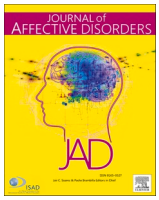
Journal of Affective Disorders 390 (2025) 119719

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad



Research paper

Benchmarking empirical severity for the Yale-Brown Obsessive Compulsive Scale-Second Edition

Caitlin M. Pinciotti [a], Juliana Avery [a], Chencheng Zhang [b], Josselyn S. Muñoz [a], Dayan Berrones [a], Vanessa Zavala Cruz [a], Andrew D. Wiese [a], Jacey L. Anderberg [a,c], Renee M. Frederick [a,d], Tomás Miño [e], Nuria Lanzagorta [f], Juan Camilo Restrepo [g], Marcos E. Ochoa-Panaifo [h], Latin American Trans-ancestry INitiative for OCD genomics (LATINO)[2], Brazilian Obsessive-Compulsive Spectrum Work Group (GTTOC)[2], Wayne K. Goodman [a], James J. Crowley [i,1], Eric A. Storch [a,*,1], Matti Cervin [a,j,1]

[a] Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, United States
[b] Shanghai Jiao Tong University School of Medicine, Shanghai, China
[c] Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE, United States
[d] Department of Psychology, University of Houston, Houston, TX, United States
[e] Centro Interdisciplinario de Neurociencia de Valparaíso, Valparaíso, Chile
[f] Investigación Clínica, Grupo Médico Carracci, Ciudad de México, Mexico
[g] Facultad de Psicología y Ciencias del Comportamiento, Universidad de La Sabana, Chía, Cundinamarca, Colombia
[h] Universidad Privada del Norte, Lima, Peru
[i] Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[j] Department of Clinical Sciences, Lund University, Lund, Sweden

ARTICLE INFO

Keywords:
Obsessive-compulsive disorder
Yale-Brown Obsessive Compulsive Scale-Second Edition
Severity benchmarks
Receiver-operating characteristic analyses
Assessment

ABSTRACT

The Yale-Brown Obsessive Compulsive Scale (Y-BOCS) is considered the primary instrument for assessing the presence and severity of obsessive-compulsive disorder (OCD). Conceptual and empirical critiques inspired the development of an updated version of the instrument, the Yale-Brown Obsessive Compulsive Scale-Second Edition (Y-BOCS-II), with a higher ceiling of OCD severity to better differentiate between severe and the most debilitating OCD presentations, among other revisions. The Y-BOCS-II has demonstrated sound psychometric properties across diverse samples. Empirically derived severity benchmarks have been proposed for the original Y-BOCS, yielding somewhat different ranges than what has been commonly used in clinical and research settings, yet severity benchmarks for the Y-BOCS-II have yet to be established. Using a diverse, pooled sample of 2982 children and adults with OCD or obsessive-compulsive and related concerns across 13 countries, receiver-operating characteristic (ROC) analyses yielded severity benchmarks that largely mirrored the original Y-BOCS at the lower range of scores and extended the previously established benchmarks at the higher range of scores, owing to the increased ceiling of the instrument. The optimal benchmark ranges were determined as: non−/sub-clinical (0–14), mild (15–21), moderate (22–34), severe (35–50). Similar benchmarks were present across sex and age groups, and their accuracy was adequate in both a holdout sample and an independent sample of OCD patients from China ($n = 78$). Limitations and implications for the use of the Y-BOCS-II in clinical and research settings are discussed.

* Corresponding author at: Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, One Baylor Plaza, MS: 350, Houston, TX 77030, United States.
E-mail address: eric.storch@bcm.edu (E.A. Storch).

[2] The collaborators of Latin American Trans-ancestry INitiative for OCD genomics (LATINO) and Brazilian Obsessive-Compulsive Spectrum Work Group (GTTOC) are listed in Appendix A.

https://doi.org/10.1016/j.jad.2025.119719
Received 28 March 2025; Received in revised form 11 June 2025; Accepted 14 June 2025
Available online 18 June 2025
0165-0327/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

1. Introduction

Obsessive-compulsive disorder (OCD) is a debilitating psychiatric disorder with a lifetime prevalence of 1–2 % (Fawcett et al., 2020; Ruscio et al., 2010). It is characterized by obsessions (i.e., intrusive and unwanted thoughts, images, impulses or urges) and compulsions (i.e., repetitive behaviors performed in response to an obsession or according to rigid rules; APA, 2013; WHO, 2019). OCD typically follows a chronic course and is associated with markedly reduced quality of life and a wide range of functional impairments as symptom severity increases (Eisen et al., 2006; Visser et al., 2014).

The Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman et al., 1989) is considered the primary instrument for assessing the presence and severity of OCD. The Y-BOCS consists of a symptom checklist, where the presence of a wide range of obsessions and compulsions is assessed, and a severity scale, which is used to assess overall severity. The severity scale is used to rate the endorsed obsessions and compulsions, respectively, with each item rated on a 0 to 4 scale, resulting in a total score ranging from 0 to 40. In the more than three decades since its development, the Y-BOCS has consistently demonstrated strong psychometric properties by way of internal consistency, convergent validity, interrater reliability, and test-retest reliability in a variety of cross-cultural and international samples (e.g., Arrindell et al., 2002; Kraus et al., 2015; Kuckertz et al., 2021; López-Pina et al., 2015; Ólafsson et al., 2010; Storch et al., 2017; Tek et al., 1995; Tolin et al., 2005; Wu et al., 2007), yet a stable factor structure and adequate loadings have not consistently been found (e.g., Storch et al., 2005).

The Y-BOCS may be clinically advantageous over other OCD assessments because of the established benchmarks that allow for clearer conceptualization of the severity of OCD for each research participant or patient. For many years, the Y-BOCS was interpreted in both research and clinical settings with benchmarks of 0–7 (subclinical), 8–15 (mild), 16–23 (moderate), 24–31 (severe), and 32–40 (extreme). Recently, Cervin et al. (2022) reexamined these benchmarks in 5140 individuals with a lifetime OCD diagnosis aged 5–82 years of age from Sweden, Brazil, South Africa, US, and India. Child (CY-BOCS; Scahill et al., 1997) and adult versions of the Y-BOCS were validated with the Clinical Global Impression-Severity (CGI-S; Guy, 1976) scale, a single measure of global disorder severity rated 1 to 7. Variation in Y-BOCS severity accounted for nearly 48 % of the variance in the CGI-S classification. In contrast to the commonly used benchmarks, the authors found support for empirically derived benchmarks across the lifespan of 0–13 (subclinical), 14–21 (mild), 22–29 (moderate), and 30–40 (severe), all of which demonstrated strong sensitivity and specificity with the CGI—S. Moderately severe OCD (associated with a CGI-S score of 4 or 5) was most prevalent in the sample (68.1 %), followed by mild (CGI-S score of 3, 15.3 %), then severe (CGI-S score of 6 or 7, 10.8 %), and lastly, subclinical OCD (CGI-S score of 1 or 2, 5.8 %). Notably, the non-empirically derived Y-BOCS cutoff of 16 that had been used as the inclusion threshold in most clinical trials was suggested to be revised to 14, a score that demonstrated excellent sensitivity and positive predictive value, and adequate specificity and negative predictive value.

Conceptual and empirical critiques of the Y-BOCS inspired the development of an updated version of the interview, the Yale-Brown Obsessive Compulsive Scale-Second Edition (Y-BOCS-II; Storch et al., 2010). The Y-BOCS-II was developed to address the ceiling effect of the Y-BOCS, which prevented differentiation between severe and the most debilitating OCD presentations, and to include relevant compulsions and active avoidance items missing from the original checklist. The Y-BOCS-II also replaced the original "resistance to obsessions" item, which was determined to be conceptually inaccurate and demonstrated the weakest psychometrics of all items (McKay et al., 1995; Woody et al., 1995), with a more theoretically consistent "obsession-free interval" item. The response range of the ten severity scale items was restructured into a 0 to 5 scale leading to a total score that could range up to ten points higher than the original Y-BOCS (0–50). Psychometric evaluation of the Y-BOCS-II has demonstrated similarly strong psychometric properties as the Y-BOCS with respect to internal consistency, validity, sensitivity, and inter-rater and test-retest reliability, including for both the interview and self-report formats, and improved psychometric support for the revised Item 2 (obsession-free interval) as well as overall factor structure (Alić et al., 2022; Castro-Rodrigues et al., 2018; Hiranyatheb et al., 2014, 2015; Melli et al., 2015; Pinciotti et al., 2024; Storch et al., 2010, 2025; Vogt et al., 2022; Wu et al., 2016).

Empirically derived severity benchmarks have yet to be established for the Y-BOCS-II. As the rating scales and total possible scores differ between the Y-BOCS and Y-BOCS-II, benchmarking the Y-BOCS-II is needed to provide a consistent definition of severity indicators in research and clinical settings. To that end, this study aims to provide empirically derived severity benchmarks for the adult and child versions of the Y-BOCS-II using a large, multinational sample of children, adolescents, and adults with obsessive-compulsive symptoms.

2. Methods

2.1. Participants and procedures

The current study consists of a pooled sample of 3061 individuals diagnosed with OCD across 12 countries. Data were collected from two subsamples. The first subsample was recruited as part of a larger research project, the Latin American Trans-ancestry INitiative for OCD genomics (LATINO), which aims to diversify OCD genomic findings and investigate clinical presentations of OCD among Latin American and Hispanic populations. Participants were recruited across the 12 countries through normal patient flow of specialist institutions and clinics as well as advertisements through social media, community engagement, and research organizations (i.e., ResearchMatch and SPARK). This sample includes 2982 participants between ages 7–89 years who are of Latin American descent from Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, El Salvador, Mexico, Paraguay, Peru, Puerto Rico, and the United States.

All participants reported OCD or obsessive-compulsive and related concerns and underwent diagnostic assessment using the MINI International Neuropsychiatric Interview (MINI/MINI-KID; Sheehan et al., 1998, 2010) or the Structured Clinical Interview for DSM-5 Disorders (SCID-5; First, 2015) administered by post-baccalaureate and graduate level assessors under the supervision of licensed psychologists/psychiatrists.[3] Exclusion criteria included cases of diagnostic uncertainty and individuals who lack the decisional capacity to consent and participate. The second subsample (i.e., holdout sample) consisted of 78 participants recruited through regular patient flow of outpatient psychiatry clinics in China as part of a research study examining the psychometric properties of the Mandarin Chinese version of the Y-BOCS-II (Zhang et al., 2019). All participants had a diagnosis of OCD determined through a clinical interview administered by a licensed psychologist or psychiatrist and were between the ages of 15–70 years. Exclusion criteria included the diagnosis of a psychotic disorder and inability to read and understand the assessment.

Both studies' procedures were reviewed and approved by their local institutional review board. Each participating non-U.S. site of the LATINO study received ethical approval from the presiding ethics committee at their institution to implement the harmonized protocol described in Crowley et al. (2024). Study data were collected and

[3] All assessors achieved certification in measurement administration by reviewing training videos developed by an expert multi-national committee. Assessors were required to meet basic proficiency for certification (i.e., >80 % Y-BOCS-II accuracy). If assessors were not independently licensed, they were required to attend weekly supervision led by the local lead investigator or an approved member of the study team. Additionally, Y-BOCS-II scores were reviewed by a data quality committee to improve reliability.

managed using REDCap electronic data capture tools hosted at Baylor College of Medicine, University of São Paulo, and Shanghai Jiao Tong University School of Medicine.

2.2. Measures

2.2.1. Clinical Global Impression-Severity (CGI-S; Guy, 1976)

The CGI-S is a single-item clinician-rated measure of psychopathology severity. The measure consists of a 7-point scale ranging from 1, which indicates no illness, to 7, which indicates among the most extremely ill. The CGI-S has demonstrated validity among clinical populations, is responsive to treatment, and has often been utilized in clinical trials (Berk et al., 2008; Hollander et al., 2003). The CGI-S also has been used to establish severity benchmarks for the original Y-BOCS/CY-BOCS (Cervin et al., 2022). Importantly, the CGI-S rating is based on all available clinical information, making it a suitable reference measure when estimating benchmarks. In accordance with previous studies using the CGI-S as a reference measure for OCD, severity scores of 1 and 2 were considered non-/sub-clinical, scores of 3 ("Mildly ill") to be mild, scores of 4 ("Moderately ill") and 5 ("Markedly ill") to be moderate, and scores of 6 ("Severely ill") and 7 ("Among the most extremely ill") to be severe. While the Y-BOCS-II is designed to differentiate the most extremely severe cases, the number of participants with a CGI-S score of 7 were too few ($n = 9$) to be analyzed in the present study.

2.2.2. Yale-Brown Obsessive Compulsive Scale – Second Edition (Y-BOCS-II; Storch et al., 2010)

The Y-BOCS-II and the corresponding Children's Y-BOCS-II (CY-BOCS-II; Storch et al., 2019) are semi-structured clinician-administered measures of OCD severity. The CY-BOCS-II was administered to those under the age of 18 in the LATINO sample. All ages of participants in the sample from China (range = 15–70 years) received the Y-BOCS-II. The Y-BOCS-II and CY-BOCS-II consist of a symptom checklist of obsessions, compulsions, and avoidance behaviors and a 10-item severity scale, with each severity item rated from 0 to 5. The current study utilized the English, Spanish, Portuguese, and Mandarin Chinese versions of the Y-BOCS-II, each of which has demonstrated strong psychometric properties (Castro-Rodrigues et al., 2018; Storch et al., 2010, 2025; Zhang et al., 2019). Access www.OCDscales.com to request access for the Y-BOCS-II and review terms of use.

2.3. Statistical analysis

We first explored the consistency of the data across sites. The CGI-S scores at one site were deemed to be of uneven quality as a very low correlation between CGI-S and Y-BOCS emerged (indicating inconsistencies in scoring) and participants from this site were omitted. Further, one participant was classified as severely ill according to the CGI—S, but had a score of 1 on the Y-BOCS-II and was omitted. As the severity groups were largely balanced, except for the smaller severe group (see Table 1), we used receiver-operating characteristic (ROC) analyses to determine the ability of the Y-BOCS-II to distinguish between the different severity groups. To identify the optimal cutoff, we used the Youden index, which maximizes the sum of sensitivity and specificity and is efficient in the case of balanced groups. To account for the possibility of uneven distributions of Y-BOCS-II scores across the severity groups, we used normal smoothing of the distributions. In the ROC analyses, the area under the curve (AUC) values and their 95 % confidence intervals were used to examine the classification performance of the Y-BOCS-II. AUC values from 0.7 to 0.8 were considered acceptable, values from 0.8 to 0.9 were considered excellent, and values above 0.9 were considered outstanding.

To examine the accuracy of the derived benchmarks, we randomly split the LATINO sample into two subsamples. The first subsample was the exploratory subsample used to derive initial benchmarks and consisted of 80 % of the LATINO participants. The second subsample was the holdout sample, which was used to test the validity of the benchmarks and consisted of the remaining 20 % of participants. The balanced subclinical, mild, and moderate groups could be adequately distinguished using the Youden index. However, the Youden index could not be used to distinguish the moderate and severe groups due to the sample size imbalance (82 [severe] vs 1035 participants [moderate]). Unbalanced groups create difficulties when establishing cutoffs. When using diagnostic accuracy measures, larger groups are prioritized over smaller groups. Conversely, maximizing the sum of sensitivity and specificity decreases accuracy due to bias from the smaller groups. To address these challenges, we aimed to maximize sensitivity in relation to the larger moderate group, while maintaining specificity of ≥0.60 for the severe group.

When benchmarks had been established, their accuracy was tested using the remaining 20 % of participants in the holdout sample and in the independent sample of participants from China. Accuracy was tested by classifying the full holdout sample according to the derived benchmarks and then comparing the classification of each participant to their CGI-S score, yielding an estimation of the overall accuracy of the full range of benchmarks.

Last, to examine the consistency of the benchmarks, we used all available LATINO data to estimate benchmarks separately in subgroups of men and women, and the following age groups: 0–20 years, 20–30 years, 30–50 years, and 50+ years. The sample with participants from China included very few participants in the non−/sub-clinical and severe categories and was excluded from these analyses. All statistical analyses were conducted in R Studio using the libraries *pROC* and *cutpointr*. The statistical script is available upon request.

Table 1
Demographic information and scores on study measures.

LATINO sample, $n = 2982$	
Women, *n* (%)	1892 (64 %)
Age, *M* (*SD*), range	30.81 (12.25), 7–86
Y-BOCS-II *M* (*SD*), range	20.44 (8.73), 0–50
CGI-S, *n* (%)	
1. Normal – not at all ill	192 (6 %)
2. Borderline mentally ill	636 (21 %)
3. Mildly ill	1037 (35 %)
4. Moderately ill	744 (25 %)
5. Markedly ill	291 (10 %)
6. Severely ill	73 (2 %)
7. Among the most extremely ill patients	9 (0.3 %)
Severity	
Non-/sub-clinical, n (%)	828 (28 %)
Mild, n (%)	1037 (35 %)
Moderate, n (%)	1035 (35 %)
Severe, n (%)	82 (3 %)
Age groups	
<20 years, n (%)	383 (12 %)
20–30 years, n (%)	1306 (44 %)
30–50 years, n (%)	1014 (35 %)
50+ years, n (%)	279 (9 %)

China sample, $n = 78$	
Women, *n* (%)	41 (52 %)
Age, *M* (*SD*), range	33.21 (13.99), 15–78
Y-BOCS-II *M* (*SD*), range	26.82 (9.50), 7–45
CGI-S, *n* (%)	
1. Normal – not at all ill	0 (0 %)
2. Borderline mentally ill	4 (5 %)
3. Mildly ill	7 (9 %)
4. Moderately ill	23 (30 %)
5. Markedly ill	30 (39 %)
6. Severely ill	13 (17 %)
7. Among the most extremely ill patients	1 (1 %)

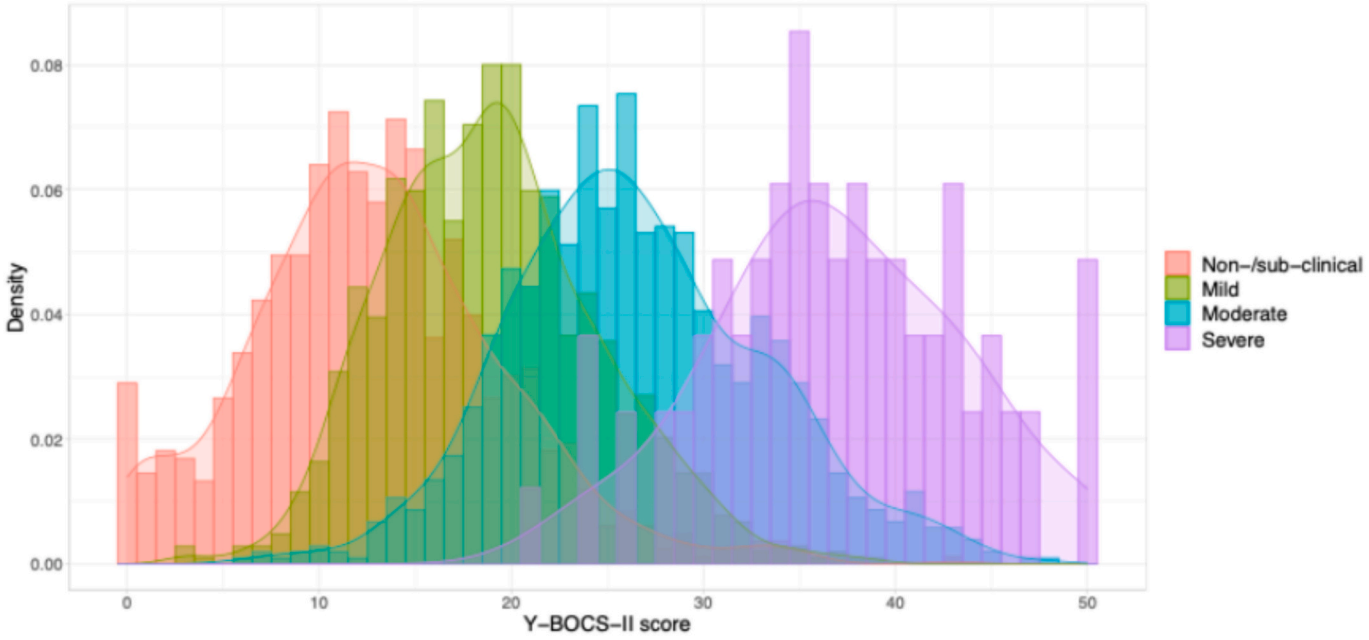


Fig. 1. Distribution of Y-BOCS-II scores across severity categories.

3. Results

3.1. Estimating benchmarks

Y-BOCS-II scores across the CGI-S-based severity groups are presented in Fig. 1. We started by identifying the optimal benchmarks in the exploratory sample. The Y-BOCS-II acceptably distinguished between the mild and the non-/sub-clinical group (AUC = 0.77 [0.74, 0.79]). The optimal cutoff was 15 points, yielding an accuracy of 71 %, a sensitivity of 78 %, and a specificity of 62 %. The Y-BOCS-II excellently distinguished between the moderate and the mild group (AUC = 0.81 [0.79, 0.83]). The optimal cutoff was 22 points, yielding an accuracy of 74 %, a sensitivity of 71 %, and a specificity of 76 %. The Y-BOCS-II excellently distinguished between the moderate and the severe group (AUC = 0.87 [0.82, 0.91]). The optimal cutoff for defining the severe group was 35 points, yielding an accuracy of 86 %, a sensitivity of 88 %, and a specificity of 63 %.

These cutoffs indicated that the non-/sub-clinical group ranged from 0 to 14 points, the mild group from 15 to 21 points, the moderate group from 22 to 34 points, and the severe group from 35 to 50 points.

3.2. Accuracy and consistency of the benchmarks

The overall accuracy of the benchmarks in the holdout sample was 59 %. A large majority of misclassifications (90 %) were either one severity category above or below the true category. The overall accuracy in the independent China sample was 63 % and again, a large majority of misclassifications (90 %) were one severity category above or below the true category. Largely similar benchmarks emerged across subsamples of men and women and age groups in the full LATINO sample. See Table 2.

Table 2
Benchmarks across subsamples.

	Non-/sub-clinical	Mild	Moderate	Severe
Women	0–14	15–21	22–34	35–50
Men	0–13	14–21	22–37	38–50
0–20 years	0–14	15–22	23–	Too few
20–30 years	0–14	15–21	22–33	34–50
30–50 years	0–13	14–22	23–34	35–50
50+ years	0–12	13–20	21–34	35–50

4. Discussion

The Y-BOCS and Y-BOCS-II are considered the primary assessment instruments for quantifying the presence and severity of OCD and are commonly used in clinical trials and treatment settings. Y-BOCS severity benchmarks have afforded researchers and clinicians a common metric to conceptualize and compare OCD severity across and within individuals. Widely accepted severity benchmarks for the Y-BOCS have recently been updated based on findings from a large sample of children, adolescents, and adults with OCD across the world (Cervin et al., 2022). The present study extended this work by establishing empirical severity benchmarks for the Y-BOCS-II using a sample of 2982 children, adolescents, and adults across 13 countries.

The severity benchmarks for the lower range of Y-BOCS-II scores – non-/sub-clinical and mild symptoms – largely mirrored the benchmarks for similar categories of the Y-BOCS, with only one point difference for either category between the current findings using the Y-BOCS-II and earlier findings using the Y-BOCS (Cervin et al., 2022). In the present study, a cutoff for mild OCD on the Y-BOCS-II was suggested to be 15 compared to 14 on the original Y-BOCS. The largest differences in benchmarks between the Y-BOCS and Y-BOCS-II were observed when differentiating moderate from severe OCD cases. For both instruments, the moderate range was found to start at a score of 22; however, in the present study using the Y-BOCS-II, this moderate range extended to 34, which is five points higher than the maximum of this category on the Y-BOCS. On the Y-BOCS-II the severe range was found to begin at 35 (compared to 30 on the Y-BOCS) and extends up to the highest possible score of 50. Importantly, scores between 30 and 34 that were previously considered to reflect severe cases of OCD using the Y-BOCS were categorized as reflecting moderate severity when using the Y-BOCS-II, suggesting a mismatch in interpretation between the two instruments. Given the increased precision offered by the Y-BOCS-II over the Y-BOCS, a conservative approach to conceptualizing a patient or participant discrepantly deemed “severe” with the Y-BOCS yet “moderate” with the Y-BOCS-II would be to defer to the Y-BOCS-II designation. However, future research on implications for treatment response and level of impairment are needed before going this route.

The upward shift of moderate and severity score ranges on the Y-

BOCS-II is intuitive given the higher ceiling of the instrument. By adding a more severe response option for each Y-BOCS-II item, the benchmarks for moderate and severe OCD presentations were pulled towards higher scores. Because the floor of the Y-BOCS-II was unchanged from the Y-BOCS, it is also intuitive that the lowest severity ranges remained unchanged. This means that the conceptualization of minimal-to-no OCD symptoms on the Y-BOCS and Y-BOCS-II appears to be largely the same and can be interpreted virtually interchangeably across the two instruments. Importantly, the benchmarks found in the present study were similar across subsamples, gender, and age, suggesting that these benchmarks can be used with diverse samples.

4.1. Strengths, limitations, and future directions

The present study is strengthened by its use of a large demographically heterogenous, multinational sample inclusive of individuals across the lifespan, offering generalizable findings regardless of culture, gender, or age. However, most participants were of Latin heritage, offering the potential that findings might not replicate to other ethnicities. Additionally, the use of an independent sample to test the accuracy of the derived benchmarks is a strength.

However, very few individuals were rated as having severity and impairment in the highest CGI-S category – those who are "among the most extremely ill." Because participants were recruited from the community and outpatient treatment settings, the sample was not representative of individuals with the most severe and impairing cases of OCD. This is a notable limitation given the scale's intention to better capture the most extreme cases and restricted our ability to explore the extent to which the upper end of the Y-BOCS-II is informative. It is unclear how inclusion of more individuals in this category might impact benchmarks, such as whether the severe range would split into "severe" and a new category of "incapacitating," as has been theoretically suggested (Pinciotti et al., 2024). We accounted for this limited subsample by favoring sensitivity in the moderate range while maintaining specificity for the severe range. Alternative methods (e.g., Youden, F1) would have resulted in substantially lower precision or a failure to identify the most severe cases altogether.

Prioritizing sensitivity is analytically justified when the aim is to avoid missing rare but clinically important severe cases — even at the cost of some false positives. As long as clinical populations largely mirror the present population, we are confident that our approach provides the best estimate of where to distinguish severe from moderate cases. As most populations will mirror the present population given the low proportion of individuals who are completely incapacitated by OCD (e.g., Alhaj et al., 2022), it is reasonable to assume that these benchmarks would translate to most populations. However, if the populations in which these benchmarks are applied differ substantially from ours — for example, with a much higher proportion of severe cases, such as samples of patients in intensive residential or inpatient treatment — the estimates may be less optimal. Still, we believe our approach offers greater generalizability than alternatives such as the Youden index, which would likely overestimate the number of severe cases in most settings. A replication study inclusive of individuals with incapacitating OCD, including those in residential or inpatient levels of care, would offer greater validity to the upper benchmark ranges suggested in the present study.

We note that while the overall classification accuracy of our benchmarks is modest at approximately 60 %, this level of performance is consistent with prior studies attempting to map dimensional severity scores onto categorical benchmarks (e.g., Cervin et al., 2022). We believe it reflects the inherent challenge of discretizing continuous clinical phenomena; the more categories applied, the lower the precision of the resulting benchmarks. However, it may also partly reflect that the Y-BOCS-II (and the original Y-BOCS) does not fully capture all aspects of clinical severity considered important by trained clinicians, which by definition can be accounted for using the CGI—S. That said, it could be argued that the CGI-S is an insufficient sole measure of severity given its subjectivity and cross-cultural variability which limit its precision. Indeed, in the present sample, data was excluded from one site due to poor CGI-S and Y-BOCS-II correlation which could indicate inter-rater inconsistency. While the CGI-S allows trained assessors to apply their judgment in rating overall severity, drawing on a holistic assessment of the patient that goes beyond symptom descriptions alone, the measure requires clearer operationalization within psychiatric research. Although this falls outside the scope of the current study, it represents a worthwhile future endeavor particularly given its frequent use as a primary or secondary outcome in clinical trials.

5. Conclusions

The present study is the first to offer updated severity benchmarks for the Y-BOCS-II. Findings from our study suggest the following benchmarks be used to categorize OCD severity when using the Y-BOCS-II: non-/sub-clinical (0–14), mild (15–21), moderate (22–34), severe (35–50). Given the moderate accuracy of these benchmarks in the holdout and independent samples, the Y-BOCS-II—like the original Y-BOCS (Cervin et al., 2022)—is recommended to be used in conjunction with other assessments, including clinical judgment and observational assessment, to establish the most comprehensive conceptualization of OCD among patients and research participants.

CRediT authorship contribution statement

Caitlin Pinciotti: Conceptualization, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Juliana Avery:** Conceptualization, Investigation, Project administration, Writing – original draft, Writing – review & editing. **Chencheng Zhang:** Data curation, Formal analysis, Writing – original draft. **Josselyn Muñoz:** Data curation, Investigation, Project administration, Writing – review & editing. **Dayan Berrones:** Data curation, Investigation, Project administration, Writing – review & editing. **Vanessa Zavala Cruz:** Data curation, Investigation, Project administration, Writing – review & editing. **Andrew Wiese:** Data curation, Investigation, Project administration, Supervision, Writing – review & editing. **Jacey Anderberg:** Data curation, Investigation, Project administration, Writing – review & editing. **Renee Frederick:** Data curation, Investigation, Project administration, Writing – review & editing. **Tomás Miño:** Data curation, Investigation, Project administration, Supervision, Writing – review & editing. **Nuria Lanzagorta:** Data curation, Investigation, Project administration, Supervision, Writing – review & editing. **Juan Camilo Restrepo:** Data curation, Investigation, Project administration, Supervision, Writing – review & editing. **Marcos Ochoa:** Data curation, Investigation, Project administration, Supervision, Writing – review & editing. **Latin American Trans-Ancestry INitiative for OCD genomics (LATINO):** Data curation, Investigation, Project Administration, Writing – review & editing. **Brazilian Obsessive-Compulsive Spectrum Work Group (GTTOC):** Data curation, Investigation, Project administration, Writing – review & editing. **Wayne Goodman:** Data curation, Investigation, Project administration, Writing – review & editing. **James Crowley:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Eric Storch:** Conceptualization, Funding acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Matti Cervin:** Data curation, Formal analysis, Investigation, Validation, Writing – original draft, Writing – review & editing.

All authors contributed to the writing of the final manuscript.

Funding

Research reported in this publication was supported by the NIMH of the National Institutes of Health under Award Numbers U01MH125062

and U01MH125050. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This study was also supported by a donation from the Gillson Longenbaugh Foundation. This research was supported by the Ancestral Population Network (APN), a National Institute of Mental Health Initiative. The APN was established by the NIMH, to support data collection and analysis of psychiatric phenotypes across diverse ancestral populations.

Declaration of competing interest

Dr. Rodriguez (in the last three years) has been a consultant for Biohaven Inc., Osmind, and Biogen; received research grant support from Biohaven Inc.; received royalties from American Psychiatric Association Publishing; and received a stipend from APA Publishing for her role as Deputy Editor at The American Journal of Psychiatry and a stipend for her role as Deputy Editor of Neuropsychopharmacology.

Dr. Arnold receives research funding from Biohaven Inc., Canadian Institute for Health Research, Ontario Brain Institute, Genome Canada, and Alberta Children's Hospital Research Foundation.

Dr. Martínez-González receives funding from NIH and FDA. She is a consultant with Abbvie, Sage Pharmaceuticals and SAMHSA.

Dr. Goodman receives research fundings from NIH, Biohaven, and the McNair Foundation and consulting fees from Biohaven. He receives royalties from Proem.

Dr. Storch reports receiving research funding to his institution from the Ream Foundation, International OCD Foundation, and NIH. He is a consultant for Brainsway and Biohaven Pharmaceuticals. He owns stock less than $5000 in NView. He receives book royalties from Elsevier, Wiley, Oxford, American Psychological Association, Guildford, Springer, Routledge, and Jessica Kingsley.

Dr. Cervin receives research support from the Kavli Foundation, the Swedish Research Council for Health, Working Life and Welfare, the Lindhaga Foundation, Stiftelsen Clas Grochinskys Minnesfond, the Crown Princess Lovisa's Association, the Kamprad Family Foundation, Fonden för Psykisk Hälsa, and Skåne University Hospital's Foundations and Donations; and financial compensation from Springer for editorial work outside of the submitted work.

All other authors report no financial disclosures.

Acknowledgements

The authors would like to thank all of the participants, clinics, and colleagues who made this project possible. We would like to acknowledge Gabrielle Armstrong, Laura Barón Castano, Sanie Marie Bulsara, Shaun Bulsara, Mabel Carrera-Olivares, María Catano-Cedillo, Kyra Cheung, Duff Dean, José Erazo, Bharathi Gadad, Katy Gathron, Miri Gitik, Letty Guerrero, Imaje Harvey, Carlos Hines, Andrea Horvath Marques, Rachel Jocson, David Keller, Minjee Kook, Daniel Macioci, María Martínez, Guanpei Ming, Carlos Miranda, Mary Beth Morgan, Jhonny Muñoz, Emily O'Bryant, Maria Ochoa Rojas, Darpan Patel, Laura Petty, Uma Ramamurthy, Vivek Ramanathan, Catherine Rast, Bradley Riemann, Juan Rivas, Santiago Romero, Emse Rosario, Brenda Patricia Salcedo Martínez, Elda Yadira Salcedo Martínez, Leanne Scott, Natasha Sefcovic, McKenzie Sluder, Sharron Smart, Keaton Soileau, Andrew Storch, Andres Viana, and Jackie Wynn.

Recruitment for this study was facilitated by the Simons Simplex Collection (SSC). We are grateful to all of the families at the participating SSC sites, as well as the principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, E. Wijsman).

Recruitment for the study included ResearchMatch, a national health volunteer registry that was created by several academic institutions and supported by the U.S. National Institutes of Health as part of the Clinical Translational Science Award (CTSA) program. ResearchMatch has a large population of volunteers who have consented to be contacted by researchers about health studies for which they may be eligible. Review and approval for this study and all procedures were obtained from Baylor College of Medicine.

Appendix A. The collaborators of the Latin American Trans-ancestry INitiative for OCD genomics (LATINO)

Victor R. Adorno[1,2], Victoria Agostini[3], William W. Aguilar[4], Cinthia Aguirre[1], Valentina Alvarado-Quiroz[5], NaEshia Ancalade[6], Maria E. Anciburo L[7], Jacey L. Anderberg[8,9], Diego Aponte[3], Alejandro A. Arellano Espinosa[10], Paul D. Arnold[11,12], Brygith Asenjo Carmelo[7], Elizabeth G. Atkinson[13], Tatevik Avanesyan[14], Juliana Avery[8], Jose N. Ayala[1,15], Hala Aziz[16], Tania L. Barbieri Aguirre[7], Cynthia N. Barrera[17], Julian A. Barrero Contreras[18], Kelly Barry[8,19], Amanda N. Belanger[20], Laura M. Benitez[1,21], Dayan Berrones[8], John R. Best[22], Tim B. Bigdeli[23], Hemamalini Bommiasamy[6], Tania Borda[24,25], Laura Boyajian[26] Diego Briceño Muñoz[7], Lauren Browning[27], Christie L. Burton[28], Carolina Busto[29], Joseph D. Buxbaum[30], Ricardo E. Calderón Rivera[3], Jennifer L. Callahan[31], Carlos A. Camacho Gomez[32,33], Laura A. Campos Ponce[7], Ernesto G. Canchari[7], Roxana Carrasco Torres[7], Matti Cervin[8,34], Alexandria Chang[27], Evelyn Chazelle[35,36], Jessica S.C. Cheng[8], Jessica M. Chire[37,38], Macarena Churruca Muñoz[5,39], David C. Cicero[40], Amanda P. Cifuentes-Espinoza[5], Pamela Claisse Quiroz[41], Gianna M. Colombo[8], Jonathan S. Comer[29], Luz Condori-Apaza[7], Jennifer Cortina[29], Jennifer Crosbie[28], James J. Crowley[6], Victor O. Cruz[37,42], Isabella A. Cruz[43], Allison Cunning[26] Guillermo J. Dager Perez[44], Megan M. Dailey[8], Brian D. Dang[8,45], Kelly L. Davalos Checco[7], Alyssa Dawson[29], Daniela del Río[41], Amelia Dev[20], Sara Diaz[3], Lucila Fay[35,46], Edith G. Figueroa[37], Letizia M. Flecha[1,21], Álvaro M. Flores-Garcia[7], Martin E. Franklin[47,48], Renee M. Frederick[8,19], Ricardo Fujita[49], Jami M. Furr[29], Alessa Gaeta[14], Olivia M. Gamarra[1,50], Camila Garay-Contreras[5,39], Nataly García-Bravo[5], Cynthia García-Jaimes[51], Sheldon R. Garrison[52], Dariana Gil-Hernández[14], Paola Giusti-Rodriguez[53], Amin Gonzalez[1,2], Ramon A. González[54], Polaris Gonzalez-Barrios[3], Wayne K. Goodman[8], Javier D. Gorosito[55], Dorothy E. Grice[30], Jerry Guintivano[56], Daniel G. Guttfreund[57], Matthew W. Halvorsen[6], Jorge A.R. Haslop Miles[58,59], Estefanía V. Hernández Núñez[54], John M. Hettema[60], Joseph D. Hovey[43], Bryan Huerta Ccarhuas[7], Jonathan Irreño-Sotomonte[18], Nisha Jagannathan[8], Matias Jensen[5], Alexandra Z. Jimenez Reynolds[8], Rozmin B. Jiwani[61], Jessica S. Johnson[56], Sonia Joutteaux[62], Joali Alexandra Juárez Lujambio[10], Sabrina Kehrer[29], James L. Kennedy[63,64], Nuria Lanzagorta[51], Gabriel Lázaro-Muñoz[65,66,67], Melanie Longhurst[62], David A. Lozada Martinez[68], Elba S. Luna[37,38,42], Anna MacLellan[22], Dora Mamani Quispe[69], Andrea H. Marques[70], Mainor R. Marquez[1], Molly Martinez[71], Galo N. Martínez[72], Mayra C. Martinez Mallen[73], Karen G. Martinez-Gonzalez[3], Maria de Los Angeles Matos[35], Manuel Mattheisen[74], Ryan J. McCarty[75], Joseph F. McGuire[27], Dean McKay[76], Ciara McOmber[6], Charlene Minaya[76], Tomás Miño[5], Sara M. Mithani[61], Alonso Morales-Rivero[77], Pablo R. Moya[5,39,78], María B. Moyano[35,79], Pavithra Mukunda[14], Josselyn S. Muñoz[8], Maria J. Muñoz[80], Humberto Nicolini[51,81], Ambar A. Núñez Bracho[54], Belinda E. Núñez Bracho[54], Diana Obando[80], Monica E. Ochoa[43], Marcos E. Ochoa-Panaifo[68], Marcelo O'Higgins[1,2], Luis A. Olavarria Castaman[68], Trinidad Olivos Balmaceda[5], María J. Olmedo Tobar[5], Ogechi C. Onyeka[8], Iliana Ortega[11,12], Vanessa S. Paredes-Linares[7], Michele T. Pato[82], Mariel Paz y Miño[83], Edilberto Peña de Leon[84], José L. Perales Orellana[32,33], Noelia S. Peralta Sossa[85], Stacey Pereira[86], Natalia A. Pereira Morales[69], Tania Pérez Duarte[17], Angel Phan[87], Maritza Placencia[38,42], Lorena Ponce[17], Kendall Poovey[26], María Belén Prieto[35], Thomas Qiao[11], Maria F. Quiroz-Padilla[80], María F. Quispe-Correa[68], Alessandra S. Ramirez Heros[7], Renato T. Ramos[88], Iaroslava C. Ramos[88], Kesley A. Ramsey[27], Diana Rancourt[26], Sebastian Rebatta Gomez[7], Rocío Celeste Reyes Espinosa[89], Peggy Richter[88], Carolyn I. Rodriguez[14], Servando

Rodriguez Barajas[8], Maritza J. Romero[80], Camilo J. Ruggero[31], Angel Ruiz-Chow[77], Melisa N. Sagarnaga[58,59], Gianella N. Salazar Montoya[7], Nadeen Salhadar[43], Russell J. Schachar[28], Lucia M. Segovia[1,21], Laura D. Seligman[43], Isaac J. Siegel[52], Samuel D. Spencer[40], Cody Staples[26], S. Evelyn Stewart[22], Eric A. Storch[8], Vissente Tapia-Cuevas[5], Peter Thompson[62,90], Kiara R. Timpano[20], Silvina Tonarelli[62] Beatriz Treviño-de la Garza[91] Angela Trujillo[80], Constanza Uribe Villar[5], Carolina Valdés[17], Alexie Vallejo-Silva[18], Katerin N. Vargas Mamani[85], Javier Vargas-Medrano[62,90], María I. Vásquez-Suyo[38,92], Anne Ver Mellstrieng[1,2], Maria V. Vergara[1,50], Guadalupe Vidal-Martinez[62,90], Mariluz V. Vilchez Huaman[7], Caroline Vollaro[29], Andrew D. Wiese[8], Sara Wilkerson[76], Gwyneth Zai[63,64], Luz M. Zappa[35,46], Vanessa Zavala Cruz[8], Raquel M. Zepeda-Burgos[93], Anthony W. Zoghbi[8]

The collaborators of the Brazilian Obsessive-Compulsive Spectrum Work Group (GTTOC)

Gustavo S. Alves[94],Gilberto S. Alves[95,96],Izabela G. Barbosa[97],Carolina Cappi[98],Maria Cecília B. Carneiro[99],Daniel L. Costa[100],Fernanda Z. Delage[99],Carolina B. Dreher[101],Ygor A. Ferrão[102],Gabriela M. Ferreira[99],Leonardo F. Fontenelle[103],Daniele T.H. Fragoso[99],Monicke O. Lima[100],Gabriela Menezes[103],Euripedes C. Miguel[100],Maria E. Moreira-de-Oliveira[103]Bárbara Perdigão Stumpf[97],Lucas C. Quarantini[94],Yana Quarantini-Alvim[94],Vanessa R. Ramos[100],Maria Cecília C. Rosario[104], Aline S. Sampaio[94],Leonardo C. Saraiva[100],Roseli G. Shavitt[100]

References

Alhaj, H., Ali, A., Askari, H., Ibrahim, H., Ali, L.M., Kassasand, L., Nasere, N.A., 2022. The prevalence of OCD like symptoms among UAE residents during COVID-19 pandemic. BJPsych Open 8 (Suppl. 1), S2. https://doi.org/10.1192/bjo.2022.78.

Alić, M., Leeuw, A., Selier, J., Megen, H., Visser, H., 2022. Responsiveness and other psychometric properties of the Yale–Brown Obsessive–Compulsive Scale Severity Scale—Second Edition in a Dutch clinical sample. Clinical Psychology and Psychotherapy 29 (4), 1355–1363. https://doi.org/10.1002/cpp.2715.

APA, 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. (Washington DC).

Arrindell, W., de Vlaming, I.H., Eisenhardt, B.M., van Berkum, D.E., Kwee, M.G., 2002. Cross-cultural validity of the Yale–Brown Obsessive Compulsive Scale. J. Behav. Ther. Exp. Psychiatry 33 (3), 159–176. https://doi.org/10.1016/S0005-7916(02)00047-2.

Berk, M., Ng, F., Dodd, S., Callaly, T., Campbell, S., Bernardo, M., Trauer, T., 2008. The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. J. Eval. Clin. Pract. 14 (6), 979–983. https://doi.org/10.1111/j.1365-2753.2007.00921.x.

Castro-Rodrigues, P., Camacho, M., Almeida, S., Marinho, M., Soares, C., Barahona-Corrêa, J.B., Oliveira-Maia, A.J., 2018. Criterion validity of the Yale-Brown Obsessive-Compulsive Scale second edition for diagnosis of obsessive-compulsive disorder in adults. Frontiers in Psychiatry 9, 431. https://doi.org/10.3389/fpsyt.2018.00431.

Cervin, M., OCD Severity Benchmark Consortium, Mataix-Cols, D., 2022. Empirical severity benchmarks for obsessive-compulsive disorder across the lifespan. World Psychiatry 21 (2), 315–316. https://doi.org/10.1002/wps.20984.

Crowley, J.J., Cappi, C., Ochoa-Panaifo, M.E., Frederick, R.M., Kook, M., Wiese, A.D., Rancourt, D., Atkinson, E.G., Giusti-Rodriguez, P., Anderberg, J.L., Latin American Trans-ancestry INitiative for OCD genomics, Brazilian Obsessive-Compulsive Spectrum Disorder Working Group, Abramowitz, J.S., Adorno, V.R., Aguirre, C., Alves, G.S., Alves, G.S., Ancalade, N., Arellano Espinosa, A.A., Storch, E.A., 2024. Latin American Trans-ancestry INitiative for OCD genomics (LATINO): study protocol. Am. J. Med. Genet. B Neuropsychiatr. Genet. 195 (4), e32962. https://doi.org/10.1002/ajmg.b.32962.

Eisen, J.L., Mancebo, M.A., Pinto, A., Coles, M.E., Pagano, M.E., Stout, R., Rasmussen, S. A., 2006. Impact of obsessive-compulsive disorder on quality of life. Compr. Psychiatry 47 (4), 270–275. https://doi.org/10.1016/j.comppsych.2005.11.006.

Fawcett, E.J., Power, H., Fawcett, J.M., 2020. Women are at greater risk of OCD than men: a meta-analytic review of OCD prevalence worldwide. J. Clin. Psychiatry 81 (4), 13075. https://doi.org/10.4088/JCP.19r13085.

First, M.B., 2015. The encyclopedia of clinical psychology structured clinical interview for the DSM (SCID). In: Cautin, R.L., Lilienfeld, S.O. (Eds.), The Encyclopedia of Clinical Psychology, 1st ed. Wiley, pp. 1–6. https://doi.org/10.1002/9781118625392.wbecp351.

Goodman, W.K., Price, L.H., Rasmussen, S.A., Mazure, C., Fleischmann, R.L., Hill, C.L., Heninger, G.R., Charney, D.S., 1989. The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. Arch. Gen. Psychiatry 46 (11), 1006–1011. https://doi.org/10.1001/archpsyc.1989.01810110048007.

Guy, W.C.G.I., 1976. Clinical global impression. In: Assessment Manual for Psychopharmacology, pp. 217–222.

Hiranyatheb, T., Saipanish, R., Lotrakul, M., 2014. Reliability and validity of the Thai version of the Yale-Brown Obsessive Compulsive Scale - Second Edition in clinical samples. Neuropsychiatr. Dis. Treat. 10, 471–477. https://doi.org/10.2147/NDT.S56971.

Hiranyatheb, T., Saipanish, R., Lotrakul, M., Prasertchai, R., Ketkaew, W., Jullagate, S., Kusalaruk, P., 2015. Reliability and validity of the Thai self-report version of the Yale–Brown Obsessive–compulsive scale. Neuropsychiatr. Dis. Treat. 2817–2824. https://doi.org/10.2147/NDT.S90276.

Hollander, E., Koran, L.M., Goodman, W.K., Greist, J.H., Ninan, P.T., Yang, H., Li, D., Barbato, L.M., 2003. A double-blind, placebo-controlled study of the efficacy and safety of controlled-release fluvoxamine in patients with obsessive-compulsive disorder. J. Clin. Psychiatry 64 (6), 640–647. https://doi.org/10.4088/JCP.v64n0604.

Kraus, S.W., Potenza, M.N., Martino, S., Grant, J.E., 2015. Examining the psychometric properties of the Yale–Brown Obsessive–Compulsive Scale in a sample of compulsive pornography users. Compr. Psychiatry 59, 117–122. https://doi.org/10.1016/j.comppsych.2015.02.007.

Kuckertz, J.M., Rhee, D.W., Schreck, M., Van Kirk, N., Dattolico, D., Falkenstein, M.J., 2021. Reevaluating the psychometric properties of symptom measures within longitudinal contexts: the Yale-Brown Obsessive Compulsive Scale and Dimensional Obsessive-Compulsive Scale. Psychol. Assess. 33 (8), 756–765. https://doi.org/10.1037/pas0001017.

López-Pina, J.A., Sánchez-Meca, J., López-López, J.A., Marín-Martínez, F., Núñez-Núñez, R.M., Rosa-Alcázar, A.I., Gómez-Conesa, A., Ferrer-Requena, J., 2015. The Yale–Brown Obsessive Compulsive Scale: a reliability generalization meta-analysis. Assessment 22 (5), 619–628. https://doi.org/10.1177/1073191114551954.

McKay, D., Danyko, S.J., Neziroglu, F., Yaryura-Tobias, J.A., 1995. Factor structure of the Yale–Brown Obsessive-Compulsive Scale: a two dimensional measure. Behav. Res. Ther. 33, 865–869. https://doi.org/10.1016/0005-7967(95)00014-O.

Melli, G., Avallone, E., Moulding, R., Pinto, A., Micheli, E., Carraresi, C., 2015. Validation of the Italian version of the Yale–Brown Obsessive Compulsive Scale–Second Edition (Y-BOCS-II) in a clinical sample. Compr. Psychiatry 60, 86–92. https://doi.org/10.1016/j.comppsych.2015.03.005.

Ólafsson, R.P., Snorrason, A., Smári, J., 2010. Yale-Brown Obsessive Compulsive Scale: psychometric properties of the self-report version in a student sample. J. Psychopathol. Behav. Assess. 32 (2), 226–235. https://doi.org/10.1007/s10862-009-9146-0.

Pinciotti, C.M., Rast, C.E., Wiese, A.D., Foshee, K.L., Spencer, S.D., Upshaw, B.M., Guzick, A.G., Goodman, W.K., Storch, E.A., 2024. Psychometric properties of the Yale-Brown Obsessive Compulsive Scale – II Self-Report. J. Psychiatr. Pract. 30 (6), 421–427. https://doi.org/10.1097/PRA.0000000000000823.

Ruscio, A.M., Stein, D.J., Chiu, W.T., Kessler, R.C., 2010. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. Mol. Psychiatry 15 (1), 53–63. https://doi.org/10.1038/mp.2008.94.

Scahill, L., Riddle, M.A., McSwiggin-Hardin, M., Ort, S.I., King, R.A., Goodman, W.K., Cichetti, D., Leckman, J.F., 1997. Children's Yale-Brown obsessive compulsive scale: reliability and validity. J. Am. Acad. Child Adolesc. Psychiatry 36 (6), 844–852. https://doi.org/10.1097/00004583-199706000-00023.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. The Journal of Clinical Psychiatry 59 (Suppl. 20), 22–33.

Sheehan, D.V., Sheehan, K.H., Shytle, R.D., Janavs, J., Bannon, Y., Rogers, J.E., Milo, K. M., Stock, S.L., Wilkinson, B., 2010. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). J. Clin. Psychiatry 71 (03), 313–326. https://doi.org/10.4088/JCP.09m05305whi.

Storch, E.A., Shapira, N.A., Dimoulas, E., Geffken, G.R., Murphy, T.K., Goodman, W.K., 2005. Yale-Brown Obsessive Compulsive Scale: the dimensional structure revisited. Depress. Anxiety 22 (1), 28–35. https://doi.org/10.1002/da.20088.

Storch, E.A., Rasmussen, S.A., Price, L.H., Larson, M.J., Murphy, T.K., Goodman, W.K., 2010. Development and psychometric evaluation of the Yale–Brown Obsessive-Compulsive Scale—Second Edition. Psychol. Assess. 22 (2), 223–232. https://doi.org/10.1037/a0018492.

Storch, E.A., Nadeau, J.M., De Nadai, A.S., Cepeda, S.L., Riemann, B.C., Seibell, P., Kay, B., 2017. Symptom correspondence between clinicians and patients on the Yale–Brown Obsessive Compulsive Scale. Compr. Psychiatry 73, 105–110. https://doi.org/10.1016/j.comppsych.2016.11.011.

Storch, E.A., McGuire, J.F., Wu, M.S., Hamblin, R., McIngvale, E., Cepeda, S.L., Schneider, S.C., Rufino, K.A., Rasmussen, S.A., Price, L.H., Goodman, W.K., 2019. Development and psychometric evaluation of the children's Yale-Brown Obsessive-Compulsive Scale Second Edition. J. Am. Acad. Child Adolesc. Psychiatry 58 (1), 92–98. https://doi.org/10.1016/j.jaac.2018.05.029.

Storch, E.A., Cheng, J.S., Higham, M., Muñoz, J.S., Zavala Cruz, V., Berrones, D., Obando, D.P., Anderberg, J.L., Frederick, R.M., Tapia-Cuevas, V., Churruca Muñoz, M., Uribe Villar, C., Moya, P.R., Ochoa-Panaifo, M.E., Martinez Mallen, M.C., Wiese, A.D., Pinciotti, C.M., Sagarnaga, M.N., McGuire, J.F., Onyeka, O.C., Moyano, M.B., Latin American Trans-Ancestry INitiative for OCD genomics (LATINO), Brazilian Obsessive-Compulsive Spectrum Disorder Working Group (GTTOC), Goodman, W.K., Crowley, J.J., 2025. Psychometric properties of the Spanish Yale-Brown Obsessive-Compulsive Scale – Second Edition. Psychiatry Res. 348, 116456. https://doi.org/10.1016/j.psychres.2025.116456.

Tek, C., Uluğ, B., Rezaki, B.G., Tanriverdi, N., Mercan, S., Demir, B., Vargel, S., 1995. Yale-Brown Obsessive Compulsive Scale and US National Institute of Mental Health Global Obsessive Compulsive Scale in Turkish: reliability and validity. Acta

Psychiatr. Scand. 91 (6), 410–413. https://doi.org/10.1111/j.1600-0447.1995.tb09801.x.

Tolin, D.F., Abramowitz, J.S., Diefenbach, G.J., 2005. Defining response in clinical trials for obsessive-compulsive disorder: a signal detection analysis of the Yale-Brown Obsessive Compulsive Scale. J. Clin. Psychiatry 66 (12), 1549–1557. https://doi.org/10.4088/JCP.v66n1209.

Visser, H.A., Van Oppen, P., Van Megen, H.J., Eikelenboom, M., Van Balkom, A.J., 2014. Obsessive-compulsive disorder; chronic versus non-chronic symptoms. J. Affect. Disord. 152, 169–174.

Vogt, G.S., Avendaño-Ortega, M., Schneider, S.C., Goodman, W.K., Storch, E.A., 2022. Optimizing obsessive-compulsive symptom measurement with the Yale-Brown Obsessive-Compulsive Scales—Second Edition. J. Psychiatr. Pract. 28 (4), 294–309. https://doi.org/10.1097/PRA.0000000000000640.

Woody, S.R., Steketee, G., Chambless, D.L., 1995. Reliability and validity of the Yale-Brown obsessive-compulsive scale. Behav. Res. Ther. 33 (5), 597–605. https://doi.org/10.1016/0005-7967(94)00076-V.

World Health Organization, 2019. International Statistical Classification of Diseases and Related Health Problems, 11th ed. https://icd.who.int/.

Wu, K.D., Watson, D., Clark, L.A., 2007. A self-report version of the Yale–Brown Obsessive–Compulsive Scale Symptom Checklist: psychometric properties of factor-based scales in three samples. J. Anxiety Disord. 21 (5), 644–661. https://doi.org/10.1016/j.janxdis.2006.10.003.

Wu, M.S., McGuire, J.F., Horng, B., Storch, E.A., 2016. Further psychometric properties of the Yale–Brown Obsessive Compulsive Scale — Second Edition. Compr. Psychiatry 66, 96–103. https://doi.org/10.1016/j.comppsych.2016.01.007.

Zhang, C.C., Gong, H., Zhang, Y., Jin, H., Yang, Y., Li, B., Li, Y., Luo, X., Liu, W., Fang, F., Li, B., Sun, B., Fletcher, T., La Buissonnière-Ariza, V., Goodman, W.K., Storch, E.A., 2019. Development and psychometric evaluation of the Mandarin Chinese version of the Yale-Brown Obsessive-Compulsive Scale – Second Edition. Brazilian Journal of Psychiatry 41 (6), 494–498. https://doi.org/10.1590/1516-4446-2018-0126.