Research Synthesis & Big Data Conference, online, May 2021

Universität Trier

Using complex network analysis based on intensive longitudinal data to predict treatment dropout in patients with mood and anxiety disorders

Brian Schwartz and Wolfgang Lutz
University of Trier

Treatment Dropout

mean dropout rate = 19.7 % (range = 0 – 74.2)

Swift & Greenberg (2012). J Consult Clin Psych, 80(4), 547–599.

Individual‘s perspective

- poor treatment outcomes
- higher hospitalization rates
- continuing psychological disorders
- strain for patients‘ relatives

Society‘s perspective

- inefficient use of clinical personnel
- strains the health system
- reduced productivity of patients
- increasing mental health costs

Predicting dropout based on a single measurement

- *N* = 707 patients treated by 66 therapists (*M* = 10.71 patients per therapist, range = 3 – 30)
- Impairment, interpersonal problems, treatment expectations, personality style, socio-demographics
- Multilevel logistic regression model

$R^2_{\text{GLMM(marginal)}}$ = .105

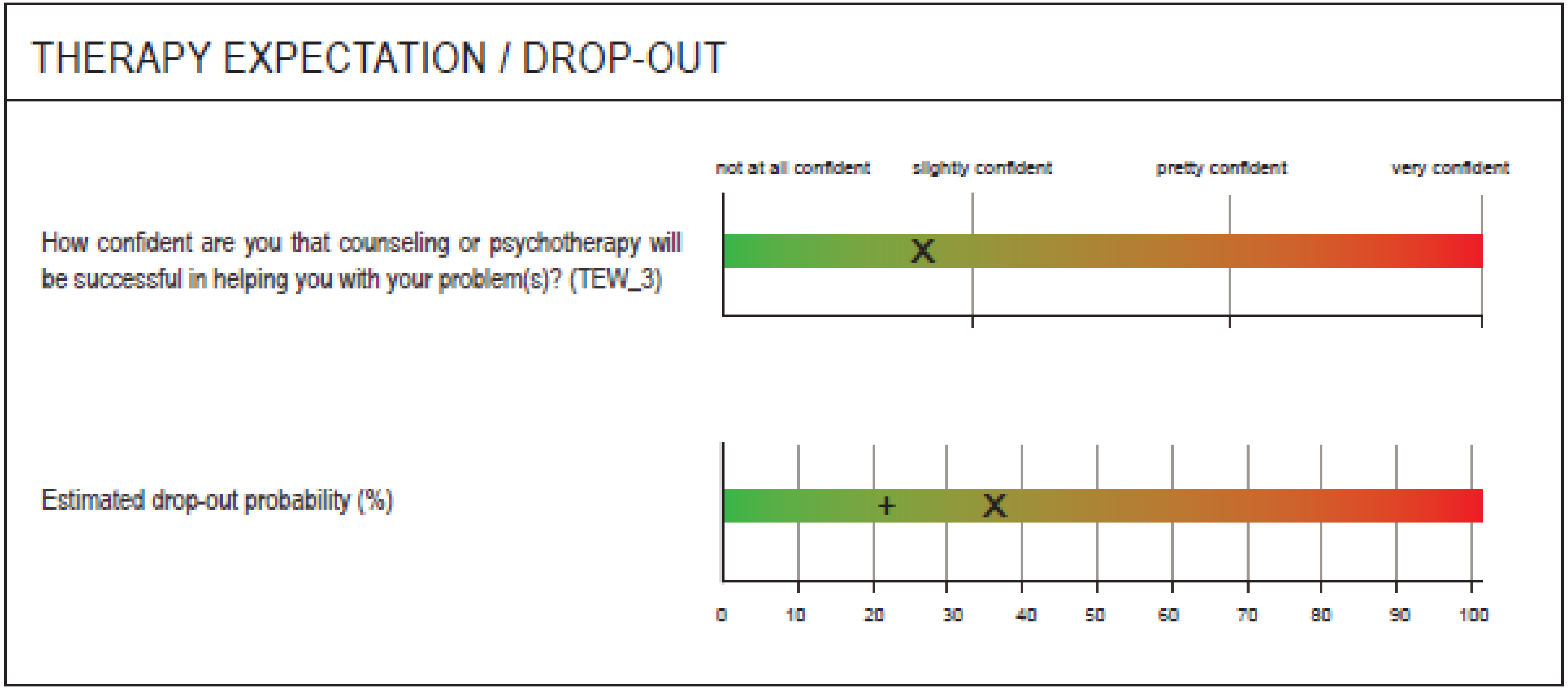
AUC = .686

Fixed effects on final model

	estimate	*P*-value
BSI – global severity index	0.396	.007
Age	–0.005	.549
Sex	0.458	.029
Education middle	–0.102	.691
Education high	-0.590	.022
PSSI-K – obsessive-compulsive	-0.390	.010
PSSI-K – histrionic	-0.395	.011
Treatment expectation 1	–0.254	.016
Treatment expectation 2	–0.457	.001

Zimmermann et al. (2017). Clin Psychol Psychot, 24(2), 312–321.

Dropout prediction in the Trier Treatment Navigator



Grant no. DFG LU 660-10/1

Dropout prediction in the Trier Treatment Navigator

Trier Treatment Navigator
Decision Support System
Trier University

- Archival data (*N* = 1234), 22.6% dropout

Variable selection

78 pre-therapy variables

Bivariate correlation with dropout

30 variables

Bootstrap ranking LASSO

7 variables

Model: Generalized linear model

Final prediction model for dropout

	BRLasso	GLM	p-value
FEP-2	-0.230	–0.697	.001
HSCL-11	0.261	0.609	.001
PSSI-K – histrionic	0.322	0.359	.001
OQ-30 – interpersonal relationships	0.411	0.530	<.001
PSSI-K – obsessive-compulsive	-0.416	–0.320	.004
Treatment expectation (therapist)	-0.509	–0.513	<.001
High school education	-0.586	–0.610	<.001

$R^2_{McFadden}$ = .120

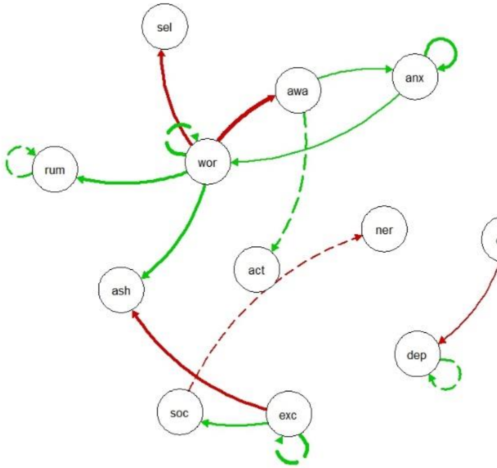
AUC = .698

Intensive longitudinal data and network analysis

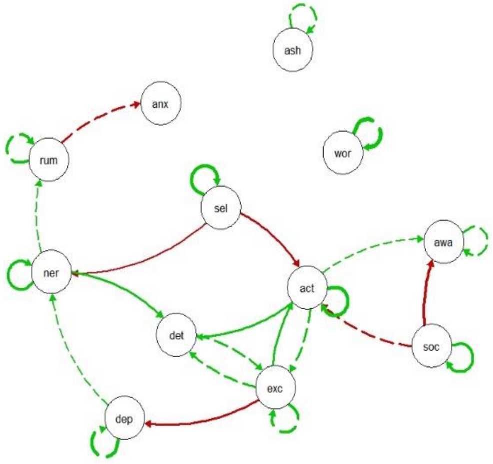
- Pre-therapy ambulatory assessments (4 times/day * 14 days)
- *N* = 3248 observations nested within *n* = 58 patients
- Positive and negative affect, rumination, worry, self-efficacy, social support (16 items)

- **Longitudinal Network analysis**:
  - Multilevel vector autoregressive (mlVAR) models
  - Each variable regressed on lagged versions of all variables (lag 1)

dropouts



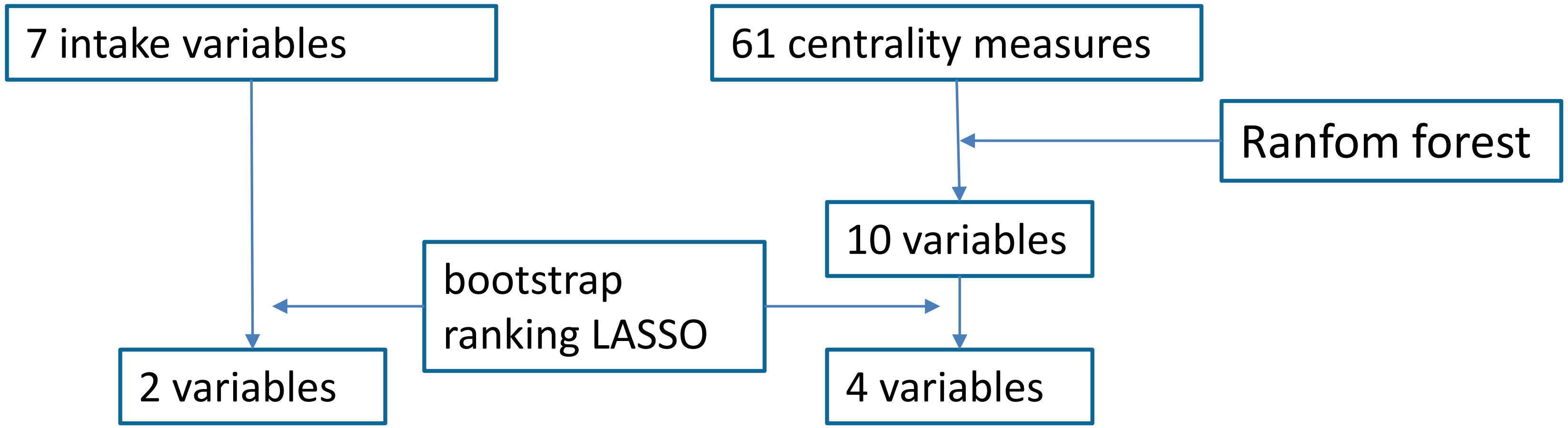
completers



Combining ML and statistical inference

- Network centrality measures: betweenness, closeness, instrength, outstrength, expected force
- 7 intake variables (e.g., impairment, sex, personality style)

Variable selection



Model: Hierarchical generalized linear model

Prediction based on centrality measures

Universität Trier

Hierarchical prediction model for dropout

	GLM			
	Block 1		Block 2	
	estimate	*p*-value	estimate	*p*-value
BSI – global severity index	0.87†	.066	0.62	.324
sex	– 0.79	.195	– 1.27	.101
nervous – betweenness			– 1.00*	.018
excited – expected force			– 0.90*	.035
active – instrength			– 1.02*	.035
social support – outstrength			– 1.00*	.029

Only cross-sectional variables (block 1)

$R^2_{\text{McFadden}} = .06$

$AUC = .64$

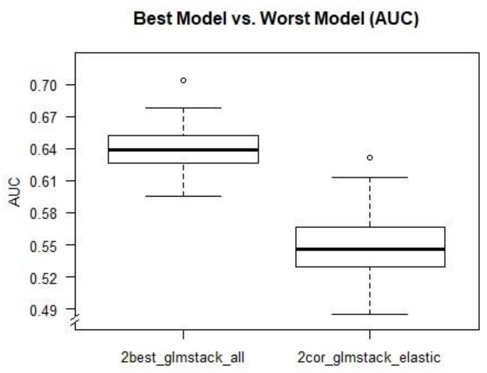
$R^2_{\text{McFadden}} = .32$

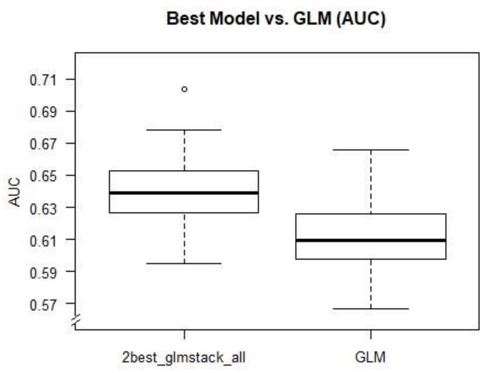
$AUC = .83$

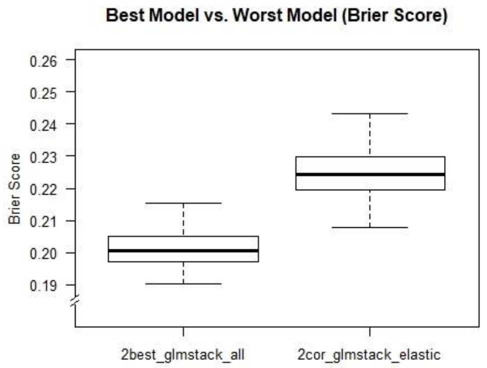
With centrality measures (block 2)

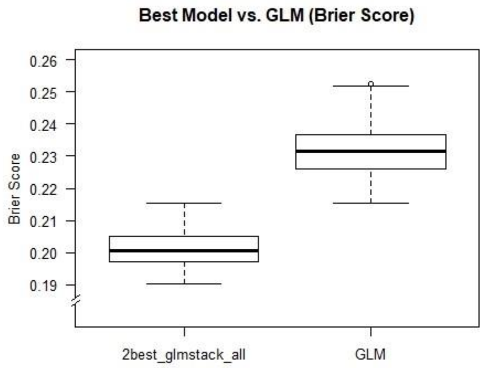
Selection of machine learning algorithms

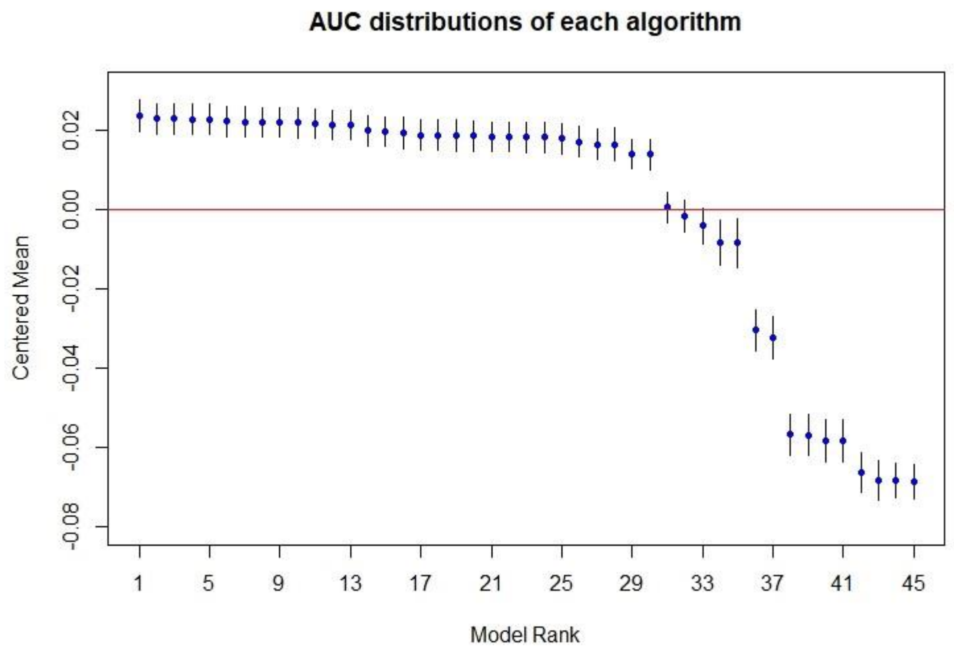
- *N* = 2043 outpatients, 22 algorithms and ensembles
- Best model: Ensemble of elasticnet and glmboost











Best model

$R^2_{McFadden}$ = .060

AUC = .640

Worst model

$R^2_{McFadden}$ = .007

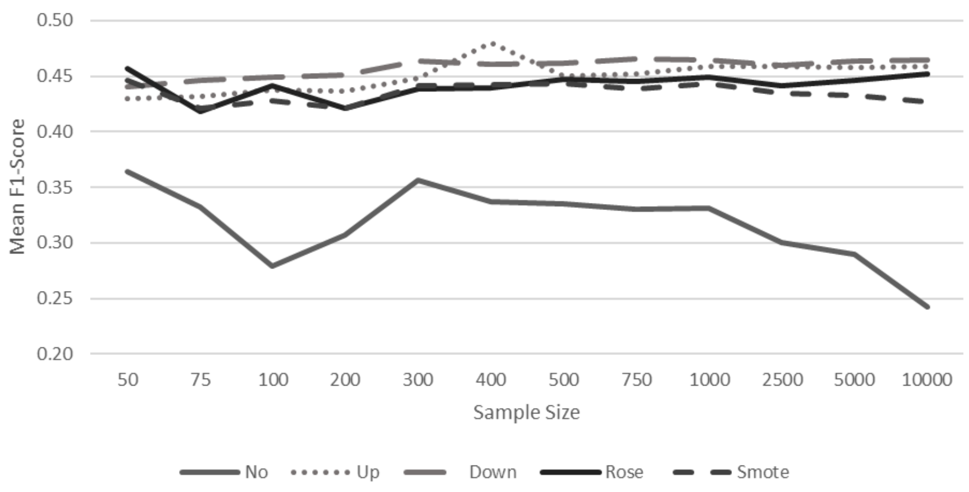
AUC = .548

Bennemann, Schwartz, Giesemann, & Lutz (submitted).

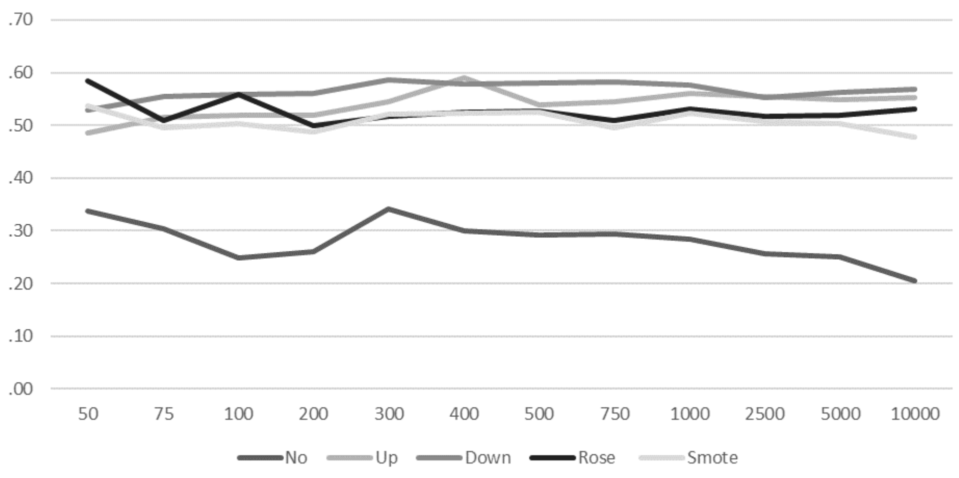
Resampling methods for imbalanced data

- *N* = 49,602 outpatients
- 5 resampling methods: no resampling, up-sampling, down-sampling, SMOTE, and ROSE
- F_1-score: accuracy of binary classifier by weighting precision and recall

Mean F_1-score across algorithms and sample sizes



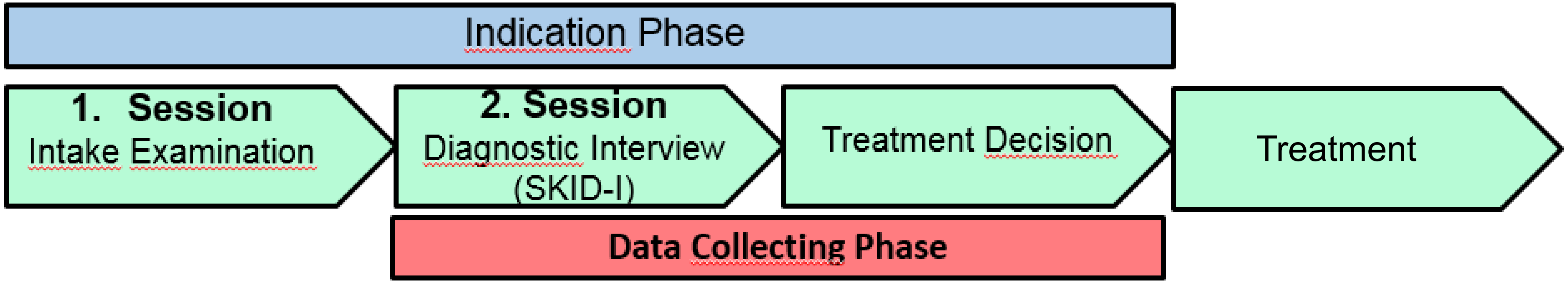
Mean sensitivity across algorithms and sample sizes



Giesemann, Delgadillo, Schwartz, Bennemann, & Lutz (submitted).

Additional data sources

- Assessing passive data from personal digital devices
- Fitness tracker: Garmin vivo smart 4
  - Stress levels (heart rate variability)
  - Sleep quality and duration
  - Activity (steps)
  - Pulse
  - …



Discussion

- Improving dropout prediction to support clinical desicion-making by scientifically trained therapists
- Longitudinal data seem to improve predictions, but implementation is challenging and psychometric training important
- Further investigation of longitudinal networks as well as ML before implementation into TTN/practice (larger data, crossvalidation, prospective evaluation)
- Unclear if centrality measures (as a summary of a network) can be meaningful predictors
  Bringmann et al. (2019). J Abnorm Psychol, 128(8), 892–903
- Limitations: Early implementation, new territory, methodological heterogenity

Thank you!

Follow this QR code to the network analysis paper

Correspondance: Brian Schwartz, M.Sc.
E-mail: schwartzb@uni-trier.de
Twitter: @Schwartz_PsyRes