Received: 5 April 2021 | Revised: 7 August 2021 | Accepted: 7 October 2021

DOI: 10.1002/mpr.1896

ORIGINAL ARTICLE

WILEY

Comparing the associations of three psychometric scales at baseline with long-term prognosis of depression over a 10-year period

Ching-I Hung[1] | Chia-Yih Liu[1] | Shih-Chieh Hsu[1,2] | Ching-Hui Yang[3]

[1]Department of Psychiatry, Chang-Gung Memorial Hospital at Linkou and Chang-Gung University College of Medicine, Tao-Yuan, Republic of China

[2]Department of Psychiatry, New Taipei Municipal Tucheng Hospital, New Taipei City, Republic of China

[3]Department of Nursing, Mackay Medical College, New Taipei City, Republic of China

Correspondence

Ching-I Hung, Department of Psychiatry, Chang Gung Memorial Hospital, 5 Fu-Shing St, Kweishan, Taoyuan 333, Republic of China.
Email: a36250@cgmh.org.tw

Funding information

Ministry of Science and Technology Research Programs, Grant/Award Number: MOST 108-2314-B-182A-070; Chang Gung Memorial Hospital Research Programs, Grant/Award Numbers: CMRPG3G1861, CMRPG3H1781, CLRPG3D0048

Abstract

Objective: The Depression and Somatic Symptoms Scale (DSSS), a free scale, includes depression (DS) and somatic (SS) subscales. This study aimed to compare the associations of the baseline DSSS, Hamilton Depression Rating Scale (HAMD) and Hospital Anxiety and Depression Scale (HADS) scores with the outcome of depression over a 10-year follow-up period.

Methods: Two hundred ninety outpatients with major depressive disorder (MDD) were enrolled and were followed-up at the 6-month, 2-year, and 10-year points. The three scales were administered at each follow-up. Multiple linear regressions were used to compare the associations.

Results: In a comparison of the HAMD, DS, and HADS-depression, the HAMD and DS scores at baseline were most strongly associated with the HAMD score at two (6-month and 2-year) and one (10-year) follow-up points, respectively. In a comparison of the HAMD, DS, SS, HADS-depression, and HADS-anxiety, the SS and HAMD scores at baseline were most strongly associated with the HAMD score at two (6-month and 10-year) and one (2-year) follow-up points, respectively.

Conclusions: The DS, SS, and HAMD scores at baseline were significantly associated with the long-term outcome of depression. Scales or subscales assessing somatic symptoms might be more strongly associated with the outcome of depression.

KEYWORDS

depression, physical symptoms, prognosis, questionnaire, somatization

1 | INTRODUCTION

The Depression and Somatic Symptoms Scale (DSSS), a free and self-administered scale, simultaneously evaluates both depression and somatic symptoms (Hung et al., 2006a). The DSSS is composed of a 12-item depression subscale (DS) and a 10-item somatic subscale (SS), which includes five pain and five non-pain somatic symptoms. The DSSS was developed because previous depressive scales failed to bring the somatic aspect of depression to a level equal to that of psychological symptoms. The items of the DS were designed based on the

Hung CI and Liu CY made equal contributions to this article.

criteria of a major depressive episode (MDE) in the DSM-IV, while the items of the SS were selected from common somatic symptoms of depression, which could reflect the severity of depression (Hung et al., 2006b). Moreover, many previous studies reported that somatic or pain symptoms among patients with depression were associated with a poor prognosis of depression (Hung et al., 2015; Jaracz et al., 2016). Therefore, a scale for depression with a somatic element might be more strongly associated with the prognosis of depression. In light of the design of the scale, we hypothesized that the DSSS could reflect the severity of depression, be significantly correlated with other depressive scales, be sensitive to changes to treatment for depression, be used as a tool for screening depression, and be associated with the prognosis of depression.

The reliability and validity of the Chinese, Korean, and English versions of the DSSS have been established (Hung et al., 2006a; Jeon et al., 2016; Tse et al., 2018). Several validities of the DSSS among patients with depression have been tested, as follows. (1) The DS score is significantly correlated with the scores of the Hamilton Depression Rating Scale (HAMD), Montgomery–Asberg Depression Rating Scale, Center for Epidemiologic Studies Depression Scale, and mental subscales of the Short-Form 36 (Hung et al., 2006a; Hung, Wang, & Liu, 2009; Jeon et al., 2016; Tse et al., 2018). (2) The DS is sensitive to pharmacotherapy—the improvement percentage of the DS score after four weeks of pharmacotherapy is correlated with that of the HAMD score (Hung et al., 2006a). (3) The cut-off points of the DS score of ≥9 and ≥19 for non-full remission and a MDE, respectively, among patients with major depressive disorder (MDD) are of good sensitivity and specificity (Hung et al., 2012). (4) The SS score is significantly correlated with the scores of the somatization subscale of the Symptom Checklist-90-Revised, the somatic component of the HAMD, and the physical subscales of the Short-Form 36 (Hung, Liu, Cheng, & Wang, 2009; Hung, Wang, & Liu, 2009; Jeon et al., 2016). (5) The SS score at baseline predicts the outcome of depression at the 2-year follow-up point (Hung et al., 2010). (6) Principal-axis factor analysis and Mokken scale analysis demonstrated that the DSSS is of an appropriate construct validity (Chou et al., 2017; Hung et al., 2006a); moreover, the validity and reliability of the DSSS in patients with lower back pain have also been established (Liu et al., 2019).

The Hospital Anxiety and Depression Scale (HADS), which includes seven items for depression (HADS-D) and seven items for anxiety (HADS-A), does not include any somatic component (Zigmond & Snaith, 1983). This design renders the HADS able to be used to evaluate anxiety and depression without being confounded by somatic symptoms. For this reason, the HADS is one of most commonly used scales in screening for anxiety and depression among patients with medical diseases (Annunziata et al., 2020; Nikayin et al., 2016). Previous studies have investigated the cut-off scores of the HADS subscales for anxiety and depression among patients with medical diseases (Annunziata et al., 2020; de Almeida Macedo et al., 2017; Nikayin et al., 2016).

Although several studies have reported on the validity of the DSSS, no study has investigated the associations of the DSSS, HAMD, and HADS with the outcome of depression over a 10-year period among patients with MDD. It is well-known that somatic symptoms, painful physical symptoms, and anxiety symptoms have negative impacts on the outcome of depression (Hung et al., 2019; Jaracz et al., 2016; Rosellini et al., 2018). This raises an interesting question: which scale or subscale is most strongly associated with the long-term outcome of depression when depression, anxiety, and somatic scales or subscales are compared? Investigation of the above is important, because (1) the results will inform physicians and psychologists as to which component of depression is most strongly associated with the long-term prognosis of depression, and (2) the results will further prove the validity of the DSSS.

Therefore, this study aimed to compare the associations of the DSSS, HAMD, and HADS with the outcome of depression over a 10-year period. We hypothesized that scales or subscales that include assessment of appropriate somatic symptoms might be more strongly associated with the long-term outcome of depression over a 10-year period, as previous studies demonstrated that somatic and pain symptoms were associated with a poor prognosis of depression (Jaracz et al., 2016).

2 | METHODS

2.1 | Subjects

The study was conducted in the psychiatric outpatient clinic of Chang Gung Memorial Hospital, a medical center in northern Taiwan. At baseline, the study enrolled patients from January 2004 to August 2007. The inclusion criteria were consecutive outpatients aged 18–65 years who (1) fulfilled the MDD criteria and were experiencing a MDE based on the DSM-IV-text revision (TR; American Psychiatric Association, 2000); and (2) had not taken antidepressants or other psychotropic drugs within the past 1 month. In order to prevent mental symptoms from being confounded, three exclusion criteria were used: (1) catatonic features, psychotic symptoms, or severe psychomotor retardation; (2) a history of substance abuse or dependence without full remission in the past one month; and (3) chronic medical diseases such as diabetes mellitus, hypertension, and others. The subjects were interviewed by a board-certified psychiatrist using the Structured Clinical Interview for DSM-IV-TR Axis I Disorders (First et al., 2002) to confirm psychiatric diagnoses.

At baseline, 290 subjects with MDD were enrolled; they were followed-up at the 6-month, 2-year, and 10-year points. The study was approved by the Institutional Review Board of the Chang Gung Memorial Hospital. The 10-year follow-up study was conducted from August 2014 to December 2016. Written informed consent was obtained from all subjects, based on the guidelines regulated in the Declaration of Helsinki.

2.2 | Psychometric scales

The 17-item HAMD, DSSS, and HADS were used (Hamilton, 1967; Hung et al., 2006a; Zigmond & Snaith, 1983), all of which have

different psychometric characteristics. First, the HAMD is an observer-rated scale and one of the most commonly used scales for the clinical evaluation of depression (Dunlop, Granros, et al., 2019; Dunlop, Parikh, et al., 2019; Nixon et al., 2020; Sawamura et al., 2018; Vindbjerg et al., 2019; Zimmerman et al., 2018), while the DSSS and the HADS are self-administered scales. Second, the three scales have different percentages of somatic components. In the 17-item HAMD, there are eight items for somatic symptoms, including initial, middle, and terminal insomnia, loss of weight, loss of appetite, loss of libido, somatic anxiety symptoms, and general somatic symptoms; therefore, the somatic symptoms component represents 34.6% (18/52) of the total possible HAMD score. The DS of the DSSS consists of 12 items, including four somatic items (insomnia, poor appetite, fatigue, and loss of sexual desire). This design allows the DS to be compatible with the criteria of a MDE and other scales for depression. All 10 items of the SS address somatic symptoms. Therefore, the somatic components of the DS and SS represent 33.3% and 100%, respectively, of the total possible subscale scores. Conversely, the HADS does not include any somatic symptoms. Owing to this characteristic, the HADS is one of the most commonly used scales for screening depression and anxiety among patients with medical illnesses (de Almeida Macedo et al., 2017; Nikayin et al., 2016). The SS therefore has the largest somatic component, followed by the HAMD and DS; the HADS has no somatic component. The total scores range from 0 to 52 for the HAMD, 0 to 36 for the DS, 0 to 30 for the SS, and 0 to 21 for both the HADS-D and the HADS-A. A higher score indicates a greater severity.

2.3 | Procedure

At baseline and each follow-up point, subjects were requested to complete the self-administered DSSS and HADS. At baseline, the HAMD score was evaluated by one of two psychiatrists, who were blind to the results of the DSSS and HADS. The two psychiatrists had been trained together in evaluating the HAMD before the study (intraclass correlation coefficient = 0.87). At follow-up, the HAMD was evaluated by the same psychiatrist.

The course of depression fluctuates, and depressive severities at the follow-up points might be unable to represent the longitudinal course of depression. At the 10-year follow-up point, subjects were asked to grossly estimate the percentages of time spent experiencing the following depressive symptoms over the past 10 years: depressed mood, anxiety, diminished motivation, insomnia, poor appetite, fatigue, decreased concentration, poor memory, and guilty feelings.

At follow-up, the HAMD score was used as the major indicator of the outcome of depression, because the HAMD is one of the gold-standard scales used for the assessment of depression (Worboys, 2013). The DS score, HADS-D score, and percentages of time spent experiencing depressive symptoms over the past 10 years as assessed at the follow-ups were considered as secondary outcomes of depression.

The subjects accepted pharmacotherapy after enrollment. Some patients might have quit pharmacotherapy at the follow-up points, but for those who still accepted pharmacotherapy, psychiatric medications were not controlled, because this study was an observational study. Pharmacotherapy without controlling for dosages and kinds of medication might confound the severity of depression at follow-up. Subjects who were undergoing pharmacotherapy at the index month of the follow-up point were classified as the treatment group, while those who were not were classified as the non-treatment group. Therefore, subjects were divided into treatment and non-treatment groups for analysis of the correlations of scale or subscale scores at baseline with the outcome of depression at follow-up. In analyzing the correlations of scale or subscale scores at baseline with the percentages of time spent experiencing the various depressive symptoms over the past 10 years, subjects were not divided into the two groups, because the percentages of time over the past 10 years were a gross estimation of the long-term course of depression, and subjects might intermittently accept treatment over the 10-year period.

At baseline and each follow-up point, the investigators checked all collected data to avoid the problem of missing data after each subject had been assessed.

2.4 | Statistical methods

At baseline, we hypothesized that 50% of the subjects would not attend the 10-year follow-up. Using G*Power v3.1.9.2 (Faul et al., 2009), the sample size was required to be greater than 123 under the conditions of alpha level = 0.05, power = 0.8, number of predictors = 11, and effect size = 0.15. Therefore, the sample size should be greater than 246 at baseline.

SPSS for Windows 20.0 was used for statistical analyses. The Chi-square test and the independent *t* test were used when appropriate. Pearson correlation or Spearman correlation were used to test the correlations of the scores of the three scales at baseline with the HAMD score at follow-up and the percentages of time with depressive symptoms over the past 10 years.

Multiple linear regressions with forward selection were employed to identify the scale or subscale score at baseline most strongly associated with the HAMD score at the three follow-up points. This method was used for two reasons: (1) it can prevent multicollinearity; and (2) the forward method will select the most powerful factor associated with the independent variable into the regression model, followed by the second most powerful factor, then the others; therefore, it can identify the scale or subscale score that is most strongly associated with the dependent variable. The Durbin–Watson test was used to detect autocorrelation in the residuals. Collinearity is an important issue for regression models (Lang & Altman, 2015); this was examined using the Variance Inflation Factor (VIF).

Two regression models were used to compare the associations of (1) three depressive scale or subscale (HAMD, DS, and HADS-D) scores at baseline, and (2) five scale or subscale (HAMD, DS,

TABLE 1 Demographic variables and psychometric scores at baseline and three follow-up points[a]

Time point	Baseline	Six months	Two years	Ten years
Case number	290	254	237	137
Follow-up participation (%)	–	87.6	81.7	47.2
Male (%)	28.6	29.1	30.8	30.7
Age (years)	30.2 (8.2)	30.6 (8.2)	32.5 (8.4)	41.0 (8.1)
Education (years)	13.2 (2.4)	13.3 (2.4)	13.3 (2.4)	13.3 (2.5)
Employed (%)	57.2	56.7	55.7	73.0
Married (%)	42.1	42.1	43.0	52.6
With pharmacotherapy (%)	0	47.6	27.4	27.7
HAMD score	23.4 (4.2)	10.6 (7.8)	10.4 (7.4)	9.4 (6.4)
DS score	25.5 (5.3)	11.4 (8.5)	10.8 (8.0)	11.2 (8.0)
SS score	16.1 (6.6)	8.2 (6.6)	8.2 (6.2)	8.9 (5.6)
HADS-D score	14.2 (3.4)	7.7 (5.0)	7.0 (5.0)	7.2 (5.1)
HADS-A score	15.0 (3.3)	8.6 (4.8)	9.0 (4.5)	8.4 (4.7)

Abbreviations: DS, depression subscale of the Depression and Somatic Symptoms Scale (DSSS); HADS-A, anxiety subscale of the HADS; HADS-D, depression subscale of the Hospital Anxiety and Depression Scale (HADS); HAMD, Hamilton Depression Rating Scale; SS, somatic subscale of the DSSS.

[a]Continuous variables are presented as the mean (SD).

SS, HADS-D, and HADS-A) scores at baseline with the outcomes of depression at the follow-up points. In the two regression models, the dependent variable was the HAMD score at the three follow-up points. In the first model, the independent variables were the scores of the HAMD, DS, and HADS-D at baseline, with pharmacotherapy or not at the follow-up point, in addition to five demographic variables, including gender, age, marital status, duration of education, and occupation. In the second model, the independent variables were all of the independent variables in the first model with the addition of the SS and HADS-A scores at baseline.

Generalized Estimating Equation (GEE) models, which were fitted with robust error estimation and an unstructured covariance matrix, were used to estimate the associations of the three psychiatric scales or subscales at baseline with the outcomes of depression at the follow-up points. The dependent variable in the GEE models was the HAMD score, while the independent variables included eight variables at baseline (gender, duration of education, marital status, HAMD score, DS score, SS score, HADS-D score, and HADS-A score) and four at each follow-up point (age, pharmacotherapy or not, employed or not, and visit). Baseline, 6-month, 2-year, and 10-year follow-up points were considered as the first, second, third, and fourth visits. Insignificant factors were removed from the model step by step until all independent variables were significant. One of the advantages of the GEE model is that it can handle imputation for missing data; that is, the GEE model remains a robust statistical method when data are missing at random. No important bias was observed with levels of loss that varied from 5% to 60% (Kristman et al., 2004; Seaman & Copas, 2009).

A p-value < 0.05 was considered statistically significant in the statistical analyses. Moreover, Bonferroni correction was used in multiple linear regressions, and a p-value < 0.017 was considered statistically significant.

3 | RESULTS

3.1 | Subjects

At baseline, 290 MDD patients with a current MDE were enrolled. Table 1 shows the percentages of participants remaining, demographic variables, and psychometric scale scores at baseline and the three follow-up points. At the 6-month and 2-year follow-up points, 36 (12.4%) and 53 (18.3%) participants did not attend, respectively, due to "being unable to be contacted by mail or phone" (n = 15 and 27, respectively) and "refusing to participate in the follow-up study" (n = 21 and 26, respectively). At the 10-year follow-up point, 153 (52.8%) subjects did not participate, for the following reasons: 99 (34.1%) could not be contacted; 49 (16.9%) refused to participate in the follow-up program; and 5 (1.7%) for other reasons. There were no significant differences in the five demographic variables between the subjects who did and did not attend follow-up at the three points, with the exception of age at the 10-year follow-up point (with vs. without follow-up: 41.3 [8.1] vs. 39.3 [8.2] years, $p = 0.04$).

For clarity, the four footnotes "(B)," "(6M)," "(2Y)," and "(10Y)" are used to represent data collected at baseline, 6 months, 2 years, and 10 years, respectively.

TABLE 2 Correlations of depression, anxiety, and somatic severities at baseline with depressive severities at three follow-up points[a]

	Without pharmacotherapy					With pharmacotherapy[b]				
	$HAMD_{(B)}$	$DS_{(B)}$	$SS_{(B)}$	$HADS\text{-}D_{(B)}$	$HADS\text{-}A_{(B)}$	$HAMD_{(B)}$	$DS_{(B)}$	$SS_{(B)}$	$HADS\text{-}D_{(B)}$	$HADS\text{-}A_{(B)}$
$HAMD_{(6M)}$	0.38**	0.26**	0.36**	0.12	0.27**	0.11	0.09	0.27**	−0.10	0.10
$HAMD_{(2Y)}$	0.31**	0.12	0.23**	−0.04	0.14	0.14	0.14	0.25*	0.14	0.06
$HAMD_{(10Y)}$	0.17	0.25*	0.31**	−0.06	0.11	0.46**	0.47**	0.35*	0.42**	0.40*
$DS_{(6M)}$	0.34**	0.25**	0.32**	0.12	0.23**	0.20*	0.21*	0.31**	−0.05	0.14
$DS_{(2Y)}$	0.31**	0.21**	0.25**	0.02	0.19*	0.10	0.17	0.20	0.04	0.05
$DS_{(10Y)}$	0.15	0.26*	0.37**	−0.04	0.12	0.39*	0.51**	0.37*	0.31	0.36*
$HADS\text{-}D_{(6M)}$	0.20*	0.20*	0.21*	0.16	0.17	0.11	0.19*	0.30**	0.04	0.13
$HADS\text{-}D_{(2Y)}$	0.22**	0.15*	0.15	0.09	0.10	0.02	0.16	0.11	0.14	0.02
$HADS\text{-}D_{(10Y)}$	0.12	0.26**	0.16	0.24*	0.01	0.39*	0.49**	0.37*	0.44**	0.25

Note: $^{**}p < 0.01$; $^{*}p < 0.05$.

Abbreviations: DS, depression subscale of the Depression and Somatic Symptoms Scale (DSSS); HADS-A, anxiety subscale of the HADS; HADS-D, depression subscale of the Hospital Anxiety and Depression Scale (HADS); HAMD, Hamilton Depression Rating Scale; SS, somatic subscale of the DSSS.

[a]"(B)," "(6M)," "(2Y)," and "(10Y)" represent data collected at baseline and at the 6-month, 2-year, and 10-year follow-up points, respectively.

[b]Subjects undergoing pharmacotherapy in the index follow-up month.

At the 6-month, 2-year, and 10-year follow-up points, the mean (SD) scores of the HAMD in the 2 groups (subjects with pharmacotherapy vs. without pharmacotherapy) were 8.6 (6.5) vs. 12.4 (8.5; $p < 0.001$), 11.5 (7.2) vs. 10.0 (7.4; $p = 0.15$), and 13.1 (6.7) vs. 8.0 (5.7; $p < 0.001$), respectively.

3.2 | Correlations of HAMD, DS, SS, HADS-D, and HADS-A scores at baseline with outcomes of depression at the three follow-up points

At baseline, the $HAMD_{(B)}$ score was significantly (all $p < 0.001$) correlated with the $DS_{(B)}$ (correlation co-efficient $r = 0.60$), $SS_{(B)}$ ($r = 0.46$), $HADS\text{-}D_{(B)}$ ($r = 0.40$), and $HADS\text{-}A_{(B)}$ ($r = 0.41$) scores. The $DSSS_{(B)}$ score was also significantly (all $p \leq 0.001$) correlated with the $HADS_{(B)}$ score ($DS_{(B)}$ and $HADS\text{-}D_{(B)}$, $r = 0.54$; $DS_{(B)}$ and $HADS\text{-}A_{(B)}$, $r = 0.46$; $SS_{(B)}$ and $HADS\text{-}D_{(B)}$, $r = 0.20$; $SS_{(B)}$ and $HADS\text{-}A_{(B)}$, $r = 0.37$).

Table 2 shows the correlations of scale or subscale scores at baseline with the outcomes of depression at the three follow-up points. In the non-treatment group, the $HAMD_{(B)}$ score was correlated with the three depressive scale (the HAMD, DS, and HADS-D) scores at the 6-month and 2-year follow-up points. The $DS_{(B)}$ score was correlated with the three depressive scale scores at the three follow-up points, with the exception of the $HAMD_{(2Y)}$ score. The $SS_{(B)}$ score was correlated with the HAMD and DS scores at the three follow-up points and the $HADS\text{-}D_{(6M)}$ score. The $HADS\text{-}D_{(B)}$ score was not significantly correlated with the scores of the three depressive scales at the three follow-up points, with the exception of the $HADS\text{-}D_{(10Y)}$ score. The $HADS\text{-}A_{(B)}$ score was significantly correlated with the $HAMD_{(6M)}$, $DS_{(6M)}$, and $DS_{(2Y)}$ scores.

In the treatment group, the $SS_{(B)}$ score was correlated with the HAMD score at the three follow-up points; however, the $HAMD_{(B)}$ and $DS_{(B)}$ scores were not significantly correlated with the $HAMD_{(6M)}$ and $HAMD_{(2Y)}$ scores. All $HAMD_{(B)}$, $DS_{(B)}$, and $SS_{(B)}$ scores were correlated with the $HAMD_{(10Y)}$, $DS_{(10Y)}$, $HADS\text{-}D_{(10Y)}$, and $DS_{(6M)}$ scores. The $HADS\text{-}D_{(B)}$ and $HADS\text{-}A_{(B)}$ scores were significantly correlated with the $HAMD_{(10Y)}$ score, but not with the three depressive scale scores at the 6-month and 2-year follow-up points.

3.3 | Correlations of the five scale or subscale scores at baseline with self-reported percentages of time spent experiencing depressive symptoms over the past 10 years

The mean (SD) self-reported percentages of time spent experiencing depressive symptoms over the past 10 years were 42.3 (26.6) for depressive mood, 38.4 (28.3) for anxiety, 37.4 (28.5) for diminished motivation, 40.0 (31.7) for insomnia, 16.6 (21.8) for poor appetite, 48.8 (28.3) for fatigue, 38.7 (28.1) for decreased concentration, 44.3 (29.9) for poor memory, and 30.2 (30.3) for guilty feelings.

Table 3 shows the correlations of the five scale or subscale scores at baseline and the 10-year follow-up point with the self-reported percentages of time spent experiencing depressive symptoms over the past 10 years. Both the $DS_{(B)}$ and $SS_{(B)}$ scores were significantly correlated with eight of the nine symptoms, while the $HAMD_{(B)}$ score was significantly correlated with five symptoms. The $HADS\text{-}D_{(B)}$ and $HADS\text{-}A_{(B)}$ scores were not significantly correlated with any of the nine symptoms, with the exception of the $HADS\text{-}A_{(B)}$ score being correlated with guilty feelings.

All of the self-reported percentages of time spent experiencing depressive symptoms over the past 10 years were significantly correlated with the scores of the five scales or subscales at the 10-year follow-up point.

TABLE 3 Correlations of depression, anxiety, and somatic scores at baseline and the 10-year follow-up point with self-reported percentages of time spent experiencing depressive symptoms over the past 10 years[a]

	$HAMD_{(B)}$	$DS_{(B)}$	$SS_{(B)}$	$HADS\text{-}D_{(B)}$	$HADS\text{-}A_{(B)}$	$HAMD_{(10Y)}$	$DS_{(10Y)}$	$SS_{(10Y)}$	$HADS\text{-}D_{(10Y)}$	$HADS\text{-}A_{(10Y)}$
Depressed mood	0.06	0.28**	0.24**	0.10	0.10	0.62**	0.66**	0.46**	0.55**	0.57**
Anxiety	0.22*	0.28**	0.28**	−0.03	0.16	0.59**	0.60**	0.53**	0.41**	0.59**
Diminished motivation	0.15	0.31**	0.25**	0.15	0.10	0.68**	0.70**	0.47**	0.60**	0.57**
Insomnia	0.14	0.16	0.23**	0.16	0.13	0.52**	0.51**	0.47**	0.39**	0.46**
Poor appetite	0.19*	0.27**	0.13	0.09	0.03	0.43**	0.39**	0.25**	0.34**	0.40**
Fatigue	0.16	0.26**	0.31**	0.08	0.10	0.62**	0.65**	0.55**	0.53**	0.50**
Decreased concentration	0.21*	0.31**	0.32**	0.12	0.16	0.62**	0.67**	0.53**	0.56**	0.57**
Poor memory	0.20*	0.34**	0.33**	0.13	0.10	0.59**	0.61**	0.46**	0.53**	0.52**
Guilty feelings	0.25**	0.35**	0.32**	0.10	0.24**	0.69**	0.71**	0.49**	0.56**	0.65**

Note: $^{**}p < 0.01$; $^{*}p < 0.05$.

Abbreviations: DS, depression subscale of the Depression and Somatic Symptoms Scale (DSSS); HADS-A, anxiety subscale of the HADS; HADS-D, depression subscale of the Hospital Anxiety and Depression Scale (HADS); HAMD, Hamilton Depression Rating Scale; SS, somatic subscale of the DSSS.

[a]"(B)" and "(10Y)" represent data collected at baseline and at the 10-year follow-up point, respectively.

3.4 | Associations of scale or subscale scores at baseline with the HAMD score at the three follow-up points

Table 4 shows the results of the first regression model. The $HAMD_{(B)}$ score appeared to be most strongly associated with the $HAMD_{(6M)}$ and $HAMD_{(2Y)}$ scores among the three depressive scale scores at baseline after controlling for demographic variables and pharmacotherapy. The $DS_{(B)}$ score was most strongly associated with the $HAMD_{(10Y)}$ score after controlling for other factors.

The second model compared the associations of the HAMD, DS, SS, HADS-D and HADS-A scores at baseline with outcomes of depression at the follow-up points (Table 5). The $SS_{(B)}$ score appeared to be an independent factor associated with the HAMD score at the three follow-up points after controlling for demographic variables and pharmacotherapy. Moreover, the $SS_{(B)}$ score was most strongly associated with (highest R square change) the $HAMD_{(6M)}$ and $HAMD_{(10Y)}$ scores. The $HAMD_{(B)}$ score was an independent factor most strongly associated with the $HAMD_{(2Y)}$ score.

In all the regression models, the values of the Durbin–Watson test ranged from 1.88 to 2.09, and the values of the VIF ranged from 1.01 to 1.29.

3.5 | Independent factors associated with outcomes of depression at follow-up

As shown in Table 6, the $SS_{(B)}$ score was still significantly associated with the HAMD scores at the follow-up points after controlling for the HAMD score at baseline, demographic variables, and pharmacotherapy.

4 | DISCUSSION

The first regression model (Table 4) demonstrated that the $HAMD_{(B)}$ score was most strongly associated with the $HAMD_{(6M)}$ and $HAMD_{(2Y)}$ scores among the three depressive scale scores at baseline after controlling for demographic variables and pharmacotherapy. The $DS_{(B)}$ score was most strongly associated with the $HAMD_{(10Y)}$ score. Moreover, the $DS_{(B)}$ score was significantly correlated with more items of the self-reported percentages of time spent experiencing depressive symptoms over the past 10 years as compared with the $HAMD_{(B)}$ score (Table 3). These results demonstrated that the $DS_{(B)}$ and $HAMD_{(B)}$ scores were associated with the long-term prognosis of depression.

In the second regression model, the $SS_{(B)}$ score appeared as a significant variable associated with the $HAMD_{(6M)}$, $HAMD_{(2Y)}$, and $HAMD_{(10Y)}$ scores and was most strongly associated with the $HAMD_{(6M)}$ and $HAMD_{(10Y)}$ scores after controlling for other factors. Moreover, the $SS_{(B)}$ score was significantly correlated with more items of the self-reported percentages of time spent experiencing depressive symptoms over the past 10 years as compared with the $HAMD_{(B)}$, $HADS\text{-}D_{(B)}$, and $HADS\text{-}A_{(B)}$ scores (Table 3). These results demonstrated that the $SS_{(B)}$ score might be more strongly associated with the long-term course of depression than the other four scale or subscale scores at baseline.

The $SS_{(B)}$ score was significantly correlated with the $HAMD_{(6M)}$, $HAMD_{(2Y)}$, and $HAMD_{(10Y)}$ scores in both the treatment and non-treatment groups (Table 2). However, the correlations of the $HAMD_{(B)}$ and $DS_{(B)}$ scores with the $HAMD_{(6M)}$ and $HAMD_{(2Y)}$ scores were not significant in the treatment group owing to confounding of pharmacotherapy. This demonstrated that the SS was a robust subscale associated with the severity of depression at the follow-up points, even under confounding of pharmacotherapy. This might result from two reasons. (1) Somatic symptoms, especially painful

TABLE 4 Independent factors associated with depressive severities in the first regression model at three follow-up points[a,b]

	Independent variable	B	R^2 change	95% B confidence interval	*p* value
$HAMD_{(6M)}$	$HAMD_{(B)}$	0.48	0.08	0.26 to 0.70	<0.001**
	Pharmacotherapy	−3.26	0.05	−5.08 to −1.45	<0.001**
	Gender	−2.2	0.02	−4.19 to −0.21	0.03*
$HAMD_{(2Y)}$	$HAMD_{(B)}$	0.43	0.07	0.21 to 0.65	<0.001**
	Education (years)	−0.59	0.02	−0.98 to −0.20	<0.01**
	Married	−2.0	0.02	−3.90 to −0.09	0.04*
$HAMD_{(10Y)}$	Pharmacotherapy	4.55	0.13	2.35 to 6.75	<0.001**
	$DS_{(B)}$	0.33	0.07	0.14 to 0.51	0.001**

Abbreviations: DS, depression subscale of the Depression and Somatic Symptoms Scale.; HAMD, Hamilton Depression Rating Scale.

[a]"(B)," "(6M)," "(2Y)," and "(10Y)" represent data collected at baseline and at the 6-month, 2-year, and 10-year follow-up points, respectively.

[b]*$p < 0.05$; **$p < 0.017$ after Bonferroni correction.

TABLE 5 Independent factors associated with depressive severities in the second regression model at three follow-up points[a,b]

	Independent variable	B	R^2 change	95% B confidence interval	*p* value
$HAMD_{(6M)}$	$SS_{(B)}$	0.29	0.11	0.14 to 0.44	<0.001**
	Pharmacotherapy	−3.37	0.05	−5.16 to −1.59	<0.001**
	$HAMD_{(B)}$	0.29	0.02	0.05 to 0.53	0.02*
$HAMD_{(2Y)}$	$HAMD_{(B)}$	0.29	0.07	0.05 to 0.54	<0.01**
	Education (years)	−0.59	0.02	−0.98 to −0.21	<0.01**
	$SS_{(B)}$	0.18	0.02	0.03 to 0.33	0.016**
	Married	−2.15	0.02	−4.04 to −0.26	0.03*
$HAMD_{(10Y)}$	$SS_{(B)}$	0.32	0.14	0.17 to 0.47	<0.001**
	Pharmacotherapy	4.39	0.09	2.23 to 6.54	<0.001**

Abbreviations: HAMD, Hamilton Depression Rating Scale; SS, somatic subscale of the Depression and Somatic Symptoms Scale.

[a]"(B)," "(6M)," "(2Y)," and "(10Y)" represent data collected at baseline and at the 6-month, 2-year, and 10-year follow-up points, respectively.

[b]*$p < 0.05$; **$p < 0.017$ after Bonferroni correction.

TABLE 6 Independent variables associated with outcomes of depression at follow-up[a,b,c]

Dependent variable	Independent variable	Estimate	Standard error	95% Wald CI	*p*
HAMD	$HAMD_{(B)}$ (one-point increment)	0.64	0.05	0.53 to 0.75	<0.001
	$SS_{(B)}$ (one-point increment)	0.13	0.03	0.07 to 0.19	<0.001
	Pharmacotherapy (yes vs. no)	−2.66	0.65	−3.93 to −1.39	<0.001
	Visit (one-visit increment)	−2.36	0.24	−2.83 to −1.89	<0.001
	Married (yes vs. no)	−1.47	0.42	−2.28 to −0.66	<0.001
	Educational $years_{(B)}$ (1-year increment)	−0.22	0.09	−0.39 to −0.05	0.01
	Age (1-year increment)	0.05	0.03	0.004 to 0.10	0.03

Abbreviations: HAMD, Hamilton Depression Rating Scale; SS, somatic subscale of the Depression and Somatic Symptoms Scale.

[a](B) Data collected at baseline.

[b]Generalized Estimating Equation models were used.

[c]Baseline, 6-month, 2-year, and 10-year follow-up points were considered as the first, second, third, and fourth visits.

physical symptoms, are more difficult to treat and have a higher probability of becoming residual symptoms (Hung et al., 2015; Jaracz et al., 2016). Moreover, MDD with painful physical symptoms is associated with a poorer prognosis of depression (Jaracz et al., 2016). (2) Patients with anxiety disorders or symptoms often suffer from several somatic symptoms, such as chest tightness, palpitation, dizziness, and muscle tension (Gelenberg, 2000; Graham et al., 2019). MDD with anxiety symptoms or comorbidities is associated with a poorer prognosis of depression (Gaspersz et al., 2017; Hung et al., 2020; van Bronswijk et al., 2018). The SS is composed of items assessing painful physical symptoms and anxiety-related somatic symptoms, which are related to a poor prognosis of depression (Gaspersz et al., 2017; Jaracz et al., 2016); therefore, the $SS_{(B)}$ might be more strongly associated with the severity of depression. Conversely, the HADS-$D_{(B)}$ and HADS-$A_{(B)}$ scores did not appear in the two regression models and had poorer correlations with the HAMD, DS, and HADS-D scores at the three follow-up points as compared with the $HAMD_{(B)}$, $DS_{(B)}$, and $SS_{(B)}$ scores. Among the five scales or subscales, the SS (100%) had the highest somatic component, followed by the HAMD (34.6%) and the DS (33.3%), while the HADS-D and HADS-A had no somatic component. Therefore, the associations of scales or subscales with the prognosis of depression might be partially associated with the percentage of the somatic component in the scale or subscale.

Three points are worthy of note. (1) The $DS_{(B)}$ and $SS_{(B)}$ scores were correlated with more items of the self-reported percentages of time spent experiencing depressive symptoms over the past 10 years as compared with the $HAMD_{(B)}$ score (Table 3). This might partially result from the fact that the design of the DS is compatible with the criteria of a MDE (Hung et al., 2006b). Moreover, in designing the SS, somatic symptoms, which are common and are associated with the prognosis of MDD, are prioritized (Hung et al., 2006b). (2) In the treatment group, the correlations of the $HAMD_{(B)}$ and $DS_{(B)}$ scores with the $HAMD_{(6M)}$ and $HAMD_{(2Y)}$ scores were not significant, but those with the $HAMD_{(10Y)}$ were significant (Table 2). This might partially result from the following reasons. At the 6-month and 2-year follow-up points, some of the subjects, who had a good response to pharmacotherapy, might still remain in the treatment group; therefore, the treatment group might be composed of subjects with a good and a poor response to pharmacotherapy. Under these conditions, pharmacotherapy had a significant confounding effect. At the 10-year follow-up point, only subjects who had a poor or limited response to pharmacotherapy and did not achieve remission remained in the treatment group. Our results showed that the HAMD score in subjects with pharmacotherapy was significantly higher than in those without pharmacotherapy (13.1 vs. 8.0) at the 10-year follow-up point. Therefore, the confounding effect of pharmacotherapy at the 10-year follow-up point might be limited. (3) The HADS-$D_{(B)}$ score had a poor association with the HAMD scores at the follow-up points. This might result from the HAMD being composed of 34.6% somatic symptoms; however, the HADS-D contains no somatic element. In fact, the correlation ($r = 0.40$) of the HADS-$D_{(B)}$ score and the $HAMD_{(B)}$ score was weaker than that ($r = 0.60$) of the $DS_{(B)}$ score and the $HAMD_{(B)}$ score. Moreover, residual symptoms of depression are commonly related to somatic symptoms, such as fatigue and insomnia (Hiranyatheb et al., 2016).

Several limitations or methodological issues should be addressed. (1) In the treatment group, the kinds and dosages of medication were not controlled. This might be one of reasons for which the correlations of the $HAMD_{(B)}$ and $DS_{(B)}$ scores with the $HAMD_{(6M)}$ and $HAMD_{(2Y)}$ scores were not significant. (2) Self-reported percentages of time spent experiencing depressive symptoms over the past 10 years were gross estimations and might have been affected by memory bias. The information presented in Table 3 shows that the recalled depressive symptoms over the past 10 years were well-correlated with the scores of the three scales at the 10-year follow-up point. This demonstrated that the recalled depressive symptoms might be affected by the severity of depression at the 10-year follow-up point. One study investigated the recall accuracy for specific symptoms of depression at 12- and 24-month follow-up points, and found that the recall accuracy for specific symptoms varied considerably, from >90% for dysphoria and anhedonia, to 55% for psychomotor and appetite/weight changes (Dunlop, Granros, et al., 2019; Dunlop, Parikh, et al., 2019). (3) Only 47.2% of the subjects attended the 10-year follow-up. Although there were no significant differences in the demographic variables, with the exception of age, between the subjects who did and did not attend follow-up, unknown bias might exist. (4) In this study, five scales or subscales were used in the statistical analysis. Multiple comparisons might increase the type I error. (5) Generalizability is an important issue (Vandenbroucke et al., 2007; von Elm et al., 2007). At baseline, the study excluded MDD subjects with psychotic symptoms, severe psychomotor retardation, substance abuse or dependence, and chronic medical diseases. Our results should be cautiously applied to patients with MDD due to these exclusion criteria.

5 | CONCLUSION

In a comparison of three depressive scale scores, the $HAMD_{(B)}$ and $DS_{(B)}$ scores were most strongly associated with the HAMD score at two (6-month and 2-year) and one (10-year) follow-up points, respectively. In a comparison of five scale or subscale scores, the $SS_{(B)}$ and $HAMD_{(B)}$ scores were most strongly associated with the HAMD score at 2 (6-month and 10-year) and one (2-year) follow-up points, respectively. The GEE models showed that the $SS_{(B)}$ score was still significantly associated with the HAMD score at the three follow-up points after controlling the $HAMD_{(B)}$ score, pharmacotherapy, and other demographic variables. The association of a scale at baseline with the prognosis of depression might be partially related to the percentage of the somatic component. Moreover, the $DS_{(B)}$ and $SS_{(B)}$ scores were significantly correlated with more items of the self-reported percentages of time spent experiencing depressive symptoms over the past 10 years as compared with the $HAMD_{(B)}$ score. Therefore, the $DS_{(B)}$ and $SS_{(B)}$ scores were associated with the long-term prognosis of depression.

ACKNOWLEDGMENTS

This study was supported in part by grants from Chang Gung Memorial Hospital Research Programs (CMRPG3G1861, CMRPG3H1781, and CLRPG3D0048) and Ministry of Science and Technology Research Programs (MOST 108-2314-B-182A-070). The funding source had no further role in study design; in the collection, analysis or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

CONFLICT OF INTEREST

All authors declare that they have no conflicts of interest.

AUTHOR CONTRIBUTIONS

Hung and Liu designed the study and wrote the protocol. Hung, Liu, Hsu, and Yang collected the data. Hung and Liu undertook the statistical analysis and wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

DATA AVAILABILITY STATEMENT

The datasets are available from the corresponding author on reasonable request.

ORCID

Ching-I Hung https://orcid.org/0000-0002-0466-5670

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Text Revision. American Psychiatric Association.

Annunziata, M. A., Muzzatti, B., Bidoli, E., Flaiban, C., Bomben, F., Piccinin, M., Gipponi, K. M., Mariutti, G., Busato, S., & Mella, S. (2020). Hospital Anxiety and Depression Scale (HADS) accuracy in cancer patients. *Supportive Care in Cancer*, *28*(8), 3921–3926. https://doi.org/10.1007/s00520-019-05244-8

Chou, Y. H., Lee, C. P., Liu, C. Y., & Hung, C. I. (2017). Construct validity of the Depression and Somatic Symptoms Scale: Evaluation by Mokken scale analysis. *Neuropsychiatric Disease and Treatment*, *13*, 205–211. https://doi.org/10.2147/NDT.S118825

de Almeida Macedo, E., Appenzeller, S., & Lavras Costallat, L. T. (2017). Assessment of the Hospital Anxiety and Depression Scale (HADS) performance for the diagnosis of anxiety in patients with systemic lupus erythematosus. *Rheumatology International*, *37*(12), 1999–2004. https://doi.org/10.1007/s00296-017-3819-x

Dunlop, B. W., Granros, M., Lechner, A., Mletzko-Crowe, T., Nemeroff, C. B., Mayberg, H. S., & Craighead, W. E. (2019). Recall accuracy for the symptoms of a major depressive episode among clinical trial participants. *Journal of Psychiatric Research*, *116*, 178–184. https://doi.org/10.1016/j.jpsychires.2019.03.008

Dunlop, B. W., Parikh, S. V., Rothschild, A. J., Thase, M. E., DeBattista, C., Conway, C. R., Forester, B. P., Mondimore, F. M., Shelton, R. C., Macaluso, M., Logan, J., Traxler, P., Li, J., Johnson, H., & Greden, J. F. (2019). Comparing sensitivity to change using the 6-item versus the 17-item Hamilton depression rating scale in the GUIDED randomized controlled trial. *BMC Psychiatry*, *19*(1), 420. https://doi.org/10.1186/s12888-019-2410-2

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders* (research version, patient edition). (SCID-I/P). Biometrics Research, New York State Psychiatric Institute.

Gaspersz, R., Lamers, F., Kent, J. M., Beekman, A. T., Smit, J. H., van Hemert, A. M., Schoevers, R. A., & Penninx, B. W. J. H. (2017). Anxious distress predicts subsequent treatment outcome and side effects in depressed patients starting antidepressant treatment. *Journal of Psychiatric Research*, *84*, 41–48.

Gelenberg, A. J. (2000). Psychiatric and somatic markers of anxiety: Identification and pharmacologic treatment. *Primary Care Companion to the Journal of Clinical Psychiatry*, *2*(2), 49–54. https://doi.org/10.4088/pcc.v02n0204

Graham, K., Searle, A., Van Hooff, M., Lawrence-Wood, E., & McFarlane, A. (2019). The value of physical symptoms in screening for post-traumatic stress disorder in the military. *Assessment*, *27*, 1139–1150, 1073191119864662. https://doi.org/10.1177/1073191119864662

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social & Clinical Psychology*, *6*(4), 278–296.

Hiranyatheb, T., Nakawiro, D., Wongpakaran, T., Wongpakaran, N., Bookkamana, P., Pinyopornpanish, M., Saisavoey, N., Wannarit, K., Satthapisit, S., & Tanchakvaranont, S. (2016). The impact of residual symptoms on relapse and quality of life among Thai depressive patients. *Neuropsychiatric Disease and Treatment*, *12*, 3175–3181. https://doi.org/10.2147/NDT.S124277

Hung, C. I., Liu, C. Y., Cheng, Y. T., & Wang, S. J. (2009). Migraine: A missing link between somatic symptoms and major depressive disorder. *Journal of Affective Disorders*, *117*(1–2), 108–115. https://doi.org/10.1016/j.jad.2008.12.015

Hung, C. I., Liu, C. Y., Wang, S. J., Juang, Y. Y., & Yang, C. H. (2010). Somatic symptoms: An important index in predicting the outcome of depression at six-month and two-year follow-up points among outpatients with major depressive disorder. *Journal of Affective Disorders*, *125*(1–3), 134–140. https://doi.org/10.1016/j.jad.2009.12.010

Hung, C. I., Liu, C. Y., Wang, S. J., Yao, Y. C., & Yang, C. H. (2012). The cut-off points of the Depression and Somatic Symptoms Scale and the Hospital Anxiety and Depression Scale in detecting non-full remission and a current major depressive episode. *International Journal of Psychiatry in Clincal Practice*, *16*(1), 33–40. https://doi.org/10.3109/13651501.2011.617456

Hung, C. I., Liu, C. Y., Yang, C. H., & Gan, S. T. (2020). Comorbidity with more anxiety disorders associated with a poorer prognosis persisting at the 10-year follow-up among patients with major depressive disorder. *Journal of Affective Disorders*, *260*, 97–104. https://doi.org/10.1016/j.jad.2019.08.085

Hung, C. I., Liu, C. Y., Yang, C. H., & Wang, S. J. (2015). The impacts of migraine among outpatients with major depressive disorder at a two-year follow-up. *PLoS One*, *10*(5), e0128087. https://doi.org/10.1371/journal.pone.0128087

Hung, C. I., Liu, C. Y., Yang, C. H., & Wang, S. J. (2019). Migraine with active headache was associated with other painful physical symptoms at two-year follow-up among patients with major depressive disorder. *PLoS One*, *14*(4), e0216108. https://doi.org/10.1371/journal.pone.0216108

Hung, C. I., Wang, S. J., & Liu, C. Y. (2009). Validation of the Depression and Somatic Symptoms Scale by comparison with the Short Form 36 scale among psychiatric outpatients with major depressive disorder. *Depression and Anxiety*, *26*(6), 583–591. https://doi.org/10.1002/da.20464

Hung, C. I., Weng, L. J., Su, Y. J., & Liu, C. Y. (2006a). Depression and Somatic Symptoms Scale: A new scale with both depression and somatic symptoms emphasized. *Psychiatry and Clinical Neurosciences*, *60*(6), 700–708. https://doi.org/10.1111/j.1440-1819.2006.01585.x

Hung, C. I., Weng, L. J., Su, Y. J., & Liu, C. Y. (2006b). Preliminary study of a scale measuring depression and somatic symptoms. *Psychological Reports*, *99*(2), 379–389. https://doi.org/10.2466/pr0.99.2.379-389

Jaracz, J., Gattner, K., Jaracz, K., & Gorna, K. (2016). Unexplained painful physical symptoms in patients with major depressive disorder: Prevalence, pathophysiology and management. *CNS Drugs*, *30*(4), 293–304. https://doi.org/10.1007/s40263-016-0328-5

Jeon, S. W., Yoon, S. Y., Ko, Y. H., Joe, S. H., Kim, Y. K., Han, C., Yoon, H. K., & Liu, C. Y. (2016). Do somatic symptoms predict the severity of depression? A validation study of the Korean Version of the Depression and Somatic Symptoms Scale. *Journal of Korean Medical Science*, *31*(12), 2002–2009. https://doi.org/10.3346/jkms.2016.31.12.2002

Kristman, V., Manno, M., & Côté, P. (2004). Loss to follow-up in cohort studies: How much is too much? *European Journal of Epidemiology*, *19*(8), 751–760. https://doi.org/10.1023/b:ejep.0000036568.02655.f8

Lang, T. A., & Altman, D. G. (2015). Basic statistical reporting for articles published in biomedical journals: The "statistical analyses and methods in the published literature" or the SAMPL guidelines. *International Journal of Nursing Studies*, *52*(1), 5–9. https://doi.org/10.1016/j.ijnurstu.2014.09.006

Liu, C. H., Fu, T. S., Lee, C. P., & Hung, C. I. (2019). Reliability and validity of the Depression and Somatic Symptoms Scale among patients with chronic low back pain. *Neuropsychiatric Disease and Treatment*, *15*, 241–246. https://doi.org/10.2147/ndt.s188277

Nikayin, S., Rabiee, A., Hashem, M. D., Huang, M., Bienvenu, O. J., Turnbull, A. E., & Needham, D. M. (2016). Anxiety symptoms in survivors of critical illness: A systematic review and meta-analysis. *General Hospital Psychiatry*, *43*, 23–29. https://doi.org/10.1016/j.genhosppsych.2016.08.005

Nixon, N., Guo, B., Garland, A., Kaylor-Hughes, C., Nixon, E., & Morriss, R. (2020). The bi-factor structure of the 17-item Hamilton Depression Rating Scale in persistent major depression; dimensional measurement of outcome. *PLoS One*, *15*(10), e0241370. https://doi.org/10.1371/journal.pone.0241370

Rosellini, A. J., Bourgeois, M. L., Correa, J., Tung, E. S., Goncharenko, S., & Brown, T. A. (2018). Anxious distress in depressed outpatients: Prevalence, comorbidity, and incremental validity. *Journal of Psychiatric Research*, *103*, 54–60.

Sawamura, J., Ishigooka, J., & Nishimura, K. (2018). Re-evaluation of the definition of remission on the 17-item Hamilton Depression Rating Scale based on recovery in health-related quality of life in an observational post-marketing study. *Health and Quality of Life Outcomes*, *16*(1), 14. https://doi.org/10.1186/s12955-018-0838-6

Seaman, S., & Copas, A. (2009). Doubly robust generalized estimating equations for longitudinal data. *Statistics in Medicine*, *28*(6), 937–955. https://doi.org/10.1002/sim.3520

Tse, P. S., Gonzalez, D. A., & Jenkins, S. R. (2018). Validating the structure of the Depression and Somatic Symptoms Scale. *Psychosomatics*, *59*(3), 277–282. https://doi.org/10.1016/j.psym.2017.11.002

van Bronswijk, S. C., Lemmens, L. H., Huibers, M. J., Arntz, A., & Peeters, F. P. (2018). The influence of comorbid anxiety on the effectiveness of cognitive therapy and interpersonal psychotherapy for major depressive disorder. *Journal of Affective Disorders*, *232*, 52–60.

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J., Egger, M., & STROBE Initiative (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *Epidemiology*, *18*(6), 805–835. https://doi.org/10.1097/EDE.0b013e3181577511

Vindbjerg, E., Makransky, G., Mortensen, E. L., & Carlsson, J. (2019). Cross-Cultural Psychometric Properties of the Hamilton Depression Rating Scale. *Canadian Journal of Psychiatry*, *64*(1), 39–46. https://doi.org/10.1177/0706743718772516

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., & STROBE Initiative (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Lancet*, *370*(9596), 1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X

Worboys, M. (2013). The Hamilton Rating Scale for Depression: The making of a "gold standard" and the unmaking of a chronic illness, 1960-1980. *Chronic Illness*, *9*(3), 202–219. https://doi.org/10.1177/1742395312467658

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370.

Zimmerman, M., Clark, H., McGonigal, P., Harris, L., Guzman Holst, C., & Martin, J. (2018). Relationship between the DSM-5 anxious distress specifier and the Hamilton Depression Rating Scale anxiety/somatization factor. *The Journal of Nervous and Mental Disease*, *206*(2), 152–154. https://doi.org/10.1097/nmd.0000000000000767

How to cite this article: Hung, C.-I., Liu, C.-Y., Hsu, S.-C., & Yang, C.-H. (2022). Comparing the associations of three psychometric scales at baseline with long-term prognosis of depression over a 10-year period. *International Journal of Methods in Psychiatric Research*, *31*(1), e1896. https://doi.org/10.1002/mpr.1896