CWI Centrum Wiskunde & Informatica

parnassia groep PHILIPS UMC Utrecht

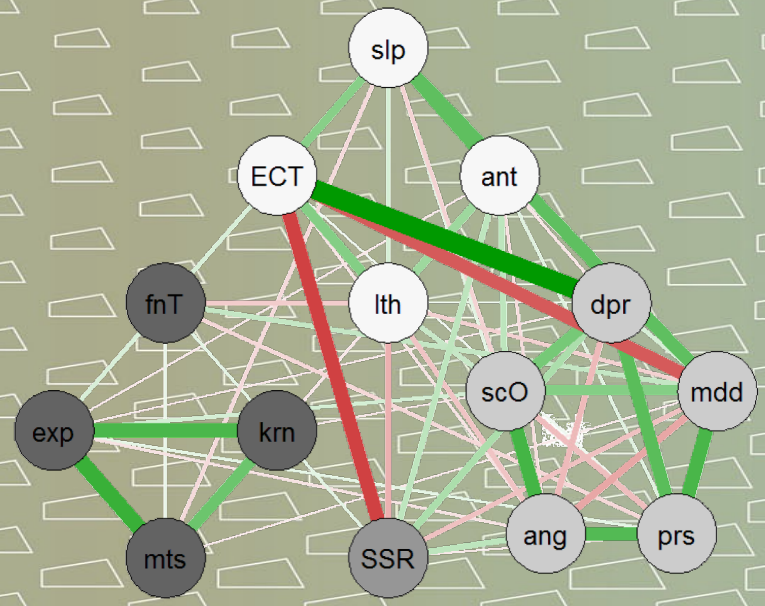
RQ: Real-time evidence collection in data streams

Updates July 2021

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)



CWI Centrum Wiskunde & Informatica

parnassia groep PHILIPS UMC Utrecht

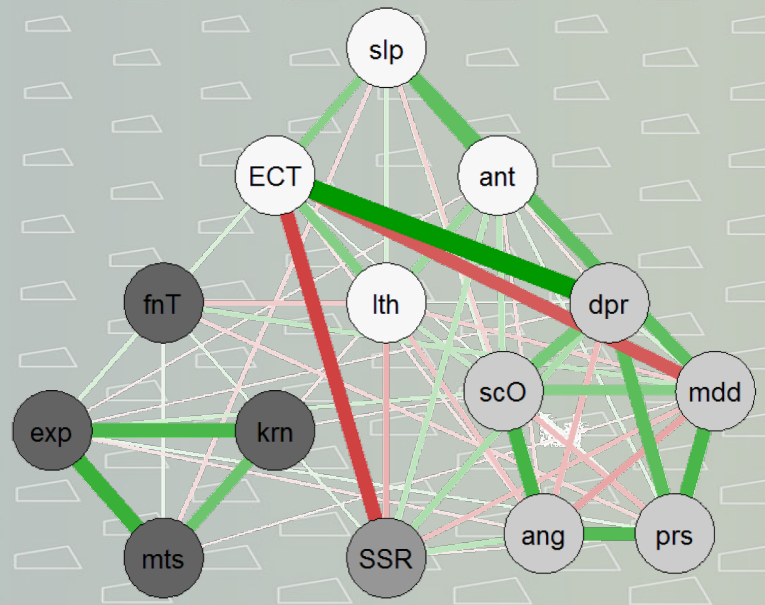
What have we made that can be used in practice so far?

1. Uitkomstmaat (outcome measures) dashboard at UMCU

parnassia groep PHILIPS UMC Utrecht

CWI Centrum Wiskunde & Informatica

Grand goal: aid clinicians and patient towards making a more informed decision, aimed at finding a useful treatment more quickly



Provide insights into (causal) relations of patient characteristics, **treatment outcome** and side effects in an app that utilizes real-time patient data

Current gold standard “treatment outcome” depressive symptoms: Hamilton questionnaire, since 1960!*

APPENDIX I

ASSESSMENT OF DEPRESSION

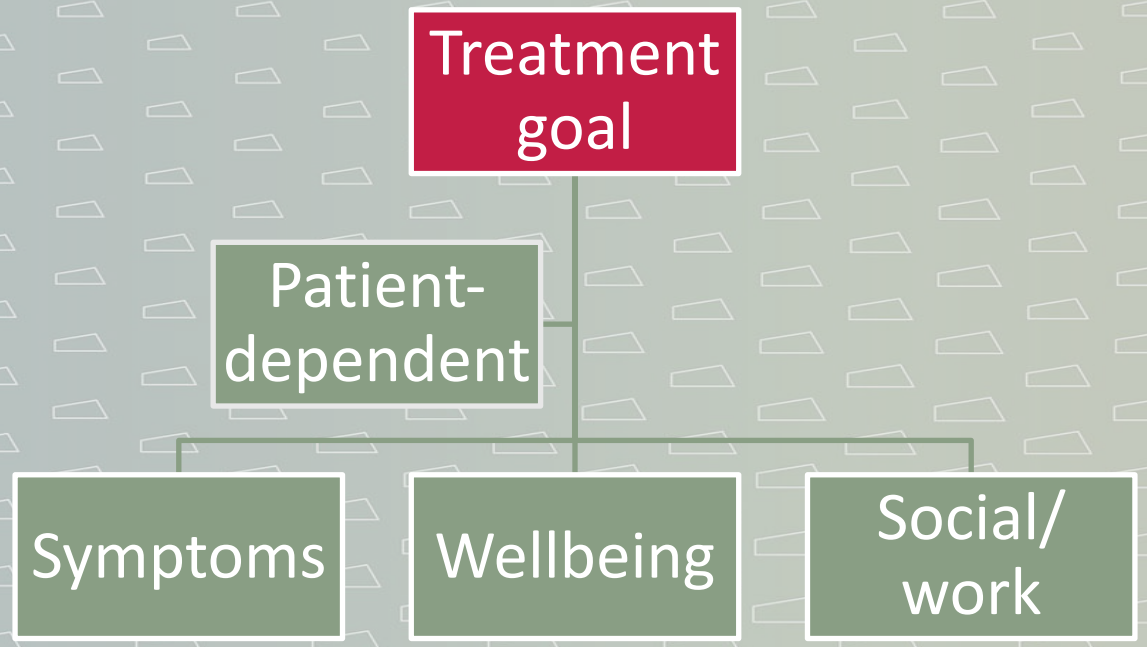
Item No.	Score Range	Symptom	Score
1	0-4	Depressed mood	
2	0-4	Guilt	
3	0-4	Suicide	
4	0-2	Insomnia, initial	
5	0-2	,, middle	
6	0-2	,, delayed	
7	0-4	Work and interests	
8	0-4	Retardation	
9	0-2	Agitation	
10	0-4	Anxiety, psychic	
11	0-4	,, somatic	
12	0-2	Somatic symptoms, gastrointestinal	
13	0-2	,, ,, general	
14	0-2	Genital symptoms	
15	0-4	Hypochondriasis	
16	0-2	Loss of insight	
17	0-2	,, ,, weight	
18	0-2	Diurnal variation { M / E	
19	0-4	Depersonalization, etc.	
20	0-4	Paranoid symptoms	
21	0-2	Obsessional symptoms	

Grading	
0	Absent
1	Mild or trivial
2, 3	Moderate
4	Severe
0	Absent
1	Slight or doubtful
2	Clearly present

- High administrative burden!
- NOT the treatment outcomes (individual) patients and clinicians are interested in

*Table from Max Hamilton, 1960

Clinicians' views on treatment goals



- 34 clinicians at UMCU were interviewed
- Thematic analysis revealed 4 major themes

Assessment of information available in EHR

Source	Theme(s)*	Structured	Availability	Relevance	Quality
Clinical notes	1, 2, 3, 4	No	High, 98-100%	High	High
Correspondence	1, 2	No	High, 87%	High	Medium
Kennedy Axis V questionnaire	1, 2	Mixed	Low, 21%	High	Medium
Crisis prevention plan	1	Mixed	Low, 1%	Medium	Low
Global Assessment of Functioning (GAF) end score	1, 2	Yes	High, 75%	Medium	Medium
Juridical status	1, 2	Yes	High, 100%	High	High
Destination after dismissal	1, 2	Yes	High, 100%	Medium	Medium
Medication use	1	Yes	High, 93%	Medium	Medium
Positive health questionnaire end score	4	Yes	Low, 17%	High	High
Lab measurements	1	Yes	High, 83%	Low	High

theme 1 = symptom reduction,
2 = social functioning,
3 = general well-being,
4 = patient's experience.

Previous work on NLP at UMCU/ PsyData by Vincent Menger

- PsyNLP library: rule- and dependency based entity and context detection
  - Word2vec model trained on UMCU clinical notes
  - Adapted slightly for this scenario

`psynlp` --- NLP functionality for psychiatric text

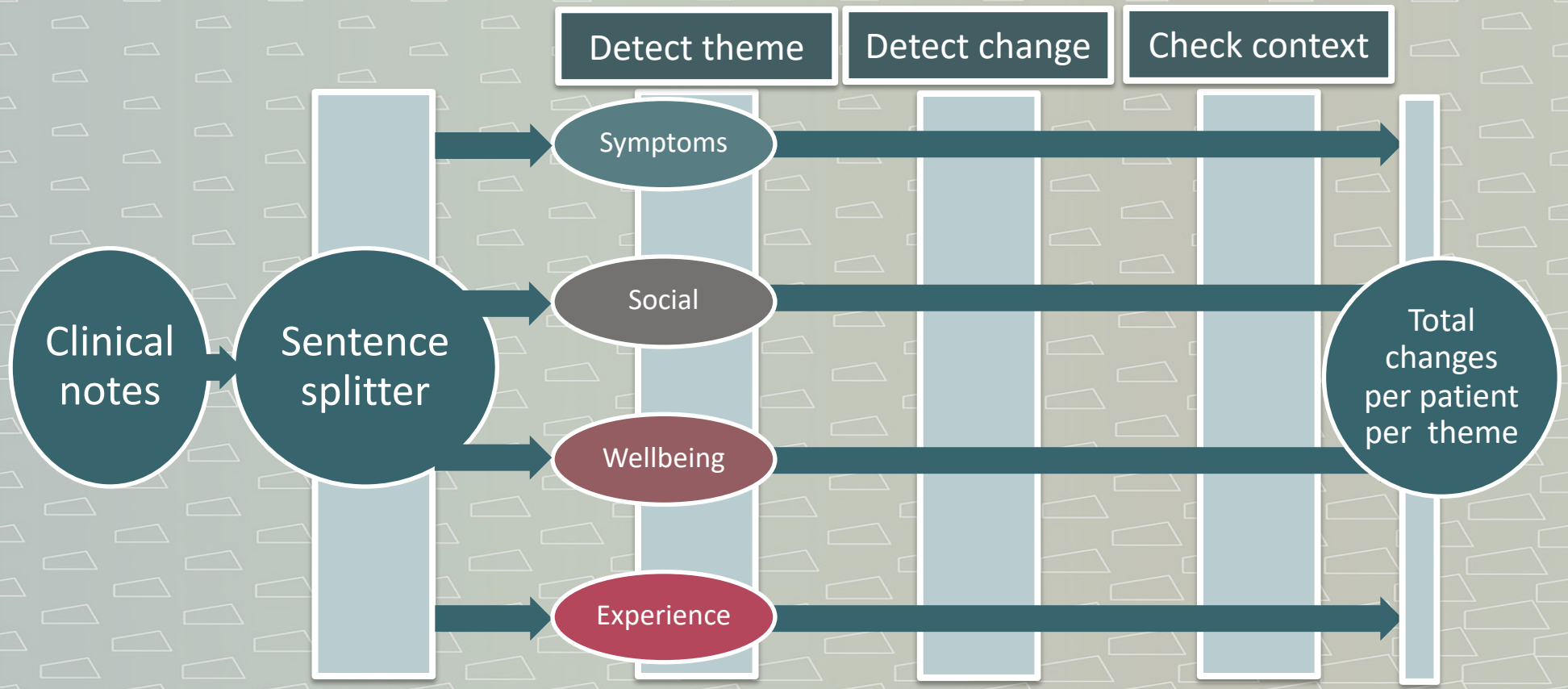
This package bundles some functionality for applying NLP (preprocessing) techniques to clinical text in psychiatry. Specifically, it contains the following `submodules`:

- `preprocessing` -- Preprocessing text
- `spelling` -- Spelling correction
- `entity` -- Entity matching
- `context` -- Detecting properties of entities (e.g. negation, plausibility) based on context

These submodules are further documented in their respective readmes, which you will find by following the links above.

https://github.com/vmenger/psynlp

Use PsyNLP to extract information on moments of change from each theme from clinical notes



Results of PsyNLP-based classifier in 2020 clinical notes

Theme	Mean number of sentences concerning theme	Mean number of sentences with relevant change in theme	Examples of sentences marked as correct	Classification accuracy of pipeline
Core symptoms	101	7.9	"Nervousness increased over the course of the day", "The patient appears drowsier than before"	0.988
Social	124	3.7	"Friendly, more in touch than yesterday"	0.997
Well-being	113	6.1	"This afternoon, the patient felt less well", "Had less energy"	0.953
Patient experience	158	9.2	"Says that it is going well, has the idea that it is going better and better"	0.993

CWI Centrum Wiskunde & Informatica

parnassia groep PHILIPS UMC Utrecht

Integration in dashboard

Work with data scientist Femke Coenen from PsyData, UMCU

Inleiding
Thema's
Afname klachten
Maatschappelijk functioneren
Algemeen welbevinden
Patiëntbeleving
Specifieke patiënt
Model
Contact PsyData

Selecteer Afdeling:

- A1 (105)
- A2 Jeugd (94)
- A2 Volwassen (162)
- A3 (114)
- A4 Kliniek (13)
- Totaal (488)

Project uitkomstmaat:

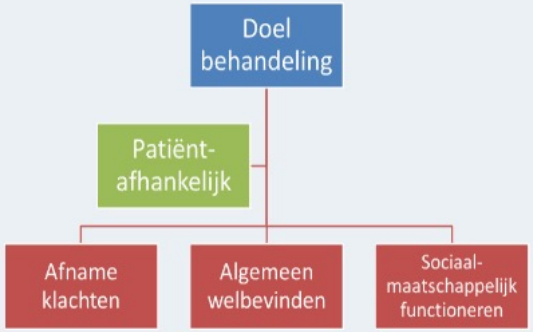
Aanleiding:
Welke patiënt is gebaat bij welke behandeling? Als we dat al vooraf beter zouden kunnen inschatten, dan zou dat de behandeling aanzienlijk verkorten. Bij PsyData proberen we dit te voorspellen op basis van bestaande EPD data met technieken uit de data science. Echter dit voorspellen is lastig omdat er verschillende (gelijkwaardige) manieren zijn om naar herstel te kijken. Om meer inzicht te krijgen in de factoren die bijdragen aan herstel, hebben we eerst een maat nodig die de uitkomst van de behandeling weergeeft.

Methode:
Graag willen we samen met zorgprofessionals en patiënten deze uitkomstmaat definiëren. Daarom hebben we tussen 23 juni en 27 juli 2020 een vragenlijst verstuurd, waarin het doel van de behandeling en definities van herstel centraal stonden. In totaal hebben 34 zorgprofessionals deze vragen beantwoord.
De respondenten bestonden uit:
- 10 psychiaters
- 9 A(N)IOS
- 12 verpleegkundigen
- 3 verpleegkundig specialisten

De antwoorden hebben we kwalitatief geanalyseerd, volgens de thematische analyse methode van Braun & Clarke. De gevonden thema's zijn weergegeven in onderstaande figuur. Vervolgens hebben we onderzocht of deze thema's terug te vinden zijn in de bestaande EPD data. Ten slotte, hebben we uitkomstmaten gedefinieerd voor al deze thema's en weergegeven in dit dashboard.

Op basis van deze uitkomstmaten maken we behandeling (zoals medicatie of ECT) onderzoekbaar en krijgen we meer inzicht in de effectiviteit van de behandelingen. Uiteindelijk kunnen we met deze inzichten betere zorg bieden aan patiënten, in de vorm van precision psychiatry.

Rosanne Turner, Femke Coenen & Karin Hagoort



Behandeluitkomsten

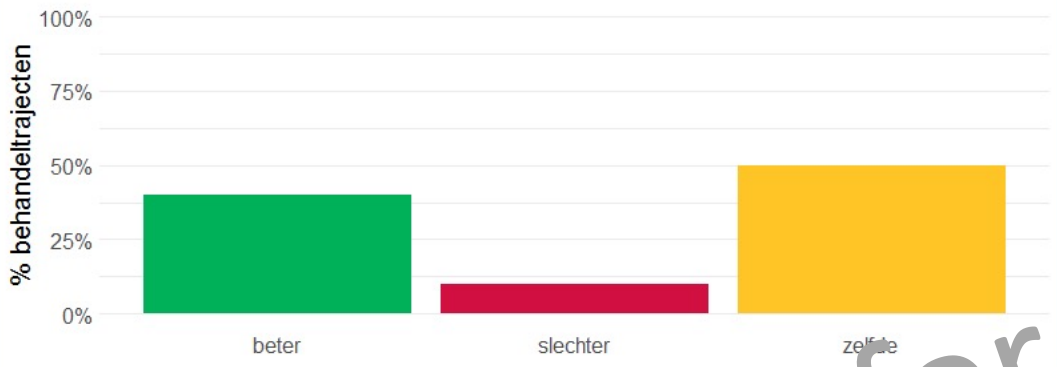
- Inleiding
- Thema's
- Afname klachten
- Maatschappelijk functioneren
- Algemeen welbevinden
- Patiëntbeleving
- Specifieke patiënt
- Model

Contact PsyData

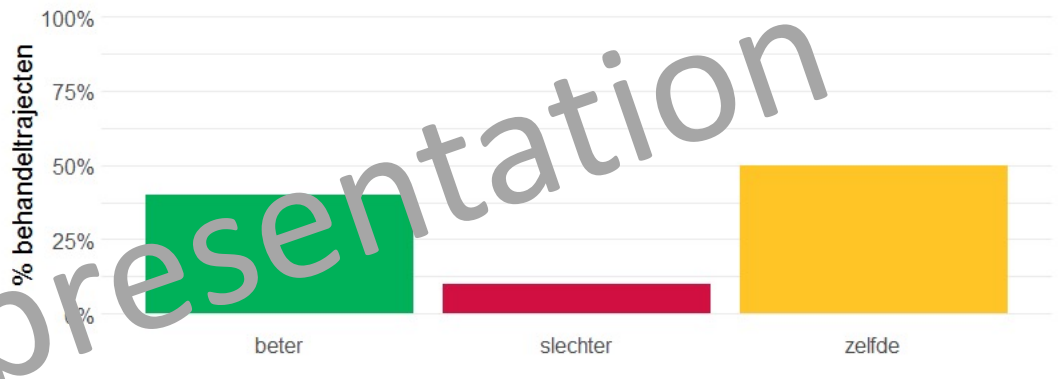
Selecteer Afdeling:

- A1 (105)
- A2 Jeugd (94)
- A2 Volwassen (162)
- A3 (114)
- A4 Kliniek (13)
- Totaal (488)

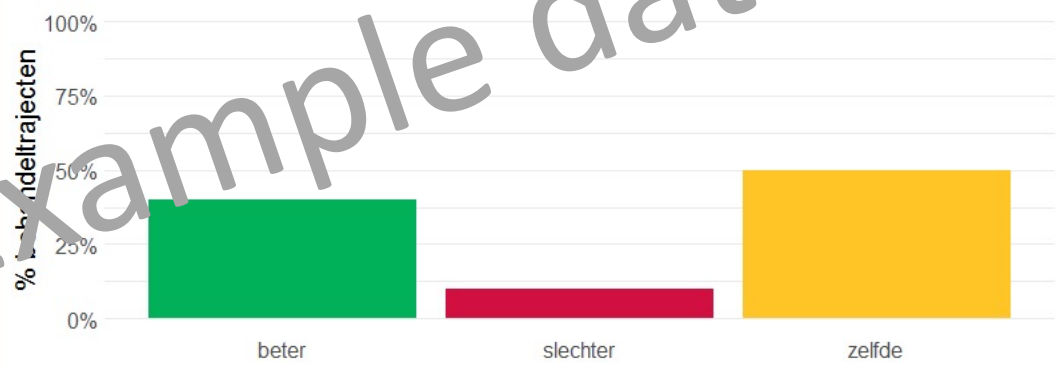
Afname klachten: totaalscore



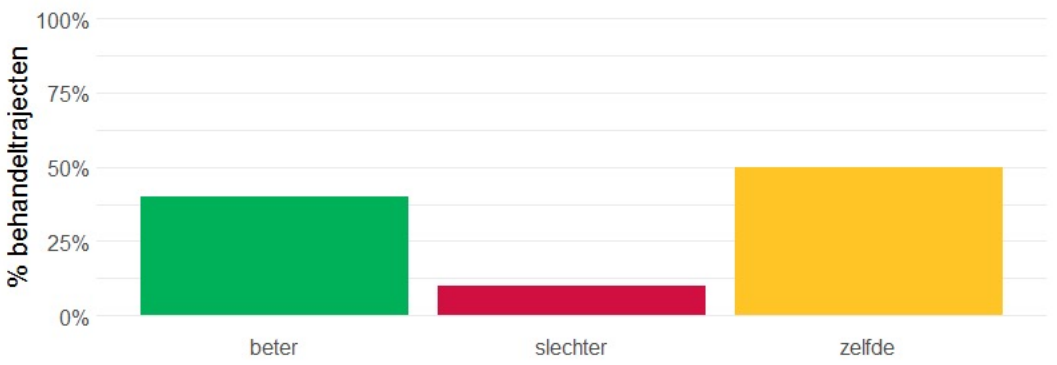
Maatschappelijk functioneren: totaalscore



Algemeen welbevinden: totaalscore



Beleving patiënt: totaalscore



Behandeluitkomsten

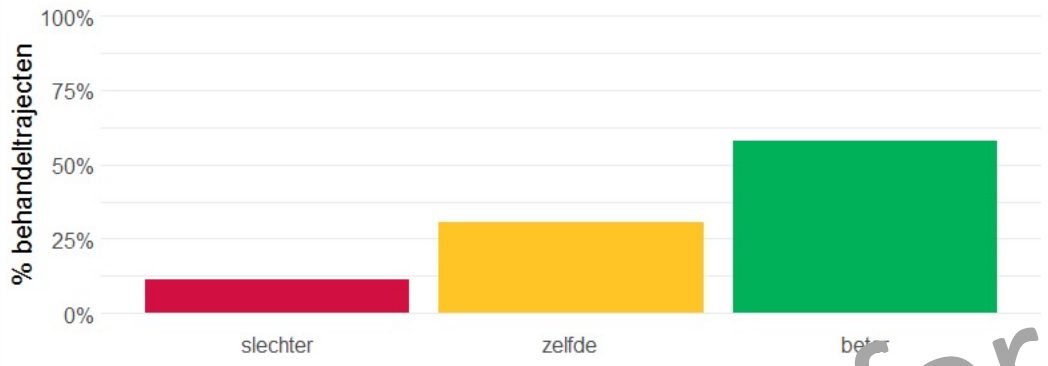
- Inleiding
- Thema's
- Afname klachten
- Maatschappelijk functioneren
- Algemeen welbevinden
- Patiëntbeleving
- Specifieke patiënt
- Model

Contact PsyData

Selecteer Afdeling:

- A1 (105)
- A2 Jeugd (94)
- A2 Volwassen (162)
- A3 (114)
- A4 Kliniek (13)
- Totaal (488)

Afname klachten in rapportages



Verbetermomenten gecorrigeerd met verslechtermomenten van symptomen in decursus

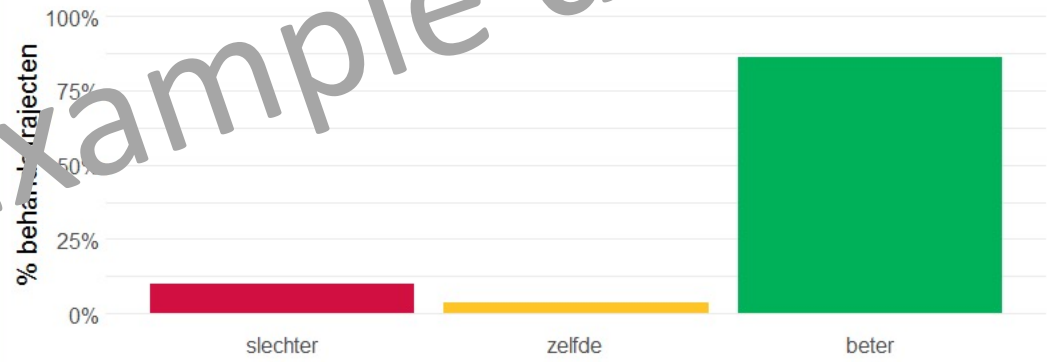
Toelichting

In de decursus werd gezocht naar zinnen die duiden op een positieve/negatieve verandering in de psychische klachten van de patiënt.

Hiervoor werden bijv. onderstaande termen gebruikt i.c.m. veranderwoorden:

patronen
aandachtsproblematiek
aanpassingsstoornis
achterdocht
achterd[illegible]
[illegible]volgingswaan
agitatie

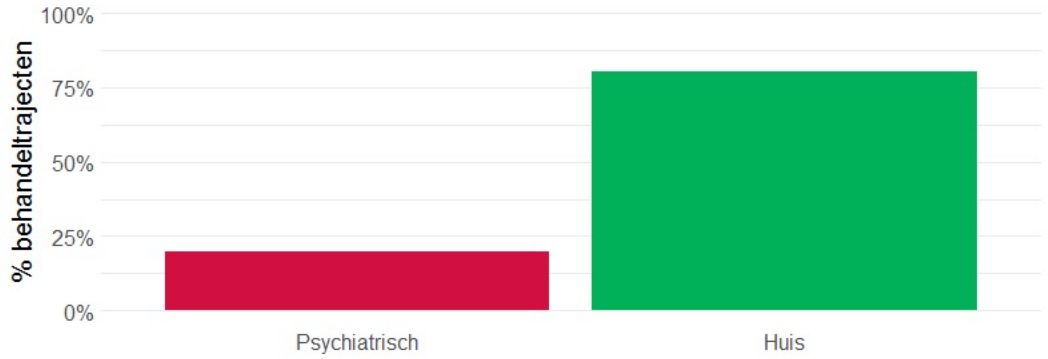
Juridische status einde traject



Uitleg:

- Eind vrijwillig/ geen informatie: groen
- Eind crisis: rood
- Eind RM na crisis of RM: geel

Bestemming na ontslag



Behandeluitkomsten

- Inleiding
- Thema's
- Afname klachten
- Maatschappelijk functioneren
- Algemeen welbevinden
- Patiëntbeleving
- Specifieke patiënt
- Model

Contact PsyData

Selecteer Afdeling:

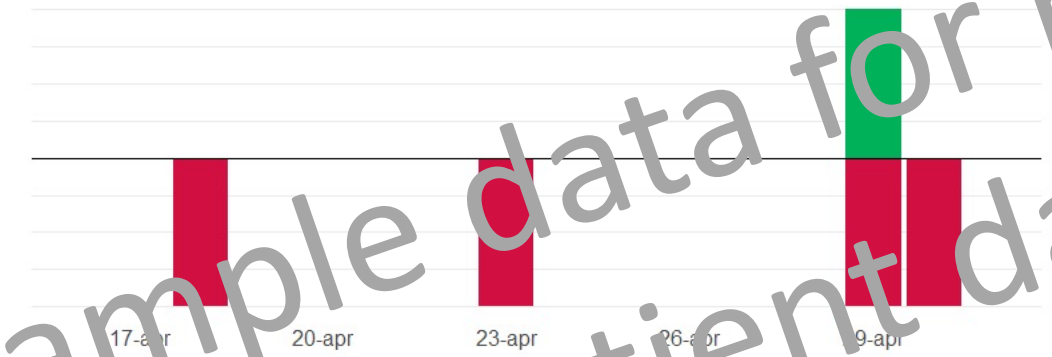
- A1 (105)
- A2 Jeugd (94)
- A2 Volwassen (162)
- A3 (114)
- A4 Kliniek (13)
- Totaal (488)

Selecteer patiënt:

Patiëntnummer:

100124575

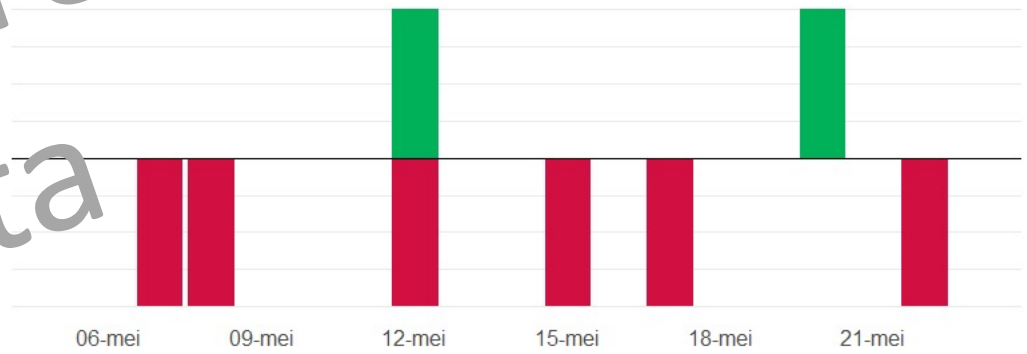
Behandeltraject 1



Selectie van verandermomenten:

DATUM	Tekst
2020-04-18	Sinds aantal dagen toenemende psychotische klachten.

Behandeltraject 2



Selectie van verandermomenten:

DATUM	Tekst
2020-05-07	Heropname i.v.m. verslechtering klachten in de thuissituatie.

parnassia groep PHILIPS UMC Utrecht

CWI Centrum Wiskunde & Informatica

This July: deployment through Rstudio Connect in collaboration with central IT department UMCU

CWI Centrum Wiskunde & Informatica

parnassia groep PHILIPS UMC Utrecht

What have we made that can be used in practice so far?

2. Safestats package for real-time data analysis

parnassia groep PHILIPS UMC Utrecht

CWI Centrum Wiskunde & Informatica

The Safestats package

github.com

AlexanderLyNL/safestats

README.md

Safestats: an R package for safe, anytime-valid inference.

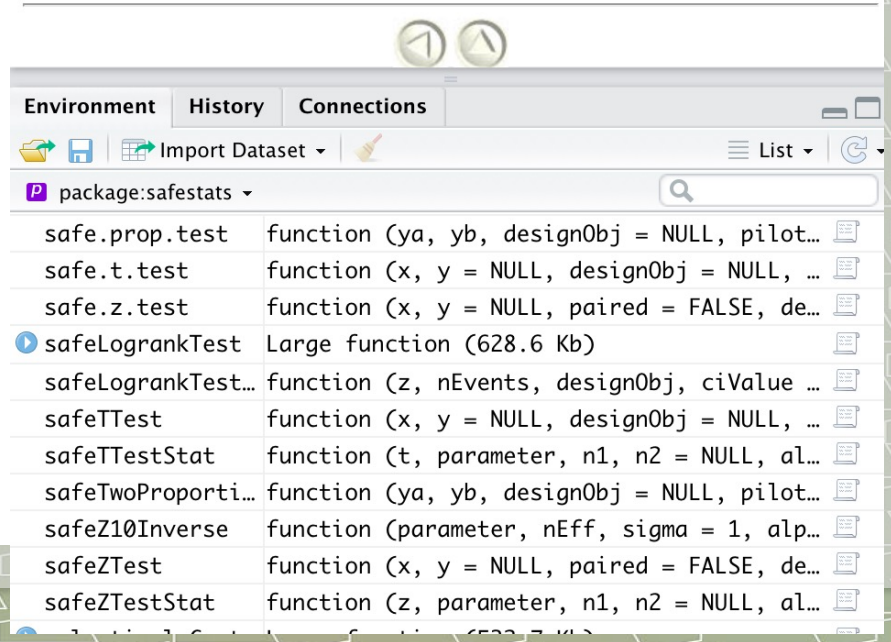
`Safestats` is for designing and applying **safe** hypothesis tests. It can be used for designing hypothesis tests in the prospective or randomised controlled trial (RCT) setting, where the tests can be used under optional stopping and experiments often can be stopped early. The 'pilot' functions in the package also enable using the safe tests in the observational/ retrospective setting. For examples and explanation about which test to choose for which study setup, our `vignette` can be used (see below). The current version includes safe t-tests and tests of two proportions. The initial paper on the theory of safe testing and a worked-out example for the t-test can be found in this paper. More on the theory behind the development of the safe tests for proportions can be found here.

Installation

The development version can be found on GitHub, which can be installed with the `devtools` package from CRAN by entering in `R`:

```
devtools::install_github("AlexanderLyNL/safestats", build_vignettes = TRUE)
```

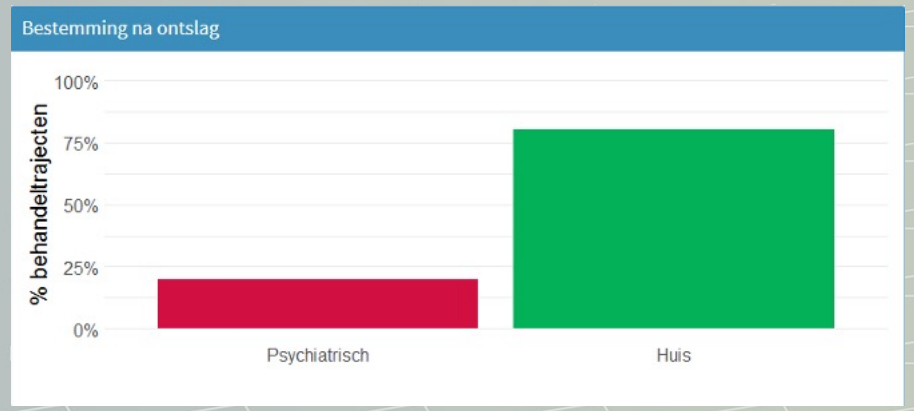
Safe Anytime-Valid Inference



https://github.com/AlexanderLyNL/safestats

Example: next location of patients with treatment-resistant depression (I)

Aim: investigate if intervention ECT is associated with increased probability of going home after treatment with a hypothesis test at significance level 0.05




```
> designSafeTwoProportions(na = 1, nb = 1, alpha = 0.05, beta = 0.8, delta = 0.2)
Simulating E values and stopping times for divergence between groups of  0.2
================================================================================

        Safe Test of Two Proportions Design

             na, nb, nBlocksPlan = 1.0, 1.0, 60.8
              minimal difference = 0.2
                     alternative = two.sided
          alternative restriction = none
                    power: 1 - beta = 0.2
 parameter: Beta hyperparameters = standard, REGRET optimal
                           alpha = 0.05
decision rule: e-value > 1/alpha = 20

Timestamp: 2021-06-24 09:21:32 CEST
```

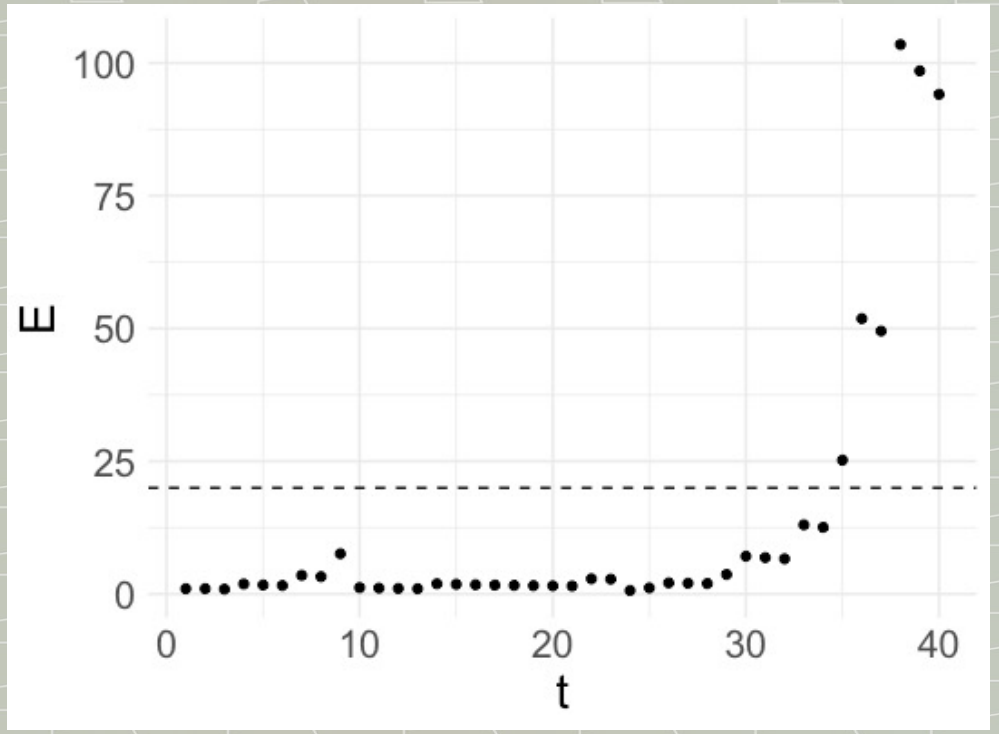
Example: next location of patients with treatment-resistant depression (II)

This example: assume probability 0.7 of going home after ECT, and 0.9 after only medication treatment

```
groupECT <- sample(c(1, 0), replace = TRUE, prob = c(0.7, 0.3), size = 60)
groupMed <- sample(c(1, 0), replace = TRUE, prob = c(0.9, 0.1), size = 60)

eValues <- numeric(length(groupECT))
for(blockNumber in seq_along(groupECT)){
  safeTestResult <- safeTwoProportionsTest(
    ya = groupECT[1:blockNumber],
    yb = groupMed[1:blockNumber],
    designObj = safeDesign
  )

  eValues[blockNumber] <- safeTestResult$eValue
}
```



Example: next location of patients with treatment-resistant depression (II)

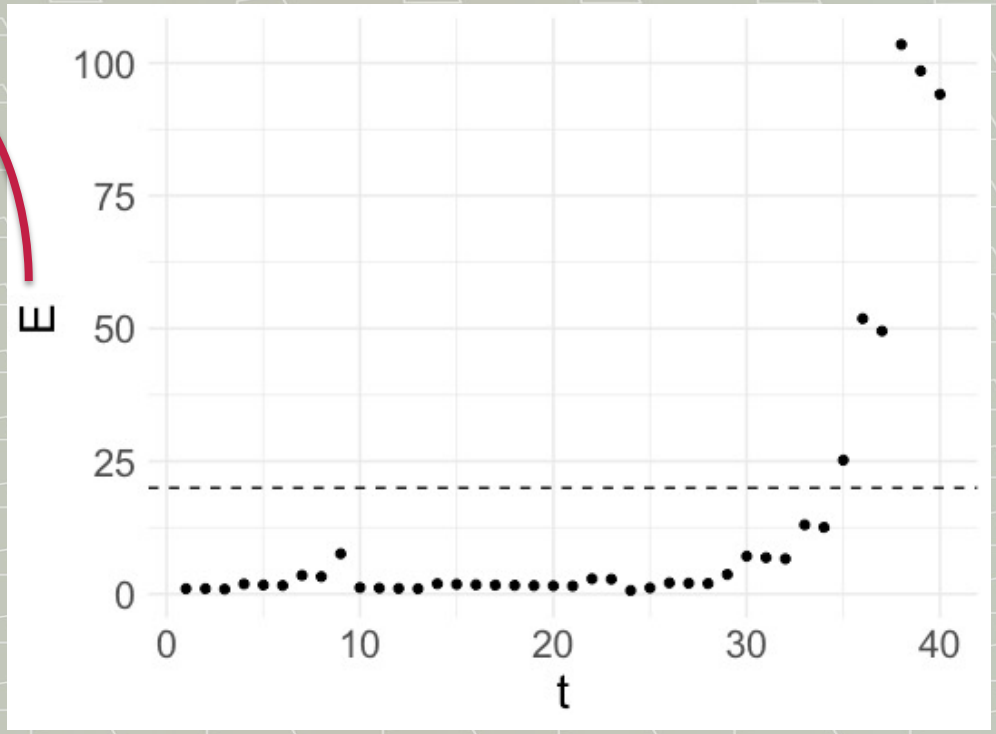
This exam
going hon
medicatio

Imagine partner clinical hubs running the same experiment: can simply exchange the E variables at any moment and multiply them to obtain collective evidence.

```
groupECT <- sample(
groupMed <- sample(c(1, 0), replace = TRUE, prob = c(0.9, 0.1), size = 60)

eValues <- numeric(length(groupECT))
for(blockNumber in seq_along(groupECT)){
  safeTestResult <- safeTwoProportionsTest(
    ya = groupECT[1:blockNumber],
    yb = groupMed[1:blockNumber],
    designObj = safeDesign
  )

  eValues[blockNumber] <- safeTestResult$eValue
}
```



Example: next location of patients with treatment-resistant depression (II)

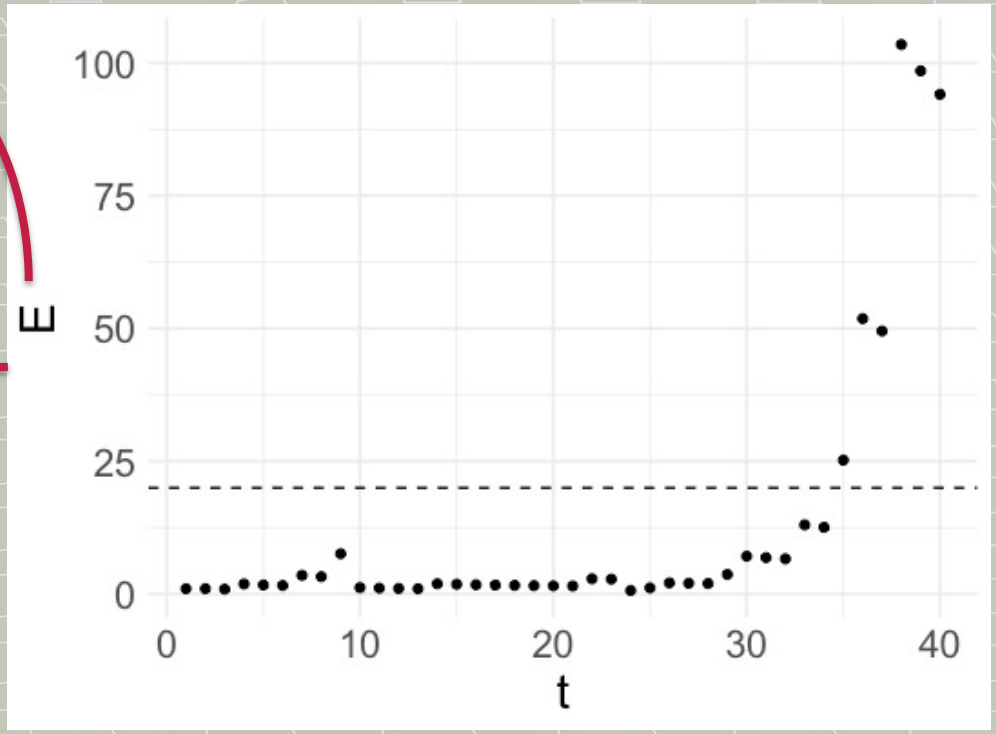
This exam
going hon
medicatio

Imagine partner clinical hubs running the same experiment: can simply exchange the E variables at any moment and multiply them to obtain collective evidence.

```
groupECT <- sample(
groupMed <- sample(c(1, 0), replace = TRUE, prob = c(0.9, 0.1), size = 60)

eValues <- numeric(length(groupECT))
for(blockNumber in seq_along(groupECT)){
  safeTestResult <- safeTwoProportionsTest(
    ya
    yb
    des
  )

  eValu
}
```



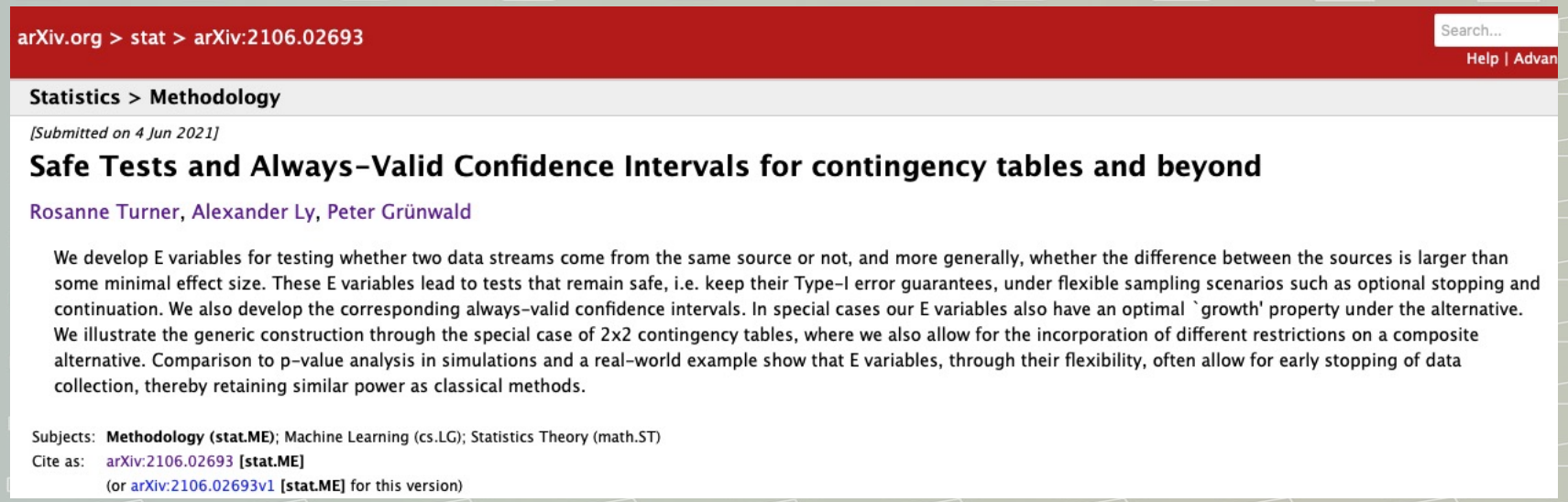
Can directly be used for scientific communication: easily convertible to common measures, such as p-values and confidence intervals

parnassia groep PHILIPS UMC Utrecht

CWI Centrum Wiskunde & Informatica

Testing for stream data paper on ArXiv

https://arxiv.org/abs/2106.02693



arXiv.org > stat > arXiv:2106.02693

Search...

Help | Advan

Statistics > Methodology

[Submitted on 4 Jun 2021]

Safe Tests and Always-Valid Confidence Intervals for contingency tables and beyond

Rosanne Turner, Alexander Ly, Peter Grünwald

We develop E variables for testing whether two data streams come from the same source or not, and more generally, whether the difference between the sources is larger than some minimal effect size. These E variables lead to tests that remain safe, i.e. keep their Type-I error guarantees, under flexible sampling scenarios such as optional stopping and continuation. We also develop the corresponding always-valid confidence intervals. In special cases our E variables also have an optimal `growth' property under the alternative. We illustrate the generic construction through the special case of 2x2 contingency tables, where we also allow for the incorporation of different restrictions on a composite alternative. Comparison to p-value analysis in simulations and a real-world example show that E variables, through their flexibility, often allow for early stopping of data collection, thereby retaining similar power as classical methods.

Subjects: **Methodology (stat.ME)**; Machine Learning (cs.LG); Statistics Theory (math.ST)
Cite as: arXiv:2106.02693 **[stat.ME]**
(or arXiv:2106.02693v1 **[stat.ME]** for this version)

CWI Centrum Wiskunde & Informatica

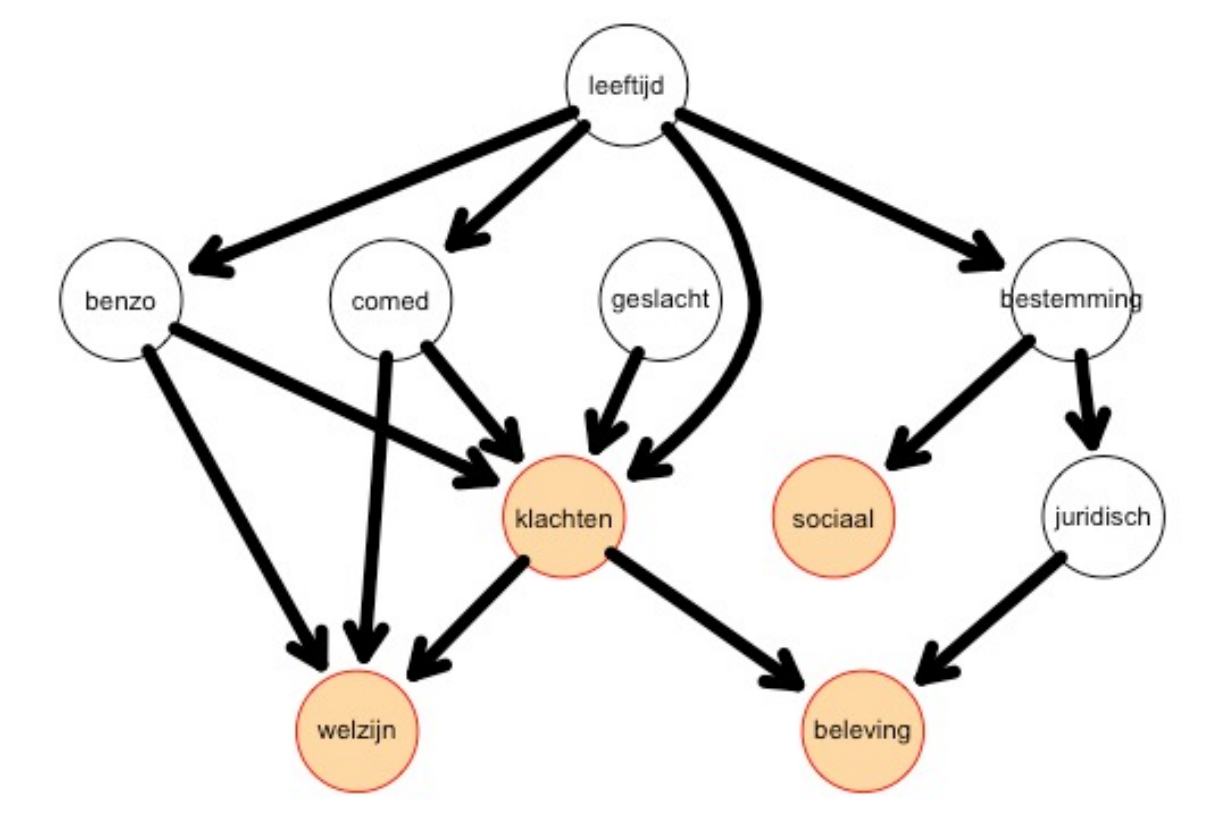
parnassia groep PHILIPS UMC Utrecht

What is the progress you plan to make during the coming half year?

Combining the two projects

(and a lot of writing..)

Networks visualizing (causal) relations between our outcomes, interventions and characteristics



- Developed for multiple patient groups:
  - Antidepressant users at UMCU
  - Antidepressant users at Parnassia
  - ECT clients at UMCU
- Directed and undirected variants: which one to use? *Or prediction only?*
- Develop versions with safe confidence measures for confidence estimation in practice