

CRITERION-BASED VALIDITY OF THE DEPRESSION SCALE OF LATVIAN CLINICAL PERSONALITY INVENTORY

Viktorija Perepjolkina

Jeļena Koļesņikova

Kristīne Mārtinsone

Ainārs Stepens

Elmārs Rancāns

Rīga Stradiņš University, Latvia

Abstract. *The main aim of this study was to evaluate the criterion validity and to estimate the cut-off score of the Depression scale (DS) and short Depression scale (DSs) for a new self-report measure – Latvian Clinical Personality Inventory (LCPI). Usefulness of DS and DSs for identifying patients with major depression were analysed based on psychometric analysis of data acquired from psychiatric inpatient sample with depressive disorder (n = 37) in comparison to randomised stratified community subsample (n = 176) selected from the overall test development sample (N = 888). The present study was carried within the framework of the National Research Program (BIOMEDICINE) 2014 – 2017 (sub-project Nr.5.8.2.). It was shown that all 24 item of DS show good to excellent discrimination power. Cronbach's alpha was 0.97 for DS and 0.95 for DSs in test development sample. For DS, the optimal cut-off score was 26 points (sensitivity 95%, specificity 91%, and positive predicted value of 69%). For DSs, the optimal cut-off was 12 points (sensitivity 92%, specificity 89%, and positive predicted value 63%). DS and DSs of LCPI is proved to have good criterion validity in detecting depression and to be a reliable and valid instrument for assessment of depression symptoms in patients with depression and in general population. Subjects scoring at least 26 on DS or 12 points on DSs constitute a target group for further diagnostic assessment in order to determine appropriate treatment.*

Keywords: *depression, criterion validity, reliability, screening, sensitivity, specificity.*

Introduction

Depression, especially if left untreated, has considerable impact on individuals' quality of life, on society and on the public health system (Mathers & Loncar, 2006) and is associated with serious consequences such as personal and

interpersonal suffering, impaired daily functioning, disturbance to interpersonal relationships, increased health care use, morbidity and an increased risk of suicide (e.g. Rihmer, 2007; Zlotnick et al., 2000). Depression is among the most commonly diagnosed mental disorders in adults. According to Latvian National Health Service, in 2015 depressive disorder was on the fifth place among all mental and behavioural disorders based on its relative frequency (Pulmanis, Japeniņa, Taube, 2016). Point prevalence of depression in the Latvian general population has been estimated to be 6.7% (Rancans, Vrublevska, Snikere, Koroleva, & Trapencieris, 2014), but 12-month prevalence of major depression has been estimated to be 7.9%, and for minor depression it was 7.7% (Vrublevska et al., 2017). Worldwide the prevalence of depressive disorders in primary care has been estimated to be between 10–20% (Mitchell, Vaze, & Rao, 2009), and in Latvia it was shown to be 10.2% in average with marked differences by gender (e.g. for male it was 6.0%, while for female it has been estimated to be 12.0%), and in lower extent with differences by age group (Rancans, Vrublevska, Kivite, et al., et al., 2016).

Because depression is often undiagnosed and untreated, active screening of depression is warranted. Use of patient-administered screening tools has increased as a quick and reliable option in the first step of depression assessment or as a treatment monitor (Gilbody, Richards, Brealey, & Hewitt, 2007). In English language, many self-assessment instruments are available for practitioners for identifying patients with major depression or dysthymia in primary care settings, but in Latvian language only some of these instruments are available, and on the moment only one of them – Patient Health Questionnaire (PHQ-9) is standardised in Latvia (see Rancans, Vrublevska, Trapencieris et al., 2016).

Screening of depression is essential not only in primary care settings, but also in occupational, educational, forensic and clinical settings. In these settings screening of general depression can be done as a part of general psychological assessment. Recently a comprehensive multi-item multi-scale self-report measure – Latvian Clinical Personality Inventory (LCPI; Perepjolkina, Koļešņikova, Mārtinsonē, & Stepens, in press) had been developed in Latvia. The Depression Scale (DS, $k = 24$) is one of nine clinical scales¹ of LCPI, along with

¹ Other clinical scales of LCPI are: Symptoms of Posttraumatic Stress Disorder (PTSD), General Anxiety Symptoms (GA), Panic Attack Symptoms (PA), Symptoms of Social Anxiety (SA), Alcohol Related Problems (ALCO), Drug Related Problems (DRUG), Somatic Symptoms (SM) and Psychotic Symptoms (PSS).

33 personality-trait scales², five functioning scales³ and five additional scales⁴. LCPI is available in a full ($k = 322$) and in a short version which is made up of first 220 items of LCPI, so also short version of DS scale (DSs) is available. DS/DSs scale for LCPI is developed for the assessment and monitoring of the severity of depression symptoms as a part of general psychological assessment in clinical, forensic and occupational settings.

Development of DS and DSs scale for LCPI

The DS scale is constructed based on Diagnostic and Statistical Manual of Mental Disorders 5th ed. (DSM–5; American Psychiatric Association [APA], 2013) diagnostic criteria for Major Depressive Disorder (GDD), employing criterion related strategy for scale construction. In the initial test development stage at least two items were formulated for every criterion symptom listed DSM-5 for major depressive disorder (APA, 2013, p. 160-161). In total, 65 items were formulated for DS scale on this stage. After discussions in a test development work-group, employing consensus approach 58 items (k_1 , Table 1) were retained for further content validity evaluation made by five external experts – three experienced practitioner psychiatrist and two clinical psychologists. As a result, 50 survived items (k_2 , Table 1) for the DS scale were included in the second preliminary version of LCPI (Perepjolkina, Kolesnikova, Mārtinsons, Stephens, 2016) for further empirical testing (see Table 1).

² LCPI personality trait model includes seven broad domains of personality trait variation – Narcissism (NR), Impulsivity (IM), Negative Affectivity (NA), Dependence (DE), Introversion (IN), Psychoticism (PSY) and Compulsivity (C) - comprising 33 specific personality trait facets (NR1: Dominance, NR2: Conceit, NR3: Attention seeking, NR4: Manipulativeness, NR5: Harshness; IM1: Aggression, IM2: Irresponsibility, IM3: Rashness, IM4: Risk taking; NA1: Intemperance, NA2: Emotional Stability, NA3: Emotional lability, NA4: Depressivity, NA5: Anxiousness, NA6: Impersistence (Distractibility), NA7: Distrustfulness; DE1: Evaluation apprehension, DE2: Submissiveness, DE3: Indecisiveness, DE4: Separation insecurity; PSY1: Cognitive dysregulation, PSY2: Dissociation proneness; PSY3: Eccentricity, PSY4: Suspiciousness, PSY5: Unusual beliefs, PSY6: Self-harm; IN1: Restricted affectivity, IN2: Social Withdrawal, IN3: Close relationship avoidance; C1: Pedantry, C2: Perseveration and C3: Perfectionism.

³ Five functioning scales of LCPI are: Sleep problems (F1), Energy level (F2), Psychomotor Retardation and Diminished Cognitive Function (F3), Problems associated with Attention Functions (F4), and Communication Problems (F5).

⁴ Additional scales of LCPI are: Suicide Ideation (SI), Self-esteem (SE), Stress Symptoms (STRS), Perceived Social Support (PSS) and Unstable and Intense Interpersonal Relationships (REL).

Table 1 Amount of Items for Every Criterion Symptom of Major Depressive Disorder Developed and Retained for Preliminary and Final Version of DS Scale for LCPI

Symptoms/diagnostic criteria of Major depressive disorder (based in DSM-5)	Amount of items			^a Item No. in LCPI
	k_1	k_2	k_3	
Depressed mood	10	8	5	145, 113, 126, 265, 289
Marked diminished interest or pleasure in all, or almost all, activities	7	7	4	74, 82, 115, 238
Significant weight loss when not dieting or weight gain	3	2	--	--
Insomnia or hypersomnia	8	7	2	119, 50
Psychomotor agitation or retardation	4	4	2	122, 62
Fatigue or loss of energy	6	5	4	134, 79, 1, 188
Feelings of worthlessness or excessive or inappropriate guilt	8	8	2	167, 290
Diminished ability to think or concentrate	6	5	4	133, 281, 288, 307
Recurrent thoughts of death or suicidal ideation	6	4	1	68

Note. k_1 = amount of items retained after evaluation in the test development work-group; k_2 = amount of items included in the second preliminary item pool of LCPI; k_3 = amount of items included in the final version of DS scale for LCPI.

^aItems included in the short version of DS scale (DSs) are presented in bold face.

Based on psychometric analysis, performed on the next test development stage, 24 best performing items (k_3 in Table 1) were selected from the second preliminary item pool of LCPI for the inclusion to the final version of Depression scale (DS) for LCPI and 12 of these 24 items were selected for the short version of this scale (DSs) (see Table 1 and Table 3 for overview). Fifteen of 24 items are included not only in DS scale, but also in some other scales of LCPI, e.g. DS scale share two items with PTSD scale, one item with Stress symptoms scale and with Self-esteem scale, five items with F2 scale: Energy level; four items with F3 scale: Psychomotor retardation and diminished cognitive function, and two items with LCPI scale F1: Sleep problems (see Table 3).

The main objective of this study was to evaluate the criterion validity and to estimate cut-off score of the Depression scale (DS) and short Depression scale (DSs) for LCPI.

Method

Participants

In total valid protocols⁵ of 888 adults (38.7 % male), who participated in the test development and primary validation study, were included in sample A (test development sample). Respondents ranged in age from 18 to 82 years with mean age of 36.23 ($SD = 16.08$) years. A part of sample A were inpatients, who received treatment in the psychiatric clinics ($n = 153$, sample P⁶, with mean age of 40.98 ($SD = 16.41$) years, 43.7 % male), 36 of whom (sample P1, with mean age of 50.73 ($SD = 13.47$) years, 24.3 % male) were inpatients with F32.0 - F33.11 diagnosis (based on ICD-10; WHO, 1992). For validation purposes a randomised community sample ($n = 176$, sample C, mean age 48.98 ($SD = 16.42$) years, 43.8 % male) stratified by age and gender according to the proportions of these demographic characteristics in Latvian population were selected from sample A.

Instruments

All participants filled in demographic questionnaire indicating background information (age, sex, income level, marital status, level of education, occupational status and area of occupation), and completed a second preliminary item pool ($k = 664$) of Latvian Clinical Personality Inventory (Perepjolkina, Koļesnikova, Mārtinsone, Stepens, 2016) (items were answered on a 4-point response format: from 0 – ‘totally disagree’ to 3 – ‘totally agree’).

The external criterion measure was a diagnosis of major depressive disorder (F32.0-F32.2 single episode (mild, moderate or severe without psychotic features) or F33.0-F33.2 recurrent (mild to severe without psychotic features)) (ICD-10; World Health Organisation, 1992) according to the medical records.

Procedure

Participation in the study was on a voluntary basis. All participants were required to meet the following criteria: (a) to be 18 years of age or older, (b) to be able to consent and complete the study protocol in Latvian. For psychiatric inpatient sample additional criteria were: (c) to endorse or exhibit current psychiatric symptoms, and (d) attending physician’s admission for a patient to

⁵ The initial sample included 936 participants, but in total 48 protocols were eliminated as invalid (32 protocols (5.3 %) in community sample, and 16 protocols (9.5 %) in psychiatric inpatients sample). LCPI protocols were considered invalid in this study if more than 20 items (3 %) were left blank or if an answer on the last control-item (*I have honestly answered to all questions*) was ‘0’ or ‘1’.

⁶ In this sample diagnoses included Mood disorder with major depressive-like episode due to known physiological condition (F06.32, $n = 1$), Personality and behavioral disorders due to known physiological condition (F07, $n = 2$), Alcohol related disorders (F10, $n = 9$), Schizophrenia (F20, $n = 55$), Schizotypal disorder (F21, $n = 32$), Brief psychotic disorder (F23, $n = 5$), Schizoaffective disorders (F25, $n = 3$), Bipolar disorder (F31, $n = 3$), Major depressive disorder, single episode (F32, $n = 9$), Major depressive disorder, recurrent (F33, $n = 28$), Generalized anxiety disorder (F41.1, $n = 1$), Reaction to severe stress, and adjustment disorders (F43, $n = 4$), and Somatoform disorders (F45, $n = 2$).

participate in the study. All psychiatric inpatients signed Informed Consent Form before participating in the study.

Participants from psychiatric inpatient sample were contacted individually and after providing informed consent, participants were provided the packet of questionnaires to complete during their own time. From community sample data were collected mostly using an online platform employing snowball sampling method, but 27 % of participants (student subsample) were assessed frontally during their personality psychology study course and filed in paper-pencil versions of questionnaires for course credit. Data were collected from January 2016 to November 2016 and all procedures were approved by the Riga Stradiņš University Ethical board. This study formed part of a LCPI development and validation research carried within the framework of Latvian National Research Programme Biomedicine for Public Health (BIOMEDICINE) 2014 – 2017 (sub-project Nr.5.8.2.).

Data analysis

Data were entered at the item level into a database and was analysed using IBM SPSS Statistics 20. Discrimination power of DS scale for LCPI on item-level was evaluated using (1) mean square contingency coefficient or *phi* (ϕ) coefficient – a correlation between an item (in dichotomized scoring format, were answers: 0 – ‘totally disagree’ and ‘1’ – ‘partly disagree’ were coded as ‘0’ and answers: ‘2’ – ‘partly agree’ and ‘3’ – ‘totally agree’ – were coded as ‘1’) and criterion variable (‘0’ – a participant is from sample C, and ‘1’ – a participant is from sample P1); (2) Pearson product-moment correlation coefficient or Pearson’s *r* – a correlation between an item (using full range answer format) and criterion variable (‘0’ – sample C, ‘1’ – sample P1); (3) by analysis of item’s Discrimination index (D) (which was calculated using dichotomized scoring format), (4) by analysing a relative frequency of positive answers (in dichotomized scoring format) in C, P and P1 sample; (5) by calculating corrected item-total correlation (using full range scoring format) and (6) by using Student’s *t*-test for comparison of mean item scores (*M*) in C and P1 sample (using full range scoring format). Cronbach’s alpha coefficient was used to assess the internal consistency of the DS and DSs scale in different subsamples.

For effect size estimation Hedges' *g* (Hedges, 1981), which provides a measure of effect size weighted according to the relative size of each sample was used⁷. The test performance (i.e. usefulness for identifying patients with major depression) of the DS and DSs scale was evaluated using receiver-operating characteristics (ROC) analyses. Areas under curve (AUC) indicated the accuracy

⁷ Hedges' *g* is an alternative measure of effect size, and in comparison to more traditionally used Cohen's *d* (which is the appropriate effect size measure if two groups have similar standard deviations and are of similar size). Hedges' *g* is more appropriate effect size measure if there are different sample sizes (Grissom & Kim, 2005; Stangroom, 2017) as it is in the case of our study.

of DS and DSs total raw scores to differentiate clinically diagnosed depression (sample P1) from potentially non-depression state (sample C), and were compared between subgroups using asymptotic tests of significance. Sensitivity, specificity, positive and negative predictive values (PPV and NPV) were calculated for different cut-off scores and the optimal cut-offs determined. For calculating positive predictive power (PPP) and negative predictive power (NPP) for various base rate indices (i.e. corrected for the prevalence or base rate of the condition in the population being tested) formulas provided by Streiner (2003, p. 213) were used.

Results

Item-level analysis

Detailed item-level analysis revealed that all 24 item of final version of DS scale for LCPI show good to excellent discrimination power (see Table 2). For example, in dichotomized scoring format, frequency of positive answers in community sample (sample C) ranged from 4.00 % to 25.60 % (11.66 % in average) in comparison to sample P1 (patients with F32 - F33 diagnosis), where the frequency of positive answers ranged from 56.80 % to 94.60 % (74.13 % in average) and to sample P (patients who received treatment in the psychiatric clinics) where the frequency of positive answers ranged from 40.30 % to 66.90 % (50.98 % in average). In total sample discrimination index ranged from 0.51 to 0.84 (0.65 in average), corrected item-total correlation indices ranged from 0.51 to 0.87 (0.74 in average). *Phi* coefficient (ϕ) – a correlation between an item (in dichotomized scoring format) and criterion ('0' – sample C. '1' – sample P1) ranged from 0.32 to 0.71 (0.58 in average), but using full range answer format, correlation between items and the same criterion ranged from 0.39 to 0.68 (0.58 in average) (Table 2).

Performed *t*-test analysis revealed that group differences (sample C vs. sample P1) on item-level (based on full-range response scale) were rather large (mean difference were 1.5 point in average) and highly significant ($p \leq .001$) (Table 2).

Scale-level analysis: Reliability and descriptive statistics

The Cronbach's alpha for internal consistency of the DS and DSs scale for LCPI in test development and validation sample (sample A) was 0.97 and 0.95 accordingly, and ranged from 0.91 to 0.97 (for DS scale) and from 0.85 to 0.95 (for DSs scale) for different subgroups (Table 2).

Descriptive statistics for full and short version of Depression scale for LCPI for different subgroups are presented in Table 2. As it could be seen, in total and in psychiatric inpatient sample (A and P sample) almost full possible range of DS scale and full range of DSs scale total score is obtained. In randomised community

sample (sample C) no one scored higher than 53 for DS scale and higher than 26 for DSs scale, but in sample P1 (patients with GDD) the lowest total score was 18 for DS scale and 4 for DSs scale. Therefore, DS scale's scores for the sample C ranged from 0 to 53, with a mean score of 10.78 ($SD = 11.17$), in contrast, DS scores for sample P ranged from 0 to 71 with a mean score of 34.26 ($SD = 18.66$) and DS scores for sample P1 ranged from 18 to 70 with a mean score of 46.49 ($SD = 13.84$) (Table 2).

Table 2 Reliability Coefficients (Cronbach's α) and Descriptive Statistics of DS and DSs Scale for LCPI

Statistics	α		$M (SD)$		Range	
	DS	DSs	DS	DSs	DS	DSs
A sample	.97	.95	20.50 (19.32)	9.75 (10.03)	0-71	0-36
C sample	.93	.87	10.78 (11.17)	4.76 (5.48)	0-53	0-26
C male (C_m)	.93	.85	7.92 (9.53)	3.27 (4.31)	0-47	0-18
C female (C_f)	.93	.87	13.00 (11.87)	5.91 (6.01)	0-53	0-26
P sample	.95	.92	34.26 (18.66)	17.05 (9.85)	0-71	0-36
P1 sample	.91	.86	46.49 (13.84)	21.97 (7.53)	18-70	4-35

Note. DS – Depression Scales for LCPI. DSs – Short Depression Scale for LCPI.

α – Cronbach's alpha. Sample A ($N = 888$) = test development total sample; sample C ($n = 176$) = community sample; sample C_m ($n = 77$) = male subsample from sample C; sample C_f ($n = 99$) = female subsample from sample C; sample P ($n = 153$) = clinical sample (patients who received treatment in the psychiatric clinics), sample P1 ($n = 37$) = patients with F32 - F33 diagnosis (based on ICD-10). Possible range for DS scale total score is from 0 to 72 and for DSs scale – from 0 to 36.

An independent samples t-test revealed that these differences were very large and statistically significant ((1) sample C vs. sample P: $t(241.12) = -13.59$, $p \leq 0.001$, $g = 1.55$ (very large effect size) and (2) sample C vs. sample P1: $t(46.36) = -14.72$, $p \leq 0.001$, $g = 3.06$ – huge effect size⁸). Very large and statistically significant differences were obtained also for DSs scale scores: sample C: $M = 4.76$ ($SD = 5.48$) vs. sample P: $M = 17.05$ ($SD = 9.85$), $t(230.18) = -13.70$, $p \leq 0.001$, $g = 1.57$ (very large effect size), and sample C: $M = 4.76$ ($SD = 5.48$) vs. sample P1: $M = 21.97$ ($SD = 7.53$), $t(44.32) = -17.13$, $p \leq 0.001$, $g = 2.92$ (huge effect size). Additional analysis revealed, that females from sample C scored significantly higher than males both for full and short version of Depression Scale for LCPI ($M_{female} = 13.00$ (11.88), $M_{male} = 7.92$ (9.53), $t(174) = 3.06$, $p < .001$, $g = 0.46$ (small effect size) for DS scale, and $M_{female} = 5.91$ (6.01), $M_{male} = 3.27$ (4.31), $t(174) = 3.25$, $p < .001$, $g = 0.49$ (small effect size) for DSs scale), but these differences were relatively small in magnitude. Such gender related differences is in line with empirically approved higher prevalence rate of GDD in females in general (e.g. Ayuso-Mateos et al.,

⁸ According to Sawilowsky (2009).

2001; Marcus et al., 2005) and in Latvian population as well (Pulmanis, Japeniņa, Taube, 2016; Rancans, Vrublevska, Kivite, et al., 2016).

Scale-level analysis: Criterion validity

In Table 4 diagnostic efficiency statistics on scale level are reported using multiple cut-off scores. Overall, the area under the curve was 0.97 for DS scale and 0.96 for DSs scale. Performed analysis revealed that DS score greater than 26 points and DSs score greater than 12 points reasonably balanced sensitivity and specificity rates. For DS scale sensitivity was 0.95 and specificity was 0.91 at cut-off of raw score = 26 points and for DSs scale sensitivity was 0.92 and specificity was 0.89 at the cut-off of raw score = 12 points.

Unlike sensitivity and specificity, positive predictive power (PPP or PPV) and negative predictive power (NPP or NPV) are not fixed characteristics of a scale and are dependent on the prevalence (i.e., base rate) of the condition being assessed (Meehl & Rosen, 1955; Streiner, 2003). Thus, in Table 4 PPP and NPP are presented for multiple base rate estimates. Base rate estimates of 5%, 10%, 15%, 20% and 25% are considered. Given that, the prevalence of general depression is likely to vary across settings and patient populations, reporting PPP and NPP allows examiners to adjust risk tolerance accordingly. Table 4 displays diagnostic efficiency statistics for multiple base rate estimates.

Discussion

Performed analysis revealed that both full and short version of Depression Scale (DS and DSs) for LCPI show good to excellent diagnostic efficiency both on item and on scale level. Determined optimal cut-off score for DS scale correctly classified 95 % of patients with GDD (F32 - F33 diagnosis) and for DSs scale proportion of correctly classified patients was 92 %. That means, that diagnostic efficiency for DSs is somewhat lower in comparison to DS, but still it is reasonably high even in comparison to other popular self-report measures of GDD. For example, the 2002 literature review (Williams, Pignone, Ramirez, & Perez Stellato, 2002) found that median sensitivity across 16 instruments, including the BDI, CES-D, MDRS, PHQ and MDI⁹, for major depression was 85%, ranging from 50-97%, while median specificity was 74%, ranging from 51-98%.

⁹ Beck Depression Inventory (BDI), Centre for Epidemiologic Studies Depression Scale (CES-D), Hamilton Depression Rating Scale (HDRS), Patient Health Questionnaire (PHQ), Major Depression Inventory (MDI).

Table 3 Item-Level Analysis of Discriminating Power for the LCPI Depression Scale (DS and DSs)

Sample	C	P1	P	C vs. P1	C vs. P1	A		C	P1	<i>t</i>
	<i>f</i> ₁ %	<i>f</i> ₂ %	<i>f</i> ₃ %	ϕ	<i>r</i>	<i>D</i>	CITC	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	
145. Depressed mood (DSs)	7.4	78.4	55.2	.68	.68	.72	.87	0.38 (0.67)	2.08 (0.86)	-13.31***
74. Inability to feel joy (DSs, PTSD, STRS)	11.4	83.8	55.8	.64	.68	.72	.85	0.40 (0.73)	2.22 (0.85)	-13.31***
126. Feelings of inner emptiness (DSs)	6.3	67.6	48.7	.62	.63	.70	.83	0.27 (0.61)	1.78 (1.08)	-8.31***
82. Indifference for everything (DSs)	6.3	64.9	40.3	.60	.66	.58	.82	0.24 (0.58)	1.81 (1.02)	-8.46***
115. Loss of interest for living (DSs)	5.7	64.9	47.4	.61	.64	.61	.83	0.20 (0.57)	1.68 (1.03)	-8.24***
289. Feelings that 'life is empty'	6.8	64.9	40.9	.59	.62	.59	.82	0.26 (0.67)	1.81 (1.08)	-11.40***
238. Loss of interest (in general) (PTSD)	5.7	67.6	43.5	.63	.60	.58	.80	0.27 (0.62)	1.70 (1.13)	-7.52***
113. Negative view on future	8.0	64.9	42.2	.56	.56	.54	.75	0.38 (0.76)	1.84 (1.04)	-8.07***
68. Do not want to live (DSs)	6.8	56.8	45.5	.52	.48	.57	.55	0.27 (0.67)	1.29 (1.21)	-9.25***
167. Feelings of worthlessness (SE) (DSs)	9.1	75.7	49.4	.62	.64	.67	.81	0.36 (0.64)	1.95 (1.00)	-12.27***
290. Excessive guilt	13.6	73.0	41.8	.53	.49	.51	.51	0.52 (0.76)	1.70 (0.97)	-12.33***
265. Feelings of loneliness	12.5	67.6	55.2	.50	.55	.72	.76	0.44 (0.77)	1.89 (1.10)	-9.24***
134. Loss of energy (DSs, F3)	16.5	94.6	66.9	.65	.65	.83	.81	0.63 (0.78)	2.35 (0.75)	-12.18***
79. Fatigue ('no reason why') (F3)	16.5	94.6	65.6	.65	.65	.82	.79	0.65 (0.85)	2.49 (0.69)	-9.02***
1. Fatigue (feel tired sooner than usually) (F3)	21.6	91.9	62.3	.56	.58	.84	.78	0.72 (0.93)	2.41 (0.72)	-8.47***
188. Fatigue & loss of energy (F3)	13.1	78.4	56.5	.58	.58	.78	.78	0.54 (0.76)	2.08 (1.04)	-7.66***
133. Diminished ability to concentrate (DSs)	10.2	78.4	54.5	.62	.60	.65	.72	0.41 (0.70)	1.86 (0.92)	-10.82***
62. 'Slow mode of thinking' (DSs, F2)	13.6	78.4	55.8	.57	.60	.68	.75	0.52 (0.77)	2.08 (0.92)	-10.78***
122. Psychomotor retardation (DSs, F2)	11.9	75.7	48.1	.57	.59	.59	.73	0.43 (0.78)	2.00 (1.03)	-10.53***
281. Absent-mindedness (F2)	14.2	73.0	51.9	.52	.51	.56	.62	0.61 (0.78)	1.89 (0.94)	-8.70***
288. Diminished cognitive function (F2)	4.0	73.0	48.7	.71	.63	.53	.70	0.34 (0.59)	1.78 (0.98)	-8.67***
307. Problems with memory (F2)	11.4	86.5	57.1	.66	.62	.73	.78	0.55 (0.74)	2.11 (0.84)	-11.40***
119. Sleep problems (F1)	25.6	64.9	46.8	.32	.39	.58	.56	0.76 (0.92)	1.81 (1.18)	-5.15***
50. Insomnia (DSs, F1)	21.6	59.5	43.5	.32	.40	.58	.56	0.64 (0.88)	1.70 (1.15)	-5.29***

Note. *f*₁, *f*₂, *f*₃% = relative frequency of positive answers (in dichotomized scoring format); ϕ = phi coefficient – a correlation between an item (in dichotomized scoring format) and criterion ('0' – sample C, '1' – sample P1); *r* = a correlation between an item and criterion (full range answer format); *D* = discrimination index (dichotomized scoring format); CITC = corrected item-total correlation; *M* = reaction index. ****p* < .001. DSs = Short Depression scale for LCPI; F1 = LCPI scale Sleep problems; F2 = LCPI scale Energy level; F3 = LCPI scale Psychomotor retardation and diminished cognitive function; SE = LCPI scale Self-esteem; PTSD = LCPI scale Posttraumatic stress disorder; STRS = LCPI scale Stress symptoms.

Table 4 Diagnostic Efficiency Statistics of the Depression scale (DS) and Short Depression scale (DSs) for LCPI for Different Cut-offs Using the Samples P1 and C

Cut-off score	Sensitivity	Specificity	PPV	NPV	LR+	LR-	Base Rate Estimates									
							0.05		0.10		0.15		0.20		0.25	
							PPP	NPP	PPP	NPP	PPP	NPP	PPP	NPP	PPP	NPP
Raw score of Depression scale (DS) for LCPI																
DP ≥ 24	0.95	0.89	0.64	0.99	8.64	17.80	0.31	1.00	0.49	0.99	0.60	0.99	0.68	0.99	0.74	0.98
DP ≥ 26	0.95	0.91	0.69	0.99	10.56	18.20	0.36	1.00	0.54	0.99	0.65	0.99	0.73	0.99	0.78	0.98
DP ≥ 28	0.92	0.92	0.71	0.98	11.50	11.50	0.38	1.00	0.56	0.99	0.67	0.98	0.74	0.98	0.79	0.97
Raw score of Short Depression scale (DSs) for LCPI																
DPs ≥ 10	0.97	0.84	0.56	0.99	6.06	28.00	0.24	1.00	0.40	1.00	0.52	0.99	0.60	0.99	0.67	0.99
DPs ≥ 11	0.95	0.86	0.59	0.99	6.79	17.20	0.26	1.00	0.43	0.99	0.54	0.99	0.63	0.99	0.69	0.98
DPs ≥ 12	0.92	0.89	0.63	0.98	8.36	11.13	0.31	1.00	0.48	0.99	0.60	0.98	0.68	0.98	0.74	0.97
DPs ≥ 13	0.89	0.91	0.67	0.98	9.89	8.27	0.34	0.99	0.52	0.99	0.64	0.98	0.71	0.97	0.77	0.96
DPs ≥ 14	0.86	0.91	0.67	0.97	9.56	6.50	0.33	0.99	0.51	0.98	0.63	0.97	0.70	0.96	0.76	0.95

Notes. Area under ROC curve (AUC) for DS scale = 0.969 (0.948 – 0.990) and for DSs scale = 0.963 (0.945 – 0.991) (figures in parenthesis indicate 95% confidence limits, asymptotic tests, under non-parametric assumptions). PPV = positive predictive value (Streiner, 2003, Formula No. 5, p. 212); NPV = negative predictive value (Streiner, 2003, formula No. 6, p. 212); LR+ = Likelihood ratio for positive tests (Sensitivity/(1-Specificity)) (Streiner, 2003, formula No. 3, p. 211); LR- = Likelihood ratio for negative test (Specificity/(1-Sensitivity)) (Streiner, 2003, formula No. 4, p. 211); PPP = positive predictive power (value) corrected for the prevalence or base rate of the condition in the population being tested (Streiner, 2003, formula No. 15, p. 213); NPP = negative predictive power (value) corrected for the prevalence or base rate of the condition in the population being tested (Streiner, 2003, formula No. 17, p. 213). Sample P1 = patients with F32 - F33 diagnosis (based on ICD-10) (n = 37). Sample C = non-clinical randomised community sample (n = 176).

While for determined optimal cut-off sensitivity in this study of DS was 95% and specificity 91%, while for DSs sensitivity was 92% and specificity 89%. In comparison sensitivity of PHQ-9 Latvian language version was only 74.7% and specificity 84.0%, correctly classifying 83.2% patients (Rancans et al, 2016).

Sensitivity, specificity, and the LR are generally seen as fixed properties of the test. That is, as long as the test is used with similar groups of people, these attributes should not change. However, if the test is used with people who have different amounts of the trait in question, then sensitivity and specificity will have to be recalculated; for example, a test validated on inpatients with mild or severe depression will likely have different properties when used with outpatients with dysthymia. So in future, it would be necessary to test diagnostic efficiency of DS and DSs in outpatient sample with dysthymia, as well as to test relative test performance (both on scale and on item-level) among different gender, age, and physical chronic illness groups, and separately for inpatients with mild, moderate and severe depression. Future studies should test whether the DS/DSs is effective

instrument for screening of depression in different contexts (e.g. educational, occupational, forensic) and is it effective tool for monitoring changes in severity of symptoms of depression during the course of treatment.

The main limitation of DS and even of DSs scale as an instrument for screening of depression is that these are no ‘independent’ scale, but are made up of items spread among other items of LCPI, which is rather long ($k = 322$ – full version and $k = 220$ – short version), so it could be difficult for patients with depression and for seniors to retain concentration and motivation and to fill-in this inventory completely and in consistent manner. So, for targeted screening of depressive disorder special screening instruments (e.g. PHQ-9) would be more appropriate.

The main limitations of this study are (1) small sample size both for inpatient with GDD group and for community sample group; (2) semi-representative sample of community population randomly selected from the non-clinical subsample of test development sample, and (3) no additional criterion was used for the control of severity of depression symptoms in community sample. Given this, it remain possible that the subjects who were included as ‘controls’ might have been GDD-positive in reality. It could be possible, especially considering their proportion in sample C, which is 9.1% both of DS and for DSs (based on DS score ≥ 26 , and DSs score ≥ 12), which corresponds reasonably well to the prevalence rate of depressive disorder estimated in general population of Latvia and in primary care settings (Rancans et al., 2014; Rancans, Vrublevska, Kivite, et al., 2016; Vrublevska et al, 2017).

Next limitation of this study is that no additional criterion, accept diagnosis based on medical records, was used in inpatient sample. On the other hand, partially verification bias was avoided in the present study because diagnosis of major depression on all the subjects from psychiatric inpatient sample was made irrespective of the results of the DS screening test. Nonetheless, the positive predictive value was high (69% for DS and 63% for DSs).

In future validation studies of DS and DSs for LCPI it would be useful to analyse different characteristics of true versus false positives (based on cut-off scores of DS/DSs scale), for example, it was shown in other studies that false positives show higher rates of anxiety symptomatology, previous depressive episodes, somatic complains, alcohol and nicotine consumption (e.g. Haringsma, Engels, Beekman, & Spinhoven, 2004).

Conclusions

DS and DSs of LCPI is proved to have good criterion validity in detecting depression while using LCPI for the overall psychological assessment, and to be a reliable and valid instrument for assessment of depression symptoms in patients

with depression and in general population. Subjects scoring ≥ 26 points on DS or ≥ 12 on DSs constitute a target group for further diagnostic assessment in order to determine appropriate treatment.

References

- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5rd edition). Washington, DC.
- Ayuso-Mateos, J. L., Vázquez-Barquero, J. L., Dowrick, C., Lehtinen, V., Dalgard, O. S., Casey, P., & Wilkinson, G. (2001). Depressive disorders in Europe: prevalence figures from the ODIN study. *The British Journal of Psychiatry*, 179, 308-316.
- Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *Journal of General Internal Medicine*, 22(11), 1596-1602.
- Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Mahwah, NJ: Erlbaum.
- Haringsma, R., Engels, G. I., Beekman, A. F., & Spinhoven, P. (2004). The criterion validity of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of self-referred elders with depressive symptomatology. *International Journal of Geriatric Psychiatry*, 19(6), 558-563.
- Hedges, L. V. (1981). Distribution theory for Glass' estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107-128. doi:10.3102/10769986006002107.
- Marcus, S. M., Young, E. A., Kerber, K. B., Kornstein, S., Farabaugh, A. H., Mitchell, J., & Rush, A. J. (2005). Gender differences in depression: Findings from the STAR*D study. *Journal of Affective Disorders*, 87(2/3), 141-150. doi:10.1016/j.jad.2004.09.008
- Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *Plos Medicine*, 3(11), e442.
- Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, 52(3), 194-216.
- Mitchell, A. J., Vaze, A., & Rao, S. (2009). Clinical diagnosis of depression in primary care: A meta-analysis. *The Lancet*, 374(9690), 609-619. Retrieved from <https://search.proquest.com/docview/199048061?accountid=32994>
- Perepjolkina, V., Koļesņikova, J., Mārtinsons, K., Stepens, A. (in press). *Latvijas Klīniskais personības tests (LKPT). Tehniskā rokasgrāma*. Rīga: RSU
- Perepjolkina, V., Kolesnikova, J., Mārtinsons, K., & Stepens, A. (2016). Latvijas Klīniskā personības testa otrās sākotnējo apgalvojumu kopas izstrāde: integratīvās pieejas pielietošana [Development of the Second Preliminary Item Pool of the Latvian Clinical Personality Inventory: an Integrative Approach]. *Society, integration, education. Proceedings of the International Scientific Conference, Volume I, May 27th-28th*, pp. 469-483. Rēzeknes Tehnoloģiju akadēmija, 2016. <http://journals.ru.lv/index.php/SIE/article/view/1521/1673> (In Latvian)
- Pulmanis, T., Japeniņa, S., Taube, M. (2016). *Psihiskā veselība Latvijā 2015. gadā. Tematiskais ziņojums, 16. izdevums* [Mental Health in Latvia in 2015: A Thematic Report] (in Latvian). Rīga: Slimību profilakses un kontroles cēns. Accessed: https://www.spkc.gov.lv/upload/Psihiska_veseliba_faili/851438_tz_pvl_2015_final_vietnem.pdf

- Rancans, E., Vrublevska, J., Kivite, A., Ivanovs, R., Logins, R., & Berze, L. (2016). P.2.h.008 - Prevalence of depression in primary care settings in Latvia – the results of the National Research Program BIOMEDICINE. *European Neuropsychopharmacology*, 26, S478-S479. doi:10.1016/S0924-977X(16)31483-3
- Rancans, E., Vrublevska, J., Snikere, S., Koroleva, I., & Trapencieris, M. (2014). The point prevalence of depression and associated sociodemographic correlates in the general population of Latvia. *Journal of Affective Disorders*, 156, 104-110. doi:10.1016/j.jad.2013.11.022
- Rancans, E., Vrublevska, J., Trapencieris, M., Snikere, S., Ivanovs, R., Logins, R., & Berze, L. (2016). P.2.h.016 - Validity of patient health questionnaire (PHQ-9) in detecting depression in primary care settings in Latvia – the results of the National Research Project BIOMEDICINE. *European Neuropsychopharmacology*, 26, S481. doi:10.1016/S0924-977X(16)31487-0
- Rihmer, Z. (2007). Suicide risk in mood disorders. *Current Opinion In Psychiatry*, 20(1), 17-22.
- Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 597 – 599.
- Stangroom. J. (2017). Effect Size Calculator for T-Test. *Social Science Statistics*. <http://www.socscistatistics.com/effectsize/Default3.aspx> [Accessed: 6 February 2017].
- Streiner, D. L. (2003). Diagnosing tests: using and misusing diagnostic and screening tests. *Journal of Personality Assessment*, 81(3), 209-219.
- Vrublevska, J., Trapencieris, M., Snikere, S., Grinberga, D., Velika, B., Pudule, I., & Rancans, E. (2017). The 12-month prevalence of depression and health care utilization in the general population of Latvia. *Journal of Affective Disorders*, 210, 204-210. doi:10.1016/j.jad.2016.12.031
- Williams, J. J., Pignone, M., Ramirez, G., & Perez Stellato, C. (2002). Identifying depression in primary care: a literature synthesis of case-finding instruments. *General Hospital Psychiatry*, 24(4), 225-237.
- World Health Organisation. (1992). *International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10)*. Geneva: WHO.
- Zlotnick, C., Kohn, R., Keitner, G., & Della Grotta, S. A. (2000). The relationship between quality of interpersonal relationships and major depressive disorder: findings from the National Comorbidity Survey. *Journal of Affective Disorders*, 59(3), 205-215.