From Clinical Trials to Highly Trustable Clinical Trials: Blockchain in Clinical Trials, a Game Changer for Improving Transparency?

Mehdi Benchoufi [1,2,3], Doug Altman [4†] and Philippe Ravaud [1,2,3,5*]

[1] METHODS Team, Center for Research in Epidemiology and Statistics Sorbonne Paris Cité (CRESS- UMR 1153), Paris, France, [2] Medical School, Paris Descartes University, Paris, France, [3] Center for Clinical Epidemiology, Hôtel-Dieu Hospital, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France, [4] Center of Statistics in Medicine, University of Oxford, Oxford, United Kingdom, [5] Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, United States

The quality of clinical research is undermined by severe misconduct, error, and fraud, which are detrimental to the trust it should arouse. In this perspective article, we show how Blockchain may trace and control processes of clinical trials to prevent these issues or at least discourage them because they would become traceable and averted. Then, we propose a short and doable program in which, amid the complex stream of events in a clinical trial, we select sensitive and misconduct-prone steps that could greatly benefit from Blockchain by its simple core features such as traceability and incorruptibility of data registration or more refined automation tools called Smart Contracts.

Keywords: clinical trials, Blockchain, transparency, reproducibility, traceability, frauds, quality of research, trusted data

INTRODUCTION

Trust is a pillar for societies. By no means does science, and in particular clinical research, escape this standard. Over the last years, scientific publications and media coverage about fraud, misconduct, detrimental research practices, lack of transparency, poor and selective reporting, and insufficient data-sharing have left the public and researchers confused about the trust they should have in clinical trials (George, 2015). Furthermore, from the viewpoint of civil society and academics, no longer can individual or private institutions (e.g., pharmaceutical company) or public institutions (e.g., hospital, university) be considered trustable by default. In short, one might say there is no trust; there is only proof of trust.

Blockchain, a secure and distributed datastore or ledger of ordered records of transactions, with incorruptibility of the data as a core feature, is a good candidate to provide proof of trust[1] Such an approach could help improve the transparency and trustworthiness of clinical trials and benefit the whole clinical research ecosystem (Benchoufi and Ravaud, 2017; Eichler and Sweeney, 2018; Gammon, 2018). In (Benchoufi and Ravaud, 2017; Benchoufi et al., 2017) we discussed the opportunities for applying Blockchain technologies to improve the quality of clinical trials; we also showed the complexity of the data streams involved in a clinical trial and how these can be traced by blockchain technologies. Here, we clarify the main principles of applying Blockchain to clinical trials, namely "time-stamping," "time-ordering," and "smart-contracting," and we elaborate a hierarchy of applications of Blockchain keeping in mind concrete and deployable solutions. We define what we considered the main clinical trial sensitive points that may benefit from

---

[1] http://scet.berkeley.edu/wp-content/uploads/BlockchainPaper.pdf.

Blockchain application, describe them briefly but functionally, and classify them in terms of the above principles. Also, we detail the operational context enabling the application of Blockchain: architectural choices, governance model, and compliance with general data protection regulations.

In fact, as much as open data has been an important social and scientific trend in the last decade, we consider that "trusted data" is a requirement for upcoming medical research data-driven assertions and that Blockchain represents the dedicated technological tool to concretize this approach, where, in the realm of health, clinical trials may become highly trustable.

BLOCKCHAIN PROPERTIES AND SOME EXAMPLES

Blockchain is a secure and distributed datastore or ledger of ordered records of transactions in which incorruptibility of the data is a core feature (Nakamoto, 2008). With Blockchain, network participants receive a copy of the datastore and can validate by consensus the transactions via a strong cryptographic process, so that trust in transaction validation becomes distributed throughout the network. Hence, parties that do not necessarily trust each other can interact without any third-party intervention, via a shared "tamper-proof" database. By contrast, centralized systems that are widely used data-management systems in clinical trials are single points of failure, do not provide strongly trustable traceability logs because they are controlled by a unique stakeholder and hence are both judge and party.

Three core functionalities that might be fundamental to quality and transparency control in clinical studies can be derived from the technical properties of Blockchain: (1) time-stamping: this data-integrity property implies secure, consistent and incorruptible proof-of-data storage; (2) time-ordering: event consistency allows for checking the integrity of all time-ordered events; and (3) smart-contracting: quality and security of automated checking allows for automatically building quality work-streams called Smart Contracts, or pieces of computer scripts executing contractual clauses on which co-contracting parties agree. Smart Contracts are pipelines of conditioned events that look like "IF Condition 1 AND Condition 2 … AND Condition N (are met) THEN DO this ELSE DO that" (Wang et al., 2018).

Blockchain technology applied to sensitive fields such as health is in its infancy, and in particular, as applied to clinical trials. In general, solutions are more at the stage of prototypes than being mature and ready to implement. One example of a live solution developed on top of a Blockchain is Embleema (https://www.embleema.com/, accessed on September 30, 2019), which are distributed electronic health record systems implemented to grant patients more control over their data and even organize a data-brokerage market place. Another is the Blockchain-based drug verification system Mediledger, powered by Chronicled (https://www.chronicled.com/, accessed on September 30, 2019), which is a solution to track counterfeit drugs. A prototype of complete clinical trials performed on top of a Blockchain is
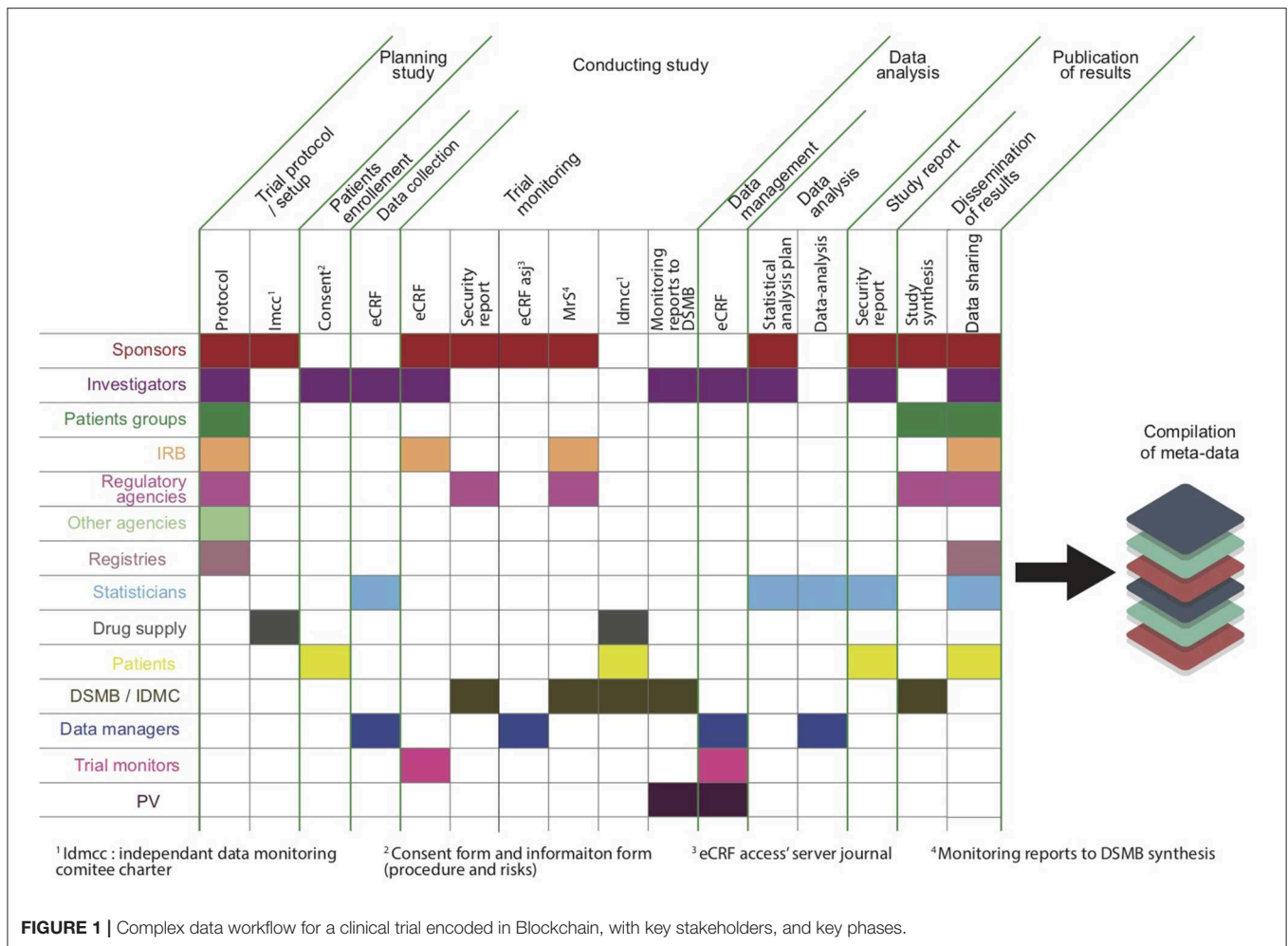
TrialChain implemented via an original Blockchain architecture, where the state of private Blockchain is synchronized to the public Ethereum Blockchain (Dai et al., 2018; Wong et al., 2019). We refer to Agbo et al. (2019) for a recent systematic review of the current Blockchain technologies applied to healthcare. The interested reader will find therein a broad overview of Health's fields in which Blockchain technologies are tested or deployed in live production, but there is no exclusive focus on clinical trials and no insights on the specific phases of clinical trials that should benefit from Blockchain. Besides, the examples that we brought above are by no means exhaustive, however further works would gain at offering a more systematic approach of such initiatives.

BLOCKCHAIN TO IMPROVE THE TRUSTABILITY OF CLINICAL TRIALS

Detrimental research practices can occur at every stage of a clinical trial, from the trial design to the report. Problems related to the poor reproducibility of clinical research have multiple examples in the scientific literature and are one of the greatest medical challenges of our time (Ioannidis, 2005; Mathieu et al., 2009). Selected outcomes, switched outcomes or lack of data-sharing are among the prominent issues (George and Buyse, 2015; Goldacre et al., 2018; House of Commons Science and Technology Committee, 2018). One key fact is that part of the research misconduct is related to "after-the-fact" modifications, so that Blockchain is the guardian of the integrity of the existence of events and their correct chronological order but allows for their traceability, so it may be an interesting tool for improving research reproducibility.

From a practical point of view, clinical trials may be seen as complex systems of intertwined data streams (Benchoufi and Ravaud, 2017) that may be corrupted by misconduct. We refer the reader to **Figure 1** (an updated and extended version of **Figure 1** in Benchoufi and Ravaud, 2017), which provides an overview of the key actors of clinical trials and the data streams in which they are involved, each of which may be subject to malpractice or errors. In **Figure 2**, for each main phase of a clinical trial, we provide key examples of Blockchain applications that may prevent some misconduct. This series of examples of Blockchain Time-stamping, Time-ordering, and Smart Contract functionalities may represent a roadmap. Here we mean by roadmap, a research and development program leading to the development of solutions for the cases that we consider sensitive and most likely to benefit from a Blockchain-type application, and that we fixed in **Figure 1**.

Precisely, blockchain could be used to store proofs of data (e.g., proofs that a consent has been signed, Benchoufi et al., 2017), metadata (e. g., different versions of the protocol, statistical analysis plan and source code) or the data themselves (e. g., data from an e-case report form). Let us consider randomization of a patient in a clinical trial before written consent: this violates basic ethical requirements and could be prohibited with a Smart Contract (Wang et al., 2018), such as "IF patient 1 signed consent form THEN DO allow patient 1 randomization by the central

FIGURE 1 | Complex data workflow for a clinical trial encoded in Blockchain, with key stakeholders, and key phases.

randomization system ELSE DO prohibit randomization." The system records time-stamping and time-ordering of events, so a proof exists that consent was signed before randomization. Patients, investigators, sponsors, and regulators can obtain the corresponding right-ordered proofs of signed consents and randomization events.

Furthermore, the clinical-trial data work-streams may benefit from secure, unalterable storage on the Blockchain, which can be remotely audited (Benchoufi and Ravaud, 2017). Metadata of a clinical trial could be attached so that these critical resources could be "compilable" as a software and provided to reviewers or regulatory authorities. One step further, data analysis could be opened to a third party and processed remotely, without breaching privacy (i.e., data remains stored locally), in a trackable way, which reinforces the credibility of claimed results. A transparency index of tracked events and Smart Contract-validated work streams could be built.

An ultimate evolution would be a Blockchain-powered repository of clinical trials with a standardized format, similar to how a medical image has its own format; i.e., the DICOM format, Graham et al., 2005). This repository could be used for data archiving and be "requestable" by using a dedicated search engine

(Blockchain-as-a-service) for data-sharing and re-use (e.g., unpublished studies, individual patient data for meta-analyses).

The architectural choices, though technical, have a structuring effect on the design of pertinent solutions and are strictly bound with high-level methodological choices. In practice, metadata and proof of data could reside in a public Blockchain via encrypted objects called hashes, whose integrity is verifiable by anyone. Personal data could be stored off chain or on a private blockchain decipherable only by agreed-upon parties. Although we advocate as much has possible the use of full public Blockchains because of their strong guarantees of data incorruptibility, private Blockchain solutions are popular; for instance, Hyperledger (Cachin, 2016) is a widely used private Blockchain. Use of private Blockchains may be legitimate for specific needs. However, when use of a private Blockchain cannot be avoided, hybrid solutions may offer good guarantees of reliability from a distributed network point of view and also combine the usability of private Blockchains. We draw attention to solutions such as Polkadot or the Cosmos network (Cosmos, 2019; Polkadot, 2019). The latter is not a Blockchain *per se* but rather an Inter-Blockchain protocol, enabling separate Blockchain services to communicate. This
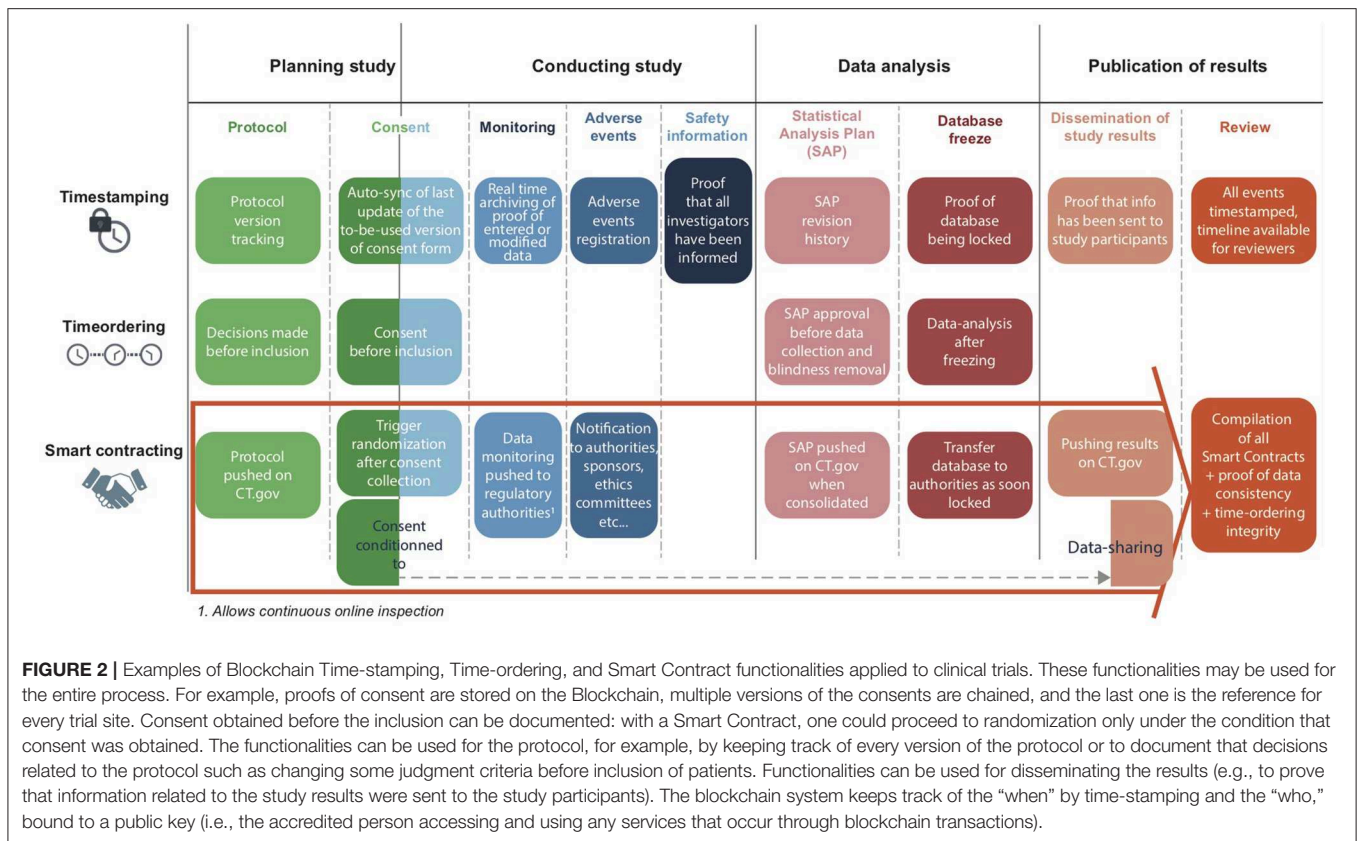
FIGURE 2 | Examples of Blockchain Time-stamping, Time-ordering, and Smart Contract functionalities applied to clinical trials. These functionalities may be used for the entire process. For example, proofs of consent are stored on the Blockchain, multiple versions of the consents are chained, and the last one is the reference for every trial site. Consent obtained before the inclusion can be documented: with a Smart Contract, one could proceed to randomization only under the condition that consent was obtained. The functionalities can be used for the protocol, for example, by keeping track of every version of the protocol or to document that decisions related to the protocol such as changing some judgment criteria before inclusion of patients. Functionalities can be used for disseminating the results (e.g., to prove that information related to the study results were sent to the study participants). The blockchain system keeps track of the "when" by time-stamping and the "who," bound to a public key (i.e., the accredited person accessing and using any services that occur through blockchain transactions).

scenario is interesting because constraints may force hosting services on private or public Blockchains, and these kinds of solutions bring a consistent layer of interoperability. With a different approach but with the aim at preserving at most the public character of Blockchain technologies, Dai et al. (2018) designed a system hosting sensitive interactions on a private Blockchain while maintaining the current state of the system on a public Blockchain.

DISCUSSION

Although Blockchain may help track more closely the series of events occurring in clinical trials and so help achieve better clinical research quality, it is not a solution that exhausts all the problems encountered when facing research reproducibility issues. For instance, some frauds such as complete data invention are not fully preventable by Blockchain, although such fraud would require cryptographic skills, and fraudulent data may be time-stamped and therefore trackable.

We argued that Blockchain may lead to better reproducibility (Leek and Peng, 2015; National Academies of Sciences Engineering and Medicine, 2019). Blockchain technologies, by a close and highly trustable monitoring of potentially any events occurring in a clinical trial, the automated safeguard that may be derived from Smart Contracts allows for replicating the exact condition in which clinical trial processes were conducted, such as data analysis, and so lead to better reproducibility.

In contrast, again tracked events, the tight control of the chronological order of the sequence of events, may raise the overall quality of processes and consequently is a strong consolidation for better replicability in a similar context.

Moreover, storing proof-of-data on a public Blockchain has a cost, although these costs are quite negligible (https://bitinfocharts.com/comparison/ethereum-transactionfees.html, accessed September 30, 2019) as compared with the overall cost of a clinical trial. There is an intense race between Blockchain solution providers to design "consensus mechanisms" that may enable the preservation of core Blockchain properties and high transaction validation efficiency at the same time so as to limit the cost of a single transaction (Baliga, 2017). The upcoming proof-of-stake protocol of the popular Ethereum platform is much anticipated (ethhub.io, 2019). Public Blockchain may also carry a cost in terms of security. Indeed, using a distributed network for clinical trials may seem counterintuitive because avoiding any breech of data privacy is mandatory in trials. However, Blockchain turns the openness of its network into an advantage: the lack of single node points of failure. It has been found quite resilient to security attacks and is constantly being improved (Li et al., in press).

At a higher level, if Blockchain offers a wide range of technical tools that may consolidate the trustworthiness of clinical trials, these tools must be integrated with consistent and coherent governance. Hopefully, Blockchain technologies offer a way to design governance systems: public, permissioned or

private Blockchains; open- or closed-source software and Smart Contracts; fixed or evaluative governance rules. Examples are separating data storage and proof-of-data storage, customizable data-sharing and data request systems, and tokenization if some user incentivization is pertinent. The ultimate distributed governance culminates in the idea of Distributed Automated Organization, whereby equal control is granted to each participant (Davidson et al., 2016). However, this concept is not well-suited for our needs because constantly evolving governance may instill uncertainty in the stability of the processes. These governance models in the realm of health must be designed in close coordination with the legal context, which varies from state to state. For instance, compliance with general data protection regulations has strong implications, mainly (1) re-identification issues because a proof of data, even shaped as a hash, is considered potentially re-identifying data as long as the raw data from which the hash was derived is stored on a database. In Europe, regarding storing proof-of-health data, some regulators in Europe recommend producing a risk analysis. Also, (2) the right to forget contradicts the append-only nature of the Blockchain data structure, from which no data can be overwritten. However, when the link between the hash and its raw data is deleted, then there is no issue for the hash to be stored as a persistent data, and since health data cannot not be stored beyond some time period depending on the state regulation, this problem becomes extinct with time (besides, such a data lifecycle could be taken care of by a Smart Contract).

Importantly, standards for clinical trials remain to be consensually defined (e.g., public or permissioned Blockchain, user control of implementation model, open source practices); this work is ongoing through Institute of Electrical and Electronics Engineers initiatives to standardize the technology https://blockchain.ieee.org/standards). The growing interest in involvement in this technology by regulatory bodies such as the US Food and Drug Administration, pharmaceutical companies, and the clinical research community would also be crucial (Vahan, 2017; Colm, 2018).

At last, let us mention that from a usability point of view, moving from proof of concept to production is always challenging, and real-life implementation will require managing some technical complexity burdens such as acculturating actors to new processes and designing simple user experiences yet requiring some crypto literacy. This last point is of importance because how the user logs onto a Blockchain system and is provided a cryptographic pair of keys is still complex and is a strong brake on adoption.

CONCLUSION

Blockchain technology, if deployed in the context of clinical trials, will help build transparent, highly trustable trials, and deserves attention from the research community. Such technology applied to critical steps of clinical trials can improve the development of trusted processes and gathering data and reinforce transparency. However, the benefits of this technology is tempered by the cost of designing a layer of a complex technology, supposes the use of new skills, and more prospectively rethinking the clinical trial workflow as an achievable and wide generalization of the seminal idea of "threading trials" (Altman et al., 2014). Beyond this conceptual approach, a great amount of work remains in building automated workflows to ensure quality of clinical trial phases and implementing them concretely.

AUTHOR CONTRIBUTIONS

MB developed the theoretical formalism and designed the work. PR identified which of clinical-trial sensitive steps should be under the Blockchain validation process. PR and DA brought their expertise to consolidate the design of work and brought synthetic insights into challenges and limits of the proposed use of Blockchain and finally reviewed the article and approved the article. DA especially put in perspective the current work with former work of DA et al. on linked publications for a single trial (Vahan, 2017).

REFERENCES

Agbo, C. C., Mahmoud, Q. H., and Eklund, J. M. (2019). Blockchain technology in healthcare: a systematic review. *Healthcare* 7:56. doi: 10.3390/healthcare7020056

Altman, D. G., Furberg, C. D., Grimshaw, J. M., and Shanahan, D. R. (2014). Linked publications from a single trial: a thread of evidence. *Trials* 15:s369. doi: 10.1186/1745-6215-15–369

Baliga, A. (2017). *Understanding Blockchain Consensus Models*. Whitepaper. Retrieved from: https://pdfs.semanticscholar.org/\penalty\z@da8a/37b10bc1521a4d3de925d7ebc44bb606d740.pdf

Benchoufi, M., Porcher, R., and Ravaud, P. (2017). Blockchain protocols in clinical trials: transparency and traceability of consent. *F1000Res.* 6:66. doi: 10.12688/f1000research.10531.1

Benchoufi, M., and Ravaud, P. (2017). Blockchain technology for improving clinical research quality. *Trials* 18:335. doi: 10.1186/s13063-017-2035-z

Cachin, C. (2016). "Architecture of the hyperledger blockchain fabric," in *Workshop on Distributed Cryptocurrencies and Consensus Ledgers*, Vol. 310 (Chicago, IL), 4.

Colm, H. (2018). *Big Pharma and Blockchain: Setting a Standard with the IEEE*. Available online at: https://www.blockchaintechnology-news.com; https://www.blockchaintechnology-news.com/2018/01/04/big-pharma-blockchain-setting-standard-ieee/

Cosmos (2019). *In Bitcoinwiki*. Retrieved from: https://en.bitcoinwiki.org/wiki/Cosmos (Accessed October 10, 2019).

Dai, H., Young, H. P., Durant, T. J., Gong, G., Kang, M., Krumholz, H. M., et al. (2018). TrialChain: a blockchain-based platform to validate data integrity in large, biomedical research studies. *arXiv 1807.03662*.

Davidson, S., De Filippi, P., and Potts, J. (2016). *Disrupting Governance: The New Institutional Economics of Distributed Ledger Technology*. Available online at: https://ssrn.com/abstract=2811995

Eichler, H. G., and Sweeney, F. (2018). The evolution of clinical trials: can we address the challenges of the future?. *Clin. Trials* 15 (Suppl. 1), 27–32. doi: 10.1177/1740774518755058

ethhub.io (2019). *Proof of Stake (PoS.)* Available online at: https://docs.ethhub.io/ ethereum-roadmap/ethereum-2.0/proof-of-stake/

Gammon, K. (2018). Experimenting with blockchain: can one technology boost both data integrity and patients' pocketbooks?. *Nat. Med.* 24, 378–381. doi: 10.1038/nm0418-378

George, S. (2015). Research misconduct and data fraud in clinical trials: prevalence and causal factors. *Int. J. Clin. Oncol.* 21, 15–21. doi: 10.1007/s10147-015-0887-3

George, S. L., and Buyse, M. (2015). Data fraud in clinical trials. *Clin. Investig.* 5, 161–173. doi: 10.4155/cli.14.116

Goldacre, B., DeVito, N. J., Heneghan, C., Irving, F., Bacon, S., Fleminger, J., et al. (2018). Compliance with requirement to report results on the EU clinical trials register: cohort study and web resource. *BMJ* 362:k3218. doi: 10.1136/bmj.k3218

Graham, R. N., Perriss, R. W., and Scarsbrook, A. F. (2005). DICOM demystified: a review of digital file formats and their use in radiological practice. *Clin. Radiol.* 60, 1133–1140. doi: 10.1016/j.crad.2005.07.003

House of Commons Science and Technology Committee (2018). *Research Integrity: Clinical Trials Transparency. Tenth Report of Session.* 2017–19 Available online at: https://publications.parliament.uk/pa/cm201719/cmselect/ cmsctech/1480/1480.pdf

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med.* 2:e124. doi: 10.1371/journal.pmed.0020124

Leek, J. T., and Peng, R. D. (2015). Opinion: reproducible research can still be wrong: adopting a prevention approach. *Proc. Natl. Acad. Sci.U.S.A.* 112, 1645–1646. doi: 10.1073/pnas.1421412111

Li, X., Jiang, P., Chen, T., Luo, X., and Wen, Q. (in press). A survey on the security of blockchain systems. *Future Gen. Comput. Syst.* doi: 10.1016/j.future.2017.08.020

Mathieu, S., Boutron, I., Moher, D., Altman, D. G., and Ravaud, P. (2009). Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 302, 977–984. doi: 10.1001/jama.2009.1242

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System.* Available online at: https://bitcoin.org/bitcoin.pdf

National Academies of Sciences Engineering and Medicine (2019). *Reproducibility and Replicability in Science.* Washington, DC: The National Academies Press.

Polkadot (2019). *In Wikipedia, The Free Encyclopedia.* Retrieved from: http:// fr.wikipedia.org/w/index.php?title=Polkadot&oldid=163039276 (Accessed October 10, 2019).

Vahan, S. (2017). *Healthcare Data Exchange Framework (HDEF) Scalable Economy Of Secure Information And Services.* Available online at: http://mdepinet.org/ wp-content/uploads/Simonyan_HDEF_MDEpiNet.pdf

Wang, S., Yuan, Y., Wang, X., Li, J., Qin, R., and Wang, F. Y. (2018). "An overview of smart contract: architecture, applications, and future trends," in *2018 IEEE Intelligent Vehicles Symposium (IV)* (IEEE). Available online at: https://en. wikipedia.org/wiki/Smart_contract

Wong, D. R., Bhattacharya, S., and Butte, A. J. (2019). Prototype of running clinical trials in an untrustworthy environment using blockchain. *Nat. Commun.* 10:917. doi: 10.1038/s41467-019-08874-y