CWI Centrum Wiskunde & Informatica

parnassia groep PHILIPS UMC Utrecht

RQ: Real-time evidence collection in data streams

Pitch at EPI general meeting, April 2022

Rosanne J. Turner

Supervisors and collaborators within EPI:

- Prof. Peter Grünwald (CWI)
- Prof. Floor Scheepers (UMCU)
- Karin Hagoort (UMCU)
- Dr. Aki Harma (Philips)
- Roel van Est (Parnassia Groep)

CWI Centrum Wiskunde & Informatica

parnassia groep PHILIPS UMC Utrecht

Aim to train a statistical recommender model sequentially, at multiple locations, with false positive rate guarantee and high power

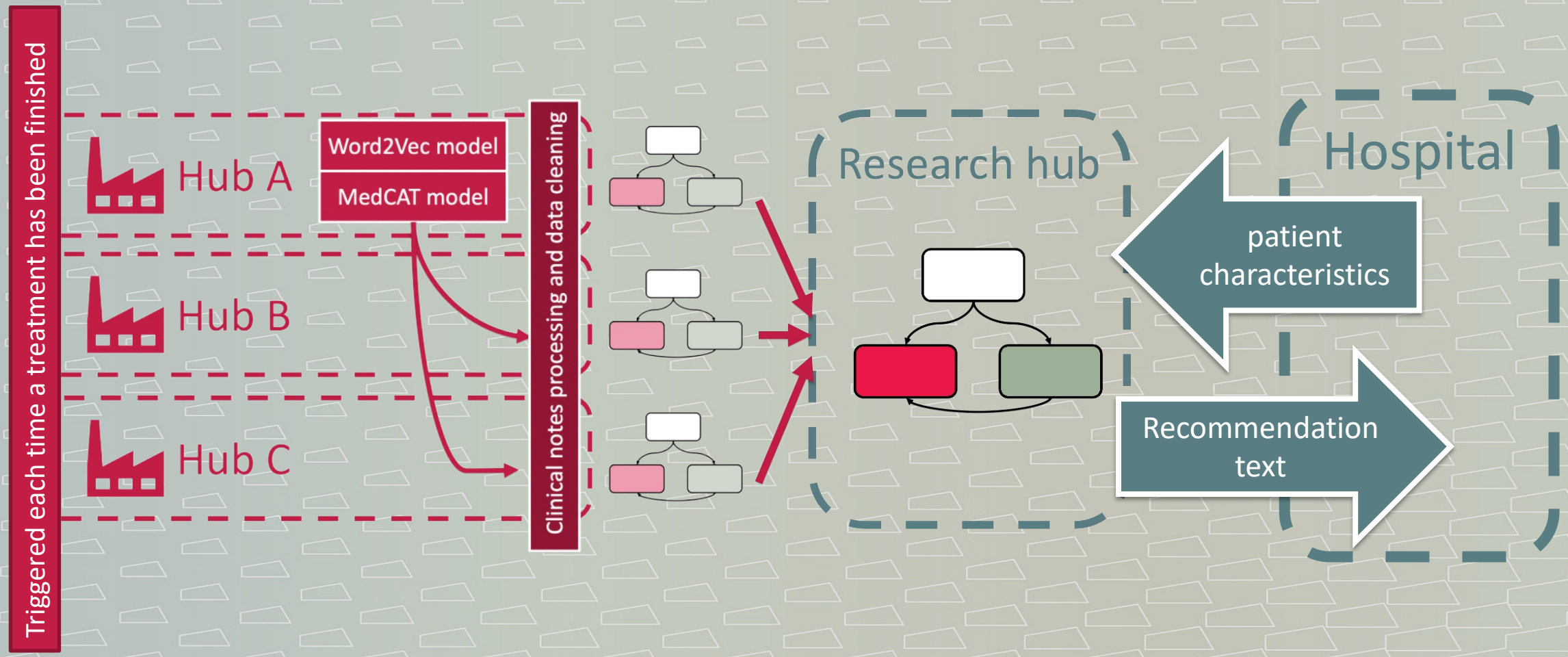
parnassia groep

PHILIPS

UMC Utrecht

CWI

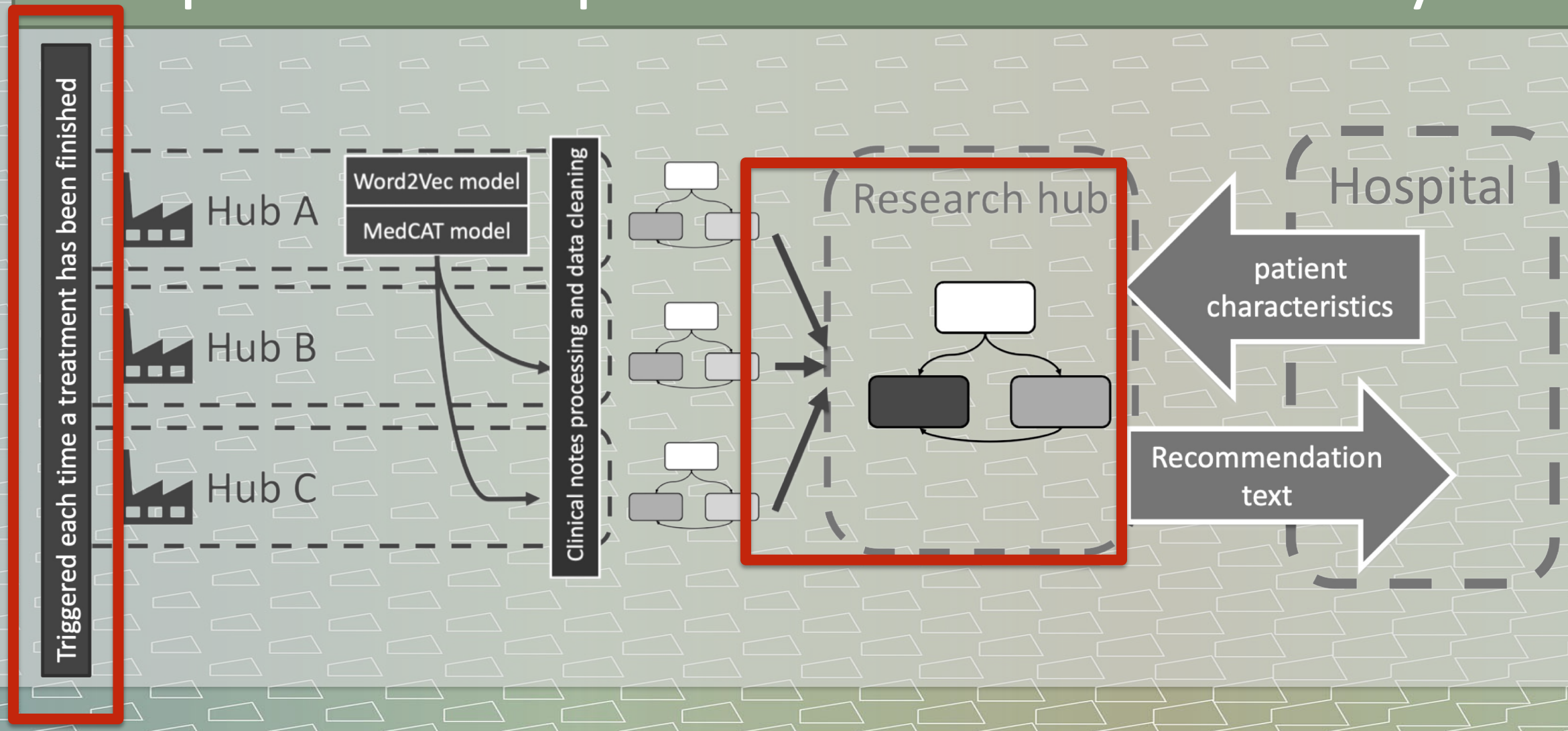
Centrum Wiskunde & Informatica



parnassia groep PHILIPS UMC Utrecht

CWI Centrum Wiskunde & Informatica

Sequential and partitioned statistical analysis



Loss of FPR guarantee in sequential or partitioned analysis

"Does choosing more invasive treatment B over standard treatment A make a difference in recovery rate of patients?"

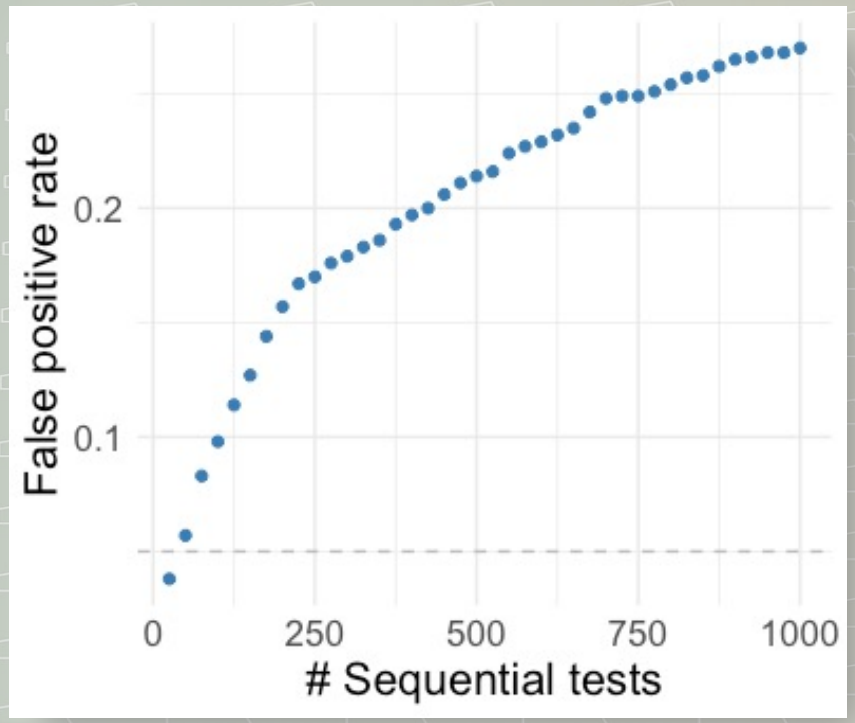


Figure adapted from Turner, Ly and Grunwald on ArXiv (2021)

Similar problem with partitioned analysis, see ter Schure and Grünwald in F1000Research (2019)

FPR guarantee is retained with e-value tests

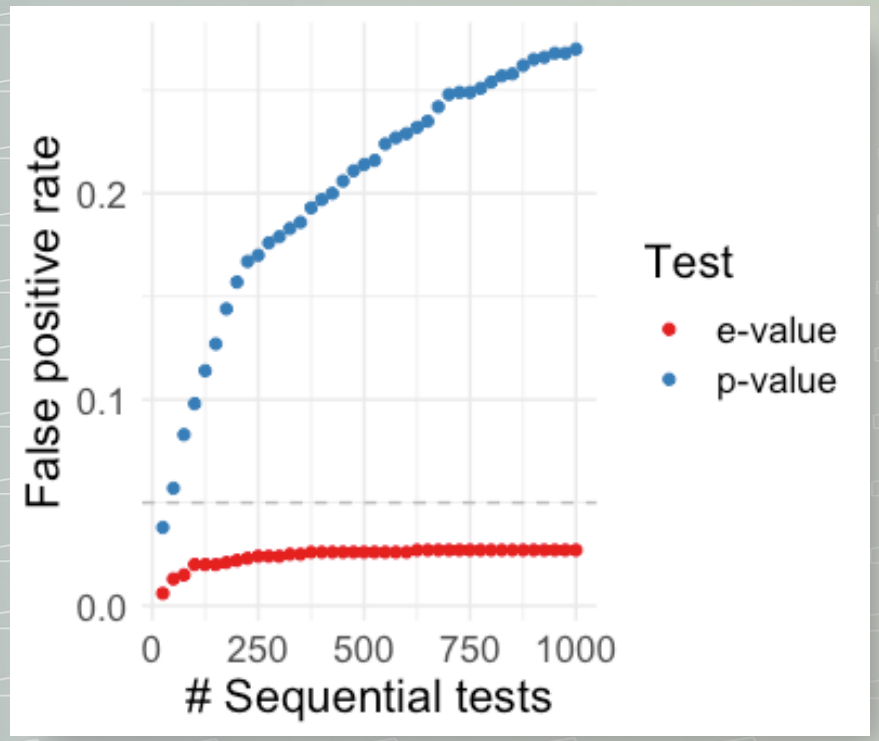


Figure adapted from Turner, Ly and Grunwald on ArXiv (2021)

E-value test:
Reject null hypothesis (no difference) when E exceeds $1/\alpha$, which gives us FPR guarantee at α under sequential and partitioned testing.*

*See Grunwald et al. on arXiv (2019) and Vovk and Wang in Annals of Statistics (2021)

Involving experts can help to decide more quickly

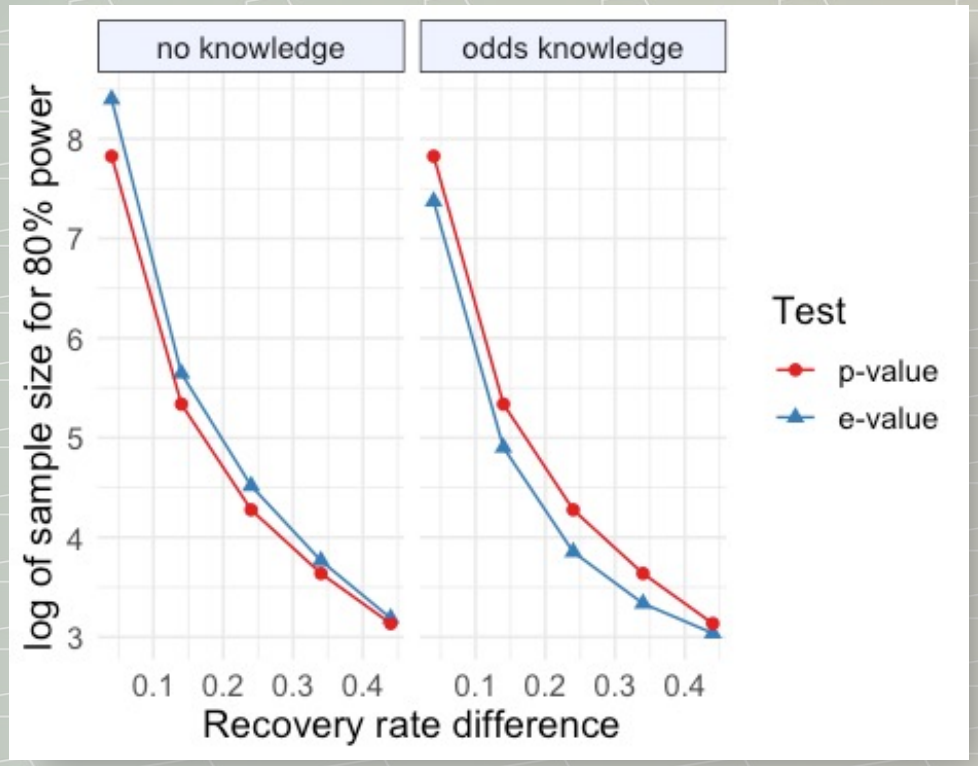


Figure adapted from Turner, Ly and Grunwald on ArXiv (2021)

Extension to anytime-valid confidence sequences

"What do we gain w.r.t. recovery rate when choosing more invasive treatment B over standard treatment A?"

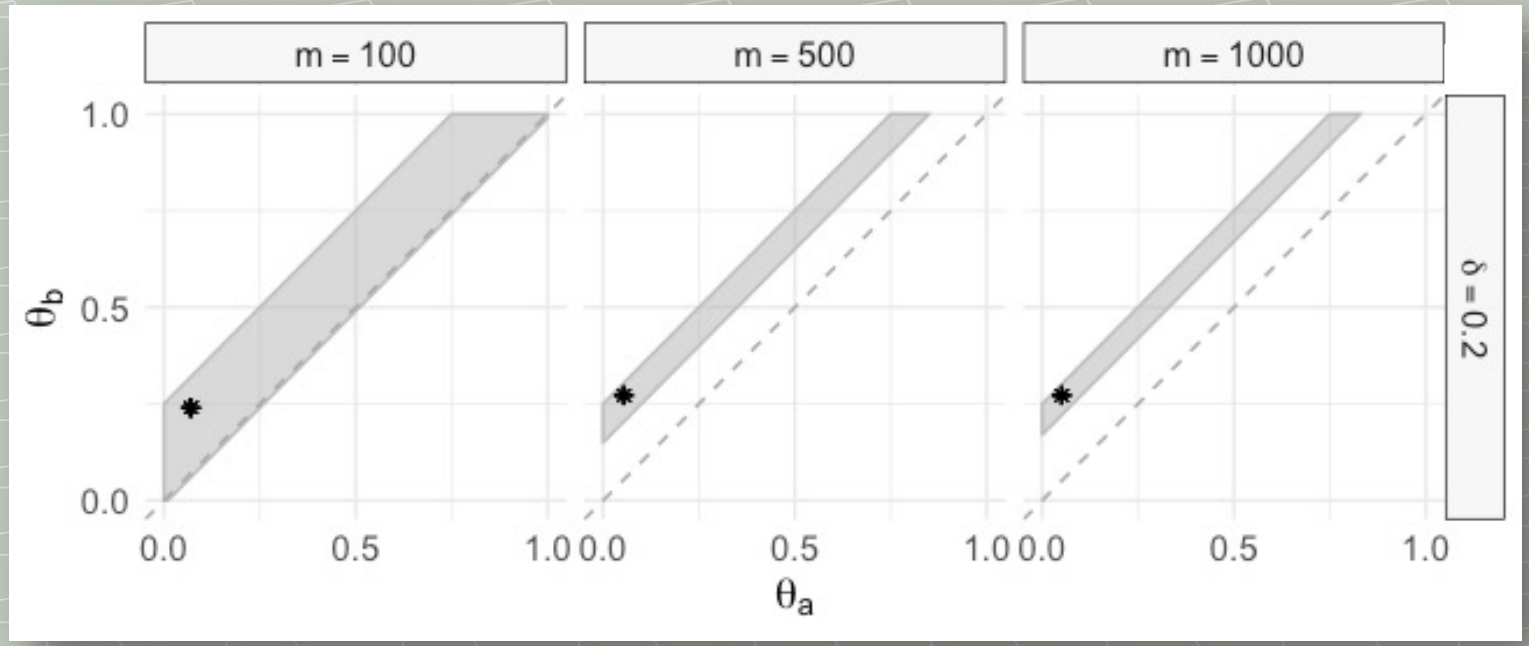


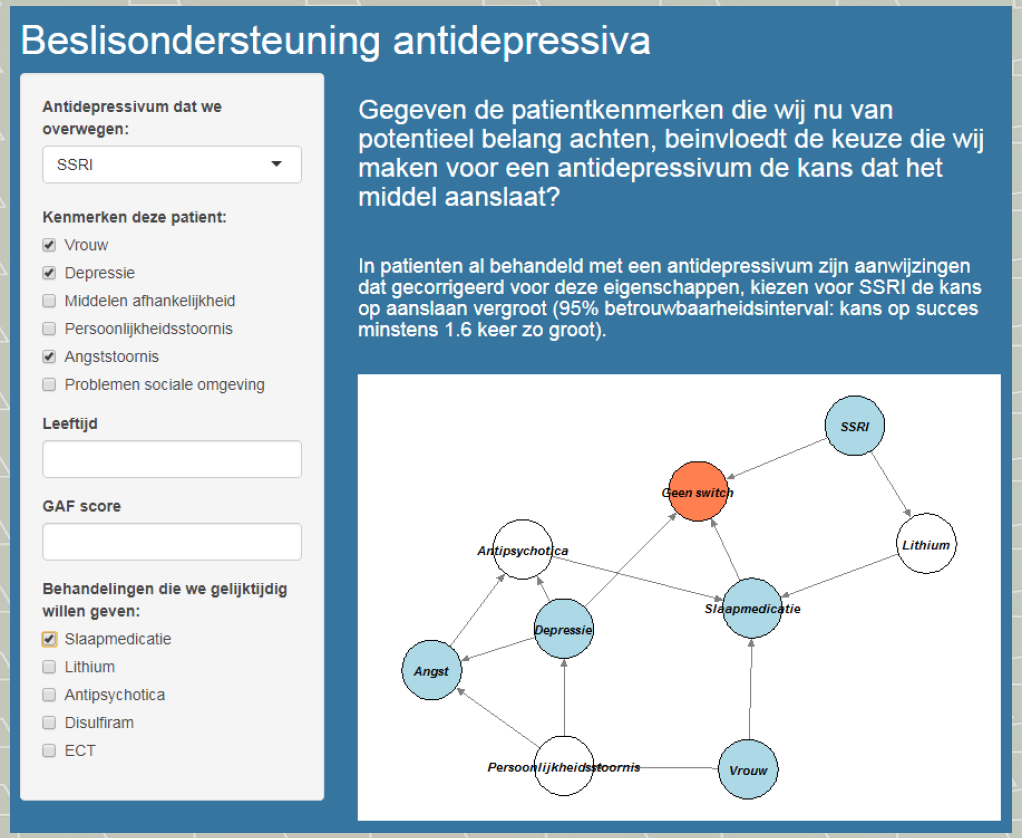
Figure adapted from Turner and Grunwald on ArXiv (2022)

parnassia groep PHILIPS UMC Utrecht

CWI Centrum Wiskunde & Informatica

Current work: combine e-values and use case

1. E-values for estimating effects for individual patients (e.g. logistic regression).
2. Combine with models developed for psychiatry use case:
   - Bayesian network for discovering conditional dependencies for antidepressant treatment
   - Hierarchical Bayesian logistic regression model with variable selection for ECT treatment
3. Proof of concept: recommender tool where uncertainty estimates are continuously updated



Pilot version, hypothetical patient

Further reading and references

Implementations of E-variables:

- R. Turner, A. Ly, and P. Grünwald, "Two-sample tests that are safe under optional stopping, with an application to contingency tables," arXiv preprint arXiv:2106.02693, 2021
- R. Turner, A. Ly, and P. Grünwald, "Anytime-valid Confidence Intervals for Contingency Tables and Beyond", arXiv preprint arXiv:2203.09785, 2022

On the theory of E-variables:

- P. Grünwald, R. de Heide, and W. Koolen, "Safe testing," arXiv preprint arXiv:1906.07801, 2019.
- Vovk and R. Wang, "E-values: Calibration, combination, and applications," Annals of Statistics, 2021

Our software package for R: https://cran.r-project.org/package=safestats