METHOD ARTICLE

REVISED How blockchain-timestamped protocols could improve the trustworthiness of medical science [version 21; referees: 32 approved]

Greg Irving[1], John Holden[2]

Author affiliations

Grant information

This article is included in the All trials matter channel.

Abstract

Trust in scientific research is diminished by evidence that data are being manipulated. Outcome switching, data dredging and selective publication are some of the problems that undermine the integrity of published research. Methods for using blockchain to provide proof of pre-specified endpoints in clinical trial protocols were first reported by Carlisle. We wished to empirically test such an approach ~~Here we report a proof-of-concept study~~ using a clinical trial protocol where outcome switching has previously been reported. Here we confirm the use of blockchain~~'blockchain'~~ as a low cost, independently verifiable method ~~that could be widely and readily used~~ to audit and confirm the reliability of scientific studies.

Corresponding author: Greg Irving

How to cite: Irving G and Holden J. How blockchain-timestamped protocols could improve the trustworthiness of medical science [version 21; referees: 32 approved]. *F1000Research* 2016, 5:222 (doi: 10.12688/f1000research.8114.21)

Competing interests:

No competing interests were disclosed.

First published: 26 Feb 2016, 5:222 (doi: 10.12688/f1000research.8114.1)Latest published: 30 Mar 2017, 5:222 (doi: 10.12688/f1000research.8114.3)

REVISED Amendments from Version 1

The method we tested here was first proposed by Carlisle in the grey literature. Clear reference to the previously described method (Reference 6) has been added throughout the revised article.

See referee responses

Editorial note:

Concerns have been raised about the overlap between Version 1 of this article and a previously published blog by Carlisle, who proposed the method 2 years earlier [Carlisle, Benjamin Gregory. "Proof of prespecified endpoints in medical research with the bitcoin blockchain", 25 August 2014], and that the correction (Version 2) published soon after the original was not sufficient to rectify the overlap.
The case has since been discussed in a Committee of Publication Ethics (COPE) Forum, and COPE advised that the correction was sufficient to correct the scientific literature.
The case has been referred to the University of Cambridge for consideration.

Introduction

Trust in scientific research is diminished by evidence that data are being manipulated[1]. Outcome switching, data dredging and selective publication are some of the problems that undermine the integrity of published research. The declaration of Helsinki states that every clinical trial must be registered in a publicly accessible database before recruitment of the first subject[2]. Yet despite the creation of numerous trial registries problems such as differences between pre-specified and reported outcomes persist[3–5]. If readers doubt the trustworthiness of scientific research then it is largely valueless to them and those they influence. Here we confirm the use of blockchain~~propose using a 'blockchain'~~ as a low cost, independently verifiable method that could be widely and readily used to audit and confirm the reliability of scientific studies.
A blockchain is a distributed, tamper proof~~permanent, timestamped~~ public ledger of timestamped transactions. It ~~transactions. In doing so it~~ provides a method for establishing the existence of a transaction~~document~~ at a particular time that can be independently verified by any interested party. When someone wishes to add to it, participants in the network – all of whom have copies of the existing blockchain – run algorithms to evaluate and verify the proposed action. Once the majority of 'nodes' confirm that a transaction is valid i.e. matches the blockchain history then the new transaction will be approved and added to the chain. Once a block of data is recorded on a blockchain ledger it is extremely difficult to change or remove it as doing so would require changing the record on many thousands computers worldwide. This prevents tampering or future revision of a submitted timestamped record. Such distributive version control has been increasingly used in fields such as software development, engineering and genetics. A method for using blockchain to provide proof of pre-specified endpoints in clinical trial protocols was first suggested by Carlisle in 2014[6]. We wished to empirically test such an approach using a clinical trial protocol where outcome switching has previously been reported ~~but to date has not been applied to the reporting of clinical studies~~.

Methods

In this ~~proof-of-concept~~ study we used publically available documentation from a recently reported randomized control trial[7,8]~~trial[6,7]~~. A copy of the clinicaltrials.gov study protocol was prepared based on it's pre-specified endpoints and planned analyses which was saved as an unformatted text file[7]. Following a method similar to that described by Carlisle the file[6] ~~(Dataset 1). The~~ document's SHA256 digest for the text was then calculated by entering text from the trial protocol into an SHA256 calculator (Xorbin©)[6]. This was then converted into a bitcoin private key and corresponding public key using a bitcoin wallet. To do this a new account was created in Strongcoin©[9,8] and the SHA256 digest used as the account password (private key)[6]. From this Strongcoin© automatically generated a corresponding Advanced Encryption Standard 256 bit public key[6]~~key~~. An arbitrary amount of bitcoin was then sent to a corresponding bitcoin address. To verify the existence of the document a second researcher was sent the originally prepared unformatted document. An

SHA256 digest was created as previously described and a corresponding private key, ~~and~~ public key and bitcoin address generated[6]~~generated~~. The exact replication of the bitcoin address~~public key~~ (1AHjCz2oEUTH8js4S8vViC8NKph4zCACXH) was then used to prove the documents existence in the blockchain using blockchain.info©[10][9]. The protocol document was then edited to reflect any changes to pre-specified outcomes as reported by the COMPare group[3]. This was used to create a further SHA256 digest and corresponding public, private key and bitcoin address[3] ~~and private keys[3]~~.

Dataset 1.**Unformatted text file.**

Downloaded data do not display as expected? **Download the data**

Results

Incorporating a transaction from the bitcoin wallet into the blockchain using a ~~public and~~ private key generated from the SHA256 digest of the trial protocol timestamped a record of the study protocol~~provided a timestamped record that the protocol was at least as old as the transaction generated~~. The transaction took under five minutes to complete. The process cost was free as the nominal bitcoin transaction could be retrieved. Researchers were able to search for the transaction on the blockchain, confirm the date when the transaction occurred and verify the authenticity of the original protocol by generating identical public and private keys. Any changes made to the original document generated different public and private keys indicating that protocol had been altered. This included assessment of an~~the~~ edited protocol reflecting pre-specified outcomes not reported or non-pre-specified outcomes ~~now~~ reported in the final paper.

Discussion

Fraud ~~or carelessness~~ in scientific methods erodes ~~the~~ confidence in medicine as a whole which is essential to performing ~~the performance of~~ its function[1]. This study demonstrates that the ~~The~~ method described by Carlisle ~~here~~ provides an immutable record of the existence, integrity and ownership of a specific trial protocol[6]~~protocol~~. It is a simple and cheap way of allowing a third party to audit and externally validate outcomes and analyses specified *a-priori* with the findings reported *a-posteriori*. It ~~The method~~ prevents researchers from changing study endpoints or analyses after seeing their study results without reporting such changes[6]~~changes~~. Transaction codes could be recorded in scientific papers, reference databases or trial registries to facilitate external verification. As discussed in the CONSORT guidelines, switching of outcomes in trials is sometimes necessary for perfectly legitimate reasons but this should be disclosed in the final report[11]. The use of blockchain timestamped protocols could facilitate trust in the reporting of this process by providing evidence of precisely when protocol changes took place. At the same time, fraudulent attempts to ~~Making changes to pre-specified text in a document or trying to bury a protocol in a trial registry would simply not be possible. Attempts to fraudulently~~ prepare multiple study protocols in advance would be technically possible but would also ~~require a considerable amount of advanced planning and would~~ leave behind a publically available trail of evidence that could not be destroyed[6]~~destroyed~~.

The blockchain offers a number of advantages over the current approaches used trial registries or publishing protocols. Firstly, the blockchain would not be confined to the validation of clinical trials. The approach could be used for a whole range of observational and experimental studies where registries do not currently exist. Secondly, the blockchain provides a real-time timestamped record of a protocol. Such precision is important given persistent problems with protocol registration after trial initiation[12]~~initiation[10]~~. Thirdly, with over 30,000 trials currently published annually and rising, manual outcome verification is simply not possible[13]~~possible[11]~~.

Conclusion

Blockchain-timestamped protocols can allow ~~The method we have described allows anyone to verify~~ the exact wording and existence of a protocol at a given point in time to be verified. They have~~. It has~~ the potential to support automated, extremely robust verification of pre-specified with~~and~~ reported outcomes. This evidence should increase trust in medical science by diminishing~~and diminish~~ suspicion in reported data and the conclusions that are drawn.