EVALUATION OF DIAGNOSTIC THRESHOLDS FOR CRITERION A IN THE ALTERNATIVE *DSM-5* MODEL FOR PERSONALITY DISORDERS

Tore Buer Christensen, MD, Benjamin Hummelen, PhD, MD,
Muirne C. S. Paap, PhD, Ingeborg Eikenaes, MD,
Sara Germans Selvik, PhD, Elfrida Kvarstein, PhD, MD,
Geir Pedersen, PhD, Donna S. Bender, PhD,
Andrew E. Skodol, MD, and Tor Erik Nysæter, PhD

The Level of Personality Functioning Scale (LPFS) of the Alternative *DSM-5* Model for Personality Disorders (AMPD) was formulated to assess the presence and severity of personality disorders (PDs). Moderate impairment (Level 2) in personality functioning, as measured by the LPFS, was incorporated into the AMPD as a diagnostic threshold for PD in Criterion A of the general criteria, as well as for the "any two areas present" rule for assigning a specific PD diagnosis. This study represents the first evaluation of the diagnostic decision rules for Criterion A, in a clinical sample (N = 282). The results indicate that an overall diagnostic threshold for PDs should be used with caution because it may not identify all *DSM-IV* PDs. The "any two areas present" rule proved to be a reasonable alternative, although this finding should be interpreted with caution because the LPFS does not measure the disorder-specific A criteria.

Keywords: personality disorder, *DSM-5*, Alternative Model for Personality Disorders, AMPD, Level of Personality Functioning Scale, LPFS, diagnostic threshold, Criterion A, SCID-5-AMPD, psychometrics

From Department of Mental Health, Sorlandet Hospital, Arendal, Norway (T. B. C., T. E. N.); Department of Personality Psychiatry, Clinic Mental Health and Addiction, Oslo University Hospital, Oslo, Norway (B. H., E. K. G. P.); Department of Research and Development, Clinic Mental Health and Addiction, Oslo University Hospital, Oslo, Norway (B. H., M. C. S. P.); Department of Special Needs, Education, and Youth Care, Faculty of Behavioral and Social Sciences, University of Groningen, Groningen, The Netherlands (M. C. S. P.); Department of Personality Psychiatry, Division of Mental Health and Addiction Treatment, Vestfold Hospital Trust, Tonsberg, Norway (I. E.); Department of Psychiatry, Hospital Namsos, Namsos, Norway (S. G. S.); Department of Mental Health , Norwegian University of Science and Technology (NTNU), Trondheim, Norway (S. G. S.); NORMENT, KG Jebsen Center for Psychosis Research, Institute of Clinical Medicine, University of Oslo, Oslo, Norway (G. P.); Department of Psychiatry and Behavioral Sciences and Counseling and Psychological Services, Tulane University, New Orleans, Louisiana (D. S. B.); and Department of Psychiatry, University of Arizona College of Medicine, Tucson, Arizona (A. E. S.).

Address correspondence to Tore Buer Christensen, Sørlandet Sykehus, Psyk. Avd Arendal Pb 783 Stoa, 4809 Arendal, Norway. E-mail: tore.buer.christensen@sshf.no

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*; American Psychiatric Association [APA], 2013) contains two diagnostic models for personality disorders (PDs). The first is the "Section II Model," a purely categorical model that is identical in format and content to the *DSM-IV* model of PDs (APA, 1994). The second is the "Section III Model" or Alternative *DSM-5* Model for Personality Disorders (AMPD), a hybrid comprising both dimensional and categorical aspects. The AMPD (Table 1) has defined PD according to impairment in personality functioning (Criterion A) as a general diagnostic criterion for PD, together with the presence of at least one pathological personality trait (Criterion B), and other relevant inclusion and exclusion criteria (i.e., Criteria C through G). Criterion A identifies the presence of PD and quantifies its severity, rated on a five-level continuum from little or no impairment (Level 0) to extreme impairment (Level 4). This continuum is the "Level of Personality Functioning Scale" (LPFS; Bender, Morey, & Skodol, 2011). The LPFS is intended to capture core features of personality pathology, manifested in impairments in self- (Identity and Self-direction) and interpersonal (Empathy and Intimacy) functioning. The four elements or "domains" of the LPFS can be further divided into subdomains based on the definitions of the elements in *DSM-5* (APA, 2013, p. 762), three for each domain, resulting in a total of 12 subdomains. For example, the LPFS domain of Identity comprises the subdomains of sense of self, self-esteem, and emotional range and regulation. The Empathy domain comprises comprehension and appreciation of others' experiences, tolerance of differing perspectives, and understanding the effects of one's own behavior on others.

Maintaining continuity between the *DSM-IV* and the AMPD through empirically based diagnostic guidelines was an important focus in the development of the AMPD (Skodol, 2014). This is in contrast to the more radical changes proposed for the ICD-11 (Bach & First, 2018; Reed, 2018; Tyrer, Mulder, Kim, & Crawford, 2019). Although the AMPD is not necessarily expected to identify all patients with a *DSM-IV* PD diagnosis, the AMPD has preserved continuity with the *DSM-IV* by retaining six specific *DSM-IV*–type hybrid diagnoses. Continuity was viewed as especially important for diagnoses with the largest bodies of research, that is, borderline PD (BPD), antisocial PD (ASPD), and schizotypal PD (STPD) (Blashfield & Intoccia, 2000; Morey & Skodol, 2013). Other *DSM-IV* PD diagnoses, such as avoidant, obsessive–compulsive, and narcissistic PDs, were also retained in the AMPD because of their clinical importance. For example, avoidant PD (AVPD), although receiving less research attention relative to other PDs, is among the most prevalent in community and clinical populations, and is associated with substantial functional impairment (Lampe & Malhi, 2018). To differentiate among these diagnoses, each PD is defined by typical impairment in personality functioning and characteristic personality traits. This assessment is the third step in the diagnostic process of the AMPD (Skodol, Morey, Bender, & Oldham, 2015), illustrated in Table 1. For Criterion A, these disorder-specific criteria—inspired by the *DSM-IV*—are intended to describe impairments characteristic for each PD, one description for each domain. Of note, these disorder-specific descriptions are not identical to the descriptions of impairments in the LPFS, which is

TABLE 1. The Stepwise Diagnostic Approach for Personality Disorders According to the Alternative *DSM-5* Model for Personality Disorders (AMPD)

	Examples of assessment instruments*
Step 1: **Assessment of levels of personality functioning (Criterion A)**	SCID-5-AMPD Module I, STiP 5.1, LPFS-SR, LPFS-BF 2.0
Identity and Self-direction domains Empathy and Intimacy domains (Self) (Interpersonal)	
Threshold: moderate or greater overall impairment	
Step 2: **Assessment of pathological personality traits (Criterion B)**	PID-5, SCID-5-AMPD Module II,
Negative affectivity, detachment, antagonism, disinhibition, psychoticism	
Threshold: one or more pathological trait is moderately or very descriptive	
Step 3a **Apply the specific diagnostic Criteria A and B for the six retained personality disorders:**	SCID-5-AMPD Module III
Antisocial, avoidant, borderline, narcissistic, obsessive-compulsive, schizotypal	
Threshold: moderate or greater impairment in two or more disorder-specific domains of personality functioning plus a disorder-specific number and configuration of pathological traits	
Step 3b: If specific criteria not fulfilled but still moderate or greater impairment in personality functioning (criterion A) and one or more pathological personality traits (criterion B):	SCID-5-AMPD Module III
Apply Criteria A and B for personality disorder–trait specified	
Threshold: moderate or greater impairment in any two or more disorder-specific domains of personality functioning plus at least one pathological trait is moderately or very descriptive	
Step 4: **Apply the other general criteria C–G:** Inflexibility and pervasiveness (C); stability and early onset (D); other mental disorder (E) , substance and medical exclusions (F); age and cultural exclusions (G)	SCID-5-AMPD Module III

Note. Adapted with permission from "The Alternative *DSM-5* Model for Personality Disorders: A clinical application," by A. E. Skodol, L. C. Morey, D. S. Bender, and J. M. Oldham, 2015, *American Journal of Psychiatry, 172*, pp. 606-613. Copyright 2015 by the American Psychiatric Association.

*References: SCID-5-AMPD Module I (Bender et al., 2018); STiP5.1 (Hutsebaut et al., 2017); LPFS-SR (Morey, 2017); LPFS-BF 2.0 (Bach & Hutsebaut, 2018); PID-5 (Krueger, Derringer, Markon, Watson, & Skodol, 2012); SCID-5-AMPD Module II (Skodol, First, Bender, & Oldham, 2018); and SCID-5-AMPD Module III (First et al., 2018)

a general measure of impairment in personality functioning, but LPFS ratings are relevant for making the diagnosis of personality disorder–trait specified (PD-TS) for patients who have a PD that does not meet the criteria for a specific PD in the AMPD (Table1, Step 3b).

Influential in the development of the AMPD was the so-called "Morey survey" (Morey, Bender, & Skodol, 2013), an online survey conducted by members of the *DSM-5* Personality and Personality Disorders Work Group. Among other objectives, one aim of the survey was to establish a common diagnostic threshold for giving a PD diagnosis based on the severity of impairment in personality functioning. Here, 337 responding clinicians provided information about one of their patients through a questionnaire developed for

the study. The questionnaire included ratings of the presence of the diagnostic criteria for all 10 *DSM-IV* PD diagnoses, as well as all parts of the AMPD. Most patients (83.4%) were reported to fulfill criteria for a *DSM-IV* PD diagnosis, with BPD as the most frequent (40.1%), followed by AVPD (27%). The study concluded that moderate (Level 2) impairment in personality functioning represented the optimal combination of sensitivity (0.85) and specificity (0.73) for identifying patients receiving any specific *DSM-IV* PD diagnosis. A receiver operating characteristic (ROC) curve analysis demonstrated an area under the curve (AUC) of 0.83, indicating a high degree of accuracy for the LPFS as a diagnostic test for the presence of a *DSM-IV* PD.

Another objective of the Morey survey (Morey & Skodol, 2013) was to investigate different criteria combinations of impairments in personality functioning for assigning a specific PD diagnosis, based on the questionnaire developed for that survey, not on an interview-based assessment of the LPFS. Here, each clinician was asked to rate each of the 24 disorder-specific A criteria for the six PDs retained in the AMPD dichotomously (present or not) for the selected patient. Based on these ratings, three possible combinations of A criterion features, that is, diagnostic algorithms, were evaluated for each of the six PD diagnoses: (a) at least one criterion present indicating impairment in self-functioning and at least one for interpersonal functioning; (b) any two criteria for either self- or interpersonal functioning; and (c) the criteria fulfilled from at least any one of the four domains. For fulfilling criteria for the corresponding *DSM-IV* diagnosis, the authors concluded that the optimal combination of specificity and sensitivity was achieved when patients met disorder-specific criteria for the combination of any two of four domains, that is, the second algorithm. When this algorithm was used, sensitivity ranged from 80.0% to 95.5% for five of the six retained PD diagnoses. The exception was ASPD, with a sensitivity estimate of 65.8%. This algorithm was incorporated in the *DSM-5* manual as a common decision rule for specific PD diagnoses in the A criteria and for the diagnosis of PD-TS (Morey & Skodol, 2013). As far as we know, the diagnostic decision rules established, based on the Morey survey, have never been reevaluated.

Regarding the general or the specific A criteria in the AMPD, we are not aware of any research published based on a clinical instrument developed for the assessment of levels of impairment in personality functioning. Currently, two clinician-rated instruments for assessing all 12 subdomains of the LPFS are available: the Semi-Structured Interview for Personality Functioning *DSM-5* (STiP 5.1; Hutsebaut, Kamphuis, Feenstra, Weekers, & De Saeger, 2017), and the Structured Clinical Interview for the *DSM-5* Alternative Model for Personality Disorders, Module I (SCID-5-AMPD Module I; Bender, Skodol, First, & Oldham, 2018). For the STiP-5.1 instrument, the global LPFS rating is based on clinical judgment: If the patient is judged to be more impaired in one component than the other, the rater is instructed to use his or her judgment to set the global score (Hutsebaut et al., 2017). In the SCID-5-AMPD Module I, however, a global LPFS rating is based on the average score of the 12 subdomains. This approach requires a discussion about how average scores are translated into different levels of impairment. For example, if "Level 2" is

defined as the interval between 1.5 and 2.5, the threshold for endorsing this level will be an average score of 1.5.

THE NORWEGIAN STUDY OF THE AMPD (NORAMP)

Given the substantial impact of diagnostic thresholds in both research and clinical practice (Balsis, Lowmaster, Cooper, & Benge, 2011), more empirical support for the proposed diagnostic thresholds in the AMPD is needed. NorAMP is a multisite study that includes a clinical sample capturing a wide severity range of personality pathology. The NorAMP aims to examine the reliability, clinical utility, and validity of the AMPD, with an emphasis on the LPFS. NorAMP is the first study to examine the new SCID-5-AMPD Module I instrument (Bender et al., 2018) for assessment of the LPFS. This instrument, developed by the architects behind the LPFS and described in more detail in this article, assesses all 12 subdomains of the scale. The current study aims to evaluate the diagnostic strategies involving the LPFS. This scale is the basis of Criterion A in the AMPD, which is the first general criterion for PD required for the diagnosis of any PD, as well as for the diagnosis of PD–trait specified (PD-TS), and may provide information useful for rating the A criteria for specific PDs

AIMS

The overall aim is to provide empirical support for the relationship between Criterion A in the AMPD in *DSM-5* and *DSM-IV* PD diagnoses. More specifically, we evaluate the diagnostic accuracy of a single LPFS threshold as a test for the presence of one or more *DSM-IV* PD diagnoses. We also assess which cutoff value of a single LPFS score gives the best combination of sensitivity and specificity when gauged against these diagnoses. To elucidate the clinical consequences of a common diagnostic threshold, these metrics will be evaluated for participants within the four main diagnostic groups in our sample: BPD, AVPD, ASPD, and obsessive–compulsive PD (OCPD). On the basis of prior research (Morey et al., 2011; Wright, Hopwood, Skodol, & Morey, 2016), we expect this threshold to be higher for BPD than for the other evaluated diagnoses.

We will also examine the "any two or more areas present" rule in the *DSM-5* Section III A criteria for specific PDs, based on ratings of all subdomains within each of the four domains. This rule will be evaluated for the four of the six specific *DSM-IV* diagnoses proposed for the AMPD with a sufficient number of participants in study: BPD, AVPD, ASPD, and OCPD. Because LPFS descriptions do not correspond exactly to the specific A criteria for these PDs, and SCID-5-AMPD Module I was not designed to assess the A criteria of the specific PDs, these analyses will not be conclusive. However, in line with prior research demonstrating a strong correlation among all LPFS subdomains (Zimmermann et al., 2015), we hypothesize that our results will

provide some empirical support for the "any two or more areas present" rule in the AMPD.

METHODS

PARTICIPANTS AND RECRUITMENT SITES

Recruitment and Sites. Between March 2015 and March 2017, a total of 286 patients were recruited from different levels of psychiatric care within six hospitals in Norway. One patient was excluded because of missing diagnostic information, and three patients were excluded because of autism spectrum disorder, diagnosed after inclusion in the study. To cover the whole spectrum of personality pathology, recruitment sites included general mental health inpatient and outpatient departments, group psychotherapy outpatient and day treatment units, and one substance abuse outpatient unit that also served a prison. All group psychotherapy units were parts of the Norwegian Network of Personality-Focused Treatment Programs (Karterud & Wilberg, 2007). This large collaborative network of clinical units specializes in PD assessment and treatment. Exclusion criteria were as follows: schizophrenia spectrum disorder (except schizotypal PD), sequelae after brain injury, pervasive developmental disorders (i.e., autism spectrum disorders), intellectual disability, severe ongoing substance abuse, and lack of understanding of Norwegian.

Sample Description. Of the 282 participants finally included in our sample, 182 were female (64.5%), ranging in age from 16 to 72 years (mean = 33; SD = 10.0). One half (50.4%) of the participants were married or lived with a partner, 43.2% were employed or a full-time student, and the average level of education was 4.1 (SD = 2.8) years after 10 years of elementary school. The mean number of diagnostic criteria met according to SCID-II was 11.1 (SD = 8.1; range = 0–49). Regarding PD diagnosis, 192 (68.1%) participants fulfilled criteria for one or more diagnoses, including PD not otherwise specified (NOS) (missing data: 7 = 2.5%). For the 159 participants with one or more specific PDs, the mean number of criteria met was 14.1 (SD = 7.8), and the mean number of PD diagnoses was 1.5 (SD = 1.0). Specific PD prevalences were as follows: AVPD 42% (n = 81), BPD 36% (n = 70), PD NOS 23% (n = 45), ASPD 16% (n = 30), paranoid PD 16% (n = 30), OCPD 11% (n = 21), and dependent 7% (n = 14). Schizotypal PD, schizoid PD, histrionic PD, and narcissistic PD each was diagnosed in less than 2% of this sample. Of note, the sum of percentages is more than 100% because of co-occuring PDs. The mean number of symptom disorders was 1.7 (SD = 1.3; range 0–8),with missing data for 27 (10%). Frequencies were as follows: major depression, 27%; social phobia, 19%; posttraumatic stress disorder, 13%; substance use disorder, 12%; generalized anxiety disorder, 10%; dysthymia, 10%; and panic disorder with agoraphobia, 9%.

Referring Procedure. Referring therapists from the different sites consisted of 46.2% psychologists, 28.8% psychiatrists, and 24% with other mental health degrees, with a mean of 14 (SD = 10) years of experience. After the local di-

agnostic assessment, candidates to be included in the project were provided written information about it and the self-report questionnaires. After giving written concent, they were interviewed with the SCID-5-AMPD Module I. Referring therapists were instructed not to refer patients in an acute crisis. For more information about the participants, procedure, and clinical evaluation, see Buer Christensen et al. (2018).

MEASURES

Diagnostic Assessment. All patients were assessed for symptom disorders according to the DSM-IV using the fifth edition of the Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998), and for PDs using the Structured Clinical Interview for Axis II Disorders (SCID-II, First, Spitzer, Gibbon, Williams, & Benjamin, 1994). Reliability and validity of the MINI and SCID-II are both considered to be adequate in clinical samples (Sheehan et al., 1998; Weertman, Arntz, Dreessen, van Velzen, & Vertommen, 2003). Although most referring therapists were experienced, all received training in using the SCID-II through courses arranged by the Department of Personality Psychiatry. Referring therapists outside the Norwegian Network of Personality-Focused Treatment Programs were trained by the National Knowledge Center for Personality Disorders at Oslo University Hospital. To enhance the quality of the assessments, consensus training of all referring therapists was arranged, using video-recorded interviews. During both the initial training and the video sessions, independent ratings and discrepancies were discussed thoroughly in the group. In cases in which the referring therapists suspected no PD, the SCID-II was performed by another independent rater. Rater agreement for diagnostic assessment was not further evaluated in this study. However, Arnevik and colleagues (2009) evaluated diagnostic reliability within the network, from which 45% of the patients were recruited. Using a training procedure similar to the one used in the current study, they found reported kappa coefficients for the three PDs as follows: AVPD, 0.75; BPD, 0.66; and paranoid PD, 0.71. These values indicated acceptable diagnostic reliability.

Assessment of the LPFS. The SCID-5-AMPD Module I is a semistructured interview covering the 12 subdomains of the LPFS (Bender et al., 2018). For the NorAMP study, members of the Department of Personality Psychiatry, Oslo University Hospital, translated both the LPFS and the SCID-5-AMPD Module I into Norwegian. Back-translation was performed post hoc by a professional translator. The back translation showed excellent correspondence with the original English version.

Before the study, Dr. Donna Bender trained seven experienced raters in using the instrument. The maximum interval between the SCID-II and SCID-5-AMPD Module I interviews was 5 weeks. Raters who performed the SCID-5-AMPD Module I were blinded to the SCID-II results.

The Module I interview starts with eight general overview questions addressing how the participant relates to himself or herself and others. The design of the interview has a "funnel structure," in that the rater is instructed to begin asking questions for each subdomain corresponding to the interviewee's

level of function as estimated from his or her responses to the overview questions. Thus, a patient who describes confusion about describing himself or herself, reports chronic low self-esteem, or exhibits poor personal boundaries at the interview's outset would not be asked questions corresponding to little or no impairment or some impairment in the Identity domain. Increasing levels of impairment are explored through one to six questions until the interviewee clearly does not qualify for a higher level of severity. The interview also included a few questions concerning demographics and former psychopathology. In addition, the interviewer had access to the original referral, which provided brief background information about the patient. Interrater reliability measured by use of video was very good: intraclass correlation coefficients for the four domains ranged from 0.89 to 0.95, while with a test–retest design, these coefficients ranged from 0.59 to 0.80 (Buer Christensen et al., 2018).

Global Assessment of Functioning. The Global Assessment of Functioning (GAF; APA, 1994) is a commonly used rating scale for assessing a patient's overall mental health, reflecting psychological, social, and occupational functioning. We used the GAF-Split version, assessing symptom (GAF-S) and function (GAF-F) scores separately (Karterud, Pedersen, Loevdahl, & Friis, 1998; Pedersen, Hagtvet, & Karterud, 2007). The referring therapists were trained in using the GAF through courses arranged by the Department of Personality Psychiatry, which included a one-day GAF rating workshop based on video interviews. We conducted no further assessment of rater agreement for the current study.

STATISTICS

In this study, our main aim was to evaluate the SCID-5-AMPD Module I as a "screener" for the presence of a *DSM-IV* PD (Step 1; Table 1). We also used this instrument to evaluate the diagnostic algorithms (Step 3), although the instrument is not designed for this purpose. This influences the use of classification accuracy statistics.

PD or Not: Sensitivity and Specificity. First, we evaluated different potential diagnostic thresholds or cutoff values of a global score, that is, the mean LPFS based on evaluation of all subdomains. We used a standard two-by-two table to find the sensitivity and specificity values for potential mean LPFS cutoff values, with the presence of one or more *DSM-IV* PD diagnoses as a dichotomous "reference standard" (Portney & Watkins, 2009; Trevethan, 2017). PD NOS was included, using *DSM-IV* PD 10 criteria as a threshold (Loranger, 1999; Pagan, Oltmanns, Whitmore, & Turkheimer, 2005). A high degree of sensitivity indicates a low risk for a false negative, and a high degree of specificity means a low risk for a false positive. When evaluating the diagnostic properties of a test, sensitivity and specificity are inversely related depending on the chosen threshold (i.e., cutoff value). We calculated Youden's (1950) index for each threshold. This index, calculated as (Sensitivity + Specificity − 1), is frequently used to choose an "optimal" cutoff value when sensitivity and specificity are equally weighted for a diagnostic threshold (Fluss, Faraggi, &

Reiser, 2005). However, a test has no universally "optimal cutoff point"; the cutoff choice depends on the target population, that is, the purpose of the test, and the clinical consequences of the results. If a test has a screening purpose in an outpatient population, the optimal cutoff criteria might be different from a test for more specific diagnostic use. For example, if the consequences of missing a diagnosis are quite serious, we will choose a lower cutoff point to increase test sensitivity. Furthermore, a high sensitivity would be preferable for a test that is the first of two steps in a diagnostic process, such as for the Criterion A for specific PD diagnoses in the AMPD. Here, the evaluation of the Criterion B, as well as the C through G criteria, is expected to result in the exclusion of some participants who scored positive in the first test. It can be expected that a new diagnostic strategy will include some individuals who do not necessarily fulfill the criteria for a disorder according to the existing diagnostic approach; but a degree of continuity between systems is desirable, because it results in less disruption in clinical practice and research.

We also performed a ROC curve analysis (Metz, 1978), still with any *DSM-IV* diagnosis as a dichotomous criterion variable. The curve was generated from paired sensitivity and 1 – specificity values for all obtained mean LPFS values, and with AUC (Hanley & McNeil, 1982) as a measure of diagnostic accuracy. An AUC value of 1 indicates that the test will result in a correct diagnosis in 100% of cases, whereas a value of 0.5 indicates that the test has no discriminatory ability. The AUC allows for comparison of how different tests perform. It also represents a summary statistic of the overall diagnostic performance of the test, including its standard error, 95% confidence intervals (CIs), and p values, with AUC = 0.5 as the null hypothesis. A commonly used interpretation for the AUC values is that estimates of 0.70 to 0.80 are "acceptable," 0.80 to 0.90 are "excellent," and 0.90 or above is "outstanding" (Hosmer, Lemeshow, & Sturdivant, 2013).

To provide important information regarding how well the thresholds will identify different specific PD diagnoses, we also evaluated sensitivity estimates for the four most important PD diagnoses in our sample. Specificity estimates for the specific PD diagnoses would be of less interest because these would be highly influenced by the high base rate of PD diagnoses in our sample, and not of primary interest in these analyses.

Because the estimates of sensitivity and specificity are conditional to the population in which they were calculated, the Positive Predictive Proportion (PPP) and Negative (NPP) Predictive Proportion would probably be more applicable for clinicians. These provide information about the probability that a patient with a positive test result has the disorder. However, because of the inconsistency of thresholds across different PD diagnoses, these estimates would be highly influenced by the distribution of PD diagnoses and are therefore not reported.

The Specific PD Diagnoses. Regarding the three algorithms evaluated in the Morey survey, we evaluated sensitivity and specificity based on the average of all three subdomains within each domain. Two-by-two tables were used for this step. Importantly, the SCID-5-AMPD Module I was not designed for evaluation of the specific diagnostic criteria, and the LPFS descriptions do not

correspond directly to the disorder-specific A diagnostic criteria. In the Morey survey, disorder-specific impairment was evaluated using a questionnaire for this purpose.

All statistics were performed with IBM SPSS Statistics 23.0 (IBM Corporation, Armonk, New York, NY).

RESULTS

DESCRIPTIVE FINDINGS

To provide an empirical backdrop for the interpretation of our results, descriptive analyses were conducted. The mean LPFS in our sample (Table 2) was 1.9 (SD = 0.9); 6% of patients were rated as having little or no impairment (Level 0), 26.6% had some impairment (Level 1), 39.4% had moderate impairment (Level 2), 23.8% had severe impairment (Level 3), and 4.3% had extreme impairment (Level 4). The number of specific *DSM-IV* PD criteria, specific *DSM-IV* PD diagnoses, and comorbid symptom diagnoses all increased with the severity of the personality function impairment, especially between Levels 3 and 4 (Table 2). For the four main specific *DSM-IV* PD diagnoses, mean LPFS varied from 2.27 for ASPD (SD = 0.87) to 2.79 (SD = 0.79) for OCPD (Table 3). For participants with no comorbid PD diagnoses (referred to as "single PD"), mean LPFS varied from 1.72 for ASPD (SD = 0.56) to 2.27 (SD = 0.47) for BPD. For participants with no PD diagnosis, the mean LPFS score was 1.14 (SD = 0.80).

To compare different PD groups, we also evaluated mean GAF-F (Table 3) as a measure of more general dysfunction. For the four main diagnostic groups (participants with comorbid PD diagnoses included), patients with ASPD demonstrated the most severe dysfunction (47.7; SD = 6.4), while for BPD and AVPD, the estimates were almost identical (51.1, SD = 8.3 for AVPD vs. 51.6, SD = 7.3 for BPD). We found the highest number of symptom disorders in the ASPD group (2.7; SD = 1.9). With respect to PD comorbidity, the OCPD group had the highest number (3.1; SD = 1.9).

Table 4 provides mean scores for all four LPFS domains, based on the ratings of all subdomains. For the whole sample, the least impaired domain

TABLE 2. Mean Number (*SD*) of PD Criteria, PD Diagnosis, and Symptom Diagnosis According to Levels of Personality Functioning (*N* = 282)

Mean LPFS[a]	% (*n*)	Mean no. PD criteria (*SD*)	Mean no. specific PDs (*SD*)	Mean no. symptom diagnosis (*SD*)
0	6 (17)	1.4 (2.7)	0	1.5 (.8)
1	26.6 (75)	6.6 (4.9)	0.4 (0.6)	1.4 (0.9)
2	39.4 (111)	10.7 (5.5)	0.8 (0.8)	1.6 (1.1)
3	23.8 (67)	16.3 (8.7)	1.5 (1.4)	2.2 (1.6)
4	4.3 (12)	26.5 (7.2)	3.1 (1.7)	3.7 (2.4)

[a]Average of 12 subdomain scores; Level 0 = 0–0.49; Level 1 = 0.50–1.49; Level 2 = 1.50–2.49; Level 3 = 2.50–3.49; Level 4 = 3.50–4.00.

TABLE 3. Mean LPFS, GAF-F, and Mean Number of Symptom Diagnoses for Some *DSM-IV* PD Diagnoses (*SD*)

	BPD	AVPD	ASPD	OCPD	Single BPD	Single AVPD	Single A SPD	Single OCPD	PD NOS	No PD
	n = 70	*n* = 81	*n* = 30	*n* = 21	*n* = 32	*n* = 47	*n* = 13	*n* = 6	*n* = 45	*n* = 83
Mean LPFS	2.67	2.29	2.27	2.79	2.47	1.88	1.72	2.22	1.95	1.14
	(0.64)	(0.80)	(0.88)	(0.79)	(0.46)	(0.53)	(0.56)	(0.86)	(0.65)	(0.80)
GAF-F	*n* = 66	*n* = 75	*n* = 30	*n* = 18	*n* = 32	*n* = 41	*n* = 13	*n* = 4	*n* = 41	*n* = 77
	51.6	51.1	47.7	49.6	54.4	54.6	49.9	56.2	54.7	57.9
	(7.3)	(8.3)	(6.4)	(7.0)	(8.0)	(7.7)	(6. 8)	(11.4)	(6.0)	(10. 6)
Number of symptom diagnoses	*n* = 63	*n* = 72	*n* = 29	*n* = 19	*n* = 30	*n* = 43	*n* = 12	*n* = 6	*n* = 36	*n* = 75
	2.2	2.0	2.7	2.4	1.8	1.7	2.2	1.8	1.5	1.3
	(1.6)	(1.6)	(1.9)	(2.0)	(1.3)	(1.9)	(1.7)	(1.8)	(1.3)	(0.8)
Number of PD diagnoses	2.17	1.86	2.37	3.10						
	(1.38)	(1.32)	(1.73)	(1.89)						

LPFS: Level of Personality Functioning Scale; GAF-F: Global Assessment of Functioning functioning score; BPD: borderline personality disorder; AVPD: avoidant PD; ASPD: antisocial PD; OCPD: obsessive-compulsive PD; PD NOS: PD not otherwise specified

based on scores was Empathy (1.44; *SD* = 1.05). With respect to specific PD diagnoses, ASPD was the only main PD category in which Empathy was not rated as less impaired than the other domains (rated almost identical to the Identity domain).

DIAGNOSTIC THRESHOLD FOR THE GENERAL CRITERION A

A ROC curve analysis was performed to estimate the accuracy of global LPFS (i.e., the average of all subdomain scores) as a single test for the presence of one or more *DSM-IV* PDs (Figure 1). An AUC of 0.84 (95% CIs [0.78, 0.90],

TABLE 4. LPFS Domain Scores, Based on Average of All Subdomain Scores (*SD*)

	Identity	Self-direction	Empathy	Intimacy
Clinical sample (282)	2.14 (0.98)	1.97 (0.97)	1.44 (1.05)	1.79 (1.11)
PD incl. NOS (192)	2.47 (0.79)	2.28 (0.81)	1.73 (1.0)	2.32 (0.94)
Clinical, No PD (83)	1.4 (0.99)	1.25 (0.98)	0.81 (0.85)	1.06 (0.97)
BPD (70)	2.91 (0.57)	2.72 (0.73)	2.24 (0.91)	2.81 (0.78)
AVPD (81)	2.64 (0.75)	2.33 (0.87)	1.71 (1.15)	2.40 (0.96)
ASPD (30)	2.10 (1.12)	2.67 (0.77)	2.11 (0.92)	2.19 (1.14)
OCPD (21)	2.33 (1.19)	2.17 (0.86)	1.94 (0.82)	2.44 (0.83)
Single BPD (32)	2.80 (0.35)	2.53 (0.59)	1.92 (0.73)	2.63 (0.67)
Single AVPD (47)	2.36 (0.62)	1.88 (0.58)	1.13 (0.88)	2.15 (0.92)
Single ASPD (13)	1.38 (0.73)	2.33 (0.63)	1.69 (0.67)	1.46 (0.93)
Single OCPD (6)	2.86 (0.92)	2.73 (0.81)	2.52 (0.90)	3.05 (0.79)

NOS: not otherwise specified; BPD: borderline personality disorder; AVPD: avoidant PD; ASPD: antisocial PD; OCPD: obsessive-compulsive PD.
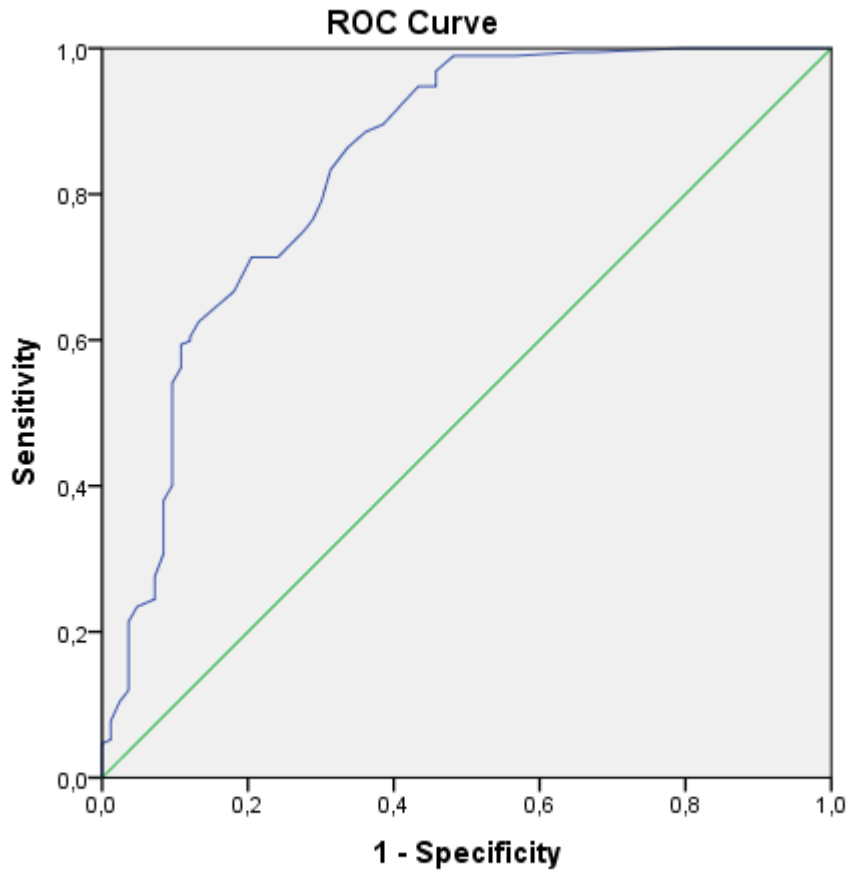
AREA UNDER CURVE (AUC) = 0.839 (95% CIS [0.781, 0.895]; *P* < 0.005)

FIGURE 1. Receiver operator characteristic (ROC) curve for all values
of mean LPFS based on average subdomain scores, with any *DSM-IV* diagnosis
(including PD NOS) set as a dichotomous variable.

p < .005) was found, indicating good accuracy, estimated with a high degree
of precision. Table 5 provides estimates for sensitivity, specificity, and Youden's
index for different possible thresholds of mean LPFS in our sample, relative
to receiving any *DSM-IV* PD diagnosis. When sensitivity and specificity were
equally weighted, a threshold value of 1.4 was found as an optimal cutoff in
our sample, with a sensitivity of .87 and specificity of .66. In other words,
using this threshold would result in 87% of participants with a *DSM-IV* PD
diagnosis being identified as cases using the global LPFS score, while 66% of
patients that do not meet criteria for a *DSM-IV* PD disorder were classified
as not meeting the threshold using the LFPS score
       When the three main specific PD diagnoses in our sample were evaluated,
estimates for sensitivity were substantially higher for BPD than for AVPD and

TABLE 5. Sensitivity (Sens), Specificity (Spec), and Youden's J Statistic (Y's J) for Different Threshold Values of Mean LPFS Scores Gauged Against Any *DSM-IV* Diagnosis, and Sensitivity for the Three Main Specific PD Diagnoses in Our Sample

	DSM IV diagnosis, PD NOS included[a]			BPD[b]	AVPD[b]	ASPD[b]	OCPD	Single BPD	Single AVPD	Single ASPD
	n = 275[c]			*n* = 70	*n* = 81	*n* = 30	*n* = 21	*n* = 32	*n* = 47	*n* = 13
Mean LPFS[d]	Sens	Spec	Y's J	Sens	Sens	Sens	Sens	Sens	Sens	Sens
1.0	.969	.542	.511	1.00	.975	.967	.952	1.00	.979	.923
1.1	.948	.566	.514	1.00	.975	.967	.952	1.00	.979	.923
1.2	.922	.590	.512	1.00	.951	.867	.952	1.0	.936	.692
1.3	.896	.614	.510	.986	.938	.800	.952	1.0	.915	.615
1.4	.865	.663	.528	.986	.864	.800	.952	1.0	.809	.615
1.5	.792	.699	.491	.943	.765	.767	.952	.969	.660	.615
1.6	.745	.723	.468	.929	.728	.700	.952	.969	.596	.538
1.7	.714	.771	.485	.943	.716	.700	.905	.969	.574	.538
1.8	.667	.819	.486	.900	.691	.667	.905	.906	.532	.462
1.9	.625	.867	.492	.871	.667	.633	.857	.875	.489	.385
2.0	.563	.892	.455	.843	.580	.567	.857	.813	.362	.308
2.1	.542	.904	.446	.843	.556	.567	.810	.813	.319	.308
2.2	.474	.904	.378	.757	.481	.533	.810	.688	.234	.231
2.5	.309	.916	.225	.552	.370	.333	.719	.387	.149	.077

Note. LPFS: Level of Personality Functioning Scale; PD NOS: PD not otherwise specified; BPD: borderline personality disorder; AVPD: avoidant PD; ASPD: antisocial PD; OCPD: obsessive-compulsive PD. [a]If specific *DSM-IV* diagnoses were applied as a reference standard, we found only minor changes in sensitivity: For mean LPFS 1.4: sens = .865/spec. = .663, for mean LPFS 1.5: sens = .792/spec. = .669; [b]Participants with co-occurrent PD diagnosis are included; [c]Number is lower than the clinical sample (*N* = 282) because of missing *DSM-IV* diagnosis for seven participants; [d]Average when all 12 subdomains are rated.

ASPD. For example, with a threshold of 1.5 for mean LPFS score, sensitivity was 0.77 for AVPD, 0.77 for ASPD, and 0.94 for BPD. The corresponding estimates for specific PD diagnoses when comorbid PD diagnoses were excluded were 0.97 for single BPD, 0.66 for single AVPD, and 0.62 for single ASPD. Specificity for the specific PD diagnosis was not relevant when the whole sample was evaluated because impairment in personality functioning will often be related to co-occuring PD diagnoses.

TABLE 6. Sensitivity (Sens) and Specificity (Spec) of Three Evaluated Algorithms as Gauged Against *DSM-IV* PD Diagnoses

Algorithms evaluated	BPD		AVPD		ASPD		OCPD	
	n = 71		*n* = 80		*n* = 30		*n* = 21	
	Sens	Spec	Sens	Spec	Sens	Spec	Sens	Spec
Any one of four	1.00	.196	.988	.201	1.00	.167	.952	.154
Any two of four	.986	.361	.926	.354	.833	.296	.905	.291
One Self and One Interpersonal	.943	.490	.802	.454	.733	.400	.905	.402

BPD: borderline personality disorder; AVPD: avoidant PD; ASPD: antisocial PD; OCPD: obsessive-compulsive PD.

DIAGNOSTIC THRESHOLDS FOR
THE DISORDER-SPECIFIC A CRITERIA

In accordance with the Morey survey, we evaluated three possible algorithms (Table 6), using the mean scores for the three subdomains within each domain as the basis for our analysis. The criterion to be fulfilled was a mean score of 1.5 or more, representing a moderate (i.e., Level 2) or more severe impairment in personality functioning.

The three algorithms tested were (a) one from each component present (i.e., requiring criteria met for at least one domain from self-functioning and one domain from interpersonal functioning); (b) any two areas present (i.e., criteria met for any two domains for either self- or interpersonal functioning); and (c) any one present (i.e., requiring criteria met for only one domain) (Table 6). Our sample allowed us to separately test four of the retained *DSM-IV* PD diagnoses in AMPD: BPD, AVPD, ASPD, and OCPD.

For the first algorithm tested, requiring criteria to be met in one domain of both self- and interpersonal functioning, sensitivities were 0.73 for ASPD, 0.80 for AVPD, 0.91 for OCPD, and 0.99 for BPD. For the second algorithm, the suggested "any two areas present" algorithm, estimates for sensitivity were 0.83 for ASPD, 0.91 for OCPD, 0.93 for AVPD, and 0.99 for BPD. For the third algorithm, requiring criteria to be met for only one domain resulted in sensitivity estimates of 0.95 for OCPD, 0.99 for AVPD, and 1.0 for both ASPD and BPD. For all PD diagnoses, specificity was, as expected, lowest for the least restrictive algorithm (the third) and highest for the most restrictive algorithm. For the most restrictive algorithm, the rates ranged from .40 (ASPD) to .49 (BPD). The best combination of sensitivity and specificity was demonstrated for the second algorithm, "any two areas present" (Table 6).

DISCUSSION

The current study is the first to evaluate the diagnostic decision rules (i.e., the cut-point for the presence of a PD and the "any two areas present" rule for specific PDs for Criterion A) of the Alternative Model of *DSM-5* Section III since the rules were developed (Morey et al., 2013; Morey & Skodol, 2013). The results provide empirical background for establishing a diagnostic threshold for receiving a PD diagnosis based on a global LPFS score. According to the ROC analysis, the diagnostic accuracy of using mean LPFS as an indicator of *DSM-IV* PD was good. When sensitivity and specificity were equally weighted, a global LPFS cutoff value slightly below 1.5 resulted in the best combination for two of the main PD diagnoses in our sample. However, we still find that moderate impairment in personality functioning, with a threshold of 1.5 for global LPFS, represents a reasonable diagnostic threshold for a PD diagnosis in the AMPD. For further diagnostic evaluation of the A criteria for specific PDs, an algorithm using the "any two areas present" rule based on subdomain scores indicated a reasonably good fit between the *DSM-IV* PD diagnoses and the Criterion A algorithm.

DIAGNOSTIC THRESHOLD FOR
THE GENERAL CRITERION A

The ROC curve analysis demonstrated excellent accuracy for global LPFS as a diagnostic test for a *DSM-IV* PD diagnosis. These results are almost identical to those of the Morey survey, despite involving a very different diagnostic assessment procedure and sample. Thus, a large percentage of persons meeting *DSM-IV* criteria also met the LPFS threshold. The continuity between *DSM-IV* and AMPD therefore was good with respect to global LPFS as a predictor of assignment of a *DSM-IV* PD diagnosis.

We evaluated possible threshold values of mean LPFS, with any *DSM-IV* PD diagnoses as the reference standard. Of interest, the optimal diagnostic threshold was 1.4, given that sensitivity and specificity were equally weighted. However, no clear consensus exists regarding an optimal combination of sensitivity and specificity. The weighting of sensitivity and specificity will depend on the consequences of the result and the prevalence of the disorder in the target population. If a test is a "screener" for further evaluation, for example, as a first step in a stepwise diagnostic process, we usually accept a low degree of specificity initially because further steps will add to the specificity of the diagnostic process. Furthermore, regarding the purpose of this study, the AMPD is expected to present a more evidence-based structure of PDs than the *DSM-IV*. This indicates that the *DSM-IV* need not be regarded as a "gold standard" for evaluating the LPFS. Although the SCID-5-AMPD Module I is not technically a screening instrument, our results provide a foundation for discussing the clinical consequences of diagnostic thresholds based on global LPFS scores in a clinical population.

If we look at AVPD, the most prevalent *DSM-IV* PD diagnosis in our sample, using 1.5 as the diagnostic threshold, would mean that up to one in four participants with this diagnosis and about one in three with a single AVPD (i.e., with AVPD only) would be excluded if a global rating of LPFS were used to screen for a possible PD. The consequence could be exclusion from further evaluation and from treatment programs. The implications could be serious, given the high prevalence of AVPD and the significant psychosocial impairment many studies have reported in people with the condition (Eikenaes, Hummelen, Abrahamsen, Andrea, & Wilberg, 2013; Lampe & Malhi, 2018). For example, in our sample, individuals with AVPD were rated with estimates of GAF-F similar to individuals with BPD. To the best of our knowledge, two other studies have evaluated the relationship between LPFS and AVPD, and only one was based on rating all subdomains (Hutsebaut et al., 2017). In the study by Hutsebaut and colleagues, a close to nonexisting correlation ($r = 0.01$) between global LPFS score and *DSM-IV* AVPD criteria was demonstrated. In a study by Few et al. (2013), LPFS ratings were based on the SCID-II, with weak to moderate correlations (rs ranging from 0.17 to 0.44) among the four domain scores of LPFS and *DSM-IV* AVPD criteria. These studies both converge with our results, indicating that the LPFS might not capture the impairment in personality functioning associated with this diagnosis well.

For individuals with an *DSM-IV* AVPD diagnosis, Empathy was the domain with the lowest impairment score. This finding is in line with the results of Few et al. (2013). One possible explanation is that raters did not find that the descriptions in the LPFS were apt for the impairment of these individuals. For example, in the subdomain understanding and appreciation of others' experiences and motivations, moderate impairment is described as "Is hyper attuned to the experience of others, but only with respect to perceived relevance to self." We find the first part of description applicable for AVPD. The second part, however, is indicative of narcissism, focusing on self-enhancing interpersonal striving (Ronningstam, 2009) rather than on the detached pattern of interpersonal interaction characteristic of AVPD (Lampe & Malhi, 2018). We did not analyze the ratings of each subdomain separately in our study, but if the raters found none of the descriptions in Level 2 to be relevant, a lower score (i.e., Level 1) could be the result. This subdomain also illustrates a discrepancy between the description in LPFS and the disorder-specific criteria in the AVPD. For impairment in Empathy, the diagnostic criterion A for AVPD is "Preoccupation with, and sensitivity to criticism and rejection, associated with distorted inference of others perspectives as negative" (APA, 2013, p. 765). We find this statement more descriptive of the typical features of AVPD.

For individuals with a *DSM-IV* BPD diagnosis, however, applying 1.5 as a diagnostic threshold led to almost 100% sensitivity. Even when comorbid PD diagnoses were excluded, the estimated sensitivity indicated that almost all participants with this diagnosis would be included for further evaluation. This result is consistent with the aim of the *DSM-5* Personality and Personality Disorders Work Group to avoid significant disruption between *DSM-IV* and the AMPD for the BPD diagnosis. The high degree of convergence between global LPFS score and this *DSM-IV* PD diagnosis that we identified here is also in line with previous findings (Few et al., 2013; Hutsebaut et al., 2017; Morey et al., 2011). We know that, in the AMPD, a PD is defined through impairment in key domains of personality functioning, operationalized through the LPFS. The strong association between global LPFS score and BPD is compatible with the perception that BPD represents a general PD factor accounting for the common variance shared across all PDs, as some authors have argued (Sharp et al., 2015; Wright et al., 2016).

Discontinuity between the AMPD and the *DSM-IV* was also a major concern for the third specific PD diagnosis we evaluated, ASPD. The low degree of sensitivity with a threshold of 1.5 calls for a comment, as most participants with an ASPD in our study were incarcerated. We are aware of two other clinical studies evaluating LPFS in incarcerated populations with ASPD, and in both studies, the estimates of LPFS domain scores were lower than in our sample (Amini, Pourshahbaz, Mohammadkhani, Khodaie Ardakani, & Lotfi, 2015; Wygant et al., 2016). The estimates for sensitivity in our sample may have several explanations. First, individuals recruited from a prison may have fewer symptoms than participants recruited from other treatment units (because incarceration itself is a result of antisocial behavior). Second, we know from another substudy of this sample (Holde, 2017) that psychopathy, as screened for by the Psychopathy Check List–Screening Version (Hart, Cox, & Hare, 1995), and LPFS were negatively correlated. In our ASPD sample, a

high degree of psychopathic traits was demonstrated. In the same substudy, there was a low rating of impairment in the subdomain understanding the effect one's behavior has on others (in the Empathy domain). Both of these findings could be explained by unreliability and lack of insight about one's own behavioral patterns, often described as key features in psychopathy (Cleckley, 1955). However, our results should be interpreted with caution because most participants in the ASPD group also fulfilled criteria for other PD diagnoses, and also because of the sample size ($N = 35$).

For participants in the fourth group we evaluated, OCPD sensitivity estimates were comparable with those for BPD, that is, higher than for the other PD diagnoses evaluated. This finding might be unexpected because some studies have demonstrated better general functioning for OCPD than for other PD diagnoses (Cramer, Torgersen, & Kringlen, 2006; Skodol et al., 2002). We also found that co-occurrence of other PD diagnoses was substantially higher in this group than for the others, which is consistent with another relevant study from the Norwegian Network of Personality-Focused Treatment Programs (Hummelen, Wilberg, Pedersen, & Karterud, 2008). However, the sample size was small ($N = 21$) so our findings should be interpreted with care.

On the whole, our findings could be considered in adjusting to a lower cutoff point for fulfilling the diagnostic criteria for a PD, for example, to improve sensitivity for AVPD. However, for the BPD and OCPD groups, the sensitivity in our study indicates that as many as 95% would be included with 1.5 as a cutoff. Rather than an adjustment of the threshold, in alignment with prior research of the alternative model (Hummelen, 2019; Zimmermann et al., 2014; Zimmermann et al., 2015), we would suggest considering some revision of the descriptions in the LPFS for future refinement. For instance, we believe that some adjustments in the empathy domain could result in better continuity with the *DSM-IV*. However, as previously underlined, the *DSM-IV* is no "gold standard," and its diagnostic thresholds have been criticized as being inconsistent (Balsis et al., 2011).

DIAGNOSTIC THRESHOLDS FOR THE DISORDER-SPECIFIC A CRITERIA

As noted, the disorder-specific A diagnostic criteria are descriptions of impairment that are distinct for each of the six PDs retained in the AMPD. These criteria do not directly correspond with the descriptions in the LPFS. The algorithms used in this study are therefore less specific than these criteria because average scores for all three subdomain scores are used instead of disorder-specific criteria.

In our evaluation of the most restrictive algorithm, requiring a moderate or more severe impairment in at least one self-domain and one interpersonal domain, the results indicated poor continuity with the *DSM-IV*, especially for ASPD and AVPD. The second algorithm we evaluated, the "any two areas present" rule incorporated into the AMPD, provided a high degree of sensitivity. For BPD, ASPD, and OCPD, all estimates of sensitivity were higher than in the Morey survey. This discrepancy could be due to methodological differences; the criteria used in the Morey survey were more specific than the

subdomain ratings we used here, thus decreasing sensitivity. Only for AVPD did we find a slightly lower sensitivity (0.926 vs. 0.955 in the Morey survey). This result could be explained by a low degree of PD comorbidity for this group. The least restrictive algorithm, in which a moderate or more severe impairment is required in only one of the four domains, demonstrated an almost 100% sensitivity. The specificity rates were in general lower than in the Morey survey, probably because our evaluation was based on the average of all subdomains, not the specific criteria. However, the best combination of sensitivity and specificity was demonstrated for the "any two present" algorithm. Although this evaluation was not based on ratings of the specific criteria, this may provide some support for the "any two areas present" rule. However, this needs to be further evaluated by another measure tailored to the assessment of the specific PDs included in the AMPD.

LIMITATIONS AND FUTURE DIRECTIONS

To the best of our knowledge, this study is the first to evaluate diagnostic decision rules for the LPFS using an instrument tailored for clinician rating on a well-defined patient sample comprising patients with few or no personality problems to those with severe PDs. The reliability of the *DSM-IV* diagnoses was not assessed in this study, but all referring therapists were experienced and thoroughly trained. Of note, the SCID-5-AMPD Module I interview was not developed for screening purposes or for assessment of the disorder-specific AMPD criteria. Like the STiP-5.1, this instrument was designed for evaluating general impairment in personality functioning, based on assessment of all 12 LPFS subdomains. Although there is theoretical and empirical support for a single rating of the LPFS, the low sensitivity for AVPD and ASPD calls into question the idea of one threshold based on a global score for screening purposes, especially for the AVPD group. A higher degree of convergence between the disorder-specific A criteria and the descriptions in the LPFS should be addressed in future revisions of the AMPD. This low degree of convergence is also the reason why our evaluation of the specific PD A criteria algorithms is not conclusive.

In a prior article (Buer Christensen et al., 2018), our research group suggested that some of the descriptions of the LPFS should be carefully examined prior to future revisions. In this study, we have suggested that some adjustment of the descriptions in the Empathy domain could result in better continuity with *DSM-IV* for the AVPD group. Regarding the ASPD group in our study, the fact that most individuals were incarcerated could limit the generalizability of our results, but our results regarding this diagnostic category should be interpreted with caution.

Furthermore, as other authors have emphasized (Morey et al., 2011; Morey & Skodol, 2013), the *DSM-IV* is only one reference standard for evaluating the AMPD. As members of the *DSM-5* Personality and Personality Disorders Work Group called for early on, studies using validators other than the *DSM-IV*, such as other measures of functional outcome, are still needed. Our study provides valuable findings that may be of use for future revisions of the AMPD, in line with the intentions of the *DSM-5* as a "living document."

Until then, more studies are needed based on tailored instruments for assessment of the LPFS in representative clinical populations.

CONCLUSION

We found that an instrument tailored for clinician ratings of the LPFS, the SCID-5-AMPD Module I, provided empirical support for moderate impairment in global personality functioning as a diagnostic threshold for identifying a PD, in keeping with the AMPD's general Criterion A. Our study also provided some support for the "any two areas present" rule for the A criteria for specific PD diagnoses, also when the rating was based on LPFS domain scores. Although continuity with the *DSM-IV* is not the sole objective of the AMPD, some revision of the LPFS could make the scale more feasible for a screening purpose. In addition, a clearer description is needed for how to evaluate the LPFS for screening purposes. Regarding the specific PD diagnosis, a new instrument has recently been published by some of the authors behind the AMPD (First, Skodol, Bender, & Oldham, 2018), which includes a Module III designed to evaluate the six specific PDs in the AMPD. This instrument should be suitable for further evaluation of the diagnostic algorithms for these disorders. We suggest that the disorder-specific A criteria and the LPFS should be better harmonized. That said, we expect the LPFS to prove its utility as a severity measure across all PD diagnoses through future revisions of the AMPD toward an even more dimensional model in the *DSM-5.1.*

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Publishing.

Amini, M., Pourshahbaz, A., Mohammadkhani, P., Khodaie Ardakani, M. R., & Lotfi, M. (2015). The *DSM-5* levels of personality functioning and severity of Iranian patients with antisocial and borderline personality disorders. *Iranian Red Crescent Medical Journal, 17*(8), e19885. https://doi.org/10.5812/ircmj.17(5)2015.19885

Arnevik, E., Wilberg, T., Urnes, O., Johansen, M., Monsen, J. T., & Karterud, S. (2009). Psychotherapy for personality disorders: Short-term day hospital psychotherapy versus outpatient individual therapy—A randomized controlled study. *European Psychiatry, 24*(2), 71–78. https://doi.org/10.1016/j.eurpsy.2008.09.004

Bach, B., & First, M. B. (2018). Application of the ICD-11 classification of personality disorders. *BMC Psychiatry, 18*(1), 351. https://doi.org/10.1186/s12888-018-1908-3

Bach, B., & Hutsebaut, J. (2018). Level of Personality Functioning Scale-Brief Form 2.0: Utility in capturing personality problems in psychiatric outpatients and incarcerated addicts. *Journal of Personality Assessment, 100*, 660–670. https://doi.org/10.1080/00223891.2018.1428984

Balsis, S., Lowmaster, S., Cooper, L. D., & Benge, J. F. (2011). Personality disorder diagnostic thresholds correspond to different levels of latent pathology. *Journal of Personality Disorders, 25*, 115–127. https://doi.org/10.1521/pedi.2011.25.1.115

Bender, D. S., Morey, L. C., & Skodol, A. E. (2011). Toward a model for assessing level of personality functioning in *DSM-5*, part I: A review of theory and methods. *Journal of Personality Assessment, 93*, 332–346. https://doi.org/10.1080/00223891.2011.583808

Bender, D. S., Skodol, A., First, M. B., & Oldham, J. (2018). *Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders (SCID-5-AMPD) Module I: Level*

of Personality Functioning Scale. Arlington, VA: American Psychiatric Association.

Blashfield, R. K., & Intoccia, V. (2000). Growth of the literature on the topic of personality disorders. *American Journal of Psychiatry, 157*, 472–473.

Buer Christensen, T., Paap, M. C. S., Arnesen, M., Koritzinsky, K., Nysaeter, T. E., Eikenaes, I., . . . Hummelen, B. (2018). Interrater reliability of the Structured Clinical Interview for the *DSM-5* Alternative Model of Personality Disorders Module I: Level of Personality Functioning Scale. *Journal of Personality Assessment, 100*, 630–641. https://doi.org/10.1080/00223891.2018.1483377

Cleckley, H. M. (1955). *The mask of sanity: An attempt to clarify some issues about the so-called psychopathic personality.* St. Louis, MO: C. V. Mosby.

Cramer, V., Torgersen, S., & Kringlen, E. (2006). Personality disorders and quality of life. A population study. *Comprehensive Psychiatry, 47*, 178–184. https://doi.org/10.1016/j.comppsych.2005.06.002

Eikenaes, I., Hummelen, B., Abrahamsen, G., Andrea, H., & Wilberg, T. (2013). Personality functioning in patients with avoidant personality disorder and social phobia. *Journal of Personality Disorders, 27*, 746–763. https://doi.org/10.1521/pedi_2013_27_109

Few, L. R., Miller, J. D., Rothbaum, A. O., Meller, S., Maples, J., Terry, D. P., . . . MacKillop, J. (2013). Examination of the Section III *DSM-5* diagnostic system for personality disorders in an outpatient clinical sample. *Journal of Abnormal Psychology, 122*, 1057–1069. https://doi.org/10.1037/a0034878

First, M. B., Skodol, A. E., Bender, D. S., & Oldham, J. M. (2018). *Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders (SCID-5-AMPD) Module III Personality Disorders (Including Personality Disorder–Trait Specified)*. Arlington, VA: American Psychiatric Publishing.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., & Benjamin, L. (1994). *Structured Clinical Interview for DSM-IV Axis II personality disorders (SCID-II, version 2.0)*. New York, NY: Biometric Research Department, New York Psychiatric Hospital.

Fluss, R., Faraggi, D., & Reiser, B. (2005). Estimation of the Youden Index and its associated cutoff point. *Biometrical Journal, 47*(4), 458–472.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36.

Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *Hare Psychopathy Checklist: Screening version (PCL: SV)*. Toronto, Canada: Multi-Heath Systems.

Holde, K. (2017). *Exploring the relationship between the* DSM-5 *Levels of Personality Functioning Scale and psychopathy in a clinical sample of patients with antisocial personality disorder* (unpublished master's thesis), University of Oslo, Oslo, Norway.

Hosmer, Jr., D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Hummelen, B. (2019). *A psychometric analysis of the Structured Clinical Interview for the* DSM-5 *Alternative Model for Personality Disorders Module I; Level of Personality Functioning Scale (SCID-5-AMPD-I)*. Manuscript submitted for publication

Hummelen, B., Wilberg, T., Pedersen, G., & Karterud, S. (2008). The quality of the *DSM-IV* obsessive-compulsive personality disorder construct as a prototype category. *Journal of Nervous and Mental Disease, 196*, 446–455. https://doi.org/10.1097/NMD.0b013e3181775a4e

Hutsebaut, J., Kamphuis, J. H., Feenstra, D. J., Weekers, L. C., & De Saeger, H. (2017). Assessing *DSM-5*-oriented level of personality functioning: Development and psychometric evaluation of the Semi-Structured Interview for Personality Functioning *DSM-5* (STiP-5.1). *Personality Disorders, 8*, 94–101. https://doi.org/10.1037/per0000197

Karterud, S., Pedersen, G., Loevdahl, H., & Friis, S. (1998). *Global Assessment of Functioning–Split Version (S-GAF): Background and scoring manual.* Oslo, Norway: Ullevaal University Hospital, Department of Psychiatry.

Karterud, S., & Wilberg, T. (2007). From general day hospital treatment to specialized treatment programmes. *International Review of Psychiatry, 19*(1), 39–49. https://doi.org/10.1080/09540260601080821

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*(9), 1879–1890. https://doi.org/10.1017/s0033291711002674

Lampe, L., & Malhi, G. S. (2018). Avoidant personality disorder: Current insights. *Psychology Research and Behavior Management, 11*, 55–66. doi:10.2147/PRBM.S121073

Loranger, A. W. (1999). *International Personality Disorder Examination: IPDE; DSM-IV and ICD-10; Interviews*. Lutz, FL: PAR.

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine, 8*(4), 283–298.

Morey, L. C. (2017). Development and initial evaluation of a self-report form of the *DSM-5*

Level of Personality Functioning Scale. *Psychological Assessment, 29*, 1302–1308. https://doi.org/10.1037/pas0000450

Morey, L. C., Bender, D. S., & Skodol, A. E. (2013). Validating the proposed *Diagnostic and Statistical Manual of Mental Disorders*, 5th edition, severity indicator for personality disorder. *Journal of Nervous and Mental Disease, 201*, 729–735. https://doi.org/10.1097/NMD.0b013e3182a20ea8

Morey, L. C., Berghuis, H., Bender, D. S., Verheul, R., Krueger, R. F., & Skodol, A. E. (2011). Toward a model for assessing level of personality functioning in *DSM-5*, part II: Empirical articulation of a core dimension of personality pathology. *Journal of Personality Assessment, 93*, 347–353. https://doi.org/10.1080/00223891.2011.577853

Morey, L. C., & Skodol, A. E. (2013). Convergence between *DSM-IV-TR* and *DSM-5* diagnostic models for personality disorder: Evaluation of strategies for establishing diagnostic thresholds. *Journal of Psychiatric Practice, 19*, 179–193. https://doi.org/10.1097/01.pra.0000430502.78833.06

Pagan, J. L., Oltmanns, T. F., Whitmore, M. J., & Turkheimer, E. (2005). Personality disorder not otherwise specified: Searching for an empirically based diagnostic threshold. *Journal of Personality Disorders, 19*, 674–689.

Pedersen, G., Hagtvet, K. A., & Karterud, S. (2007). Generalizability studies of the Global Assessment of Functioning–Split version. *Comprehensive Psychiatry, 48*, 88–94. https://doi.org/10.1016/j.comppsych.2006.03.008

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (3rd ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.

Reed, G. M. (2018). Progress in developing a classification of personality disorders for ICD-11. *World Psychiatry, 17*, 227–229. https://doi.org/10.1002/wps.20533

Ronningstam, E. (2009). Narcissistic personality disorder: Facing *DSM-V*. *Psychiatric Annals, 39*, 111–121.

Sharp, C., Wright, A. G., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., & Clark, L. A. (2015). The structure of personality pathology: Both general ('g') and specific ('s') factors? *Journal of Abnormal Psychology, 124*, 387–398. https://doi.org/10.1037/abn0000033

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . 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. *Journal of Clinical Psychiatry, 59*(Suppl. 20), 22–33; quiz 34–57.

Skodol, A. E. (2014). Personality disorder classification: Stuck in neutral, how to move forward? *Current Psychiatry Reports, 16*(10), 480. https://doi.org/10.1007/s11920-014-0480-x

Skodol, A. E., First, M. B., Bender, D. S., & Oldham, J. M. (2018). *Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders (SCID-5-AMPD) Module II Personality Traits*. Arlington, VA: American Psychiatric Publishing.

Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S., . . . Oldham, J. M. (2002). Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *American Journal of Psychiatry, 159*, 276–283. https://doi.org/10.1176/appi.ajp.159.2.276

Skodol, A. E., Morey, L. C., Bender, D. S., & Oldham, J. M. (2015). The Alternative *DSM-5* Model for Personality Disorders: A clinical application. *American Journal of Psychiatry, 172*, 606–613. https://doi.org/10.1176/appi.ajp.2015.14101220

Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health, 5*, 307. https://doi.org/10.3389/fpubh.2017.00307

Tyrer, P., Mulder, R., Kim, Y. R., & Crawford, M. J. (2019). The development of the ICD-11 classification of personality disorders: An amalgam of science, pragmatism, and politics. *Annual Review of Clinical Psychology, 15*, 481–502. https://doi.org/10.1146/annurev-clinpsy-050718-095736

Weertman, A., Arntz, A., Dreessen, L., van Velzen, C., & Vertommen, S. (2003). Short-interval test-retest interrater reliability of the Dutch version of the Structured Clinical Interview for *DSM-IV* personality disorders (SCID-II). *Journal of Personality Disorders, 17*, 562–567.

Wright, A. G., Hopwood, C. J., Skodol, A. E., & Morey, L. C. (2016). Longitudinal validation of general and specific structural features of personality pathology. *Journal of Abnormal Psychology, 125*, 1120–1134. https://doi.org/10.1037/abn0000165

Wygant, D. B., Sellbom, M., Sleep, C. E., Wall, T. D., Applegate, K. C., Krueger, R. F., & Patrick, C. J. (2016). Examining the *DSM-5* alternative personality disorder model operationalization of antisocial personality disorder and psychopathy in a male correctional sample. *Personality Disorders: Theory, Research, and Treatment, 7*, 229–239. https://doi.org/10.1037/per0000179

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32–35.

Zimmermann, J., Benecke, C., Bender, D. S., Skodol, A. E., Schauenburg, H., Cierpka, M., & Leising, D. (2014). Assessing *DSM-5* level of personality functioning from videotaped clinical interviews: A pilot study with untrained and clinically inexperienced students. *Journal of Personality Assessment, 96*, 397–409. https://doi.org/10.1080/00223891.2013.852563

Zimmermann, J., Bohnke, J. R., Eschstruth, R., Mathews, A., Wenzel, K., & Leising, D. (2015). The latent structure of personality functioning: Investigating criterion A from the alternative model for personality disorders in *DSM-5*. *Journal of Abnormal Psychology, 124*, 532–548. https://doi.org/10.1037/abn0000059