Psychiatry Research 279 (2019) 71–76

Contents lists available at ScienceDirect

Psychiatry Research

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

Identifying non-affective psychosis in first admission patients: MMPI-2, structured diagnostic interview, and consensus lifetime best estimate

John W. Carter[a,b], Julie Nordgaard[c,*], Josef Parnas[d,e]

[a] Department of Counseling and Human Services, School of Education, University of Redlands, 1200 E. Colton Ave., Redlands, CA 92373, USA
[b] Carter Strategic Research, Granada Hills, California, USA
[c] Mental Health Center Amager, University Hospital Copenhagen, Gammel Kongevej 33, Copenhagen 1610, Denmark
[d] Center for Subjectivity Research, University of Copenhagen, Karen Blixens Plads 8, Copenhagen 2300, Denmark
[e] Mental Health Center Glostrup, University of Copenhagen, Brøndbyøstervej 160, Brøndby 2605, Denmark

ARTICLE INFO

Keywords:
Schizophrenia
SCID
MMPI-2
Goldberg Index
Incremental validity

ABSTRACT

Purpose: We aimed to evaluate the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) as a potential means of improving on the SCID's diagnostic efficacy.
Methods: 76 first-admission patients were assigned DSM-IV consensus diagnoses by two experienced psychiatrists using all available information, then dichotomized into non-affective psychosis and other mental illness groups. The patients were also given the SCID and the MMPI-2. The diagnostic performance of the MMPI-2 was compared to that of the SCID to assess both diagnostic accuracy and incremental validity.
Results: MMPI-2 scales 8 (Schizophrenia) and BIZ (Bizarre Mentations) correctly identified 58% and 56% respectively of non-affective psychotic patients. The Goldberg Index had an overall correct classification rate of 70%, but only identified 49% of the psychosis group. The SCID had a correct classification rate of 66% but correctly identified only 25% of the non-affective psychosis patients. Three MMPI-2 scales combined with the SCID resulted in an overall correct classification rate of 73%, and identification of 66% of the non-affective psychosis patients.
Conclusion: The results suggest that the MMPI-2 may identify early psychosis at least as well as the SCID. Furthermore, using a combination of the MMPI-2 and the SCID shows incremental validity over using the SCID alone.

1. Introduction

The last decades have seen worldwide efforts and research targeting early detection of schizophrenia-spectrum psychosis. This effort was initiated in Australia (Yung et al., 2003) and was motivated by the recognition of the fact that many help-seeking individuals with subtle psychotic features failed to meet the full criteria for the schizophrenia diagnosis. We have elsewhere discussed in detail the high diagnostic threshold for schizophrenia in DSM-IV, DSM-5 and ICD-10 (Parnas, 2002, 2012).

In the studies on early detection a variety of diagnostic approaches have been used, ranging from self-report measures to structured interviews. Kendler et al. (1996) observed in the US National Comorbidity Survey that the structured interview has limited utility in detecting psychosis. In our own study of 100 consecutive first admission patients, we found that a fully structured psychiatric interview, the Structured Clinical Interview for DSM-IV-TR or SCID, displayed high specificity but low sensitivity for schizophrenia and other non-affective psychotic disorders (Nordgaard et al., 2012). Other evidence has shown an overall low level of agreement between the SCID and clinicians' diagnoses (Steiner et al., 1995).

We have consistently advocated for thorough psychiatric assessment by experienced clinicians as the "gold standard" for all schizophrenia-related diagnostic efforts (Leckman et al., 1982; Nordgaard et al., 2012) including early detection (Nordgaard et al., 2013; Parnas, 2015). Core criteria include the use of multiple sources of information such as historical records, psychological tests, informants' observations, and following the patient over time (Meehl, 1959). We recognize the practical reality that in some clinic settings this gold standard may not be feasible, either due to a lack of economic resources or staff expertise. In this paper, we explore the possibility that using the MMPI-2 (Minnesota Multiphasic Personality Inventory-2) as an adjunct to the SCID can improve its diagnostic accuracy for schizophrenia-related psychosis at first admission. Since the MMPI-2 is a self-complete instrument that can

* Corresponding author.
E-mail address: Julie_nordgaard@dadlnet.dk (J. Nordgaard).

https://doi.org/10.1016/j.psychres.2019.07.010
Received 26 April 2019; Received in revised form 5 July 2019; Accepted 7 July 2019
Available online 08 July 2019
0165-1781/

be administered with limited staff time and training, its addition would not pose an undue burden on clinic resources.

One impediment to its use in early detection paradigms is that the MMPI-2 cannot be used "right off the shelf" to assign psychotic diagnoses like the SCID can. Although there is a traditional clinical practice of profile interpretation, where scale elevations above a standard cut point are sorted to identify characteristic 2- or 3-point code types, this practice has not yet produced a set of validated differential diagnosis decision rules using current diagnostic criteria.

Instead, the research literature has advanced several alternative approaches for identifying patients with schizophrenia[1] among mixed psychiatric populations. One method is to use the Clinical scales, which reflect diagnostic categories in use at the time the MMPI was developed in the 1940s; the most consistent results for schizophrenia patients are a pattern of higher scores on Scale 8 (Schizophrenia) and lower scores on Scales 2 (Depression) and 7 (Psychasthenia) (Bagby et al., 2005; Ben-Porath et al., 1991; Graham, 2006; Greenblatt and Davis, 1999; Wetzler et al., 1998). Alternatively, the more recently constructed Content Scales (Butcher et al., 1990) may be used; here, the most consistent results for schizophrenia patients are a pattern of higher scores on the Bizarre Mentation (BIZ) scale and lower scores on the Depression (DEP) and Anxiety (ANX) scales (Bagby et al., 2005; Ben-Porath et al., 1991; Butcher et al., 1990; Graham, 2006; Munley et al., 1997; Wetzler et al., 1998). Another approach has been to construct an index using an arithmetic combination of several different scales. Goldberg developed an index using the original MMPI that outperformed expert clinicians in predicting psychotic vs. non-psychotic discharge diagnoses (Goldberg, 1965, 1968); an MMPI-2 update of the Goldberg Index was successful in discriminating schizophrenia from major depression (Egger et al., 2003). It is worthwhile to note that although Goldberg's Index made more efficient use of MMPI scale data than clinicians' judgments, the gold-standard criterion that both attempted to predict was the discharge diagnosis assigned by the treating clinician.

Our overall aim in the current study is to test the ability of the above MMPI-2 measures to identify non-affective psychosis in a sample of first-admission psychiatric patients, and then to determine if the best-performing MMPI-2 measure can add to the diagnostic accuracy of the SCID alone.

2. Methods

2.1. Sample

The sample comprised consecutive first-admission patients at the Psychiatric Centre Hvidovre, a department of the University Hospital of Copenhagen. The department provides psychiatric service to a population of 150,000 in a catchment area of the City of Copenhagen as there are no private psychiatric in-patient facilities in Denmark. All consecutive first admissions from June 2009 to November 2010 were screened for eligibility. Due to ethical concerns, involuntarily admitted and legal patients were excluded. Participants had to be considered capable of tolerating lengthy interviews, which naturally excluded severely aggressive, agitated or severely psychotic patients; substance abuse, mental retardation, and organic patients were also excluded, as were those over 65 (for details see Nordgaard et al. (2012)). Out of a total of 231 first-admission patients, 122 were invited to enroll in the study. Of these, 16 patients declined to participate (clinical diagnoses: 4 with schizophrenia, 1 with schizotypal disorder, 9 with major depression, 1 with anxiety and 1 with deferred diagnosis). Six patients had to be excluded after enrollment because they were found to not meet the inclusion criteria ($n = 3$), did not show up for the interview appointments ($n = 2$), or withdrew consent after completed interviews ($n = 1$). The final sample consisted of 34 men and 66 women with a mean age of 27.7 years (range 18–60 years), representing 82% of the patients initially invited to participate. All patients gave informed consent to participate; the study was approved by the Medical Ethical Committee for the University of Copenhagen and conforms to the ethical standards set forth in the Declaration of Helsinki.

2.2. Assessments

All patients were interviewed twice within the same week: first with the SCID interview by a non-psychiatrist, and second with a semi-structured conversational interview (SSCI) by an experienced psychiatrist. All interviews were videotaped. The mean time from admission to the first interview was 13 days (range 2–71).

2.2.1. The SCID

The SCID interview comprised the Structured Clinical Interview for the DSM-IV-TR axis I disorders (SCID-I) and the Schizotypal Personality Disorder module from the SCID-II (First et al., 2002). The interview was conducted in a fully structured way by a master's level clinician trained and certified as a SCID interviewer by the University of California Los Angeles Center for Neurocognition and Emotion in Schizophrenia (Ventura et al., 1998). The interviewer assigned a DSM-IV diagnosis based on the SCID's specific scoring rules. An experienced research psychiatrist supervised the performance of the interviews and the allocations of diagnoses in order to prevent errors.

2.2.2. The semi-structured conversational interview (SSCI)

This interview comprised a thorough psychosocial history and multiple item sets relevant to the differential diagnosis of psychosis and detection of self-disturbance[2] (for a complete description of items covered, see Nordgaard et al. (2012). The interviewer explored these items in a sequence that seemed appropriate and adequate to the participant's own concerns and responses, according to the phenomenological principles proposed by Jaspers and others (Jasper, 1963; Parnas et al., 2002).

The SSCI interview was conducted by one of us (JN), an experienced clinical psychiatrist with psychometric research experience. On the basis of the interview, JN allocated a DSM-IV diagnosis.

The project director (JP), a senior research clinician, independently reviewed the diagnostic material elicited in the SSCI and allocated his own DSM-IV diagnoses. The diagnostic agreement between JN and JP was 93%.

Finally, the consensus lifetime best estimate (CLBE) diagnosis was allocated to each patient at a consensus meeting of JP and JN, using all clinical information available on each patient (e.g. the SSCI diagnoses, videos of these interviews, notes, hospital charts, informant descriptions). This diagnosis was the study's gold standard.

2.2.3. The MMPI-2

Within the same week as the interviews were conducted, all

[1] We recognize that the terms "schizophrenia," "schizophrenia-related psychosis," and "non-affective psychosis" are used somewhat interchangeably throughout this paper. While we are interested in detecting the full class of non-affective psychotic disorders, the research evidence principally concerns schizophrenia (the most prevalent and perhaps most prototypical disorder in the class). Based on our experience, it is acceptable to extend psychometric inferences based on schizophrenia research to non-affective psychotic populations more generally, whereas it is important to maintain the distinction between non-affective psychosis and bipolar or depressive psychosis. The scattered research that is available suggests that those with schizophreniform and brief psychotic disorders show similar but less extreme MMPI profiles than patients with schizophrenia (Rusaka and Rancans, 2014; Walters, 1984).

[2] Notably, self-disturbances are not diagnostic of any specific psychiatric disorders in the DSM-IV nor in the ICD-10, and self-disturbances are present to the same degree in non-psychotic schizotypal personality disorder and schizophrenia (Nordgaard and Parnas, 2014).

participants were asked to complete the MMPI-2 via computer administration (the Hogrefe Test System (HTS) program). 93 of the 100 first-admission patients in our sample completed the MMPI-2. Of the 7 who refused the test, 3 were diagnosed with schizophrenia, 2 with major depression, 1 with schizotypal personality disorder, and 1 with an unspecified personality disorder. These 93 MMPI-2 profiles were reviewed for validity, with reference to the recommended criteria of Graham (2006). Profiles were considered invalid if any of the following were true:

(a) VRIN (Variable Response Inconsistency) raw score greater than or equal to 13
(b) TRIN (True Response Inconsistency) raw score either less than or equal to 5 or greater than or equal to 13
(c) T-score on the L (Lie) scale of at least 80
(d) T-score on the K (Defensiveness) scale of at least 65
(e) T-score of 80 or more on either the F (Infrequency) or the Fb (Back F) scale, but ONLY if the Fp (F Psychopathology) T-score was at least 100 as well. This caveat prevents otherwise valid psychotic profiles from being excluded due to infrequent responses related to psychotic symptoms—see Greenblatt and Davis (1999)
(f) missing 25% or more of the test items (i.e. ? (Cannot Say) score greater than or equal to 142)

A total of 17 profiles were discarded due to invalidity, leaving 76 participants with valid MMPI-2 profiles. Consensus diagnoses of the 17 with invalid MMPI-2 profiles are as follows: 3 organic disorders, 8 schizophrenia, 3 major depression, 1 panic disorder, 1 schizotypal personality disorder, and 1 personality disorder unspecified.

It will be noted that our criterion (f) allows a much higher number of Cannot Say responses than is standard (i.e. 30). However, the standard recommendation is based on the assumption that the MMPI/MMPI-2 is administered in a standard fashion—namely, that the response sheet includes only two possible responses for each item (True or False) and that the test administrator gives specific instructions discouraging subjects from leaving items blank (i.e. Cannot Say)—and thus that profiles with more than 30 Cannot Say responses are relatively rare. In our sample, 13% of patients taking the MMPI-2 had more than 30 Cannot Say responses. Upon review of the HTS format for administering the test, we identified that the computer program deviated from standard administration practices by presenting the user with a large 'No Answer' button for each item in addition to 'True' and 'False' buttons. We judged that this format would logically affect the relative likelihood of 'No Answer' (Cannot Say) responses across the board, and thus simply invalidating a large number of otherwise valid profiles due to Cannot Say responses would not be appropriate. Although we did not find any literature specifically bearing on the Hogrefe program, differences in Cannot Say scores between computerized and pencil-and-paper administrations have been reported (Lambert et al., 1987).

To compensate for the relatively high frequency of missing items (i.e. Cannot Say responses), we prorated scale scores based on the proportion of missing items, with scales scored as missing if more than 25% of the scale's items were missing. This is a more conservative version of the scoring procedure we employed successfully in other MMPI studies (Carter et al., 1999, 2002) and follows the general approach of Moldin et al. (1987) and Greene (1980).

2.3. Analysis

Diagnoses were categorized into 2 groups: (1) non-affective psychosis, including schizophrenia and other Axis I psychotic disorders outside of the affective disorders category, and (2) other mental illness, including all disorders in the affective disorders category (even affective psychoses), anxiety disorders, and all other Axis I and Axis II disorders. The broader diagnostic category of non-affective psychosis was selected instead of schizophrenia to provide increased statistical power and a cleaner dichotomous criterion (as recommended by Ganellen (1996)). Among the sample of 76 participants with valid MMPI-2 profiles, there were 33 in the CLBE psychosis (PSY) group[3] and 43 in the CLBE other mental illness (OMI) group.[4] Presence or absence of a valid MMPI-2 profile was not significantly related to diagnosis (Chi-square = 0.85, df = 1, NS).

MMPI-2 T-scores were calculated from raw scores using the norms in Graham (2006). The Goldberg Index was calculated by summing T-scores for Scale L (Lie), Scale 6 (Paranoia), and K-corrected Scale 8 and subtracting scores for Scale 3 (Hysteria) and K-corrected Scale 7. K-correction was only used here to preserve continuity with the original index; uncorrected scale scores were used in the remainder of the analyses.

Logistic regression (LR) was used for all hypothesis testing. We first tested the prediction of diagnostic group (PSY vs. OMI) for each of the following MMPI-2 scales: 2, 7, 8, DEP, ANX, and BIZ. Bonferroni correction was used to control for Type-I error associated with multiple statistical tests. We next included all six scales in another LR analysis and used backwards selection to identify the set of MMPI-2 scales most predictive of PSY. We similarly tested the Goldberg Index's prediction of PSY diagnosis. Finally, the best-performing MMPI-2 index (either the LR-derived set of scales or the Goldberg Index) was added to SCID diagnosis in a hierarchical LR to test its incremental validity (Haynes and Lench, 2003). Classification results for each LR analysis were based on a 0.5 probability cutoff (i.e. all cases predicted to have more than a 50% chance of being diagnosed PSY were classified as PSY; all others were classified as OMI).

3. Results

Table 1 presents demographic and hospitalization information about the sample as a whole ($N = 76$), as well as by CLBE diagnostic group. The PSY ($N = 33$) and OMI ($N = 43$) groups did not differ from each other in terms of gender, age, marital status, education level, employment status prior to hospitalization, nor duration of hospitalization prior to the interview.

Among the LR tests of individual scales (presented in Table 2), only 3 were significant: Scales 7, 8, and BIZ. Examination of the group means shows that the PSY group scored higher than the OMI group on each of these scales. Of the three scales, both Scale 8 and the BIZ scale correctly identified a slight majority of PSY patients (58% and 56% respectively), with the BIZ scale risking fewer false positives to reach a 75% correct classification rate (vs. 66% correct classification for Scale 8).

Backwards elimination of the 3 Clinical and 3 Content scales predicting CLBE diagnosis resulted in a LR equation consisting of 3 predictors: Scales 8, ANX, and BIZ (Table 3). The overall model was highly significant. The classification results showed an overall correct classification rate of 72% while correctly identifying 66% of PSY patients. Higher scores on both Scale 8 and BIZ increased the risk for non-affective psychosis, while higher ANX scores decreased the risk for psychosis.

The LR results for the Goldberg Index are reported in Table 2. The Goldberg Index (GI) was significantly associated with CLBE diagnosis. Examination of the groups means shows that PSY scored higher on the Index (mean GI = 63.0) than OMI (mean GI = 46.7). The GI had an overall correct classification rate of 70%, but could only correctly identify 49% of the psychosis group.

The final analysis was a hierarchical LR to test the incremental effect of adding MMPI-2 score information to the SCID. On its own, the

[3] 29 schizophrenia, 2 brief psychotic disorder, 1 schizophreniform disorder, 1 delusional disorder.

[4] 20 schizotypal personality disorder, 9 major depression (2 with psychotic features), 6 personality disorder NOS, 5 adjustment disorder, 1 bipolar disorder with psychotic features, 1 generalized anxiety disorder, 1 no diagnosis.

Table 1
Sample demographic and hospitalization characteristics.

	Total (n = 76)	PSY (n = 33)	OMI (n = 43)	Significance Test
Gender -% female	66%	61%	70%	$\chi^2(1) = 0.70, p = .40$
Age in years - Mean (SD)	27.1 (8.7)	26.3 (6.2)	27.8 (10.2)	$t(74) = 0.73, p = .47$
Marital status -% unmarried	54%	55%	54%	$\chi^2(1) = 0.01, p = .93$
Education level -% high school degree or less	78%	73%	81%	$\chi^2(1) = 0.81, p = .37$
Employment -% employed	86%	85%	86%	$\chi^2(1) = 0.02, p = .88$
Days of hospitalization - Mean (SD)	15.5 (14.6)	16.7 (19.6)	14.6 (9.2)	$t(74) = -0.62, p = .54$

PSY, Non-affective psychosis; OMI, Other mental illness.

Table 2
Univariate logistic regression of individual MMPI-2 scales and Goldberg Index on CLBE diagnosis.

	PSY (n = 33) Mean (SD)	OMI (n = 43) Mean (SD)	Significance Test	Goodness of Fit	Classification Results
Clinical scale 2	80.5 (12.3)	79.5 (12.2)	$\chi^2(1) = 0.12, p = .73$	N. $R^2 = 0.00$	–
Clinical scale 7	79.3 (10.7)	70.2 (13.8)	$\chi^2(1) = 9.49, p = .002$	N. $R^2 = 0.16$	TPR = 52%, TNR = 72%, PPV = 59%, NPV = 66%, CCR = 63%
Clinical scale 8	82.4 (12.3)	68.8 (13.9)	$\chi^2(1) = 17.51, p < .001$	N. $R^2 = 0.28$	TPR = 58%, TNR = 72%, PPV = 61%, NPV = 69%, CCR = 66%
DEP scale	79.2 (10.4)	75.0 (10.6)	$\chi^2(1) = 3.04, p = .08$	N. $R^2 = 0.05$	–
ANX scale	72.9 (10.9)	69.4 (10.7)	$\chi^2(1) = 1.99, p = .16$	N. $R^2 = 0.04$	–
BIZ scale	68.7 (15.2)	55.0 (9.4)	$\chi^2(1) = 20.03, p < .001$	N. $R^2 = 0.32$	TPR = 56%, TNR = 88%, PPV = 78%, NPV = 73%, CCR = 75%
Goldberg Index[a]	63.0 (21.0)	46.7 (17.7)	$\chi^2(1) = 12.42, p < .001$	N. $R^2 = 0.20$	TPR = 49%, TNR = 86%, PPV = 73%, NPV = 69%, CCR = 70%

MMPI-2, Minnesota Multiphasic Personality Inventory-2; CLBE, Consensus lifetime best estimate; PSY, Non-affective psychosis; OMI, Other mental illness; DEP, Depression; ANX, Anxiety; BIZ, Bizarre mentation; N. R^2, Nagelkerke R^2; TPR, True positive rate (Sensitivity) for psychosis; TNR, True negative rate (Specificity) for psychosis; PPV, Positive predictive value; NPV, Negative predictive value; CCR, Correct classification rate (Accuracy).

[a] Goldberg Index modified for use with MMPI-2.

Table 3
Multivariate logistic regression of MMPI-2 Scales for CLBE non-affective psychosis diagnosis.

	Significance Test	Goodness of Fit	Classification Results	Odds Ratio (95% CI)
Overall LR model	$\chi^2(3) = 27.30, p < .001$	N. $R^2 = 0.41$	TPR = 66%, TNR = 76%, PPV = 68%, NPV = 74%, CCR = 72%	
Intercept (Constant)	W. $\chi^2(1) = 4.40, p = .04$			0.02
Clinical scale 8	W. $\chi^2(1) = 4.80, p = .03$			1.10 (1.01–1.20)
ANX scale	W. $\chi^2(1) = 6.15, p = .01$			0.90 (0.82–0.98)
BIZ scale	W. $\chi^2(1) = 3.78, p = .05$			1.08 (1.00–1.16)

MMPI-2, Minnesota Multiphasic Personality Inventory-2; CLBE, Consensus lifetime best estimate; LR, Logistic regression; ANX, Anxiety; BIZ, Bizarre mentation; C.I., Confidence interval; N. R^2, Nagelkerke R^2; TPR, True positive rate (Sensitivity) for psychosis; TNR, True negative rate (Specificity) for psychosis; PPV, Positive predictive value; NPV, Negative predictive value; CCR, Correct classification rate (Accuracy); W. χ^2, Wald χ^2.

Table 4
Hierarchical logistic regression adding MMPI-2 scales to SCID for CLBE non-affective psychosis diagnosis.

	Significance Test	Goodness of Fit	Classification Results	Odds Ratio (95% CI)
Step 1: SCID diagnosis[a]	$\chi^2(1) = 9.34$, $p = .002$	N. $R^2 = 0.16$	TPR = 25%, TNR = 98%, PPV = 89%, NPV = 63%, CCR = 66%	
Step 2: Clinical scale 8, ANX scale, BIZ scale	$\chi^2(3) = 21.92$, $p < .001$	ΔN. $R^2 = 0.30$		
Final LR model	$\chi^2(4) = 31.26$, $p < .001$	N. $R^2 = 0.46$	TPR = 66%, TNR = 79%, PPV = 70%, NPV = 75%, CCR = 73%	
Intercept (Constant)	W. $\chi^2(1) = 4.63$, $p = .03$			0.02
SCID diagnosis[a]	W. $\chi^2(1) = 3.06$, $p = .08$			7.98 (0.78–81.90)
Clinical scale 8	W. $\chi^2(1) = 3.66$, $p = .06$			1.09 (1.00–1.19)
ANX scale	W. $\chi^2(1) = 5.35$, $p = .02$			0.90 (0.82–0.98)
BIZ scale	W. $\chi^2(1) = 3.71$, $p = .05$			1.09 (1.00–1.18)

MMPI-2, Minnesota Multiphasic Personality Inventory-2; SCID, Structured Clinical Interview for DSM-IV-TR; CLBE, Consensus lifetime best estimate; ANX, Anxiety; BIZ, Bizarre mentation; LR, Logistic regression; C.I., Confidence interval; N. R^2, Nagelkerke R^2; ΔN. R^2, Change in Nagelkerke R^2; TPR, True positive rate (Sensitivity) for psychosis; TNR, True negative rate (Specificity) for psychosis; PPV, Positive predictive value; NPV, Negative predictive value; CCR, Correct classification rate (Accuracy); W. χ^2, Wald χ^2.

[a] SCID diagnosis dichotomized into non-affective psychosis (1) and other mental illness (0).

SCID diagnosis was significantly associated with CLBE diagnosis ($P = .002$). The SCID LR had an overall correct classification rate of 66% but correctly identified only 25% of the PSY patients (Table 4). The addition of the 3 MMPI-2 scales from Table 3 resulted in substantial incremental validity above and beyond the SCID diagnosis. The overall model showed an increased significance ($P < .001$) and a higher correct classification rate of 73%—furthermore, it correctly identified 66% of the PSY group.

4. Discussion

The level of criterion validity demonstrated by the MMPI-2 in our study is roughly in line with the results from two important meta-analyses in the literature. One report looked at 31 studies that investigated any sort of relationship between MMPI scores and theoretically-related variables (such as between the MMPI Scale 2 (Depression) and the Beck Depression Inventory or between MMPI Scale 8 and days of psychiatric hospitalization). The overall average correspondence was $r = 0.30$, which the authors translated to an estimated 65% correct classification rate (Hiller et al., 1999). The second report examined a smaller number of studies that specifically tested the MMPI's accuracy in identifying schizophrenia. The correct classification rate for a single scale was 44%, but this figure was based on only a single study; this improved to an average of 72% when multiple scales were used (Ganellen, 1996). Using equivalent metrics, we saw an average correct classification rate of 68% for the three single scales (7, 8, and BIZ) and an average of 71% for the two indices (Goldberg Index, LR equation of 8, BIZ, and ANX).

One of the challenges in using the MMPI-2 as a diagnostic aid is the sheer number of scales that have been developed and, at various times, held to discriminate schizophrenia (and other non-affective psychoses) from non-psychotic diagnoses. What our findings can recommend in this context is that we focus on a much smaller set of scales that show broad support in the literature. Elevations on the 8 and BIZ scales are indicative of non-affective psychosis both in the present study as well as in virtually every other report we reviewed. Discerning a consistent role for the ANX scale is more problematic. In other reports (Bagby et al., 2005; Ben-Porath et al., 1991; Graham, 2006; Greenblatt and Davis, 1999), ANX scores are generally lower for the schizophrenia group than they are for the comparison group, usually major depression. In our results, the schizophrenia/non-affective psychosis group actually scored higher on ANX than those with other psychiatric diagnoses; however, the logistic regression results revealed that it was the relative differences in elevation between the psychotic scales (8 and BIZ) and the anxiety scale (ANX) that was most diagnostically important. This relative result (psychotic scores higher than anxiety scores) would also hold true in many of the above reports, and is contained in the algorithm for the Goldberg Index—which is successful in discriminating psychosis both here and in other samples (Egger et al., 2003).

Our study focused on only the most commonly-researched Clinical and Content scales. However, there is a sizable body of work in the MMPI-2 literature that advocates use of the Restructured Clinical (RC) scales (Tellegen et al., 2006) which were constructed to reduce shared variance between scales in order to improve their discriminating power (Tellegen et al., 2009). Therefore, we ran an alternate analysis with a set of 5 scales—RCd (Restructured Demoralization), RC2 (Restructured Depression), RC6 (Restructured Paranoia), RC7 (Restructured Psychasthenia), and RC8 (Restructured Schizophrenia)—associated with discriminating schizophrenia and/or non-affective psychosis in other studies (Sellbom et al., 2012; Tellegen et al., 2009). In a logistic regression analysis similar to that used to test the combination of Clinical and Content scales, backwards elimination resulted in a single variable, RC8, predicting non-affective psychosis diagnosis ($P < .001$; Nagelkerke $R^2 = 0.28$). The classification results showed about the same correct classification rate (73%) as the main analyses, but only correctly identified 56% of the non-affective psychosis patients (i.e. sensitivity; also, specificity = 86%, positive predictive value = 75%, negative predictive value = 72% for this analysis). In short, the result using the RC scales is similar but not superior to the result we obtained using the Clinical and Content scales.

The performance of the SCID in this study (and in previous reports of the same sample; e.g. Nordgaard et al. (2012)) is notable for its high

specificity but low sensitivity: only one case without psychosis was misclassified, but the SCID criteria were not sensitive enough to detect the majority of psychotic individuals. This finding is in line with the results from other studies, e.g. Benazzi (2001) observed that the sensitivity of the SCID for bipolar II was 29.4% and the specificity was 90.7%. Kendler and colleagues concluded that standard structured psychiatric interviews are a questionable method for detecting psychosis (Kendler et al., 1996). It is likely that the SCID's threshold scoring system is particularly ill-suited for first episodes of psychosis—where patients' self-reports of subtle self-disturbances and perceptual abnormalities may not satisfy the frequency, duration, and severity criteria necessary to be scored as "delusions" and "hallucinations", but may nevertheless indicate a psychotic Gestalt to the experienced clinician. This raises the interesting possibility that the SCID response data could be scored in a continuous (to what extent) fashion instead of a categorical (all or nothing) manner, with the resulting dimensional scores used to develop more sensitive thresholds for early identification applications.

We considered several potential objections to our interpretation of the study findings. It could be argued that the inclusion criterion of being able to tolerate lengthy interviews may have artificially restricted the study sample. This may have had the effect of excluding some of the most 'easy-to-diagnose' patients (as some florid psychotic conditions are so clear-cut that specialized interviews are not needed to classify them). However, even if this requirement resulted in a more difficult-to-diagnose sample, then this would be the case for both kinds of interviews and the relationship between the SCID and the gold-standard CLBE diagnoses would remain unchanged. Another objection could be that the SCID was performed by a certified junior clinician and that the SCID would have performed better if conducted by a more experienced clinician. However, a study testing SCID interrater reliability and diagnostic accuracy between trained novice raters and clinically experienced raters found no significant differences (Ventura et al., 1998).

In conclusion, our results suggest that the use of the SCID to diagnose psychosis in early identification samples is not sufficient on its own. In the absence of other sources of clinical information (such as those utilized in a comprehensive psychiatric assessment), we showed that the MMPI-2 can add incremental validity and increase diagnostic accuracy for psychosis over the SCID alone. Although promising, these results are only a first step; multivariate equations from a single sample are often unstable and sample-specific error can distort the true relationships between predictors and criterion. Additional research on independent samples is necessary to cross-validate these findings; furthermore, any prediction algorithm that comes out of such cross-validation should be experimentally tested vs. the SCID alone before it be recommended for diagnostic use in early detection and/or early intervention clinical settings.

Declarations of interest

None

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2019.07.010.

References

Bagby, R.M., Marshall, M.B., Basso, M.R., Nicholson, R.A., Bacchiochi, J., Miller, L.S., 2005. Distinguishing bipolar depression, major depression, and schizophrenia with the MMPI-2 clinical and content scales. J. Pers. Assess. 84 (1), 89–95.

Ben-Porath, Y.S., Butcher, J.N., Graham, J.R., 1991. Contributions of the MMPI-2 content scales to the differential diagnosis of schizophrenia and major depression. Psychol. Assess 3, 634–640.

Benazzi, F., 2001. Sensitivity and specificity of clinical markers for the diagnosis of bipolar II disorder. Compr. Psychiatry 42 (6), 461–465.

Butcher, J.N., Graham, J.R., Williams, C.L., Ben-Porath, Y.S., 1990. Development and Use of the MMPI-2 Content Scales. University of Minnesota Press, Minneapolis.

Carter, J.W., Parnas, J., Cannon, T.D., Schulsinger, F., Mednick, S.A., 1999. MMPI variables predictive of schizophrenia in the Copenhagen High-Risk Project: a 25-year follow-up. Acta Psychiatr. Scand. 99 (6), 432–440.

Carter, J.W., Schulsinger, F., Parnas, J., Cannon, T., Mednick, S.A., 2002. A multivariate prediction model of schizophrenia. Schizophr. Bull. 28 (4), 649–682.

Egger, J.I., Delsing, P.A., De Mey, H.R., 2003. Differential diagnosis using the MMPI-2: Goldberg's index revisited. Eur. Psychiatry 18 (8), 409–411.

First, M., Gibbon, M., Spitzer, R., Williams, J., 2002. User's Guide for SCID-I Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research version. Biometrics Research Department, New York State Psychiatric Institute, New York.

Ganellen, R.J., 1996. Comparing the diagnostic efficiency of the MMPI, MCMI-II, and Rorschach: a review. J. Pers. Assess. 67 (2), 219–243.

Goldberg, L.R., 1965. Diagnosticians vs. diagnostic signs; the diagnosis of psychosis and neurosis from the MMPI. Psychol. Monogr. 1–28.

Goldberg, L.R., 1968. Simple models or simple processes? Some research on clinical judgments. Am. Psychol. 23 (7), 483–496.

Graham, J.R., 2006. MMPI-2: Assessing Personality and Psychopathology. Oxford University Press, Oxford.

Greenblatt, R.L., Davis, W.E., 1999. Differential diagnosis of PTSD, schizophrenia, and depression with the MMPI-2. J. Clin. Psychol. 55 (2), 217–223.

Greene, R.L., 1980. The MMPI: An Interpretive Manual. Grune & Stratton, New York.

Haynes, S.N., Lench, H.C., 2003. Incremental validity of new clinical assessment measures. Psychol. Assess. 15 (4), 456–466.

Hiller, J.B., Rosenthal, R., Berry, D.T.R., Brunell-Neuleib, S., 1999. A comparative meta-analysis of Rorschach and MMPI validity. Psychol. Assess. 11 (3), 278–296.

Jasper, K., 1963. General Psychopathology. The John Hopkins University Press, London.

Kendler, K.S., Gallagher, T.J., Abelson, J.M., Kessler, R.C., 1996. Lifetime prevalence, demographic risk factors, and diagnostic validity of nonaffective psychosis as assessed in a US community sample. The National Comorbidity Survey. Arch. Gen. Psychiatry 53 (11), 1022–1031.

Lambert, M.E., Andrews, R.H., Rylee, K., Skinner, J.R., 1987. Equivalence of computerized and traditional MMPI administration with substance abusers. Comput. Hum. Behav. 3 (2), 139–143.

Leckman, J.F., Sholomskas, D., Thompson, W.D., Belanger, A., Weissman, M.M., 1982. Best estimate of lifetime psychiatric diagnosis: a methodological study. Arch. Gen. Psychiatry 39 (8), 879–883.

Meehl, P., 1959. Some ruminations on the validation of clinical procedures. Can. J. Psychol. 13, 102–128.

Moldin, S.O., Gottesman, I.I., Erlenmeyer-Kimling, L., 1987. Psychometric validation of psychiatric diagnoses in the New York High-Risk Study. Psychiatry Res. 22 (2), 159–177.

Munley, P., Busby, R., Jaynes, G., 1997. MMPI-2 findings in schizophrenia and depression. Psychol. Assess. 9, 508–511.

Nordgaard, J., Parnas, J., 2014. Self-disorders and the schizophrenia spectrum: a study of 100 first hospital admissions. Schizophr. Bull. 40 (6), 1300–1307.

Nordgaard, J., Revsbech, R., Saebye, D., Parnas, J., 2012. Assessing the diagnostic validity of a structured psychiatric interview in a first-admission hospital sample. World Psychiatry 11 (3), 181–185.

Nordgaard, J., Sass, L.A., Parnas, J., 2013. The psychiatric interview: validity, structure, and subjectivity. Eur. Arch. Psychiatry Clin. Neurosci. 263 (4), 353–364.

Parnas, J., 2002. On defining schizophrenia. In: May, M., Sartorius, N. (Eds.), Schizophrenia, 2nd ed. Wiley, pp. 43–45.

Parnas, J., 2012. The nature of the psychiatric object and classification. In: Kendler, K., Parnas, J. (Eds.), Philosophical Issues in Psychiatry II: Nosology. Oxford University Press, Oxford, pp. 118–123.

Parnas, J., 2015. Differential diagnosis and current polythetic classification. World Psychiatry 14 (3), 284–287.

Parnas, J., Bovet, P., Zahavi, D., 2002. Schizophrenic autism: clinical phenomenology and pathogenetic implications. World Psychiatry 1 (3), 131–136.

Rusaka, M., Rancans, E., 2014. A prospective follow-up study of first-episode acute transient psychotic disorder in Latvia. Ann. Gen. Psychiatry 13 (1), 4.

Sellbom, M., Bagby, R.M., Kushner, S., Quilty, L.C., Ayearst, L.E., 2012. Diagnostic construct validity of MMPI-2 restructured form (MMPI-2-RF) scale scores. Assessment 19 (2), 176–186.

Steiner, J.L., Tebes, J.K., Sledge, W.H., Walker, M.L., 1995. A comparison of the Structured Clinical Interview for DSM-III-R and clinical diagnoses. J. Nerv. Ment. Dis. 183 (6), 365–369.

Tellegen, A., Ben-Porath, Y.S., Sellbom, M., 2009. Construct validity of the MMPI-2 restructured clinical (RC) scales: reply to Rouse, Greene, Butcher, Nichols, and Williams. J. Pers. Assess. 91 (3), 211–221.

Tellegen, A., Ben-Porath, Y.S., Sellbom, M., Arbisi, P.A., McNulty, J.L., Graham, J.R., 2006. Further evidence on the validity of the MMPI-2 restructured clinical (RC) scales: addressing questions raised by Rogers, Sewell, Harrison, and Jordan and Nichols. J. Pers. Assess. 87 (2), 148–171.

Ventura, J., Liberman, R.P., Green, M.F., Shaner, A., Mintz, J., 1998. Training and quality assurance with the Structured Clinical Interview for DSM-IV (SCID-I/P). Psychiatry Res. 79 (2), 163–173.

Walters, G.D., 1984. Empirically derived characteristics of psychiatric inpatients with DSM-III diagnoses of schizophreniform disorder. J. Abnorm Psychol. 93 (1), 71–79.

Wetzler, S., Khadivi, A., Moser, R.K., 1998. The use of the MMPI-2 for the assessment of depressive and psychotic disorders. Assessment 5 (3), 249–261.

Yung, A.R., Organ, B.A., Harris, M.G., 2003. Management of early psychosis in a generic adult mental health service. Aust. N. Z. J. Psychiatry 37 (4), 429–436.