Available online at www.sciencedirect.com

SciVerse ScienceDirect

Comprehensive Psychiatry 54 (2013) 556–561

Comprehensive PSYCHIATRY

www.elsevier.com/locate/comppsych

Temperament and character profiles of Japanese university student suicide completers

Nobuyuki Mitsui[a,*], Satoshi Asakura[a,b], Takeshi Inoue[a], Yusuke Shimizu[a], Yutaka Fujii[a], Yuki Kako[a], Teruaki Tanaka[a], Nobuki Kitagawa[c], Ichiro Kusumi[a]

[a]*Department of Psychiatry, Hokkaido University Graduate School of Medicine, Sapporo, Japan*
[b]*Hokkaido University's Health Care Center, Sapporo, Japan*
[c]*Department of Clinical Social Work, Health Sciences University of Hokkaido School of Nursing and Social Services, Tobetsu, Japan*

Abstract

Objective: The aims of this study were to investigate the personality traits of suicide completers using the Temperament and Character Inventory (TCI) scale.

Methods: Newly enrolled students who enrolled at Hokkaido University in 1999–2002 and 2004–2007 completed the TCI. Among these students, twenty subjects (2 females and 18 males) later completed suicide. We compared the TCI scales of these subjects with those of 60 (6 females and 54 males) well-matched controls. The controls were matched for age, gender, university department and year of enrollment in the university. Because the number of females was too small, the statistical analyses for the TCI subscales and logistic regression analysis were performed only with the 18 males.

Results: A univariate analysis of seven personality dimensions on the TCI revealed higher scores of harm avoidance (HA) in subjects with suicide completion ($P=0.034$). Analysis of the male subjects showed that suicide completers had higher scores for anticipatory worry (HA1, $P=0.007$) and fear of uncertainty (HA2, $P=0.036$) and lower scores for spiritual acceptance (ST3, $P=0.038$) than did the controls. A multivariate analysis, which was performed to adjust confounding factors, demonstrated significantly higher scores for HA1 among suicide completers ($P=0.01$, OR=1.32).

Conclusions: These results suggest that higher HA scores may predict suicide completion.

1. Introduction

An increased number of completed suicides is a serious social problem in present-day Japan. Although nationwide efforts have been made to prevent suicide, the number of suicides has not decreased. Moreover, the number of youth suicides in Japan has increased in recent years (http://www8.cao.go.jp/jisatsutaisaku/whitepaper/w-2012/pdf/gaiyou/pdf/p2-6.pdf). Kessler et al. found the highest risk of initial suicide ideation, plans, and attempts in individuals' late teens and early 20s [1]. It is important to identify predictors of suicide in young adults to prevent suicide attempts and suicide completion. However, at the present time, it is not possible to predict suicide with any degree of accuracy [2]. The identification of the predictors of suicide remains challenging.

Personality traits are one of the most important candidates for predictors of suicide because personality represents individuals' pervasive attitude toward their environment. The personality traits that are risk factors for suicide include neuroticism, introversion and hopelessness [3]. Protective factors can also be found within the structure of the personality. However, there are methodological difficulties in conducting research on personality traits related to suicide completion [3]. One important difficulty in investigating suicide completion is that we cannot reaffirm psychiatric states and personality traits retrospectively. Psychological autopsy is one of the major methods of studying suicide completion [4,5], but the retrospective nature of this approach limits its validity [6].

* Corresponding author. Department of Psychiatry, Hokkaido University Graduate School of Medicine, Sapporo 060–8638, Japan. Tel.: +81 11 716 1161x5973; fax: +81 11 706 5081.
E-mail address: nmitsui@med.hokudai.ac.jp (N. Mitsui).

0010-440X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.comppsych.2012.11.002

Prospective longitudinal studies using standardized clinical rating scales are necessary to investigate the personality traits related to suicide completion, although it is very difficult to perform these studies. Some prospective studies have investigated the relevance of suicide completion to scales such as the Beck Hopelessness Scale (BHS) for outpatients [7] and the Minnesota Multiphasic Personality Inventory (MMPI) for non-clinical college students [8]. However, no studies have used the Temperament and Character Inventory (TCI) scale to investigate the personality traits of young adults who complete suicide. The TCI is a widely used scale for assessing personality among adult samples in Japan. The TCI is a self-report measure of four dimensions of temperament (novelty seeking (NS), harm avoidance (HA), reward dependence (RD), and persistence (P)) and three dimensions of character (self-directedness (SD), cooperativeness (C), and self-transcendence (ST)).

In the present study, we investigated the personality traits of subjects who completed suicide using the TCI scale with a naturalistic follow-up design. The purpose of the present study was to compare the TCI scores of university students who completed suicide with those of controls who had similar backgrounds but did not complete suicide.

2. Methods

2.1. Subjects

Subjects were Hokkaido University students who enrolled between 1999 and 2002 and between 2004 and 2007. A total of 20,919 students enrolled at Hokkaido University during these periods. Among these students, 16,343 students (78.13%) completed the TCI (Table 1). We excluded students who did not agree to answer the questionnaire or did not complete the TCI.

In this population, 20 students had completed suicide by March 2011. All information, including the death certificates of suicide completers, was collected in the health care center of Hokkaido University. The 20 suicide completers were the subjects in the suicide completion group in this study. All subjects were Japanese, and only 2 of the 20 subjects were women.

Table 1
TCI response rates.

Year	University enrollment (N)	TCI respondents (N)	Rate of TCI respondents	Suicide completers (N)
1999	2633	2406	91.4%	3
2000	2560	2249	87.9%	1
2001	2584	2416	93.5%	3
2002	2566	2489	97.0%	4
-	-	-	-	-
2004	2662	1288	48.4%	2
2005	2643	1759	66.6%	4
2006	2627	1804	68.7%	2
2007	2644	1932	73.1%	1
Total	20919	16343	78.1%	20

TCI: Temperament and Character Inventory.

These 20 suicide completers had been assessed using the TCI at the time of enrollment at Hokkaido University. Using the full case records from Hokkaido University's health care center, we judged that these suicide completers had no history of psychiatric disorders at the time of enrollment. However, before the suicides occurred, 5 (25.0%) of the 20 suicide completers had a history of psychiatric treatment. The clinical diagnoses were schizophrenia in two students, depression in two students and social anxiety disorder in one student.

Because this study was designed as a case–control study, we randomly selected three students for each suicide completer as controls from the same population. To avoid effects of age, gender, intelligence and year of enrollment, the control group was matched for age, gender, university department and year of enrollment with the suicide completion group.

Written informed consent was obtained from all subjects prior to completion of the TCI scale. This study was approved by the Ethical Committee of Hokkaido University Graduate School of Medicine and was conducted in accordance with the ethical standards established in the 1964 Declaration of Helsinki (amended in Seoul, October 2008).

2.2. Measures

At the time of enrollment in Hokkaido University, students' personalities were assessed using the TCI scale. The TCI is a self-report measure of personality based on a theory proposed by Cloninger [9]. In this study, we used the 125-item version of the TCI with a 4-point scale because Kijima et al. showed that a 4-point scale was superior to a dichotomous scale in terms of internal consistency, as expressed by Cronbach's α coefficient [10]. This Japanese version of the TCI is a valid and reliable measure of temperament and character in a young adult population [11].

2.3. Statistical analyses

The scores of the dimensions of the TCI were compared between suicide completers and matched controls using Student's *t*-test. A logistic regression analysis was used to identify which subscale was the most important contributing personality factor for suicide completion. A stepwise analysis was conducted as a backward-stepping procedure and a forward-stepping procedure based on a likelihood ratio test with $P<0.05$. HA1 and ST3 were used for this analysis because these two subscales were significantly different between suicide completers and matched controls in the univariate analysis. Pearson's correlation was used to confirm the relationship between HA1 and ST3 in each group. A P value of <0.05 was adopted for the *t*-test and logistic regression analysis, and <0.01 was adopted for the correlated coefficients. All tests were two-tailed. We did not apply Bonferroni's correction for the TCI scores because of

the small sample size. All data were analyzed using SPSS version 17.0 (SPSS Japan, Tokyo).

3. Results

3.1. Demographic data

Demographic data are presented in Table 2. The mean age at the time of completing the TCI was 19.1±1.5 years old, and the mean age at suicide was 22.0±2.6 years old. Thus, the duration between the completion of the TCI and suicide was 33.4±23.2 months. With the exception of five subjects, the suicide completers' level of intelligence, family relationships and psychiatric assessments were not assessed.

3.2. Temperament and character profiles

The mean scores of suicide completers and matched controls for seven personality scales of the TCI are presented in Table 3. Of the seven personality scales, only the HA score was significantly higher among suicide completers than among the matched controls. Because gender differences have been reported for the TCI [12,13], and there were few female suicide completers in this study, we further analyzed only male subjects (N=18). The mean scores of the seven personality scales and their subscales among the male subjects are presented in Table 4. The HA score was also significantly higher among male suicide completers than among matched controls. In the analysis of the subscales, higher HA1 (P=0.007) and HA2 (P=0.036) scores and lower ST3 scores (P=0.038) were observed among male suicide completers compared with the controls.

3.3. Logistic regression analysis

The significantly different subscales between the two male groups in the univariate analysis, HA1 and ST3, were included in a backward stepwise logistic regression model to identify the most important risk factor for suicide completion. Prior to the logistic regression analysis, Pearson's correlation was performed to confirm the correlation between the three subscales HA1, HA2 and ST3. HA2 was excluded from the logistic regression analysis because it had a high correlation with HA1 (r=0.77). The correlation coefficients of HA1 and ST3 were r=−0.31 in suicide completers and r=−0.43 in the matched controls. The results of the backward stepwise logistic regression analysis are presented in Table 5. In step 1, the Hosmer-Lemeshow test was 0.58, and HA1 and ST3 were not significantly different. In step 2, the Hosmer-Lemeshow test was 0.61, and the HA1 score of suicide completers was significantly lower than the score of the controls (P=0.01, 95% CI; 1.06-1.63). The same result was obtained using a forward selection procedure.

Table 2
Demographic features.

	Controls N=60	Suicide completers N=20
Gender		
Male	54	18
Female	6	2
Age, years	19.1±1.5	19.1±1.5
Age at suicide, years	-	22.0±2.6
Duration between TCI and suicide, months	-	33.4±23.2

Controls were matched by gender, age, university department and year of enrollment.

Table 3
Comparisons of TCI dimensions between controls and suicide completers.

TCI	Controls N(f/m)=60(6/54)	Suicide completers N(f/m)=20(2/18)	t	P
	Mean SD	Mean SD		
Temperament				
Novelty seeking (NS)	49.5±6.8	47.7±6.8	1.03	0.305
Harm avoidance (HA)	56.3±10.8	62.2±9.9	−2.16	0.034*
Reward dependence (RD)	42.0±6.3	41.6±6.4	0.30	0.767
Persistence (P)	10.3±2.8	10.3±2.2	−0.07	0.943
Character				
Self-directedness (SD)	86.4±9.9	84.2±9.6	0.90	0.372
Cooperativeness (C)	71.4±8.3	72.1±8.7	−0.30	0.766
Self-transcendence (ST)	13.2±7.2	11.4±6.9	0.97	0.335

* P<0.05, SD, standard deviation; t, t-value; f, female; m, male.

4. Discussion

The main findings of this study were that suicide completers had a high HA score. High HA1 and HA2 and low ST3 scores were observed in the TCI subscales of male suicide completers. However, the multivariate analysis demonstrated that only a high HA1 score was a predictor for suicide completion. Although few studies have examined the personalities of suicide completers, some studies on suicide attempts and suicidal ideation have used the TCI [4,14–17]. In previous studies, high HA scores in the temperament dimensions have been reported in psychiatric patients who had histories of suicide attempts [14,15,17]. In contrast, in the character dimensions, low SD and high ST scores [15] or low SD and low C scores [14] have been reported in suicide attempters. The results of this study showed that only high HA was a common feature for suicide attempters and completers. This long-term follow-up study is the first report on the TCI scores of suicide completers. Although HA has not previously been investigated in suicide completers, a similar tendency toward a high HA score was found in suicide completers in this study and in suicide attempters in other studies.

Table 4
Comparisons of TCI subscales between male controls and male suicide completers.

TCI subscale		Controls N=54	Suicide completers N=18	Significance test and p value	
		Mean SD	Mean SD	*t*	*P*
Temperament					
NS	Novelty seeking	49.7±6.6	47.5±7.1	1.19	0.237
NS1	Exploratory excitability	12.8±2.2	12.1±2.1	1.11	0.272
NS2	Impulsiveness	10.5±2.3	10.0±2.5	0.81	0.422
NS3	Extravagance	16.9±2.4	16.9±2.7	0.00	1.000
NS4	Disorderliness	9.5±2.5	8.5±2.6	1.46	0.150
HA	Harm avoidance	55.6±10.7	62.5±10.4	−2.40	0.019*
HA1	Anticipatory worry	13.9±2.7	16.0±3.1	−2.75	0.007*
HA2	Fear of uncertainty	13.2±3.3	15.1±2.5	−2,13	0.036*
HA3	Shyness	14.6±3.6	16.1±3.4	−1.54	0.127
HA4	Fatigability and asthenia	13.9±2.9	15.4±2.8	−1.95	0.055
RD	Reward dependence	41.9±6.5	40.9±6.2	0.55	0.583
RD1	Sentimentality	9.5±2.7	8.9±2.8	0.82	0.413
RD3	Attachment	13.2±3.3	12.7±2.9	0.62	0.538
RD4	Dependence	19.1±2.9	19.3±2.2	−0.25	0.804
P	Persistence	10.3±2.8	10.2±2.1	0.16	0.877
Character					
SD	Self-directedness	86.8±9.6	83.4±9.9	1.26	0.211
SD1	Responsibility	21.0±2.5	19.9±3.1	1.50	0.138
SD2	Purposeful	16.4±2.7	17.0±2.8	−0.85	0.396
SD3	Resourcefulness	16.5±2.6	16.1±3.0	0.53	0.597
SD4	Self-acceptance	17.1±4.0	15.3±3.5	1.63	0.108
SD5	Congruent	15.9±2.7	15.1±2.2	1.06	0.293
C	Cooperativeness	72.0±8.3	70.8±8.1	0.50	0.616
C1	Social acceptance	16.6±2.4	15.5±2.9	1.55	0.125
C2	Empathy	9.4±2.0	9.0±2.3	0.75	0.453
C3	Helpfulness	15.1±2.7	15.7±2.3	−0.82	0.417
C4	Compassion	15.7±2.8	15.2±2.3	0.77	0.445
C5	Pure-hearted	15.1±2.6	15.4±2.2	−0.49	0.623
ST	Self-transcendence	13.4±7.4	10.7±6.1	1.41	0.163
ST1	Self-forgetful	5.5±2.8	4.7±2.7	1.15	0.252
ST2	Transpersonal	3.7±2.8	3.6±2.3	0.13	0.900
ST3	Spiritual acceptance	4.2±3.4	2.4±2.0	2.11	0.038*

* $P<0.05$.

In another study using the NEO-PI, neuroticism, anxiety and shyness were found to be associated with suicide completion [5]. Because HA is positively correlated with neuroticism [18], the results of the NEO-PI for suicide completion are consistent with high HA in this study.

Table 5
Stepwise logistic regression analysis in male subjects with suicide completion.

	Beta	SE	Wald	*P*	OR	95% CI	Hosmer–Lemeshow test
Step 1							
HA1	0.23	0.12	3.83	0.05	1.25	1.00-1.57	0.58
ST3	−0.18	0.14	1.70	0.19	0.84	0.64-1.09	
Constant	−3.93	1.91	4.23	0.04	0.02		
Step 2							
HA1	0.28	0.11	6.35	0.01	1.32	1.06-1.63	0.61
Constant	−5.21	1.71	9.33	0.00	0.01		

SE, standard error; Wald, Wald statistic; OR, odds ratio; CI, confidence interval; HA, harm avoidance; ST, self-transcendence.

Moreover, considering that NEO-PI neuroticism is strongly correlated with both high HA scores and low SD scores [18], high HA scores may be a more specific predictor of suicide risk than neuroticism given that SD was not shown to be a significant predictor in this study. Thus, suicide attempters have more personality disorders, which are characterized by low SD scores and low C scores [19], whereas suicide completers in this study were more likely to be at risk for severe mood disorders with high HA scores and not low SD scores. High HA is known to be associated with various mental disorders, such as depression [20–24], bipolar disorder [16,25], eating disorders [17], obsessive-compulsive disorder [26–28], panic disorder [29,30], posttraumatic stress disorder [31,32], and schizophrenia [33,34]. In a recent review [22], high HA was associated with current depressive symptoms and depressive traits. Moreover, studies of Japanese university students have reported that depressed patients had high HA [35,36]. Thus, a high score for this dimension may reflect some type of mental disorder or a personality risk of developing a mental disorder. We

speculate that suicide completers in this study might already have had mental disorders or a risk of developing mental disorders at the time of enrollment in the university.

This study found low scores for ST3 in male suicide completers. ST generally refers to identification with the aspects of the personality that are understood as essential and consequential parts of a unified whole. This understanding involves a state of "unitive consciousness" in which everything is part of one totality [9]. Individuals who have lower scores for ST can be characterized as repressive, practical, dualistic, skeptical, and materialistic [37]. The ST3 scale measures a person's apprehension of phenomena that cannot be explained by objective demonstration. A low ST3 score indicates a tendency toward materialism. In general, materialists tend to regard people and the world as deterministic machines [9]. Conversely, subjects with high scores in ST3 tend to endorse extrasensory perception and ideation, whether deities or a common unifying force. A high ST3 score seems to be an advantageous trait when individuals are confronted with critical life events and suffering.

Previous studies reported higher ST scores in depressed suicide attempters [14,15] in contrast to our results. Depressed patients also had higher ST scores in several previous reports [20,21,38]. These previous studies suggested that high ST scores are associated with the tendency toward psychosis when combined with poor development of one or both of the other character dimensions [20]. Although a history of suicide attempts is known to be the strongest risk factor for suicide completion, most attempters do not commit suicide, and suicide completers are known to have a profile that differs in important ways from that of suicide attempters [39]. Suicide completers tend to be male, use more lethal methods, and die on the first attempt [39]. A Finnish study reported that 56% of suicide completions were associated with a first attempt [40]. Low ST3 in suicide completers and high ST in suicide attempters may be an important difference between the two groups. Multivariate analysis does not support the important contribution of ST3 to suicide completion. Thus, further studies with more subjects are needed to clarify the role of ST3 in suicide completion and attempts.

A recent study that investigated the relationship between character dimensions and well-being indicated that ST has a strong impact on the awareness of participation in something beyond the individual self, which increases the experience of positive emotions [41]. Lower well-being may have been present in suicide completers at enrollment in the university. These considerations suggest that a lower ST3 score may be a characteristic that is specific to suicide completers.

This study has some limitations. First, we did not assess a psychiatric diagnosis for each case prior to the suicide. Second, the sample size was small, and the statistical power was weak. These factors may lead to an underestimation of the personality factors of suicide completers. Third, it is difficult to generalize the present results to all suicide completers because all of the subjects in this study were young Japanese students.

In conclusion, suicide completers had high HA scores, especially HA1 and HA2 on the HA subscales for male completers. These features may be associated with suicide in the future.

Acknowledgment

We are deeply grateful to Dr. Manabu Musashi, the director of Hokkaido University's health care center, for supporting this study.

This study was partly supported by "Integrated research on neuropsychiatric disorders" carried out under the Strategic Research Program for Brain Sciences by the Ministry of Education, Culture, Sports, Science and Technology of Japan, a Research Grant 24–2 for Nervous and Mental Disorders from the Ministry of Health, Labour and Welfare, Japan and a grant for Interdisciplinary Project for Psychosomatological Research in Hokkaido University.

References

[1] Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. Arch Gen Psychiatry 1999;56:617-26.
[2] Paris J. Predicting and preventing suicide: do we know enough to do either? Harv Rev Psychiatry 2006;14:233-40.
[3] Brezo J, Paris J, Turecki G. Personality traits as correlates of suicidal ideation, suicide attempts, and suicide completions: a systematic review. Acta Psychiatr Scand 2006;113:180-206.
[4] Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. Psychol Med 2003;33: 395-405.
[5] Duberstein PR, Conwell Y, Caine ED. Age differences in the personality characteristics of suicide completers: preliminary findings from a psychological autopsy study. Psychiatry 1994;57:213-24.
[6] Pouliot L, De Leo D. Critical issues in psychological autopsy studies. Suicide Life Threat Behav 2006;36:491-510.
[7] Beck AT, Brown G, Berchick RJ, Stewart BL, Steer RA. Relationship between hopelessness and ultimate suicide: a replication with psychiatric outpatients. Am J Psychiatry 1990;147:190-5.
[8] Yen S, Siegler IC. Self-blame, social introversion, and male suicides: prospective data from a longitudinal study. Arch Suicide Res 2003;7: 17-27.
[9] Cloninger CR, Svrakic DM, Przybeck TR. A psychobiological model of temperament and character. Arch Gen Psychiatry 1993;50:975-90.
[10] Kijima NSR, Takeuti M, Yoshino A, Ono Y, Kato M, et al. Cloninger's seven-factor model of temperament and character and Japanese version of Temperament and Character Inventory (TCI). Arch Psychiatr Diagn Clin Eval 1996;7:379-99.
[11] Takeuchi M, Miyaoka H, Tomoda A, Suzuki M, Lu X, Kitamura T. Validity and reliability of the Japanese version of the Temperament and Character Inventory: a study of university and college students. Compr Psychiatry 2011;52:109-17.
[12] Hansenne M, Delhez M, Cloninger CR. Psychometric properties of the temperament and character inventory-revised (TCI-R) in a Belgian sample. J Pers Assess 2005;85:40-9.
[13] Pelissolo A, Lepine JP. Normative data and factor structure of the Temperament and Character Inventory (TCI) in the French version. Psychiatry Res 2000;94:67-76.

[14] Calati R, Giegling I, Rujescu D, Hartmann AM, Moller HJ, De Ronchi D, et al. Temperament and character of suicide attempters. J Psychiatr Res 2008;42:938-45.
[15] Conrad R, Walz F, Geiser F, Imbierowicz K, Liedtke R, Wegener I. Temperament and character personality profile in relation to suicidal ideation and suicide attempts in major depressed patients. Psychiatry Res 2009;170:212-7.
[16] Engstrom C, Brandstrom S, Sigvardsson S, Cloninger R, Nylander PO. Bipolar disorder: I. Temperament and character. J Affect Disord 2004; 82:131-4.
[17] Bulik CM, Sullivan PF, Joyce PR. Temperament, character and suicide attempts in anorexia nervosa, bulimia nervosa and major depression. Acta Psychiatr Scand 1999;100:27-32.
[18] De Fruyt F, Van de Wiele L, Van Heeringen C. Cloninger's psychobiological model of temperament and character and the five-factor model of personality. Pers Individ Differ 2000;29:441-52.
[19] Cloninger CR, Zohar AH, Cloninger KM. Promotion of well-being in person-centered mental health care. Focus 2010;8:165-79.
[20] Cloninger CR. A systematic method for clinical description and classification of personality variants. A proposal. Arch Gen Psychiatry 1987;44:573-88.
[21] Hansenne M, Reggers J, Pinto E, Kjiri K, Ajamier A, Ansseau M. Temperament and character inventory (TCI) and depression. J Psychiatr Res 1999;33:31-6.
[22] Kampman O, Poutanen O. Can onset and recovery in depression be predicted by temperament? A systematic review and meta-analysis. J Affect Disord 2011;135:20-7.
[23] Celikel FC, Kose S, Cumurcu BE, Erkorkmaz U, Sayar K, Borckardt JJ, et al. Cloninger's temperament and character dimensions of personality in patients with major depressive disorder. Compr Psychiatry 2009;50: 556-61.
[24] Chien AJ, Dunner DL. The Tridimensional Personality Questionnaire in depression: state versus trait issues. J Psychiatr Res 1996;30: 21-7.
[25] Allen MH, Chessick CA, Miklowitz DJ, Goldberg JF, Wisniewski SR, Miyahara S, et al. Contributors to suicidal ideation among bipolar patients with and without a history of suicide attempts. Suicide Life Threat Behav 2005;35:671-80.
[26] Pfohl B, Black D, Noyes Jr R, Kelley M, Blum N. A test of the tridimensional personality theory: association with diagnosis and platelet imipramine binding in obsessive-compulsive disorder. Biol Psychiatry 1990;28:41-6.
[27] Lyoo IK, Lee DW, Kim YS, Kong SW, Kwon JS. Patterns of temperament and character in subjects with obsessive-compulsive disorder. J Clin Psychiatry 2001;62:637-41.
[28] Kim SJ, Kang JI, Kim CH. Temperament and character in subjects with obsessive-compulsive disorder. Compr Psychiatry 2009;50:567-72.
[29] Wachleski C, Salum GA, Blaya C, Kipper L, Paludo A, Salgado AP, et al. Harm avoidance and self-directedness as essential features of panic disorder patients. Compr Psychiatry 2008;49:476-81.
[30] Kennedy BL, Schwab JJ, Hyde JA. Defense styles and personality dimensions of research subjects with anxiety and depressive disorders. Psychiatr Q 2001;72:251-62.
[31] Evren C, Dalbudak E, Cetin R, Durkaya M, Evren B. Relationship of alexithymia and temperament and character dimensions with lifetime post-traumatic stress disorder in male alcohol-dependent inpatients. Psychiatry Clin Neurosci 2010;64:111-9.
[32] Ruchkin VV, Eisemann M, Hagglof B. Juvenile male rape victims: is the level of post-traumatic stress related to personality and parenting? Child Abuse Negl 1998;22:889-99.
[33] Guillem F, Bicu M, Semkovska M, Debruille JB. The dimensional symptom structure of schizophrenia and its association with temperament and character. Schizophr Res 2002;56:137-47.
[34] Hori H, Noguchi H, Hashimoto R, Nakabayashi T, Saitoh O, Murray RM, et al. Personality in schizophrenia assessed with the Temperament and Character Inventory (TCI). Psychiatry Res 2008;160:175-83.
[35] Tanaka E, Kijima N, Kitamura T. Correlations between the Temperament and Character Inventory and the Self-rating Depression Scale among Japanese Students. Psychol Rep 1997;80:251-4.
[36] Naito M, Kijima N, Kitamura T. Temperament and Character Inventory (TCI) as predictors of depression among Japanese college students. J Clin Psychol 2000;56:1579-85.
[37] Cloninger CR. Feeling good: the science of well being. New York: Oxford University Press; 2004.
[38] Nery FG, Hatch JP, Nicoletti MA, Monkul ES, Najt P, Matsuo K, et al. Temperament and character traits in major depressive disorder: influence of mood state and recurrence of episodes. Depress Anxiety 2009;26:382-8.
[39] Beautrais AL. Suicides and serious suicide attempts: two populations or one? Psychol Med 2001;31:837-45.
[40] Isometsa ET, Lonnqvist JK. Suicide attempts preceding completed suicide. Br J Psychiatry 1998;173:531-5.
[41] Cloninger CR, Zohar AH. Personality and the perception of health and happiness. J Affect Disord 2011;128:24-32.

Journal of Affective Disorders 150 (2013) 546–550

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Preliminary communication

Development and validation of a screening questionnaire for present or past (hypo)manic episodes based on *DSM-IV-TR* criteria

Rie Kameyama [a,1], Takeshi Inoue [a,*,1], Mai Uchida [b], Teruaki Tanaka [a], Yuji Kitaichi [a], Yasuya Nakato [a], Yoshiyuki Hayashishita [a], Yukiei Nakai [a], Shin Nakagawa [a], Ichiro Kusumi [a], Tsukasa Koyama [a]

[a] Department of Psychiatry, Hokkaido University Graduate School of Medicine, North 15, West 7, Kita-ku, Sapporo 060-8638, Japan
[b] Department of Psychiatry, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA

ARTICLE INFO

Article history:
Received 2 December 2012
Accepted 29 January 2013
Available online 7 March 2013

Keywords:
Manic episode screening questionnaire
Bipolar disorder
Self-report
Diagnosis
Misdiagnosis

ABSTRACT

Background: We developed a self-reported questionnaire, the Manic Episode Screening Questionnaire (MES), based on the eight diagnostic criteria items of *DSM-IV-TR* (hypo)manic episodes. This study was designed to determine the optimal screening methods to identify bipolar disorders among mood disorder patients of a psychiatric specialty clinic.
Methods: In 95 mood disorder patients, we assessed the operational characteristics of the MES as a screening and diagnostic instrument using a *DSM-IV-TR* diagnosis by a trained psychiatrist as a reference standard. The reference criteria were bipolar disorders. MES was used with two methods: the diagnostic algorithm and the one-question method (question #1 only). The diagnostic algorithm was regarded as fulfilled if the answers to question #1 and three or more of questions #2 to #8 were "yes", corresponding to the *DSM-IV-TR* (hypo)manic episode criteria. In different subjects, the test-retest reliability of the MES was examined.
Results: The two methods of the MES showed high specificity (0.93–0.94), high positive predictive value (0.81–0.83) and high negative predictive value (0.88–0.90), but the sensitivity scored lower (0.68–0.75). The test-retest reliability was moderate: 0.75 for the diagnostic algorithm and 0.68 for the one-question method.
Limitations: This study includes a small number of bipolar I patients. The findings might not be generalized to patients outside of this patient population.
Conclusions: The MES is useful for the screening and diagnosis of bipolar disorders among mood disorder patients in psychiatric specialty clinics. The one-question method of the MES is more convenient to use than prior questionnaires and is here recommended.

1. Introduction

The misdiagnosis or underdiagnosis of bipolar disorders, caused by the oversight of a particular past hypomanic episode, has been noted (Ghaemi et al., 2000b; Hirschfeld et al., 2003). In some reports, the rates of bipolar disorder misdiagnosed as major depressive disorder or treatment-resistant major depressive disorder have been 37% and 59%, respectively (Ghaemi et al., 2000b; Sharma et al., 2005). A delay in the correct diagnosis inevitably leads to long-term incorrect treatment plans that are not recommended by recent treatment guidelines for bipolar disorder (Yatham et al., 2009). As a result, some bipolar disorder patients cannot receive adequate treatment for long periods of time (Hirschfeld et al., 2003), resulting in chronic and/or treatment-resistant depressive episodes (Sharma et al., 2005) or rapid cycling (Ghaemi et al., 2000b). This clinical problem is more prevalent in primary care than in psychiatric specialty clinics because the diagnosis of hypomanic or manic episodes, which is sometimes puzzling even to psychiatric specialists, can be quite challenging for primary care physicians (Smith et al., 2011). To avoid the misdiagnosis of unipolar depression for bipolar disorder cases, it is necessary not only to carefully and continuously screen through interviews regarding current or past (hypo)manic episodes but also to provide information for patients and their families based on psychoeducation (Colom and Vieta, 2006).

Three self-reported questionnaires for bipolar disorder have been developed to prevent bipolar disorders from being overlooked and to obtain sufficient information from patients. Two of

* Correspondence to: Department of Psychiatry, Neural Function, Hokkaido University Graduate School of Medicine North 15, West 7, Sapporo 060-8638, Japan. Tel.: +81 11 716 1161; fax: +81 11 706 5081.
E-mail address: tinoue@med.hokudai.ac.jp (T. Inoue).
[1] Equally contributed to this paper and will be designated as co-first authors.

0165-0327/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jad.2013.01.034

these questionnaires (Mood Disorder Questionnaire, MDQ; Hypomania Checklist-32, HCL-32) were designed to screen from a lifetime history of (hypo)manic syndromes (Angst et al., 2005; Hirschfeld et al., 2000), and one (Bipolar Spectrum Diagnostic Scale, BSDS) was designed to assess mood fluctuations, such as high or low mood, and to detect the milder portions of the bipolar spectrum (Ghaemi et al., 2005). The application of these questionnaires not only to psychiatric but also to primary care clinics has been tested previously (Hirschfeld et al., 2005; Smith et al., 2011). All three questionnaires showed relatively high sensitivity (MDQ and BSDS, 0.73; HCL-32, 0.80) and specificity (MDQ and BSDS, 0.90; HCL-32, 0.51) among mood disorder patients. Although the items of the MDQ and HCL-32 overlap with the diagnostic criteria items of a *DSM-IV-TR* (hypo)manic episode (American Psychiatric Association, 2000), the quantity of items is much larger than the eight *DSM-IV-TR* criteria items (MDQ, 13 items; BSDS, 18 items and HCL-32, 32 items). Hence, these prior questionnaires complicate the understanding of the *DSM* concept of "(hypo)manic episode" and its use for psychoeducation.

We developed a self-reported, single page, paper-and-pencil questionnaire, the Manic Episode Screening Questionnaire (MES), which is based on the eight diagnostic criteria items of a *DSM-IV-TR* (hypo)manic episode. The concept of the MES is similar to that of the PHQ-9, which is based on the nine diagnostic criteria items of a *DSM-IV-TR* major depressive episode and is widely used to screen major depressive episodes and to evaluate the severity of depression (Furukawa, 2010). The MES was designed to screen a lifetime history of (hypo)manic episodes and can be used for psychoeducation and self-evaluation due to its facilitation of understanding the concept of a *DSM-IV-TR* (hypo)manic episode. The present study was designed to determine the optimal screening methods for identifying bipolar disorders among mood disorder patients and to assess the sensitivity and specificity of those methods using a diagnosis of bipolar disorder by a mood disorder specialist as the standard criteria.

2. Methods

2.1. Subjects

From February 2008 to March 2011, 293 outpatients who visited the Department of Psychiatry, Hokkaido University Hospital as new patients were consecutively included in the study. Among them, we included 95 patients who had been diagnosed with either major depressive disorder or bipolar disorder using the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (*DSM-IV-TR*) (American Psychiatric Association, 2000) by a mood disorder specialist psychiatrist (T.I.). T.I. was blinded to the MES results and had more than 20 years of clinical experience in the field of psychiatry. The Japanese version of the MES was administered to all 293 patients during their waiting time as a routine clinical task. This study was performed in accordance with the Declaration of Helsinki and was approved by the institutional review board of Hokkaido University Hospital.

2.2. MES

The Japanese version of MES (Appendix B) was originally designed by one of the authors (T.I.) to detect current or past episodes of mania or hypomania that fulfill the *DSM-IV-TR* criteria in psychiatric patients. The English version of the MES (Appendix A) was translated from its Japanese version by a bilingual psychiatrist (M.U.) and approved by other authors. The validity of the English version for the screening of current or past (hypo)manic episodes has not been confirmed. The MES consists of eight yes/no items derived from the eight diagnostic criteria items of the *DSM-IV-TR*, namely, elevated, expansive or irritable mood, inflated self-esteem, decreased need for sleep, manic speech, flight of ideas, distractibility, increase in goal-directed activity, and excessive involvement in pleasurable activities with high potential for painful consequences. The MES was self-completed by the patient in a written form. (Hypo)manic episodes were diagnosed in two ways using the MES: the diagnostic algorithm and the one-question method. The threshold of the diagnostic algorithm for diagnosing a current or past (hypo)manic episode was regarded as fulfilled if the answers to question #1 and three or more of questions #2–#8 were "yes". For the one-question method, a "yes" answer to question #1 was considered a positive test. Question #1 assesses "elevated, expansive or irritable mood" and represents an essential symptom for the *DSM-IV-TR* diagnostic criteria of (hypo)manic episodes.

2.3. Psychiatric evaluations

The *DSM-IV-TR* diagnoses of mood disorders, including major depressive disorder, minor depressive disorder and bipolar disorder, were made by a psychiatrist specializing in mood disorders (T.I.) using the *Quick Reference to Diagnostic Criteria from the DSM-IV-TR* on the same day on which the patients answered the MES. The average interview duration was 60 min. In each case, the presence of a current or past major depressive episode or a current or past (hypo)manic episode was identified.

2.4. Test-retest reliability

To assess reliability across time, a sample of 52 subjects, who were different from the 95 patients described above, was retested approximately 4–8weeks after an initial testing from April 2011 to July 2011. All subjects had been diagnosed with major depressive disorder or bipolar disorder by trained psychiatrists using the *DSM-IV-TR* criteria before the administration of the MES.

2.5. Data analysis

With respect to the validity of the criteria, we investigated the sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio for the diagnostic algorithmic threshold and the one-question method. The *DSM-IV-TR* diagnosis of "bipolar disorder" was the criterion standard.

Test-retest reliability was measured using κ coefficients of agreement.

All continuous data are presented as the means with standard deviations or 95% confidence intervals (CIs).

3. Results

3.1. Demographic characteristics and DSM-IV-TR diagnoses of subjects

The demographic characteristics and *DSM-IV-TR* diagnoses of 95 subjects are presented in Table 1. Seventy-one percent of subjects were diagnosed with major depressive disorder, and 29% of subjects were diagnosed with bipolar disorder (5% bipolar I, and 24% bipolar II). The chief complaints of most of mood disorder subjects were depressive symptoms (89%), and most were in a current major depressive episode that fulfilled the diagnostic criteria.

Table 1
Characteristics and *DSM-IV-TR* diagnoses of 95 patients

Characteristic	Value
Sex	
Female, *n* (%); male, *n* (%)	39 (41); 56 (59)
Age, mean ± SD (yr)	44.3 ± 17.4
Range	17–81
DSM-IV-TR diagnosis, *n* (%)	
Major depressive disorder	67 (71)
Bipolar disorder	28 (29)
Bipolar I disorder	5 (5)
Bipolar II disorder	23 (24)
Current episode diagnosis, *n* (%)	
Depressive episode	85 (89)
Hypomanic episode	4 (4)
Euthymic state	6 (6)

Table 2
Sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and overall accuracy for the diagnostic algorithm and the one-question method (item 1 only) of Manic Episode Screening Questionnaire (MES).

	Diagnostic algorithm of MES	One-question method of MES
True positive, *n* (%)	19 (20)	21 (22)
False positive, *n* (%)	4 (4)	5 (5)
False negative, *n* (%)	9 (9)	7 (7)
True negative, *n* (%)	63 (66)	62 (65)
Sensitivity (95% CI)	0.68 (0.55–0.76)	0.75 (0.62–0.84)
Specificity (95% CI)	0.94 (0.89–0.97)	0.93 (0.87–0.96)
Positive predictive value (95% CI)	0.83 (0.67–0.92)	0.81 (0.67–0.90)
Negative predictive value (95% CI)	0.88 (0.83–0.91)	0.90 (0.85–0.93)
Positive likelihood ratio (95% CI)	11.37 (4.82–28.66)	10.05 (4.78–21.62)
Negative likelihood ratio (95% CI)	0.34 (0.25–0.51)	0.27 (0.17–0.44)
Overall accuracy (95% CI)	0.86 (0.79–0.91)	0.87 (0.80–0.92)

95% CI, 95% confidence intervals.

3.2. *Validity of the MES to screen for bipolar disorder*

Table 2 (left half) presents the operational characteristics of the diagnostic algorithmic threshold of the MES for "bipolar disorder". This threshold had a satisfactory specificity (0.94) and a positive likelihood ratio (11.37), as well as a relatively low sensitivity (0.68). The negative likelihood ratio (0.34) was a little higher than desired.

Table 2 (right half) presents the operational characteristics of the one-question method of the MES for "bipolar disorder". The one-question method had a satisfactory specificity (0.93) and positive likelihood ratio (10.05), as well as a relatively low sensitivity (0.75). The negative likelihood ratio (0.27) was a little higher than desired. Compared with the diagnostic algorithmic threshold, the one-question method showed relatively more valid operational characteristics for bipolar disorder screening.

The *DSM-IV-TR* diagnoses of false-negative cases (9 for the diagnostic algorithm and 7 cases for the one-question method) were all bipolar II disorder. Conversely, the *DSM-IV-TR* diagnoses of false-positive cases (4 for the diagnostic algorithm and 5 cases for the one-question method) were major depressive disorder with cyclothymic disorder in 1 case and atypical depression in 1 case. However, 2 false-positive cases analyzed by both methods were converted to bipolar disorder within 1–2 years after the MES was administered. Therefore, the actual number of false positive cases was lower than that shown in Table 2; the actual values of sensitivity, specificity and positive predictive value confirmed by follow-up were higher by 1–8% than those in Table 2.

Because fewer cases were diagnosed with bipolar I disorder (5 cases), the comparison of the utility of the MES in bipolar I and bipolar II disorders was impossible. However, the operational characteristics of the MES for both disorders showed similar tendencies (data not shown).

3.3. *Test-retest reliability of the MES for the screening of bipolar disorder*

To assess reliability across time, a different sample of 52 subjects was retested approximately 4–8 weeks after an initial testing. All subjects had been diagnosed with major depressive disorder (29 cases, 56%) or bipolar disorder (23 cases 44%; 2 bipolar I and 21 bipolar II). The kappa values for test-retest reliability were moderate: 0.75 for the diagnostic algorithm and 0.68 for the one-question method.

4. Discussion

This study revealed that the MES has a high sensitivity (0.68–0.75), specificity (0.93–0.94), positive predictive value (0.81–0.83), and negative predictive value (0.88–0.90) for the screening of bipolar disorder among mood disorder patients in the setting of a clinic specializing in psychiatry. These values are comparable with those from the conventional self-report questionnaires, the MDQ, BSDS, and HCL-32 (Angst et al., 2005; Ghaemi et al., 2005; Hirschfeld et al., 2000). Moreover, because both the one-question method and the diagnostic algorithm of MES showed similarly excellent operational

characteristics, the one-question method, which includes a single question to screen for bipolar disorders, is more advantageous and efficient than other screening methods. The test-retest reliability was excellent (0.68–0.75). Therefore, we conclude that the MES questionnaire, particularly the one-question method, is a simple, easy, and reliable screening tool for bipolar disorders.

As expected by the relatively low sensitivity of the one-question method of the MES, there were many false-negative results. This is one limitation on the use of the MES to screen for bipolar disorder. This limitation is, however, also true of the MDQ, BSDS, and HCL-32 (Angst et al., 2005; Ghaemi et al., 2005; Hirschfeld et al., 2000). One possible explanation for the low sensitivity is that reliance on patient self-reports may contribute to underawareness of mania because their insight is more impaired in mania than in depression (Ghaemi et al., 2000a). False-negative cases were all diagnosed as bipolar II disorders, and 4 of 7 and 2 of 7patients were depressed and hypomanic at testing, respectively. In the 2 hypomanic patients, their insight on hypomanic episodes might be impaired, but in the 4 depressed patients, not only poor insight into hypomania but also ignorance of the concept of a hypomanic episode might contribute to a false negative. Future study as to whether patients or their families are even really familiar with the concept of (hypo)manic episodes should be performed.

In this study, there were significantly fewer bipolar I patients than bipolar II patients, and this constitutes another limitation. Accordingly, we could not compare the utility of the MES for bipolar I and bipolar II patients. The utility of the MES for both bipolar disorders will be examined in the future. In addition, although such an easy-to-use questionnaire is useful in primary care as described in the introduction, one cannot extrapolate the clinical utility of the MES in primary care from the results of this study. For example, the positive predictive value may be lower in primary care, in which the prevalence of bipolar disorder is expected to be lower than that in psychiatric clinics (Akobeng, 2007). Hence, further study to test the operational characteristics of the MES using the one-question method and the diagnostic algorithm threshold is needed in primary care settings. Finally, the English version of the MES in Appendix A has been approved by us but has not been validated for the screening of bipolar disorders in native English-speaking patients. A validation of the English version of the MES will be necessary.

The MES shares some content with previous self-report questionnaires, the MDQ and the HCL-32, which ask patients about a lifetime history of (hypo)manic syndromes. As described in the introduction, that the MES includes a smaller number of items constitutes an advantage over the MDQ and HCL-32, particularly because the 8 items of the MES correspond to the *DSM-IV-TR* criteria and can be easily used for psychoeducation on (hypo)manic episodes. Moreover, this study indicates that the first question alone is enough to screen for a lifetime history of (hypo)manic episodes. To the best of our knowledge, the MES is the first one-question case-finding instrument for the screening of (hypo)manic episodes.

In conclusion, this study clarified that question 1 alone of the simple self-reported questionnaire MES is useful for the screening of (hypo)manic episodes. The validity of the MES as a screening tool, through both the diagnostic algorithm threshold and the one-question method, needs to be confirmed in a primary care setting, and the validity of the English version of the MES needs to be examined. All eight items can be used for both screening and psychoeducation. The MES is clinically useful as a diagnostic tool to prevent the misdiagnosis or underdiagnosis of bipolar disorder and can promote the recognition of undiagnosed bipolarity. However, the gold standard for the diagnosis of bipolar disorder remains the *DSM* criteria. For this reason, caution should be exercised against overestimating the accuracy of this screening tool.

Role of funding source

This study was partly supported by "Integrated research on neuropsychiatric disorders" carried out under the Strategic Research Program for Brain Sciences by the Ministry of Education, Culture, Sports, Science and Technology of Japan, a Research Grant 24-2 for Nervous and Mental Disorders from the Ministry of Health, Labour and Welfare, Japan and a grant for Interdisciplinary Project for Psychosomatological Research in Hokkaido University.

Conflict of interest

The authors report no financial or other relationship that is relevant to the subject of this article.

TI has received honoraria from GlaxoSmithKline, Pfeizer, Astellas, Eli Lilly, Mitsubishi Tanabe Pharma, Mochida Pharmaceutical, Otsuka Pharmaceutical, Meiji Seika Pharma, Asahi Kasei Pharma, Shionogi, Janssen Pharmaceutical, Takeda Pharmaceutical and Yoshitomi Pharmaceutical, has received research/grant support from Otsuka Pharmaceutical, and is a member of the advisory boards of GlaxoSmithKline, Eli Lilly, Mochida Pharmaceutical and Mitsubishi Tanabe Pharma.

SN has received honoraria from GlaxoSmithKline, Eisai, Pfeizer, Daiichi-Sankyo, Meiji Seika Pharma, Ono Pharmaceutical and Eli Lilly, and has received research/grant support from Pfeizer, Eli Lilly, Eisai and Ono Pharmaceutical.

IK has received honoraria from Astellas, Eli Lilly, and Dainippon Sumitomo Pharma, and has received research/grant support from Otsuka Pharmaceutical, Astellas and Dainippon Sumitomo Pharma, and is a member of the advisory board of Dainippon Sumitomo Pharma.

TK has received honoraria from GlaxoSmithKline, Astellas, and Eli Lilly, has received research/grant support from Astellas and GlaxoSmithKline, and is a member of the advisory boards of GlaxoSmithKline and Mitsubishi Tanabe Pharma.

The other authors declare that they have no actual or potential conflict of interest.

Appendix A. English version of Manic Episode Screening Questionnaire (MES)

Please answer yes or no to the following questions and circle your responses.

1. Have you ever had episodes of being extremely happy, energized or irritable, or have you felt that your condition is much better than usual for longer than a few days?	Yes	No
If you answered yes to No.1, please answer the following questions.		
2. During these episodes, were you more confident than usual?	Yes	No
3. During these episodes, were you able to operate without getting much sleep?	Yes	No
4. During these episodes, were you more talkative than usual?	Yes	No
5. During these episodes, did you come up with lots of ideas, one after another?	Yes	No
6. During these episodes, did your interests constantly shift?	Yes	No
7. During these episodes, were you active and enthusiastic about engaging in activities?	Yes	No
8. During these episodes, did you shop, gamble, make financial investments, or pursue romantic or sexual relationships more than usual?	Yes	No

以下の質問があなたにあてはまりましたら「はい」に○を、あてはまらなければ「いいえ」に○をつけてください。

1. これまでの人生で、気分高揚し、ハイテンションで、怒りっぽく、普段の調子(100%)を超えた時期が数日以上続いたことがありますか？	はい　いいえ
1で「はい」に○をつけた方は以下の質問にお答え下さい	
2. その時、いつもより自信がありましたか？	はい　いいえ
3. その時、あまり寝なくても平気でしたか？	はい　いいえ
4. その時、いつもよりよくしゃべりましたか？	はい　いいえ
5. その時、いろいろな考えが次々に思いつきましたか？	はい　いいえ
6. その時、次々に関心や興味がうつりましたか？	はい　いいえ
7. その時、活発・精力的に活動できましたか？	はい　いいえ
8. その時、買い物・賭け事・投資・異性との交際などが多くなりましたか？	はい　いいえ

Appendix B. Japanese version of Manic Episode Screening Questionnaire (MES)

References

Akobeng, A.K., 2007. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. Acta Paediatrica 96, 338–341.

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR. American Psychiatric Association, Washington, DC.

Angst, J., Adolfsson, R., Benazzi, F., Gamma, A., Hantouche, E., Meyer, T.D., Skeppar, P., Vieta, E., Scott, J., 2005. The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients. Journal of Affective Disorders 88, 217–233.

Colom, F., Vieta, E., 2006. Psychoeducation Manual for Bipolar Disorder. Cambridge University Press, Cambridge.

Furukawa, T.A., 2010. Assessment of mood: guides for clinicians. Journal of Psychosomatic Research 68, 581–589.

Ghaemi, N., Sachs, G.S., Goodwin, F.K., 2000a. What is to be done? Controversies in the diagnosis and treatment of manic-depressive illness. World Journal of Biological Psychiatry 1, 65–74.

Ghaemi, S.N., Boiman, E.E., Goodwin, F.K., 2000b. Diagnosing bipolar disorder and the effect of antidepressants: a naturalistic study. Journal of Clinical Psychiatry 61, 804–808.

Ghaemi, S.N., Miller, C.J., Berv, D.A., Klugman, J., Rosenquist, K.J., Pies, R.W., 2005. Sensitivity and specificity of a new bipolar spectrum diagnostic scale. Journal of Affective Disorders 84, 273–277.

Hirschfeld, R.M., Cass, A.R., Holt, D.C., Carlson, C.A., 2005. Screening for bipolar disorder in patients treated for depression in a family medicine clinic. Journal of the American Board of Family Practice 18, 233–239.

Hirschfeld, R.M., Lewis, L., Vornik, L.A., 2003. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. Journal of Clinical Psychiatry 64, 161–174.

Hirschfeld, R.M., Williams, J.B., Spitzer, R.L., Calabrese, J.R., Flynn, L., Keck Jr., P.E., Lewis, L., McElroy, S.L., Post, R.M., Rapport, D.J., Russell, J.M., Sachs, G.S., Zajecka, J., 2000. Development and validation of a screening instrument for bipolar spectrum disorder: the mood disorder questionnaire. American Journal of Psychiatry 157, 1873–1875.

Sharma, V., Khan, M., Smith, A., 2005. A closer look at treatment resistant depression: is it due to a bipolar diathesis? Journal of Affective Disorders 84, 251–257.

Smith, D.J., Griffiths, E., Kelly, M., Hood, K., Craddock, N., Simpson, S.A., 2011. Unrecognised bipolar disorder in primary care patients with depression. British Journal of Psychiatry 199, 49–56.

Yatham, L.N., Kennedy, S.H., Schaffer, A., Parikh, S.V., Beaulieu, S., O'Donovan, C., MacQueen, G., McIntyre, R.S., Sharma, V., Ravindran, A., Young, L.T., Young, A.H., Alda, M., Milev, R., Vieta, E., Calabrese, J.R., Berk, M., Ha, K., Kapczinski, F., 2009. Canadian network for mood and anxiety treatments (CANMAT) and international society for bipolar disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2009. Bipolar Disorder 11, 225–255.

Neuropsychiatric Disease and Treatment

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Differences between bipolar and unipolar depression on Rorschach testing

Hiromi Kimura
Akemi Osaki
Rui Kawashima
Takeshi Inoue
Shin Nakagawa
Katsuji Suzuki
Satoshi Asakura
Teruaki Tanaka
Yuji Kitaichi
Takuya Masui
Nobuki Kitagawa
Yuki Kako
Tomohiro Abekawa
Ichiro Kusumi
Hiroyoshi Yamanaka
Kenzo Denda
Tsukasa Koyama

Department of Psychiatry, Hokkaido University Graduate School of Medicine, Kita-ku, Sapporo, Japan

Background: The bipolar-unipolar distinction in patients with a major depressive episode is the most important issue related to the diagnosis and treatment of mood disorders, but remains unresolved. This study was undertaken to compare bipolar and unipolar depression on Rorschach testing using the Comprehensive System with reference to healthy Japanese controls.

Methods: Patients with bipolar or unipolar depression who had undergone the Rorschach test for routine clinical purposes were followed up naturalistically for a long period. Based on diagnostic confirmation after long-term follow-up, scores on this test for patients with bipolar and unipolar depression were compared with those published elsewhere for healthy Japanese controls.

Results: The bipolar depression group showed significantly higher scores or positive findings in five variables of the Rorschach test, ie, WSum6, DR2 > 0, (CF + C) > FC + 2, PureC > 1, and Populars > 7, as assessed using the Comprehensive System, than did the unipolar depression group and healthy controls. These scores did not differ between the unipolar depression and control groups.

Conclusion: The results of this study show thought disorder or cognitive slippage and marked laxness in modulating emotion in bipolar depression, indicating the psychopathological characteristics of bipolar disorder.

Keywords: bipolar depression, bipolar disorder, Rorschach test, thought disorder, unipolar depression

Introduction

The bipolar–unipolar distinction in patients with a major depressive episode is crucial for diagnosis and treatment of mood disorders. The index episode in two thirds of bipolar patients is a depressive one,[1] and unipolar-to-bipolar conversion in most bipolar patients with depressive onset occurs within 5–9 years.[2] Therefore, major depressive disorder is likely to include false unipolar depression, ie, unrecognized bipolar disorder, the prevalence of which is about 10%.[2] Because the current guidelines for major depressive disorder and bipolar disorder differ greatly,[3–5] misdiagnosis should be avoided. Introduction of antidepressant therapy without a mood stabilizer may cause manic switch and rapid cycling. Early diagnosis of unrecognized bipolar disorder before the first manic episode is necessary.

Clinical symptoms, natural course, personality, and genetics might differentiate bipolar depression from unipolar depression. Family history of bipolar disorder in a first-degree relative is the most reliable predictor of bipolarity aside from occurrence and recognition of manic episodes.[6] However, differential diagnosis of bipolar and unipolar depression remains difficult clinically. Several investigators have attempted

Correspondence: Takeshi Inoue
Department of Psychiatry, Hokkaido University Graduate School of Medicine, North 15, West 7, Sapporo 060-8638, Japan
Tel +811 1706 5160
Fax +811 1706 5081
Email tinoue@med.hokudai.ac.jp

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/NDT.S42782

to discriminate bipolar–unipolar using biological tests.[6] Among such studies, serotonin-induced platelet intracellular calcium mobilization[7] and increased cerebral blood flow in the frontal lobe assessed using multichannel near-infrared spectroscopy[8] showed clear differences between bipolar and unipolar depression. In addition, self-questionnaires for temperament and personality, such as the TEMPS-A (Temperament Evaluation of Memphis, Pisa, Paris and San Diego-autoquestionnaire version) and MMPI (Minnesota Multiphasic Personality Inventory), might be useful.[9,10]

The Rorschach test is a sensitive performance-based personality assessment instrument that has been used in psychiatric practice for nearly a century, and has been introduced as a standard battery in the American Psychiatric Association textbook.[11] There have been controversies over the validity of the Rorschach test. For example, use of Comprehensive System normative data as a control group as reported by Exner, nonblinded judgments of coding, poor interrater reliability, selective reporting of results, and failure to control alpha level, have been criticized.[12] Only two studies using the Rorschach test for comparison of bipolar depression with unipolar depression have been reported in the relevant literature.[13,14] The reports describe more responses based on external characteristics of a blot, more color-naming responses, and more cognitive slippage by subjects with bipolar depression than by those with unipolar depression. However, these studies had some drawbacks, ie, the possibility exists that major depressive disorder unintentionally included latent bipolar disorder because periods from the onset of mood disorder to the final observation were not described, and the studies lacked a healthy control group that would have determined which of the unipolar or bipolar groups differs from healthy controls. Furthermore, because one study did not use the Comprehensive System developed by Exner,[15–17] which is widely used as an interpretation and quantitative method of the Rorschach test now, it is not possible to compare their data with those of other studies.[14]

The aim of this study was to compare bipolar and unipolar depression in the Rorschach test using the Comprehensive System developed by Exner,[15–17] with reference to data from a healthy control group in Japan.[18] No reported study has compared these three groups using the Rorschach test. The Comprehensive System might be appropriate as an examination for differential diagnosis of unipolar and bipolar depression if results of the Rorschach test differ between those of people with bipolar and unipolar depression. In these patients, periods from the onset of mood disorder to the final observation were 5 years or longer. Therefore, diagnosis of unipolar or bipolar disorder was confirmed in our subjects to some extent, although not completely. In addition, to limit the drawbacks of the Rorschach test mentioned previously, coding and scoring were blind, and two groups of bipolar and unipolar depression were statistically compared.

Materials and methods

Subjects

This study was a naturalistic retrospective study of adult patients with major depressive disorder (unipolar depression group, n = 20), bipolar I disorder (most recent episode depressed, n = 10), or bipolar II disorder (n = 10), (the latter two designated as bipolar depression group), who had been treated at the Department of Psychiatry, Hokkaido University Hospital. Diagnoses of mood disorder were made in accordance with the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition) by experienced psychiatrists responsible for these patients.[19] The patients underwent Rorschach testing during 1996–2005 when their depression had become slight to moderate with pharmacotherapy. The severity of depression was evaluated at the time of testing using the Clinical Global Impression scale. At the time of testing, the results of this test were used clinically for the understanding of personality, but did not affect the bipolar–unipolar distinction. They had a total intelligence quotient of at least 70 on the WAIS-R (Wechsler Adult Intelligence Scale-Revised). Depressed patients with brain magnetic resonance imaging or electroencephalographic evidence of organic brain disease or a concurrent severe medical problem were excluded from this study. Patients who were diagnosed with bipolar spectrum disorder[20] or substance-induced mood disorder[19] were excluded from the unipolar depression group. These patients were followed up until December 2007. Their diagnoses were confirmed during the follow-up period, which was 5 years or more from onset of their mood disorder, given that much bipolar conversion takes place during the first 5 years.[2] This study was performed in accordance with the Declaration of Helsinki and approved by the institutional review board of Hokkaido University Hospital.

Rorschach test

The full Rorschach test (10 blots) was ordered by attending psychiatrists for diagnosis and treatment in routine clinical work. The test was administered to the participants by clinical psychologists. Both the administration and inquiry phases of the test were conducted in accordance with standard procedures as outlined in the Comprehensive System by Exner.[15–17] All test responses were handwritten

by the examiner and recorded on one location sheet per participant using the Structural Summary Blank published by Rorschach Workshops. Standard seating procedures were followed.[17] Scoring of protocols were conducted individually in accordance with the most current available version of the Comprehensive System[16] by two expert clinical psychologists (HK, AO, or RK) who had been trained extensively in the Comprehensive System and were blinded to each subject's identity and disease status. The coding of each was reviewed. Substantial coding disagreement was discussed and resolved between the senior reviewers. The protocol was not included in the sample if the total responses (R) were fewer than 14. Scores on this test for the unipolar depression and bipolar depression groups were compared with those of the healthy Japanese controls reported by Nakamura et al.[18]

Data analyses

Continuous data are presented as the mean ± standard deviation. Statistical analyses were done using JMP 8 software (SAS Institute Inc, Cary, NC, USA), and GraphPad Prism 4, (GraphPad Software, La Jolla, CA, USA). For dichotomous variables, the Kruskal–Wallis test between the three groups (unipolar depression, bipolar depression, and controls) was used to calculate the P values followed by Dunn's multiple comparison test. For all other continuous variables, if the homogeneity of variance was confirmed by Levene's test, one-way analysis of variance and the Tukey–Kramer post hoc test were used. However, if the heterogeneity of variance was confirmed by Levene's test, then Welch's analysis of variance and Scheffe's post hoc test were used.

Univariate logistic regression analysis was used to screen variables that were significantly associated with a bipolar diagnosis. Multivariate logistic regression analysis was performed using features that were significant in the univariate analysis to identify independent factors for a bipolar or unipolar diagnosis and to control for confounding and interactions. Excel Statistics for Macintosh (Esumi Co, Ltd, Tokyo, Japan) was used for logistic regression analysis. Differences were considered to be statistically significant at $P < 0.05$. Inter-rater reliability (κ) was analyzed for various variables.

Results

Clinical and patient data

Patient clinical and demographic characteristics for the unipolar depression and bipolar depression groups are shown in Table 1. The interval between onset of the mood disorder to Rorschach testing and the final observation was longer in the bipolar depression group than in the unipolar depression group. Duration from onset to the final observation was adequate in both groups to confirm

Table 1 Clinical and demographic characteristics for the bipolar depression, unipolar depression, control subjects

	BD group	UD group	
Subjects (n)	20	20	
Final diagnosis (n)	Bipolar I (n = 10) Bipolar II (n = 10)	MDD, single episode (n = 14) MDD, recurrent (n = 6)	
Inpatient/outpatient at the Rorschach test			
Inpatient	14	15	NS
Outpatient	6	5	
Period from onset of mood disorder, years, mean ± SD			
To Rorschach test	8.2 ± 6.0	3.3 ± 2.1	t-test, $P < 0.01$
To final observation	13.6 ± 7.1	9.3 ± 2.3	t-test, $P < 0.05$
CGI, mean ± SD	2.5 ± 0.69	2.85 ± 0.88	Mann–Whitney U test, NS
Total IQ of WAIS-R, mean ± SD	94.0 ± 13.6	95.7 ± 15.5	NS

	BD group	UD group	Control group	
Subjects (n)	20	20	240	
Gender				
Male	10	7	89	NS
Female	10	13	151	
Age	38.2 ± 12.9	45.1 ± 16.6	31.6 ± 10.5	Welch's ANOVA, $P = 0.001$ U**, B* > C
Education, mean ± SD (years)	14.0 ± 2.1	13.8 ± 1.9	12.5 ± 5.9	NS

Notes: **$P < 0.01$; *$P < 0.05$.

Abbreviations: ANOVA, analysis of variance; NS, not significant; BD, bipolar depression; IQ, intelligence quotient; MDD, major depressive disorder; WAIS-R, Wechsler Adult Intelligence Scale-Revised; CGI, Clinical Global Impression scale (1, normal; 2, slightly ill; 3, mildly ill; 4, moderately ill; 5, markedly ill; 6, severely ill; 7, extremely ill); SD, standard deviation; B, BD, bipolar depression; U, UD, unipolar depression; C, control.

bipolarity (average > 9 years). Depressive symptoms were slight to moderate in both groups. Further, the severity of the Clinical Global Impression score was not different between the two groups. Compared with the control group, subjects with bipolar depression or unipolar depression were older, but gender ratios and education did not differ (Table 1).

Rorschach data

κ coefficients for inter-rater reliability were 0.86 for whole responses, 0.92 for location and space (two variables), 0.83 for DQ (+, o, v/+, v), 0.88 for determinants (11 variables), 0.73 for FQ (+, o, u, −, none), 0.92 for pairs, 0.86 for content (27 variables), 0.94 for popular, 0.83 for Z score, and 0.79 for special scores (14 variables). Inter-rater reliability in this study was excellent for most variables, except for special scores, and was comparable with that in a previous study.[18]

The average total number of responses (R) did not differ between groups (Table 2) although, if significant, it might affect the results of other analyses. The statistical analyses of continuous and dichotomous variables of the Rorschach test are shown in Tables 3–5. None of the special indices (S-CON, PTI, DEPI, CDI, HVI, and OBS) differed between the groups (data not shown).

Table 2 Descriptive statistics and *F* values for continuous variables of the Rorschach test

	BD Mean (SD)	UD Mean (SD)	Control Mean (SD)	*F*	*P*	Post hoc
R	25.10 (11.00)	20.70 (9.30)	26.25 (9.97)	2.89	0.057	Not done
Controls						
M	4.05 (3.35)	3.65 (3.08)	5.62 (3.14)	5.49	0.005	U* < C
Sum Shd	2.10 (1.71)	2.40 (2.96)	3.70 (2.88)	4.58	0.011	B* < C
EA	8.50 (6.09)	6.05 (3.63)	8.98 (4.20)	4.28	0.015	U* < C
es	7.40 (5.27)	6.60 (4.08)	9.11 (4.50)	3.88	0.022	U* < C
Affect						
Blends	3.20 (3.12)	2.65 (2.43)	4.32 (3.03)	3.88	0.022	U* < C
Processing						
D	8.55 (5.08)	6.45 (3.86)	9.97 (6.00)	3.74	0.025	U* < C
DQ+	5.90 (4.83)	5.10 (2.92)	7.83 (3.57)	7.30	0.001	U** < C
Zf	16.20 (8.04)	12.75 (5.21)	16.16 (5.71)	3.14	0.045	U* < C
Mediation						
FQo	11.95 (5.34)	9.40 (2.41)	12.88 (4.21)	6.57	0.002	U** < C
Ideation						
MQo	2.70 (2.00)	2.05 (1.79)	3.57 (1.99)	6.80	0.001	U** < C
Ma	2.00 (2.58)	1.45 (1.50)	2.75 (2.14)	4.28	0.015	U* < C
active	4.80 (4.88)	3.15 (2.41)	5.32 (3.14)	4.23	0.015	U* < C
Sum6	5.35 (3.87)	3.35 (2.58)	2.97 (2.85)	6.20	0.002	B** > C
WSum6	14.40 (12.25)	7.65 (7.34)	8.19 (8.69)	4.62	0.011	B > U*, C**
Self-perception and interpersonal perception						
Hd	0.85 (0.93)	1.45 (1.54)	2.05 (2.00)	4.27	0.015	B* < C
GHR	4.15 (2.46)	3.50 (1.64)	5.01 (2.39)	4.74	0.009	U* < C
Others						
Cg	1.40 (2.70)	1.10 (1.29)	3.13 (2.35)	11.34	0.000	B, U < C**
Hh	1.35 (1.84)	0.80 (1.01)	0.70 (1.06)	3.08	0.047	B* > C
Idio	0.40 (0.99)	0.65 (0.75)	1.36 (1.49)	6.07	0.003	B* < C

Notes: The following variables were not statistically significant between groups: (Controls) R, FM, m, F, Lambda, FM + m, D Score, AdjD; (Affect) FC, CF, Sum Color, WSumC, Blend%, Afr; (Processing) W, DQo, Zd; (Mediation) FQ+, FQu, FQ-, Popular, XA%, WDA%, X+%, X–%, Xu%, PureF%; (Ideation) MQ+ MQu, Mp, Passive, Intellect, Art; (Self-perception and interpersonal perception) Isolate, H, (H), (Hd), HC, (H)+Hd+(Hd), AG ,COP, PHR, PER, (2), 3r+(2)/R; (Others) A, Bt, Ls. R: total number of responses to any location in the blot. M: human movement, a determinant reflecting kinesthesis or activity involving humans. Sum Shd (Sum Shading): All of the shading and achromatic determinants in a record. EA: Experience Actual, the addition of Sum M and WSumC (weighted sum color). ES: Experienced Stimulation, the sum of FM + m (addition of all animal movement and all inanimate, inorganic, or insensate movement) and Sum Shd. Blends: A response that contains two or more determinants, eg, "a person with a black (ie, achromatic) hat is dancing (ie, human movement)". D: location code indicating a response involving a major detail, one identified in the Card's location table. DQ+: a developmental quality code reflecting a synthesized answer. Zf: The total number of Z scores (organizational activity) in a record. FQo: All the Form Quality ordinary in a record. MQo: All the Form Quality ordinary for Human Movement (M) in a record. Ma: human movement that is active. Active: superscript for movement determinants denoting a higher level of behavioral output than p (passive). Sum6: All of the cognitive special scores in a record (DV, INC, DR, FAB, ALOG, and CONTAM), which are described in the footnotes of Table 3. WSum6: weighted Sum6, the weighted sum of the six cognitive special scores (DV, INC, DR, FAB, ALOG, and CONTAM). Hd: content code for reference to a real human detail. GHR: good human representational response. Cg: content code for the use of clothing. Hh: content code for reference to household items. Idio (idiographic): content that is not captured by other content categories. **$P < 0.01$; *$P < 0.05$. Adapted with permission Glossary of Comprehensive System Scores. *J Pers Assess*. 2007;89 Suppl 1:S217–S220.[22]

Abbreviations: B, BD, bipolar depression; U, UD, unipolar depression; C, control; SD, standard deviation.

Table 3 Descriptive statistics and *P* values for dichotomous variables of six cognitive special scores, from which WSum6 is calculated [n (%)]

	BD (n = 20)	UD (n = 20)	Control (n = 240)	Kruskal–Wallis test, *P*	Post hoc
DV1 > 0	12 (60%)	10 (50%)	65 (27%)	0.002	B** > C
DV2 > 0	1 (5%)	0 (0%)	1 (0%)	0.061	
INC1 > 0	15 (75%)	17 (85%)	151 (63%)	0.089	
INC2 > 0	1 (5%)	1 (5%)	17 (7%)	0.889	
DR1 > 0	12 (60%)	5 (25%)	54 (23%)	0.001	B** > C
DR2 > 0	6 (30%)	1 (5%)	4 (2%)	0.000	B > U**, C**
FAB1 > 0	8 (40%)	6 (30%)	108 (45%)	0.408	
FAB2 > 0	0 (0%)	0 (0%)	11 (5%)	0.386	
ALOG > 0	1 (5%)	0 (0%)	19 (8%)	0.389	
CONT > 0	0 (0%)	0 (0%)	2 (1%)	0.846	

Notes: Some variables are subdivided to level 1 and level 2. Level 2 is more severe than level 1. DV, Deviant Verbalization, use of a neologism, an individualized meaning, or redundancy; INC (INCOM), Incongruous Combination, attribution of some aspect or activity to a response that is out of keeping with that response; DR, Deviant Response, use of an inappropriate phrase or circumstantiality; FAB (FABCOM), Fabulized Combination, use of an implausible relationship; ALOG, Autistic Logic, use of strained reasoning; CONT (CONTAM), Contamination, the merging or blending or both of two contents within one blot area. **$P < 0.01$; *$P < 0.05$. Adapted with permission Glossary of Comprehensive System Scores. *J Pers Assess.* 2007;89 Suppl 1:S217–S220.[22]

Abbreviations: B, BD, bipolar depression; U, UD, unipolar depression; C, control.

Of the ideation cluster, MQo, Ma, and active scores in the unipolar depression group were significantly lower than those of controls. The Sum6 score in the bipolar depression group was significantly higher than that in controls (Table 2). In addition, Lv2 > 0 was more frequent in the bipolar depression group than in controls (Table 4). However, these parameters were not different between the bipolar depression and unipolar depression groups. The weighted sum of the six cognitive special scores (DV, INC, DR, FAB, ALOG, and CONT), as shown in Table 3. WSum6 in the bipolar depression group differed significantly from that in the unipolar depression group and in

Table 4 Descriptive statistics and *P* values for dichotomous variables of the Rorschach test [n, (%)]

	BD (n = 20)	UD (n = 20)	Control (n = 240)	Kruskal–Wallis test, *P*	Post hoc
Controls					
Extratensive	5 (25%)	1 (5%)	16 (7%)	0.012	B* > C
Affect					
(CF + C) > FC + 2	8 (40%)	2 (10%)	40 (17%)	0.021	B > U*, C*
PureC > 1	5 (25%)	0 (0%)	14 (6%)	0.002	B > U**, C**
Mediation					
WDA% < 0.85	18 (90%)	12 (60%)	148 (62%)	0.038	B* > C
Populars > 7	6 (30%)	1 (5%)	27 (11%)	0.029	B > U*, C*
Ideation					
Lv2 > 0	6 (30%)	2 (10%)	24 (10%)	0.025	B* > C
Self perception and interpersonal perception					
SumT = 0	19 (95%)	19 (95%)	134 (56%)	0.000	B*, U* > C
COP = 0	9 (45%)	14 (70%)	89 (37%)	0.014	U* > C

Notes: The following variables were not statistically significant between groups: (Controls) Introversive, Ambitent, Pervasive Extratensive, Pervasive Introversive, Avoidant, D Score > 0, D Score = 0, D Score < 0, D Score < −1, Adj D Score > 0, Adj D Score = 0, Adj D Score < 0, Adj D Score < −1, R < 17, R > 27, FM + m < Sum Shd; (Processing) Zd > +3.0, Zd < −3.0, DQv > 2; (Mediation) XA% > 0.89, XA% < 0.70, WDA% < 0.75, X+% < 0.55, X+% > 0.20, X–% > 0.20, X–% > 0.30, Populars < 4, *P* > a + 1, Mp > Ma, MOR > 2; (Affect) FC > (CF + C) + 2, FC > (CF + C) +1, (CF + C) > FC + 1, S > 2, PureC > 0, Afr < 0.40, Afr < 0.50; (Self-perception and interpersonal perception) Ego. Index < 0.33, Ego. Index > 0.44, Fr + rF > 0, SumT ≥ 1, COP > 2, AG = 0, AG > 2, GHR > PHR, PureH = 0, PureH < 2. C: a determinant reflecting use of only color in generating the response. CF: a determinant reflecting the emphasis of color over form in generating the response. FC: a determinant reflecting the emphasis of form over color in generating the response. WDA%: a form quality calculation derived by adding FQ+, FQo, and FQu for W (Whole) and D (Detail) locations divided by the total number of W + D responses. A high score indicates that form quality for whole and detail locations is appropriate. Populars: a response that occurs with unusually high frequency, at least once in every three protocols. Lv2: Level 2 Special Score; assigned to Cognitive Special Scores (Table 3) that contain a bizarre or severe quality. H: Content code for reference to a real, whole human figure. SumT: All of the Texture determinants in a record. COP: Cooperative Movement, a score reflecting the use of movement (M, FM, m) that is positive or collaborative. **$P < 0.01$; *$P < 0.05$. Adapted with permission Glossary of Comprehensive System Scores. *J Pers Assess.* 2007;89 Suppl 1:S217–S220.[22]

Abbreviations: B, BD, bipolar depression; U, UD, unipolar depression; C, control.

Table 5 Odds ratios for a bipolar or unipolar diagnosis according to Rorschach variables and other features

Rorschach variables and other features	Univariate analysis: OR (95% CI), *P* value	Multivariate analysis: OR (95% CI), *P* value
Age	0.97 (0.93–1.01), 0.15	Not done
Gender	0.54 (0.15–1.92), 0.339	Not done
Period from onset to Rorschach	1.38 (1.08–1.77), 0.009**	1.40 (1.09–1.81), 0.009**
R	1.05 (0.98–1.12), 0.190	Not done
(CF + C) > FC + 2	6.00 (1.08–33.27), 0.040*	8.61 (1.28–58.02), 0.027*
Populars > 7	8.14 (0.88–75.48), 0.065	Not done
PureC > 1	97787 (0– very large), 0.876	Not done
DR2 > 0	8.14 (0.88–75.48), 0.065	Not done
WSum6	1.07 (1.00–1.15), 0.054	Not done

Notes: **$P < 0.01$; *$P < 0.05$. Values calculated by multivariate logistic regression analysis using variables that showed statistical significance in univariate analysis as independent variables.

Abbreviations: CI, confidence interval; OR, odds ratio.

the control group (Table 2). These special cognitive scores indicate specific verbal expressions, such as deviant verbalization, inappropriate combinations, deviant response, and inappropriate logic. WSum6 expresses the severity of thought disorder and cognitive slippage, as well as primary process thinking.[13,16,21] As Table 3 shows, DR2 > 0 (deviant responses, level 2) was significantly more frequent in the bipolar depression group than in the unipolar depression and control groups. Level 2 responses for special scores indicate that inappropriate thought or deviation from the theme is more marked.[16] Other special scores (DV1 and DR1) were more frequent in the bipolar depression group than in the control group (Table 3). Consequently, more frequent DR2 responses in conjunction with DV1 and DR1 responses contributed to the higher WSum6 scores in the bipolar depression group than in other groups.

In the affect cluster, two scores, ie, (CF + C) > FC + 2 and PureC > 1, were significantly more frequent in the bipolar depression group than in the other groups (Table 4). In fact, C codes a determinant reflecting use of color in generating the response, CF codes a determinant reflecting emphasis of color over form in generating the response, and FC codes a determinant reflecting emphasis of form over color in generating the response.[22] Consequently, control of the color response was laxer in the bipolar depression group. Another score, ie, Blends, was lower in the unipolar depression group, but this score did not differ between the two depressed groups (Table 2).

In the mediation cluster, Populars > 7 was significantly more frequent in the bipolar depression group than in the other groups (Table 4). Populars is a response that occurs with unusually high frequency, at least once in every three protocols.[22] Other scores, ie, FQo and WDA% < 0.85, were less frequent in the unipolar depression group and more frequent in the bipolar depression group, but these scores did not differ between the two depressed groups (Tables 2 and 4).

In other clusters (controls, processing and self-perception, and interpersonal perception), some scores differed between the bipolar depression or unipolar depression group and the control group, but did not differ between the two depressed groups (Tables 2 and 4).

To control for potential confounders, such as gender, age, years from onset of the illness to Rorschach testing, or R (number of responses to the cards), we performed a logistic regression analysis. The results of univariate and multivariate analysis are shown in Table 5. Univariate logistic regression analysis was used to screen variables that were significantly associated with a bipolar diagnosis. Selected variables were clinical features (age, gender, period from onset of illness to Rorschach test), R, and Rorschach variables that had shown significant differences between bipolar and unipolar depression. Periods from onset of illness to Rorschach testing and (CF + C) > FC + 2 were positive by univariate logistic regression analysis, but other factors were negative (Table 5). Multivariate logistic regression analysis was performed using two variables that were significant in univariate analysis to identify independent factors for a bipolar or unipolar diagnosis and to control for confounding and interactions. (CF + C) > FC + 2 was revealed to be an independent predictor of a bipolar diagnosis.

Discussion

The main finding of this study is that the bipolar depression group had significantly higher scores or positive findings in five variables of the Rorschach test, ie, WSum6, DR2 > 0, (CF + C) > FC + 2, PureC > 1, and Populars > 7, as assessed using the Comprehensive System,[16] than the unipolar depression group. These scores were also significantly different between the bipolar depression group and healthy control group, suggesting that the positive findings or higher scores observed in patients with bipolar depression are aberrant. The time interval between onset of illness to the final observation was, on average, 9.3 years or longer in both the bipolar depression and unipolar depression groups, and was of adequate duration to confirm the differential diagnosis of bipolar and unipolar depression. Therefore, it is less likely that the unipolar depression group included false unipolar depression, ie, latent bipolar disorder. The results of this

study indicate the specific findings for bipolar depression in the Rorschach test.

The WSum6 scores, which reflect cognitive slippage or thought disorder, were not different between the unipolar depression and control groups, but were higher in the bipolar depression group than in the other groups. The bipolar depression group showed a higher frequency of DR2 (deviant response, level 2) among these special scores than did other groups. Actually, DR2 represents an inappropriate phrase and circumstantiality that is bizarre or outside the bounds of reality, and also indicates thought disorder by which subjects stray from a task before them.[22] Deviant response is often observed in affective disorders and suggests severe dysfunction.[15] The Rorschach test is a sensitive instrument for detecting mild thought disorders that cannot be detected by interviews. Thought disorder and cognitive slippage are important for understanding of the psychopathology of bipolar depression.

Singer and Brabender also reported that WSum6 scores were higher in cases of bipolar depression than in cases of unipolar depression, and that bipolar mania showed the highest scores.[13] Some differences are apparent between their studies and our own, ie, DR1 and INC1 among special scores included in the WSum6 were more frequent in the bipolar depression group than in the unipolar depression group in their study, but only DR2 was more frequent in ours. In their study, DR2 was more frequent in bipolar mania than in bipolar and unipolar depression. However, it was difficult to determine whether bipolar depression and unipolar depression were aberrant or not because these findings were not compared with those in healthy controls. On the other hand, Osher et al reported higher WSum6 scores in euthymic bipolar patients and healthy children of bipolar parents than in healthy controls.[21,23] They suggested that thought disorder detected by the Rorschach test might be a psychological marker or endophenotype. Unfortunately, because they did not compare bipolar depression with unipolar depression, it was not clear whether high WSum6 scores are specific for bipolar depression. Taken together with our results, WSum6 scores are high in either a manic, depressive, or euthymic state of bipolar disorder and might be a psychological marker that can differentiate between bipolar depression and unipolar depression.

Patients who had (CF + C) > FC + 2 or PureC > 1 and those who had Populars > 7 were observed more often in the bipolar depression group than in other groups. According to Exner's interpretation, the former show a laxness in modulating emotion, and the latter show an orientation toward oversimplification and correctness.[15] These findings were not observed in a previous study of bipolar depression and mania,[13] or examined in previous studies of bipolar disorder patients and their children,[21,23] although healthy offspring of bipolar parents showed lower FC scores, which might result in more (CF + C) > FC + 2. Univariate and multivariate logistic regression analysis showed that (CF + C) > FC + 2 was as an independent predictor of a bipolar diagnosis. Because the small sample size in this study limits the possible number of independent variables in logistic regression analysis, future studies with a larger sample size are necessary to identify independent predictors among several variables. An earlier study by Donnelly et al showed that primary response to color is observed in bipolar depression but not in unipolar depression.[14] Because their study did not use the Comprehensive System, comparison with our study is difficult, but their study might point to a similar phenomenon, and probably a more PureC response. More Populars > 7 and DR2 > 0 seen in the bipolar depression group seems to be counterintuitive. DR2 is defined as an inappropriate phrase and circumstantiality that is bizarre or outside the bounds of reality, but is not always bizarre.[16,22] DR2 may include responses associated with distractibility and flight of ideas that are diagnostic for (hypo)manic episodes and are often observed, not only in (hypo)manic episodes but also in bipolar depressive episodes.[19] Thus, more Populars > 7 reflects conventional cognition or perception and seems to be able to coexist with more DR > 2.

Only slight differences were observed in psychiatric symptoms between bipolar and unipolar depression.[6,24] Therefore, it can be difficult to differentiate bipolar depression from unipolar depression by psychiatric interview only. A personal history of manic or hypomanic episodes and a family history of bipolar disorder in a first-degree relative are useful for the differential diagnosis, but cannot differentiate unrecognized latent bipolar depression, ie, without a history of manic or hypomanic episodes, from unipolar depression. The Rorschach test is a psychological instrument that needs no special facilities. It is clinically useful and can be used easily for the differential diagnosis of bipolar and unipolar depression, as reported in this study.

Limitations to the present findings of our study include a small sample size and the use of Japanese control data from another study.[18] Recruitment of local, matched nonclinical subjects as a control group is needed. A further limitation is that we did not recruit and test subjects for the Rorschach test prospectively. Using charts from our department,

we retrospectively investigated patients who had already undergone Rorschach testing and had been followed up for a long period. Test–retesting and assessment of mood/cognitive factors are fundamental for chronic, relapsing disorders such as bipolar and unipolar depression. Further, a difference was found between depressed patients and healthy controls in terms of their average age. It is unclear whether age affects the Rorschach data in adults, and this effect has not been reported. Future studies should use a larger sample size and a prospective study design for comparison with age-matched controls and should examine Rorschach test–retesting and include more precise assessment of mood/cognitive factors that can influence the Rorschach test to evaluate the over-time stability of scores and to what extent the contingent psychological condition affects the results of Rorschach testing. Finally, as described in the Introduction, the validity of the Rorschach test has been debated for decades. A recent systematic review showed that the strength of the validity evidence of most variables in the Comprehensive System, including WSum6, DR2, (CF + C) > FC + 2, and Populars, is excellent or good, but that of some variables, such as PureC, is low.[25] Therefore, the validity of each variable is heterogeneous and this heterogeneity must be noted by quoting a recent systematic review.[25]

Conclusion

Our study showed thought disorder or cognitive slippage and marked laxness in modulating emotion, ie, strong emotional expression in bipolar depression. These are not usually observed in unipolar depression or in healthy controls. Taken together with previous findings, our results indicate the psychopathological characteristics of bipolar disorder. Consequently, the Rorschach test might be useful for diagnosis of unipolar depression and false unipolar depression, ie, unrecognized bipolar depression, which is a salient issue in modern psychiatry. The predictive value of the Rorschach test for this purpose warrants further investigation.

Acknowledgments

This study was partly supported by "Integrated Research on Neuropsychiatric Disorders" carried out under the Strategic Research Program for Brain Sciences by the Ministry of Education, Culture, Sports, Science and Technology of Japan, research grant 24-2 for Nervous and Mental Disorders from the Ministry of Health, Labour and Welfare, Japan, and a grant for Interdisciplinary Project for Psychosomatological Research at Hokkaido University.

Disclosure

The authors declare that they have no conflict of interest in this work.

References

1. Daban C, Colom F, Sanchez-Moreno J, García-Amador M, Vieta E. Clinical correlates of first-episode polarity in bipolar disorder. *Compr Psychiatry*. 2006;47:433–437.
2. Akiskal HS, Maser JD, Zeller PJ, et al. Switching from 'unipolar' to bipolar II. An 11-year prospective study of clinical and temperamental predictors in 559 patients. *Arch Gen Psychiatry*. 1995;52:114–123.
3. Lam RW, Kennedy SH, Grigoriadis S, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. III. Pharmacotherapy. *J Affect Disord*. 2009;117:S26–S43.
4. Yatham LN, Kennedy SH, Schaffer A, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2009. *Bipolar Disord*. 2009;11:225–255.
5. Tränkner A, Sander C, Schönknecht P. A critical review of the recent literature and selected therapy guidelines since 2006 on the use of lamotrigine in bipolar disorder. *Neuropsychiatr Dis Treat*. 2013;9:101–111.
6. Goodwin FK, Jamison KR. *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression*, 2nd ed. New York, NY: Oxford University Press; 2007.
7. Suzuki K, Kusumi I, Sasaki Y, Koyama T. Serotonin-induced platelet intracellular calcium mobilization in various psychiatric disorders: is it specific to bipolar disorder? *J Affect Disord*. 2001;64:291–296.
8. Kameyama M, Fukuda M, Yamagishi Y, et al. Frontal lobe function in bipolar disorder: a multichannel near-infrared spectroscopy study. *Neuroimage*. 2006;29:172–184.
9. Donnelly EF, Murphy DL, Goodwin FK. Cross-sectional and longitudinal comparisons of bipolar and unipolar depressed groups on the MMPI. *J Consult Clin Psychol*. 1976;44:233–237.
10. Mendlowicz MV, Akiskal HS, Kelsoe JR, Rapaport MH, Jean-Louis G, Gillin JC. Temperament in the clinical differentiation of depressed bipolar and unipolar major depressive patients. *J Affect Disord*. 2005;84:219–223.
11. Clarkin JF, Howieson DB, McClough J. The role of psychiatric measures in assessment and treatment. In: Hales RE, Yudofsky SC, Gabbard GO, editors. *The American Psychiatric Publishing Textbook of Clinical Psychiatry*, 5th ed. Washington, DC: American Psychiatric Publishing Inc; 2009.
12. Garb HN, Wood JM, Nezworski MT, Grove WM, Stejskal WJ. Toward a resolution of the Rorschach controversy. *Psychol Assess*. 2001;13:433–448.
13. Singer HK, Brabender V. The use of the Rorschach to differentiate unipolar and bipolar disorders. *J Pers Assess*. 1993;60:333–345.
14. Donnelly EF, Murphy DL, Scott WH. Perception and cognition in patients with bipolar and unipolar depressive disorders. A study in Rorschach responding. *Arch Gen Psychiatry*. 1975;32:1128–1131.
15. Exner J Jr. *The Rorschach, Basic Foundations and Principles of Interpretation*, 4th ed. Hoboken, NY: John Wiley & Sons; 2003:1.
16. Exner J Jr. *A Rorschach Workbook for the Comprehensive System*, 5th ed. Asheville, NC: Rorschach Workshops; 2001.
17. Exner J Jr. *A Rorschach Workbook for the Comprehensive System*, 4th ed. Asheville, NC: Rorschach Workshops; 1995.
18. Nakamura N, Fuchigami Y, Tsugawa R. Rorschach Comprehensive System data for a sample of 240 adult nonpatients from Japan. *J Pers Assess*. 2007;89 Suppl 1:S97–S102.
19. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*. Washington, DC: American Psychiatric Association; 1994.

20. Ghaemi S, Ko J, Goodwin F. The bipolar spectrum and the antidepressant view of the world. *J Psychiatr Pract.* 2001;7:287–297.
21. Osher Y, Bersudsky Y. Thought disorder in euthymic bipolar patients: a possible endophenotype of bipolar affective disorder? *J Nerv Ment Dis.* 2007;195:857–860.
22. Glossary of Comprehensive System Scores. *J Pers Assess.* 2007;89 Suppl 1:S217–S220.
23. Osher Y, Mandel B, Shapiro E, Belmaker RH. Rorschach markers in offspring of manic-depressive patients. *J Affect Disord.* 2000;59: 231–236.
24. Kiejna A, Rymaszewska J, Hadryś T, Suwalska A, Łojko D, Rybakowski JK. Bipolar or unipolar? The question for clinicians and researchers. *J Affect Disord.* 2006;93:177–183.
25. Mihura JL, Meyer GJ, Dumitrascu N, Bombel G. The validity of individual Rorschach variables: systematic reviews and meta-analyses of the Comprehensive System. *Psychol Bull.* August 27, 2012. [Epub ahead of print.]