

Title: The added value of the CGI-I scale in assessing global severity: a cost/benefit analysis using data from four Phase III MDD trials

Authors: Nations KR¹, Gandy-Don Sing Z¹, Spiridonescu L¹, Reinhold C¹, Brewer C², Baker RA²

¹INC Research, ²Otsuka Pharmaceutical Development and Commercialization

The Methodological Question Being Addressed: Does the Clinical Global Impression- Improvement scale(CGI-I) convey added benefit over the CGI-Severity scale (CGI-S) in clinical trials, considering relative psychometric performance as well as resource costs associated with scale administration and data cleaning?

Introduction: Considered a gold standard of global disease evaluation, the CGI is ubiquitously selected as a key secondary measure in CNS trials. Both the CGI-S and CGI-I are well-validated, highly sensitive to change, and frequently deployed in the same trial. However, the literature is mixed as to whether both scales are warranted, and trial designers may fail to appreciate the significant time and cost associated with administering, monitoring, and data cleaning a measure even as simple as the CGI. In MDD trials, response is typically defined as a CGI-I score of 1 or 2 (much or very much improved), while a CGI-S rating of 1 or 2 is used to define remission (normal or borderline ill). We undertook to understand whether the CGI-S could reasonably address response, in addition to global severity and remission, obviating any need for the CGI-I.

Methods: The current analysis examined the relative predictive and discriminative properties of the CGI-S and CGI-I, as well as cost/resource burden, using data from four Phase III adjunctive brexpiprazole trials in major depressive disorder (MDD; POLARIS-NCT01360632, PYXIS-NCT01360645, DELPHINUS-NCT01727726, SIRIUS-NCT02196506). Subjects who responded to antidepressant (ADT) monotherapy in the lead-in period continued ADT for an additional 6 weeks. For the current analysis, these subjects were pooled across the four trials to examine the CGI-S vs CGI-I on predictive/discriminative validity (likelihood ratios, diagnostic odds ratio, receiver operating characteristics), reliability (internal consistency), and resource demands (time and cost).

Results: A total of 2374 subjects completed 14 weeks of ADT, 70% of whom were MADRS responders at Week14 (defined as decrease from baseline (CFB) of $\geq 50\%$). The relationship between MADRS and CGI-S was significantly stronger than MADRS and CGI-I (CGI-S $R^2 = 0.82$; CGI-I $R^2 = 0.73$; $p < 0.001$). Of four Week14 CGI-S response definitions [CGI-S CFB ≥ 1 ; CGI-S CFB ≥ 2 ; CGI-S score ≤ 3 (mildly ill or better); CGI-S score ≤ 2 (borderline ill or better)], CGI-S absolute score ≤ 3 was a more accurate measure of response ($\chi^2 = 368.68$, $p < 0.001$, Diagnostic Odds Ratio (DOR) = 56.86) than any other CGI-S definition, and was more accurate than the commonly used CGI-I score of 1 or 2 ($\chi^2 = 219.23$, $p < 0.001$, DOR: 32.95). The nearly 500 data and medical queries for CGI-I in this program were 21% higher than for CGI-S, and translate to \$11,500 in medical/data personnel time for cleaning, not counting more significant costs associated with scale production, investigator administration, data entry, and database/edit check programming. The majority of errors were related to the lookback period (rater incorrectly comparing severity to last visit rather than Baseline).

Conclusions: The results from the current analysis suggest that the CGI-I scale does not necessarily offer added value beyond what can be measured using the CGI-S alone. The CGI-S's superior predictive/discriminative accuracy, taken together with lower resource time/cost associated with data cleaning, suggest that CGI-S as a sole measure of global severity and response is a viable, valid, and cost effective option in trial design for MDD, with potential for broader applicability in other therapeutic areas.

Disclosures: Funding for the POLARIS, PYXIS DEPHINUS and SIRIUS trials was provided by Otsuka Pharmaceutical Development and Commercialization (OPDC) and H. Lundbeck A/S. RB and CB are employees of OPDC. INC Research was the Contract Research Organization responsible for execution of all four trials. KN, ZG, LS and CR are employees of INC Research.