Estimating heterogeneity of treatment effects in clinical trials

Joshua S. Siegel[1], Jinglin Zhong[1], Sasagu Tomioka[1], Ajay Ogirala[1], Stephen V. Faraone[2], Steven T. Szabo[1], Kenneth S. Koblan[1], Seth C. Hopkins[1]

[1]Sumitomo Pharma America Inc., [2]SUNY Upstate Medical University, Syracuse, NY, USA

KEY FINDINGS

- We present a measure of estimated heterogeneity of treatment effect (**eHTE**) in clinical trials
- Testing this approach with large published placebo-controlled RCTs revealed reproducible instances of heterogeneity.
- eHTE provides useful info to clinicians, patients, and drug developers (can facilitate enrichment)
- Results suggest 'variability ratio' does not sufficiently capture treatment heterogeneity

Supported by funding from Sumitomo Pharma America, Inc. Marlborough, MA 01752

The International Society for CNS Clinical Trials and Methodology, 20th Annual Scientific Meeting; Washington D.C., February 21-23, 2024

INTRODUCTION

- Outcomes from large clinical trials are reported with group means and effect sizes.
- This hides heterogeneity of treatment effects (HTE) and obfuscates attempts to determine if subgroups are present.
- For example, the population effect of SSRI is ~2 points on the HAMD-17. This is below the clinically relevant individual effect (3-7 points). But it is often assumed that subgroups exist who receive larger benefit. However, heterogeneity of treatment effect **has not conclusively been shown to exist**.
- Volkmann 2020 & Maslej 2021 questions the existence of heterogeneity of antidepressant response based on meta-analysis showing 'variability ratio' close to 1.

DATASET

- 11 psychiatric clinical trial datasets with 23 active treatment arms (Table 1) for which participant-level outcomes were available
- Multiple treatments showed significant heterogeneity of treatment effect ($P_{eHTE} < 0.05$ uncorrected).

Dataset	NCTID	Outcome	Arms (N)	eHTE	P_{eHTE}
Dasotraline in adults with binge-eating disorder (McElroy et al., 2020)	NCT02564588	Binge Days Per Week	Dasotraline 4-8mg (159) Placebo (160)	0.25	0.002
Dasotraline in adults with binge-eating disorder: (Grilo et al., 2021)	NCT03107026	Binge Days Per Week	Dasotraline 4mg (173) Dasotraline 6mg (150) Placebo (163)	0.22 0.26	0.011 0.003
Dasotraline in children with attention deficit disorder (Findling et al., 2019)	NCT02428088	ADHD RS-IV	Dasotraline 2mg (111) Dasotraline 4mg (115) Placebo (116)	0.14 0.16	0.49 0.32
Dasotraline for the Treatment of ADHD in Adults (Koblan et al., 2015)	NCT01692782	ADHD RS-IV	Dasotraline 4mg (116) Dasotraline 8mg (115) Placebo (110)	0.14 0.19	0.48 0.17
Adler et al 2021 Dasotraline in adults with ADHD (Adler et al., 2021)	NCT02276209	ADHD RS-IV	Dasotraline 4mg (219) Dasotraline 6mg (210) Placebo (219)	0.10 0.10	0.52 0.59
Lurasidone or Olanzapine for Schizophrenia (, Study D1050231) (Meltzer et al., 2011)	NCT00615433	PANSS	Luras.40mg (79) Luras.120mg (68) Olanzapine 15mg (87) Placebo (73)	0.14 0.23 0.21	0.79 0.18 0.20
Lurasidone for acute schizophrenia: a 6-week RCT (Nasrallah et al., 2013)	NCT00549718	PANSS	Lurasidone 40mg (84) Lurasidone 80mg (88) Lurasidone 120mg (86) Placebo (75)	0.12 0.21 0.22	0.83 0.19 0.15
Lurasidone in the treatment of schizophrenia (Loebel et al., 2013)	NCT00790192	PANSS	Lurasidone 80mg (89) Lurasidone 160mg (96) Quetiap. XR 600mg (99) Placebo (77)	0.15 0.16 0.19	0.50 0.37 0.19
Psilocybin vs Escital. for Depression (Carhart-Harris et al., 2021)	NCT03429075	MADRS	Psilocybin (28) Escitalopram (29)	0.36	0.153
Psilocybin for Treatment of Major Depressive Disorder (Raison et al., 2023)	NCT03866174	HAM-D	Psilocybin (50) Niacin (44)	0.33	0.068
Dasotraline or venlafaxine for 8 weeks in adults with MDD (Hopkins et al., 2013)	NCT0058497	HAMD-17	Dasotraline 0.5mg (101) Dasotraline 2mg (110) Venlafaxine 150mg (107) Placebo (114)	0.21 0.16 0.25	0.045 0.24 0.034

METHODS

The distribution of treatment responses in each arm is represented by sorting patients in each arm ordinally by response and then plotting response across percentiles. Let $P(x)$ represent the cumulative response function for the placebo group and $T(x)$ represent the same for the treatment group, over some range of percentiles $x \in [x1, x2, \ldots, x99]$. For each x_i, we calculate the difference between the treatment and placebo cumulative responses:

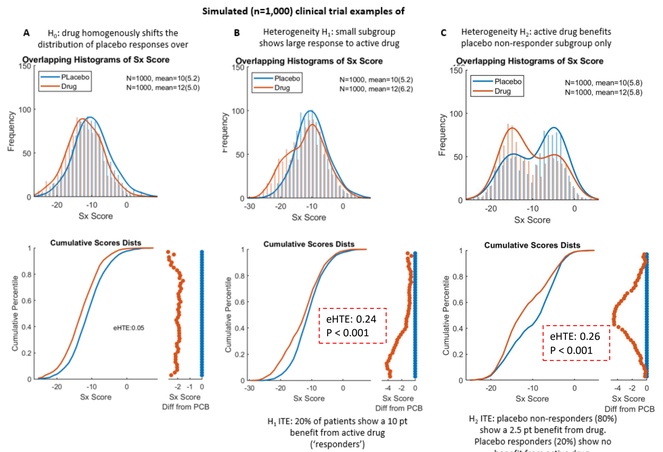
$$D(x_i) = T(x_i) - P(x_i) \qquad \text{Eq. 1}$$

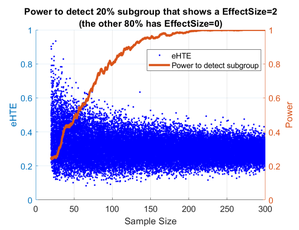
Then, eHTE is calculated as the standard deviation of the differences across percentiles D(xi) divided by the standard deviation of placebo response:

$$eHTE = {SD(D(xi))}/{SD(placebo)} \qquad \text{Eq. 2}$$

This ratio represents the relative variability in responses between the two arms, normalized by the variability observed within the placebo arm. Importantly, eHTE scales with the standard deviation in the ITE (thus it scales with HTE) and it is unitless (thus it can be compared across different clinical/outcome scales). A high value of eHTE indicates a substantial heterogeneity in responses between the treatment arms (relative to overall variability in placebo response).

TOY CASES & POWER

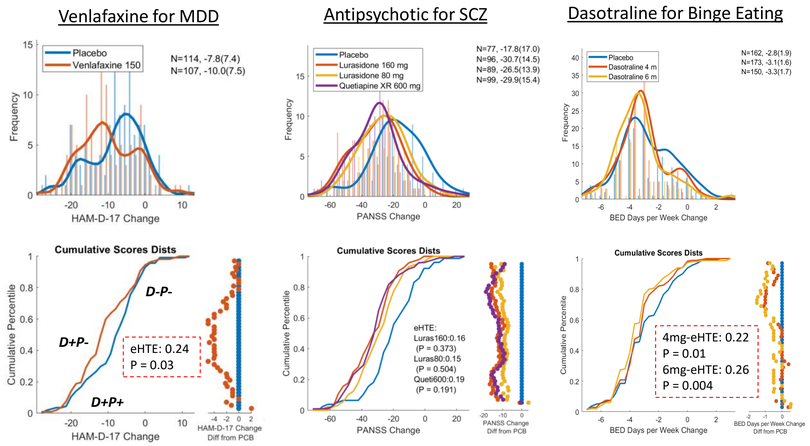




Power to detect subgroups.

Based on simulation B – Blue dots represent eHTE values across a range of sample sizes. Orange line depicts power to detect a significant heterogeneity with alpha (two-sided) set to p<0.05.

EVIDENCE IN REAL RCTs



- eHTE provides evidence for heterogeneity in MDD response to SNRI (despite variability ratio close to 1!)
- Results reflects response subgroups previously hypothesized to exist: D+P+, D+P-, D-P- (Fava 2015)

Phase III study of Lurasidone 40/80/120mg did not find heterogeneity (P = 0.83, 0.19, 0.15)

Phase II study of Dasotraline for BED: eHTE = 0.25 (P = 0.003)

- Heterogeneity replicates across multiple trials of the same/indication (e.g., Antipsychotics for Schizophrenia, Dasotraline for binge eating disorder).

CONCLUSIONS & FUTURE DIRECTIONS

- Multiple clinical trial datasets demonstrated that heterogeneity may exist (despite variability ratio close to 1). This contrasts to the variance ratio-based method that does not find antidepressant heterogeneity (Maslej et al., 2021; Volkmann 2020)
- eHTE provides valuable info to clinicians, patients, and drug developers (without compromising efficacy-based endpoints)
- Compare eHTE to baseline covariates. If covariate explains heterogeneity, this can be used for enrichment!
- Test additional datasets (eHTE code **available in Python/Matlab/SAS/R** on request!)

KEY REFERENCES

1. Kravitz, R. L., Duan, N., & Braslow, J. (2004). Evidence-Based Medicine, Heterogeneity of Treatment Effects, and the Trouble with Averages. The Milbank Quarterly, 82(4), 661–687.
2. Senn, S. (2018). Statistical pitfalls of personalized medicine. Nature, 563(7733), 619–621.
3. Fava, M. (2015). Implications of a Biosignature Study of the Placebo Response in Major Depressive Disorder. JAMA Psychiatry, 72(11), 1073–1074.
4. Maslej, M. M., et al., (2021). Individual Differences in Response to Antidepressants: A Meta-analysis of Placebo-Controlled Randomized Clinical Trials. *JAMA Psychiatry*, *78*(5), 490–497.