data validation, and many other features for clinical trials. EDC tools are widely used in all phases of clinical trials. EDC platforms do not allow central data cleaning and cannot detect if data is falsified in the middle of data collection. This function will be added by integration of EDC platforms with blockchain technology.

Challenge 5: Support the oversight role of the FDA.

The FDA would have full control of the RPC servers and the blockchain nodes hosted there. The FDA would have the right to audit any blockchain transaction. They also can check the reports from trial subjects directly instead of receiving all information from trial sponsors.

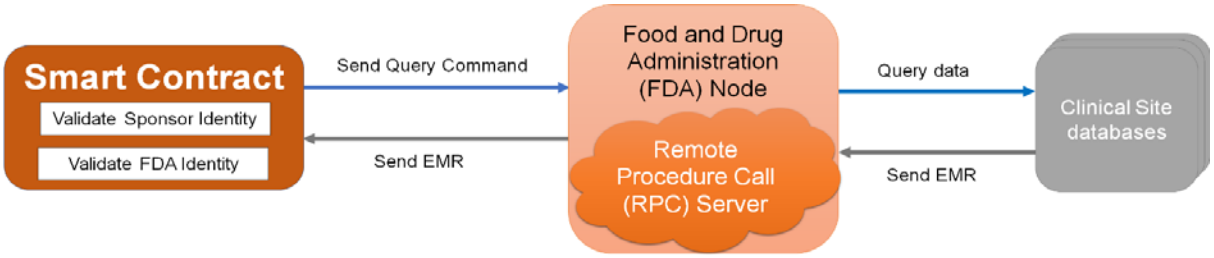

Figure 2 RPC server interact with smart contract and databases

Reference Model:

Private blockchain could be used for persistent monitoring. Participating clinical sites and clinical trial sponsors would join the private chain. Clinical sites would provide computing

resources as “miners” to provide automatic validation of blockchain integrity. Distributed miners could also prevent 51% of known hacking problems [10]. Clinical trial contracts would be structured to reward the contribution of blockchain mining resources. All health records would be structured by HL7 standards [11]. All users inside the private blockchain would audit all transactions inside the system. Any falsification in a transaction would change the sender of the transaction and would be seen by all other user in the private blockchain.

The model includes different phases of the clinical trial. Phases II and III need recruitment first, after sponsors have finalized the subjects list. The “smart contract” would then validate the conductors’ identity and requests from others would be declined. The smart contract would send health data query requests to the RPC servers, which would be fully controlled by the FDA’s node within the private blockchain. The RPC servers would query patient records from the clinical sites’ databases. (Figure 2)

The greatest benefit of using blockchain technology for clinical trials communication is that the FDA could receive raw data from different healthcare providers in real time, without barriers, and without data corruption. Blockchain would ensure immutable audit trails and data provenance. After querying patient records from the trial subjects, RPCs would push data back to a smart contract for further operation. The FDA would usually get updated data in less than 12 seconds.

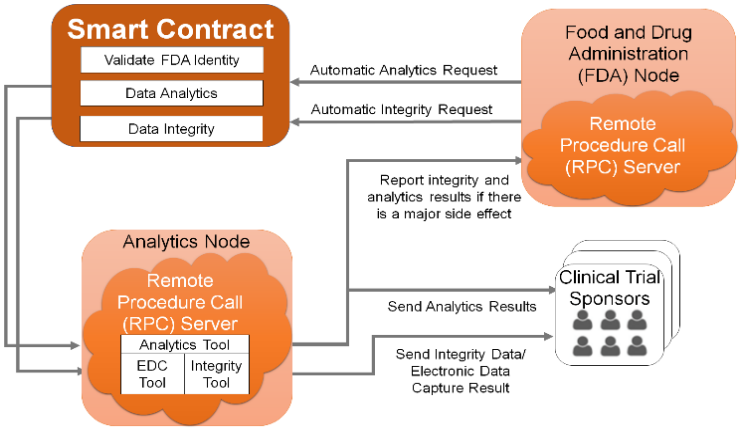

Figure 3 Data integrity and analytics

Data integrity and data analytics would be done on other nodes which contain artificial intelligence (AI) and machine learning components (Figure 3). The whole process would process automatically without mediation from other parties. Since clinical trial subjects may be distributed among different healthcare providers, with their data in different formats, data integrity engines would process the data into a standard format. Data analytics tools would aggregate patterns of demographics and adverse events on integrated data over time. Data analytics reports would be sent to clinical trial sponsors through analytics RPC servers. The FDA would be able to check any reports during clinical trials, and reports of major adverse events would also be sent directly to the FDA. Legacy EDC platforms currently require clinical trial sites to type results data manually. Blockchain technology would automate that process within the analytics node, allow central data cleaning, and detect if the data is falsified in the middle of data collection. Within our model, this process would be done inside the analytics node automatically by integrating the raw data, and blockchain technology would ensure data integrity.

During phase IV, the new drug is approved by the FDA and widely used in multiple clinical sites. This phase may not require recruitment of new patients but does need the reports from patients who have been taking the new drug or treatment for a considerable amount of time. Participating clinical sites would need to provide patient lists to smart contracts through clinical trial sponsors. Smart contract would query the data from the list at regular intervals and go through the same integrity and analytics process.

In our model, subjects could also participate in the trial (Figure 4). They could report their symptoms or adverse events to a smart contract through mobile app. After the smart contract validates their identities, their information could be sent to the FDA. When many subjects in one area have similar symptoms or a major adverse event, the analytics report would be sent to the FDA and the study sponsors.

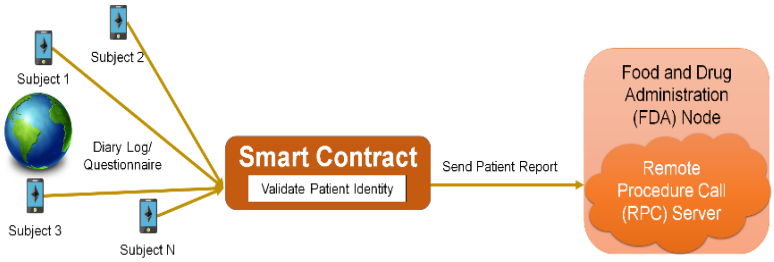

Figure 4 Individuals interact with FDA

In the current system, the FDA is unable to audit every single step during the clinical trial according to the complex process. Our proposed model could help the FDA play an important role during the whole process of any clinical trial. They could check the raw data and any analytics reports, subjects could also communicate with the FDA which cannot be done in the current system.

Issues: All clinical sites must provide some computing resource for mining work. A security issue may arise if one party controls over 50% of computing resources. Clinical sites may not report when there is a minor adverse event according to different clinical trial policies.

Future work

Our approaches touch on representative issues only. More in-depth exploration of the values of blockchain in clinical trials remains to be made by exchange of views, ideas, practical problems and unmet needs between IT and clinical trial specialties.

References

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," ed, 2008.

[2] A. Ekblaw, A. Azaria, J. D. Halamka, and 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*, 2016.

[3] C. P. Gross, R. Mallory, A. Heiat, and H. M. Krumholz, "Reporting the recruitment process in clinical trials: who are these patients and how did they get there?," *Annals of Internal Medicine,* vol. 137, pp. 10-16, 2002.

[4] M. C. Moken, "Fake pharmaceuticals: How they and relevant legislation or lack thereof contribute to consistently high and increasing drug prices," *Am. JL & Med.,* vol. 29, p. 525, 2003.

[5] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," in *Security and Privacy (SP), 2016 IEEE Symposium on*, 2016, pp. 839-858.

[6] J. A. DiMasi, "Risks in new drug development: approval success rates for investigational drugs," *Clinical Pharmacology & Therapeutics,* vol. 69, pp. 297-307, 2001.

[7] N. J. Schork, "Personalized medicine: time for one-person trials," *Nature,* vol. 520, pp. 609-611, 2015.

[8] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde, "Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support," *Journal of biomedical informatics,* vol. 42, pp. 377-381, 2009.

[9] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum Project Yellow Paper,* vol. 151, 2014.

[10] D. Bradbury, "The problem with Bitcoin," *Computer Fraud & Security,* vol. 2013, pp. 5-8, 2013.

[11] R. H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron*, et al.*, "HL7 clinical document architecture, release 2," *Journal of the American Medical Informatics Association,* vol. 13, pp. 30-39, 2006.