Are we able to conclude on antidepressant efficacy for depression with the current methodology of RCT?

**BACKGROUND**

- Antidepressant’s efficacy in depression is debated because their trials are criticized.
- Several pitfalls may limit the statistical power of antidepressant RCTs and their ability to conclude:
  - Sample heterogeneity due to poor reproducibility of DSM diagnosis
  - Sample heterogeneity in disease severity
  - Use of controversial outcomes
  - We found no study analysing their simultaneous impact on the power of RCTs.

**OBJECTIVES**

- To study the impact of several design’s element on the statistical power of RCT on antidepressant in depression

**METHODS**

- Modelisation of one subject’s response to the medication for each item of the Hamilton Depression Rating Scale (HDRS-one of the most used main outcome for trial on antidepressant in depression)

- The model takes into account:
  - Pre-treatment scores
  - Item quotation reproducibility
  - Response to placebo and natural variation of scores
  - Within-subject correlation
  - Between-subject variability

- Calibration of the model was done using literature data

- Simulation of 960 scenarios:
  - 7 different types of drug
  - 4 various case-mix of differential diagnoses in the sample:
    - 5 sample sizes
    - 3 types of outcome measures: response (HDRS<8), remission (HDRS score decrease by at least 50%) and continuous (t-test).
    - 15 different HDRS minimal scores for patient’s inclusion

**RESULTS**

**Sensitivity of statistical power to sample heterogeneity**

We simulated 4 samples:
1. In black: Reference population is composed by: 75% depression, 8% bipolar disorder, and a mix of adaptive disorder (AD), substance abuse disorder (SAD), post traumatic stress disorder (PTSD), generalized anxiety disorder (GAD) and schizophrenia.
2. In yellow: 100% depression
3. In blue: 55% depression, 25% bipolar disorder and a mix of GAD, SAD, AD, PTSD and schizophrenia
4. In red: 55% depression, 20% adaptive disorder and a mix of bipolar disorder, GAD, SAD, PTSD and schizophrenia

**Response** outcome was highly dependent to sample heterogeneity whereas **remission** outcome was not

**Sensitivity of statistical power to the type of drug**

- Sample sizes under 650 subjects yielded a power under 90% whatever the scenario
- HDRS score at inclusion has marginal impact on the statistical power
- Trials on drugs that only have an elective effect on few items (sleep, sedative, mood and stimulating medication) had lower power

**DISCUSSION**

- Sample size of most studies is too low due to methodological limitations such as heterogeneity and outcomes
- Lack of power leads to studies that are unable to provide adequate evidence
- Improvements in trial design of antidepressant medications should be made in order to limit waste in research

**Disclosure:** The authors report no conflicts of interest for this work