

Original Paper

Evaluating the Construct Validity and Sensitivity to Change of the Klenico Depression Domain in Psychotherapeutic Inpatient Care: Instrument Validation Study

Stefan Reutimann¹, MSc; Jasmin Steiner¹, MSc; Noah Hübscher^{1,2}, MSc; Ulrich Voderholzer^{3,4}, Prof Dr Med; Adrian Meule⁵, PhD; Mareike Augsburg¹, PhD

¹University of Zurich Start-Up, Klenico Health AG, Zürich, Switzerland

²Department of Health Sciences and Technology, Swiss Federal Institute of Technology, Zurich, Switzerland

³Department of Psychiatry and Psychotherapy, Ludwig Maximilian University of Munich, Munich, Germany

⁴Schoen Clinic Roseneck, Prien am Chiemsee, Germany

⁵Department of Psychology, University of Regensburg, Regensburg, Germany

Corresponding Author:

Jasmin Steiner, MSc

University of Zurich Start-Up

Klenico Health AG

Förllibuckstrasse 190

Zürich, CH-8005

Switzerland

Phone: 41 44 533 13 90

Email: jasmin.steiner@klenico.com

Abstract

Background: The accurate diagnosis of mental disorders, such as depression, requires comprehensive, valid, and reliable tools to ensure evidence-based treatments and effective outcome monitoring. Existing diagnostic practices often lack standardization, leading to missed comorbidities and variable diagnostic accuracy. The Klenico system is an innovative, web-based diagnostic tool that integrates patient self-reports with clinical validations by mental health professionals. This system covers a broad spectrum of mental disorders, including depression.

Objective: This research aimed to evaluate the psychometric properties of the Klenico Depression Domain (KDD), the component of the Klenico system that measures depressive symptomatology, in a real-world clinical setting. Specifically, the evaluation focused on the assessment of its construct validity, internal consistency, and sensitivity to change in symptom severity.

Methods: Anonymized data from 496 inpatients with mental disorders collected between 2019 and 2022 were analyzed. Patients completed the KDD alongside parts of the Patient Health Questionnaire (PHQ), Beck Depression Inventory (BDI-II), and Satisfaction With Life Scale (SWLS) at both admission and discharge. Internal consistency was measured using Cronbach α . Exploratory factor analysis was conducted to examine the factor structure. Construct validity was assessed via Pearson correlations with PHQ-9 and BDI-II, while divergent validity was tested against the PHQ Somatic Symptoms Scale (PHQ-15), PHQ-Generalized Anxiety Disorder-7, and SWLS. Sensitivity to change was evaluated using paired 1-tailed t tests, effect sizes, and repeated measures correlations.

Results: The KDD demonstrated excellent internal consistency (Cronbach α =0.91 at admission and 0.93 at discharge). Factor analysis revealed a 7-factor structure encompassing dimensions like “inadequacy,” “anhedonia,” and “self-hatred,” aligning with core depressive symptoms outlined in the *International Statistical Classification of Diseases, Tenth Revision*. The correlations with the convergent questionnaires PHQ-9 ($r=0.68$; $P<.001$) and BDI-II ($r=0.70$; $P<.001$) were high. While the KDD showed a moderate correlation with the divergent PHQ-15 ($r=0.35$; $P<.001$), it was more strongly associated with the divergent SWLS ($r=-0.51$; $P<.001$) and Generalized Anxiety Disorder-7 ($r=0.51$; $P<.001$). Sensitivity to change was high, with significant reductions in KDD scores for patients with improved symptoms ($t_{27}=5.36$, $P<.001$; Cohen $d=0.79$) and high repeated measures correlation with both the BDI-II ($r=0.61$; $P<.001$) and the PHQ-9 ($r=0.59$; $P<.001$).

Conclusions: The KDD shows promise as a reliable and valid instrument for diagnosing depression and monitoring treatment outcomes in psychotherapeutic settings. Its alignment with *International Statistical Classification of Diseases, Tenth Revision*

diagnostic criteria and sensitivity to symptom change underlines its potential utility. These findings highlight the Klenico system's potential to enhance clinical diagnostics by addressing current gaps in mental health care, thus improving diagnostic accuracy and consistency. Further research is recommended to validate its performance across different populations and settings.

(*JMIR Form Res* 2025;9:e50504) doi: [10.2196/50504](https://doi.org/10.2196/50504)

KEYWORDS

depression diagnosis; Klenico system; Klenico Depression Domain; sensitivity to change; psychometric properties; construct validity; mental health assessment; clinical outcome monitoring

Introduction

Mental health disorders represent a broad spectrum of conditions with significant impacts on individuals' lives globally [1]. These disorders impair both personal and social functioning [2] and pose substantial challenges to public health systems [3]. For example, depression affects a significant portion of the European population, with a lifetime prevalence of 12% and an annual prevalence of about 7% [4,5], imposing considerable economic burdens of roughly €3000 (≈US \$3549) per patient per year [6]. This highlights the critical need for qualitative diagnosis and treatment strategies across the spectrum of mental disorders. However, the diagnosis of mental disorders remains challenging due to common comorbidities among different disorders [7,8] and their complex structures. For instance, depression is a highly heterogeneous disorder characterized by a wide variety of different presentations and a broad range of symptoms [9].

In current clinical practice, the quality of diagnostic procedures is often compromised by a lack of standardization and comprehensiveness. Health professionals predominantly tend to rely on unstructured, open interviews for diagnosis and classification [10-12], which compromises validity and reliability [13,14], thereby impeding evidence-based diagnoses [15]. Standardized tools like structured interviews are rarely used in real settings, as they are considered too time-consuming [10,11]. Alongside structured interviews, self-report questionnaires have emerged as an additional standardized method of assessment, seeing more frequent application in routine diagnostic procedures. Although these questionnaires are valid and time-saving tools when implemented for their intended fields of use, such as screening, severity assessment, and course measurement, they do generally not cover the whole spectrum of symptoms [16]. Focusing solely on one particular disorder domain, such as depression, using a disorder-specific questionnaire may lead to missing critical comorbidities. Moreover, self-report questionnaires lack consistency, as they measure a wide variety of different constructs and show inhomogeneity in the represented diagnostic criteria. For instance, the most common depression self-report questionnaires revealed only a 39% similarity in their covered symptoms [17]. Therefore, self-assessments and their cutoff scores alone should not be relied on for diagnosis [18]. It is necessary to obtain further clarifications on comorbidities and symptom duration through structured diagnostic interviews [19]. Alongside initial diagnostics, there is a significant lack of standardized measurement techniques for assessing outcomes [20-22]. However, this step is indispensable for the precise determination of treatment success [23].

The described lack of quality standardization and comprehensiveness in clinical practice causes mental disorders to often be unrecognized or misdiagnosed in different settings [15,24-28]. For instance, studies comparing routine diagnoses made by unstructured, unstandardized diagnostic procedures with standardized diagnoses made by semistructured interviews—the gold standard in diagnostics—among psychiatric inpatients and outpatients found that depression was correctly identified in only half of the cases. The agreement for most other disorders was also limited or poor [14,15,29]. Additionally, comorbidity is often missed in routine diagnostics [30]; for example, the frequent comorbidity between depression and anxiety was detected in fewer than half of the cases using unstandardized routine diagnostics [29].

These inadequate diagnostic practices, among other possible reasons such as poor access to therapy [31], can lead to ineffective therapeutic interventions. This situation may culminate in unwarranted patient distress and increased demands on health care resources [32]. This is reflected by the fact that only 60% of diagnosed depression patients are provided with evidence-based therapy according to guidelines in primary care in Germany [33], and only one-third of mental health treatments in Spain meet minimal adequacy criteria in both specialty and general medical care [34], with similar findings in other countries and for other mental disorders [35-37].

To enable efficient and adequate treatments, diagnostic procedures need to assess the entire symptomatology of mental disorders in a standardized way, allow conclusions about individual symptoms, and be able to comprehensively record potential comorbidities by querying all relevant disorders. In addition, it is important to adequately measure the course of symptoms to evaluate the success of treatments [23]. However, they also have to be practical, efficient, user-friendly, and feasible in clinical practice due to time constraints [38].

Digital technologies offer a promising solution for the described situation. These tools not only facilitate the diagnostic process but also offer the capability for self-administered assessments, thereby reducing the burden on healthcare professionals and enhancing the accessibility of diagnostic services by overcoming conventional barriers. Through the application of adaptive testing, a wide array of symptoms can be efficiently evaluated, optimizing both time and comprehensiveness. Furthermore, digital tools can empower professionals in standardized diagnostics by supporting their decision-making processes.

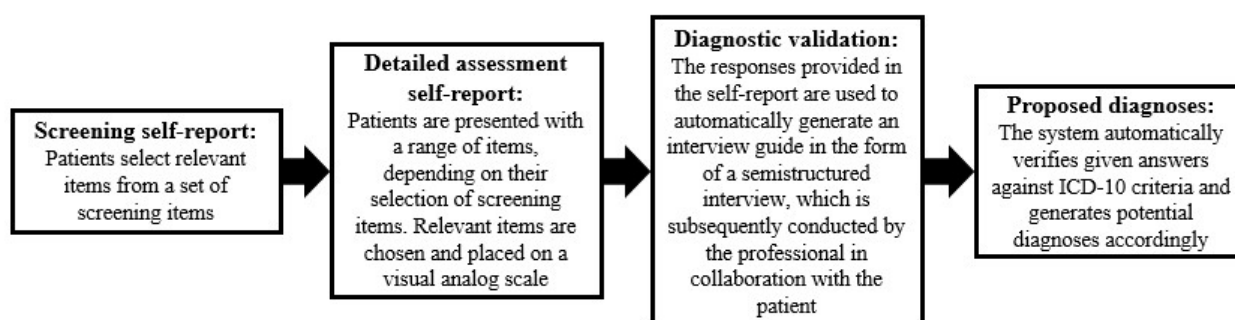
To this end, the Klenico system was developed as a comprehensive, web-based diagnostic tool that covers the following mental disorder domains: anxiety disorders, disorders

of adult personality and behavior, eating disorders, mental and behavioral disorders due to psychoactive substance use, depression, mania, obsessive-compulsive disorder, reaction to severe stress and adjustment disorders, and psychotic disorders. Furthermore, the system includes aspects of autism spectrum disorder, attention-deficit/hyperactivity disorder, and dementia. Its items are directly derived from the diagnostic criteria of the *International Statistical Classification of Diseases, Tenth Revision (ICD-10)* [39]. The Klenico Depression Domain (KDD) investigated in this research covers all *ICD-10* depressive disorders under F32 and F33, as well as dysthymia (F34.1), providing a comprehensive representation of depressive symptomatology.

More specifically, Klenico consists of a web-based, adaptive self-report module with a total of 496 items, including a preceding screening as well as subsequent clinical validation in the form of a diagnostic interview conducted by mental health professionals. The detailed process of the Klenico system is illustrated in Figure 1. Lustig [40] provides more details on the development of the Klenico system. For clinical validation, the

software creates an interactive interview guide based on responses to self-reports. This allows the professional to assess the symptomatology reported by the patient and reevaluate it externally through their expertise by asking questions and making observations. The combination of self-report and external assessment is crucial for standardized, multimodal diagnostics and serves as the ideal procedure. This integration facilitates efficient diagnoses of depression and other disorders that are easily implemented in everyday clinical practice and can relieve the burden on professionals while still maintaining a standardized and comprehensive diagnostic approach. Based on the results of the clinical validation, the system generates final automated suggestions for the most likely diagnoses, ensuring strict adherence to the diagnostic criteria outlined in the *ICD-10*. Moreover, the manifestations of the individual symptoms are clearly visually presented on a so-called symptom map, which enables targeted assessment at the symptom level and reveals important comorbidities. Furthermore, the assessments can be repeatedly performed throughout the course of treatment, whereby symptom change is made visible, thereby enabling treatment outcomes to be monitored.

Figure 1. Flowchart depicting the entire process of the Klenico system. ICD-10: International Statistical Classification of Diseases, Tenth Revision.



The validity of the Klenico self-report module was preliminarily confirmed on an outpatient sample in a previous study. The study showed that the diagnostic results from the Klenico self-report module aligned well with those obtained from the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders (DSM)*, *Fourth Edition*, thus demonstrating its criterion validity. Furthermore, preliminary indications of its convergent and divergent validity were also established through the comparison with common self-report questionnaires [41]. The purpose of this assessment was to confirm these findings and further validate the Klenico system by laying the focus on its depression domain (KDD). Research was focused on the KDD's reliability, construct validity, and sensitivity to change by considering the measured severity of the symptomatology. This is of substantial importance for ensuring the quality of the diagnostic tool and thus enabling valid, adequate, and standardized diagnostics of depression and mental disorders in general.

Methods

Procedure

The data used in the present research were collected between 2019 and 2022 in a German psychotherapeutic inpatient clinic, which specializes in treating patients from a range of diagnostic

groups. The clinic's standard diagnostic procedure was administered to all patients, which provided the foundation for the data used in this study. The Klenico system was added provisionally to these standard procedures as an extension to the usual care.

Upon admission, all inpatients received their routine diagnoses from the clinic based on an initial interview and received routine clinical self-report questionnaires (Beck Depression Inventory [BDI-II], Patient Health Questionnaire [PHQ]-9, PHQ-15, PHQ-Generalized Anxiety Disorder-7 [GAD-7], and Satisfaction With Life Scale [SWLS]) to further clarify their symptoms. Afterward, all patients were informed about the procedure and were asked to sign informed consent to allow their routine data and data from the Klenico assessment to be used for scientific studies and publications. Patients then received access to the Klenico self-report module.

Patients then completed the Klenico self-report module on a device of their choice (computer or tablet) and the clinics' routine self-report questionnaires within the first 2 weeks of their stay. The German version of the Klenico system was used. After completion, mental health professionals conducted the Klenico clinical validation module together with the patient in their next session. At discharge, the procedure was repeated while patients again were asked to answer the clinic's routine

self-questionnaires and the Klenico self-report module. Additionally, the Klenico clinical validation module was performed again. No diagnoses were made in the data collection based on the Klenico system.

Mental health professionals consisted of licensed psychotherapists, psychotherapists in training under supervision (with a minimum of 1 year of professional experience), and cotherapists. Professionals received an introductory session and a demonstration of the use of the Klenico system prior to their initial diagnostic validation. They were also provided with instructions on how to interpret and elucidate the results. The data were fully anonymized.

Measures

BDI-II Questionnaire

The BDI-II was part of the clinic's routine diagnostic procedure to evaluate depressive symptom severity and to serve as a measure of progress at discharge. The BDI-II [42] evaluates the severity of symptoms of depression in the past 2 weeks. It includes 21 items that are summed up to a maximum total score of 63. The BDI-II has good validity and reliability [43], including test-retest reliability [44], and is sensitive to symptom change [45]. The sum score was calculated. Cronbach α coefficient was 0.91 in this analysis.

PHQ-D Questionnaire

The PHQ-D [46] is a self-report instrument for recording mental disorders. The questionnaire demonstrates construct validity based on the number of doctor visits, psychosocial impairment, and incapacity to work, as well as excellent criterion validity (85% sensitivity and 70% specificity) when compared with Structured Clinical Interview for *DSM-IV*, Axis I diagnoses. Both professionals and patients show high acceptance of the PHQ-D, with a satisfaction rate greater than 90% [47]. The short form (option C according to the manual), which screens for depressive (PHQ-9), generalized anxiety (GAD-7), and somatoform (PHQ-15) symptoms, was used in the data collection as part of the clinic's routine diagnostic procedure at admission and discharge to evaluate symptom severity and to measure progress. The PHQ-9 has good test-retest reliability [48] as well as sensitivity to symptom change [49]. Sum scores of the PHQ-9 (maximum score=27), PHQ-15 (maximum score=30), and GAD-7 (maximum score=21) were calculated. Cronbach α coefficients for the PHQ-9, PHQ-15, and GAD-7 were 0.84, 0.78, and 0.83, respectively.

SWLS Questionnaire

The SWLS was part of the clinic's routine diagnostic procedure upon admission and is a 5-item questionnaire for measuring life satisfaction [50]. Items are summed up to a maximum score of 35, indicating extreme satisfaction. The SWLS has factorial validity as well as reasonable convergent and divergent validity indicated through correlations with other indicators of quality of life [51,52]. The sum score was calculated. Cronbach α coefficient was 0.86.

Klenico (Including the KDD)

The Klenico system is a web-based software designed to assist professionals in diagnosing mental disorders. It comprises 2 modules: a self-report followed by clinical validation. The general process of the Klenico system is depicted in Figure 1. Initially, the self-report module presents patients with 55 screening items, from which they select the relevant ones. Depending on the screening items selected, up to 496 items may then be adaptively presented to the patient for a more detailed assessment. The patient again selects the applicable ones and rates each previously selected item by dragging and dropping it on a visual analog scale (ranging from 0=low to 100=high) to assess symptom severity. Upon completion by the patient, the process transitions to the professional. The results from the self-report are graphically displayed on a symptom map, and an interview guide is generated for use in the subsequent clinical validation. This clinical validation involves a semistructured interview where the professional assesses the self-report from an external perspective. This includes posing questions based on the self-reported items, now reformulated for external inquiry, alongside additional external assessment items. These items are rated on a 4-point Likert scale. The system then compares the responses with *ICD-10* diagnostic criteria and provides the professional with suggested diagnoses.

In the data collection used for this assessment, the Klenico study version was used, in which all available 496 items are presented in the detailed assessment of the self-report, regardless of the previously selected screening items.

The KDD assessed in this assessment consists of a total of 31 items, with 29 items included in the self-report module. The clinical validation module reiterates these 29 items as external interview questions, with 2 additional observational items (on psychomotoric agitation and retardation) made by the professional. Means of the KDD for the clinical validation module were calculated where possible. For patients who solely answered the self-report module, means for the self-report were calculated consequently. Cronbach α coefficient was 0.91. Item names, means and SDs of the different KDD items, as well as an exemplary selection of the item formulations, are presented in Multimedia Appendix 1.

Sample

Fully anonymized data from 496 adult inpatients with mental disorders were used (337 female patients and 159 male patients). Data from discharge were available for 59 patients. Data from the Klenico clinical validation module were available for 445 patients at admission and 42 patients at discharge. The mean age was 38.9 (SD 15) years, and the mean duration of stay was 70.2 (SD 38.6) days. A total of 391 patients received a diagnosis of depression. The frequencies of the different depression diagnoses are reported in Table 1 (for a full overview of all given diagnoses, see Multimedia Appendices 2 and 3 [somatic diagnoses]).

Table 1. Overview of the given depression diagnosis with corresponding International Statistical Classification of Diseases, Tenth Revision (ICD-10) codes. The diagnoses were established through the initial interview as part of the routine diagnostic procedures followed at the clinic. Listed frequencies are relative to the number of total given depression diagnoses.

Depression diagnosis	ICD-10 code	Frequency (n=391), n (%)
Mild depressive episode	F32.0	4 (1)
Moderate depressive episode	F32.1	75 (19.2)
Severe depressive episode without psychotic symptoms	F32.2	41 (10.5)
Severe depressive episode with psychotic symptoms	F32.3	1 (0.3)
Recurrent depressive disorder, current episode mild	F33.0	7 (1.8)
Recurrent depressive disorder, current episode moderate	F33.1	135 (34.5)
Recurrent depressive disorder, current episode severe without psychotic symptoms	F33.2	115 (29.4)
Recurrent depressive disorder, current episode severe with psychotic symptoms	F33.3	2 (0.5)
Recurrent depressive disorder, currently in remission	F33.4	1 (0.3)
Dysthymia	F34.1	10 (2.6)

Statistical Analysis

Data analysis was performed using RStudio software (version 2021.9.1.372; Posit, PBC). *P* values <.05 were considered statistically significant.

Exploratory factor analyses (EFA) were conducted to explore the underlying factor structure of the KDD using the *lavaan*, *SEMtools*, and *psych* packages. Kaiser Meyer Olkin (KMO) and Bartlett tests were used to test the adequacy of the factor analysis using the functions *cortest.bartlett* and *KMO*. Parallel analysis and fit estimation were used to identify the optimal number of factors using the functions *efaUnrotate* and *fa.parallel*. A maximum likelihood estimator with robust standard errors was applied. Oblique rotation was applied by using the *obliqueRotate* function (the R code for the EFA can be found in [Multimedia Appendix 4](#)). The model fit was considered adequate when the comparative fit index >0.90, Tucker Lewis coefficient >0.90, standardized root mean square residual <0.10, and root mean square error of approximation <0.08 [53]. Items were assigned to a specific factor if they demonstrated a loading greater than 0.30 on that factor.

Convergent validity for the KDD was tested by performing Pearson correlation analysis with both the PHQ-9 and BDI-II, while divergent validity was assessed by comparison with the PHQ-15, GAD-7, and SWLS. An *r* value from >0.1 to <0.3 was considered to indicate a small effect, from ≥0.3 to <0.5 was considered to indicate a medium effect, and ≥0.5 was considered to indicate a large effect [54]. Bonferroni-adjustment was performed to correct for multiple testing.

Sensitivity to change of the KDD was examined by categorizing each patient into groups of either improved or unchanged symptoms based on changes in the PHQ-9 or BDI-II sum scores. KDD means for admission and discharge were compared for both improved and unchanged symptom groups by paired 2-tailed *t* tests. Effect sizes and standardized response means were calculated. Additionally, repeated measures correlations between the KDD and the BDI-II and between the KDD and the PHQ-9 were calculated using the *rmcorr* package [55].

Bonferroni adjustment was performed to correct for multiple testing.

A dropout analysis was conducted to compare patients who completed the Klenico system and the PHQ-9 and BDI-II questionnaires at discharge with those who did not complete these assessments. The analysis examined potential differences in demographic characteristics, questionnaire scores (KDD, PHQ-9, BDI-II, and SWLS), and length of stay using 2-tailed *t* tests or Wilcoxon rank sum tests. This approach aimed to identify any systematic biases between the groups.

Post hoc power analyses were conducted to assess the statistical power of the results (using the functions *pwr.r.test*, *pwr.t.test*, and *power.rmcorr*).

Ethical Considerations

Ethical approval was not required for the studies because according to German law, no ethical approval was required for the use of the Klenico system in routine diagnostic procedures. With reference to the Council Directive 93/42/EEC of the European Union and the according German legislation (Medizinproduktegesetz (MPG) § 23b), the test-related use of a medical device with certification (Conformité Européenne marking) does not require an ethics vote if it is carried out within the scope of its intended purpose [56]. Furthermore, the use of the Klenico system in routine diagnostic procedures falls within the scope of the responsibility of the attending physician (Berufsordnung der Ärztekammern §15). According to the ethics committee of the Ludwig Maximilian University Munich, such data collection with the aim of evaluating the performance of an accepted, evidence-based practice (quality assurance) does not require ethical approval. Finally, no ethical approval was required for analysis of the data, since retrospective studies using already available, anonymized data are exempt from needing ethics approval, according to the guidelines of the institutional review board at the Ludwig Maximilian University Munich. The study was conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Data were anonymized. No compensation was provided.

Results

Reliability

Excellent internal consistency was demonstrated by the KDD both at admission (Cronbach $\alpha=0.91$) and at discharge (Cronbach $\alpha=0.93$).

Exploratory Factor Analysis

A KMO value of 0.91 and a significant Bartlett test indicated eligibility for EFA. Parallel test and fit estimation favored a 7-factor solution ([Multimedia Appendices 5 and 6](#)). The 7-factor

model revealed good fit ($\chi^2_{269}=382.912$, comparative fit index=0.971, Tucker Lewis coefficient=0.950, standardized root mean square residual=0.025, and root mean square error of approximation=0.031). Based on the rotated loadings presented in [Table 2](#), the 7 factors were interpreted and labeled as “inadequacy,” “sleep disturbance,” “affect,” “anhedonia,” “somatization,” “appetite decrease,” and “self-hatred.” The item “feelings of worthlessness” was cross-loaded onto 2 factors. Four items (“psychomotor agitation,” “psychomotor retardation,” “inner tension,” and “indecision”) did not load adequately on any factor.

Table 2. Exploratory factor analysis (EFA) factor structure with rotated loadings of the 31 items of the Klenico Depression Domain (KDD). All reported factor loadings are statistically significant at $P<.05$, unless otherwise specified.

	Inadequacy	Sleep disturbance	Affect	Anhedonia	Somatization	Appetite decrease	Self-hatred
Depressed mood	— ^a	—	–0.338 ^b	–0.171	–0.239	—	—
Loss of interest	—	—	—	–0.229	–0.380 ^b	—	—
Loss of pleasure	—	—	—	–1.111 ^b	—	—	—
Loss of energy	—	—	—	–0.108	–0.723 ^b	—	—
Tiredness	—	—	–0.274	—	–0.493 ^b	—	—
Difficulty falling asleep	0.143	–0.350 ^b	–0.135	—	–0.184	—	0.157
Disturbed sleep	—	–0.833 ^b	—	—	—	—	—
Lack of restful sleep	0.220	–0.497 ^b	–0.165	—	–0.221	—	—
Early morning awakening	—	–0.602 ^b	—	—	—	—	—
Hypersomnia	0.163	0.147	—	—	–0.435 ^b	—	—
Weight loss	—	—	—	—	—	0.455 ^b	—
Loss of appetite	—	—	—	—	—	0.591 ^b	—
Increased appetite	—	—	—	—	–0.416 ^b	–0.241	—
Psychomotor agitation	—	—	—	—	—	—	—
Psychomotor retardation	—	—	—	—	0.150	—	—
Inner tension	—	–0.236	—	—	–0.210	—	—
Indecision	–0.264	–0.137	—	—	–0.228	—	—
Diminished ability to think	—	—	–0.194	—	–0.310 ^b	—	—
Rumination	–0.233	–0.148	–0.301 ^b	—	—	—	—
Feelings of hopelessness	—	—	–0.673 ^b	—	—	—	—
Feelings of worthlessness	–0.395 ^b	—	—	—	—	—	0.300 ^b
Loss of self-esteem	–0.547 ^b	—	—	—	—	—	—
Unreasonable feelings of guilt	—	—	—	—	—	—	0.816 ^b
Unreasonable feelings of self-reproach	—	—	—	—	—	—	0.869 ^b
Low self-esteem	–0.529 ^b	—	—	—	—	—	0.184
Crying	–0.183	—	–0.263	—	—	—	0.128
Despair	—	—	–0.780 ^b	—	—	—	—
Loss of affective reactivity	—	—	–0.186	–0.389 ^b	—	—	—
Circadian rhythms: worse in the morning	—	–0.239	—	—	–0.210	—	—
Loss of libido	—	–0.161	—	—	–0.291	—	—
Suicidality	—	—	–0.297	—	—	—	—

^aNot applicable.^bFactor loadings greater than 0.3.

Convergent and Divergent Validity

Regarding convergent validity, the KDD showed very high correlations with both the BDI-II ($r=0.70$; $P<.001$) and PHQ-9

($r=0.68$; $P<.001$). Regarding divergent validity, moderate associations were apparent with the PHQ-15 ($r=0.35$; $P<.001$). In contrast, the GAD-7 ($r=0.51$; $P<.001$) and the SWLS ($r=-0.51$; $P<.001$) were strongly correlated with the KDD.

Sensitivity to Change

The KDD mean scores of the improved symptom group (n=28) demonstrated a significant difference ($t_{27}=5.36, P<.001$) between admission (mean 36.85, SD 16.43) and discharge (mean 23.90, SD 18.54), with a high effect size (Cohen $d=0.79$) and a standardized response mean of 1.01. The means of the other measurements can be found in [Multimedia Appendix 7](#). Values of the unchanged symptom group could not be reliably calculated due to the small group size (n=4). A total of 27 patients were excluded from the analysis due to missing BDI-II and PHQ-9 scores or deteriorated symptoms. Additionally, the KDD revealed a high repeated measures correlation with both the BDI-II ($r=0.61; P<.001$) and the PHQ-9 ($r=0.59; P<.001$). The dropout analysis revealed significant differences between patients who completed the Klenico system and standard questionnaires at discharge and those who did not. Completers showed higher scores in the KDD, BDI-II, and SWLS, and a longer length of stay (see [Multimedia Appendix 7](#) for detailed results). Post hoc analysis revealed statistical power above 0.80.

Discussion

Principal Findings

This research assessed the validity of the KDD by focusing on its construct validity, reliability, and sensitivity to change.

Regarding KDDs’ underlying structure, the EFA suggested a 7-factor model ([Table 2](#)). The items on “psychomotor retardation” and “agitation” as well as for “inner tension,” “indecision,” “crying,” “circadian rhythms: worse in the morning,” “loss of libido,” and “suicidality,” did not adequately load on any of the factors.

The 7 factors of the KDD roughly reflect the 10 diagnostic criteria for depression as outlined in the classification of *ICD-10*. [Table 3](#) summarizes which KDD factors reflect which *ICD-10* diagnostic criteria. The KDD factors “inadequacy,” “sleep disturbance,” “self-hatred,” “anhedonia,” “affect,” and “appetite decrease” each reflect a single *ICD-10* criterion. The factor “somatization” maps on 3 criteria.

Table 3. Overview of the International Statistical Classification of Diseases, Tenth Revision (*ICD-10*) diagnostic criteria reflected by the Klenico Depression Domain (KDD) factors. Overview of the *ICD-10* diagnostic criteria reflected by the KDD factors.

KDD factors	<i>ICD-10</i> diagnostic criteria
Inadequacy	• Loss of confidence or self-esteem
Sleep disturbance	• Sleep disturbance any type
Anhedonia	• Loss of interest or pleasure in activities that are normally pleasurable
Self-hatred	• Unreasonable feelings of self-reproach or excessive and inappropriate guilt
Affect	• Depressed mood
Somatization	• Decreased energy or increased fatiguability • Complaints or evidence of diminished ability to think or concentrate... • Change in appetite (decrease or increase) with corresponding weight change
Appetite decrease	• Change in appetite

These results indicate that the *ICD-10* criteria for depression are subject to a complex and heterogeneous structure composed of different symptom dimensions. This is underlined by the fact that the latent structure of depression is widely disputed [57]. A recent study on the factor structure of the *DSM-IV* major depression criteria showed comparable results to our findings, as it demonstrated a multidimensional symptom structure with 5 factors for the 24 criteria [58].

The multiple mappings of the “somatization” factor on the *ICD-10* criteria and the inadequate loading of some items can be explained by the sample-specific characteristics, the Klenico system, and the structure of depression.

The factors “somatization” and “appetite decrease” reflect those *ICD-10* criteria categorized as somatic symptoms in depression that have been assigned to a specific somatic factor in various previous studies [57]. This confirms that the heterogeneous group of somatic symptoms has an underlying common structure

that is separate from the rest of the affective and cognitive symptoms of depression.

A possible explanation for the split into 2 separate somatic factors is that loss of appetite and the resulting weight loss are related to the overall severity of depression and may be determined by the severity of other symptoms, such as anhedonia, which is usually most pronounced in severely affected patients, potentially leading to a small variance and a possible floor effect [59]. Finally, the separation found between general somatic symptoms and sleep disturbance does not contradict the factorial validity of the KDD, since previous studies have shown that the 2 symptom dimensions impose separate factors [58].

Effects comparable to those at the factor level may also have occurred at the item level. The lack of significant factor loadings for “psychomotor agitation” and “retardation” may be due to their inclusion only in the clinical validation and not in the Klenico self-report module. These symptoms are more

accurately assessed through external observation rather than self-report and are often more pronounced in severe depression [23,60], which can lead to reduced variability and a resulting floor effect in the data. This smaller variance might hinder their ability to align clearly with a specific factor in the analysis. The same reasoning accounts for the items for “suicidality” and “loss of libido,” both of which were just below the cutoff of 0.3—“suicidality” with loads just not on “affect” and “loss of libido” with loads just not on “somatization.” Their prevalences are correlated with the severity of depression [59,61], and acute suicidality is mainly an indication for treatment in a psychiatric clinic. Since the present sample comes from a psychotherapeutic clinic in which suicidality is an exclusion criterion, it is to be expected that such patients were less present.

“Inner tension,” “indecision,” “circadian rhythms: worse in the morning,” and “crying” did not load strongly on a single factor, instead showing distributed loadings across multiple factors. “Inner tension” aligns with both the “sleep disturbance” and “somatization” factors, indicating its role in exacerbating both sleep issues and physical symptoms. “Indecision” correlates with the “somatization” and “inadequacy” factors, highlighting its connection to cognitive impairments and reduced self-confidence typical in depressive episodes. “Circadian rhythms: worse in the morning” describes the symptom of morning mood lows in depression. Due to the wording, it is likely that this item is answered in a rather undifferentiated way and can therefore be associated by patients with both poor sleep and physical effects. “Crying” is a universal symptom that may not load strongly on any specific factor, likely due to its general association with various forms of distress across the depressive spectrum. It possibly associated with multiple factors such as “affect” and “self-hatred,” reflecting its role as a common expression of emotional pain, sadness, and feelings of low worth. This pattern suggests these items intersect various dimensions of depression, making their specific factor alignment more complex. Overall, the findings confirm that depression is a multidimensional construct and suggest that the KDD is determined by the heterogeneous structure of the *ICD-10* criteria. Since the structure of the KDD faithfully mirrors the *ICD-10* criteria, this offers initial support for its factorial validity.

Concerning convergent validity, large associations were found between the KDD and the corresponding BDI-II and PHQ-9, as expected. Additionally, the KDD showed medium associations with divergent PHQ-15. Exceptions were the GAD-7 and SWLS, with high-range correlations.

Similar to the structure of depression, the factor structure for different diagnostic tools is widely debated. The number of factors found for the BDI-II ranges between 2 and 4 [44] and for the PHQ-9 ranges between 1 and 2 [62]. For both, various compositions of the different items loading on the different factors were found. However, although the KDD reflects the factor structure according to *ICD-10* criteria, it showed significantly high associations with these 2 instruments, thus indicating the measurement of a comparable construct of depression.

The results regarding divergent validity are not surprising, as comorbidity between depression and various anxiety disorders is highly frequent [63], and depressive symptoms show structural overlap with anxiety disorders [64]. This is further supported by other studies demonstrating strong intercorrelations between the GAD-7 and PHQ-9 [65], as well as significant correlations between the GAD-7 and BDI-II. [66]

Moreover, depressive disorders are strongly negatively associated with life satisfaction and subjective well-being [67]. The strong negative correlation with the SWLS aligns with theoretical expectations and highlights the substantial impairment in life satisfaction that accompanies depressive symptomatology.

The results therefore do not counteract the construct validity of the KDD but rather support its ability to capture the broader context of psychological distress and well-being.

Regarding sensitivity to change, the KDD was able to detect improvements in symptom severity according to PHQ-9 or BDI-II sum scores and showed strong associations with changes in the BDI-II and PHQ-9 sum scores. This demonstrates that the KDD is highly sensitive to symptom change and is therefore able to monitor treatment outcomes, similar to the commonly used self-report questionnaires.

Overall, the results show that the KDD is a reliable and valid instrument for the diagnosis of depression according to *ICD-10* criteria while being able to adequately recognize improvements in symptom severity. In contrast to routine diagnostic procedures, the KDD enables a state-of-the-art diagnostic procedure that includes both self-report and clinical validation by experts and is additionally able to assign comorbidities by covering all relevant mental disorders (n=12).

A major strength of this study is that the KDD was tested in a large clinical sample under real-world conditions. It could thus be demonstrated that the KDD measures adequately in patients with frequent comorbidities in clinical reality, providing real-life evidence of its validity.

Limitations and Outlook

A limitation of the study was the small sample size at discharge. Despite the fact that the post hoc power analysis has shown sufficient power for the *t* tests and the repeated measurement correlation, the findings regarding the sensitivity of change should be considered preliminary. Furthermore, the analysis of the sensitivity of change in the unchanged symptom group was not feasible due to the limited sample size and requires further research. In addition, the patients who used the Klenico tool at discharge exhibited greater symptom severity than those who did not. A reason for this higher representation may be that those more heavily burdened patients perceived greater relevance in completing the exit assessment, leading to higher motivation. The composition of the sample also imposed certain limitations. On the one hand, less severe cases were to be expected in the psychotherapeutic setting, and on the other hand, the design of the study could further lead to the possible exclusion of severe depression cases. This was shown by the fact that, for example, “severe depressive episode with psychotic symptoms” were extremely rare, but also “mild depressive

episode” were relatively uncommon. Since the completion of the Klenico assessment requires a certain level of motivation and cognitive abilities, patients affected by severe concentration problems and anhedonia may have faced difficulties in participation. These circumstances led to certain restrictions in the factor analysis, for example, the nonloading of the item “psychomotor retardation.” Additionally, the sample exhibited a gender imbalance, with a majority (68%) being female. This mirrors the real-world scenario where women are more commonly affected by mental disorders compared with men [68]. However, this imbalance could also impact the outcome, given that different genders may exhibit variations in the manifestation of depressive symptoms [69]. Moreover, the age distribution of our participants, concentrated primarily in the younger adult segment and in the segment aged around 55 years, suggests a demographic skew that may have influenced the clinical profile observed. For these reasons, the use of the psychotherapeutic sample in this study might restrict the generalizability of the findings to broader populations, particularly individuals with severe depression. Therefore, it would be valuable to assess the functionality and generalizability of the KDD in different settings that involve patients with more severe symptoms and with a balanced gender and age distribution. This would provide a more comprehensive understanding of the measure’s effectiveness across diverse populations.

This analysis focused on examining Klenico assessment of symptom severity, and Klenico-suggested diagnoses were not examined. Therefore, this should be added in future analyses. Finally, the current validation was limited to the German version of the Klenico system. Further validation studies are necessary to assess the system in other languages, such as English. While the EFA provided initial insights into the factor structure of the KDD, these results need to be confirmed through a confirmatory factor analysis. Additionally, the factor structures of other Klenico domains also require assessment to ensure the comprehensive validity of the system across different mental health conditions.

Conclusions

The KDD presents itself as an accurate tool for measuring depressive symptoms and a valuable tool for measuring depressive symptoms and monitoring treatment outcomes in psychotherapeutic settings. As an integral part of the Klenico system, it comprehensively covers the whole range of depression symptoms according to *ICD-10* and might therefore be a comprehensive alternative for the diagnosis and course measurement of depression. The current assessment represents an important next step in validating the psychometric properties of the Klenico system. Therefore, it may help to avoid misdiagnosis and incorrect therapies through standardized and evidence-based diagnostic procedures, thus reducing health care costs and minimizing the burden on the patients.

Acknowledgments

This work was funded by Klenico Health AG.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SR was responsible for formal analysis, drafting the original manuscript, and reviewing and editing the manuscript. JS contributed to project administration, drafting the original manuscript, and reviewing and editing the manuscript. NH participated in drafting the original manuscript. MA oversaw supervision, validation, conceptualization, and reviewing and editing of the manuscript. UV contributed to the conceptualization and provided resources. AM handled data curation and participated in reviewing and editing the manuscript.

Conflicts of Interest

SR, NH, and MA were formerly employed by Klenico Health AG, while JS remains currently employed. We declare that this did not affect data collection or analysis. UV and AM report no conflicts of interest.

Multimedia Appendix 1

Item names, means and SDs of the different KDS items, as well as an exemplary selection of the item formulations.

[\[PDF File \(Adobe PDF File\), 142 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Frequencies of ICD-10 diagnostic categories. Groups were formed by assigning each individual diagnosis of each patient to the corresponding disorder domain. Multiple diagnoses are possible. Listed frequencies are relative to the number of total given diagnoses. ICD-10: International Statistical Classification of Diseases, Tenth Revision.

[\[PDF File \(Adobe PDF File\), 126 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

The most frequent somatic diagnoses in the sample, in descending order.

[\[PDF File \(Adobe PDF File\), 96 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

R code for the exploratory factor analyses.

[\[DOCX File , 24 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Exploratory factor analysis fit estimations for different number of factors.

[\[PDF File \(Adobe PDF File\), 100 KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Scree plot of exploratory factor analysis factor estimation.

[\[PDF File \(Adobe PDF File\), 134 KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

Means, SDs, medians, and IQRs of demographic parameters and standard questionnaires at admission for the whole patient sample, patients who were measured at discharge, and patients who were not measured at discharge.

[\[PDF File \(Adobe PDF File\), 138 KB-Multimedia Appendix 7\]](#)

References

1. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int J Epidemiol*. 2014;43(2):476-493. [\[FREE Full text\]](#) [doi: [10.1093/ije/dyu038](https://doi.org/10.1093/ije/dyu038)] [Medline: [24648481](#)]
2. Hirschfeld RMA, Montgomery SA, Keller MB, Kasper S, Schatzberg AF, Möller HJ, et al. Social functioning in depression: a review. *J Clin Psychiatry*. 2000;61(4):268-275. [doi: [10.4088/jcp.v61n0405](https://doi.org/10.4088/jcp.v61n0405)] [Medline: [10830147](#)]
3. Christensen MK, Lim CCW, Saha S, Plana-Ripoll O, Cannon D, Presley F, et al. The cost of mental disorders: a systematic review. *Epidemiol Psychiatr Sci*. Aug 18, 2020;29:e161. [\[FREE Full text\]](#) [doi: [10.1017/S204579602000075X](https://doi.org/10.1017/S204579602000075X)] [Medline: [32807256](#)]
4. Busch MA, Maske UE, Ryl L, Schlack R, Hapke U. Prävalenz von depressiver Symptomatik und diagnostizierter Depression bei Erwachsenen in Deutschland. Prevalence of depressive symptoms and diagnosed depression among adults in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. Federal Health Gazette Health Research Health Protection. May 2013;56(5-6):733-739. [doi: [10.1007/s00103-013-1688-3](https://doi.org/10.1007/s00103-013-1688-3)] [Medline: [23703492](#)]
5. Lim GY, Tam WW, Lu Y, Ho CS, Zhang MW, Ho RC. Prevalence of depression in the community from 30 countries between 1994 and 2014. *Sci Rep*. Feb 12, 2018;8(1):2861. [\[FREE Full text\]](#) [doi: [10.1038/s41598-018-21243-x](https://doi.org/10.1038/s41598-018-21243-x)] [Medline: [29434331](#)]
6. Eden JL, Konnopka A, König H-H. Costs of depression in Germany - systematic review. *Psychiatr Prax*. Sep 2021;48(6):290-300. [doi: [10.1055/a-1399-4430](https://doi.org/10.1055/a-1399-4430)] [Medline: [33773504](#)]
7. Avenevoli S, Swendsen J, He JP, Burstein M, Merikangas KR. Major depression in the national comorbidity survey-adolescent supplement: prevalence, correlates, and treatment. *J Am Acad Child Adolesc Psychiatry*. Jan 2015;54(1):37-44.e2. [\[FREE Full text\]](#) [doi: [10.1016/j.jaac.2014.10.010](https://doi.org/10.1016/j.jaac.2014.10.010)] [Medline: [25524788](#)]
8. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. Jun 2005;62(6):617-627. [\[FREE Full text\]](#) [doi: [10.1001/archpsyc.62.6.617](https://doi.org/10.1001/archpsyc.62.6.617)] [Medline: [15939839](#)]
9. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D. What are 'good' depression symptoms? comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *J Affect Disord*. 2016;189:314-320. [\[FREE Full text\]](#) [doi: [10.1016/j.jad.2015.09.005](https://doi.org/10.1016/j.jad.2015.09.005)] [Medline: [26458184](#)]
10. Bruchmüller K, Margraf J, Suppiger A, Schneider S. Popular or unpopular? therapists' use of structured interviews and their estimation of patient acceptance. *Behav Ther*. 2011;42(4):634-643. [doi: [10.1016/j.beth.2011.02.003](https://doi.org/10.1016/j.beth.2011.02.003)] [Medline: [22035992](#)]
11. Aboraya A. Use of structured interviews by psychiatrists in real clinical settings: results of an open-question survey. *Psychiatry (Edgmont)*. 2009;6(6):24-28. [\[FREE Full text\]](#) [Medline: [19724758](#)]
12. Anderson DA, Paulosky CA. A survey of the use of assessment instruments by eating disorder professionals in clinical practice. *Eat Weight Disord*. 2013;9(3):238-241. [doi: [10.1007/bf03325075](https://doi.org/10.1007/bf03325075)]

13. Miller PR. Inpatient diagnostic assessments: 2. interrater reliability and outcomes of structured vs. unstructured interviews. *Psychiatry Res.* 2001;105(3):265-271. [doi: [10.1016/S0165-1781\(01\)00318-3](https://doi.org/10.1016/S0165-1781(01)00318-3)]
14. Andreas S, Theisen P, Mestel R, Koch U, Schulz H. Validity of routine clinical DSM-IV diagnoses (Axis I/II) in inpatients with mental disorders. *Psychiatry Res.* 2009;170(2-3):252-255. [doi: [10.1016/j.psychres.2008.09.009](https://doi.org/10.1016/j.psychres.2008.09.009)] [Medline: [19896721](#)]
15. Shear MK, Greeno C, Kang J, Ludewig D, Frank E, Swartz HA, et al. Diagnosis of nonpsychotic patients in community clinics. *Am J Psychiatry.* 2000;157(4):581-587. [doi: [10.1176/appi.ajp.157.4.581](https://doi.org/10.1176/appi.ajp.157.4.581)] [Medline: [10739417](#)]
16. Serra F, Spoto A, Ghisi M, Vidotto G. Formal psychological assessment in evaluating depression: a new methodology to build exhaustive and irredundant adaptive questionnaires. *PLoS One.* 2015;10(4):e0122131. [FREE Full text] [doi: [10.1371/journal.pone.0122131](https://doi.org/10.1371/journal.pone.0122131)] [Medline: [25875359](#)]
17. Newson JJ, Hunter D, Thiagarajan TC. The heterogeneity of mental health assessment. *Front Psychiatry.* 2020;11:76. [FREE Full text] [doi: [10.3389/fpsy.2020.00076](https://doi.org/10.3389/fpsy.2020.00076)] [Medline: [32174852](#)]
18. Ferentinos P, Paparrigopoulos T, Rentzos M, Zouvelou V, Alexakis T, Evdokimidis I. Prevalence of major depression in ALS: comparison of a semi-structured interview and four self-report measures. *Amyotroph Lateral Scler.* 2011;12(4):297-302. [doi: [10.3109/17482968.2011.556744](https://doi.org/10.3109/17482968.2011.556744)] [Medline: [21428731](#)]
19. Maurer DM. Screening for depression. *Am Fam Physician.* 2012;85(2):139-144. [FREE Full text] [Medline: [22335214](#)]
20. Zimmerman M, McGlinchey JB. Why don't psychiatrists use scales to measure outcome when treating depressed patients? *J Clin Psychiatry.* 2008;69(12):1916-1919. [doi: [10.4088/jcp.v69n1209](https://doi.org/10.4088/jcp.v69n1209)] [Medline: [19192467](#)]
21. Hatfield DR, Ogles BM. Why some clinicians use outcome measures and others do not. *Adm Policy Ment Health.* 2007;34(3):283-291. [doi: [10.1007/s10488-006-0110-y](https://doi.org/10.1007/s10488-006-0110-y)] [Medline: [17211715](#)]
22. Jensen-Doss A, Haimes EMB, Smith AM, Lyon AR, Lewis CC, Stanick CF, et al. Monitoring treatment progress and providing feedback is viewed favorably but rarely used in practice. *Adm Policy Ment Health.* 2018;45(1):48-61. [FREE Full text] [doi: [10.1007/s10488-016-0763-0](https://doi.org/10.1007/s10488-016-0763-0)] [Medline: [27631610](#)]
23. DGPPN BÄK, KBV, AWMF, Härter M, Schorr S, Schneider F, editors. S3-Leitlinie/Nationale Versorgungsleitlinie Unipolare Depression-Langfassung. S3 Guideline/National Care Guideline on Unipolar Depression-Long Version. Berlin, Heidelberg. Springer; Mar 10, 2017.
24. J. Rakofsky J, W. Dunlop B. The over-under on the misdiagnosis of bipolar disorder: a systematic review. *CPSR.* 2015;11(4):222-234. [doi: [10.2174/157340051266615117211523](https://doi.org/10.2174/157340051266615117211523)]
25. Vermani M, Marcus M, Katzman MA. Rates of detection of mood and anxiety disorders in primary care. *Prim Care Companion CNS Disord.* 2011;27211. [doi: [10.4088/pcc.10m01013](https://doi.org/10.4088/pcc.10m01013)]
26. Takara K, Kondo T, Kuba T. How and why is autism spectrum disorder misdiagnosed in adult patients? - from diagnostic problem to management for adjustment -. *Ment Health Fam Med.* Nov 30, 2015;11(02):73-88. [FREE Full text] [doi: [10.2514/1756-8358.1102011](https://doi.org/10.2514/1756-8358.1102011)]
27. Wahl K, Kordon A, Kuelz KA, Voderholzer U, Hohagen F, Zurowski B. Obsessive-compulsive disorder (OCD) is still an unrecognised disorder: a study on the recognition of OCD in psychiatric outpatients. *Eur Psychiatry.* 2010;25(7):374-377. [doi: [10.1016/j.eurpsy.2009.12.003](https://doi.org/10.1016/j.eurpsy.2009.12.003)] [Medline: [20627468](#)]
28. Fusar-Poli L, Brondino N, Politi P, Aguglia E. Missed diagnoses and misdiagnoses of adults with autism spectrum disorder. *Eur Arch Psychiatry Clin Neurosci.* 2022;272(2):187-198. [FREE Full text] [doi: [10.1007/s00406-020-01189-w](https://doi.org/10.1007/s00406-020-01189-w)] [Medline: [32892291](#)]
29. Zander E, Wyder L, Holtforth MG, Schnyder U, Hepp U, Stulz N. Validity of routine clinical diagnoses in acute psychiatric inpatients. *Psychiatry Res.* 2018;259:482-487. [doi: [10.1016/j.psychres.2017.11.004](https://doi.org/10.1016/j.psychres.2017.11.004)] [Medline: [29154169](#)]
30. Zimmerman M. A review of 20 years of research on overdiagnosis and underdiagnosis in the Rhode Island methods to improve diagnostic assessment and services (MIDAS) project. *Can J Psychiatry.* 2016;61(2):71-79. [FREE Full text] [doi: [10.1177/0706743715625935](https://doi.org/10.1177/0706743715625935)] [Medline: [27253697](#)]
31. Herzog DP, Wagner S, Ruckes C, Tadic A, Roll SC, Härter M, et al. Guideline adherence of antidepressant treatment in outpatients with major depressive disorder: a naturalistic study. *Eur Arch Psychiatry Clin Neurosci.* 2017;267(8):711-721. [doi: [10.1007/s00406-017-0798-6](https://doi.org/10.1007/s00406-017-0798-6)] [Medline: [28421334](#)]
32. Stahnke B. A systematic review of misdiagnosis in those with obsessive-compulsive disorder. *J Affect Disord Rep.* 2021;6:100231. [doi: [10.1016/j.jadr.2021.100231](https://doi.org/10.1016/j.jadr.2021.100231)]
33. Trautmann S, Beesdo-Baum K. The treatment of depression in primary care. *Dtsch Arztebl Int.* 2017;114(43):721-728. [FREE Full text] [doi: [10.3238/arztebl.2017.0721](https://doi.org/10.3238/arztebl.2017.0721)] [Medline: [29143731](#)]
34. Fernández A, Haro JM, Codony M, Vilagut G, Martínez-Alonso M, Autonell J, et al. Treatment adequacy of anxiety and depressive disorders: primary versus specialised care in Spain. *J Affect Disord.* 2006;96(1-2):9-20. [doi: [10.1016/j.jad.2006.05.005](https://doi.org/10.1016/j.jad.2006.05.005)] [Medline: [16793140](#)]
35. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet.* 2007;370(9590):841-850. [FREE Full text] [doi: [10.1016/S0140-6736\(07\)61414-7](https://doi.org/10.1016/S0140-6736(07)61414-7)] [Medline: [17826169](#)]
36. Wang PS, Demler O, Kessler RC. Adequacy of treatment for serious mental illness in the United States. *Am J Public Health.* 2002;92(1):92-98. [doi: [10.2105/ajph.92.1.92](https://doi.org/10.2105/ajph.92.1.92)] [Medline: [11772769](#)]

37. Lora A, Conti V, Leoni O, Rivolta AL. Adequacy of treatment for patients with schizophrenia spectrum disorders and affective disorders in Lombardy, Italy. *Psychiatr Serv*. 2011;62(9):1079-1084. [doi: [10.1176/ps.62.9.pss6209_1079](https://doi.org/10.1176/ps.62.9.pss6209_1079)] [Medline: [21885588](https://pubmed.ncbi.nlm.nih.gov/21885588/)]
38. Aboraya A, Nasrallah HA, Elswick DE, Ahmed E, Estephan N, Aboraya D, et al. Measurement-based care in psychiatry-past, present, and future. *Innov Clin Neurosci*. 2018;15(11-12):13-26. [FREE Full text] [Medline: [30834167](https://pubmed.ncbi.nlm.nih.gov/30834167/)]
39. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. World Health Organization. 1992. URL: <https://apps.who.int/iris/handle/10665/37958> [accessed 2025-02-26]
40. Lustig L. Die Klenico-Items: theoretische Herleitung und empirische Fundierung eines neuen Ansatzes in der psychiatrischen Diagnostik. The Klenico Items: Theoretical Derivation and Empirical Foundation of a New Approach in Psychiatric Diagnostics. Zurich, Switzerland. University of Zurich; 2020.
41. Reutimann S, Milanovic D, Gulewitsch MD, Augsburg M. Preliminary validation of the Klenico diagnostic software self-report module through comparison with the diagnostic gold standard in an outpatient routine clinical sample. *Health Psychol Behav Med*. 2023;11(1):2244576. [FREE Full text] [doi: [10.1080/21642850.2023.2244576](https://doi.org/10.1080/21642850.2023.2244576)] [Medline: [37663014](https://pubmed.ncbi.nlm.nih.gov/37663014/)]
42. Beck AT, Steer RA, Brown GK. Beck Depression Inventory (BDI-II). London, UK. Pearson; 1996.
43. Sprinkle SD, Lurie D, Insko SL, Atkinson G, Jones GL, Logan AR, et al. Criterion validity, severity cut scores, and test-retest reliability of the Beck Depression Inventory-II in a university counseling center sample. *J Couns Psychol*. 2002;49(3):381-385. [doi: [10.1037/0022-0167.49.3.381](https://doi.org/10.1037/0022-0167.49.3.381)]
44. Wang YP, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Braz J Psychiatry*. 2013;35(4):416-431. [FREE Full text] [doi: [10.1590/1516-4446-2012-1048](https://doi.org/10.1590/1516-4446-2012-1048)] [Medline: [24402217](https://pubmed.ncbi.nlm.nih.gov/24402217/)]
45. Meesters Y, Sijbrandij JJ, Visser E, de Beurs E. Sensitivity to change of the Beck Depression Inventory versus the inventory of depressive symptoms. *J Affect Disord*. 2021;281:338-341. [FREE Full text] [doi: [10.1016/j.jad.2020.12.036](https://doi.org/10.1016/j.jad.2020.12.036)] [Medline: [33341647](https://pubmed.ncbi.nlm.nih.gov/33341647/)]
46. Löwe B, Spitzer RL, Zipfel S, Herzog W. Gesundheitsfragebogen für patienten (PHQ-D). komplettversion und kurzform. testmappe mit manual, fragebögen, schablonen. 2. auflage. Health Questionnaire for Patients (PHQ-D). complete version and short form. test folder with manual, questionnaires, templates. 2nd edition. GESIS – Leibniz Institute for the Social Sciences. 2002. URL: <https://search.gesis.org/publication/zis-L%C3%B6weSpitzerZipfelHerzog2002Gesundheitsfragebogen> [accessed 2025-01-31]
47. Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer störungen mit dem “Gesundheitsfragebogen für Patienten (PHQ-D)“. Screening for mental disorders using the ‘Health Questionnaire for Patients (PHQ-D)’. *Diagnostica*. 2004;50(4):171-181. [doi: [10.1026/0012-1924.50.4.171](https://doi.org/10.1026/0012-1924.50.4.171)]
48. Zuithoff NP, Vergouwe Y, King M, Nazareth I, van Wezep MJ, Moons KG, et al. The Patient Health Questionnaire-9 for detection of major depressive disorder in primary care: consequences of current thresholds in a crosssectional study. *BMC Fam Pract*. Dec 13, 2010;11(1):98. [FREE Full text] [doi: [10.1186/1471-2296-11-98](https://doi.org/10.1186/1471-2296-11-98)] [Medline: [21144018](https://pubmed.ncbi.nlm.nih.gov/21144018/)]
49. Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord*. 2004;81(1):61-67. [doi: [10.1016/S0165-0327\(03\)00198-8](https://doi.org/10.1016/S0165-0327(03)00198-8)] [Medline: [15183601](https://pubmed.ncbi.nlm.nih.gov/15183601/)]
50. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. *J Pers Assess*. Feb 1985;49(1):71-75. [doi: [10.1207/s15327752jpa4901_13](https://doi.org/10.1207/s15327752jpa4901_13)] [Medline: [16367493](https://pubmed.ncbi.nlm.nih.gov/16367493/)]
51. Hinz A, Conrad I, Schroeter ML, Glaesmer H, Brähler E, Zenger M, et al. Psychometric properties of the Satisfaction With Life Scale (SWLS), derived from a large German community sample. *Qual Life Res*. Jun 2018;27(6):1661-1670. [doi: [10.1007/s11136-018-1844-1](https://doi.org/10.1007/s11136-018-1844-1)] [Medline: [29589212](https://pubmed.ncbi.nlm.nih.gov/29589212/)]
52. Glaesmer H, Grande G, Braehler E, Roth M. The German version of the Satisfaction With Life Scale (SWLS). *Eur J Psychol Assess*. 2011;27(2):127-132. [doi: [10.1027/1015-5759/a000058](https://doi.org/10.1027/1015-5759/a000058)]
53. Marsh HW, Hau KT, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Struct Equ Model*. Jul 2004;11(3):320-341. [doi: [10.1207/s15328007sem1103_2](https://doi.org/10.1207/s15328007sem1103_2)]
54. Cohen J. Statistical Power Analysis for the Behavioral Sciences. London, United Kingdom. Routledge; 2013.
55. Bakdash JZ, Marusich LR. Repeated measures correlation. *Front Psychol*. 2017;8:456. [FREE Full text] [doi: [10.3389/fpsyg.2017.00456](https://doi.org/10.3389/fpsyg.2017.00456)] [Medline: [28439244](https://pubmed.ncbi.nlm.nih.gov/28439244/)]
56. Medical Devices Act. Federal Ministry of Justice and Consumer Protection. 2002. URL: <https://tinyurl.com/443hfw8w> [accessed 2025-06-213]
57. Bi Y, Wang L, Cao C, Fang R, Li G, Liu P, et al. The factor structure of major depressive symptoms in a sample of Chinese earthquake survivors. *BMC Psychiatry*. Jan 28, 2021;21(1):59. [FREE Full text] [doi: [10.1186/s12888-020-02993-3](https://doi.org/10.1186/s12888-020-02993-3)] [Medline: [33509157](https://pubmed.ncbi.nlm.nih.gov/33509157/)]
58. van Loo HM, Aggen SH, Kendler KS. The structure of the symptoms of major depression: factor analysis of a lifetime worst episode of depressive symptoms in a large general population sample. *J Affect Disord*. Jun 15, 2022;307:115-124. [FREE Full text] [doi: [10.1016/j.jad.2022.03.064](https://doi.org/10.1016/j.jad.2022.03.064)] [Medline: [35367501](https://pubmed.ncbi.nlm.nih.gov/35367501/)]

59. Zimmerman M, Balling C, Chelminski I, Dalrymple K. Understanding the severity of depression: which symptoms of depression are the best indicators of depression severity? *Compr Psychiatry*. Nov 2018;87:84-88. [doi: [10.1016/j.comppsy.2018.09.006](https://doi.org/10.1016/j.comppsy.2018.09.006)] [Medline: [30282058](https://pubmed.ncbi.nlm.nih.gov/30282058/)]
60. Angst J, Gamma A, Benazzi F, Ajdacic V, Rössler W. Does psychomotor agitation in major depressive episodes indicate bipolarity? evidence from the Zurich study. *Eur Arch Psychiatry Clin Neurosci*. 2009;259(1):55-63. [doi: [10.1007/s00406-008-0834-7](https://doi.org/10.1007/s00406-008-0834-7)] [Medline: [18806921](https://pubmed.ncbi.nlm.nih.gov/18806921/)]
61. Williams K, Reynolds MF. Sexual dysfunction in major depression. *CNS Spectr*. Aug 2006;11(8 Suppl 9):19-23. [doi: [10.1017/s1092852900026729](https://doi.org/10.1017/s1092852900026729)] [Medline: [16871134](https://pubmed.ncbi.nlm.nih.gov/16871134/)]
62. Lamela D, Soreira C, Matos P, Morais A. Systematic review of the factor structure and measurement invariance of the Patient Health Questionnaire-9 (PHQ-9) and validation of the portuguese version in community settings. *J Affect Disord*. 2020;276:220-233. [doi: [10.1016/j.jad.2020.06.066](https://doi.org/10.1016/j.jad.2020.06.066)] [Medline: [32697702](https://pubmed.ncbi.nlm.nih.gov/32697702/)]
63. Li Y, Shi S, Yang F, Gao J, Li Y, Tao M, et al. Patterns of co-morbidity with anxiety disorders in Chinese women with recurrent major depression. *Psychol Med*. Nov 30, 2011;42(6):1239-1248. [doi: [10.1017/s003329171100273x](https://doi.org/10.1017/s003329171100273x)]
64. Forbes MK, Sunderland M, Rapee RM, Batterham PJ, Calear AL, Carragher N, et al. A detailed hierarchical model of psychopathology: from individual symptoms up to the general factor of psychopathology. *Clin Psychol Sci*. 2021;9(2):139-168. [FREE Full text] [doi: [10.1177/2167702620954799](https://doi.org/10.1177/2167702620954799)] [Medline: [33758691](https://pubmed.ncbi.nlm.nih.gov/33758691/)]
65. Villarreal-Zegarra D, Paredes-Angeles R, Mayo-Puchoc N, Arenas-Minaya E, Huarcaya-Victoria J, Copez-Lonzoy A. Psychometric properties of the GAD-7 (General Anxiety Disorder-7): a cross-sectional study of the Peruvian general population. *BMC Psychol*. Apr 02, 2024;12(1):183. [FREE Full text] [doi: [10.1186/s40359-024-01688-8](https://doi.org/10.1186/s40359-024-01688-8)] [Medline: [38566138](https://pubmed.ncbi.nlm.nih.gov/38566138/)]
66. Johnson SU, Ulvenes PG, Øktedalen T, Hoffart A. Psychometric properties of the General Anxiety Disorder 7-Item (GAD-7) scale in a heterogeneous psychiatric sample. *Front Psychol*. 2019;10:1713. [FREE Full text] [doi: [10.3389/fpsyg.2019.01713](https://doi.org/10.3389/fpsyg.2019.01713)] [Medline: [31447721](https://pubmed.ncbi.nlm.nih.gov/31447721/)]
67. Gigantesco A, Fagnani C, Toccaceli V, Stazi MA, Lucidi F, Violani C, et al. The relationship between satisfaction with life and depression symptoms by gender. *Front Psychiatry*. 2019;10:419. [FREE Full text] [doi: [10.3389/fpsyg.2019.00419](https://doi.org/10.3389/fpsyg.2019.00419)] [Medline: [31258495](https://pubmed.ncbi.nlm.nih.gov/31258495/)]
68. Boyd A, van de Velde S, Vilagut G, de Graaf R, O'Neill S, Florescu S, et al. Gender differences in mental disorders and suicidality in Europe: results from a large cross-sectional population-based study. *J Affect Disord*. 2015;173:245-254. [doi: [10.1016/j.jad.2014.11.002](https://doi.org/10.1016/j.jad.2014.11.002)] [Medline: [25462424](https://pubmed.ncbi.nlm.nih.gov/25462424/)]
69. Winkler D, Pjrek E, Heiden A, Wiesegeger G, Klein N, Konstantinidis A, et al. Gender differences in the psychopathology of depressed inpatients. *Eur Arch Psychiatry Clin Neurosci*. Aug 2004;254(4):209-214. [doi: [10.1007/s00406-004-0471-8](https://doi.org/10.1007/s00406-004-0471-8)] [Medline: [15309388](https://pubmed.ncbi.nlm.nih.gov/15309388/)]

Abbreviations

BDI-II: Beck Depression Inventory
DSM: Diagnostic and Statistical Manual of Mental Disorders
EFA: exploratory factor analyses
GAD-7: Generalized Anxiety Disorder-7
ICD-10: International Statistical Classification of Diseases, Tenth Revision
KDD: Klenico Depression Domain
KMO: Kaiser Meyer Olkin
PHQ: Patient Health Questionnaire
SWLS: Satisfaction With Life Scale

Edited by A Mavragani; submitted 04.07.23; peer-reviewed by V Sideropoulos, C Haag, B Guo; comments to author 31.01.24; revised version received 29.04.24; accepted 27.12.24; published 24.07.25

Please cite as:

Reutimann S, Steiner J, Hübscher N, Voderholzer U, Meule A, Augsburg M
 Evaluating the Construct Validity and Sensitivity to Change of the Klenico Depression Domain in Psychotherapeutic Inpatient Care:
 Instrument Validation Study
JMIR Form Res 2025;9:e50504
 URL: <https://formative.jmir.org/2025/1/e50504>
 doi: [10.2196/50504](https://doi.org/10.2196/50504)
 PMID:

©Stefan Reutimann, Jasmin Steiner, Noah Hübscher, Ulrich Voderholzer, Adrian Meule, Mareike Augsburg. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 24.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.