A practical guide to adding patient heterogeneity into Phase III trials: Results from IMI GetReal WP2

Karcher H1, Fu S2, Efthimiou O3, Schneeweiss S1, Nordon C4, Abenham L5

1LASER Analytica, London, United Kingdom, 2LASER Analytica, Lorraine, France, 3LASER Research, Paris, France, 4University of Ioannina, Ioannina, Greece, 5Harvard Medical School, Boston, MA, USA

1. CONTEXT & OBJECTIVES

- Phase III trials typically exclude patients with certain baseline characteristics, such as older age or co-morbidities, and thereby hamper learning of new drugs’ effectiveness in real-life.
- A modification to both design and analyses is proposed to address this issue.
- Schizophrenia is taken as case example.

2. DATA SOURCE

SOHO - a prospective, observational study on 10,218 schizophrenia patients
- from 10 European countries
- followed over 3 years
- who received antipsychotic treatment

3. OUTCOME

We used mean CGI-S at 3 months (change from baseline) as outcome. This outcome was evaluated in patients taking the most frequently used drug (blinded). CGI-S score (Clinical Global Impression-Severity):
- Assesses severity of patient’s mental illness at time of rating with one question
- 7-point scale: from 1 (not at all ill) to 7 (extremely ill)
- In SOHO cohort study, most patients have CGI-S values of 4 or 5

4. METHODS

1. A “base synthetic RCT” was created by applying Phase III exclusion criteria.
2. A series of “enriched synthetic RCTs” were defined by replacing patients with SOHO patients that were initially excluded due to various factors.
3. The real-life drug effect was predicted from schizophrenia using regression models and was compared with SOHO.

5. RESULTS

Exclusion (enrichment) criteria:
- age > 65 years, duration of illness < 3 years, patients with few previous suicide attempts, patients with history of alcohol or substance abuse, and patients treated at private practices

Exclusion (enrichment) in the real-world sample

Figure 1: Real-world patients typically excluded from Phase III by type of exclusion criterion

Final regression model, used for real-life predictions:

¬∆CGI-S at 3 months (age + chronicity + gender + BMI + duration of hospitalization + number of admissions in hospital + depression score + QOL score + patient compliance + country + work status + housing condition + social activity + relationship + negative symptom at baseline + positive symptom at baseline + cognitive symptom at baseline + dosage DDD dose) if (initiated the drug at baseline)

Evaluation of prediction accuracy:
MSE = mean[(predicted CGI-S - from RCT data - real-life observed CGI-S)]^2
Coverage of 95% confidence interval = % of times that the 95% predicted interval contains the true real-life value, which should ideally be close to 95%

Figure 2: Distribution of number of suicide attempts in synthetic base RCT, SOHO and two RCTs enriched with patients who had 1-5 suicide attempts

Figure 3: Predicted error and CI using different RCTs enriched with few “suicide attempts”

Table 2: Which factors to enrich Phase III? A comparison between the terms of the benefit their addition in terms real-life prediction accuracy, and associated trial sample size

6. Conclusion

- A simulation study was performed to guide addition of patient heterogeneity to standard Phase III trials in schizophrenia. The impact of the following changes was assessed:
  - Trial design: add a few patients with selected factors through stratification, without change in sample size (“enriched RCTs”)
  - Trial analysis: use predictive modeling to estimate real-life effectiveness
- The best choice of enrichment factor to predict real-life effects was found to be driven by:
  - Size of the excluded population in real-life. Excluding “number of past suicide attempts > 1” left out the greatest schizophrenia population from Phase 3 trials.
  - Change in outcome in patients with this factor. Patients with a practice type “private” and disease chronicity < 5 years had the most different outcome from typical Phase 3 patients.
- Enriching typical Phase 3 with selected factors improved the representability of real-life and as a result, improved predictions of the real-life effects of the investigated drug.