

# Longitudinal diagnostic efficiency of DSM-IV criteria for obsessive–compulsive personality disorder: a 2-year prospective study

Grilo CM, Skodol AE, Gunderson JG, Sanislow CA, Stout RL, Shea MT, Morey LC, Zanarini MC, Bender DS, Yen S, McGlashan TH. Longitudinal diagnostic efficiency of DSM-IV criteria for obsessive–compulsive personality disorder: a 2-year prospective study. Acta Psychiatr Scand 2004; 110: 64–68. © Blackwell Munksgaard 2004.

**Objective:** To examine the *longitudinal* diagnostic efficiency of the DSM-IV criteria for obsessive–compulsive personality disorder (OCPD).

**Method:** At baseline, criteria and diagnoses were determined using diagnostic interviews, and blinded assessments were performed 24 months later with 550 participants. Diagnostic efficiency indices (conditional probabilities, total predictive power, and kappa) were calculated for each criterion determined at baseline, using the independent OCPD diagnosis at follow-up as the standard.

**Results:** *Longitudinal* diagnostic efficiencies for the OCPD criteria varied; findings suggested the overall predictive utility of ‘preoccupied with details’, ‘rigid and stubborn’, and ‘reluctant to delegate’.

**Conclusion:** These findings suggest the predictive validity of three cognitive-interpersonal OCPD criteria.

C. M. Grilo<sup>1</sup>, A. E. Skodol<sup>2</sup>,  
J. G. Gunderson<sup>3</sup>, C. A. Sanislow<sup>1</sup>,  
R. L. Stout<sup>4</sup>, M. T. Shea<sup>4</sup>,  
L. C. Morey<sup>5</sup>, M. C. Zanarini<sup>3</sup>,  
D. S. Bender<sup>2</sup>, S. Yen<sup>4</sup>,  
T. H. McGlashan<sup>1</sup>

<sup>1</sup>Department of Psychiatry, Yale University School of Medicine, New Haven, CT, <sup>2</sup>New York State Psychiatric Institute, Columbia University College of Physicians and Surgeons, New York, NY, <sup>3</sup>McLean Hospital, Harvard Medical School, Belmont, MA, <sup>4</sup>Department of Psychiatry and Human Behavior, Brown University, Providence, RI, <sup>5</sup>Department of Psychology, Texas A & M University, TX, USA

Key words: obsessive–compulsive personality disorder; personality disorders; diagnostic efficiency; assessment; longitudinal follow-up

Carlos M. Grilo, Yale Psychiatric Research, Yale University School of Medicine, 301 Cedar Street, PO Box 208098, New Haven, CT 06520, USA.  
E-mail: carlos.grilo@yale.edu

Accepted for publication January 15, 2004

## Introduction

Obsessive–compulsive personality disorder (OCPD) is prevalent in community (1) and clinical samples (2). The DSM-IV (3) OCPD diagnostic construct evolved considerably from earlier editions of the DSM and from the clinical literature (4). While ‘pre-DSM-III-R’ OCPD received considerable attention [see (5)], the past 20 years have witnessed little empirical attention paid to OCPD (4, 6).

Empirical support for the disorder-level classification of OCPD in the DSM-IV was recently reported (7). Additional psychometric work on the structure and validity of OCPD represents an important avenue for the continued refinement of the diagnostic construct (8, 9). Grilo (10) reported good internal consistency for OCPD criteria and,

using factor analysis, identified potential components of the construct. Another relevant approach is the study of diagnostic efficiency, which refers to the extent to which criteria are able to discriminate individuals with a specific diagnosis from those without that diagnosis, by the application of conditional probabilities. Such analyses, in addition to informing the refinement of nosological systems (11), have direct relevance to clinicians as guides to their diagnostic decision-making practices.

Such analyses have contributed to the refinement of some psychiatric diagnoses (12) and have been recently applied to selected personality disorders (13, 14). As part of the DSM-IV Work Group, Pfohl and Blum (4) noted – in relation to OCPD – ‘only limited data are available for DSM-III-R (pp. 267)’, and reviewed the relevant data from unpublished

data sets, only some of which utilized structured interviews. Noting variability in the performance of the DSM-III-R OCPD criteria within and across the four unpublished data sets, the Work Group (4) highlighted the need for additional research. Since then, for OCPD, only three studies of diagnostic efficiency based on semi-structured diagnostic interviews have been published, all of which examined DSM-IV criteria (15–17).

In terms of DSM-IV studies, two found that some of the OCPD criteria failed to have positive predictive power (PPP) greater than 50/50 chance (15, 16). Farmer & Chapman (15) found that five criteria had PPP less than 50/50 chance in a heterogeneous community ( $n = 149$ ) sample of respondents to ads. Grilo (16) reported similar findings based on an analysis of a homogeneous treatment-seeking patient group ( $n = 211$ ) with binge eating disorder. Specifically, Grilo (16) found five OCPD criteria performed either poorly (pack-rat, miserly and rigid and stubborn) or marginally (inflexible about morality and ethics and reluctant to delegate) as inclusion criteria. Grilo (16) reported that ‘perfectionism’ and ‘preoccupied with details’ performed best as inclusion criteria and had the best total predictive values (TPV). In contrast, are findings from our initial report (17) based on a large ( $n = 668$ ) heterogeneous patient group. Grilo *et al.* (17) reported that all OCPD criteria at baseline performed better than 50/50 chance for predicting the OCPD diagnosis, with a mean PPP of 0.73 (ranging from 0.62 to 0.83).

#### Aim of the study

The aim of this study was to examine prospectively the *longitudinal* diagnostic efficiency of the DSM-IV criteria for OCPD using semi-structured diagnostic interviews administered 2-years apart. We aimed to determine the efficiency of each criterion for predicting the presence and/or absence of the OCPD diagnosis 2 years later.

## Material and methods

### Subjects

Subjects were 550 participants from the Collaborative Longitudinal Personality Disorders Study (CLPS) (18, 19), an ongoing prospective repeated-measures study designed to examine the natural course and outcome of patients with PDs compared with patients with major depressive disorder without any PD. Most participants were patients recruited from diverse in- and out-patient clinical programs at each of the four CLPS recruitment sites (Brown,

Columbia, Harvard, and Yale Universities) with the intent to obtain a diverse and clinically representative study group. Advertising was also used to recruit participants with current or past psychiatric treatment. The current report is based on 550 of the overall ( $n = 688$ ) CLPS study group (18, 19), for whom complete data through the 24-month follow-up assessment were available at the time of analysis. Participants in this study for whom follow-up data were available did not differ from those not included in the current analyses in age, gender, ethnicity, or PD group.

At baseline, 222 (40%) of the 550 participants met criteria for OCPD, while at the 24-month follow-up, 120 (22%) of the 550 participants received the OCPD diagnosis. Of the 550 participants, 352 (64%) were female and 198 (36%) were male. The majority of participants were Caucasian ( $n = 419$ , 76%) and 131 (24%) were non-Caucasian. The mean age of participants was 32.9 (SD = 8.3) years. The mean Global Assessment of Functioning (GAF) at baseline for the past month was 51.5 (SD = 9.7). Co-occurring Axis I and Axis II disorders were common. In the present study, at baseline participants met criteria for a mean of 2.7 (SD = 1.6) axis I psychiatric disorders and for a mean of 3.7 (SD = 2.2) axis II PDs.

### Procedures and assessments

All participants signed written informed consents following a description of study procedures. The study protocol was approved by each collaborating site’s Institutional Review Board. Participants were interviewed in person by experienced interviewers. The research interviewers underwent extensive standardized training to achieve reliability in the use of the diagnostic measures (20). Interviewers were monitored and received regular on-going supervision by the investigators at each site, as well as regular supervision across sites to maintain reliability and prevent drift over time.

The Diagnostic Interview for DSM-IV Personality Disorders (DIPD-IV) (21), a semi-structured interview, was administered to assess PDs. Each of the criteria for all disorders is assessed with one or more questions, which are then rated on a three-point scale (0, not present; 1, present but of uncertain clinical significance; 2, present and clinically significant). The time frame covered is the prior 2 years, but traits or behaviors must be characteristic of the person for most of his or her adult life in order to be counted toward a diagnosis.

*Reliability.* Inter-rater reliability (based on 84 pairs of raters using taped interviews) kappa coefficients

ranged from 0.58 to 1.00 (median = 0.68) for all PDs; kappa was 0.71 for OCPD. Test–retest reliability (based on 52 cases interviewed independently twice a week apart) ranged from kappa = 0.69 to 0.74 (for OCPD). Kappa coefficients for inter-rater and for test–retest reliability, respectively, for the specific OCPD criteria were: 0.57/0.59 (preoccupied with details), 0.64/0.64 (perfectionism), 0.61/0.54 (workaholic), 0.69/0.46 (inflexible morality and ethics), 0.65/0.80 (packrat), 0.60/0.58 (reluctant to delegate), 0.44/0.25 (miserly), and 0.79/0.43 (rigid and stubborn).

*Follow-up evaluations.* Participants were re-interviewed at 6, 12, and 24 months following the baseline assessment. The 24-month assessment included a separate, independent administration of the DIPD-IV to assess all PDs. This DIPD-IV was given by an interviewer who was blind to baseline and all interim assessments.

Statistical analysis

*Longitudinal* diagnostic efficiency of the OCPD criteria at baseline entry into the study was considered against the OCPD diagnosis obtained at 2-year follow-up. Efficiency of the criteria (ratings of ‘2’ on the DIPD-IV) was examined using four types of conditional probabilities. *Sensitivity* (true-positive rate) is the proportion of patients with the diagnosis who have the criterion – or the probability of having the criterion, given that one has the diagnosis. *Specificity* (true-negative rate) is the probability of not having the criterion, given that one does not have the diagnosis. *PPP* is the probability of having the diagnosis, given that one has the criterion. Lastly, *negative predictive power* (NPP) is the probability that one does not have the diagnosis, given that one does not have the criterion. In addition, *TPV*, a measure of percent agreement (not a conditional probability) was calculated as an overall indicator of a criterion’s utility in making a correct diagnosis. Kappa coefficient, which corrects for chance agreement, was also calculated.

Positive predictive power indicates whether a criterion will have utility as an inclusion criterion, and the relative values of the PPPs for various criteria of a diagnosis can provide information about which are the best predictors of the disorder. All else being equal, PPP increases while NPP decreases with increasing diagnosis base rates (22). TPV has an advantage over PPP and NPP since it relates to both inclusion and exclusion. While kappa provides similar information as TPV, it has

the important advantage in that its scale does not inflate a criterion’s prediction of diagnosis.

Results

Table 1 shows the base rates of the eight OCPD criteria at baseline. ‘Reluctant to delegate’ and ‘rigid and stubborn’ were the most frequent of the OCPD criteria, both occurring in 47% of all participants, followed by ‘perfectionism,’ which occurred in 43% of participants. The least frequent criterion was ‘miserly,’ which occurred in only 15% of all participants.

For each criterion, the longitudinal diagnostic efficiency indices for predicting the diagnosis of OCPD at the 24-month blinded DIPD-IV assessment are also shown in Table 1. For predicting OCPD at 24-months, given the base-rate of 0.22, all eight criteria performed better than ‘chance’ given the base-rate of the diagnosis, suggesting that they do possess some degree of predictive utility. In this regard, ‘preoccupied with details’ had the highest PPP (0.41). Taking both PPP and NPP into account, the TPVs indicate that ‘preoccupied with details’ and ‘miserly’ had the most overall utility in correctly diagnosing OCPD. Here, it is important to note that the kappa findings, which take into account chance, suggest the overall utility of ‘preoccupied with details’ (kappa = 0.29) but not ‘miserly’ (kappa = 0.12), and also suggest the utility of two other criteria (rigid and stubborn and reluctant to delegate).

A logistic regression analysis (stepwise forward method) was used to examine the joint contributions of the OCPD criteria at baseline for predicting the OCPD diagnosis at 24-month follow-up. The overall model was significant ( $\chi^2 = 120.67$ ,  $P < 0.0001$ ) and revealed five criteria as making

Table 1. Base rates and diagnostic efficiency indices for DSM-IV criteria for OCPD in 550 Patients\*

OCPD criterion	BR <sub>s</sub>	Sensitivity	Specificity	PPP	NPP	TPV	Kappa
Preoccupied with details	0.30	0.55	0.78	0.41	0.86	0.73	0.29
Perfectionism	0.43	0.68	0.64	0.34	0.88	0.65	0.23
Workaholic	0.27	0.42	0.77	0.34	0.83	0.69	0.17
Inflexible morality and ethics	0.31	0.53	0.75	0.37	0.85	0.70	0.24
Packrat	0.40	0.58	0.65	0.32	0.85	0.64	0.18
Reluctant to delegate	0.47	0.78	0.61	0.36	0.91	0.65	0.28
Miserly	0.15	0.23	0.88	0.34	0.80	0.73	0.12
Rigid and stubborn	0.47	0.78	0.62	0.37	0.91	0.66	0.29
Mean	0.35	0.57	0.71	0.36	0.86	0.68	0.23
SD	0.11	0.19	0.10	0.03	0.04	0.04	0.06

\*Obsessive–compulsive personality disorder (OCPD) diagnosis base rate at 2-year follow-up = 0.22. BR<sub>s</sub>, criteria base rate at baseline entry; PPP, positive predictive power; NPP, negative predictive power; TPV, total predictive value.

Table 2. Summary of logistic regression findings for criteria at baseline predicting OCPD diagnosis at 2-year follow-up in 550 patients\*

OCPD criterion	Odds ratio	95% confidence interval	P-value
Rigid and stubborn	3.01	1.77–5.14	<0.0001
Preoccupied with details	2.65	1.66–4.23	<0.0001
Reluctant to delegate	2.41	1.39–4.18	<0.002
Inflexibility about morality	1.84	1.14–2.96	<0.02
Packrat	1.77	1.11–2.82	<0.02

\*Obsessive-compulsive personality disorder (OCPD) diagnosis base rate at 2-year follow-up = 0.22.

Only criteria with significant effects are shown.

significant independent contributions (summarized in Table 2).

**Discussion**

This study – using data obtained with reliably administered, semi-structured research interviews – examined the *longitudinal* diagnostic efficiency of the DSM-IV criteria for OCPD for *prospectively* predicting the OCPD diagnosis at 24-month follow-up. The best criterion for predicting inclusion in OCPD assessed 2 years later (highest PPP) was ‘preoccupied with details’, which was also, overall, the most predictive criterion (TPV). Kappa coefficients, which take into account chance, suggest the overall utility of three criteria for predicting the diagnosis of OCPD 2 years later: ‘preoccupied with details’, ‘rigid and stubborn’, and ‘reluctant to delegate’. A logistic regression analysis predicting the OCPD diagnosis revealed that these three criteria made significant contributions and showed odds ratios greater than 2.0.

How to best operationalize the constructs of disorders of personality remains a topic of clinical and research importance (6). The DSM-IV criteria for OCPD appear to reflect behavioral, cognitive, and interpersonal domains. Our findings here could be taken to suggest that OCPD be defined primarily by psychological and interpersonal behaviors. The three best performing criteria are cognitive or interpersonal in nature whereas the three least efficient criteria reflect behaviors considered socially excessive. In this regard, we note some convergence across studies. Our cross-sectional analysis at baseline (17) and a similar analysis with a different patient group (16) revealed that ‘preoccupied with details’ performed best as an inclusion criterion for OCPD.

To provide a context for our findings, we note several issues to consider. The diagnostic interviews did not use randomized assessment of the PD criteria; thus a ‘halo’ effect (interviewers might rate criteria similarly within a given diagnosis) is

possible, which could artificially inflate the degree to which the criteria appear to be inter-related and thus make them appear to be more similar in their efficiency. On the other hand, this assessment method of determining criteria grouped by disorder is more clinically ecological because criteria are evaluated within the context of the disorder. Assessments were administered by reliably trained interviewers who were monitored to prevent drift. Importantly, the variability in reliability for the specific OCPD criteria at baseline was not related to the patterning of diagnostic efficiency indices. It is also important to keep in mind that variability in diagnostic base rates across patient groups can influence PPP/NPP (22). Generally, one can expect higher PPP and lower NPP values in study groups with higher base rates of the diagnosis in question. Our OCPD diagnosis base rate (0.22) at 2-year follow-up is only slightly higher than the rate of 0.15 reported by Grilo (16) for a different patient group, and is similar to that reported (2) for other clinical samples (0.16) and falls near the middle of the range for the unpublished DSM-III-R data sets reviewed by the DSM-IV Work Group (4).

The study recruited patients and treatment-seeking participants given our focus on clinical entities. We acknowledge that there exist inherent complexities in studying treatment-seeking patients and that our findings may not generalize to community or non-treatment seeking groups. The pattern of diagnoses were comparable with those reported for other clinical samples [see (19)] suggesting its relevance for psychiatric patient groups. Furthermore, our sampling across diverse clinical settings affiliated with universities in four northeastern urban settings produced a heterogeneous adult study group that approximates US norms for ethnicity. Since we targeted adults aged 18–45 years, our findings may not generalize to different age groups (23).

Our findings here suggest the particular relevance of the presence or absence of three criteria (preoccupied with details, rigid and stubborn, and reluctant to delegate). Of course, we emphasize that deriving diagnoses based only on partial criteria or information is not advisable and, indeed, may prove to be erroneous (24). Instead, we offer these findings with the hope that they inform future diagnostic schemes (9).

Three criteria showed significant overall efficiency for predicting the presence of OCPD at blind follow-up assessments performed 2-years later. It is noteworthy that the base rate of OCPD at the 24-month follow-up was 22%, which represents a substantial decline from the 40% base rate at baseline; the 24-month course and

stability PD findings are detailed elsewhere (25). Thus, the *longitudinal* diagnostic efficiency findings presented here address the important clinical question of which OCPD criteria (which had undergone substantial revisions from earlier editions of the DSM) are most predictive of primarily stable cases of OCPD. In essence, three cognitive-interpersonal criteria, which demonstrate predictive validity for the diagnosis 2 years later, may represent the most central or core aspect of the OCPD. Similar analyses with criteria and dimensional measures with additional and varied patient groups (26, 27) need to be performed to provide complementary data.

### Acknowledgements

Supported by NIH grants MH 50837, 50838, 50839, 50840, 50850, and K05 MH 01645.

This publication has been reviewed and approved by the Publications Committee of the Collaborative Longitudinal Personality Disorders Study.

### References

1. TORGERSEN S, KRINGLEN E, CRAMER V. The prevalence of personality disorders in a community sample. *Arch Gen Psychiatry* 2001;**58**:590–596.
2. STUART S, PFOHL B, BATAGLIA M, BELLODI L, GROVE W, CADORET R. The co-occurrence of DSM-III-R personality disorders. *J Pers Dis* 1998;**12**:302–315.
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn (DSM-IV). Washington, DC: APA, 1994.
4. PFOHL B, BLUM N. Obsessive-compulsive personality disorder. In: LIVESLEY WJ, ed. *The DSM-IV personality disorders*. New York: Guilford Press, 1995.
5. POLLAK JM. Obsessive-compulsive personality: a review. *Psychol Bull* 1979;**86**:225–241.
6. GRILO CM, McGLASHAN TH. Stability and course of personality disorders. *Curr Opin Psychiatry* 1999;**12**:157–162.
7. SANISLOW CA, MOREY LC, GRILO CM et al. Confirmatory factor analysis of DSM-IV borderline, schizotypal, avoidant, and obsessive-compulsive personality disorders: findings from the Collaborative Longitudinal Personality Disorders Study. *Acta Psychiatr Scand* 2002;**105**:28–36.
8. BLASHFIELD R, DRUGUNS J. Evaluative criteria for psychiatric classification. *J Abn Psychol* 1976;**85**:140–150.
9. BERTELSEN A. Wanted: validation studies on the current diagnostic classifications. *Acta Psychiatr Scand* 2002; **106**: 81–82.
10. GRILO CM. Factor structure of DSM-IV criteria for obsessive compulsive personality disorder in patients with binge eating disorder. *Acta Psychiatr Scand* 2004;**109**:64–69
11. BALDESSARINI RJ, FINKELSTEIN S, ARANA GW. The predictive power of diagnostic tests and the effect of prevalence of illness. *Arch Gen Psychiatry* 1983;**40**:569–573.
12. WALDMAN ID, LILIENFELD SO. Diagnostic efficiency of symptoms for oppositional defiant disorder and attention-deficit hyperactivity disorder. *J Consult Clin Psychol* 1991;**59**:732–738.
13. BECKER DF, GRILO CM, EDELL WS, McGLASHAN TH. Diagnostic efficiency of borderline personality disorder criteria in hospitalized adolescents: comparison with hospitalized adults. *Am J Psychiatry* 2002;**159**:2042–2047.
14. GRILO CM, BECKER DF, ANEZ LM, McGLASHAN TH. Diagnostic efficiency of DSM-IV Criteria for borderline personality disorder: an evaluation in Hispanic men and women with substance use disorders. *J Consult Clin Psychol* 2004;**72**:126–131.
15. FARMER RF, CHAPMAN AL. Evaluation of DSM-IV personality disorder criteria as assessed by the Structured Clinical Interview for DSM-IV personality disorders. *Compr Psychiatry* 2002;**43**:285–300.
16. GRILO CM. Diagnostic efficiency of DSM-IV criteria for obsessive compulsive personality disorder in patients with binge eating disorder. *Behav Res Ther* 2004;**35**:80–85.
17. GRILO CM, McGLASHAN TH, MOREY LC et al. Internal consistency, intercriterion overlap, and diagnostic efficiency of criteria sets for DSM-IV personality disorders. *Acta Psychiatr Scand* 2001;**104**:264–272.
18. GUNDERSON JG, SHEA MT, SKODOL AE et al. The Collaborative Longitudinal Personality Disorders Study: development, aims, design, and sample characteristics. *J Personal Disord* 2000;**14**:300–315.
19. McGLASHAN TH, GRILO CM, SKODOL AE et al. The Collaborative Longitudinal Personality Disorders Study: baseline Axis I/II and II/II diagnostic co-occurrence. *Acta Psychiatr Scand* 2000;**102**:256–264.
20. ZANARINI MC, SKODOL AE, BENDER D et al. The collaborative longitudinal personality disorders study: reliability of axis I and axis II diagnoses. *J Personal Disord* 2000;**14**:291–299.
21. ZANARINI MC, FRANKENBURG FR, SICKEL AE, YONG L. *The Diagnostic Interview for DSM-IV Personality Disorders*. Belmont, MA: McLean Hospital and Harvard Medical School, 1996.
22. MEEHL PE, ROSEN A. Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychol Bull* 1955;**52**:194–216.
23. GRILO CM, BECKER DF, EDELL WS, McGLASHAN TH. Stability and change of DSM-III-R personality disorder dimensions in adolescents followed up 2 years after psychiatric hospitalization. *Compr Psychiatry* 2001;**42**:364–368.
24. MOREY LC, OCHOA ES. An investigation of adherence to diagnostic criteria: clinical diagnosis of the DSM-III personality disorders. *J Personal Disord* 1989;**3**:180–192.
25. GRILO CM, SHEA MT, SANISLOW CA et al. Two-year stability and change in schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *J Consult Clin Psychol* (in press).
26. IKETANI T, KIRIKE N, STEIN MB et al. Relationship between perfectionism, personality disorders, and agoraphobia in patients with panic disorder. *Acta Psychiatr Scand* 2002; **106**:171–178.
27. SVRAKIC DM, DRAGANIC S, HILL K, BAYON C, PRZYBECK TR, CLONINGER CR. Temperament, character, and personality disorders: etiologic, diagnostic, treatment issues. *Acta Psychiatr Scand* 2002;**106**:189–195.