UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

AI IN CLINICAL TRIALS: IS THERE A ROLE FOR STATISTICS?

EFSPI Regulatory Statistics Workshop 2025

Tim Friede

Institut für Medizinische Statistik

Universitätsmedizin Göttingen

IS THERE A ROLE FOR STATISTICS IN AI?

- **DAGStat (German Consortium of Statistics)**
  - Association of 15 professional and learned societies and the Destatis (Federal Statistical Office in Germany)
  - Homepage https://www.dagstat.de/

Advances in Data Analysis and Classification
https://doi.org/10.1007/s11634-021-00455-6

REGULAR ARTICLE

Is there a role for statistics in artificial intelligence?

Sarah Friedrich, et al. *[full author details at the end of the article]*

Received: 13 September 2020 / Revised: 6 July 2021 / Accepted: 7 July 2021

STATISTICAL STUDY DESIGN

- **Experiment - observational study - convenience sample**
  - AI utilizes often convenience samples (e.g. routine data) since large data sets and accessible at low costs
  - Target trial emulation
- **Naive expectation that sufficiently large data automatically lead to representativity**
  - Careful planning required
  - Example: Apple Heart Study (Perez et al (2019) NEJM) to assess Apple Smartwatch‘s ability to identify atrial fibrillation; Large-scale assessment including more than 400,000 participants; Caveat: Average age of participants 41 years, but AF most prevalent in older patients (>65 years of age)

STATISTICAL APPROACHES FOR VALIDATION

- **Internal vs. external validation**
- **Bench marking** (computer science) vs. **simulations** (statistics) (see also Friedrich & Friede (2023) https://doi.org/10.1002/bimj.202200212)
  - Experience with structuring, reporting and interpreting simulation studies in statistics
- **Sample sizes**
  - Sample size planning common task in clinical trials, but not routinely performed in AI / ML applications
  - Requirements depending on dimensionality, sparsity, non-linearity, …
- **Challenges include fast development cycles with AI technologies** (often faster than validation studies)

"GARBAGE IN, GARBAGE OUT"

- AI need to be trained and evaluated on **'fit for purpose' data** (includes relevance, completeness, availability, timeliness, meta-information, documentation and context-dependent expertise) (Duke-Margolis, 2018)

Fig. 3 Data relevancy and quality are equivalent components of a fit-for-purpose real-world data set. Figure according to Duke-Margolis (2018)

- **More data ≠ more information** (e.g. random-effects meta-analysis; Jackson and Turner, 2017)

CAUSALITY AND ASSOCIATION

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

- **AI / ML excellent at discovering associations**
- Important to acknowledge in the interpretation of results: **Association not necessarily due to causal relationship**
- **Lessons to be learned from other fields**
  - Clinical epidemiology: Bradford Hill criteria
  - Statistics: Counterfactual framework by Rubin (1974)

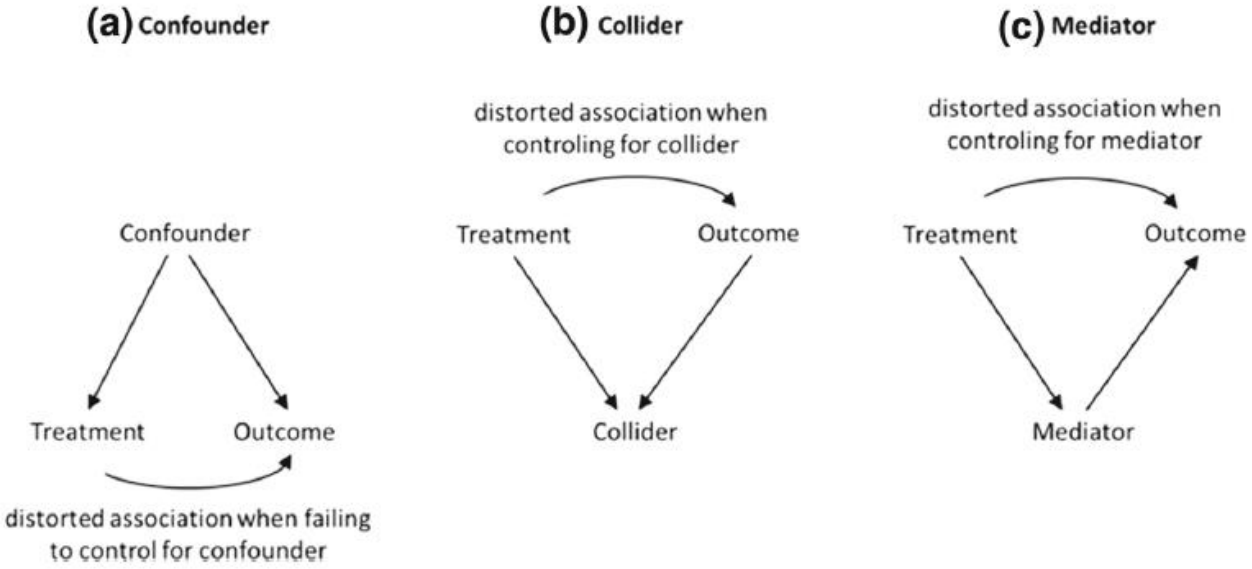


Fig. 4 Covariate effects in observational data, according to Catalogue of bias collaboration (2019)

EVALUATING UNCERTAINTY

- **Uncertainty quantification is often neglected in AI applications**
- With large data sets **sampling variation** might be small
- However: **Model uncertainty** might remain
- Derivation / computation of measures of uncertainty such as standard errors or prediction intervals can be tricky with complex analysis methods
- More recently, application of **resampling techniques** such as bootstrapping or jack knifing (theoretical properties not always established)
- Alternative approaches embed algorithmic methods in statistical models

APPLICATIONS OF AI/ML IN CARDIOVASCULAR MEDICINE: A SYSTEMATIC REVIEW

- DZHK (German Center for Cardiovascular Research)
- Project group AI / ML

DZHK
DEUTSCHES ZENTRUM FÜR HERZ-KREISLAUF-FORSCHUNG E.V.

ESC European Society of Cardiology
European Heart Journal Digital Health
doi:10.1093/ehjdh/ztab054

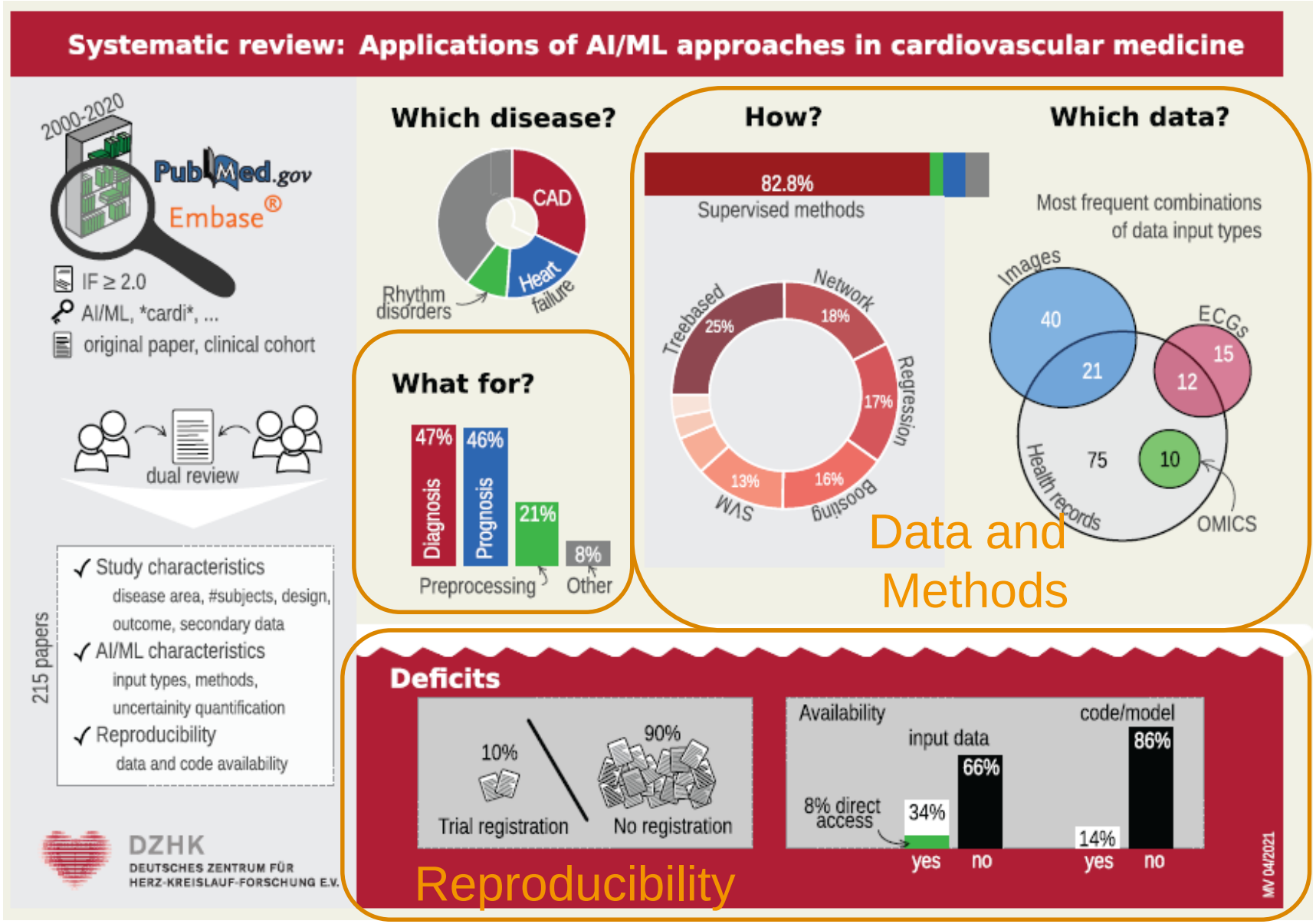
REVIEW

Applications of artificial intelligence/machine learning approaches in cardiovascular medicine: a systematic review with recommendations

Sarah Friedrich [1]*, Stefan Groß [2,3], Inke R. König[4,5], Sandy Engelhardt[6,7,8], Martin Bahls[2,3], Judith Heinz[1], Cynthia Huber[1], Lars Kaderali[3,9], Marcus Kelm [10,11,12,13], Andreas Leha[1,14], Jasmin Rühl[1], Jens Schaller[10,13], Clemens Scherer [15,16], Marcus Vollmer[3,9], Tim Seidler[14,17], and Tim Friede [1,14]

AI/ML IN CV MEDICINE

UNIVERSITÄTSMEDIZIN GÖTTINGEN : UMG



IN SILICO CLINICAL TRIALS

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

Received: 8 February 2021 | Revised: 27 May 2021 | Accepted: 27 May 2021

DOI: 10.1002/psp4.12669

ASCPT

WHITE PAPER

Scientific and regulatory evaluation of mechanistic *in silico* drug and disease models in drug development: Building model credibility

Flora T. Musuamba[1,2,3] | Ine Skottheim Rusten[1,4] | Raphaëlle Lesage[5,6] | Giulia Russo[7] | Roberta Bursi[8] | Luca Emili[8] | Gaby Wangorsch[1,9] | Efthymios Manolis[1,10] | Kristin E. Karlsson[1,11] | Alexander Kulesza[12] | Eulalie Courcelles[12] | Jean-Pierre Boissel[12] | Cécile F. Rousseau[13] | Emmanuelle M. Voisin[13] | Rossana Alessandrello[14] | Nuno Curado[15] | Enrico Dall'ara[16] | Blanca Rodriguez[17] | Francesco Pappalardo[7] | Liesbet Geris[5,6,18]

In silico trials: "Class of trials for pharmacological therapies or medical devices based on modelling and simulation technologies. Such trials produce digital evidence that can serve in complement to or replacement of in vivo clinical trials for the development and regulatory evaluation of medical therapies."

IN-SILICO CLINICAL TRIALS IN DRUG DEVELOPMENT: A SYSTEMATIC REVIEW

- Systematic review of PubMed and clinicaltrials.gov for published articles and registered clinical trials related to in-silico clinical trials
- Preprint https://arxiv.org/abs/2503.08746

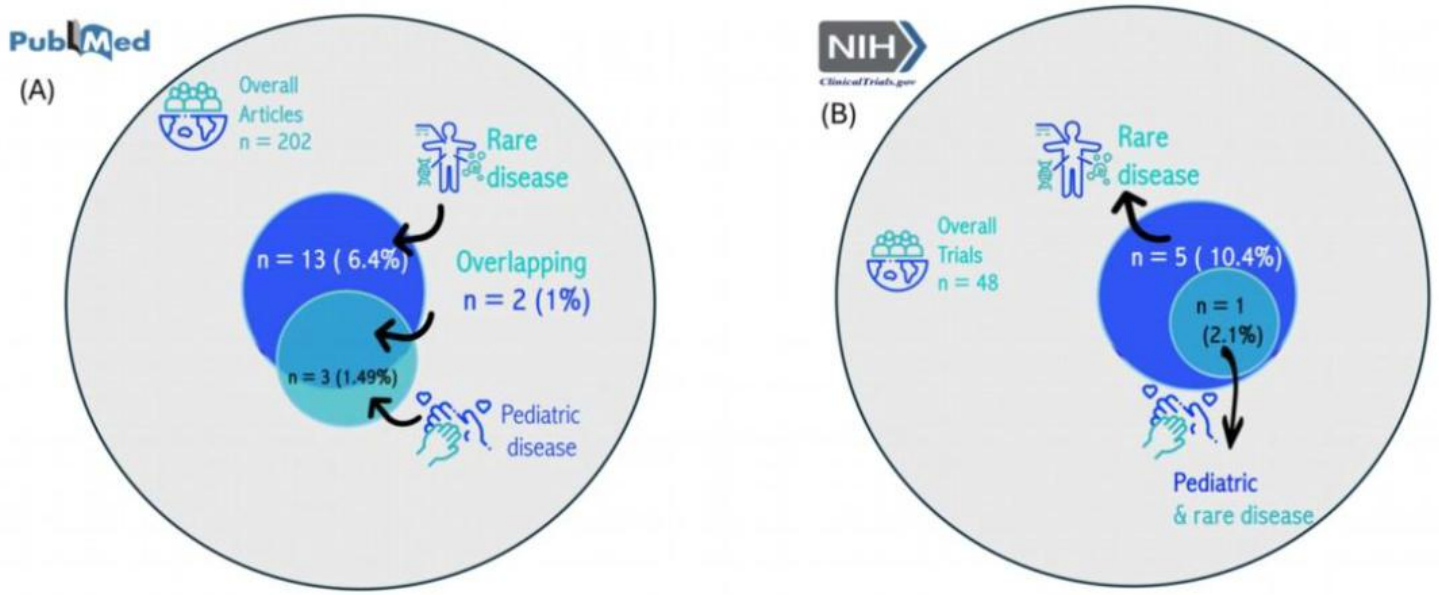


Figure 6. Panels (A) and (B) showthe number of rare and pediatric diseases out of all the pulications and trials.

CONCLUSIONS AND DISCUSSION

- **Increasing number of applications of AI/ML** in clinical research, e.g. AI based adjudication of clinical events, subgroup identification in clinical trials
- **Is there a role for statistics in AI / ML? Yes!**
  - Statistics contributes to **methods**, but impact on **applications** in my view even more important
- **AI fuelling innovative approaches such as in-silico clinical trials and digital twins**
- **Interdisciplinary collaboration and network**

FROM BLACK BOX TO PANDORA'S BOX

Pandora's box

43 languages

Article Talk | Read Edit View history Tools

From Wikipedia, the free encyclopedia

This article is about the mythological artifact. For other uses, see Pandora's box (disambiguation).

Pandora's box is an artifact in Greek mythology connected with the myth of Pandora in Hesiod's c. 700 B.C. poem *Works and Days*.[1] Hesiod related that curiosity led her to open a container left in the care of her husband, thus releasing curses upon mankind. Later depictions of the story have been varied, with some literary and artistic treatments focusing more on the contents than on Pandora herself.

The container mentioned in the original account was actually a large storage jar, but the word was later mistranslated. In modern times an idiom has grown from the story meaning "Any source of great and unexpected troubles",[2] or alternatively "A present which seems valuable but which in reality is a curse".[3]

https://en.wikipedia.org/wiki/Pandora%27s_box