EXAMINATION OF THE SCREENING PROPERTIES OF THE PERSONALITY DIAGNOSTIC QUESTIONNAIRE-4+ (PDQ-4+) IN A NON-CLINICAL SAMPLE

Martine Bouvard, Marie Vuachet, Caroline Marchand

Abstract

Given the time-consuming nature of interviewing in diagnosing personality disorders (PDs), it is important to establish the qualities of the Personality Diagnostic Questionnaire (PDQ-4+). The PDQ-4+ is a self-report questionnaire derived from the personality disorders section of the DSM-IV. The clinician should assess the clinical significance for any of the disorders for which the patient meets criteria with the Clinical Significance Scale. The aims of the study are to establish the efficacy of the PDQ-4+ Clinical Significance Scale as a screening instrument. The SCID-II was used as the criterion. A group of 251 French undergraduate students were administered the PDQ-4+, the SCID-II and the M.I.N.I., a structured diagnostic interview for Axis I disorders. *Results*: More than half of the participants were classified as showing one or more PDs when using the SCID-II and PDQ-4+ questionnaires. However, only 55 students (21.19%) met the criteria for one or more PDs following the Clinical Significance Scale of the PDQ-4+ (a mini interview), and 35 students (13.94%) met the criteria for one or more PDs following the SCID-II interview. The simultaneous use of the SCID-II interview and the Clinical Significance Scale (a mini interview of the PDQ4+) could allow clinicians to reduce the number of false positive identified by these measures. The level of agreement between the SCID-II and PDQ-4+ was moderate for 2 PDs (borderline, obsessive-compulsive) and light for 5 PDs (dependent, passive aggressive, antisocial, depressive, avoidant). Four PDs (borderline, obsessive-compulsive, dependent, passive aggressive) emerged as relatively efficient with positive predictive power in the moderate range, identifying a moderate proportion of the students (sensitivity in the moderate range). *In conclusion,* the use of the Clinical Significance Scale of the PDQ-4+ and the awareness of Axis I diagnosis, improved the agreement with the SCID-II structured interview.

Key words: personality disorders, questionnaire, interview, PDQ-4+, SCID-II

Declaration of interest: none

Martine Bouvard[1], Marie Vuachet[1], Caroline Marchand[2]

[1] = LIP, Université de Savoie, Chambéry, France
[2] = clinique Lyon Lumière, 33B rue du 8 mai 1945, 69330, Meyzieu, France

Institution:
Laboratoire LIP, Université de Savoie, BP 1104, 73011 Chambéry, France.
Phone: 33 (0)4 79 75 83 42
Email: martine.bouvard@univ-savoie.fr

Introduction

The nature of the relationships among personality, personality disorders (PDs), and other mental disorders, has been a recurrent theme in the history of psychiatry. In some cases, comorbidity seems to be an artefact, such as in social phobia and avoidant personality disorders, or dysthymia and depressive personality disorders (Petot 2007). The difficulty to establish clear boundaries between Axis I and Axis II disorders is also present in the case of schizotypal personality disorder and a minor form of schizophrenia (Petot 2007), since this PD is only recognized in the diagnostic criteria of the American Psychiatric Association (APA 1994). Finally, research has shown the influence of depression and anxiety on the evaluation of PDs. Interviews and questionnaires on personality detect fewer PDs after successful treatment of depression or anxiety (Stuart et al. 1991, Ricciardi et al. 1992, Fava et al. 1994, McKay et al. 1996). Participants' responses in a depressive or anxious phase appear to be strongly modified by their clinical state and clinicians need a great deal of experience and ability to make the difference between Axis I and Axis II diagnosis.

Several interviews or questionnaires are available to clinicians who desire to evaluate PDs, but no "gold standard" exists for the diagnosis of PDs. The Structured Interview for DSM-III Personality Disorders (SIDP) was the first structured interview to offer an extensive evaluation of the PDs (Pfohl et al. 1982). Its most recent version, the SIDP-IV (Pfohl et al. 1995) allows, in addition to the 10 officially recognized PDs in the DSM-IV, the assessment of depressive personality and passive-aggressive (negativistic) disorders, as included in

Submitted December 2010, Accepted March 2011

Annex B of DSM-IV. It also assesses for self-defeating personality disorder, which had been introduced and proposed for further studies in DSM-III-R but has not been retained in DSM-IV (Cloos et al. 2006). The International Personality Disorders Examination (IPDE) was developed, from 1985, within the Joint Program for the Diagnosis and Classification of Mental Disorders of the World Health Organization (WHO) and U.S. National Institutes of Health (NIH) and provides a uniform approach for assessing PDs for both the DSM-IV and the ICD-10 classification systems. The instrument was translated in many languages and was based on worldwide field trials carried out in 14 centers of 11 countries in North America, in Europe, in Africa and Asia (Loranger et al. 1991, Loranger et al. 1994). Modifications were then introduced to adapt the IPDE to the DSM-III-R, and later with the DSM-IV. Because of the length of the interview, it was decided to publish it in two different modules (a DSM-IV module and an ICD-10 module). The two IPDE modules contain both a self-administered screening questionnaire and a semi-structured interview. The complete interview makes it possible to evaluate all the PDs described in the ICD-10 (WHO 1993) and the DSM-IV (APA 1994).

The Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II), (First et al. 1997) covers all 10 PDs of the DSM-IV (APA 1994), as well as Personality Disorder Not Otherwise Specified, passive aggressive personality disorder and depressive personality disorder. The first part of the SCID-II consists in a 119-items screening self-report questionnaire in which participants answer "yes" or "no" to a series of questions that reflect the diagnostic criteria of the DSM-IV PDs. "Yes" answers indicate the presence of criterion for a given PD. The SCID-II has a clinical interview version in which a trained interviewer rates positive PD diagnosis and inquires specifically about the disorders that reached the threshold for the number of items screened as positive. In clinical settings, the SCID-II can be used in at least 3 ways: (1) the clinician does his or her usual clinical interview and then uses a portion of the SCID-II to confirm and document suspected DSM-IV PD diagnoses (cluster A, B or C); (2) the SCID-II is administered as an Axis II intake procedure; (3) the SCID-II screening self-report questionnaire can be administered as a screening tool to shorten the time required for the clinician to conduct the SCID-II structured interview. Thus, when the SCID-II questionnaire is administered, the interviewer only needs to inquire, following the user's guide, about either the items screened positive on the questionnaire or about the disorders for which the number of items screened positive reached the threshold. Administration and scoring of the structured interviews require at least two hours (usually more) for trained individuals, and therefore this measure would be cumbersome for routine clinical work.

Interviews are considered to be more "valid" than are questionnaires for clinical diagnosis, but they generally require more time than questionnaires. Among the questionnaires, the Personality Diagnostic Questionnaire-4 (PDQ-4) is most frequently used in clinical practice. The PDQ-4 (Hyler 1998) is a self-report questionnaire, using a true/false response format (with the "true" statement being a pathological response) designed to assess the 10 PDs of the DSM-IV. The PDQ-4 items are specifically keyed to the diagnostic criteria for the 10 personality disorders of the DSM-IV, yielding diagnoses in accord with these criteria. Its latest version, the PDQ-4+, also takes into account the two additional PDs that were included in the Appendix of the DSM-IV (passive-aggressive and depressive disorders). The first part of the PDQ-4+ consists of a 99-item screening self-report questionnaire in which participants answer "yes" or "no" to a series of questions that reflect the diagnostic criteria of the DSM-IV PDs. Scoring the specific personality diagnosis is a two-step process. The scoring key allows the clinician to establish whether the patient is reporting any clusters of behaviours consistent with DSM-IV Axis II criteria for PDs. This key guides the clinician to use the Clinical Significance Scale of the PDQ-4 (or 4+) that will confirm or rule out the diagnosis of a PD. This requires that the interviewer ensure with the patient that: (A) There was no mistake in endorsing the items; (B) The traits have been present since age 18 or for the past several years; (C) The traits are not due primarily to Axis I conditions, such as an anxiety disorder, mood disorder, substance/alcohol abuse or physical condition, and either; (D) The traits have caused significant difficulty for the patient at home, at work (or school) or in his/her relationships, or (E) The patient himself is bothered by his/her own traits. After completion of the Clinical Significance Scale, the interviewer either confirmed or ruled out the presence of each personality disorder. The total PDQ-4 (or 4+) score can be used as an index of overall personality disturbance. A total score of 30 or more indicates a substantial likelihood that the subject has a significant personality disturbance (Hyler 1998). The PDQ-4+ (questionnaire and Clinical Significance Scale) takes 30-40 minutes to complete.

Each criteria of the DSM has a corresponding question on the PDQ and SCID, so the two questionnaires have the same number of items. For example, for criteria A(1) of Avoidant personality disorder, "avoidance of occupational activities because of fears of criticism, disapproval or rejection", the corresponding PDQ item is "I avoid working with others who may criticize me" and the SCID item is "Have you avoided jobs or tasks that involved having to deal with a lot of people ?" Since the administration of an interview demands more time and clinical experience than the administration of a self-report questionnaire, the latter may be a useful alternative during the initial screening phase. Only those cases that have indications of the presence of diagnostic criteria on the questionnaire scales would then need to be verified with an interview. After completion of the SCID interview, each criterion of personality disorder is confirmed or ruled out, resulting in two possible scores: a categorical score (yes or no PD) and a dimensional score (number of "yes" items confirmed). After completion of the Clinical Significance Scale of the PDQ (mini interview), only a categorical score is possible (yes or no PD). Given the time-consuming nature of interviewing in diagnosing personality disorders, it is important to establish the qualities of the last version of the Personality Diagnostic Questionnaire (PDQ-4+).

To date, there have been few published studies on the version DSM-IV of the PDQ-4+ and SCID-II. The PDQ and PDQ-R (the predecessors of the PDQ-4 and PDQ-4+) have been studied in clinical and non-clinical samples. Several of these studies have shown that the agreement between the DSM-III-R version questionnaire (PDQ-R) and an interview is modest or poor (Hyler et al. 1990, Hyler et al. 1992). The properties of the current version, the PDQ-4+ has only been reported by four studies. Fossati et al. (1998) found that the best total cut-off score for screening for the presence of PDs differed from the one suggested in the previous version, with a psychiatric sample. Generally, levels of agreement with interview-based diagnoses and symptoms were significant but quite modest in magnitude. Yang et al. (2000) replicated the results in psychiatric patients. When the Personality Disorders Interview (PDI-IV) was used as the diagnostic standard, agreement between the two instruments was low (Kappa values ranged from .02 to .33). Overall, the PDQ-4+ generated many false-positive diagnoses but few false-negative ones. Wilberg et al. (2000) compared the PDQ-4+ with Longitudinal, Expert, All Data (LEAD) in a clinical sample. Diagnostic agreement was poor between the two assessment methods, with Kappa values of .05 to .26 for specific PDs. Their conclusion was also that the PDQ-4+ yielded many false-positive and few false-negative diagnoses. The last study (Davison et al. 2001) studied the PDQ-4+ in a prison population. The PDQ-4+ appeared to have suitable properties to be used as a screening instrument, particularly when screening for the presence or absence of PD rather than for individual PD categories. None of these four studies used the Clinical Significance Scale of the PDQ-4+.

Hypotheses

The present study evaluated the Personality Diagnostic Questionnaire-4+ (PDQ-4+) and the Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II) in a non-clinical sample, in order to investigate the diagnostic agreement between the two assessment methods and the diagnostic efficiency of the PDQ-4+. Therefore, the principal aim was to compare the completed versions of the SCID interview and the Clinical Significance Scale of the PDQ-4+ (mini-interview). Diagnostic efficiency is defined as the relative usefulness/value of symptoms for diagnosis. In order to *improve* the validity of the PDQ-4+, we carried out the Clinical Significance Scale recommended by Hyler (1998) after a structured interview of Axis I. We also used the SCID-II screening self-report questionnaire to make it more compatible with the requirements of clinical practice. To our knowledge, it is the first report to compare the PDQ-4+ after the Clinical Significance Scale and the SCID-II structured interview. Moreover, these are the first data reported for a French-speaking population.

Material and methods

A group of 251 French undergraduate students were administered the PDQ-4+. The SCID-II was used as an external diagnostic standard for PD assessment. The SCID-II (interviewer A) was administered blind to PDQ-4+ (interviewer B). The order of assessment was counterbalanced. The self-report personality questionnaires of the two instruments were first administered. After the PDQ-4+ questionnaire was scored, the interviewer used the Clinical Significance Scale to assess each PDQ-4 diagnosis that met the threshold in order to confirm or rule out the diagnosis. Once the SCID-II self-report questionnaire was completed, the interviewer only probed questionnaire items that were answered "yes" for disorders that met the threshold. All the participants were first interviewed using a structured diagnostic interview for assessing principal DSM-IV axis I (M.I.N.I.; Mini International Neuropsychiatric Interview, French version 5); (Lecrubier et al. 1998). In doing so, the interviewers could be aware of the results of the M.I.N.I. during the administration of the Clinical Significance Scale (PDQ-4+) and the SCID-II interview.

Participants were undergraduate students at the University of Savoie (France). Of the 264 students approached, 251 (226 women and 25 men) completed the questionnaires and interviews. Their mean age was 20.77 years (range 17-49). The M.I.N.I. assessment of current Axis I disorders for these 251 subjects showed that 25% (N=63) met the criteria for one or more Axis I diagnoses: 18.72% (N=47) had anxiety disorders [generalized anxiety disorder (N= 15), panic disorder with agoraphobia (N=13), panic disorder (N= 12), social phobia (N= 4), posttraumatic stress disorder (N=2), obsessive-compulsive disorder (N= 1)], 6.77% (N=17) had substance or alcohol dependence or abuse, 6.77% (N=7) had major depression or dysthymia, 1.19% (N= 3) met the criteria for an eating disorder and 0.03% (N= 1) met the criteria for psychotic disorders. The participants were predominantly free of Axis I diagnosis (75%, N=188).

Statistical analysis

The screening properties of a test are usually expressed in terms of degree of agreement, sensitivity, specificity, and positive and negative predictive powers. The sensitivity refers to the proportion of those individuals with a particular condition who are correctly identified as positive by the instrument in question. It is the probability that a person with PDQ-4+ personality disorder will have a SCID-II diagnosis. The specificity refers to the proportion of individuals who do not suffer from the condition who are correctly identified as negative by the instrument. It is the probability that a person without PDQ-4+ personality disorder will not have SCID-II diagnosis. The positive predictive power (PPP) refers to the proportion of those who test positive who are correctly identified as suffering from the condition in question. The PPP is the probability that a subject has a PD (according to the SCID) if he or she tests positive on the PDQ. The negative predictive power (NPP) refers to the proportion of those who test negative who are correctly identified as not suffering from the condition in question. The NPP is the probability that a subject has no PD (according to the

Table 1a. *Prevalence of DSM-IV Personality Disorders according to two diagnostic "interviews" (PDQ.-4+, SCID-II)*

	Presence of at least one diagnosis on Axis II according to PDQ-4+	No diagnosis according to PDQ-4+	Chi2 et p
Presence of at least one diagnosis on Axis II according to SCID-II	19	16	24.91 P<.0001
No diagnosis according to SCID-II	36	180	

Table 1b. *Prevalence of DSM-IV Personality Disorders according to two diagnostic "interviews" (PDQ-4+, SCID-II)*

Personality disorder (PD)	Interview of the SCID-II	Clinical Significance Scale of the PDQ-4+
Avoidant	5 (1.99 %)	30 (11.95%)
Dependent	2 (0.79 %)	5 (1.99%)
Obsessive-Compulsive	11 (4.38 %)	14 (5.57%)
Passive-Aggressive	3 (1.19 %)	3 (1.19%)
Depressive	9 (3.58 %)	16 (6.37%)
Paranoid	9 (3.58 %)	17 (6.77%)
Schizotypal	0	3 (1.19%)
Schizoid	0	4 (1.59%)
Histrionic	0	3 (1.19%)
Narcissistic	2 (0.79 %)	1 (0.39%)
Borderline	7 (2.78 %)	8 (3.18%)
Antisocial	9 (3.58 %)	4 (1.59%)

SCID) if, according to the PDQ, he or she has no PD. The positive and negative predictive powers describe the success of the instrument in a population. The sensitivity and specificity depend on the degree of overlap between cases and non-cases, and the balance between them will depend on the cut-off used. Diagnostic agreement was measured by the Kappa value (Landis and Koch 1977), and diagnostic efficiency was evaluated by sensitivity, specificity, positive predictive power (PPP), and negative predictive power (NPP). Kappa statistic measures the agreement of two scales with correction for chance factors (Fleiss 1981, Davis and Fleiss 1982). Fleiss (1981) characterized kappas over 0.75 as excellent, from 0.40 to 0.70 as fair to good and below 0.40 as poor. Values for diagnostic efficiency indices were calculated according to the formulas shown in Pina et al. (2002). In accordance with Pina and colleagues, conditional probability (i.e. sensitivity, specificity, PPP, NPP) values ranging from .00 to .29 were considered to be low, values ranging from .30 to .69 were considered moderate, and values ranging from .70 to 1.00 were considered to be high (Pina et al. 2002). All other data analyses were performed with the statistical program SPSS.

Results

The prevalence rates of DSM-IV PDs as diagnosed by the self-report of the SCID-II was 64 % (n= 161) and of the PDQ-4+ was 62 % (n= 155). More than half of the participants were identified as having one or more PDs when using the two questionnaires only. Table 1a shows prevalence rates of DSM-IV PDs as diagnosed by the two instruments, after the use of the Clinical Significance Scale of the PDQ-4+ or the use of the structured interview of the SCID-II. Following the interview of the SCID-II, 35 students (13.94%) were identified as having one or more PDs. Following the Clinical Significance Scale of the PDQ-4+, 55 students (21.19%) were identified as having one or more PDs. The Clinical Significance Scale of the PDQ-4+ diagnosed more subjects as having at least one PD than did the SCID-II interview. Using the SCID-II as a criterion (Table 1b), all the PD diagnostic categories had a prevalence rate of less than 5 %. Following the interview with the SCID-II, no subject met the criteria for schizotypal, schizoid or histrionic personality disorders. Finally, the mean number of PDs diagnosed in patients with at least one diagnosis was 1.92 for the

Table 2. *Prevalence of DSM-IV Axis I and II disorders according to the M.I.N.I. and the two diagnostic instruments (PDQ-4+, SCID-II)*

M.I.N.I.	Presence of at least one diagnosis on Axis II according to SCID-II	Presence of at least one diagnosis on Axis II according to PDQ-4+ and SCID-II	Presence of at least one diagnosis on Axis II according to PDQ-4+	Chi2 and p
Positive (presence of at least one Axis I diagnosis) N = 21	9	13	12	6.71 P = .03
Negative (no diagnosis on Axis I) N = 31	7	6	24	

Table 3. *Kappa, Sensitivity, Specificity, and Positive and Negative Predictive Power of the PDQ-4+ and the SCID-II*

Personality disorder	Kappa	Sensitivity	Specificity	PPP	NPP
Avoidant	0.26	1	0.89	0.16	1
Dependent	0.39	0.50	0.99	0.33	0.99
Obsessive- Compulsive	0.45	0.54	0.96	0.42	0.98
Passive- Aggressive	0.32	0.33	0.99	0.33	0.99
Depressive	0.28	0.44	0.95	0.25	0.97
Paranoid	0.03	0.11	0.93	0.05	0.96
Narcissistic	-0.005	0	0.99	0	0.99
Borderline	0.51	0.57	0.98	0.50	0.98
Antisocial	0.29	0.22	0.99	0.50	0.97
Any	0.30	0.54	0.83	0.34	0.91

PPP: Positive Predictive Power
NPP: Negative Predictive Power

PDQ-4+ and 1.62 for the SCID-II.

A total of 180 subjects (72%) were free of PDs as diagnosed by the two instruments, of whom 29 subjects (12%) met the criteria for one or more Axis I diagnoses. Table 2 shows prevalence rates of DSM-IV as diagnosed by the M.I.N.I. and the two instruments (SCID-II and PDQ-4+). Nineteen students (7.3%) met the criteria for one or more Axis II disorder as diagnosed by the SCID-II and the PDQ-4+. The PDQ-4+ diagnosed more subjects (N=36; 15%) as having at least one PD than did the SCID-II (N= 16; 5.7%). The PDQ-4+ diagnosed significantly more subjects with a negative score on the M.I.N.I. (N = 24) than did the SCID-II.

Table 3 shows the chance-corrected agreement (Kappa) between the two instruments, sensitivity, specificity, and positive negative predictive powers of the PDQ-4+. This table shows that agreement varied from 0.51 for borderline to 0.00 for paranoid and narcissistic PDs. The agreement between the instruments for any PD regardless of subtype was slight (0.30). The agreement between the SCID-II and PDQ-4+ was moderate for 2 PDs (borderline, obsessive-compulsive) and slight for 5 PDs (dependant, passive-aggressive, antisocial, depressive and avoidant). The agreement was poor for 2 PDS (paranoid and narcissistic). Four PDs emerged as relatively efficient with positive predictive power in the moderate range (borderline, obsessive-compulsive, dependent and passive-aggressive) identifying a moderate proportion of the students (sensitivity in the moderate range). One PD (antisocial) was inadequate to identify the students (sensitivity = 0.22). Four PDs had lower positive predictive power (avoidant, depressive, paranoid, narcissistic). These four PDs generated many false positive diagnoses but very few false negatives. All the negative predictive powers were in the high range (0.91 to 1) and the specificities were also in the high range. Generally, the probability to have no PDQ-4+ diagnosis if no SCID-II diagnosis was present was excellent.

Discussion

The aim of the research was to examine the screening properties of the PDQ-4+ in a French-speaking non-clinical sample. To do so, we administered the Structured Clinical Interview for DSM-IV Personality Disorder (SCID-II) and a structured diagnostic interview processing principal DSM-IV Axis I disorders (M.I.N.I.; Lecrubier et al. 1998). The M.I.N.I. assessment showed that 25% (N=63) of participants met the criteria for one or more current Axis I diagnoses. Anxiety disorders were the most frequent [generalized anxiety disorder (N= 15), panic disorder with agoraphobia (N=13), panic disorder (N= 12), social phobia (N= 4), posttraumatic disorder (N=2), obsessive-compulsive disorder (N= 1)], followed by substance or alcohol dependence disorders, or substance or alcohol abuse disorders. We compared the PDQ-4+ to the SCID-II first on the self-report questionnaires and second on the interviews (Clinical Significance Scale of the PDQ-4 or SCID-II structured interview). At the time of comparison of the self-report questionnaires of the PDQ-4+ and SCID-II, more than half of the participants were identified as having one or more PDs. The questionnaire versions of the PDQ-4+ and the SCID-II seem inadequate to be used in an undergraduate population. The two self-reports reveal a substantial number of false positives. Following the Clinical Significance Scale of the PDQ-4+, 55 students (21.19%) were identified as having one or more PDs and only 35 students (13.94%) were identified as having one or more PDs following the interview of the SCID-II. The awareness of the diagnoses on Axis I and an interview (Clinical Significance Scale or structured interview) significantly affects the number of individuals meeting the criteria for a PD disorder. The PDQ-4+ Clinical Significance Scale diagnosed more individuals as having at least one PD than did the SCID-II interview. However, in terms of time spent with the participants, the Clinical Significance Scale is quicker to use than the SCID-II structured interview and thus more adaptable to clinical practice. Finally, 180 subjects (72%) were free of PDs as diagnosed by the two instruments, but 29 subjects (12%) met the criteria for one or more Axis I diagnoses. Nineteen students (7.30%) met the criteria for one or more Axis II diagnosed by the SCID-II and PDQ-4+. The distribution of the disorders of Axis I is not comparable in the group of participants meeting the criteria for a PD with the PDQ-4+ and those meeting criteria for a PD with the SCID-II. The PDQ-4+ diagnosed significantly more subjects who had a negative score on the M.I.N.I. than did the SCID-II.

The prevalence of PDs in the general population varies from 9% to 14.8% (Ekselius et al. 2001, Samuels et al. 2002, Grant et al. 2004). In our study, 35 undergraduate students (13.94%) were classified as having one or more PDs following the interview with the SCID-II. In a study of Nelson-Gray et al. (2004), which was also conducted with undergraduate students, 30% of the subjects met criteria for at least one PD following SCID-II interview (DSM-III-R). In an exploratory study of the PDQ-4+ on a group of French undergraduate students, which was conducted without the use of a structured interview for controlling for Axis I comorbid disorders (Bouvard and Cosma 2008), 27.13% of the subjects were identified as having one or more PDs. In the present study, 21.19% were identified as having one or more PDs following the Clinical Significance Scale of the PDQ-4+ and a structured interview of Axis I. The proportion of young adults and undergraduate students meeting criteria for one or several PDs is more important than the proportion of subjects from the general population, whatever was the instrument used for assessment (interview SCID or questionnaire PDQ). This is congruent with the results of a study conducted by Ekselius et al. (2001) that showed that the age (between 18 and 34 years) and the professional status (students or not employed) were categories where the number of PDs was most frequent.

The agreement between the instruments for any PD regardless of subtype was small (0.30). The agreement between the SCID-II and PDQ-4+ was moderate for 2 PDs [borderline (0.51), obsessive-compulsive (0.45)] and small for 5 PDs [(dependent (0.39), passive aggressive (0.32), antisocial (0.29), depressive (0.28) and avoidant (0.26)]. The agreement was poor for only 2 PDS (paranoid and narcissistic). It seems that the use of the Clinical Significance Scale improves the agreement between the PDQ-4+ and the SCID-II. Diagnostic agreement was poor between the PDQ-4+ questionnaire and Longitudinal Expert All Data (LEAD), with K values of 0.05 to 0.26 for specific PDs (Wilberg et al. 2000). Low to modest agreement between the PDQ-4+ questionnaire and PDI-IV (an interview) was observed for PD evaluations in a Chinese population (Yang et al. 2000). In a prison population (Davison et al. 2001), the agreement between the instruments for any PD subtype was 0.47. Only antisocial and borderline personality disorders showed better agreement. The agreement was moderate for 3 other PDs [avoidant (0.40), paranoid (0.35), depressive (0.31)] and low for the other PDs. In a mixed psychiatric sample (Fossati et al. 1998), the agreement between the instruments for any PD regardless of subtype was poor (0.18). The chance-corrected agreement was in the poor range [depressive (0.03) to antisocial (0.28)]. Only 3 PDs (antisocial, narcissistic and dependent) had slight agreement between the two instruments. In general, diagnostic categories with a prevalence rate less than 5 % were excluded from agreement and efficiency analyses, as Kappa, PPP and NPP are prevalence-dependant (Wilberg et al. 2000). However, it was difficult to apply this rule in our study, given that the prevalence rates varied from 0.79 % (dependant and narcissistic PDs) to 4.38 % (obsessive compulsive PD).

Four PDs emerged as relatively efficient with positive predictive power in the moderate range (borderline, obsessive-compulsive, dependent and passive-aggressive) identifying a moderate proportion of the students (sensitivity in the moderate range). One PD (antisocial) was inadequate to identify the students (sensitivity = 0.22). Four PDs had lower positive predictive power (avoidant, depressive, paranoid, narcissistic). These four PDs generated many false positive diagnoses but very few false negatives. All the negative predictive powers were in the high range (0.91 to 1) and the specificities were also in the high range. In a Chinese population (Yang et al. 2000), only 3 PDs

had a moderate positive predictive power associated with sensitivity in the high-moderate range (avoidant, antisocial and depressive). All the other PDs had lower positive predictive power, but the probability to have no PDQ4+ diagnosis if no SCID diagnosis was excellent. Many false positive but few false negative diagnoses for PDQ-4+ scales were also present in Fossati's results (1998). A tendency towards the improvement of the indices of diagnosis efficiency seems to appear when using the Clinical Significance Scale of the PDQ-4+.

The principal limitation of our study was the small number of subjects; therefore, it will be necessary to replicate with a much larger sample. The use of the SCID-II as a criterion also limits the generalizability of the results. Finally, the prevalence of PDs was less than 5 %, so it would be informative to investigate whether these findings would replicate in a clinical sample.

In summary, according to these results, the two self-reports (PDQ-4+ and SCID-II) are inadequate to diagnose PDs in an undergraduate sample. However, an aim of self-report questionnaires is to generate few false-negative diagnoses and more false-positives. The use of the Clinical Significance Scale of the PDQ4+ and the knowledge of Axis I diagnoses, improve the agreement and the diagnostic efficiency between an interview and a questionnaire but the indices are modest to moderate. The administration of the Clinical Significance Scale is quicker than the structured interview of the SCID-II. When the SCID-II was used as the standard diagnostic tool, the PDQ-4+ showed higher negative predictive power than positive predictive power. The lack of agreement upon the "gold standard" instrument for PD diagnoses (Yeung et al. 1993) limits the generalizability of these findings. A short screening test for DSM-IV PDs would be extremely helpful, but some adjustments should be carried out before PDQ-4+ can be safely used as a screening instrument.

Acknowledgments

This research was support by a grant from the Pays de Savoie (BQR 2006). The authors thank Poquin D. for his assistance in the supervision of the interviewers and Carton M., Chambard S., Fourcade C., Largès D., Le Bihan D., Marc C., Mariaz R., Neyroud S., Perez C., and Sigel L. for their involvement in the collection of data.

References

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of mental Disorders*, 4th ed. American Psychiatric Press, Washington, DC.

Bouvard M, Cosma P (2008). Etude exploratoire d'un questionnaire sur les troubles de la personnalité. *L'Encéphale* 34, 517-525.

Cloos JM, Pull-Erpelding MC, Pull CB (2006). *Entretiens diagnostiques structurés en psychiatrie*. Encyclopédie Médico Chirurgicales, Elsevier SAS, Psychiatrie, 37-102-B-10, Paris.

Davis M, Fleiss JL (1982). Measuring agreement for multinomial data. *Biometrics* 38, 1047-1051.

Davison S, Leese M, Taylor PJ (2001). Examination of the screening properties of the personality diagnostic questionnaire 4 + (PDQ-4 +) in a prison population. *Journal of Personality Disorders* 15, 180-194.

Ekselius L, Tillfors M, Furmark T, Fredrikson M (2001). Personality disorders in the general population: DSM-IV and ICD-10 defined prevalence as related to sociodemographic profile. *Personality and Individual Differences* 30, 311-320.

Fava M, Bouffides E, Fava JA, McCarthy MK, Steingard RJ, Rosenbaum JF (1994). Personality disorder comorbidity with major depression and response to Fluoxetine treatment. *Psychotherapy and Psychosomatics* 62, 160-167.

First MB, Gibbon M, Spitzer RL, Williams JB, Benjamin LS (1997*). User's guide for the Structured Clinical Interview for DSM-IV axis II Personality Disorder*s. American Psychiatric Press, Washington, DC.

Fleiss JL (1981). *Statistical methods for rates and proportions.* Wiley, New York, 2ed.

Fossati A, Maffei C, Bagnato M, Donati D, Donini M, Fiorilli M, Novella L, Ansoldi M (1998). Brief communication: criterion validity of the Personality Diagnostic Questionnaire-4 plus (PDQ-4 +) in a mixed psychiatric sample. *Journal of Personality Disorders* 12, 172-178.

Grant BF, Hasin DS, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP (2004). Prevalence, correlates, and disability of Personality Disorders in the United States: results from the National Epidemiological Survey on Alcohol and related conditions. *Journal of Clinical Psychiatry* 65, 948-958.

Hyler SE (1998*). PDQ-4 and PDQ-4+, instructions for use.* New York State Psychiatric Institute, New York.

Hyler SE, Skodol AE, Kellman HD, Oldham J, Rosnick L (1990). The validity of the Personality Diagnostic Questionnaire Revised: a comparison with two structured interviews. *American Journal of Psychiatry* 147, 1043-1048.

Hyler SE, Skodol AE, Oldham J, Kellman HD, Doidge N (1992). The validity of the Personality Disorder Questionnaire Revised: a replication in an out-patient sample. *Comprehensive Psychiatry* 33, 73-77.

Landis JR, Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159-174.

Lecrubier Y, Hergueta T, Sheehan D (1998*). M.I.N.I., Mini International Neuropsychiatric Interview*, French version 5.0.0. Inserm U302, Hôpital de la Salpétrière, Paris.

Loranger AW, Sartorius N, Andreoli A, Berger P, Buchheim P, Channabasavanna SM, Cold B, Dahl A, Diekstra RFW, Ferguson B, Jacobsberg LB, Mombour W, Pull C, Ono Y, Régier DA (1994). The International Personality Disorder Examination. The World Health Organisation/alcohol, drug abuse and mental health administration international pilot study of personality disorders. *Archives of General Psychiatry* 51, 215-224.

Loranger AW, Hirschfeld RMA, Sartorius N, Regier DA (1991). The WHO/ ADAMHA international pilot study of personality disorders: background and purpose. *Journal of Personality Disorders,* 5, 296-306.

McKay D, Neziroglu F, Todaro J, Yaryura Tobias JA (1996). Changes in personality disorders following behavior for compulsive disorder. *Journal of Anxiety Disorders* 10, 47-57.

Nelson-Gray RO, Huprich SK, Kissling GE, Ketchum K (2004). A preliminary examination of Beck's cognitive theory of personality disorders in undergraduate analogues. *Personality and Individual Differences* 36, 219-233.

Petot JP (2007). Consensus et divergences sur l'évaluation de la personnalité : approche catégorielle ou approche dimensionnelle? Conférence à l'Association Française et de Recherche des Troubles de la personnalité (AFERTP), Paris.

Pfohl B, Blum M, Zimmerman M (1995*). The Structured Interview for DSM-IV personality disorders (SIDP-IV).*

University of Iowa, Iowa City.
Pfohl B, Stangl D, Zimmerman M (1982). *The Structured Interview for DSM-III personality disorders (SIPD).* University of Iowa, Iowa City.
Pina AA, Silverman WK, Alfano CA, Saavreda LM (2002). Diagnostic efficiency of symptoms in the diagnosis of DSM-IV: generalized anxiety disorder in youth. *Journal of Child Psychology and Psychiatry* 43, 959-967.
Ricciardi JN, Baer L, Jenike JA, Fisher SC, Sholtz D, Buttolph ML (1992). Changes of DSM-III-R Axis II diagnoses following treatment of obsessive compulsive disorders. *American Journal of Psychiatry* 149, 829-831.
Samuels J, Eaton WW, Bienvenu III J, Brown C, Costa PT, Nestadt G (2002). Prevalence and correlates of personality disorders in a community sample. *British Journal of Psychiatry* 180, 536-542.
Stuart S, Simons AD, Thase M, Pilkonis P (1991). Are personality assessments valid in acute major depression? *Journal of Affective Disorders* 24, 281-290.
Wilberg T, Dammen T, Friis S (2000). Comparing Personality Diagnostic Questionnaire 4+ with Longitudinal, Expert, All Data (LEAD) Standard Diagnoses in a sample with a high prevalence of Axis I and Axis II disorders. *Comprehensive Psychiatry* 41, 295-302.
World Health Organization (WHO) (1993). *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research*. World Health Organization, Genève.
Yang J, McCrae RR, Costa PT, Yao S, Dai X, Cai T, Gao B (2000). The cross-cultural generalizability of axis II constructs: an evaluation of two personality disorder assessment instruments in the people's republic of China. *Journal of Personality Disorders* 14, 249-263.
Yeung AS, Lyons MJ, Waternaux CM, Faraone SV, Tsuang MT (1993). Empirical determination of the thresholds for case identification: validation of the personality diagnostic questionnaire-revised. *Comprehensive Psychiatry* 34, 384-391.