

Clinical Global Impression–Stimulation Outcome Monitoring Tool (CGI-STIM): Development and Preliminary Evaluation of a Pragmatic Framework for Outcome Monitoring in Electroconvulsive Therapy

Sarah Kayser

sarah.kayser@med.uni-tuebingen.de

University of Tübingen

Christian Plewnia

University of Tübingen

Andreas Fallgatter

University of Tübingen

Research Article

Keywords: Electroconvulsive therapy, cognitive tolerability, outcome monitoring, Clinical Global Impression, measurement-based care

Posted Date: July 24th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10308211/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at European Archives of Psychiatry and Clinical Neuroscience on September 26th, 2026. See the published version at <https://doi.org/10.1007/s00406-026-02387-8>.

Abstract

Objectives

Electroconvulsive therapy (ECT) is highly effective for severe and treatment-resistant mental disorders, but concerns about cognitive side effects and enduring stigma continue to hinder its broader acceptance. Although treatment response and cognitive tolerability are both clinically important, their systematic assessment and documentation remain challenging in routine clinical practice. To address this gap, we developed the Clinical Global Impression–Stimulation Outcome Monitoring Tool (CGI-STIM), a clinician-rated framework that integrates both domains into a structured and clinically interpretable outcome assessment.

Methods

We retrospectively analyzed 21 ECT series in 15 patients. Two independent clinicians applied the CGI-STIM using routine clinical documentation to classify treatment outcomes based on clinical response and cognitive tolerability. Feasibility and interrater agreement were evaluated.

Results

The CGI-STIM was successfully applied to all 21 ECT series. Outcome classification was feasible in all cases using routinely available clinical information. Average completion time was approximately 5 minutes. Interrater agreement was excellent (quadratically weighted Cohen's $\kappa = 0.84$). The resulting classifications reflected varying combinations of clinical improvement and cognitive tolerability, illustrating the instrument's ability to capture multidimensional treatment outcomes.

Conclusions

The CGI-STIM is a feasible and pragmatic outcome monitoring tool for ECT that facilitates structured documentation of both treatment response and cognitive tolerability. Preliminary findings support its reliability, feasibility, and applicability in routine clinical practice. Future studies should prospectively evaluate the CGI-STIM in larger ECT samples and examine its validity, sensitivity to change, and utility for routine outcome monitoring.

Introduction

ECT remains among the most effective interventions for severe mental disorders, including treatment-resistant depression (TRD), catatonia, mania, psychosis, and dementia-related behavioral disturbances 1–3. Despite its well-established efficacy and favorable safety profile, ECT remains underutilized, largely because of persistent stigma and concerns regarding cognitive adverse effects 4–6.

Cognitive side effects are among the most frequently discussed limitations of ECT and remain highly relevant for patients, caregivers, and clinicians 6–8. Consequently, current guidelines increasingly emphasize cognitive tolerability as an important treatment outcome alongside clinical effectiveness 7, 9–12. At the same time, contemporary concepts such as measurement-based care and personalized psychiatry emphasize systematic outcome monitoring to support clinical decision-making and treatment optimization in routine practice 13, 14.

The Clinical Global Impressions (CGI) scale 15 is a brief and widely used clinician-rated instrument that captures illness severity (CGI-S) and global improvement (CGI-I). Owing to its brevity, ease of administration, and broad applicability across diagnostic groups, the CGI remains one of the most frequently used clinician-rated outcome measures in psychiatry 16. However, neither the original CGI nor its disorder-specific adaptations explicitly capture treatment-related cognitive tolerability, which represents a particularly important outcome dimension in ECT. Likewise, disorder-specific adaptations of the CGI, including the Clinical Global Impressions Scale for Bipolar Disorder (CGI-BP) 17, the Clinical Global Impressions Scale for Schizophrenia (CGI-SCH) 18, and the Clinical Global Impressions Scale for Neurocognitive Disorders (CN-CGI) 19, do not specifically address cognitive side effects associated with ECT.

This gap highlights the need for pragmatic tools that support routine documentation of both treatment response and cognitive tolerability following ECT. Although both domains are clinically important, their systematic assessment in routine clinical practice may be challenging despite increasing emphasis on structured outcome monitoring and measurement-based care 14.

To address this need, we developed the Clinical Global Impression–Stimulation Outcome Monitoring Tool (CGI-STIM), a clinician-rated framework designed to facilitate structured outcome monitoring after ECT. The CGI-STIM combines independent ratings of clinical response and cognitive tolerability within a predefined outcome matrix, generating clinically interpretable outcome categories suitable for routine care. It is intended to complement, rather than replace, standardized symptom scales and neuropsychological assessments.

The present study describes the development of the CGI-STIM and provides a preliminary evaluation of its feasibility and interrater agreement in a retrospective sample of patients receiving ECT.

Materials and Methods

Study Design and Participants

This retrospective, single-center study was conducted at the Department of Psychiatry III, Rheinessen-Fachklinik Alzey, Germany. All ECT treatment series administered with the Thymatron System IV (Somatics, LLC, Lake Bluff, IL, USA) between February 2020 and March 2021 were eligible for inclusion. Data included demographic characteristics, diagnoses, number of ECT sessions, documented adverse effects, and available clinical outcome measures, including CGI 15 ratings and Global Assessment of

Functioning (GAF) 20 scores. The study was approved by the Ethics Committee of the State Medical Association of Rhineland-Palatinate, Mainz, Germany (approval number: 2021–15792).

Clinical Rationale and Development of the CGI-STIM

The CGI-STIM was developed as a pragmatic clinician-rated outcome monitoring tool for routine clinical practice. Its purpose is to support structured documentation of two clinically central domains after ECT: therapeutic response and cognitive tolerability.

CGI-STIM Outcome Monitoring Tool

The CGI-STIM consists of two independently rated domains: clinical response and cognitive tolerability. Ratings were performed after completion of each ECT series to allow an overall assessment of therapeutic benefit and cognitive tolerability across the treatment course.

Clinical response was rated on a four-point scale:

- 1 = Non-response
- 2 = Partial response
- 3 = Response
- 4 = Remission

Cognitive tolerability was rated on a four-point scale:

- 1 = No cognitive impairment
- 2 = Mild cognitive impairment
- 3 = Moderate cognitive impairment
- 4 = Severe cognitive impairment

The two domain ratings were integrated using a predefined outcome matrix that assigns each combination of clinical response and cognitive tolerability to a clinically interpretable outcome category.

Table 1
Clinical Global Impression – Stimulation Outcome Monitoring Tool (CGI-STIM)

Cognitive impairment / Clinical outcome	No	Mild	Moderate	Severe
Remission	Optimal	Favorable	Mixed	Limited
Response	Favorable	Favorable	Mixed	Limited
Partial response	Acceptable	Acceptable	Mixed	Unfavorable
Non-response	No benefit	No benefit	Unfavorable	Unfavorable

Outcome categories represent global clinical judgments integrating therapeutic benefit and cognitive tolerability. The resulting CGI-STIM outcome category was used as the primary descriptive output of the tool.

Feasibility and Interrater Reliability

Two experienced clinicians independently applied the CGI-STIM to the available clinical documentation. Feasibility was assessed based on the time required for completion and the ability to derive ratings from routinely available clinical documentation without additional formal testing.

Because the CGI-STIM outcome categories represent ordinal classifications, interrater agreement was assessed using quadratically weighted Cohen's kappa. Interpretation of agreement followed commonly used benchmarks for weighted kappa values.

Statistical Analysis

Analyses were primarily descriptive, reflecting the retrospective design and small sample size. Demographic and clinical characteristics were summarized using means, standard deviations, frequencies, and percentages, as appropriate. The distribution of CGI-STIM domain ratings and outcome categories was described across ECT series. Interrater agreement was assessed using weighted Cohen's kappa. No confirmatory psychometric validation was performed.

Results

Sample Characteristics

Fifteen patients (73% female; mean age 58.8 ± 10.7 years) underwent a total of 21 ECT treatment series. Diagnoses included major depressive disorder ($n = 13$), schizoaffective disorder ($n = 4$), bipolar depression ($n = 2$), catatonic schizophrenia ($n = 1$), and psychotic depression ($n = 1$). The mean number of ECT sessions per series was 9.7 ± 3.2 . Mean CGI-S scores decreased from 5.4 to 3.8 following

treatment, while GAF scores improved from 48.9 to 65.0. The mean CGI-I score was 2.25. Detailed demographic and clinical characteristics are presented in Table 2.

Table 2
Demographic and clinical characteristics of the sample

Variable	Mean	SD
Age (years)	58.81	10.67
Female sex (%)	73.3	–
Number of ECT sessions	9.74	3.16
CGI-I	2.25	1.89
CGI (pre-treatment)	5.39	0.50
CGI (post-treatment)	3.75	0.96
GAF (pre-treatment)	48.85	11.32
GAF (post-treatment)	65.00	5.00

Demographic and clinical characteristics of the sample (n = 15). Values are means and standard deviations (SD), unless otherwise indicated. Pre- and post-treatment scores refer to clinical assessments conducted before and after the ECT series, respectively. ECT, electroconvulsive therapy; CGI-I, Clinical Global Impressions–Improvement scale; CGI, Clinical Global Impressions–Severity scale; GAF, Global Assessment of Functioning.

CGI-STIM Outcome Categories

The CGI-STIM was successfully applied to all 21 ECT series. Outcome categories were assigned independently by two independent clinicians using the predefined CGI-STIM outcome matrix (see Table 3).

For Rater 1, outcome categories were distributed as follows: Favorable (n = 6, 28.6%), Acceptable (n = 8, 38.1%), Mixed (n = 3, 14.3%), No Benefit (n = 3, 14.3%), and Unfavorable (n = 3, 14.3%).

For Rater 2, outcome categories were distributed as follows: Favorable (n = 5, 23.8%), Acceptable (n = 5, 23.8%), Mixed (n = 5, 23.8%), No Benefit (n = 2, 9.5%), and Unfavorable (n = 6, 28.6%).

No treatment series were assigned to the categories Optimal or Limited by either rater.

Table 3
Distribution of CGI-STIM outcome categories across 21 ECT treatment series.

CGI-STIM Outcome Category	Rater 1 n (%)	Rater 2 n (%)
Optimal	0 (0%)	0 (0%)
Favorable	6 (28.6%)	5 (23.8%)
Acceptable	8 (38.1%)	5 (23.8%)
Mixed	3 (14.3%)	5 (23.8%)
Limited	0 (0%)	0 (0%)
No Benefit	3 (14.3%)	2 (9.5%)
Unfavorable	3 (14.3%)	6 (28.6%)

Outcome categories were assigned independently by two raters using the CGI-STIM (Clinical Global Impression–Stimulation Outcome Monitoring Tool) matrix, which integrates treatment response and cognitive tolerability into a single clinically interpretable outcome classification. Favorable outcomes indicate meaningful clinical improvement with minimal cognitive burden; Acceptable outcomes reflect partial or clinically relevant benefit with limited cognitive impairment; Mixed outcomes indicate a balance of therapeutic benefit and cognitive side effects; No Benefit indicates insufficient clinical improvement without relevant cognitive burden; and Unfavorable outcomes indicate limited therapeutic benefit accompanied by clinically relevant cognitive impairment. Values are presented as number (%) of ECT series classified within each category.

Feasibility and Interrater Reliability

The CGI-STIM could be completed using routine clinical documentation, including physician notes, progress reports, and patient or caregiver reports. Average completion time was approximately 5 minutes. Outcome classification was feasible for all 21 ECT series.

Interrater agreement for the CGI-STIM outcome categories was excellent. Quadratically weighted Cohen’s kappa was 0.84, indicating a high degree of consistency between independent clinicians despite minor differences in category assignments.

Discussion

Clinical Relevance of the CGI-STIM

The present study provides preliminary evidence that the CGI-STIM is a feasible and clinically meaningful framework for structured outcome monitoring following ECT. By integrating treatment response and cognitive tolerability into clinically interpretable outcome categories, the CGI-STIM addresses two

domains that are central to ECT evaluation 6, 7, but are often documented inconsistently in routine clinical practice.

Rather than replacing detailed symptom assessments or neuropsychological testing, the CGI-STIM is intended to complement routine clinical documentation. Its matrix-based structure allows clinically relevant response–tolerability constellations to be summarized within a single outcome framework, facilitating communication of treatment outcomes among clinicians, patients, and caregivers. Importantly, the matrix-based classification preserves clinically distinct combinations of treatment response and cognitive tolerability that might otherwise be obscured by a single summary score. By assigning clinically meaningful outcome categories to specific response–tolerability constellations, the CGI-STIM retains information that may be relevant for clinical interpretation and shared decision-making.

Feasibility and Reliability

The primary objective of this study was not to establish psychometric validity, but to determine whether the CGI-STIM could be applied consistently using routine clinical documentation. The instrument could be completed using routinely available information and required approximately five minutes per ECT series, supporting its feasibility for clinical use.

Interrater agreement was excellent (quadratically weighted Cohen's $\kappa = 0.84$), indicating a high degree of consistency between independent raters despite minor differences in category assignments. This finding is particularly relevant because ratings were derived from routine clinical documentation rather than standardized research assessments. The observed agreement suggests that the CGI-STIM can be applied reliably in naturalistic clinical settings and supports further prospective evaluation.

Importantly, the CGI-STIM was specifically developed for environments in which comprehensive neuropsychological testing is not routinely available. In such settings, structured documentation of both treatment response and cognitive tolerability may improve longitudinal outcome monitoring and support clinical decision-making 9, 12, 13. Given the frequent concerns regarding cognitive side effects of ECT 6, 7, a framework that systematically incorporates cognitive tolerability alongside therapeutic benefit may also facilitate shared decision-making and patient-centered care.

Alignment with Current Clinical Practice

Current ECT guidelines recommend the systematic assessment of both treatment effectiveness and cognitive tolerability throughout the course of treatment 9-12. At the same time, contemporary concepts such as measurement-based care emphasize the routine use of structured outcome measures to support clinical decision-making 13, 14. However, the implementation of comprehensive symptom ratings and standardized cognitive assessments remains challenging in routine clinical practice because of limited time, personnel, and resource availability. Consequently, outcome monitoring often falls short of guideline recommendations, and cognitive outcomes are frequently documented inconsistently or not assessed using standardized procedures.

The CGI-STIM was developed to address this implementation gap. By relying on information routinely available in clinical documentation, it provides a pragmatic framework for structured outcome monitoring that may support both clinical practice and research. If validated prospectively, the CGI-STIM could facilitate broader implementation of measurement-based care in ECT and may facilitate implementation of existing guideline recommendations.

Comparison with Existing Instruments

Several instruments are available to assess clinical improvement, global functioning, or cognition in psychiatric practice, including the CGI-I (Guy 1976), the GAF (APA 2000), and the MMSE (Folstein et al. 1975). However, these instruments assess individual outcome domains separately and do not provide an integrated evaluation of therapeutic benefit and cognitive tolerability.

The CGI-STIM was developed to address this practical gap. Rather than replacing established symptom or cognitive measures, it provides a structured framework for documenting treatment response and cognitive tolerability using information that is typically available in routine clinical documentation. By integrating both domains within a single clinically interpretable outcome matrix, the CGI-STIM may facilitate systematic outcome monitoring while remaining feasible for routine clinical use.

Limitations

Several limitations should be considered when interpreting these findings. First, the study was based on a small retrospective sample from a single center, limiting generalizability. Second, ratings of cognitive tolerability were derived from routine clinical documentation and were not supported by standardized neuropsychological assessments. Objective cognitive outcomes were not systematically available in this naturalistic dataset. Although baseline MMSE scores were documented for a subset of patients, post-treatment MMSE data were available for only one case and therefore precluded quantitative analysis. This reflects the challenges associated with implementing comprehensive neuropsychological assessments in routine clinical settings. Third, the present study was not designed as a formal psychometric validation study. The primary objective was to evaluate feasibility and interrater agreement in a real-world clinical setting. Consequently, conclusions regarding construct validity, criterion validity, and sensitivity to clinical change remain preliminary. Finally, the CGI-STIM outcome categories were developed pragmatically and have not yet undergone formal psychometric evaluation.

Future Directions and Conclusions

Future studies should prospectively evaluate the CGI-STIM in larger and more diverse ECT samples and examine its validity, sensitivity to change, and utility for routine outcome monitoring. Comparisons with standardized symptom ratings and neurocognitive measures will be important to determine the extent to which CGI-STIM classifications correspond to established clinical outcomes.

In summary, the CGI-STIM represents a feasible and pragmatic clinician-rated outcome monitoring tool that integrates treatment response and cognitive tolerability within a structured framework. By

supporting consistent and clinically meaningful outcome documentation, the CGI-STIM may contribute to routine outcome monitoring and quality improvement in ECT care.

Declarations

Conflicts of Interest and Source of Funding

The authors declare no known competing financial interests or personal relationships that could have influenced the work.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contributions (CRediT Taxonomy)

Conceptualization: SK; Methodology: SK, CP; Formal analysis: SK; Investigation: SK; Data curation: SK; Writing – original draft: SK; Writing – review & editing: SK, CP, AJF; Visualization: SK; Supervision: AJF; Project administration: SK; Funding acquisition: None.

Ethical standards

The study was approved by the Ethics Committee of the Medical Faculty of the University of Mainz (approval number: 2021-15792). All participants provided written informed consent prior to study participation. The study was conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Declaration of Helsinki and its later amendments.

Acknowledgments

The authors thank the clinical staff of the Department of Psychiatry III, Rheinhessen-Fachklinik Alzey, for their support in patient care and documentation, which made this retrospective analysis possible.

Use of Artificial Intelligence (AI) Tools

Language editing assistance was supported by AI-based tools (ChatGPT, OpenAI). All content was reviewed and approved by the authors, who take full responsibility for the manuscript's accuracy and integrity.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. Due to the retrospective clinical nature of this study, the data contain potentially sensitive patient information and cannot be shared publicly.

References

1. Weiner RD, Reti IM (2017) Key updates in the clinical application of electroconvulsive therapy. *Int Rev Psychiatry* 29:54–62. <https://doi.org/10.1080/09540261.2017.1309362>
2. Stella F, Radanovic M, Gallucci-Neto J, Forlenza OV (2023) Electroconvulsive therapy for treating patients with agitation and related behavioral disorders due to dementia: a systematic review. *Dement Neuropsychol* 17:e20230007. <https://doi.org/10.1590/1980-5764-DN-2023-0007>
3. Kayser S, Fallgatter A, Funer F (2026) Electroconvulsive Therapy for Aggression and Agitation in Dementia: A Systematic Review With Exploratory Meta-analytic and Ethical Perspectives. *Am J Geriatr Psychiatry* 34:624–638. <https://doi.org/10.1016/j.jagp.2026.01.004>
4. Fink M (2001) Convulsive therapy: a review of the first 55 years. *J Affect Disord* 63:1–15. [https://doi.org/10.1016/s0165-0327\(00\)00367-0](https://doi.org/10.1016/s0165-0327(00)00367-0)
5. Zatorski T, Szczepańska M, Kabała D, Głód M, Jaskulski A, Zapałowska A (2024) M. B Addressing patient concerns: Exploring Electroconvulsive Therapy (ECT) in primary care: Review of the literature. *Medical Science* 28: e119ms3430. <https://doi.org/https://doi.org/10.54905/disssi.v28i150.e119ms3430>
6. Semkovska M, McLoughlin DM (2010) Objective cognitive performance associated with electroconvulsive therapy for depression: a systematic review and meta-analysis. *Biol Psychiatry* 68:568–577. <https://doi.org/10.1016/j.biopsych.2010.06.009>
7. Semkovska M, Knittle H, Leahy J, Rasmussen JR (2023) Subjective cognitive complaints and subjective cognition following electroconvulsive therapy for depression: A systematic review and meta-analysis. *Aust N Z J Psychiatry* 57:21–33. <https://doi.org/10.1177/00048674221089231>
8. Robertson H, Pryor R (2006) Memory and cognitive effects of ECT: informing and assessing patients. *Adv Psychiatr Treat* 12(3):228–238. <https://doi.org/doi:10.1192/apt.12.3.228>
9. Thirthalli J, Sinha P, Sreeraj VS (2023) Clinical Practice Guidelines for the Use of Electroconvulsive Therapy. *Indian J Psychiatry* 65:258–269. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_491_22
10. DGPPN (2022) Stellungnahme der DGPPN zu den Indikationen der Elektrokonvulsionstherapie. *Nervenheilkunde* 41(10):706–709. <https://doi.org/10.1055/a-1917-9004>
11. APA (2024) In: Association AP (ed) *The Practice of Electroconvulsive Therapy: Recommendations for Treatment, Training, and Privileging*, 3 edn. American Psychiatric Association Publishing, Washington, DC
12. Weiss A, Hussain S, Ng B, Sarma S, Tiller J, Waite S, Loo C (2019) Royal Australian and New Zealand College of Psychiatrists professional practice guidelines for the administration of electroconvulsive therapy. *Aust N Z J Psychiatry* 53:609–623. <https://doi.org/10.1177/0004867419839139>
13. APA. APA Task Force on Professional Practice Guidelines on Measurement-Based Care (2025) [
14. Whitmyre ED, Esposito-Smythers C, Lopez R Jr., Goldberg DG, Liu F, Defayette AB (2024) Implementation of Measurement-Based Care in Mental Health Service Settings for Youth: A Systematic Review. *Clin Child Fam Psychol Rev* 27:909–942. <https://doi.org/10.1007/s10567-024-00498-z>

15. Guy W ECDEU Assessment Manual for Psychopharmacology. U.S. Department of Health E, and Welfare, editor. Rockville1976
16. Busner J, Targum SD (2007) The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont)* 4:28–37
17. Spearing MK, Post RM, Leverich GS, Brandt D, Nolen W (1997) Modification of the Clinical Global Impressions (CGI) Scale for use in bipolar illness (BP): the CGI-BP. *Psychiatry Res* 73:159–171. [https://doi.org/10.1016/s0165-1781\(97\)00123-6](https://doi.org/10.1016/s0165-1781(97)00123-6)
18. Haro JM, Kamath SA, Ochoa S, Novick D, Rele K, Fargas A, Rodriguez MJ, Rele R, Orta J, Kharbeng A, Araya S, Gervin M, Alonso J, Mavreas V, Lavrentzou E, Lontos N, Gregor K, Jones PB, Group SS (2003) The Clinical Global Impression-Schizophrenia scale: a simple instrument to measure the diversity of symptoms present in schizophrenia. *Acta Psychiatr Scand Suppl* 16–23. <https://doi.org/10.1034/j.1600-0447.107.s416.5.x>
19. Isserles M, Daskalakis ZJ, Kumar S, Rajji TK, Blumberger DM (2017) Clinical Effectiveness and Tolerability of Electroconvulsive Therapy in Patients with Neuropsychiatric Symptoms of Dementia. *J Alzheimers Dis* 57:45–51. <https://doi.org/10.3233/JAD-161000>
20. APA (2000) In: Association AP (ed) Global Assessment of Functioning (GAF) Score. Fourth Edition, Text Revision (DSM-IV-TR). American Psychiatric Association, Washington, DC