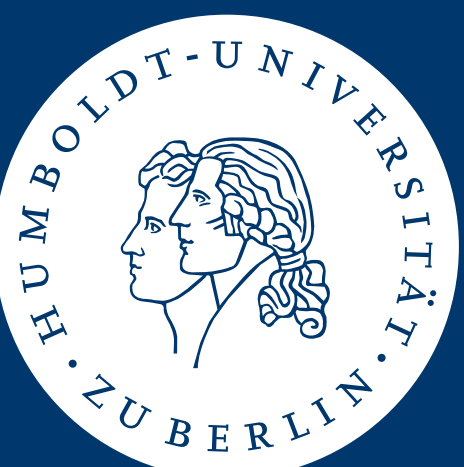# Predicting treatment outcome based on baseline resting-state fMRI functional connectivity? A realistic view

Charlotte Meinke[1], Sarah Wellan[2], Stephan Heinzel[3], Kati Roesmann[4], Joscha Boehnlein[5], Bettina Gathmann[6], Martin J. Herrmann[7], Markus Junghoefer[5,8], Hanna Schwarzmeier[7], Fabian R. Seeger7, Niklas Siminski[7], Thomas Straube[5,6], Udo Dannlowski[5], Elisabeth J. Leehr[5], Ulrike Lueken[1,9], Henrik Walter[2], Kevin Hilbert[10]

[1] Humboldt-Universität zu Berlin, Department of Psychology
[2] Charité Universitätsmedizin Berlin, Klinik für Psychiatrie und Psychotherapie CCM
[3] TU Dortmund University, Institute of Psychology
[4] University of Osnabrück, Institute of Psychology
[5] University of Münster, Institute for Translational Psychiatry
[6] University of Münster, Institute of Medical Psychology and Systems Neuroscience
[7] University Hospital of Würzburg, Center for Mental Health, Department of Psychiatry, Psychosomatics, and Psychotherapy
[8] University of Münster, Institute for Biomagnetism and Biosignalanalysis
[9] German Center for Mental Health (DZPG), partner site Berlin-Potsdam
[10] Health and Medical University Erfurt, Department of Psychology

FOR 5187
PREACT

BACKGROUND

- **Predicting treatment outcome prior to treatment** initiation could be a crucial step towards a more **personalized treatment selection**.
- **Machine learning** is particularly suitable for this endeavor as it enables the accurate prediction of new data on a single-subject level.
- One data modality that appeared promising **is resting-state fMRI functional connectivity** (rs FC), which describes the covariation of the BOLD (blood-oxygenation level dependent) signal in voxels or brain regions over time.

Aim 1: Giving an **estimate** of the rs FC` predictive performance from prior machine-learning studies and highlighting **current challenges**.
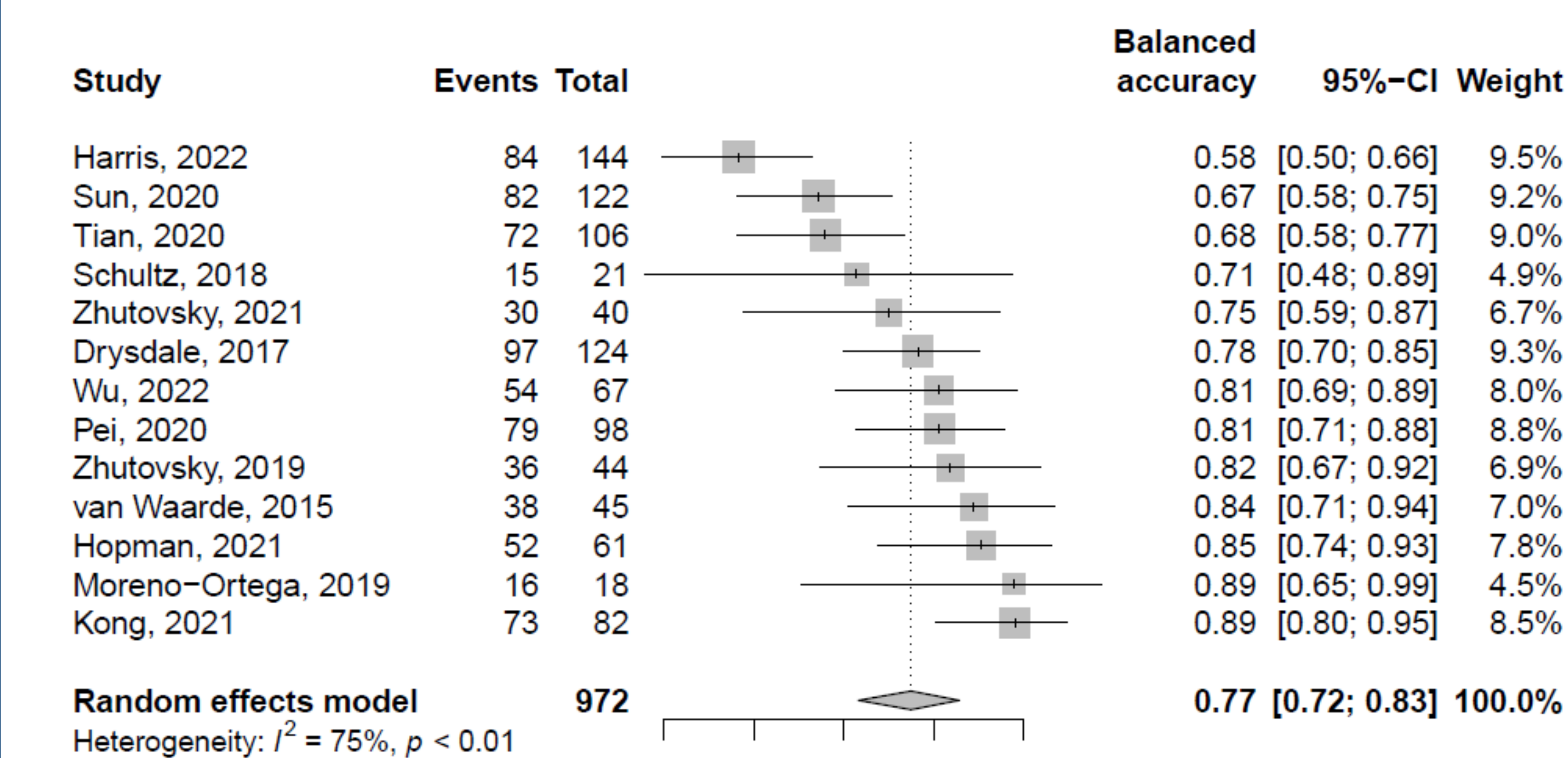
Aim 2: Examining **the predictive performance of rs FC in original data** when tackling some of the problems identified under Aim 1.
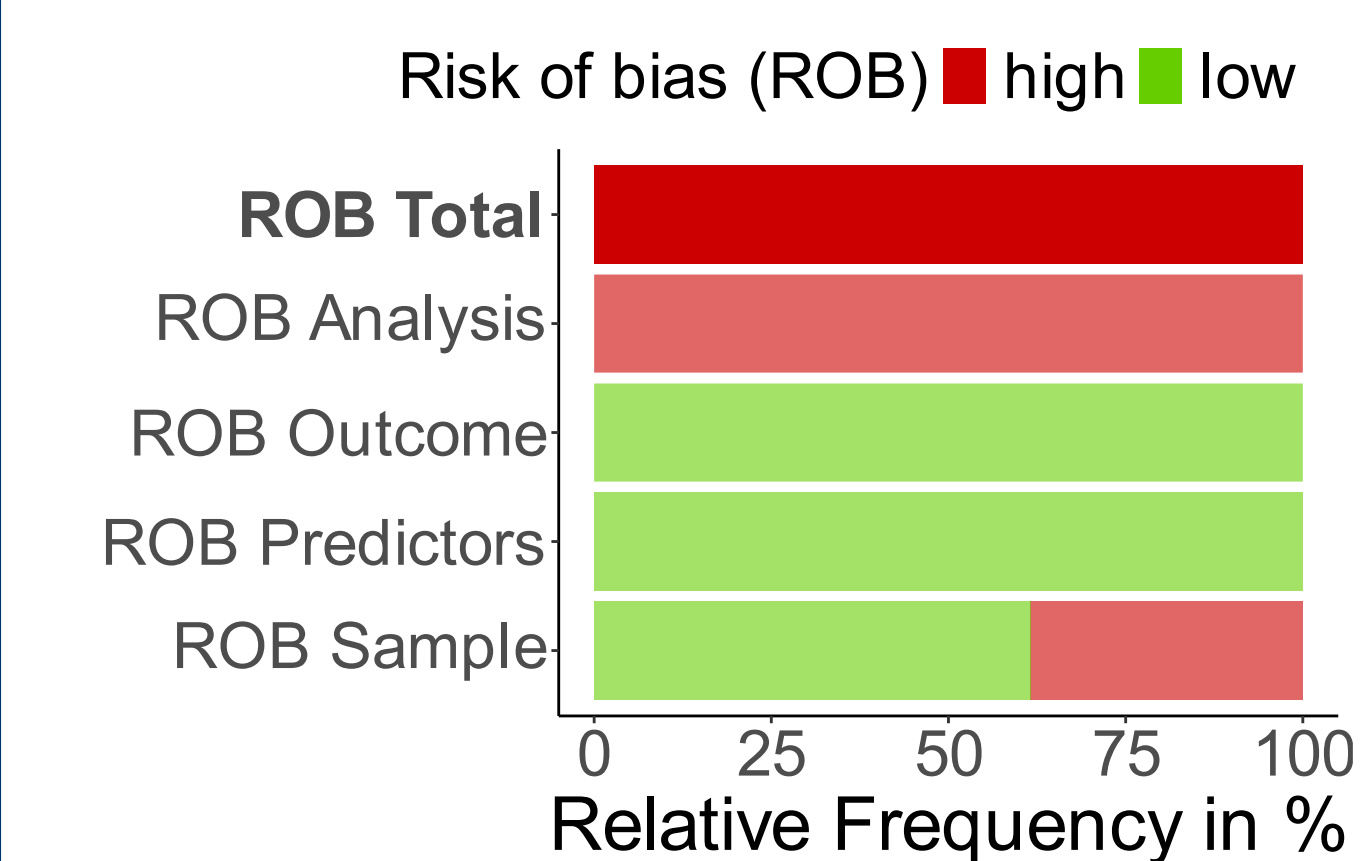
METHODS & RESULTS

1. SYSTEMATIC REVIEW & META-ANALYSIS

- Search in Scopus, PubMed and PsycINFO (12th of December 2022)
  Result: 13 articles (n = 972 patients)
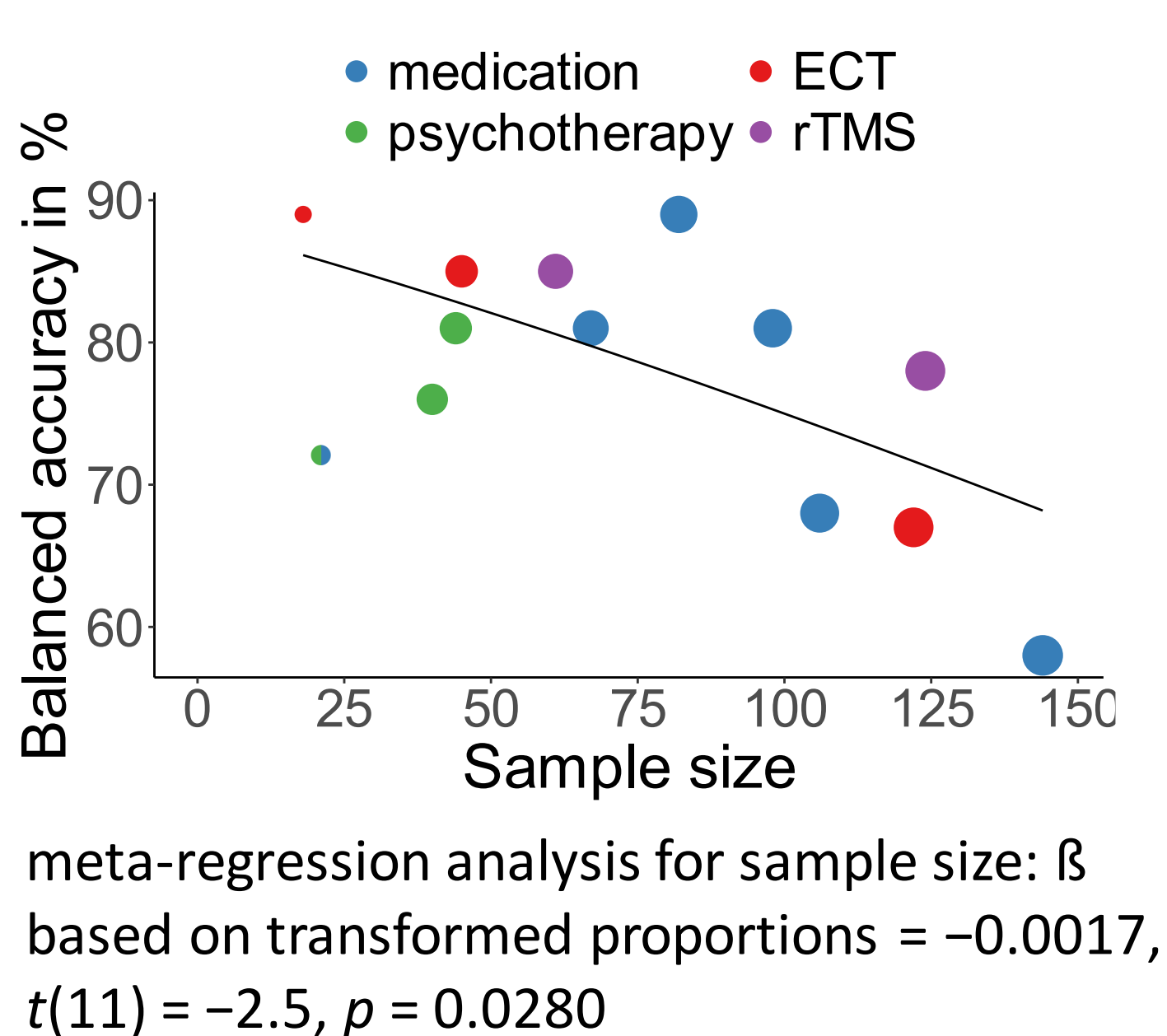- Random effect meta-analysis for proportions

The mean estimated balanced accuracy was 77%, suggesting a good predictive ability of rs FC.

Study	Events	Total	Balanced accuracy	95%−CI	Weight
Harris, 2022	84	144	0.58	[0.50; 0.66]	9.5%
Sun, 2020	82	122	0.67	[0.58; 0.75]	9.2%
Tian, 2020	72	106	0.68	[0.58; 0.77]	9.0%
Schultz, 2018	15	21	0.71	[0.48; 0.89]	4.9%
Zhutovsky, 2021	30	40	0.75	[0.59; 0.87]	6.7%
Drysdale, 2017	97	124	0.78	[0.70; 0.85]	9.3%
Wu, 2022	54	67	0.81	[0.69; 0.89]	8.0%
Pei, 2020	79	98	0.81	[0.71; 0.88]	8.8%
Zhutovsky, 2019	36	44	0.82	[0.67; 0.92]	6.9%
van Waarde, 2015	38	45	0.84	[0.71; 0.94]	7.0%
Hopman, 2021	52	61	0.85	[0.74; 0.93]	7.8%
Moreno−Ortega, 2019	16	18	0.89	[0.65; 0.99]	4.5%
Kong, 2021	73	82	0.89	[0.80; 0.95]	8.5%
Random effects model		**972**	**0.77**	**[0.72; 0.83]**	**100.0%**

Heterogeneity: I^2 = 75%, $p < 0.01$

BUT…

… the risk of bias assessed with PROBAST[1] was high

Risk of bias (ROB) ■ high ■ low
ROB Total / ROB Analysis / ROB Outcome / ROB Predictors / ROB Sample
Relative Frequency in %

… studies with small sample sizes did likely overestimate predictive performance

● medication ● ECT
● psychotherapy ● rTMS
Balanced accuracy in %
Sample size

meta-regression analysis for sample size: ß based on transformed proportions = −0.0017, $t(11)$ = −2.5, p = 0.0280

2. EMPIRICAL ANALYSIS

COMBINED DATASET (larger than any of the studies included in our review)

|	SPIDER PHOBIA[2] (n = 196; 103 Responders)	DEPRESSION[3] (n = 77; 34 Responders)
Treatment	1 session virtual exposure	25 sessions CBT (+ exercise)
Response	↓30% Spider Phobia Questionnaire	BSI-GSI < 0.56 + reliable change

FEATURES EXTRACTED FROM RS FC
42 Graph metrics (AUC across several thresholds):
- Local (for the anterior cingulate cortex)
- Global
- Network-based

LOW-BIAS MACHINE LEARNING PIPELINE
Preregistered:
- 100 * random shuffle split (80/20)
- feature selection with elastic net
- Final classifier: random forest

Exploratory: varying preprocessing of rs FC, graph metrics summary measure, stratification, hyperparameter tuning, classifier

The maximal balanced accuracy reached by pipelines employed (preregistered + exploratory) was 0.52.
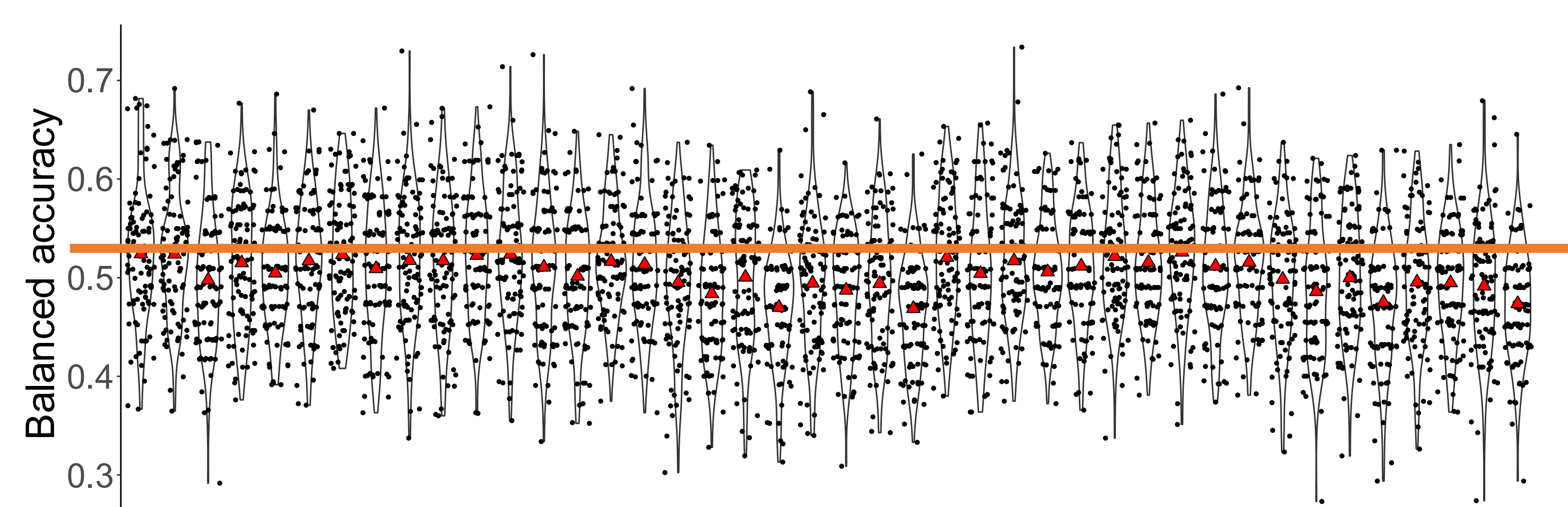

Balanced accuracy

DISCUSSION

➢ In regard of the studies´ high risk of bias, the mean estimated balanced accuracy of 77% should be considered as an **optimistic upper limit of potential predictive performance**.

➢ In our own **empirical analysis**, where we mitigated some of the biasing factors identified (e.g., sample size), predictive performance was **not above chance**.

➢ Other recent state-of-the machine-learning corroborate that the predictive ability of rs FC is still unclear[4,5].

FIND OUR REVIEW!
FIND OUR PREREGISTRATION !
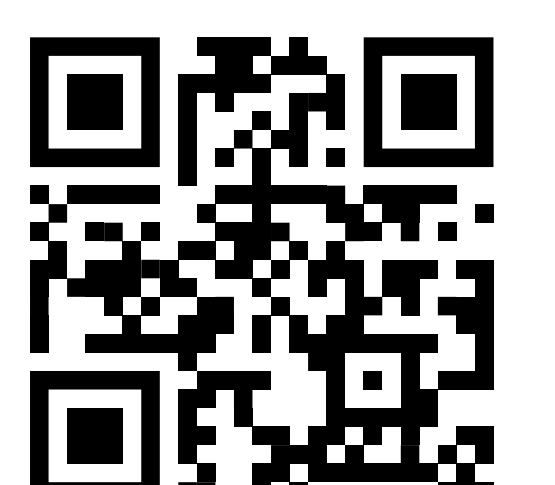
References

[1] Moons, Karel G. M.; Wolff, Robert F.; Riley, Richard D.; Whiting, Penny F.; Westwood, Marie; Collins, Gary S. et al. (2019): PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. In: Annals of Internal Medicine 170 (1), S. 51–58. DOI: 10.7326/M18-1376.
[2] Schwarzmeier, H., Leehr, E. J., Böhnlein, J., Seeger, F. R., Roesmann, K., Gathmann, B., Herrmann, M. J., Siminski, N., Junghöfer, M., Straube, T., Grotegerd, D., & Dannlowski, U. (2020). Theranostic markers for personalized therapy of spider phobia: Methods of a bicentric external cross-validation machine learning approach. International Journal of Methods in Psychiatric Research, 29(2), e1812. https://doi.org/10.1002/mpr.1812
[3] Heinzel, S., Rapp, M. A., Fydrich, T., Ströhle, A., Terán, C., Kallies, G., Schwefel, M. & Heissel, A. (2018). Neurobiological mechanisms of exercise and psychotherapy in depression: The SPeED study— Rationale, design, and methodological issues. Clinical Trials (London, England), 15(1), 53–64. https://doi.org/10.1177/1740774517729161
[4] Harris, J.K., Hassel, S., Davis, A.D., Zamyadi, M., Arnott, S.R., Milev, R., et al., 2022. Predicting escitalopram treatment response from pre-treatment and early response resting state fMRI in a multi-site sample: a CAN-BIND-1 report. Neuroimage Clin. 35, 103120.
[5] Hilbert, K., Böhnlein, J., Meinke, C., Chavanne, A. V., Langhammer, T., Stumpe, L., Winter, N., Leenings, R., Adolph, D., Arolt, V., Bischoff, S., Cwik, J. C., Deckert, J., Domschke, K., Fydrich, T., Gathmann, B., Hamm, A. O., Heinig, I., Herrmann, M. J., . . . Lueken, U. (2024). Lack of evidence for predictive utility from resting state fMRI data for individual exposure-based cognitive behavioral therapy: A machine learning study in two large multi-site samples in anxiety disorders. NeuroImage, 295, 120639. https://doi.org/10.1016/j.neuroimage.2024.120639