The concurrent validity of a Web-based self-report assessment for personality disorders

Personality and Mental Health
Personality and Mental Health
13: 53–62 (2019-02)
Published online 27 December 2018 in Wiley Online Library
(wileyonlinelibrary.com) DOI 10.1002/pmh.1438

The concurrent validity of a Web-based self-report assessment for personality disorders

FARID CHAKHSSI[1,†], IES DIJKSMAN[2†], MARA L. VELMANS[1], J. MONIQUE ZOET[1], JEANINE M. OOSTENDORP[1], GEERT-JAN DINANT[2] AND MARK SPIGT[2], [1]Scelta, Apeldoorn, The Netherlands; [2]Maastricht University, Maastricht, The Netherlands

ABSTRACT

Incorporating online tools in clinical practice could help improve routine assessments of personality disorders and their co-occurring clinical disorders. TeleScreen, a Web-based self-report questionnaire for DSM-IV disorders, has not yet been compared with well-validated structured interviews for clinical and personality disorders. Patients with personality disorders (n = 89) were assessed with TeleScreen and independently interviewed with Structured Clinical Interview for DSM-IV Axis-I disorders (SCID-I) and personality disorders (SCID-II). The concurrent validity was examined using sensitivity, specificity, and positive and negative predictive values. Five personality disorders had a prevalence rate higher than 4% and could be examined in the analyses: borderline, obsessive–compulsive, dependent, avoidant and paranoid personality disorders. TeleScreen showed moderate to good validity for borderline personality disorder and obsessive–compulsive personality disorder but suboptimal validity for the dependent, avoidant and paranoid personality disorders. Clinical disorders showed moderate to good values, except for social phobia, dysthymia and eating disorders. These findings provide preliminary evidence for the concurrent validity of TeleScreen for some personality disorders, such as the borderline personality disorder, and pave the way for larger studies to confirm these results. © 2018 John Wiley & Sons, Ltd.

Introduction

Personality disorders (PDs) are among the most frequent psychiatric disorders treated by mental health professionals. About a quarter of patients in primary care and 45% in psychiatric outpatient settings meet the criteria for one or more PDs.[1,2] PDs are associated with a wide range of clinical disorders,[3,4] increased mortality,[5] poor treatment outcome, extensive health-care use,[6,7] high burden of disease[8] and high economic burden on society.[9] Patients with PDs that remain undetected in clinical practice might be given treatments that are ineffective or harmful, and therefore, timely and accurate assessment of PDs is important.[10]

To assess PDs and their co-occurring clinical disorders, the Structured Clinical Interview for *DSM-IV* Axis I disorders (SCID-I[11]) and for Axis II disorders (SCID-II[12]) are widely regarded as valid and reliable measures.[13,14] However, PDs are not routinely assessed in clinical practice.[15–19] One of the most important reasons is the absence of quick and valid assessment instruments for primary and specialized mental health-care settings.[10] For example, semi-structured assessments for PDs and their co-occurring clinical disorders using the SCID-I and SCID-II usually take up to 2–6 h.[13] Another

†These two authors contributed equally to this manuscript.
Scelta is a Dutch psychotherapy center specialized in the treatment of personality disorders.

important reason is the presenting co-occurring clinical disorders that may mask the personality pathology. Patients with PDs are likely to have co-morbid disorders, such as depressive or anxiety disorders, that may dominate the clinical picture.[4,10] To improve the assessment of PDs and their co-morbid clinical disorders in clinical practice, there is a need for novel methods that are less time-consuming and easy to use for both patient and professional. The majority of patients with PDs rarely present themselves without one or more co-morbid clinical disorders,[3,4] and therefore, it is important to examine the concurrent validity of TeleScreen for these highly co-morbid clinical disorders as well.

Incorporating web-based tools, such as a web-based self-report version of the SCID-II[20] or a web-based self-report questionnaire for *DSM-IV* clinical disorders and PDs,[21] into routine intake procedures in mental health care could help improve routine assessments of PDs. In contrast to most of the widely used (self-report) outcome measures, these web-based self-report questionnaires do not require meetings with clinicians (to fill out questionnaires), nor do they require clinician time for scoring and interpretation, as they result in automated preliminary *DSM* classifications of clinical and/or PDs. This screening would be highly valuable in clinical practice for making decisions about further more in-depth assessment (or in research studies to indicate which patients are likely to meet the eligibility criteria). Also, the web-based tools can be used across all mental health settings, including those that may not have the clinical resources for extensive diagnostic assessment, such as primary care. Previous studies have indicated that clinicians preferred using these web-based tools above semi-structured interviews, that it reduced time needed by clinicians substantially and that patients were satisfied with the web-based version.[20,21]

Previously, the validity of a web-based self-report questionnaire for *DSM-IV* clinical disorders and PDs, called TeleScreen,[22] was examined in a sample of 675 adult primary care patients. The results of TeleScreen, in terms of a *DSM-IV* classification, were compared with the clinical judgement of psychologists who had consulted the patients by phone. Results showed modest to high concurrent validity for the *DSM-IV* classification of clinical disorders and PDs. Although these results were encouraging, the concurrent validity of TeleScreen with valid and reliable measures of assessing PDs and their co-morbid clinical disorders, that is, semi-structured clinical interviews, remains inconclusive.

The current study compared the classifications of TeleScreen, with independent semi-structured assessments of trained raters using the SCID-I and SCID-II in a sample of patients with PDs.

Method

Design

We examined the concurrent validity of TeleScreen when compared with the SCID-I and SCID-II, among patients referred for specialized treatment of personality pathology. The first 31 patients were recorded on digital voice recorders, in order to investigate the interrater reliability between the raters of the SCID-I and SCID-II.

Participants

Participants were recruited at two inpatient and one day-hospital unit from a Dutch psychotherapy centre specializing in the treatment of PDs, called Scelta. All patients were referred for specialized treatment of PD pathology. All patients who were consecutively admitted between June 2016 and July 2017 were invited to participate in the study. Excluded for this study were patients with insufficient understanding of the Dutch language, who showed signs of alcohol or drugs intoxication or who showed acute psychotic symptoms during SCIDs assessments.

Raters

The group of raters consisted of 10 first raters and two second raters. All raters were bachelor-level or master-level psychologists with a mean age of 24.4 years (standard deviation (SD) = 2.2). The

DOI: 10.1002/pmh

raters were trained in administrating the SCID-I and SCID-II by two of the co-authors (JMO and JMZ), who had extensive experience in SCIDs assessments (average years of experience is approximately 10 years). The 1-day training consisted of theory on SCID assessments, scoring case material of patients with different PDs previously assessed by the trainers and role-playing interview situations. In addition, the raters accompanied the trainers for one time when performing SCID-I and SCID-II assessments, after which the raters performed at least two SCID-I and SCID-II assessments accompanied by the trainers. Finally, raters received feedback on all consecutive SCIDs assessments from the SCID trainers.

Materials

The Structured Clinical Interview for DSM-IV Disorders (SCID-I). The SCID-I[11] is a diagnostic interview designed to assess the most common clinical disorders as described in the *DSM-IV*.[23] Test–retest and interrater reliability for the SCID-I is high.[13] The Structured Clinical Interview for DSM-IV personality disorders (SCID-II)[12] consists of modules for all 11 PDs in which questions are grouped per PD. The assessment of each PD starts with an open question, after which further explanations are asked. For each disorder, all criteria are addressed and scored on a three-point scale (1 = omission or the criterion is incorrect; 2 = criterion is doubtful; 3 = criterion is present or correct). Previous research demonstrated that the SCID-II yields reliable PD assessments.[13]

TeleScreen. TeleScreen was previously described in two other studies.[21,22] TeleScreen is an automated web-based adaptive self-report questionnaire to assess the likelihood of *DSM-IV* clinical disorders and PDs based on the SCID-I,[11] MINI-International Neuropsychiatric Interview[24] version 5.0.0 and SCID-II.[12] For each clinical disorder, there are entrance questions and follow-up questions. To improve efficiency, follow-up questions are skipped according to the answers to the entrance questions. For example, when participants answer negatively on one of the core symptoms of major depressive disorder (i.e. depressed mood or a pronounced loss of interest or pleasure), the follow-up questions for major depressive disorder are automatically skipped, and the participants are classified as not fulfilling the criteria for major depression. To assess PDs, TeleScreen presents a fixed set of 125 statements representing *DSM-IV* PD traits that can be scored with 'correct' or 'incorrect'. For example, participants are asked to answer correct or incorrect to the statement *'Throughout the largest extent of your life, but especially the last 5 years: as soon I sense that someone will abandon me, I will take extreme precautions to prevent that from happening. (For instance, by calling the person excessively, become overly nice, or beg the other person not to leave me)'*, which directly corresponds with *DSM-IV* borderline PD criterion regarding frantic efforts to avoid real or imagined abandonment. For each PD, all PD traits are assessed. When a patient fulfils the required number of PD traits based on the *DSM-IV*, the presence of a PD is indicated. In addition, patients' demographics, somatic complaints, psychosocial problems (e.g. problems with primary support group, occupational or economic problems) and Global Assessment of Functioning score (GAF-score) are assessed, by using both open and closed questions, to provide a *DSM-IV* multi-axial classification.

Procedure

Approval for the study was obtained from the local institutional research committee and the ethical board of the University of Twente (#16058). After participants provided written informed consent, the first 50 patients were administered the SCID-I and the SCID-II in two or three 2-h sessions. Three weeks after SCIDs assessments, participants received a link to TeleScreen via email. For the last 39 patients, the assessment of SCIDs and TeleScreen was reversed. These patients first received a link to TeleScreen, and 3 weeks later, SCIDs assessments took place. In the first 31 cases,

DOI: 10.1002/pmh

both SCIDs were administrated face to face by the first raters and recorded on digital voice recorders. The second raters listened to recorded interviews and assessed SCIDs criteria independent and blind of the first raters' scores.

Statistical analysis

Interrater reliability. Cohen's κ, their 95% confidence intervals and percentage agreement were calculated to assess agreement between the raters of the categorical judgements (present/absent) of SCIDs clinical disorders and PDs. According to Fleiss,[25] Cohen's κ values below 0.40 indicate poor agreement, values between 0.40 and 0.75 as fair to good and values above 0.75 as excellent.

Concurrent validity. The categorical presence (absent/present) of *DSM-IV* clinical disorders and PDs according to SCIDs interviews were compared with categorical presence of the *DSM-IV* clinical disorders and PDs based on TeleScreen. Sensitivity, specificity, positive predictive values (PPVs), and negative predictive values (NPVs) were calculated to investigate the concurrent validity between SCIDs and TeleScreen. Sensitivity is the ability of TeleScreen to correctly identify patients who, according to the SCIDs, meet the criteria of a clinical and/or PD. Specificity is the ability of TeleScreen to correctly identify those who, according to the SCIDs, do not meet the criteria for a clinical and/or PD. PPVs are the proportion of patients who, according to TeleScreen, met the criteria for a disorder and who were actually classified as having the disorder on the basis of the SCIDs. NPVs are the proportion of patients who did not meet the criteria for a disorder according to TeleScreen and who were actually classified as not having the disorder on the basis of SCIDs. Analyses were only performed for categories of disorders when they had a prevalence rate of more than 4%, according to SCIDs, because the results, especially the PPVs and NPVs, could be influenced by the low base rate of the disorders e.g.[26]

Results

TeleScreen assessment

In total, 95 patients were invited to participate, five patients refused and one patient was excluded because he or she met an exclusion criterion (i.e. active psychotic symptoms during SCIDs assessments). The remaining 89 patients who agreed to participate included 72 (80.9%) women and 17 (19.1%) men, with an average age of 31.8 years (SD = 10.8; range 18–59). In the current sample, the average length of TeleScreen assessments was 52 min (SD = 20 min, range 22–120 min). The mean number of (login-) sessions was 2.3 (SD = 1.8, range 1–11 sessions).

Interrater reliability

The interrater reliability between the raters of SCIDs for clinical disorders and PDs is shown in Table 1. For all classifications, fair to good or excellent agreement was found (Table 1). The percentage agreement for all clinical disorders and PDs was high.

Concurrent validity of the personality disorders

The classifications according to TeleScreen and SCID-II are shown in Table 2. Most prevalent PDs according to the SCID-II in the current sample were borderline (40.5%), avoidant (40.5%) and obsessive compulsive PDs (25.8%). Only five PDs were prevalent with a rate of more than 4% according to the SCID-II and were examined in the analyses: borderline, obsessive–compulsive, dependent, avoidant and paranoid PDs. Sensitivity was moderate to good (range 0.70–0.89) for the prevalent PDs with the exception of dependent PD (0.38) and paranoid PD (0.50). Specificity was also moderate to good (range 0.71–0.82) for the prevalent PDs with the exception of avoidant PD (0.52) and paranoid PD (0.59). The classification of the borderline PD by TeleScreen showed the highest sensitivity and specificity, 0.86 and 0.71 respectively.

DOI: 10.1002/pmh

Table 1: Agreement of SCIDs clinical disorders (n = 31) and personality disorders (n = 30)

DSM-IV categories	Total classifications Rater 1	Total classifications Rater 2	Cohen's κ (95% CI)	Percentage agreement
Major depression	26	22	0.64 (0.33–0.95)	0.87
Post-traumatic stress disorder	15	15	0.74 (0.50–0.98)	0.87
Panic disorder	4	5	0.87 (0.62–1.12)	0.97
Substance disorder	3	4	0.84 (0.53–1.14)	0.97
Eating disorder	11	9	0.71 (0.44–0.97)	0.87
Paranoid PD	6	3	0.62 (0.24–1.00)	0.90
Schizoid PD	—	—	—	—
Schizotypal PD	—	—	—	—
Antisocial PD	—	—	—	—
Borderline PD	9	8	0.92 (0.76–1.08)	0.97
Histrionic PD	—	—	—	—
Narcissistic PD	—	—	—	—
Avoidant PD	9	9	0.84 (0.63–1.05)	0.93
Dependent PD	—	—	—	—
Obsessive–compulsive PD	8	8	1.00 (1.00–1.00)	1.00

Note: analyses were only performed for categories of disorders with three or more observations. PD, Personality disorder.

Table 2: Concurrent validity of TeleScreen compared with SCID-II for personality disorders (n = 89)

Personality disorder	Total TS	Total SCID	TP	FP	FN	TN	Sens	Spec	PPV	NPV
Paranoid PD[1]	37	6	3	34	3	48	0.50	0.59	0.08	0.94
Schizoid PD[1]	15	0	0	15	0	73	—	—	—	—
Schizotypal PD[1]	12	1	1	11	0	76	—	—	—	—
Antisocial PD[1]	13	1	1	12	0	75	—	—	—	—
Borderline PD[1]	46	36	31	15	5	37	0.86	0.71	0.67	0.88
Histrionic PD[1]	2	0	0	2	0	86	—	—	—	—
Narcissistic PD[1]	2	2	1	1	1	85	—	—	—	—
Avoidant PD[1]	57	36	32	25	4	27	0.89	0.52	0.56	0.87
Dependent PD[1]	18	8	3	15	5	65	0.38	0.81	0.17	0.93
Obsessive–compulsive PD[1]	28	23	16	12	7	53	0.70	0.82	0.57	0.88

Note: analyses were only performed for categories of disorders with a prevalence rate higher than 4% according to Structured Clinical Interview for *DSM-IV* Axis II disorders (SCID-II); PD, Personality disorder; Total TS, total number of TeleScreen classifications; Total SCID, total number of SCID classifications; TP, true positive; FP, false positive; FN, false negative; TN, true negative; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; PD, personality disorders.
[1]There was one missing value.

The PPVs were of poor strength for the paranoid PD (0.08) and dependent PD (0.17) and moderate for the avoidant PD (0.56), obsessive–compulsive PD (0.57) and borderline PD (0.67), corresponding with a high proportion of false positives (Table 2). In other words, the lower the PPVs, the more likely patients were to meet the criteria for a PD based on TeleScreen than based

DOI: 10.1002/pmh

on the SCID-II. On the other hand, the NPVs were high for all the PDs prevalent in the sample (≥0.87). So when patients did not meet the criteria for a PD on TeleScreen, they also did not meet the criteria for that specific PD on the SCID-II, with a high rate of accordance.

Concurrent validity of the clinical disorders

The classifications of clinical disorders by TeleScreen and based on the SCID-I are shown in Table 3. Most prevalent disorders according to the SCID-I in the sample were major depression (67.4%) and post-traumatic stress disorder (56.2%). Sensitivity for the clinical disorders was moderate to good (range 0.67–0.80), with the exception of social phobia (sensitivity 0.21), dysthymia (sensitivity 0.20), eating disorder (sensitivity 0.29), specific phobia (sensitivity 0.43) and other drug abuse/dependence (sensitivity 0.50). Specificity was moderate to high for the clinical disorders (range 0.77–0.94), with the exception of major depression (specificity 0.48).

The PPVs were of poor strength for dysthymia (0.08) ranging to alcohol abuse/dependence (0.58) and moderate to good strength for major depression (0.76) ranging to post-traumatic stress disorder (0.85). The NPVs were high for all clinical disorders (≥0.85) with the exception of major depression (0.54), post-traumatic stress disorder (0.69) and eating disorder (0.78). The classification of post-traumatic stress disorder by TeleScreen showed the highest PPV and NPV of 0.85 and 0.69 respectively.

Discussion

We examined the concurrent validity of TeleScreen, a Web-based adaptive self-report version for *DSM-IV* clinical disorders and PDs, compared with independent semi-structured assessments using the SCID-I and SCID-II in a sample of patients with PDs. TeleScreen showed moderate to good validity for the borderline and obsessive–compulsive PD but suboptimal validity for the other PDs prevalent in the sample: the avoidant, dependent and paranoid PDs. However, the moderate to good validity for the classification of the borderline PD is a remarkable finding and comparable with that of previous studies on convergent validity of specifically developed self-report measures with semi-structured interviews for borderline personality pathology.[27–29]

Table 3: Concurrent validity of TeleScreen compared with SCID-I for the prevalent clinical disorders in the sample (n = 89)

Clinical disorder	Total TS	Total SCID	TP	FP	FN	TN	Sens	Spec	PPV	NPV
Major depression	63	60	48	15	12	14	0.80	0.48	0.76	0.54
Dysthymia	12	5	1	11	4	73	0.20	0.87	0.08	0.95
Post-traumatic stress disorder	41	50	35	6	15	33	0.70	0.85	0.85	0.69
Panic disorder	28	12	10	18	2	59	0.83	0.77	0.36	0.97
Social phobia	15	14	3	12	11	63	0.21	0.84	0.20	0.85
Obsessive–compulsive disorder	13	9	6	7	3	73	0.67	0.91	0.46	0.96
Specific phobia	10	7	3	7	4	75	0.43	0.91	0.30	0.95
Alcohol abuse/dependence	12	9	7	5	2	75	0.78	0.94	0.58	0.97
Other drug abuse/dependence	13	4	2	11	2	74	0.50	0.87	0.15	0.97
Eating disorder	13	24	7	6	17	59	0.29	0.91	0.54	0.78

Note: analyses were only performed for categories of disorders with a prevalence rate higher than 4% according to Structured Clinical Interview for *DSM-IV* Axis I disorders (SCID-I); Total TS, total number of TeleScreen classifications; Total SCID, total number of SCID classifications; TP, true positive; FP, false positive; FN, false negative; TN, true negative; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value.

DOI: 10.1002/pmh

Overall, the findings for PDs were comparable with those previously reported in primary care[22] (i.e. moderate to good validity with a high proportion of false positives), indicating that future research of web-based self-report in larger samples of patients with PDs is warranted. Our findings suggest that TeleScreen can be easily implemented in a specialized mental health setting and that the adherence to fill out the questionnaire by the respondents was high (the mean time to fill out TeleScreen was 52 min). No patients dropped out during the study, and no adverse experiences were reported by the respondents. Thus, these observations pave the way for further research on, and further development of, TeleScreen as a web-based self-report for *DSM-IV/5* clinical and PDs.

The concurrent validity of the co-occurring clinical disorders reached moderate to good values, with the exception of social phobia, dysthymia and eating disorders, which is comparable with that of previous research.[22] Most of the existing online outcome measures for clinical disorders focus on one clinical disorder or the most common clinical disorders.[30–35] The electronic Psychological Assessment and Screening System (e-PASS)[26] is an exception and has a considerably wider diagnostic reach, also taking into account less frequent clinical disorders. When published data of the e-PASS[26] are compared with the findings of TeleScreen in the current and previous study,[22] e-PASS showed better results in terms of concurrent validity in screening for social phobia, panic disorder, eating disorders and major depression, while TeleScreen showed better results for post-traumatic stress disorder, obsessive compulsive disorders and alcohol abuse/dependence. e-PASS, however, does not assess PDs and has not been examined in samples with PDs.

Our findings show that TeleScreen compared with SCIDs seems to over-diagnose patients, except for dysthymia, social phobia and eating disorder. Screening can result in over-diagnosis, leading to classifying patients as having disorders when that may not actually be the case, yet simultaneously indicating an effective screening procedure under certain circumstances. As the screening procedure in the study setting was aimed at identifying individuals who are likely to have one or more PDs and clinical disorders and might need further assessment to make a full diagnosis, over-diagnosis in these settings may be preferred to under-diagnosis. Over-diagnosis may lead, after further assessment, to an accurate (or no) diagnosis, whereas under-diagnosis reduces the ability to provide timely and accurate assessment and possible withholding optimal treatment for these patients. Further development of TeleScreen should aim at improving accuracy for disorders that were under-diagnosed in the current sample, such as social phobia and eating disorder, but also for disorders that are over-diagnosed, such as the paranoid PD, so that the follow-up assessment can be performed more efficiently.

The current study was conducted in a specialized mental health setting that contributes to the external validity, but this also ensued several limitations that should be taken into consideration. First, the number of included patients was limited to treatment-seeking patients. Different prevalence rates of other PDs such as schizoid, schizotypal, antisocial or narcissistic PD may be expected in individuals who are recruited in a general population or in forensic psychiatry. Second, the time frame between the administration of SCIDs and TeleScreen was 3 weeks. It could be that the presentation of the clinical disorders between the two assessments had changed and may have influenced the results. This time frame, however, was chosen to minimize the burden of consecutive assessments. Third, although the raters were trained in SCIDs assessment and were supervised during the study, most of them were bachelor-level or master-level psychologists with little clinical experience. The interrater reliabilities were, however, fair to excellent. This is in line with a previous study that showed that clinically inexperienced students, even those that are untrained, are able to reliably assess patients' level

DOI: 10.1002/pmh

of personality pathology from clinical interview material that converge with expert ratings.[36]

The analyses in the study were limited to categorical agreement (i.e. present/not present) between TeleScreen and the SCIDs, and this approach gave us the first preliminary results on the concurrent validity of TeleScreen for PDs in a relatively small sample. However, dimensional comparisons (i.e. severity scores) would be more informative for clinical and research purposes and is in line with the alternative dimensional approach for diagnosing personality pathology of the *DSM-5*,[37] but for the current study, where we examined different PDs, this would have required a much larger sample size. A next step would be to examine the dimensional assessment of one PD by TeleScreen compared with a well-validated dimensional measure for this PD, for example, a comparison of TeleScreen borderline PD criterion scores with the borderline PD severity index—fourth edition (BPDSI-IV).[38] Based on the data reported here on the concurrent validity of TeleScreen for the borderline PDs, such a study is underway. Also, the present study focused on classifications of TeleScreen, based on self-report; future research could examine whether the results of TeleScreen augmented with a follow-up interview by a mental health professional are comparable with results of semi-structured assessments (e.g. SCIDs).

Finally, the introduction of the *DSM-5*, when compared with the *DSM-IV*, implies several changes. The *DSM-5* criteria for PDs remained largely unchanged, and the results of the current study remain relevant for *DSM-5* PDs. However, several substantial changes have been made to the clinical disorders. To match the *DSM-5* changes for the clinical disorders, a new version of the TeleScreen (version 5.0) was developed after the data of this study were gathered. This new version of TeleScreen also re-examined several questions underlying the classifications of clinical and PDs that showed low validity in the current and previous studies. Future research should re-examine TeleScreen (version 5.0) with reliable and validated measures of *DSM-5* clinical and PDs, such as the Structured Clinical Interview for *DSM-5* (SCID-5)[39] and its Personality Disorders Version (SCID-5-PD).[40] Future research might also consider how TeleScreen could capture the alternative dimensional trait model for PDs that is included in Section III (Emerging models and measures) of the *DSM-5*.

Taken together, the current study provides preliminary evidence for the concurrent validity of TeleScreen for some PDs, and their co-occurring clinical disorders, and paves the way for larger studies to confirm these results.

Financial disclosure

Ies Dijksman has a part-time position as a Research and Development Manager of TelePsy and a part-time position as a researcher at Maastricht University.

References

1. Moran P, Jenkins R, Tylee A, Blizard R, Mann A. The prevalence of personality disorder among UK primary care attenders. *Acta Psychiatr Scand* 2000; **102**(1): 52–57.
2. Zimmerman M, Rothschild L, Chelminski I. The prevalence of DSM-IV personality disorders in psychiatric outpatients. *Am J Psychiatry* 2005; **162**(10): 1911–8.
3. Huang Y, Kotov R, de Girolamo G, Preti A, Angermeyer M, Benjet C et al. DSM-IV personality disorders in the WHO World Mental Health Surveys. *Br J Psychiatry* 2009; **195**(1): 46–53.
4. Lenzenweger MF, Lane MC, Loranger AW, Kessler RC. DSM-IV personality disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 2007; **62**(6): 553–64.
5. Fok M, Hayes R, Chang C, Stewart R, Callard F, Moran P. Life expectancy at birth and all-cause mortality among people with personality disorder. *J Psychosom Res* 2012; **73**(2): 104–7.
6. Bender DS, Dolan RT, Skodol AE, Sanislow CA, Dyck IR, McGlashan TH et al. Treatment utilization by patients with personality disorders. *Am J Psychiatry* 2001; **158**(2): 295–302.
7. Moran P, Rendu A, Jenkins R, Tylee A, Mann A. The impact of personality disorder in UK primary care: a 1-year follow-up of attenders. *Psychol Med* 2001; **31**(8): 1447–54.

8. Soeteman DI, Verheul R, Busschbach JJ. The burden of disease in personality disorders: diagnosis-specific quality of life. *J Personal Disord* 2008; **22**(3): 259–68.
9. Soeteman DI, Hakkaart-van Roijen L, Verheul R, Busschbach JJ. The economic burden of personality disorders in mental health care. *J Clin Psychiatry* 2008; **69**(2): 259–65.
10. Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence, and effect of personality disorder. *Lancet* 2015; **385**(9969): 717–26.
11. First M, Spitzer R, Gibbon M, Williams J. *Structured Clinical Interview for DSM-IV Clinical Version (SCID-I/CV)*. Washington DC: American Psychiatric Press, 1997.
12. First M, Spitzer R, Gibbon M, Williams J, Benjamin L. *Structured Clinical Interview for DSM-IV Axis II Personality Disorders*. Washington: American Psychiatric Press, 1997.
13. Lobbestael J, Leurgans M, Arntz A. Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). *Clin Psychol Psychother* 2011; **18**(1): 75–79.
14. Widiger TA, & Samuel DB. Evidence-based assessment of personality disorders. *Psychol Assess* 2005; **17**(3): 278–87.
15. Newton-Howes G, Tyrer P, Anagnostakis K, Cooper S, Bowden-Jones O, Weaver T et al. The prevalence of personality disorder, its comorbidity with mental state disorders, and its clinical significance in community mental health teams. *Soc Psychiatry Psychiatr Epidemiol* 2010; **45**(4): 453–60.
16. Beckwith H, Moran PF, Reilly J. Personality disorder prevalence in psychiatric outpatients: a systematic literature review. *Personal Ment Health* 2014; **8**(2): 91–101.
17. Gunderson JG. Clinical practice. Borderline personality disorder. *N Engl J Med* 2011; **364**(21): 2037–42.
18. Westen D. Divergences between clinical and research methods for assessing personality disorders: implications for research and the evolution of Axis II. *Am J Psychiatry* 1997; **154**(7): 895–903.
19. Zimmerman M, & Mattia JI. Differences between clinical and research practices in diagnosing borderline personality disorder. *Am J Psychiatry* 1999; **156**(10): 1570–4.
20. Brodey BB, First M, Linthicum J, Haman K, Sasiela JW, Ayer D. Validation of the NetSCID: an automated web-based adaptive version of the SCID. *Compr Psychiatry* 2016; **66**: 67–70.
21. Dijksman I, Dinant GJ, Spigt MG. eDiagnostics: a promising step towards primary mental health care. *Fam Pract* 2013; **30**(6): 695–704.
22. Dijksman I, Dinant GJ, Spigt M. The concurrent validity of a new eDiagnostic system for mental disorders in primary care. *Fam Pract* 2016; **33**(6): 607–16.
23. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition. Washington: American Psychiatric Association, 2000.
24. Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Sheehan KH et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur Psychiatry* 1997; **12**(5): 224–31.
25. Fleiss J. *Statistical Methods for Rates and Proportions*, 2nd Edition. New York: Wiley, 1981.
26. Nguyen DP, Klein B, Meyer D, Austin DW, Abbott JA. The diagnostic validity and reliability of an internet-based clinical assessment program for mental disorders. *J Med Internet Res* 2015; **17**(9): e218.
27. Zanarini MC, Weingeroff JL, Frankenburg FR, Fitzmaurice GM. Development of the self-report version of the Zanarini Rating Scale for Borderline Personality Disorder. *Personal Ment Health* 2015; **9**(4): 243–9.
28. Bloo J, Artz A, Schouten E. The Borderline Personality Disorder Checklist: psychometric evaluation and factorial structure in clinical and nonclinical samples. *Roczniki Psychologiczne/Ann Psychol* 2018; **20**(2): 311–36.
29. Calvo N, Valero S, Arntz A, Andión Ó, Matalí J, Navascues V et al. Validation of the Spanish version of the Borderline Personality Disorder Checklist (BPD Checklist) in a sample of BPD patients: study of psychometric properties. *Eur J Psychiatry* 2018; **32**(1): 26–35.
30. Donker T, van Straten A, Marks I, Cuijpers P. A brief Web-based screening questionnaire for common mental disorders: development and validation. *J Med Internet Res* 2009; **11**(3): e19.
31. Donker T, Comijs H, Cuijpers P, Terluin B, Nolen W, Zitman F et al. The validity of the Dutch K10 and extended K10 screening scales for depressive and anxiety disorders. *Psychiatry Res* 2010; **176**(1): 45–50.
32. Lin CC, Bai YM, Liu CY, Hsiao MC, Chen JY, Tsai SJ et al. Web-based tools can be used reliably to detect patients with major depressive disorder and subsyndromal depressive symptoms. *BMC Psychiatry* 2007; **7**: 12.
33. Farvolden P, McBride C, Bagby RM, Ravitz P. A Web-based screening instrument for depression and anxiety disorders in primary care. *J Med Internet Res* 2003; **5**(3).
34. Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV, 3rd, Hahn SR et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA* 1994; **272**(22): 1749–56.
35. Spitzer RL, Kroenke K, Williams JW. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. *JAMA* 1999; **282**(18): 1737–44.
36. Zimmermann J, Benecke C, Bender DS, Skodol AE, Schauenburg H, Cierpka M et al. Assessing DSM-5

level of personality functioning from videotaped clinical interviews: a pilot study with untrained and clinically inexperienced students. *J Pers Assess* 2014; **96**(4): 397–409.
37. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edition. Arlington, VA: American Psychiatric Association, 2013.
38. Arntz A, van den Hoorn M, Cornelis J, Verheul R, van den Bosch WM, de Bie AJ. Reliability and validity of the borderline personality disorder severity index. *J Personal Disord* 2003; **17**(1): 45–59.
39. First M, Williams J, Karg R, Spitzer R. *Structured Clinical Interview for DSM-5 Research Version (SCID-5-RV)*. Arlington, VA: American Psychiatric Association, 2015.
40. First M, Williams J, Benjamin L, Spitzer R. *Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD)*. Arlington, VA: American Psychiatric Association, 2016.

Address correspondence to: Farid Chakhssi, Scelta, GGNet, PO Box 928, 7301 BD, Apeldoorn, The Netherlands. Email: f.chakhssi@ggnet.nl