OPEN TrialChain: Privacy-Preserving Blockchain Design for Diabetes Clinical Trials Using Open Algorithms

Kim, A. Milan C. Alsalamah, S. and Pentland, A.
Agenda

• Clinical Trials
• Clinical Trial Problems
• Type II Diabetes Case Study
• OPEN TrialChain Model
• Adopting OPEN TrialChain
• Challenges
Clinical Trials
Introduction

- Governments and pharmaceutical companies (Pharma) spend $100-800 millions on clinical trials for each drug candidate, and yet up to 25-50% go unpublished.

- Section 801 of the Food and Drug Administration Amendments Act (FDAAA) of 2007 requires the submission of "basic results" for certain clinical trials,
  - The ambiguity of “basic results” leads to rather scant result submissions.

- It is in the best interest of pharmaceutical companies to report as little as possible
  - Gain a competitive advantage over other pharmaceutical companies developing drugs in the same space
  - Maintain patient privacy.
Clinical Trials- Ecosystem

- Patients
- FDA
- Secondary Researchers
- Quintiles IMS
- Research Sites
- National Clinical Registries
- Other Countries
- Pharma
- Families
- Other Countries
- National Clinical Registries
- Research Sites
- Quintiles IMS
- Secondary Researchers
- FDA
- Patients
Clinical Trials- Ecosystem

- **FDA**: Approve all clinical trials (but don’t always enforce publication of trials)
- **Quintiles IMS**: third-party that plans the logistics of clinical trials, and are often the data owners
- **Research Sites**: physical clinics that report patient metrics
- **Other Countries**: if clinical trials are done internationally (to cut cost), countries may intervene in the exportation of the data by order of the General Data Protection Regulation (GDPR)
- **Pharmaceutical companies**: develop the drugs that are being tested. Their IP is on the line and has to be protected
- **Patients**: In the case of GDPR, they have to have revoke access
- **Secondary Researchers**: In order to learn and make retrospective studies on drug development
- **Families**: In the case of pediatric trials, families may be consulted for consent
Clinical Trial Problems
Clinical Trial Problems

1. Pharmaceutical companies are **discouraged from publishing** their trial results as they do not want to lose ownership over it.

2. Even when they do publish, **reported results** are **very limited** for competitive advantage.

3. **Underrepresented groups** have **no significance** in the **results** of trials.

4. **Results** are **not illustrated** in a **meaningful** way because they could be cross-referenced and track it back to the subject under study resulting in **privacy invasion**.
Case Study: Type II Diabetes in Adults
Type II Diabetes Clinical Trials- Dataset in Adults

2 Genders
237 Subjects Started
4 Adverse Event Categories
1 Death

January 31, 2018
OPEN TrialChain
## Diabetes Type II Clinical Trials - Dataset in Adults

<table>
<thead>
<tr>
<th></th>
<th>Metformin XR</th>
<th>Canagliflozin 100 Milligram (mg)</th>
<th>Canagliflozin 300 mg</th>
<th>Canagliflozin 100 mg + Metformin XR</th>
<th>Canagliflozin 300 mg + Metformin XR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Started</strong></td>
<td>237</td>
<td>237</td>
<td>238</td>
<td>237</td>
<td>237</td>
</tr>
<tr>
<td><strong>Adverse Event</strong></td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td><strong>Death</strong></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
OPEN TrialChain Model
OPEN TrialChain Model

• What is it?
  ▶ It is a **Blockchain-based** data sharing **infrastructure** that uses open algorithms between stakeholders of the federation in the clinical trials ecosystem.

• Benefits
  ▶ OPEN TrialChain balances between the **sharing** of clinical **data** and the need for subject’s **privacy protection**.

• How?
  ▶ By allowing queries on decentralized raw datasets from which returns aggregated safe answers that are blinded (i.e. anonymized).
OPEN TrialChain: Components

1. **Data Providers**
   - Users owning clinical trial datasets owners who monitors other’s queries over their data.

2. **Vetted algorithms**
   - Allow users to query for specific demographics across multiple clinical trials without compromising individuals’ data.

3. **Queries Blockchain**
   - A tamper-proof, time-stamped ledger for audits and monitoring purposes.
   - FDA could use for study auditing.
   - Stakeholders use it to allow well-informed decision making processes.
OPEN TrialChain - Key Principles

1. Keeps data encrypted at all times
2. Allows only vetted algorithms
3. Moves the algorithm to the data, (not vice versa)
4. Aggregates blinded data from distributed repositories in a decentralized infrastructure
5. Returns “safe answers”
OPEN TrialChain - Model Architecture

Choose Algorithms

Querier

Algorithms

Safe Answers

Coordinator Node

Log (e.g. Blockchain/DLT)

Auditor

MPC Node

Data Repository

MPC Node

Data Repository

MPC Node

Data Repository

Multi Party Computation
OPEN TrialChain - Data Flow

Clinical Trial Settings

Data Providers/Owners

Raw Datasets

Decentralized Discrete Data Centers

OPEN TrailChain

Monitor

Queriers

Query

“Safe” Answer

Raw Data

Blinded Data
OPEN TrialChain - Blockchain For Clinical Trials

New Query

New block (transaction)

Signed transaction added to the sequence

Transaction validation
Adopting OPEN TrialChain Model
OPEN TrialChain - Potential Users

› **Doctors** - query for information relevant to their patients

› **FDA** - audit the blockchain and usage of information

› **Secondary Researchers** - conduct primary retroactive research for other potential drug developments

› **Pharmaceutical Companies** - monitor the blockchain for usage of their data in user queries
OPEN TrialChain - In Type II Diabetes in Practice

- User (Doctor)
  - A doctor who has a diabetic patient who is an African American female
  - Although there have been many diabetes trials with immediate-release Metformin, each trial doesn’t have a very significant cohort of African American women.
  - Therefore, the granularity of the studies’ results is very coarse in order to protect the privacy of these few patients.
OPEN TrialChain- In Type II Diabetes in Practice

‣ Question / Algorithm

• What are the adverse events for diabetic African American females taking Metformin?

‣ Unsafe Answer

Detailed answers about a subject from one clinical trial that would compromise the subject’s health data.

‣ Safe Answer (without OPEN TrialChain)

Vague answers in disjoint tables that protect the few African American females in a single study.
OPEN TrialChain- In Type II Diabetes in Practice

Question / Algorithm

- What are the adverse events for diabetic African American females taking Metformin?

Safe Answer (with OPEN TrialChain)

- A safe answer would provide a blinded summary statistics (histogram or scattered plot) that illustrates the adverse events for this particular demographic.
What are the adverse events for diabetic African American females taking Metformin?
OPEN TrialChain - Aggregated vs. Blinded Results

Aggregated Less Informative Disjoint Table Results

OPEN TrialChain Interface Layer

Application Layer “Algorithms”

Blockchain Layer

Physical Layer “Server”

Aggregated Informative Blinded Visual Results

January 31, 2018
Addressing Clinical Trial Problems
Addressing the Problems

1. Encourages **reporting trial results** with higher fidelity **audited by the blockchain**
2. Incentives **more detailed results** in exchange for their peers' detailed results.
3. Report **safe answers** in summary statistics about the salient information (ex. effects of Metformin) without linking “unflattering” results to individual studies.
4. Returns **meaningful visualizations** that can help understand complex results and derive insights by comparing the models/diagrams.
Impact and Challenges
Impact on Ecosystem

• Eventually, OPEN TrialChain is expected to **optimize the clinical trial ecosystem** by
  • improving patient safety,
  • saving lives,
  • cutting drug development costs,
  • encouraging transparent results,
  • preserving patients privacy, and
  • maintaining pharmaceuticals integrity.
Adoption Challenges

• Early adopters would be hesitant to provide details of their data. Even if trial details would only be disclosed in safe answers, it still **dulls a company’s competitive edge.**

• **Setting up OPEN TrialChain servers** for each dataset would be work that pharmaceutical companies may not be willing to invest in.

• At this time, OPEN TrialChain is **not GDPR compliant** and does not give individual subjects the ability to revoke access to their data.
Thank you