Can machine learning identify untrustworthy clinical trials?

ISB Science Transforming Health

TRANSLATOR Funded by NIH's National Center for Advancing Translational Sciences

Jackie Lindstrom[1,2], Gwênlyn Glusman[2]
[1]University of Michigan, Ann Arbor, MI [2]Institute for Systems Biology, Seattle, WA

Introduction to Clinical Trial Data:

- clinicaltrials.gov is the publicly accessible database managed by the NIH and National Library of Medicine
- Contains over 500,000 trials
- We have > 900 snapshots starting from 2017
- Used in research, but not quality controlled:

The U.S. government does not review or approve the safety and science of all studies listed on this website.

Read our full disclaimer for details.

*"The study **sponsor** or investigator submits information about their study to ClinicalTrials.gov and is **responsible** for the safety, science, and **accuracy** of any study they list"*[1]

Untrustworthy Trials:

Medicine is plagued by untrustworthy clinical trials. How many studies are faked or flawed?

Investigations suggest that, in some fields, at least one-quarter of clinical trials might be problematic or even entirely made up, warn some researchers. They urge stronger scrutiny.

nature

- An analysis of 150 randomized control trials showed 44% trials contained "at least some flawed data: impossible statistics, incorrect calculations or duplicated numbers or figures"[2]
- In a review of 110 trials, publication and clinicaltrial.gov information differed, where 80% of secondary outcome and 35% of serious adverse events reporting had inconsistencies[3]

Clinical Trial Data Use in Translator:

- Translator ingests data from clinicaltrials.gov

Trial is Phase 4 and FDA approved:
assertion that *intervention treats disease*

Trial exists in clinicaltrials.gov:
prediction that *intervention treats disease*

Ideally, Translator would *NOT* ingest untrustworthy trials

Objective: Classify clinical trials into *trustworthy* or *untrustworthy (faked, sloppy)* using machine learning

Selection criteria:

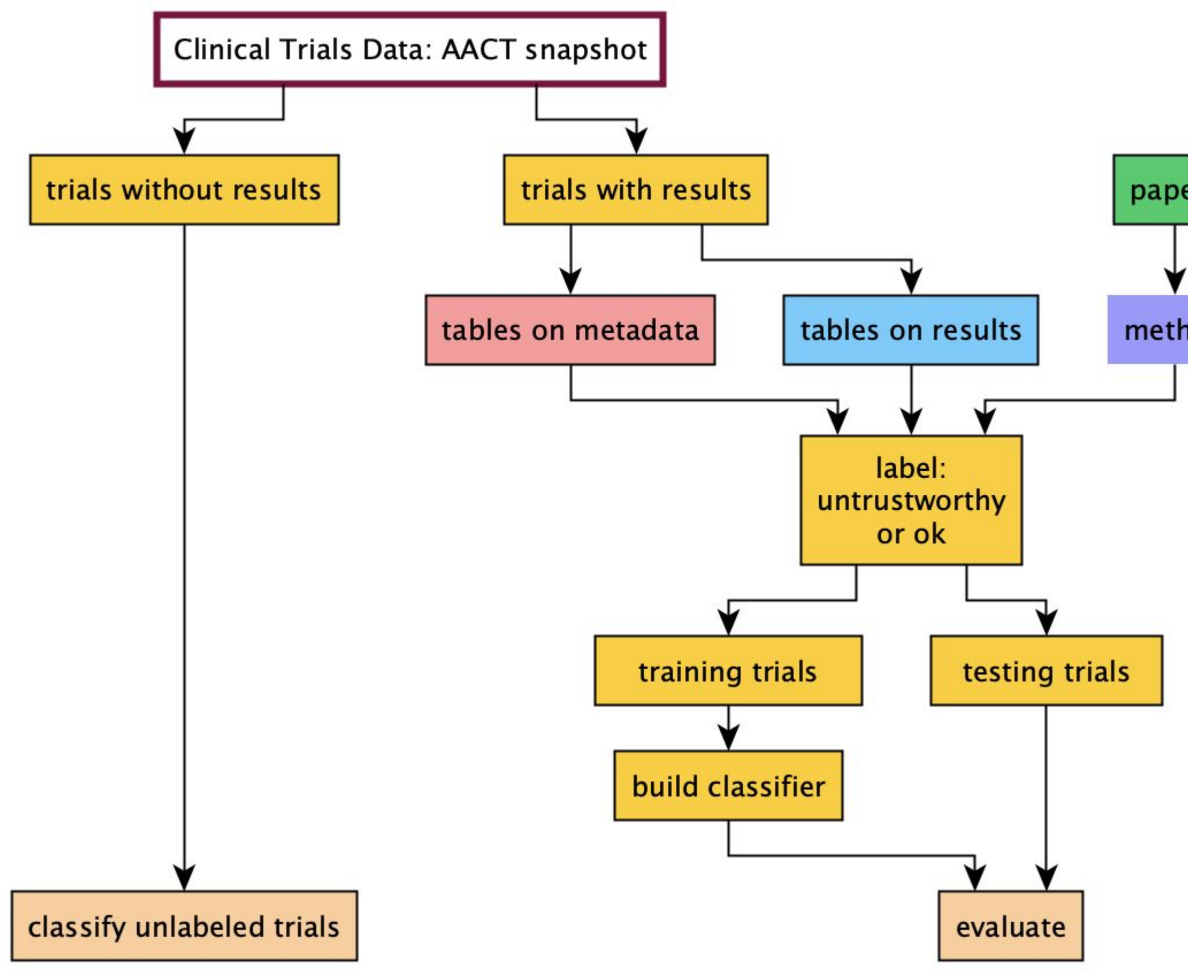
1. Was an associated paper retracted?
2. Was the study information posted retroactively?
3. Do key features have low stability scores?

Stability: $S = 1 - \frac{R-1}{N-1}$

R: # times value *changed*
N: # of snapshots (value *could* have changed

For every column in every trial, out of all the times a value could change, how often did it?

Workflow:

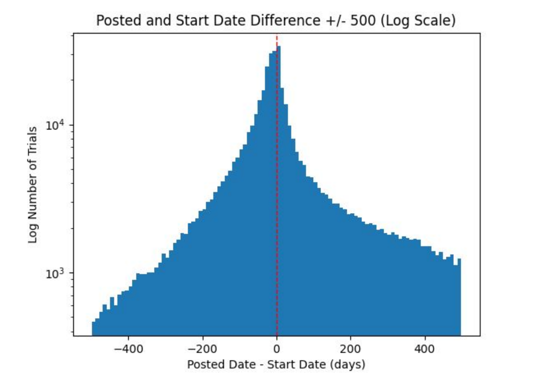


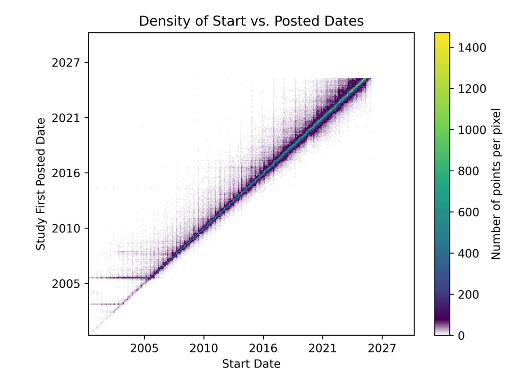
What would you consider untrustworthy?

Please leave a sticky note- your thoughts are welcome

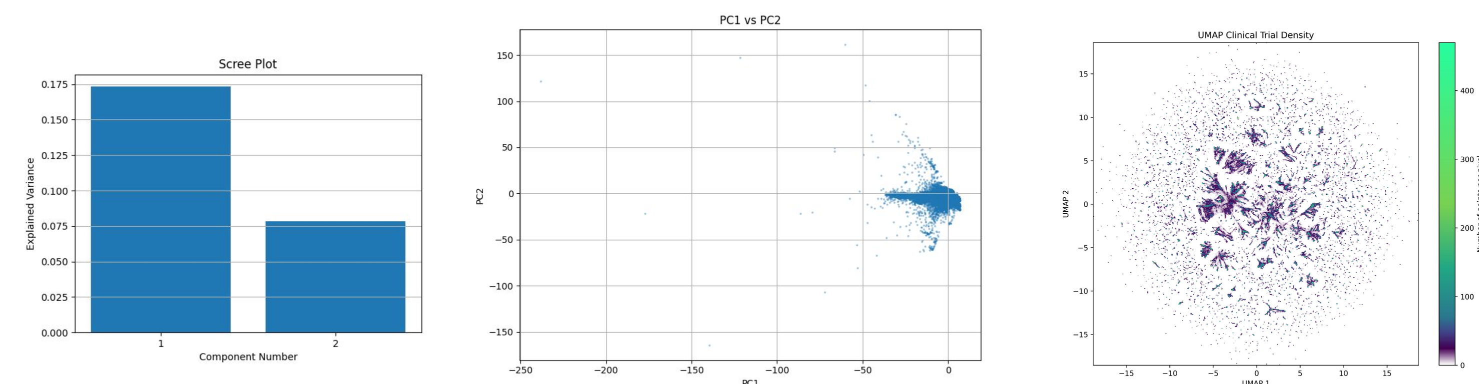
Labeling Trustworthy or Ok (In Progress)

1) Using PubMed database, link trial to retracted paper
2) Data visualization: examined retroactive studies

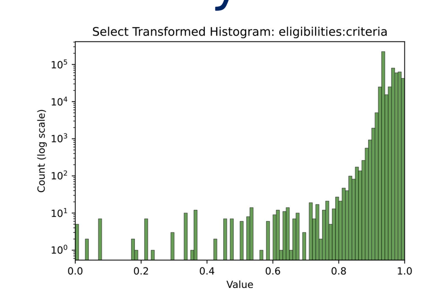




1) PCA with stability score data for feature selection→ PC1 & PC2 capture only small fraction of variance
Next, UMAP applied to examine structure



4) Identify outliers from stability data



- 132 plots, one for each feature
- outlying trials identified as having more outlying features
- features expected to change (such as updated_at date) disregarded

Future Work:

Eventually after examining or filtering based on the trustworthiness of the trials, we could classify phase 3 trials into success/ failure.

Definition	Source
"Successful patient registration is defined as clinical trial success"	• **Factors Affecting Success of New Drug Clinical Trials (1)**
"Success and "Failure" is in the context of achieving approval"	• **Machine Learning with Statistical Imputation for Predicting Drug Approvals (2)** • **A Tool for Predicting Regulatory Approval After Phase II Testing of New Oncology Compounds (3)**

Acknowledgements:

- Hood Lab, Institute for Systems Biology
- Dr. Gwzênlyn Glusman and Skye Goetz
- NIH NCATS Biomedical Data Translator Award #OT2 TR005706

Sources:

1) U.S. National Library of Medicine. *ClinicalTrials.gov*. https://clinicaltrials.gov.
2) Van Noorden, R. *Medicine is plagued by untrustworthy clinical trials. How many studies are faked or flawed?* **Nature 2023**, *619* (7970), 454–458. DOI: 10.1038/d41586-023-02299-W.
3) Hartung, D.; Zarin, D. A.; Guise, J.-M.; McDonagh, M.; Paynter, R.; Helfand, M. *Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications.* **Ann. Intern. Med. 2014**, *160* (7), 477–483. DOI: 10.7326/M13-0480. PMC4617780
4) Kim, E.; Yang, J.; Park, S.; Shin, K. Factors Affecting Success of New Drug Clinical Trials. Ther. Innov. Regul. Sci. 2023, 57 (4), 737–750. DOI: 10.1007/s43441-023-00509-1.
5) Lo, A. W.; Siah, K. W.; Wong, C. H. Machine Learning with Statistical Imputation for Predicting Drug Approvals. Harvard Data Sci. Rev. 2019, 1 (1). DOI: 10.1162/99608f92.5c5f0525.
6) DiMasi, J. A.; Hermann, J. C.; Twyman, K.; Kondru, R. K.; Stergiopoulos, S.; Getz, K. A.; Rackoff, W. A Tool for Predicting Regulatory Approval after Phase II Testing of New Oncology Compounds. Clin. Pharmacol. Ther. 2015, 98 (5), 506–513. DOI: 10.1002/cpt.194.