MASTER MEASUREMENT REPOSITORY DOCUMENT

Andrea Barnes, DBH, MSCN, RD, LDN, FAND, ACPMC

COLLABORATIVE FAMILY HEALTH ASSOCIATION

Table of Contents

Adult ADHD Self-Report Scale (ASRS v1.1/ASRS-5)[1-9]

1. Objective

To evaluate the psychometric reliability, validity, clinical utility, and feasibility of the Adult ADHD Self-Report Scales (ASRS v1.1 and updated ASRS-5) for the screening and assessment of adult Attention-Deficit/Hyperactivity Disorder (ADHD) in primary care, psychiatric, and community settings across diverse populations.

---

2. Measure Evaluated

ASRS v1.1 – A 6- and 18-item self-report screening tool based on DSM-IV criteria, widely used for identifying probable ADHD in adults.

ASRS-5 – A 6-item updated version calibrated to DSM-5 criteria using machine learning methods to enhance diagnostic accuracy and brevity.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- Internal consistency:
  - ASRS v1.1: α = 0.86–0.88 (full), 0.82 (Inattention), 0.78 (Hyperactivity/Impulsivity)
  - ASRS-5: α = 0.88 (German version)
- Test–retest reliability:
  - ASRS-18: ICC = 0.77; ASRS-6: ICC = 0.76
- Convergent validity:
  - Correlation with structured diagnostic interviews r = 0.73–0.79
- Diagnostic accuracy:
  - ASRS-6 (DSM-5 version): Sensitivity = 91.4%, Specificity = 96%, AUC = 0.94
  - ASRS-6 (original): Sensitivity ≈ 68–83%, Specificity ≈ 80–87%, AUC ≈ 0.90
- Cross-cultural validations:
  - Turkish, German, South African, and multinational U.S. samples

Pragmatic Properties:

- **Acceptability:** Very brief (<2 min for ASRS-6); positive evaluations by clinicians and patients
- **Feasibility:** Easy EHR integration; suitable for high-volume settings
- **Interpretability:** Cutoffs—ASRS-18 ≥30; ASRS-6 ≥4 positives; DSM-5 ASRS-5 uses weighted scoring thresholds
- **Equity:** Successfully validated across different languages, psychiatric populations, and primary care cohorts
- **Sustainability:** Recommended for routine primary care screening in updated guidelines (e.g., Germany, U.S.)

---

4. Application to Integrated Primary Care

The ASRS is highly effective for adult ADHD screening in primary care. It enables early identification in patients presenting with complex psychiatric and functional concerns. ASRS-5's brevity and accuracy make it particularly suitable for routine screening in busy practices, telehealth platforms, and integrated behavioral health models.

---

5. Strengths and Limitations

Strengths:

- High reliability and diagnostic accuracy
- Validated short forms (ASRS-6, ASRS-5) appropriate for clinical use
- Cross-culturally adapted and tested
- Easy to administer, score, and interpret

Limitations:

- Risk of over-identification without structured diagnostic follow-up
- Sensitivity lower in severe psychiatric comorbidities (e.g., heavy SUD)
- Primarily validated in adult populations; less studied in elderly or adolescent transition cases
- Some minor cultural nuances in self-reporting behaviors across populations

---

6. Recommendations for Practice and Research

Practice:

- Screen all adults with attention, executive dysfunction, or emotional regulation concerns using ASRS-6 or ASRS-5
- Integrate with clinician-administered confirmation (e.g., DIVA-5) before diagnosing ADHD
- Provide staff training on interpreting scores and discussing findings sensitively

Research:

- Validate ASRS-5 further in underserved and multicultural populations
- Study longitudinal responsiveness of ASRS scores to treatment (psychotherapy, medication)
- Explore digital adaptations for asynchronous telehealth screening

---

Digital Repository Format

Measure: Adult ADHD Self-Report Scale (ASRS v1.1, ASRS-5)

Type: Symptom Screening + Diagnostic Support

Languages: Multilingual (English, Turkish, German, Afrikaans, Spanish, others)

Validated Populations: U.S. general, psychiatric outpatients, primary care, substance use disorder populations, South African, Turkish, German adults

Cutoffs:

- ASRS-6: ≥4 positive items
- ASRS-18: ≥30 total score

- ASRS-5 (DSM-5 weighted): Algorithmic cutoff (machine learning-derived)

Psychometrics:

– Sensitivity: 83–91% (varies by version)
– Specificity: 80–96%
– AUC: 0.90–0.94
– Internal consistency α = 0.86–0.88

Setting: Primary care, psychiatry, occupational health, telehealth, community mental health
Use Case: Early identification of probable adult ADHD, diagnostic triage, treatment monitoring

Alcohol Use Disorder Identification Test- (AUDIT-C/S)[10-19]

1. Objective

To evaluate the psychometric validity, reliability, and pragmatic utility of the full AUDIT and its shorter versions—especially the AUDIT-C—for identifying alcohol misuse and alcohol use disorders (AUDs) in diverse primary care populations across global and U.S. settings.

---

2. Measure Evaluated

Alcohol Use Disorders Identification Test (AUDIT) – A 10-item screening tool developed by WHO to assess alcohol consumption, drinking behaviors, and alcohol-related problems.
AUDIT-C – A 3-item abbreviated version focusing on alcohol consumption, increasingly used for its brevity in primary care.
Alcohol Symptom Checklist – An 11-item DSM-5-based tool used in electronic health records to support AUD diagnosis and clinical decision-making.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **AUDIT:** Excellent sensitivity (up to 0.95) and specificity (up to 0.94) across cultures and genders. Internal consistency (Cronbach's α = 0.83–0.94).
- **AUDIT-C:** Strong validity (AUC 0.88–0.94), high test-retest reliability (ICC = 0.87–0.95), and consistent performance across ethnic/racial groups.
- **Alcohol Symptom Checklist:** Supported one-dimensionality and measured AUD severity consistently across demographics.

Pragmatic Properties:

- **Acceptability:** Widely accepted and feasible in real-world practice; minimal burden.
- **Feasibility:** AUDIT-C and checklists are efficient (<3 min) and usable in EHRs or patient portals.
- **Interpretability:** Clear scoring cutoffs (AUDIT-C ≥4 for men, ≥3 for women; adjusted for populations).

- **Equity:** Validated across diverse groups (White, Hispanic, African American, AI/AN, multiracial, Mozambican, Japanese, and online users).
- **Sustainability:** Routinely implemented in U.S. primary care, including Kaiser Permanente and VA systems.

---

4. Application to Integrated Primary Care

The AUDIT-C and Alcohol Symptom Checklist support behavioral health integration through routine alcohol screening, identification of AUD, and decision support within EHRs. Especially effective for screening in Federally Qualified Health Centers (FQHCs) and clinics serving diverse, underserved populations.

---

5. Strengths and Limitations

Strengths:

- Valid across cultures and formats (in-clinic and digital)
- High reliability and clinical relevance
- Can support shared decision-making in care

Limitations:

- Slight demographic variance in performance (e.g., AI/AN)
- AUDIT may be too long for fast-paced settings
- AUDIT-C is less sensitive to identifying severe AUD without follow-up tools

---

6. Recommendations for Practice and Research

Practice:

- Use AUDIT-C for universal alcohol misuse screening in primary care.
- Implement follow-up symptom checklists to assess severity.
- Embed into EHRs and patient portals for streamlined workflows.

Research:

- Study optimal cutoffs by race/ethnicity, gender, and care settings.
- Develop and validate culturally adapted tools in underserved U.S. populations.
- Explore longitudinal utility for treatment monitoring.

Digital Repository Format

Measure: AUDIT / AUDIT-C / Alcohol Symptom Checklist
Type: Screening + Diagnostic Support
Languages: Multilingual (English, Spanish, Japanese, Portuguese, etc.)
Validated Populations: U.S. general, AI/AN, Hispanic, multiracial, Mozambican, Japanese
Cutoffs: AUDIT-C ≥4 (men), ≥3 (women); Full AUDIT ≥8
Psychometrics: High sensitivity (0.78–0.95), specificity (0.74–0.94), reliability ($\alpha > 0.8$)

Setting: Integrated primary care, behavioral health, EHR-enabled practices
Use Case: Early identification and management of alcohol misuse and AUD

Behavioral Health Measure (BHM-20)[20-22]

1. Objective

To assess the reliability, validity, clinical utility, and outcome responsiveness of the Behavioral Health Measure-20 (BHM-20) for tracking patient symptoms, functioning, and well-being in integrated primary care settings, emphasizing brief behavioral health interventions.

---

2. Measure Evaluated

Behavioral Health Measure-20 (BHM-20) – A 20-item, patient self-report tool designed to measure three domains of mental health: well-being (subjective distress and life satisfaction), symptoms (depression, anxiety, etc.), and life functioning (social, occupational, daily activities). It also provides a Global Mental Health (GMH) index.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- Internal consistency:
  - Total scale (GMH): α = 0.90
  - Well-being: α = 0.84
  - Symptoms: α = 0.88
  - Life Functioning: α = 0.79
- Test-retest reliability: ICC not directly reported, but stability observed over short intervals
- Construct validity: Strong support aligning with the phase model of psychotherapy
- Clinical sensitivity: Detects change across domains after brief interventions (e.g., 2–3 behavioral health consultant [BHC] visits)
- Convergent validity: Aligns with psychotherapy phase change models and outcome monitoring practices

Pragmatic Properties:

- **Acceptability:** Very high; brief (completed in <5 minutes); easily integrated into clinic workflows
- **Feasibility:** Paper-and-pencil administration or electronic; designed for routine clinical use with minimal disruption
- **Interpretability:** Clinical ranges defined (Healthy, At-Risk, Distressed) per Jacobson-Truax method
- **Equity:** Used across diverse adult patient populations in military, community health, and family medicine settings; further validation in broader civilian populations recommended

• **Sustainability:** Designed for ongoing tracking of patient progress within integrated primary care and collaborative behavioral health models

---

4. Application to Integrated Primary Care

The BHM-20 is a core outcome tool for monitoring patient progress during integrated behavioral health interventions. It enables real-time assessment of subjective well-being, symptom severity, and functional capacity, allowing PCPs and behavioral health consultants (BHCs) to rapidly evaluate treatment efficacy, adjust care plans, and document behavioral health outcomes within primary care workflows.

---

5. Strengths and Limitations

Strengths:

• Captures simultaneous improvements in well-being, symptoms, and life functioning
• Sensitive to change after as few as two brief BHC sessions
• Provides actionable insights into whether patients are moving toward recovery or need referral to specialty mental health care
• Very brief and non-burdensome, ideal for fast-paced settings

Limitations:

• Limited racial/ethnic-specific normative data
• No direct diagnostic capability (designed for monitoring rather than diagnosing)
• Not specifically validated for pediatric populations
• Limited longitudinal predictive studies beyond initial episodes of care

---

6. Recommendations for Practice and Research

Practice:

• Administer BHM-20 at each BHC or behavioral health visit to track recovery trajectories
• Use clinical range (Healthy, At-Risk, Distressed) designations to guide stepped-care decisions
• Integrate BHM-20 scores into EHR dashboards to enhance population health management

Research:

• Further validate BHM-20 across diverse racial, ethnic, and socioeconomic groups
• Compare performance against PHQ-9, GAD-7, and SF-12 for benchmarking
• Study BHM-20 responsiveness to various brief therapeutic modalities (e.g., ACT, CBT, mindfulness interventions)
• Explore predictive value for long-term patient outcomes like relapse, hospitalization, or chronic disease exacerbation

---

Digital Repository Format

Measure: Behavioral Health Measure-20 (BHM-20)
Type: Symptom, Functioning, and Well-being Tracking
Languages: English (translation adaptations needed for broader use)
Validated Populations: Adults in military, integrated primary care, and behavioral health consultant settings
Cutoffs: Clinical ranges: Healthy, At-Risk, Distressed (domain-specific scoring thresholds available)
Psychometrics: α = 0.79–0.90; strong clinical sensitivity; aligns with psychotherapy phase models
Setting: Integrated primary care, family medicine clinics, behavioral health consultation models
Use Case: Monitoring mental health outcomes during brief primary care-based interventions; guiding stepped-care or referral decisions

Brief Addiction Monitor (BAM/BAM-R)[23-27]

1. Objective
To evaluate the psychometric validity, longitudinal reliability, and pragmatic clinical utility of the Brief Addiction Monitor-Revised (BAM-R)—a 17-item tool used to monitor substance use behaviors, recovery supports, and risk factors among individuals with substance use disorders (SUD), particularly in Veteran and non-Veteran outpatient populations receiving integrated behavioral health services.

---

2. Measure Evaluated
Brief Addiction Monitor-Revised (BAM-R) – A multidimensional, 17-item self-report or clinician-administered tool assessing alcohol and drug use, stressors, protective factors, recovery-oriented behaviors, and risks associated with relapse. Initially validated in Veterans Affairs (VA) SUD clinics, BAM-R is being explored in broader non-VA outpatient populations and integrated care settings.
The tool includes both continuous and Likert-type item responses and is used to track recovery progress and guide treatment adjustments. A shortened version with 5 clinically sensitive items has been proposed to improve feasibility in primary care settings.

---

3. PAPERS Framework Evaluation
Psychometric Properties:
• Confirmed 4-factor structure (Alcohol Use, Stressors, Risk, Stability) in large Veteran samples.
• Excellent internal consistency for Alcohol Use subscale (α = .92–.96).
• Questionable-to-poor internal consistency for Risk, Stress, and Stability (α = .35–.68).
• Sensitivity to change demonstrated in key items: Craving, Mood, Drug Use, Heavy Alcohol Use,

and Self-Help participation.
• Longitudinal measurement invariance supported in large VA datasets.

Pragmatic Properties:

• **Acceptability:** Widely used within VA and adaptable for non-specialty outpatient care.
• **Feasibility:** Administered in <5 minutes; a shortened version with 5 items improves efficiency.
• **Interpretability:** No fixed diagnostic cutoffs; score changes used for progress tracking.
• **Equity:** Tested across race, gender, and mandated vs. voluntary referral groups; factor loadings varied slightly by race.
• **Sustainability:** Mandated use in all VA SUD clinics; electronic integration supported in EHR systems.

---

4. Application to Integrated Primary Care

The BAM-R supports measurement-based care (MBC) in integrated behavioral health models by offering a structured tool to monitor treatment progress and risk factors in individuals with SUDs. While originally used in specialty addiction care, BAM-R shows growing relevance in primary care settings—particularly for populations with co-occurring mental health and substance use needs. The 5-item abbreviated BAM-R may be ideal for fast-paced clinical environments and aligns with stepped care and SBIRT (Screening, Brief Intervention, and Referral to Treatment) workflows.

---

5. Strengths and Limitations

Strengths:

• Proven feasibility and sensitivity to change in real-world addiction treatment.
• Supports longitudinal monitoring in both specialty and primary care settings.
• Recognized by the Kennedy Forum, Joint Commission, and VA MBC initiative.

Limitations:

• Only the Alcohol Use subscale has high reliability.
• Mixed validity across non-Veteran and racially diverse populations.
• Lacks standard clinical cutoffs; interpretation requires familiarity with item-level trends.
• Stability and Protective Factors subscales show poor psychometric strength.

---

6. Recommendations for Practice and Research

Practice:

• Use BAM-R to support longitudinal tracking in SUD recovery programs.
• Incorporate the abbreviated 5-item BAM-R in integrated primary care to reduce burden.
• Train clinicians to interpret item-level changes to guide treatment plans.

Research:

• Validate BAM-R short-form in diverse, non-Veteran populations.

• Explore psychometric refinement of low-reliability subscales (Risk, Stability).
• Investigate cultural adaptations and equity in scoring across racial and ethnic groups.
• Evaluate predictive validity for treatment retention, relapse, and mortality outcomes.

Digital Repository Format
Measure: Brief Addiction Monitor-Revised (BAM-R)
Type: Monitoring + Measurement-Based Care Support
Languages: English
Validated Populations: U.S. Veterans, Non-veteran outpatient SUD patients
Cutoffs: No fixed clinical cutoffs; item trends used to track recovery progress
Psychometrics: High reliability for Alcohol Use (α ~.94); low for others; sensitive to change for key items
Setting: Integrated primary care, VA, and non-VA outpatient addiction treatment, EHR-enabled systems
Use Case: Progress monitoring, relapse risk detection, treatment adjustment in SUD care

Child and Adolescent Trauma Screen (CATS/CATS-2)[28-37]

1. Objective
To evaluate the psychometric validity, clinical applicability, and cultural adaptability of the Child and Adolescent Trauma Screen (CATS and CATS-2) for screening trauma exposure, posttraumatic stress symptoms (PTSS), and functional impairment in children and adolescents according to DSM-5 and ICD-11 criteria.

---

2. Measure Evaluated
CATS (Original) – A 20-item PTSD symptom screening tool plus trauma exposure checklist, aligned with DSM-5, available in self-report and caregiver-report versions.
CATS-2 – An updated version measuring both DSM-5 and ICD-11 PTSD and Complex PTSD (CPTSD), validated internationally for cross-system assessment.

---

3. PAPERS Framework Evaluation
Psychometric Properties:
• Internal consistency:
  – CATS self-report: α = 0.88–0.94
  – CATS caregiver-report: α = 0.84–0.91
• Test–retest reliability: ICC = 0.57–0.68 (moderate stability)
• Factor structure:
  – Four-factor DSM-5 PTSD model (re-experiencing, avoidance, negative mood/cognition,

hyperarousal) confirmed across samples
  – ICD-11 PTSD and CPTSD are distinguishable via CATS-2
• Criterion validity:
  – Strong correlations with depression, anxiety, and trauma measures (r = 0.62–0.82)
  – CATS scores discriminated between clinical vs. non-clinical groups
• Diagnostic accuracy (CATS-2 cutoffs):
  – DSM-5 PTSD: ≥21 (screening), ≥25 (diagnosis)
  – ICD-11 PTSD: ≥7 (screening), ≥9 (diagnosis)

Pragmatic Properties:

• **Acceptability:** Widely used across clinical, research, and humanitarian contexts
• **Feasibility:** Brief, child- and caregiver-friendly; available at no cost
• **Interpretability:** Clear symptom cluster and total scores; screening and diagnostic cutoffs available
• **Equity:** Validated internationally (U.S., Germany, Norway, Sweden, Ukraine, Turkey, refugee populations)
• **Sustainability:** Ongoing use in trauma-focused clinical trials, refugee assessments, and community screenings

---

4. Application to Integrated Primary Care

The CATS and CATS-2 are ideal for early identification of trauma exposure and PTSD symptoms in pediatric primary care, behavioral health integration, school-based health centers, and refugee health programs.

• Brief and developmentally appropriate for children as young as 7 (self-report) and 3 (caregiver-report)
• Helps triage patients for trauma-focused cognitive behavioral therapy (TF-CBT) or other evidence-based treatments
• Useful for outcome tracking in stepped care models and culturally sensitive care approaches

---

5. Strengths and Limitations

Strengths:

• Strong cross-cultural validation and multi-language availability
• Updated to DSM-5 and ICD-11 criteria, allowing flexible diagnostic alignment
• Parallel caregiver and child versions for comprehensive assessment
• Free for clinical and research use

Limitations:

• Some variability in preschool form psychometrics (needs more research)
• Not designed to diagnose all trauma-related disorders (e.g., dissociative subtypes)

• Possible reliance on caregiver-report in young or illiterate populations introduces bias
• Some minor cross-cultural scoring adjustments may be necessary

---

6. Recommendations for Practice and Research

Practice:

• Routinely screen for trauma and PTSD symptoms in pediatric primary care, especially among refugees and underserved populations
• Combine self- and caregiver reports for the highest accuracy when possible
• Use CATS-2 for streamlined DSM-5 and ICD-11 diagnosis considerations

Research:

• Further validate preschool CATS versions across languages and cultures
• Study longitudinal responsiveness to trauma interventions (e.g., TF-CBT outcomes)
• Develop digital CATS administration platforms for school and telehealth settings
• Examine CATS use in complex trauma and dissociation screening

---

Digital Repository Format

Measure: Child and Adolescent Trauma Screen (CATS, CATS-2)
Type: Trauma Exposure and PTSD Symptom Screening
Languages: English, German, Norwegian, Turkish, Ukrainian, Swedish, Spanish (others in progress)
Validated Populations: U.S. general, refugee youth (Germany, Sweden), Turkish preschoolers, Ukrainian war-affected children, trauma-exposed adolescents
Cutoffs:
  – CATS-2 DSM-5 PTSD: ≥21 (screening), ≥25 (diagnostic)
  – ICD-11 PTSD: ≥7 (screening), ≥9 (diagnostic)
Psychometrics:
  – Internal consistency α = 0.84–0.94
  – Strong convergent validity (r = 0.62–0.82)
  – Confirmed factor structure matching DSM-5 and ICD-11 models
Setting: Primary care, pediatric behavioral health, refugee health services, school-based mental health
Use Case: Early identification and monitoring of PTSD symptoms, triage for trauma-focused interventions, research on trauma epidemiology in youth populations.

Child Outcome Rating Scale (CORS)[37-38]

Child Outcome Rating Scale (CORS)

1. Objective

To evaluate the psychometric robustness and pragmatic utility of the Child Outcome Rating Scale (CORS) as a brief, child self-report measure of psychosocial functioning, distinct from psychiatric symptoms, and its applicability for community and clinical populations, particularly in integrated health settings.

---

2. Measure Evaluated

Child Outcome Rating Scale (CORS) – A 4-item visual analogue scale (VAS) self-report tool for children aged 6–12 years, designed to assess psychosocial functioning across domains including self, family, school, and overall wellbeing.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **Reliability:** Cronbach's α = .84–.87; McDonald's ω = .85, indicating strong internal consistency.
- **Validity:**
  - Concurrent validity with the Youth Outcome Questionnaire (YOQ): r = .61.
  - Construct validity supported through significant correlations with emotional and behavioral health measures and resilience factors.
- **Sensitivity:** ROC curve AUC = 0.78 (SE= .01, 95% CI: .76–.79); clinical cutoff scores identified at 28 (for adolescents 10–15) and 32 (for younger children 6–12).
- **Differential Item Functioning (DIF):** Minimal bias across gender, socio-economic status, SEN status, language, and age.

Pragmatic Properties:

- **Acceptability:** Highly accepted due to brevity, simplicity, and relevance to both clinicians and young clients.
- **Feasibility:** Takes under 5 minutes to complete; minimal training required; easily integrated into clinical workflows.
- **Interpretability:** Simple visual scoring (0–40 scale); clear clinical thresholds aid quick decision-making.
- **Equity:** Demonstrated validity across diverse socioeconomic and educational backgrounds; feasible for multilingual populations with basic reading comprehension.
- **Sustainability:** Requires a license. The measure is free to download and use in 36 languages for individual use, but users must agree to the license and register at https://betteroutcomesnow.com/download-ors-srs/. There is a fee for group licenses.

---

4. Application to Integrated Primary Care

The CORS offers a scalable, child-centered tool for integrated behavioral health teams to quickly assess psychosocial functioning alongside traditional symptom measures. It supports

collaborative care, facilitates shared decision-making, and monitors progress within schools, primary care, and outpatient behavioral health services, including underserved and multicultural U.S. populations.

---

5. Strengths and Limitations

Strengths:

- Very brief (4 items) yet psychometrically robust.
- Strong alignment with value-based care emphasizing patient voice.
- Useful for longitudinal tracking of psychosocial functioning across care episodes.

Limitations:

- Less specific than longer diagnostic tools; does not assess psychopathology severity.
- Slight variability in measurement across subgroups; statistical adjustments recommended for research comparisons.
- Limited initial validation in non-English languages, though growing adoption internationally.

---

6. Recommendations for Practice and Research

Practice:

- Integrate CORS as a routine outcome measure in pediatric primary care and behavioral health programs.
- Use alongside disorder-specific tools to provide a holistic view of youth wellbeing.
- Train staff on engaging youth in interpreting and using CORS results collaboratively.

Research:

- Further validation needed across broader cultural and linguistic populations.
- Explore digital/electronic administration formats for greater reach.
- Study longitudinal sensitivity to detect change over treatment courses in real-world integrated settings.

Digital Repository Format

Measure: Child Outcome Rating Scale (CORS)

Type: Patient-Reported Outcome Measure (PROM); Psychosocial functioning assessment

Languages: English (validated); translated to 36 languages and although widely used, are not validated.

Validated Populations: U.S., UK, multilingual school-based samples, outpatient pediatric behavioral health populations

Cutoffs:

- Children 6–12 years: Clinical cutoff <28

- Adolescents 10–15 years: Clinical cutoff <28
  Psychometrics:
- Internal consistency (α = .84–.87)
- ROC AUC = 0.78; Sensitivity = .73, Specificity = .70
- Reliable Change Index: 6 points
  Setting: Integrated primary care, pediatric behavioral health, schools, community health centers
  Use Case: Rapid assessment of psychosocial functioning to monitor outcomes and inform integrated behavioral health interventions for youth

Collaborative Assessment and Management of Suicidality (CAMS)[39-48]

1. Objective

To assess the clinical effectiveness, psychometric reliability, feasibility, and patient-centered benefits of the Collaborative Assessment and Management of Suicidality (CAMS) for reducing suicidal ideation, psychological distress, and suicide attempts across outpatient, inpatient, and transitional care settings.

---

2. Measure Evaluated

Collaborative Assessment and Management of Suicidality (CAMS) – A therapeutic framework using the **Suicide Status Form (SSF)** for assessment, treatment planning, tracking, and clinical documentation of suicidal risk. CAMS targets patient-identified "suicidal drivers" through collaborative, individualized intervention, integrating therapeutic techniques (e.g., CBT, DBT elements) without allegiance to a single psychotherapy school.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

• Internal consistency of SSF Core Assessment: Acceptable; domain-specific reliability supported
• Validity: CAMS significantly reduces suicidal ideation (effect sizes d = 0.25–0.88) compared to treatment as usual (TAU)
• Sensitivity to change: CAMS produces large effect size reductions in suicidal ideation and symptom distress within as few as 6–8 sessions
• Therapeutic alliance: CAMS consistently rated higher than TAU in patient-reported alliance measures
• Test-retest reliability of tracking tools (Suicide Status Form - Tracking) acceptable across multiple settings

Pragmatic Properties:

• **Acceptability:** Highly acceptable to both patients and providers; improves satisfaction and treatment engagement

• **Feasibility:** Implementable in outpatient, inpatient, and "next-day appointment" crisis models; manageable training requirements

• **Interpretability:** Standardized SSF scoring allows for tracking suicide risk, psychological pain, hopelessness, and therapeutic progress

• **Equity:** Successfully tested across veterans, college students, civilian inpatient/outpatient settings, and in Europe and the U.S.

• **Sustainability:** Increasing integration into national crisis intervention models, especially post-hospitalization settings

---

4. Application to Integrated Primary Care

While primarily tested in mental health, CAMS offers clear applications to integrated primary care models for suicide prevention:

• Facilitates suicide-specific assessment beyond generic depression screening

• Supports stepped care transitions (e.g., from ED to outpatient behavioral health)

• Aligns with collaborative care principles—shared goal setting, patient-centered planning, real-time tracking of symptoms

• Useful for primary care-based behavioral health consultants managing suicide risk without extensive specialty referral delays

---

5. Strengths and Limitations

Strengths:

• Targets suicide risk directly rather than mental illness symptoms alone

• Strong focus on patient collaboration, therapeutic alliance, and individualized drivers

• Demonstrates sustained reductions in suicidal ideation and distress up to 12 months post-intervention

• Adaptable across diagnoses, settings, and provider types

Limitations:

• CAMS training still requires time investment (although less intensive than DBT)

• Some RCTs show comparable efficacy to enhanced TAU in certain settings (e.g., immediately post-hospitalization)

• Sample sizes in some studies are small; larger multi-site trials are ongoing

• Suicidal behavior outcomes (e.g., suicide attempts) show mixed or small effects

---

6. Recommendations for Practice and Research

Practice:

• Train behavioral health consultants and integrated care teams on CAMS basics and SSF completion

• Use CAMS for stepped-care transitions post-ED discharge or psychiatric hospitalization

• Implement CAMS for patients presenting with acute suicidal ideation but manageable risk profiles suitable for outpatient management

Research:

• Conduct large-scale pragmatic trials in primary care and FQHC settings

• Study CAMS integration with digital health tools (tele-CAMS, SSF electronic platforms)

• Investigate CAMS for culturally specific suicide prevention adaptations (e.g., tribal health, rural health)

• Examine cost-effectiveness compared to standard mental health referrals over time

---

Digital Repository Format

Measure: Collaborative Assessment and Management of Suicidality (CAMS)

Type: Suicide-specific Assessment, Treatment Planning, and Outcome Monitoring

Languages: English (formal translations underway)

Validated Populations: U.S. veterans, civilians (college, outpatient, inpatient), European psychiatric populations, crisis settings

Cutoffs: N/A; clinical improvement tracked via SSF scores across psychological pain, hopelessness, self-hate, and overall risk

Psychometrics:

- Effect sizes d = 0.25–0.88 for suicidal ideation reduction
- Therapeutic alliance consistently stronger than TAU
- SSF domain internal reliability acceptable to high

Setting: Outpatient mental health, integrated primary care, crisis clinics, inpatient psychiatry

Use Case: Suicide risk assessment, suicide-specific treatment planning, symptom tracking, recovery monitoring, stepped-care stabilization

Columbia-Suicide Severity Rating Scale (C-SSRS)[49-57]

1. **Objective**
   To evaluate the psychometric validity, reliability, and pragmatic utility of the Columbia-Suicide Severity Rating Scale (C-SSRS) for assessing the severity of suicidal ideation and behavior across diverse primary care and clinical research populations globally and within the U.S.
2. **Measure Evaluated**
   Columbia-Suicide Severity Rating Scale (C-SSRS) – A semi-structured tool developed to

standardize the assessment of suicidal ideation and behavior severity, intensity, and lethality across clinical and research settings.

3. **PAPERS Framework Evaluation**

Psychometric Properties:

• **Reliability:** Internal consistency, Cronbach's α ranged from 0.73 to 0.95 across populations.

• **Validity:** Demonstrated good convergent and divergent validity with related measures (e.g., SSI).

• **Sensitivity/Specificity:** High sensitivity (up to 0.95) and specificity (up to 0.95) for detecting suicidal behaviors.

• **Predictive Validity:** Baseline severity and intensity scores predicted future suicide attempts and behaviors over time.

Pragmatic Properties:

• **Acceptability:** Widely accepted across settings (primary care, psychiatric, emergency) and age groups.

• **Feasibility:** Can be clinician-administered or self-reported; electronic versions (e-CSSRS) available.

• **Interpretability:** Clear behavioral categories (e.g., actual attempt, interrupted attempt) enhance utility for diverse clinical teams.

• **Equity:** Validated in multiple languages and across culturally diverse samples, including the U.S., Mexico, China, Turkey, Lebanon, and others.

• **Sustainability:** Routinely integrated into healthcare and research protocols for suicide prevention globally.

4. **Application to Integrated Primary Care**

The C-SSRS enables standardized suicide risk assessments within integrated primary care models, facilitating early identification and intervention. Especially valuable in Federally Qualified Health Centers (FQHCs), VA medical centers, and practices serving socioeconomically diverse, high-risk populations.

5. **Strengths and Limitations**

Strengths:

• Strong psychometric robustness across diverse demographic and clinical populations.

• Practical adaptability for primary care and behavioral health integration.

• Predicts suicide attempts better than some other leading scales (e.g., SSI).

Limitations:

• Requires training for optimal administration, especially for nuanced suicidal behavior definitions.

• Electronic and online versions (e.g., e-CSSRS) have limited validation in some low-resource settings.

• Slight variance in performance when adapted for adolescent versus adult samples.

6. **Recommendations for Practice and Research**

Practice:

• Implement the C-SSRS universally for suicide risk screening in integrated behavioral health settings.

• Train primary care and behavioral health clinicians to use the full spectrum of ideation, behavior, and lethality subscales.

• Electronic versions (eC-SSRS) are used to streamline documentation and monitoring.

Research:

• Conduct more cross-cultural validation studies in underrepresented U.S. racial/ethnic groups (e.g., AI/AN, Black/African American).

• Examine long-term outcomes of C-SSRS screening integrated into population health management strategies.

• Validate adaptations specifically tailored for pediatric primary care settings.

Digital Repository Format

Measure: Columbia-Suicide Severity Rating Scale (C-SSRS)

Type: Suicide Risk Screening + Severity and Behavior Classification

Languages: Multilingual (English, Spanish, Chinese, Turkish, Arabic, etc.)

Validated Populations: U.S. general, Mexican, Turkish adolescents, Chinese adults with depression, Lebanese adults, juvenile justice-involved youth, Veterans

Cutoffs: Behavioral classification system (e.g., actual attempt, aborted attempt); no single numeric cutoff

Psychometrics: High reliability (α = 0.73–0.95), sensitivity (up to 0.95), specificity (up to 0.95), strong predictive validity for future suicide attempts

Setting: Primary care, emergency departments, psychiatric care, behavioral health, community settings

Use Case: Standardized identification, tracking, and management of suicidal ideation and behavior risk across clinical and research environments.

CRAFFT 2.0[58-67]

1. Objective

To evaluate the psychometric validity, reliability, and pragmatic utility of the CRAFFT 2.0 screening tool for identifying substance use risk (alcohol, cannabis, and other drugs) among adolescents in primary care and diverse community settings.

---

2. Measure Evaluated

CRAFFT 2.0 – A 9-item substance use screening tool for adolescents, including three opening questions about recent use and six core yes/no questions (Car, Relax, Alone, Forget, Family/Friends, Trouble). Developed to detect substance use risk and inform need for further assessment or brief intervention.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **Sensitivity/Specificity**:
  - CRAFFT 2.0 (cutoff ≥2) showed high sensitivity (0.80–0.98) and specificity (0.73–0.94) for identifying alcohol, cannabis, and substance use disorders.
  - For heavy cannabis use detection specifically, sensitivity was 0.92 and specificity was 0.75 when using a revised "Car" item.
- **Internal Consistency**:
  - Cronbach's alpha ranged from 0.74 to 0.81 across samples.
- **Area Under Curve (AUC)**:
  - Ranged from 0.88 to 0.90 across diverse populations.

Pragmatic Properties:

- **Acceptability**:
  - Highly acceptable to adolescents and providers; preferred computer self-administration for greater disclosure.
- **Feasibility**:
  - Brief (completion time <2 minutes); feasible in high-volume primary care and school settings.
- **Interpretability**:
  - Clear cutoffs (≥2 positive responses) for need for further assessment or intervention.
- **Equity**:
  - Demonstrated good performance across diverse racial/ethnic groups (African American, Latino, White) and socioeconomic statuses, with potential need for adjusting thresholds for minority youth to reduce disparities.
- **Sustainability**:
  - Widely recommended by the American Academy of Pediatrics and integrated into SBIRT models.

---

4. Application to Integrated Primary Care

The CRAFFT 2.0 enables early identification of adolescent substance use within pediatric and adolescent primary care settings. Its brief format and strong validity make it ideal for routine screening as part of preventive services, behavioral health integration, or risk assessment

workflows. Use in electronic health records and self-administered tablet versions improves efficiency and confidentiality, supporting health equity goals.

---

5. Strengths and Limitations

Strengths:

- Brief, easy to administer.
- High sensitivity and specificity for multiple substances.
- Good cross-cultural validity (e.g., validated in American Indian/Alaska Native, Hispanic, European, and Asian adolescents).
- Useful across settings: clinics, schools, community health centers.

Limitations:

- Slightly lower specificity at lower cutoff points in minority populations (tradeoff for higher sensitivity).
- Possible social desirability bias in clinician-administered formats.
- Limited predictive data for cannabis use disorder specifically compared to alcohol use.

---

6. Recommendations for Practice and Research

Practice:

- Administer CRAFFT 2.0 universally to adolescents aged 12–21 during primary care visits.
- Prefer computer self-administration when feasible to improve honesty and reduce bias.
- Use cutoff ≥2 for general populations; consider lower threshold (≥1) for high-risk racial/ethnic groups with further assessment to avoid disparities.

Research:

- Conduct further studies validating CRAFFT 2.0 thresholds for specific racial/ethnic groups and low-SES populations.
- Explore optimal ways to combine CRAFFT screening with EHR alerts and telehealth platforms.
- Investigate longitudinal use of CRAFFT 2.0 in treatment tracking and outcomes monitoring.

Digital Repository Format

Measure: CRAFFT 2.0

Type: Screening (Substance Use Risk)

Languages: English, Spanish, and validated in multiple other languages (Norwegian, Persian, Spanish)

Validated Populations: U.S. adolescents (White, African American, Hispanic, AI/AN), Spanish, Argentine, Native American adolescents

Cutoffs: Score ≥2 for positive screen (general); Score ≥1 for minority youth or high-risk

populations (consider with caution)
Psychometrics: Sensitivity 0.80–0.98; Specificity 0.73–0.94; Cronbach's α ~0.74–0.81; AUC 0.88–0.90
Setting: Pediatric and adolescent primary care, school-based health centers, community health settings, integrated behavioral health programs
Use Case: Early detection of alcohol and drug risk among adolescents to prompt early intervention, brief counseling, or referral

Depression Anxiety Stress Scale (DASS-21)[68-77]

1. Objective

To evaluate the psychometric validity, cross-cultural applicability, and clinical utility of the DASS-21 as a brief self-report measure for assessing symptoms of depression, anxiety, and stress in clinical and non-clinical populations globally, with emphasis on its relevance to U.S. integrated primary care settings.

---

2. Measure Evaluated

DASS-21 – A 21-item, short-form version of the original DASS-42, developed by Lovibond & Lovibond (1995), comprising three 7-item subscales: Depression, Anxiety, and Stress. Each item is rated on a 4-point Likert scale reflecting the frequency of symptoms experienced over the past week.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **Internal Consistency**: Cronbach's α typically ranges from:
    - Depression: 0.91–0.97
    - Anxiety: 0.81–0.92
    - Stress: 0.88–0.95.
- **Test-Retest Reliability**: ICCs of 0.70–0.81 over 2 weeks in clinical samples and 0.39–0.46 over 6 months in non-clinical samples.

- **Construct Validity**: Supported by high correlations with BDI (Depression), BAI (Anxiety), and factor analyses confirming a consistent 3-factor or bifactor structure in clinical and non-clinical populations across multiple languages and cultures.
- **Discriminant Validity**: Subscales differentiate between anxiety, depression, and general distress, though some overlap exists with the Stress scale tapping general negative affectivity (NA).

Pragmatic Properties:

- **Acceptability**: Freely available; widely accepted by patients and providers; easy to administer and score.
- **Feasibility**: Self-administered in <5 minutes; appropriate for routine use in primary care and behavioral health.
- **Interpretability**: Subscale scores classified as normal, mild, moderate, severe, or extremely severe using normative percentile cutoffs.
- **Equity**: Validated across diverse cultures and translated into 20+ languages (e.g., Persian, Chinese, Nepali, Swedish, Spanish, Russian).
- **Sustainability**: Supported by a large body of global normative data (>70,000 individuals); endorsed in public mental health and research protocols.

---

4. Application to Integrated Primary Care

DASS-21 is ideal for integrated behavioral health settings, enabling rapid screening and monitoring of emotional distress in patients. Particularly useful for identifying comorbid mental health concerns in primary care, chronic illness management, geriatrics, and telehealth models where symptom burden must be routinely assessed. Its use can inform stepped care approaches and triage decisions.

---

5. Strengths and Limitations

Strengths:

- Brief and freely available; strong international validation
- Distinguishes depression, anxiety, and stress symptoms with high reliability

- Effective in primary care, psychiatric, and community populations, including older adults

Limitations:

- Stress subscale may reflect general negative affectivity (NA) rather than discrete stress symptoms
- Some inter-scale correlation may reduce distinctiveness in complex comorbid presentations
- Limited diagnostic precision for DSM disorders; best used as a screening and symptom tracking tool

---

6. Recommendations for Practice and Research

Practice:

- Use DASS-21 for routine mental health screening in primary care, especially in integrated care or stepped care models.
- Repeat assessments over time to evaluate intervention outcomes or worsening symptoms.
- Combine with clinical interviews for full diagnostic evaluation, particularly for older adults or somatically focused patients.

Research:

- Continue validating cutoffs and factorial structure across age, culture, and diagnostic groups.
- Explore predictive validity for clinical outcomes (e.g., hospitalization, suicide risk).
- Examine use in digital and telehealth settings to support scalable mental health access.

Digital Repository Format

Measure: DASS-21 (Depression Anxiety Stress Scale – 21 Item Version)
Type: Symptom severity scale (Depression, Anxiety, Stress)
Languages: Multilingual (validated in English, Persian, Chinese, Swedish, Spanish, Nepali, Russian, Turkish, etc.)

Validated Populations: U.S. general population, healthcare workers, students, primary care, older adults, global clinical and community samples
Cutoffs (Subscale Score Ranges for Severity):

- Normal: 0–4 (Dep), 0–3 (Anx), 0–7 (Str)
- Mild: 5–6 (Dep), 4–5 (Anx), 8–9 (Str)
- Moderate: 7–10 (Dep), 6–7 (Anx), 10–12 (Str)
- Severe: 11–13 (Dep), 8–9 (Anx), 13–16 (Str)
- Extremely Severe: 14+ (Dep), 10+ (Anx), 17+ (Str)
  Psychometrics: Cronbach's α = .81–.97; ICCs = .39–.81; 3-factor model consistently supported
  Setting: Integrated primary care, mental health clinics, schools, community centers, geriatrics
  Use Case: Brief screening and monitoring of depression, anxiety, and stress symptoms in diverse populations

Drug Abuse Screening Test (DAST-10)[78-87]

1. Objective
To evaluate the psychometric and pragmatic properties of the DAST and its short forms (DAST-10 and DAST-20) across various cultural and clinical populations for the purpose of drug use disorder screening.

---

2. Measure Evaluated
The Drug Abuse Screening Test (DAST) is a self-report instrument comprising 28, 20, and 10 items. It assesses problems related to drug abuse and is designed for both clinical and research settings. Responses are dichotomous (yes/no), and higher scores indicate greater severity of drug use-related issues.

---

3. PAPERS Framework Evaluation

- **Psychometric Properties:**
  - *Reliability:* High internal consistency reported across versions (Cronbach's α: 0.74–0.93).
  - *Validity:* Confirmed construct, criterion, and concurrent validity. Validated against DSM criteria and urine drug tests. Persian and Arabic versions maintain structural integrity with adequate sensitivity and specificity.

    - *Factor Structure:* Primarily unidimensional across studies, supporting its use as a general screening measure.
- **Pragmatic Properties:**
    - *Ease of Use:* Brief (DAST-10 takes ~5 minutes), self-administered, or interview format.
    - *Acceptability and Compatibility:* Culturally adapted and translated into Arabic and Persian with validation studies supporting cross-cultural use.
    - *Scoring and Interpretation:* Straightforward scoring with established cutoffs for intervention thresholds.
    - *Accessibility:* Freely available for non-commercial clinical and research use.
    - *Actionability:* Supports early identification and triage to appropriate treatment levels.

---

4. Application to Integrated Primary Care

DAST is a feasible screening tool for identifying substance use issues in primary care settings, including among patients with comorbid conditions like ADHD. The tool's brevity and ease of use make it suitable for diverse, underserved populations, facilitating timely referrals and integrated behavioral health interventions. Its adaptability across languages supports application in multicultural U.S. settings.

---

5. Strengths and Limitations

- **Strengths:** Strong psychometric support; multiple validated translations; suitable for varied clinical settings; promotes brief, standardized substance use screening.
- **Limitations:** Limited validation in the justice system or settings where underreporting may be significant; cultural nuances may affect responses even in translated versions.

---

6. Recommendations for Practice and Research

- Clinicians in primary care should integrate DAST-10 as part of routine behavioral health screenings, especially in high-risk populations.
- Future research should explore DAST use in telehealth settings, among pediatric populations transitioning to adult care, and in Spanish-speaking U.S. populations.
- Additional validation studies in justice-involved or rural primary care populations would enhance generalizability.

Measure: DAST-10 (Drug Abuse Screening Test – 10 item)
Type: Screening + Diagnostic Support
Languages: Multilingual (English, Arabic, Persian, Spanish, Mandarin)
Validated Populations: U.S. general, ADHD patients, Saudi, Iranian, psychiatric, and substance use disorder treatment populations

Cutoffs: ≥3 (moderate problems), ≥6 (substantial/severe problems)
Psychometrics: Sensitivity 91.5–96%, Specificity 57–92.5%, Reliability (α = 0.75–0.93), Test-retest ICC > 0.99
Setting: Integrated primary care, behavioral health, mental health specialty clinics
Use Case: Identification of drug misuse severity, triage for brief intervention, or referral to substance use treatment

Duke Health Profile (DUKE)[88-96]

1. Objective
To evaluate the psychometric reliability, cross-cultural validity, clinical utility, and application of the DUKE Health Profile as a brief, multidimensional measure of health-related quality of life (HRQoL) in general and clinical populations. The DUKE is especially useful in research, public health surveillance, and primary care settings.

---

2. Measure Evaluated
DUKE Health Profile (DUKE) – A 17-item self-report tool designed to assess HRQoL across 10 domains: 6 health function scales (physical, mental, social, general, perceived health, and self-esteem) and 4 dysfunction scales (anxiety, depression, pain, and disability). The tool measures patient health over the past week and produces standardized scores from 0 (worst) to 100 (best for functional scales; reverse for dysfunction scales). It is available in multiple languages and can be used for individuals aged 12+.

---

3. PAPERS Framework Evaluation
Psychometric Properties:
• Internal consistency (Cronbach's α):
  – General health: 0.71 (France)
  – Physical: 0.62, Mental: 0.63, Social: 0.34–0.53, Self-esteem: 0.46–0.47
  – Anxiety/Depression: 0.57–0.62
• Test–retest reliability: ICC = 0.64–0.88 (Persian version)
• Content validity (Persian version): I-CVI = 88–100%, S-CVI = 94–96%
• Confirmatory factor analysis supports the construct validity of functional and dysfunction domains
Pragmatic Properties:
• **Acceptability:** Excellent – brief (≤5 minutes), acceptable to older adults and adolescents
• **Feasibility:** Suitable for self or interviewer administration; validated via phone, paper, or clinical visit formats
• **Interpretability:** Scores by domain, allowing granular assessment of HRQoL; can be summed

for general HRQoL score

• **Equity:** Translated into 17+ languages, including French, Persian, and Spanish; validated in adolescents and older adults

• **Sustainability:** Used in public health surveillance (e.g., French National Health Barometer) and research studies worldwide

---

4. Application to Integrated Primary Care

DUKE is a highly feasible tool for integrated primary care, particularly in chronic disease management, behavioral health monitoring, and care transitions. Its utility in quick HRQoL profiling makes it a valuable option for health coaches, care managers, and interdisciplinary care teams. It complements clinical interviews and can guide shared decision-making in cases of multimorbidity, especially in aging and adolescent populations.

---

5. Strengths and Limitations

Strengths:

• Very brief with broad domain coverage
• Suitable for adolescents and older adults
• High test–retest reliability
• Validated for cross-cultural use
• Good balance of functional and dysfunction measures

Limitations:

• Lower internal consistency in social and self-esteem domains
• Less commonly used in the U.S. compared to SF-36 or PROMIS tools
• Requires manual scoring if not integrated digitally
• Some scales contain only one item (e.g., pain, disability), limiting reliability

---

6. Recommendations for Practice and Research

Practice:

• Use DUKE in primary care and public health for rapid HRQoL screening
• Incorporate in annual wellness visits or chronic disease reviews
• Pair with specific condition-based tools (e.g., PHQ-9, MoCA) for multidimensional assessment

Research:

• Validate updated versions across racial/ethnic minorities and adolescents in the U.S.
• Explore DUKE's utility in mental health and social prescribing interventions
• Study longitudinal responsiveness to interventions and patient-centered outcomes

---

Digital Repository Format

Measure: Duke Health Profile (DUKE)
Type: Health-Related Quality of Life Assessment (HRQoL)
Languages: Multilingual (English, French, Persian, Spanish, German, Dutch, Korean, etc.)
Validated Populations: French and Iranian general populations, U.S. adults, adolescents, older adults, chronic disease cohorts
Cutoffs: No standard cutoffs; scores reported 0–100 for each domain (higher = better for function, worse for dysfunction)
Psychometrics: Cronbach's α = 0.34–0.71; Test–retest ICC = 0.64–0.88; CFA-supported structure
Setting: Primary care, public health, chronic disease management, adolescent and geriatric care
Use Case: Brief HRQoL screening across physical, mental, and social domains for individual and population health monitoring

Edinburgh Postnatal Depression Scale (EPDS)[97-106]

1. Objective
To assess the psychometric and pragmatic properties of the EPDS and related maternal mental health screening tools across diverse populations and settings, especially among underserved groups in LMICs and U.S. primary care contexts.

---

2. Measure Evaluated
Edinburgh Postnatal Depression Scale (EPDS): A 10-item self-report tool used globally to screen for symptoms of postnatal depression. Evaluations included culturally adapted versions in multiple languages (English, Spanish, Malay, Chinese, Amharic, Kiswahili).

---

3. PAPERS Framework Evaluation
Psychometric Properties:

- **Reliability**: Internal consistency (Cronbach's α) ranged from 0.71–0.90 across studies, indicating good reliability.
- **Validity**: Strong construct, content, and criterion validity established. Sensitivities/specificities were generally >75%. Local validation improved diagnostic accuracy (e.g., Kenya PDEPS: 90%/90%).
- **Factor Structure**: Supported one- or two-factor models (e.g., depression and anxiety components in Hispanic, Danish, and rural U.S. populations).

Pragmatic Properties:

- **Acceptability**: High completion rates; well-tolerated across cultural and literacy levels when adapted.
- **Feasibility**: Easily administered by non-specialists; requires minimal training.
- **Cost & Time**: Low cost, ~5-10 minutes to administer.

- **Interpretability**: Clear scoring guidelines; varying cut-off scores based on cultural norms (e.g., ≥11 in Denmark; ≥13 in India).
- **Sustainability**: Can be integrated into primary care routines with appropriate staff support.

---

4. Application to Integrated Primary Care

EPDS demonstrates strong relevance for integrated behavioral health in U.S. primary care settings, especially among:

- **Rural populations** (validated in U.S. rural women)
- **Hispanic communities** (validated bilingual tool)
- **Immigrant groups** (validated in Malay, Amharic, Chinese) The tool's adaptability and strong psychometrics make it suitable for systematic mental health screening, facilitating early identification and intervention in underserved settings.

---

5. Strengths and Limitations

Strengths:

- Validated across diverse global populations.
- Flexible, brief, and self-administered.
- Accurately discriminates depressive symptoms from somatic postpartum symptoms.

Limitations:

- Cultural variations in symptom expression can affect cut-off scores and interpretation.
- Less effective in some rural LMIC contexts unless locally adapted.
- Limited data on performance across the full perinatal spectrum (antenatal, postpartum, extended).

---

6. Recommendations for Practice and Research

Practice:

- Integrate the EPDS in routine care for perinatal women in FQHCs and community clinics.
- Use culturally validated versions with tailored cut-off scores.
- Train primary care staff in use and referral pathways.

Research:

- Further validate and compare EPDS adaptations for specific underserved U.S. populations (e.g., Native American, immigrant communities).
- Investigate the effectiveness of digital administration and follow-up integration.
- Develop and test brief tools that blend idioms of distress with global criteria (e.g., PDEPS in Kenya).

Digital Repository Format

Measure: Edinburgh Postnatal Depression Scale (EPDS)
Type: Screening Tool
Population Validated: Global (incl. U.S. rural, Hispanic, African, and Asian women)
Psychometrics: Cronbach's α 0.71–0.90, sensitivity/specificity ≥75%
Languages: Multilingual (English, Spanish, Malay, Amharic, Chinese, Kiswahili)
Cut-offs: Varies (≥9 to ≥13)
Clinical Utility: High – suitable for integrated primary care
Cultural Considerations: Requires local validation for optimal use

Epworth Sleepiness Scale (ESS)[107-118]

1. Objective

To evaluate the psychometric properties, cross-cultural utility, and clinical relevance of the ESS as a subjective self-report tool for measuring excessive daytime sleepiness (EDS) in adult populations across diverse global and U.S. primary care settings.

---

2. Measure Evaluated

Epworth Sleepiness Scale (ESS) – An 8-item, self-administered questionnaire that assesses the likelihood of dozing in various routine daily situations (e.g., watching TV, sitting quietly after lunch). Each item is scored from 0 (no chance) to 3 (high chance), with total scores ranging from 0 to 24.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **Internal Consistency**:
    - Meta-analytic Cronbach's α = 0.82 (CI: 0.798–0.832) based on 63 estimates from over 92,000 participants.
    - Reliability in specific populations:
        - Obstetric sample: α = 0.75 (Factor 1 = 0.75, Factor 2 = 0.52)
        - Older adults: Good reliability among both Black and White women aged 70+
        - Hindi (India): α = 0.89; excellent test–retest reliability
- **Construct Validity**:
    - ESS scores correlate significantly with objective sleep disorder severity (e.g., OSA, narcolepsy) and subjective complaints of daytime sleepiness.
    - Validated through polysomnography, PSQI correlations, and sleep latency measures across multiple settings.
- **Discriminant Validity**:

    - Differentiates between individuals with and without clinically diagnosed sleep disorders (e.g., sleep apnea, narcolepsy).

Pragmatic Properties:

- **Acceptability**: Highly accepted and frequently used in sleep clinics, primary care, geriatrics, and research.
- **Feasibility**: Takes 2–4 minutes to complete. Easy to score manually or via EHR.
- **Interpretability**:
    - Scores >10 indicate clinically significant daytime sleepiness.
    - Scores >16 suggest severe EDS, often linked with narcolepsy or OSAS.
- **Equity**:
    - Successfully translated and validated in dozens of languages (e.g., Spanish, Hindi, Persian, Turkish, Chinese, Portuguese, and Colombian Spanish), with cultural adaptations for driving vs. non-driving populations.
- **Sustainability**: Routinely integrated into sleep medicine, geriatrics, behavioral health, and research protocols globally.

---

4. Application to Integrated Primary Care

ESS enables early identification of EDS within primary care, especially for patients at risk of obstructive sleep apnea (OSA), depression, or cardiovascular disease. Its integration into EHRs or pre-visit screenings supports care coordination and referrals to behavioral health or sleep specialists. Also effective in population health screenings and prenatal care settings.

---

5. Strengths and Limitations

Strengths:

- Brief, easy-to-administer, cost-free tool
- Validated across clinical and non-clinical populations
- High sensitivity and test–retest reliability
- Widely translated and culturally adaptable

Limitations:

- Some ambiguity in items (e.g., passenger vs. driver in traffic)
- Ceiling effects in severe cases
- May underperform in populations with atypical routines (e.g., rural, non-driving, or culturally distinct groups) without local adaptation

Generalized Anxiety Disorder Scale (GAD-7 and GAD-2)[119-127]

1. Objective

To assess the psychometric and pragmatic utility of the Generalized Anxiety Disorder scales (GAD-7 and GAD-2) across diverse cultural and linguistic contexts, including underserved populations in the U.S., for use in primary care behavioral health integration.

---

2. Measure Evaluated

GAD-7: A 7-item self-report measure assessing symptoms of generalized anxiety disorder over the past two weeks.

GAD-2: A 2-item ultra-brief version of the GAD-7, focusing on core anxiety symptoms for rapid screening.

---

3. PAPERS Framework Evaluation

Psychometric Properties

- **Reliability**: High internal consistency (GAD-7 $\alpha = 0.87–0.92$; GAD-2 $\alpha = 0.75–0.91$)
- **Validity**: Strong construct and criterion validity across populations (AUCs typically >0.90)
- **Cross-Cultural Validation**: Successfully validated in multiple languages (Spanish, Korean, Latvian, Kinyarwanda, Icelandic, Chinese), including rural and immigrant/refugee populations

Pragmatic Properties

- **Acceptability**: Well-tolerated and favorably received by patients
- **Feasibility**: Extremely brief (GAD-2 <2 min; GAD-7 <5 min); easily self-administered
- **Scoring**: Clear cut-offs; GAD-7 ≥10 suggests moderate anxiety; GAD-2 cut-off often ≥3
- **Interpretability**: Direct correlation with functional impairment and quality of life
- **Cost & Accessibility**: Free, open-access tools available in multiple languages

---

4. Application to Integrated Primary Care

The GAD-7 and GAD-2 are highly suited for integrated primary care due to their brevity, ease of administration, and robust psychometric support. Their effectiveness has been confirmed across diverse U.S. populations, including Hispanic Americans and Kinyarwanda-speaking African refugees, and are especially helpful in resource-limited and high-need settings like FQHCs and rural clinics.

---

5. Strengths and Limitations

Strengths:

- Consistently high psychometric quality across cultures
- Strong clinical utility in primary care
- Wide accessibility and ease of integration

Limitations:

- Slight variations in optimal cut-offs between populations

- Less effective as a diagnostic tool for anxiety subtypes beyond GAD

---

6. Recommendations for Practice and Research

Practice:

- Use GAD-7 for initial screening and monitoring of anxiety severity.
- Employ GAD-2 in high-volume or time-limited settings as a first-line screener.
- Translate and validate for local dialects when working with immigrant populations.

Research:

- Investigate digital and culturally adapted implementation strategies.
- Evaluate longitudinal responsiveness to treatment in diverse populations.
- Develop adjunct tools to complement GAD screening in multilingual populations.

Digital Repository Format

Measure: GAD-7 / GAD-2

Type: Screening Tool

Validated Populations: Global and U.S. (Hispanic, Kinyarwanda-speaking, rural)

Languages: English, Spanish, Korean, Latvian, Russian, Icelandic, Chinese, Kinyarwanda

Cut-offs: GAD-7 (≥10), GAD-2 (≥3)

Psychometrics: Excellent reliability/validity ($\alpha \geq 0.87$, AUC > 0.90)

Use Case: Anxiety screening in primary care and integrated behavioral health

Format: Free, self-administered; available in print and electronic formats

Insomnia Severity Index (ISI)[128-137]

1. Objective

To evaluate the psychometric properties, clinical utility, and cultural adaptability of the Insomnia Severity Index (ISI) as a brief self-report instrument to assess the presence, severity, and impact of insomnia symptoms in diverse clinical and non-clinical adult populations.

---

2. Measure Evaluated

Insomnia Severity Index (ISI) – A 7-item self-report tool assessing the nature, severity, and distress caused by insomnia. Each item is rated on a 5-point Likert scale (0–4), yielding a total score range of 0 to 28. Domains include difficulty falling asleep, staying asleep, early waking, dissatisfaction with sleep, and daytime functional impairment.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

Internal Consistency:

- Meta-analytic α across studies ranges from **0.74 to 0.94**, demonstrating high reliability.
- Veterans with TBI: α = **0.91**; two-factor model supported (Insomnia Impact, Sleep Onset/Maintenance).
- Young adult cancer survivors: α = **0.90**; strong diagnostic validity compared to SCID-5.
- Hindi version: α = **0.90**; test–retest reliability = **0.90** (ICC).
- Chinese version: α = **0.81**; good test–retest and construct validity.

Construct Validity:

- ISI scores positively correlated with objective sleep latency, poor sleep hygiene, and subjective sleep dissatisfaction.
- Correlated with PSQI, HADS, BDI-II, and other psychological distress measures in cross-cultural validations (e.g., Portuguese, Mexican, Chinese).

Discriminant Validity:

- Differentiates between insomnia and normal sleepers or those with other sleep or psychiatric disorders.
- SCID-5 comparison revealed high AUC of **0.86**, sensitivity = **0.88**, specificity = **0.85**.

Pragmatic Properties:

Acceptability:

- Widely used in clinical practice (primary care, oncology, veteran health, psychiatry) and research.

Feasibility:

- Self-administered, paper or digital. Takes **5 minutes** or less to complete.

Interpretability:

- Total Scores:
    - 0–7 = No clinically significant insomnia
    - 8–14 = Subthreshold insomnia
    - 15–21 = Clinical insomnia (moderate)
    - 22–28 = Clinical insomnia (severe)

Equity:

- Validated in more than 10 languages, including Hindi, Portuguese, Chinese, and Spanish. Appropriate for use in diverse cultural contexts.

Sustainability:

- Integrated into behavioral health protocols, oncology clinics, veteran services, and population health research.

---

4. Application to Integrated Primary Care

ISI provides a rapid, standardized way to assess and monitor insomnia symptoms in primary care, behavioral health, oncology, and specialty clinics. It supports stepped-care interventions,

referral decisions, and tracking outcomes in evidence-based behavioral therapies (e.g., CBT-I). ISI results are actionable within integrated electronic health records and can prompt medication reviews or sleep hygiene counseling.

---

5. Strengths and Limitations

Strengths:

- Brief, easy-to-use, and free for clinical and academic use
- Strong psychometric support across global populations
- Useful for both screening and severity tracking
- Validated in medically complex patients (e.g., cancer, TBI, veterans)

Limitations:

- Limited sensitivity for sleep architecture features (e.g., REM latency)
- Does not assess sleep apnea or parasomnia risk
- Ceiling effects in psychiatric populations with multiple comorbidities

---

6. Recommendations for Practice and Research

Practice:

- Integrate ISI into routine screenings in primary care, oncology, and behavioral health.
- Use as a monitoring tool in CBT-I and pharmacotherapy.
- Combine with PSQI or STOP-Bang for differential diagnosis of insomnia vs. OSA.

Research:

- Expand validation in adolescent and older adult populations.
- Examine sensitivity to treatment effects across digital and in-person CBT-I interventions.
- Explore dimensional factor structures in racial/ethnic subgroups and low-literacy settings.

Digital Repository Format

- **Measure**: Insomnia Severity Index (ISI)
- **Type**: Symptom Screening – Insomnia
- **Languages**: English, Hindi, Portuguese, Chinese, Spanish, Korean, Nepali
- **Validated Populations**: Primary care, oncology, veterans, Chinese, Indian, Portuguese, nurses, TBI, cancer survivors
- **Cutoffs**:
    - ISI ≥ 15 = Clinical insomnia
    - ISI ≥ 22 = Severe insomnia
- **Psychometrics**: α = 0.74–0.94; strong test–retest, construct and discriminant validity
- **Setting**: Integrated care, oncology, veterans, sleep research

- **Use Case**: Screening, monitoring, and treatment planning for insomnia and sleep-related distress in diverse patient populations.

Montreal Cognitive Assessment (MoCA)[138-147]

1. Objective

To evaluate the reliability, diagnostic accuracy, and cross-cultural adaptability of the Montreal Cognitive Assessment (MoCA) for detecting mild cognitive impairment (MCI), major neurocognitive disorder (dementia), and cognitive decline in diverse populations, including those with sensory impairments, low education, and ethnically diverse older adults.

---

2. Measure Evaluated

Montreal Cognitive Assessment (MoCA) – A 30-point, one-page cognitive screening tool assessing multiple domains: visuospatial/executive function, naming, memory, attention, language, abstraction, and orientation. Originally validated in older adults to detect MCI, MoCA has since been adapted into over 60 languages, with special versions for sensory impairments (e.g., MoCA-H for hearing-impaired).

Versions reviewed include:

- MoCA-H (Hearing Impaired)
- MoCA-C (Chinese), MoCA-R (Rwandese), Tamil MoCA (Sri Lanka), Portuguese MoCA
- Validation in American Indian, Ethiopian, and Eastern Chinese populations

---

3. PAPERS Framework Evaluation

Psychometric Properties:

• Internal consistency: α = .79–.89 across languages and populations
• Inter-rater reliability: ICC ≥ 0.92 in most studies (e.g., Ethiopia, MoCA-H)
• Sensitivity and specificity:
  – Original MoCA: Sensitivity = 90%, Specificity = 87% for MCI (Nasreddine et al.)
  – Tamil MoCA: Sensitivity = 84.7%, Specificity = 76.4% (cutoff = 23/24)
  – Ethiopian MoCA: Sensitivity = 87.2%, Specificity = 74.0% (cutoff = ≤21)
  – MoCA-H: AUC = 0.973, Sensitivity = 92.8%, Specificity = 90.8% (cutoff = 24)
• Construct validity confirmed across cultures using ROC, CFA, Rasch analysis
• Education and depression significantly affect scores—highlighting the need for contextual interpretation

Pragmatic Properties:

• **Acceptability:** Easy to administer (~10 minutes); minimal burden; translated into 60+ languages
• **Feasibility:** Suitable for community, clinic, and low-resource settings; validated in LMICs

• **Interpretability:** Traditional cutoff = 26; updated evidence supports lower cutoffs (e.g., 23–24) depending on population
• **Equity:** Demonstrated cross-cultural adaptability; specific versions developed for American Indian, African, and hearing-impaired populations
• **Sustainability:** Recommended in national guidelines (Canada, UK); integrated into Alzheimer's clinical trials and EHRs

---

4. Application to Integrated Primary Care

MoCA is highly applicable in integrated and primary care settings for early detection of cognitive impairment. It supports diagnosis, referral, and shared decision-making in older adults and high-risk populations. The MoCA-H version expands accessibility for patients with hearing impairment, and culturally adapted versions increase appropriateness for BIPOC, immigrant, and multilingual communities.

---

5. Strengths and Limitations

Strengths:

• High diagnostic utility across conditions (MCI, AD, vascular dementia, Parkinson's)
• Validated in more than 60 languages
• Freely accessible and endorsed by Alzheimer's and geriatric societies
• Adaptable for sensory impairments (e.g., MoCA-H) and cross-cultural needs

Limitations:

• Original cutoff (26) may yield false positives in low-education or older adults
• Cultural differences affect item difficulty (e.g., naming tasks, clock drawing)
• Less effective in individuals with severe sensory deficits or limited education without adjustment

---

6. Recommendations for Practice and Research

Practice:

• Use MoCA as a primary screen for MCI and dementia in integrated primary care and geriatrics
• Adjust scoring for education (e.g., +1 point for ≤12 years of education)
• Use localized or adapted versions (e.g., MoCA-H, MoCA-Tamil, MoCA-Rwandese) when applicable
• Interpret scores in the context of functional status, language fluency, and comorbidities

Research:

• Further refine culturally relevant items and develop new norms in underrepresented populations
• Study longitudinal utility for tracking progression or treatment outcomes
• Validate digital or app-based delivery modes, including in telehealth environments

Digital Repository Format

Measure: Montreal Cognitive Assessment (MoCA) / MoCA-H / MoCA-C / MoCA-Tamil / MoCA-R
Type: Cognitive Screening
Languages: Multilingual (English, Tamil, Chinese, Portuguese, Rwandese, French, Spanish, etc.)
Validated Populations: Global older adults, American Indian, Sri Lankan, Ethiopian, Eastern Chinese, hearing-impaired, low-education
Cutoffs: Traditional = 26; population-specific cutoffs = 21–24
Psychometrics: Sensitivity 84–97%; Specificity 74–91%; Reliability α = .79–.89; ICC ≥ 0.92
Setting: Primary care, geriatrics, memory clinics, community health, LMICs
Use Case: Early detection of MCI and dementia, cognitive screening in diverse populations, memory clinic triage

Mood Disorder Questionnaire (MDQ)[148-156]

1. Objective

To assess the reliability, validity, cross-cultural applicability, and clinical utility of the Mood Disorder Questionnaire (MDQ) as a screening instrument for bipolar disorder (BD), especially bipolar II and spectrum disorders, across diverse international and U.S. primary care and psychiatric populations, including trauma-exposed, non-English-speaking, and low-resource settings.

---

2. Measure Evaluated

Mood Disorder Questionnaire (MDQ) – A 13-item self-report screening tool developed to identify lifetime symptoms of mania and hypomania, including a symptom checklist, symptom co-occurrence, and impairment severity assessment.
Cultural and linguistic adaptations examined: Chinese (Mainland and Hong Kong), Rwandese (Kinyarwanda), Italian, Spanish, Korean, Thai, Arabic (Tunisian), and Finnish versions.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

• Internal consistency: α = .80–.91 across translations (Chinese, Korean, Rwandese, Tunisian).
• Sensitivity: Ranges from 0.61 to 0.94 depending on population and version; strongest for BD-I, weaker for BD-II.
• Specificity: Ranges from 0.58 to 0.97; cultural and clinical setting dependent.
• ROC AUC: Rwandese version = 0.99; Chinese = ~0.75–0.80; Korean = 0.77; Tunisian = high NPV (0.92).

• Factor Structure: Two- and three-factor models validated (e.g., acceleration, energy, imprudence dimensions).

Pragmatic Properties:

• **Acceptability:** Self-administered; takes 5–7 minutes; no cost to use.

• **Feasibility:** Successfully used in outpatient psychiatry, primary care, community, and low-resource settings.

• **Interpretability:** Original scoring requires ≥7 symptoms plus co-occurrence and impairment; many versions drop the latter for improved BD-II detection.

• **Equity:** Validated across diverse ethnic, linguistic, and geographic groups including African American, Chinese, Rwandese, Thai, Korean, Arabic-speaking, and Italian populations.

• **Sustainability:** Used globally; available in digital (e.g., MDCalc) and print formats.

---

4. Application to Integrated Primary Care

MDQ is highly relevant for integrated primary care practices as a first-line screen for bipolar disorder, particularly in patients presenting with depression or trauma history. While effective for BD-I detection, care must be taken in interpreting results for BD-II and culturally diverse patients. Revised scoring (e.g., ignoring co-occurrence or impairment questions) enhances sensitivity. Useful for collaborative care models, behavioral health integration, and universal screening initiatives in FQHCs.

---

5. Strengths and Limitations

Strengths:

• Validated internationally with robust psychometric properties.

• Simple, fast, and low-burden; enables early detection of BD.

• Cross-culturally adaptable; available in many languages.

Limitations:

• Lower sensitivity for BD-II if all three scoring criteria are applied.

• High false-positive rates in trauma-exposed or non-psychiatric settings (e.g., PPV ~17% in trauma-exposed African American sample).

• Factor structure inconsistencies; symptom overlap with other disorders (e.g., ADHD, PTSD).

• Best used as part of a broader diagnostic workflow, not standalone.

---

6. Recommendations for Practice and Research

Practice:

• Administer the MDQ in primary care for patients presenting with depressive symptoms.

• Use revised scoring methods (symptom checklist only) to improve sensitivity for BD-II.

• Combine with clinical interview and structured tools (e.g., SCID) for diagnosis.

• Train providers on symptom interpretation, especially in diverse populations.

Research:

• Validate MDQ adaptations for Black, Indigenous, and People of Color (BIPOC) in U.S. primary care.

• Test digital delivery via EHRs and patient portals for scalable implementation.

• Study predictive validity for treatment response, cost reduction, and suicide prevention.

Digital Repository Format

Measure: Mood Disorder Questionnaire (MDQ)

Type: Screening + Diagnostic Support

Languages: Multilingual (English, Chinese, Thai, Italian, Korean, Arabic, Finnish, Rwandese)

Validated Populations: Global psychiatric, trauma-exposed primary care, African American, East Asian, Sub-Saharan African, Mediterranean

Cutoffs: ≥7 symptoms endorsed (original); some versions use ≥6 without co-occurrence/impairment for better BD-II detection

Psychometrics: Sensitivity 0.61–0.94, Specificity 0.58–0.97, Reliability α = .80–.91

Setting: Psychiatric care, primary care, global mental health clinics, low-resource contexts

Use Case: Screening for bipolar disorder (especially BD-I and BD spectrum), early detection in patients with depressive symptoms

Opioid Risk Tool (ORT/ORT-OUD)[157-164]

1. Objective

To evaluate the predictive and diagnostic capacity, psychometric reliability, and clinical relevance of the original and revised versions of the Opioid Risk Tool (ORT) in diverse healthcare settings, including pain clinics, oncology, primary care, and community pharmacy, with particular attention to chronic noncancer pain and opioid use disorder (OUD) prevention.

---

2. Measure Evaluated

Opioid Risk Tool (ORT) – A 10-item brief screening tool assessing personal and family history of substance use, age, psychological comorbidities, and (in the original) preadolescent sexual trauma. Administered in primary care and pain settings to stratify risk for opioid misuse.

ORT-OUD (Revised ORT) – A simplified 9-item tool designed specifically to predict the risk of developing Opioid Use Disorder (OUD), eliminating gendered scoring and controversial items for improved reliability and equity.

Both tools are designed for clinician or patient administration and stratify risk as low (0–3), moderate (4–7), or high (≥8).

---

3. PAPERS Framework Evaluation

Psychometric Properties:

• Original ORT: Mixed validity; AUCs ranged from .35 to .73 in predictive studies; reliability inconsistently reported.

• ORT-OUD: Superior predictive performance (AUC .88), strong sensitivity (0.85), specificity (0.85), and PPV/NPV in CNMP cohorts.

• Validity is higher with clinician administration; patient self-report may be biased by social desirability.

Pragmatic Properties:

• **Acceptability:** Easy to use (<1 min); low burden; widely disseminated.

• **Feasibility:** Implementable across settings (EHR, paper, digital apps like MDCalc).

• **Interpretability:** Simple score bands (0–3 = low, 4–7 = moderate, 8+ = high); revised tools remove ambiguity.

• **Equity:** Revised versions (e.g., ORT-OUD) reduce gender bias and improve accessibility in multilingual settings.

• **Sustainability:** Used in national pharmacy programs, oncology clinics, and behavioral health integration initiatives.

---

4. Application to Integrated Primary Care

ORT is increasingly used in primary care as part of opioid stewardship initiatives to identify patients at risk of opioid misuse or developing OUD. The tool's brevity makes it suitable for high-volume practices, including Federally Qualified Health Centers (FQHCs) and rural clinics. Integrating ORT into EHR workflows enhances communication and decision-making across interdisciplinary teams, especially when combined with follow-up assessments and opioid agreements.

---

5. Strengths and Limitations

Strengths:

• Extremely brief and easy to score.

• Strong predictive validity when revised (ORT-OUD) and used in CNMP populations.

• Facilitates risk stratification and targeted patient education.

Limitations:

• Original version criticized for gender-biased scoring and cultural insensitivity.

• Mixed psychometric performance in Spanish- and Arabic-speaking populations.

• Predictive accuracy declines in self-report format; requires clinician involvement for optimal accuracy.

• May not reflect dynamic risk; better used as an initial screen than as a standalone decision tool.

---

6. Recommendations for Practice and Research

Practice:

• Use ORT-OUD in integrated primary care and community pharmacy settings for patients initiating opioid therapy.

• Administer via EHR with clinician oversight to reduce self-report bias.

• Pair with additional monitoring tools (e.g., COMM, PDMP, urine screening) for longitudinal tracking.

Research:

• Further validation in racially/ethnically diverse, non-English-speaking, and oncology populations.

• Study comparative performance of ORT vs. SOAPP-R and other tools in predicting OUD.

• Evaluate longitudinal outcomes from ORT-informed clinical decisions in pain and primary care clinics.

Digital Repository Format

Measure: Opioid Risk Tool (ORT) / Revised ORT (ORT-OUD)

Type: Screening + Risk Stratification

Languages: English, Spanish, Arabic

Validated Populations: U.S. chronic pain patients, oncology patients, Lebanese general population, Spanish-speaking populations

Cutoffs: Low risk (0–3), Moderate (4–7), High (≥8)

Psychometrics: ORT-OUD: AUC = 0.88, Sensitivity = 0.85, Specificity = 0.85; ORT: Mixed (AUC 0.35–0.73)

Setting: Integrated primary care, community pharmacy, pain management, oncology

Use Case: Predicting risk for opioid misuse/OUD, supporting opioid stewardship, guiding prescribing decisions

Outcome Rating Scale (ORS)[165-173]

1. **Objective**
   To evaluate the psychometric properties, cross-cultural adaptability, and clinical effectiveness of the Outcome Rating Scale (ORS) as a brief, client-reported tool for tracking well-being and therapeutic progress across diverse global populations in behavioral health and primary care settings.
2. **Measure Evaluated**
   Outcome Rating Scale (ORS) – A 4-item, visual analog self-report tool measuring a client's sense of well-being across individual (personal well-being), interpersonal (family, close relationships), social (work, school), and general domains. Each item is scored from

0–10, with higher scores indicating greater perceived well-being. Total scores range from 0–40.

3. **PAPERS Framework Evaluation**

Psychometric Properties:

- *Internal Consistency:*
  - Initial development study: Cronbach's α = 0.87–0.96 in U.S. community mental health samples.
  - Spanish version (Spain): α = 0.91; test–retest reliability = 0.82.
  - Dutch version: α = 0.93 in adults seeking psychotherapy.
  - Czech version: α = 0.87; excellent construct validity across clinical subscales.
- *Construct Validity:*
  - Correlated significantly with longer instruments such as the Outcome Questionnaire-45 (OQ-45) and CORE-OM.
  - Sensitive to symptom changes over time and aligned with client-perceived progress.
- *Discriminant Validity:*
  - Differentiates clinical from non-clinical populations and tracks change over time better than longer measures in brief settings.

Pragmatic Properties:

- *Acceptability:* Very high across languages and cultures due to brevity and visual format.
- *Feasibility:* Completion time is under 1 minute. Minimal training required; can be self-administered or clinician-guided.
- *Interpretability:*
  - Scores <25 typically indicate clinical concern.
  - A 6+ point change from baseline is considered clinically significant improvement.
- *Equity:*
  - Validated and culturally adapted in multiple languages (Spanish, Dutch, Czech, Portuguese, etc.)
  - Adapted for youth, adults, and diverse health literacy levels.
- *Sustainability:* Embedded in measurement-based care frameworks globally. Frequently used in integrated behavioral health, substance use treatment, and school-based interventions. Free for individual use, but users must agree to the license and register at https://betteroutcomesnow.com/download-ors-srs/. There is a fee for group licenses.

4. **Application to Integrated Primary Care**

The ORS is ideal for integrated behavioral health settings due to its brevity, ability to track therapeutic outcomes session-to-session, and client-centered design. Its use enhances collaborative care by offering real-time feedback for care planning, particularly

useful in warm hand-offs, behavioral health consultant models, and stepped care programs.

5. **Strengths and Limitations**

Strengths:

- Ultra-brief and highly acceptable to clients.
- Strong psychometric performance in clinical and non-clinical samples.
- Encourages client-clinician collaboration and shared decision-making.
- Excellent cross-cultural adaptability.

Limitations:

- Visual analog format may be difficult for some individuals with visual or fine motor impairments.
- May lack depth for complex diagnostic evaluations.
- Interpretation of domains may vary slightly across cultures without local adaptation.

6. **Recommendations for Practice and Research Practice:**

- Use ORS routinely at the beginning of behavioral health sessions to monitor client functioning.
- Integrate with EHRs for longitudinal tracking and case management with license.
- Pair with the Session Rating Scale (SRS) to evaluate alliance and client experience.
- Software support for MBC is available through licensing with the developer.

Research:

- Expand longitudinal studies validating ORS change sensitivity across treatment modalities.
- Validate in low-literacy, neurodiverse, and rural populations.
- Further assess predictive utility for treatment dropout and functional improvement.

Digital Repository Format

- **Measure:** Outcome Rating Scale (ORS)
- **Type:** Patient-Reported Outcome Measure (PROM)
- **Languages:** English, Spanish, Dutch, Czech, Portuguese and 30+ others
- **Validated Populations:** Adults, youth, primary care, behavioral health, multicultural samples
- **Cutoffs:** ORS <25 indicates clinical concern; 6+ point increase indicates meaningful improvement
- **Psychometrics:** Cronbach's $\alpha = 0.87–0.96$; valid across global settings
- **Setting:** Integrated behavioral health, primary care, school-based care, community mental health
- **Use Case:** Routine outcome monitoring in brief interventions, tracking well-being across therapy sessions

Pain, Enjoyment of Life, and General Activity Scale (PEG)[174--176]

1. Objective

To evaluate the psychometric properties, cultural relevance, and clinical applicability of the 3-item PEG scale—a brief measure derived from the Brief Pain Inventory (BPI)—for assessing pain intensity and its impact on function across diverse primary care settings, including Spanish-speaking populations.

---

2. Measure Evaluated

PEG Scale – A 3-item self-report tool assessing:

1. **Pain intensity**,
2. **Pain interference with enjoyment of life**, and
3. **Pain interference with general activity**.
   Each item is scored on a 0–10 numeric scale (higher scores = worse symptoms), with a composite score calculated as the average of the three items.

---

3. PAPERS Framework Evaluation

Psychometric Properties

Internal Consistency:

- **Spanish-speaking U.S. primary care patients:** Cronbach's α = 0.82 (95% CI: 0.77–0.86)
- **Original English version (study 1):** α = 0.73

Convergent Validity:

- Strong correlations with other validated pain measures:
  - BPI Interference: r = 0.79
  - BPI Severity: r = 0.68
  - GCPS Intensity: r = 0.69
  - GCPS Disability: r = 0.69

Discriminant Validity:

- PEG scores showed lower correlation with depressive symptoms (PHQ-9: r = 0.53), supporting construct specificity for pain.

Responsiveness:

- PEG was shown to have **comparable responsiveness** to longer legacy tools like the BPI in chronic pain populations.

---

4. Application to Integrated Primary Care

The PEG offers an efficient, validated method to screen for pain intensity and functional interference within short visit timeframes. Its integration supports care planning, goal setting,

and outcome monitoring—especially in populations with limited literacy or non-English language preference. It is highly applicable in behavioral health integration, chronic disease management, and opioid stewardship programs.

---

5. Strengths and Limitations

Strengths:

- Ultra-brief and validated in both English and Spanish
- Suitable for in-person or telephone administration
- Sensitive to clinical change; useful for goal setting
- Free to use and easy to score

Limitations:

- Composite score may mask individual domain variability
- Lacks information on pain location or quality
- Requires contextual interpretation in patients with multiple chronic conditions

---

6. Recommendations for Practice and Research

Practice:

- Use PEG routinely during primary care or behavioral health visits to assess pain burden.
- Integrate into electronic health records (EHRs) for longitudinal tracking.
- Translate and culturally adapt for other linguistic groups beyond Spanish and English.

Research:

- Further validation in rural, non-English speaking, and geriatric populations.
- Evaluate test–retest reliability and minimally important difference (MID) values across conditions.
- Investigate its utility in digital health tools and mobile self-report platforms.

Digital Repository Format

Measure: Pain, Enjoyment of Life, and General Activity Scale (PEG)

Type: Pain Intensity and Interference Screener

Languages: English, Spanish

Validated Populations: U.S. primary care, Hispanic/Latino Spanish-speaking adults, chronic pain patients

Cutoffs: No standardized clinical cutoff; use mean score (0–10) to monitor change

Psychometrics:

- Spanish version Cronbach's $\alpha = 0.82$
- Convergent validity $r = 0.68–0.79$
- Discriminant validity vs. PHQ-9 $r = 0.53$

  Setting: Primary care, behavioral health, chronic pain management

Use Case: Tracking pain-related outcomes, assessing treatment effectiveness, setting patient-centered goals

Panic Disorder Severity Scale (PDSS)[177-186]

1. Objective

To evaluate the psychometric properties, diagnostic utility, and cultural adaptability of the Panic Disorder Severity Scale (PDSS and PDSS-SR) for assessing symptom severity, functional impairment, and treatment response in individuals with panic disorder (with or without agoraphobia) across clinical and community settings.

---

2. Measure Evaluated

Panic Disorder Severity Scale (PDSS) – A 7-item clinician-administered scale assessing core symptoms of panic disorder: panic frequency, panic-related distress, anticipatory anxiety, agoraphobic avoidance, interoceptive avoidance, and functional impairment in work/social domains.

PDSS-SR – Self-report version with identical items, rated on a 5-point Likert scale (0–4), with modified recall period (past week). Total scores range from 0–28.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **Internal Consistency:**
  - Adult PDSS: α = 0.78–0.92 across studies
  - Swedish PDSS and PDSS-SR: α = 0.88–0.91
  - PDSS-C (Child version): α = 0.82
- **Test–Retest Reliability:**
  - PDSS-C (1-day): r = 0.79
  - PDSS-SR: r = 0.83 (Houck et al.)
- **Construct Validity:**
  - Strong correlations with anxiety and panic-specific scales such as:
    - Anxiety Sensitivity Index (ASI)
    - Clinical Severity Ratings (ADIS)
    - Multidimensional Anxiety Scale for Children (MASC)
- **Discriminant Validity:**
  - Weak or non-significant correlations with depression measures (e.g., CDI) suggest discriminant validity from general emotional distress
- **Factor Structure:**
  - Mixed findings:

        - Original 2-factor model: Panic symptoms (items 1–2) vs. impairment and avoidance (items 3–7)
        - Other studies support a unidimensional model
        - Factor structure varies slightly by language/culture

Pragmatic Properties:

- **Acceptability:**
    - Brief and widely used in both clinical trials and community mental health settings
- **Feasibility:**
    - Takes ~5 minutes to administer
    - Both versions are easy to score and interpret
- **Interpretability:**
    - PDSS cutoffs proposed:
        - 0–3 = Normal
        - 4–9 = Mild
        - 10–13 = Moderate
        - 14–15 = Marked
        - ≥16 = Severe
- **Equity:**
    - Validated translations in multiple languages:
        - French, French-Canadian, Spanish, Korean, Swedish, Chinese
- **Sustainability:**
    - Routinely used in psychiatric and primary care trials for panic disorder
    - Integrated into treatment monitoring protocols

---

4. Application to Integrated Primary Care

The PDSS supports identification and monitoring of panic disorder symptoms in integrated primary care and behavioral health settings. PDSS-SR offers a self-report alternative for primary care, pediatrics, and telehealth, enabling early screening and referral to mental health services. The PDSS-C adaptation facilitates pediatric assessments in schools or outpatient settings.

---

5. Strengths and Limitations

Strengths:

- Validated across diverse populations (adults, adolescents, global translations)
- Sensitive to treatment-related changes
- Includes clinician- and self-report versions for flexibility
- Strong psychometric performance across versions

Limitations:

- Inconsistent factor structure across cultures and versions

- May require contextual adaptation (e.g., school vs. work in PDSS-C)
- Potential underreporting in self-report formats vs. clinician-administered

---

6. Recommendations for Practice and Research

Practice:

- Use PDSS or PDSS-SR in routine behavioral health screenings for panic symptoms
- Employ PDSS-C for adolescents in primary care or school-based health
- Integrate into progress monitoring during treatment of anxiety disorders

Research:

- Further investigate factor structure across global populations
- Evaluate utility in underserved or low-literacy populations
- Examine predictive validity for treatment response and relapse prevention

---

Digital Repository Format

- **Measure:** Panic Disorder Severity Scale (PDSS, PDSS-SR, PDSS-C)
- **Type:** Symptom Severity (Panic Disorder)
- **Languages:** English, Spanish, French, French-Canadian, Korean, Swedish, Chinese
- **Validated Populations:** Adults (clinical and general), adolescents, Swedish, Chinese, French-Canadian, Spanish-speaking populations
- **Cutoffs:**
  - 0–3 = Normal
  - 4–9 = Mild
  - 10–13 = Moderate
  - 14–15 = Marked
  - ≥16 = Severe
- **Psychometrics:**
  - Adult PDSS: α = 0.78–0.92; r (test–retest) = 0.83
  - PDSS-C: α = 0.82; strong construct validity
- **Setting:** Behavioral health, primary care, psychiatry, pediatrics, research trials
- **Use Case:** Diagnostic assessment and treatment monitoring of panic disorder in both adults and adolescents

Patient Health Questionnaire-9 (PHQ-9) and PHQ-2[187-195]

1. Objective

To evaluate the psychometric validity, reliability, and pragmatic utility of the PHQ-9 and PHQ-2 for depression screening and case detection in primary care, with a focus on application in diverse global and U.S. underserved populations.

2. Measure Evaluated

PHQ-9: A 9-item self-administered screening tool based on DSM criteria for major depressive disorder (MDD).

PHQ-2: An ultra-brief screener comprising the first two PHQ-9 items (depressed mood and anhedonia).

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **PHQ-9: Sensitivity: 0.74–0.95; Specificity: 0.76–0.94 at cut-off ≥10. Internal consistency: α = 0.78–0.89 across populations (Chile, Taiwan, Spain, Japan).**
- **PHQ-2: Optimal cut-offs of ≥2 or ≥3 yielded sensitivity of 0.80–0.91 and specificity of 0.70–0.95 (Colombia, Chiapas, Japan).**
- **Factor Structure: One-factor model supported; some populations (elderly Chileans) found overlap in somatic and affective dimensions.**
- **Criterion Validity: Strong correlations with structured clinical interviews (SCID, MINI).**

Pragmatic Properties:

- **Acceptability: Highly acceptable; brief administration time (~2–5 mins).**
- **Feasibility: Effective when administered by non-specialists (e.g., health workers in India and Mexico).**
- **Interpretability: Standardized cut-offs available with cultural adaptations.**
- **Equity: Validated in multiple languages and settings including rural, elderly, low-literacy, and immigrant populations.**
- **Sustainability: Embedded in U.S. clinical workflows and EHR systems; used for screening, treatment planning, and monitoring.**

---

4. Application to Integrated Primary Care

The PHQ-9 and PHQ-2 are well-suited for routine use in U.S. integrated primary care, particularly for screening underserved groups such as Hispanic/Latinx, rural, and refugee populations. Their brevity and reliability make them ideal for use by behavioral health providers, PCPs, and allied staff in settings such as FQHCs.

---

5. Strengths and Limitations

Strengths:

- **Strong cross-cultural validation**
- **High reliability and diagnostic accuracy**
- **Simple administration and scoring**

Limitations:

- **Cultural variability in expression of symptoms may affect cut-off sensitivity**
- **PHQ-2 alone may yield false positives without follow-up**
- **Diagnostic confirmation required (not a stand-alone diagnostic tool)**

---

6. Recommendations for Practice and Research

Practice:

- **Use PHQ-2 for initial screening; follow up with PHQ-9 or clinical interview for diagnostic confirmation.**
- **Adapt cut-off scores to local populations (e.g., ≥9 in elderly; ≥3 for PHQ-2 in Spanish-speaking settings).**
- **Integrate into EHR workflows and team-based care models.**

Research:

- **Further study optimal thresholds across age, gender, ethnicity**
- **Explore longitudinal responsiveness to treatment**
- **Examine digital delivery and patient self-report adaptations in U.S. underserved settings**

Digital Repository Format
Measure: PHQ-9 / PHQ-2
Type: Depression Screening Tool
Languages: English, Spanish, Chinese, Japanese, Malayalam, others
Validated Populations: U.S., Spain, Taiwan, Chile, India, Colombia, Mexico
Cut-offs: PHQ-9 (≥10), PHQ-2 (≥2 or ≥3)
Psychometrics: Sensitivity 0.74–0.95; Specificity 0.70–0.95; α > 0.78
Use Case: Universal depression screening in integrated primary care
Format: Self-administered, paper/digital

Patient Health Questionnaire Modified for Adolescents (PHQ-A)[196-204]

1. Objective

To assess the reliability, validity, and utility of the PHQ-A as a brief, standardized depression screening tool for adolescents (ages 11–21) across diverse global contexts and primary care, mental health, and school settings.

---

2. Measure Evaluated

PHQ-A – A 9-item self-report questionnaire adapted from the PHQ-9 to assess DSM-IV/DSM-5 symptoms of major depressive disorder in adolescents. Items are rated on a 4-point Likert scale

(0 = “Not at all” to 3 = “Nearly every day”), yielding a total score ranging from 0 to 27. Includes an additional item on functional impairment.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

Internal Consistency:

- Thai version: α = 0.92
- Chinese version: α = 0.89
- Swahili (Kenya): α = 0.84
- Nepalese: α = 0.84; test–retest ICC = 0.90

Construct & Criterion Validity:

- Strong convergent validity with CDI and CES-D in Thai study (r = 0.83 and r = 0.87)
- High correlation with clinical diagnosis and SCID in multiple settings
- One-factor structure confirmed in EFA/CFA across translations (Chinese, Thai, Nepalese, Swahili)

Sensitivity & Specificity:

- Thai version: AUC = 0.88; optimal cutoff ≥10 for moderate/severe depression with strong sensitivity and specificity
- Chinese sample: Cutoff ≥10 with sensitivity 0.84, specificity 0.85
- U.S. adolescents: Cutoff ≥11 for optimal detection in primary care

---

Pragmatic Properties

Acceptability:

- Widely accepted by adolescents across school, primary care, and psychiatric settings.
- Easily understood, brief, and translated into multiple languages.

Feasibility:

- Completion time ~2–4 minutes.
- Freely available and suitable for integration into EHR or paper screening.

Interpretability:

- Scores ≥10 indicate moderate depression.
- Scores ≥15 typically warrant further evaluation or treatment.
- Functional impairment item adds clinical relevance.

Equity:

- Validated in diverse settings: U.S., China, Kenya, Nepal, Thailand.
- Adapted for use in Swahili, Thai, Nepali, Chinese, and English.
- Appropriate for low-resource, rural, and underserved populations.

Sustainability:

- Recommended by U.S. Preventive Services Task Force.

- Routinely used in schools, community mental health, and adolescent medicine.

---

4. Application to Integrated Primary Care

The PHQ-A is an efficient, evidence-based screening tool for early identification of depressive symptoms in adolescents in primary care, school-based health centers, and integrated behavioral health settings. It supports decision-making for referral, treatment initiation, and symptom monitoring. Its brevity and widespread validation enhance utility in high-volume, low-resource environments.

---

5. Strengths and Limitations

Strengths:

- Strong internal consistency across cultures
- Short, freely available, and easy to administer
- High validity and reliable cutoffs for clinical decision-making
- Culturally adapted and translated globally

Limitations:

- Self-report bias in some adolescents
- Lacks diagnostic precision—intended as a screening tool
- Some translations need further validation in low-literacy or highly stigmatized populations

---

6. Recommendations for Practice and Research

Practice:

- Use routinely in primary care, school-based clinics, and community health.
- Combine with clinical interviews or SCID for diagnostic confirmation.
- Integrate into depression care pathways and follow-up protocols.

Research:

- Examine longitudinal predictive validity for outcomes (e.g., suicidality, functional decline).
- Validate in underrepresented populations (e.g., LGBTQ+, Indigenous, migrant youth).
- Explore digital administration and cross-platform integration.

Digital Repository Format

Measure: Patient Health Questionnaire – Adolescent (PHQ-A)
Type: Symptom Screening (Depression)
Languages: English, Chinese, Thai, Swahili, Nepali, Spanish, Hindi
Validated Populations: U.S. adolescents, Chinese, Thai, Kenyan, Nepalese, Spanish-speaking,

primary care, psychiatric and school populations

Cutoffs:

- PHQ-A ≥10 = Moderate Depression
- PHQ-A ≥15 = Moderately Severe Depression

Psychometrics:

- Internal consistency: α = 0.84–0.92
- AUC = 0.88 (vs. clinical diagnosis)

Settings: Primary care, adolescent psychiatry, schools, community mental health

Use Case: Screening, severity assessment, and treatment monitoring for adolescent depression

Patient Mania Questionnaire-9 (PMQ-9)[205-206]

1. Objective

To develop and evaluate a brief, valid, and pragmatic self-report measure of manic symptoms for use in routine clinical care—especially in integrated and primary care settings—paired with the PHQ-9 for bipolar disorder monitoring.

---

2. Measure Evaluated

Patient Mania Questionnaire–9 (PMQ-9) – A 9-item, patient-reported outcome measure assessing the frequency of manic/hypomanic symptoms over the past week. Uses the same 4-point Likert response scale (0–3) and scoring system as the PHQ-9. Total scores range from 0–27, with higher scores indicating more severe manic symptoms.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- **Internal Consistency:**
  - Cronbach's α = 0.88 in participants with psychiatrist-diagnosed bipolar disorder
- **Test-Retest Reliability:**
  - Pearson correlation = 0.85 (excellent)
- **Concurrent Validity:**
  - PMQ-9 vs. Internal State Scale (Activation subscale): Pearson r = 0.70
  - PMQ-9 vs. Altman Self-Rating Mania Scale: r = 0.26 (likely due to differences in context and symptom timeframe)
- **Factor Structure:**
  - Confirmatory factor analysis showed unidimensional structure distinct from PHQ-9
- **Sensitivity to Change:**

- Established; minimally important difference (MID) estimated at 3 points (range 2–4)

Pragmatic Properties:

- **Acceptability:**
  - Rated highest among 8 clinician-assessed bipolar disorder measures for acceptability and usefulness by a diverse panel (psychiatrists, psychologists, PCPs, and social workers)
- **Feasibility:**
  - Easy to score and interpret alongside the PHQ-9; suitable for primary care teams and telehealth use
- **Interpretability:**
  - Can be used with PHQ-9 to classify mood states (e.g., high depression/high mania)
  - Score ≥10 indicates high manic symptoms
- **Equity:**
  - Designed for use in safety-net primary care settings and inclusive of low-resource populations
- **Sustainability:**
  - Used in large, federally funded studies (e.g., SPIRIT trial); applicable in routine measurement-based care workflows

---

4. Application to Integrated Primary Care

The PMQ-9 enables primary care and behavioral health providers to monitor manic symptoms in patients with bipolar disorder—especially in contexts where the PHQ-9 is already integrated. The measure supports shared decision-making, treatment adjustments, and detection of relapse or mixed mood states. Its compatibility with collaborative and telehealth models makes it particularly suited for integrated primary care systems.

---

5. Strengths and Limitations

Strengths:

- Mirrors PHQ-9 in format and scoring, easing clinical implementation
- Strong psychometric properties (reliability, validity, sensitivity to change)
- Supported by clinician and patient preferences
- Effective in measurement-based care and longitudinal tracking

Limitations:

- Still undergoing wider dissemination and cross-cultural validation
- Less suited to detect acute mania in inpatient/emergency settings
- Limited comparative studies against clinician-administered structured interviews

6. Recommendations for Practice and Research

Practice:

- Use PMQ-9 in conjunction with PHQ-9 to classify and monitor mood states in patients with diagnosed or suspected bipolar disorder
- Incorporate into collaborative care models, including primary care and behavioral health integration
- Train diverse clinical team members on interpretation and follow-up protocols

Research:

- Validate PMQ-9 across racial/ethnic populations and in adolescent or geriatric cohorts
- Examine predictive validity for relapse or functional decline
- Test cross-cultural adaptations and translations for global primary care use

---

Digital Repository Format

- **Measure:** Patient Mania Questionnaire–9 (PMQ-9)
- **Type:** Symptom Monitoring (Mania)
- **Languages:** English (additional translations under development)
- **Validated Populations:** U.S. adults with bipolar disorder in primary care
- **Cutoffs:**
  - PMQ-9 ≥10 = High manic symptom severity
  - MID = 3 points
- **Psychometrics:** α = 0.88; test-retest r = 0.85; concurrent validity (r = 0.70)
- **Setting:** Primary care, telepsychiatry, behavioral health integration
- **Use Case:** Ongoing monitoring of manic symptoms in individuals with bipolar disorder, used in tandem with PHQ-9 for full mood state tracking.

Patient-Reported Outcomes Measurement Information System (PROMIS)[207-215]

1. Objective

To assess the psychometric quality, cross-population validity, and clinical utility of the PROMIS item banks and short forms for evaluating physical, emotional, and social health across diverse populations and clinical settings, especially primary care, behavioral health, and chronic illness management.

---

2. Measure Evaluated

PROMIS – A suite of validated self-report instruments developed by the NIH to measure patient-reported health domains such as depression, anxiety, pain interference, physical functioning,

fatigue, sleep disturbance, and social participation. Available as computerized adaptive tests (CAT) and fixed-length short forms.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

Internal Consistency:

- PROMIS Depression short forms (4-, 6-, 8-items): Cronbach's $\alpha > 0.90$ across White, Black, Latino, and Asian American samples
- PROMIS Physical Function: ICC = 0.98, SEM = 2.0 in Dutch–Flemish sample
- PROMIS-57 Norwegian version: $\alpha = 0.84–0.95$ across domains

Construct Validity:

- PROMIS Depression and Anxiety scales highly correlated with legacy measures like PHQ-9 and GAD-7.
- Strong convergent validity in patients with HIV, chronic pain, cancer, and multiple chronic conditions

Discriminant Validity:

- PROMIS physical function differentiated between cancer patients of varying stages and conditions .
- PROMIS Depression scores distinguished depression severity in chronic pain patients.

Pragmatic Properties:

Acceptability:

- Developed with extensive stakeholder input, widely accepted in both clinical and research settings.
- Preferred over legacy tools for its clarity, adaptability, and cultural sensitivity.

Feasibility:

- Short forms (4–10 items) and CAT versions require minimal time and adapt to respondent needs.
- Integrated in EHRs and supported by the PROMIS HealthMeasures toolkit.

Interpretability:

- T-scores standardized to mean = 50, SD = 10.
- Higher scores represent greater severity in symptom domains or better functioning, depending on the scale.

Equity:

- Extensively translated and validated across languages and cultural contexts (Dutch–Flemish, Norwegian, ASL, Spanish, etc.)

Sustainability:

- Supported by the NIH, PROMIS tools are continuously updated, freely available, and promoted for national use in clinical care and population health tracking.

---

4. Application to Integrated Primary Care

PROMIS offers a standardized framework to screen and monitor multiple domains of health—mental, physical, and social—in diverse primary care populations. Its modular approach allows targeting specific symptoms (e.g., depression, fatigue, pain) while facilitating shared decision-making, interdisciplinary referrals, and tracking treatment response.

PROMIS is particularly helpful in:

- Behavioral health integration
- Chronic disease management (e.g., HIV, cancer, diabetes)
- Geriatric and palliative care settings
- Underserved populations through culturally validated versions

---

5. Strengths and Limitations

Strengths:

- NIH-developed and psychometrically superior to legacy tools
- Efficient, adaptable (CAT/short forms), and patient-centered
- Universally interpretable via T-scores
- Broad applicability across languages, ages, and conditions

Limitations:

- Requires digital infrastructure for CATs
- Some domains (e.g., pain behavior) may require further validation across cultures
- Norms may vary across subgroups—local calibration may be necessary

---

6. Recommendations for Practice and Research

Practice:

- Integrate PROMIS CATs or short forms in EHR workflows for physical and behavioral health.
- Use for routine outcome monitoring, shared decision-making, and quality improvement.
- Apply culturally validated versions in multilingual care teams.

Research:

- Investigate longitudinal responsiveness in low-income or rural populations.
- Validate cross-cultural equivalence in ASL, indigenous languages, and underrepresented U.S. ethnic groups.
- Examine predictive utility of PROMIS for hospitalization, morbidity, and cost-of-care outcomes.

Digital Repository Format

- **Measure:** Patient-Reported Outcomes Measurement Information System (PROMIS)
- **Type:** Comprehensive Health Domains (depression, anxiety, pain, fatigue, function, sleep)
- **Languages:** English, Spanish, Dutch, Norwegian, ASL, Flemish, and more
- **Validated Populations:** Chronic pain, cancer, HIV, older adults, primary care, Deaf, diverse U.S. ethnic groups
- **Cutoffs:** T-score > 60 (elevated symptoms); T-score < 40 (low function)
- **Psychometrics:** α > 0.90 across most domains; cross-cultural validity established
- **Setting:** Primary care, specialty care, research, behavioral health, population health
- **Use Case:** Symptom tracking, risk stratification, outcome monitoring, and treatment planning across diverse populations

Pediatric Symptom Checklist-17 (PSC-17)[216-225]

1. Objective

To evaluate the psychometric performance, cross-cultural validity, and practical utility of the PSC-17 as a brief parent-report screener for emotional and behavioral difficulties in children ages 4–15, within U.S. and international pediatric primary care and community-based settings.

---

2. Measure Evaluated

Pediatric Symptom Checklist-17 (PSC-17) – A 17-item, parent-completed tool assessing psychosocial functioning in children. It includes three subscales:

- **Internalizing** (5 items)
- **Externalizing** (7 items)
- **Attention Problems** (5 items)
  Each item is rated as:
  0 = "Never," 1 = "Sometimes," 2 = "Often."
  Cut-off for total score: **≥15 indicates psychosocial impairment**.

---

3. PAPERS Framework Evaluation

Psychometric Properties

Internal Consistency:

- U.S. national outpatient sample: Cronbach's α = 0.89 (Total); Test–retest reliability = 0.85
- Urban primary care (Philadelphia): Internalizing = 0.70, Externalizing = 0.84, Attention = 0.67

Construct Validity:

- Confirmatory factor analysis supported the 3-subscale structure across diverse samples.

- Cross-cultural findings indicate partial validity concerns in subscale loadings among urban minority populations.

Criterion Validity:

- Correlates with ADHD, depression, and behavioral disorder diagnoses.
- Compared favorably with longer tools like PSC-35 and structured interviews (e.g., K-SADS).

Pragmatic Properties

Acceptability:

- Widely accepted across pediatric and school-based settings. Endorsed by Medicaid and national programs like Head Start.

Feasibility:

- Takes <5 minutes to complete; suitable for integration in paper or electronic workflows (e.g., CHADIS platform).

Interpretability:

- Cut-off score ≥15 flags need for further assessment. Subscale cut-offs:
  - Internalizing: ≥5
  - Externalizing: ≥7
  - Attention: ≥7

Equity:

- Validated and translated in multiple cultural contexts including Spanish (Spain, U.S. Latinx), Korean, Sinhala (Sri Lanka), and Swahili (Kenya), with mixed findings regarding subscale structure equivalency.
- Some differences in item performance across racial/ethnic and socio-economic groups warrant culturally informed interpretation.

Sustainability:

- Routinely embedded in pediatric screening guidelines and widely adopted in primary care, school-based health, and global child health initiatives.

---

4. Application to Integrated Primary Care

The PSC-17 serves as a valuable behavioral health screener in pediatric primary care, supporting early identification of children at risk for emotional or behavioral disorders. Its integration into EHR systems (e.g., CHADIS) allows for real-time triage and referral to mental health professionals. Effective in medical homes and community outreach programs addressing behavioral health disparities.

---

5. Strengths and Limitations

Strengths:

- Brief, validated, and scalable across systems
- Subscale breakdown enhances specificity for targeted referrals
- High reliability and clinical utility

Limitations:

- Subscale validity may vary in low-income, minority populations
- Potential misclassification due to context effects in abbreviated format
- Limited sensitivity for anxiety disorders

---

6. Recommendations for Practice and Research

Practice:

- Use routinely in pediatric checkups, school-based health centers, and community screening.
- Follow up high PSC-17 scores with diagnostic assessments (e.g., structured interviews or broadband behavioral tools).

Research:

- Further evaluate dimensional and cross-cultural validity in diverse and underserved populations.
- Explore predictive validity for mental health treatment outcomes and long-term academic or behavioral trajectories.

Digital Repository Format

- **Measure:** Pediatric Symptom Checklist–17 (PSC-17)
- **Type:** Behavioral Health Screening (Parent-Report)
- **Languages:** English, Spanish, Korean, Swahili, Sinhala
- **Validated Populations:** U.S. general pediatric, African American, Latino, Korean, HIV-infected youth, low-income communities
- **Cutoffs:** Total Score ≥15 = clinical concern; subscale-specific thresholds apply
- **Psychometrics:** α = 0.89 (Total); Confirmed 3-factor structure; moderate validity in urban populations
- **Setting:** Pediatric primary care, school-based clinics, low-resource community health
- **Use Case:** Screening for internalizing, externalizing, and attention difficulties in pediatric behavioral health integration initiatives

PTSD Checklist for DSM-5 (PCL-5)[226-235]

1. Objective

To evaluate the psychometric validity, reliability, diagnostic utility, and international applicability of the PCL-5 for identifying posttraumatic stress disorder (PTSD) in clinical and nonclinical

populations, including primary care, low- and middle-income countries (LMICs), patients with comorbid conditions (e.g., chronic pain, HIV), and those undergoing trauma-focused treatment.

---

2. Measure Evaluated

PCL-5 (PTSD Checklist for DSM-5) – A 20-item self-report tool aligned with DSM-5 PTSD criteria, covering four symptom clusters: intrusion, avoidance, negative alterations in cognition/mood, and hyperarousal. Respondents rate each item based on the past month using a 0–4 Likert scale. Versions include full-length, abbreviated (4- and 8-item), and "past-day" versions adapted for massed treatments.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

• Internal consistency:
  – α = 0.90–0.96 (full version) across multiple populations (Brazilian, Zimbabwean, Danish, U.S.)
• Test–retest reliability:
  – ICC = 0.84–0.94 (full and abbreviated versions)
• Construct validity:
  – Confirmatory factor analysis supports DSM-5 4-factor and 7-factor hybrid models across settings
• Convergent validity:
  – High correlations with clinician-administered PTSD scale (CAPS-5), MADRS, PSS, and depression screeners
• Diagnostic accuracy:
  – AUC = 0.78–0.84 (Brazil, Denmark, Zimbabwe); sensitivity ≈ 74–85%, specificity ≈ 70–90%

Pragmatic Properties:

• **Acceptability:** Highly acceptable; used across military, primary care, trauma clinics, and LMICs
• **Feasibility:** Self-administered in <10 min; literacy-sensitive options needed for some populations
• **Interpretability:** Recommended cutoff = 31–33; abbreviated versions suggest cutoffs of 7 (4-item) or 13 (8-item)
• **Equity:** Validated across racial, cultural, and medical populations (e.g., Black Americans, HIV-positive individuals, Danish chronic pain patients, Zimbabweans)
• **Sustainability:** Extensively used in research, clinical trials, FQHCs, VAs, and global mental health initiatives

---

4. Application to Integrated Primary Care

The PCL-5 is a core tool in trauma-informed primary care, particularly effective for identifying PTSD among socioeconomically vulnerable and medically complex patients (e.g., HIV, TBI, chronic pain). It supports early intervention, facilitates mental health referrals, and provides measurable outcomes for trauma-focused therapy. The "Past Day" version enables daily tracking in massed or intensive PTSD treatment settings.

---

5. Strengths and Limitations

Strengths:

- Strong psychometric performance across languages, countries, and clinical populations
- Abbreviated and daily-use versions are validated
- Highly correlated with gold-standard CAPS-5 and clinician-administered tools
- Widely adopted in global and underserved settings

Limitations:

- Full version may be too long for low-literacy or fast-paced settings without adaptation
- Limited use among pediatric populations
- Requires clinician interpretation for diagnosis – not a standalone diagnostic tool
- Cultural variability in response to symptom expressions

---

6. Recommendations for Practice and Research

Practice:

- Use PCL-5 (cutoff ≥33) in primary care to screen trauma-exposed adults
- Apply "Past Day" PCL-5 in intensive or massed trauma treatment
- Use abbreviated forms when time or literacy is constrained
- Incorporate into trauma-informed workflows with follow-up care protocols

Research:

- Continue psychometric studies in LMICs, especially among women and youth
- Explore cultural adaptation needs and linguistic nuance in symptom reporting
- Validate further with comorbidities (e.g., TBI, chronic illness, perinatal populations)
- Study longitudinal responsiveness to treatment in diverse real-world contexts

---

Digital Repository Format

Measure: PTSD Checklist for DSM-5 (PCL-5)

Type: Symptom Severity Screening + Diagnostic Support

Languages: Multilingual (English, Spanish, Portuguese, Shona, Danish, etc.)

Validated Populations: Primary care patients, HIV+ individuals, chronic pain, TBI, veterans, trauma-exposed civilians, LMIC settings (e.g., Zimbabwe, Brazil, Mozambique)

Cutoffs: Full version ≥33; 8-item ≥13; 4-item >7

Psychometrics: α = .90–.96; ICC = .84–.94; AUC = .78–.84; Sensitivity 74–85%, Specificity 70–

90%
Setting: Primary care, trauma clinics, LMICs, FQHCs, Veterans Affairs, massed treatment programs
Use Case: Probable PTSD diagnosis, treatment planning, trauma symptom monitoring, research trials

Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE)[236-240]

1. Objective
To evaluate the reliability, validity, clinical implementation challenges, and health outcome relevance of PRAPARE as a standardized tool to assess social determinants of health (SDoH) among diverse patient populations across federally qualified health centers (FQHCs), community health centers (CHCs), and specialty clinics, with specific attention to its impact on care coordination, chronic disease management, and health equity.

---

2. Measure Evaluated
PRAPARE Tool – A standardized social risk assessment tool developed by the National Association of Community Health Centers (NACHC) to evaluate core SDoH domains: housing, employment, income, education, insurance, transportation, stress, and social support, among others. The tool contains 21 core and optional items and is available in over 20 languages. It is integrated into major EHRs and widely used in Medicaid, FQHC, and equity-focused initiatives.

---

3. PAPERS Framework Evaluation
Psychometric Properties:
• Internal consistency: KR-20 = 0.76 (cardiac patients); item correlations with PHQ-9 and functional outcomes support validity
• Test-retest reliability: r = 0.88 over short-term periods in clinical trials (heart failure, CHD)
• Factor Structure: 3 composite clusters (social background, social insecurities, insurance/employment) and 3 stand-alone (housing, isolation, poverty); construct validity demonstrated using EFA/CFA in >11,000 patients
Pragmatic Properties:
• **Acceptability:** High acceptability among patients; 81.8% of clinicians using PRAPARE found it helpful
• **Feasibility:** Successfully implemented across FQHCs, CHCs, and oncology/OB clinics; workflow integration with CHWs and case managers supported efficient use
• **Interpretability:** Cluster-based and item-level scoring; risk level stratification allows tailored referrals

• **Equity:** Culturally relevant across populations (e.g., Hispanic, NHPI, African American, pregnant women, gynecologic oncology)
• **Sustainability:** Supported by national initiatives; embedded in EHR platforms (Epic, eClinicalWorks) and used by Medicaid managed care organizations

---

4. Application to Integrated Primary Care

PRAPARE is broadly implemented in integrated primary care, OB/GYN, pediatrics, and specialty settings to identify unmet social needs and inform referrals. The tool enhances multidisciplinary care coordination, supports patient-centered interventions, and promotes health equity. Its utility is pronounced in CHCs serving diverse, low-income populations, and it has been shown to correlate with outcomes like delayed prenatal care, chronic disease control, and access to COVID-19 testing.

---

5. Strengths and Limitations

Strengths:

• Validated across diverse populations and chronic disease cohorts
• Integrates with EHRs for real-time care coordination
• Supports both individual risk identification and population-level social risk stratification
• Correlates with critical outcomes (e.g., diabetes, hypertension, prenatal care delays)

Limitations:

• Implementation barriers include staff training, lack of referral pathways, and time constraints
• Not originally developed for research; variable data completeness depending on administration protocol
• Limited validation in non-FQHC systems or international populations

---

6. Recommendations for Practice and Research

Practice:

• Use PRAPARE as part of routine intake or annual wellness visits in primary care and OB/GYN settings
• Employ community health workers or case managers to administer and respond to PRAPARE screenings
• Create referral workflows and resource databases to act on identified social needs

Research:

• Further psychometric validation in non-FQHC populations and specialty care (e.g., oncology, cardiology)
• Develop and test implementation strategies to improve fidelity and clinician training
• Explore longitudinal impacts of PRAPARE-informed care on health disparities and outcomes

---

Digital Repository Format
Measure: PRAPARE (Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences)
Type: Social Risk Screening + Care Coordination Support
Languages: Multilingual (English, Spanish, Marshallese, and 20+ others)
Validated Populations: Medicaid, FQHC patients, pregnant women, patients with chronic diseases (diabetes, hypertension, CHD, HF), Black and Hispanic communities
Cutoffs: No fixed cutoff; total scores or cluster-based risk classification used for care targeting
Psychometrics: KR-20 = 0.76; Test-retest reliability r = 0.88; construct validity supported by EFA/CFA
Setting: Primary care, OB/GYN, cardiology, community health centers, behavioral health
Use Case: Identifying and addressing social needs, risk stratification for chronic disease, prenatal care planning, population health management

Quality of Life Scale (QoLS)[241-249]

1. Objective
To assess the general quality of life across physical, psychological, and social domains using a global, non-disease-specific, patient-reported measure. The QoLS is widely used in chronic illness populations and in aging and rehabilitation contexts to evaluate subjective well-being.

---

2. Measure Evaluated
Quality of Life Scale (QoLS) – Originally developed by John Flanagan in 1978, the QoLS was later adapted for healthcare and research purposes. It is a 16-item or 15-item self-report instrument (depending on version), with items rated on a 7-point Likert scale (1 = terrible to 7 = delighted), evaluating domains such as material well-being, health, social relationships, and personal development.

---

3. PAPERS Framework Evaluation
Psychometric Properties:

- **Internal Consistency:**
  - Cronbach's α ranges from 0.82 to 0.92 across various populations, including those with chronic illness and older adults.
  - In Alzheimer's caregivers and health professionals, α = 0.82 (Spanish version) .
- **Reliability in Specific Populations:**
  - Alzheimer's patients, caregivers, and clinicians show strong agreement with QoLS ratings, supporting its reliability in dementia contexts.

    - Swedish and Dutch studies reported stable test–retest reliability and internal consistency above 0.80 for RA and SLE patients.
- **Construct Validity:**
    - Principal component and confirmatory factor analyses support a one- or two-factor structure (e.g., personal development and social functioning) across cultural groups.
    - Convergent validity is supported by positive correlations with well-being indices, ADLs, and negative correlations with depressive symptoms.

Pragmatic Properties:

- **Acceptability:** High acceptability across clinical, community, and aging populations; low literacy burden due to simple item phrasing.
- **Feasibility:** Self-administered in 5–7 minutes. Appropriate for in-person, paper, or electronic administration.
- **Interpretability:**
    - Higher scores = better perceived quality of life. No universally accepted cutoff, but change scores ≥ 10% are used to indicate meaningful improvement in intervention studies.
- **Equity:**
    - Translated and validated in Swedish, Spanish, Dutch, and other languages. Adapted versions demonstrate strong psychometric stability across age, gender, chronic illness, and caregiving roles.
- **Sustainability:** Used internationally across geriatric, mental health, chronic illness, rehabilitation, and caregiver support settings.

---

4. Application to Integrated Primary Care

QoLS is a useful tool for assessing overall well-being, particularly among patients with chronic conditions, caregivers, and older adults. It complements disease-specific assessments in integrated primary care by capturing social and psychological health. The scale supports behavioral health integration and care coordination for populations with complex psychosocial needs.

---

5. Strengths and Limitations

Strengths:

- Brief, generalizable across disease states
- Strong internal consistency and cultural adaptability
- Captures multidimensional aspects of QoL not addressed by symptom-based scales

Limitations:

- Does not capture acute clinical symptoms (e.g., depression or anxiety)

- May require contextual adaptation for younger populations
- Some cultural variability in domain salience (e.g., religious participation vs. personal autonomy)

---

6. Recommendations for Practice and Research

Practice:

- Include QoLS as part of comprehensive biopsychosocial assessments for chronic illness, geriatrics, and integrated care teams.
- Use in routine follow-up for evaluating life satisfaction and rehabilitation outcomes.

Research:

- Further evaluate QoLS responsiveness in diverse patient populations (e.g., non-Western, youth, and disability).
- Explore item response theory (IRT) methods to optimize scoring precision.
- Investigate cultural domain weighting to improve contextual sensitivity.

Digital Repository Format

- **Measure:** Quality of Life Scale (QoLS)
- **Type:** Patient-Reported Outcome – Quality of Life
- **Languages:** Spanish, Swedish, Dutch, English, others
- **Validated Populations:** Adults, older adults, patients with chronic illness (RA, SLE, Alzheimer's), caregivers
- **Cutoffs:** No standardized cutoff; higher scores indicate higher perceived QoL
- **Psychometrics:** Cronbach's α = 0.82–0.92; validated factor structure
- **Setting:** Primary care, geriatrics, chronic illness management, caregiving, mental health
- **Use Case:** Assessing subjective quality of life in patient-centered care models and integrated behavioral health settings

Screening for Child Anxiety Related Disorders (SCARED)[250-255]

1. Objective

To assess the psychometric properties, cultural adaptability, and clinical relevance of the SCARED as a brief self-report and parent-report screening tool for detecting anxiety-related disorders in children and adolescents (ages 8–18), across global and U.S. primary care, school, and integrated care settings.

2. Measure Evaluated

Screen for Child Anxiety Related Emotional Disorders (SCARED) – A 41-item (original) or 38-item (revised) questionnaire measuring five anxiety domains: somatic/panic, generalized

anxiety, separation anxiety, social phobia, and school phobia. Available in child self-report and parent-report formats. Responses are scored from 0 ("not true") to 2 ("very true/often true"), with higher scores indicating greater symptom severity.

3. PAPERS Framework Evaluation

Psychometric Properties:

- **Internal Consistency:**
  - Cronbach's α across studies = 0.74–0.96
  - Subscale reliabilities typically >0.70
  - Spanish sample (clinical): α = 0.93 (total)
  - Korean version: total α = 0.95; subscales range from 0.75–0.90
- **Test-Retest Reliability:**
  - Moderate to excellent agreement over 2–4 weeks (r = 0.60–0.90)
- **Construct Validity:**
  - Confirmatory factor analysis supports the 5-factor structure in multiple settings
  - Adapted versions (e.g., Korean, Dutch, Sinhala) show structural equivalence with minor cultural variation
- **Discriminant Validity:**
  - Effectively distinguishes between anxiety and other psychiatric disorders (e.g., ADHD, depression)
  - Youth self-report tends to yield higher anxiety levels than parent-report
- **Sensitivity/Specificity:**
  - Parent version: Sensitivity = 0.86, Specificity = 0.74 in African American and Non-Hispanic White youth
  - Cross-informant agreement (child vs. parent): κ = 0.20–0.58

Pragmatic Properties:

- **Acceptability:**
  - Routinely used in school-based screenings, pediatric and primary care, and behavioral health evaluations
  - Recognized for high engagement from adolescents and caregivers
- **Feasibility:**
  - Completion time: ~10 minutes
  - Suitable for self-administered or interviewer-assisted formats
- **Interpretability:**
  - A total score ≥ 25 suggests the presence of an anxiety disorder
  - Subscale-specific cutoffs guide differential diagnosis (e.g., social phobia, GAD)
- **Equity:**

    - Successfully translated and validated in over a dozen languages: Spanish, Korean, Sinhala, Dutch, Iranian Persian, French, Swahili, and more
    - Demonstrated cultural sensitivity in format and structure across regions (e.g., Malta, Sri Lanka, Korea)
- **Sustainability:**
    - Broadly implemented in research, primary care, community health, and school systems worldwide
    - Free to access and widely adaptable for integrated care models

---

4. Application to Integrated Primary Care

SCARED facilitates early detection of anxiety in youth within primary care and school-linked behavioral health models. Its child and parent versions enable cross-informant comparison, supporting triage to psychological services. Especially valuable in integrated pediatric settings and for tailoring anxiety-focused interventions in educational or low-resource environments.

---

5. Strengths and Limitations

Strengths:

- Strong internal consistency and validity across diverse cultural contexts
- Identifies multiple anxiety subtypes
- Available in multiple languages; suitable for school and clinical use
- Informant flexibility (child and parent forms)
- Freely accessible, non-proprietary

Limitations:

- Informant discrepancies (parent vs. child scores) may limit diagnostic accuracy without clinician input
- Possible overlap in symptom presentation with depressive disorders
- Some adaptations report factor instability (e.g., fewer than five domains)
- Limited sensitivity for very young children (under age 8)

---

6. Recommendations for Practice and Research

Practice:

- Use SCARED in annual behavioral health screening in pediatric and school-based primary care settings
- Implement both child and parent versions when possible, to account for symptom perception differences
- Apply culturally adapted versions for multilingual or immigrant youth populations

Research:

- Investigate SCARED's predictive validity for long-term mental health outcomes

- Further explore structural invariance across racial, ethnic, and gender identities
- Validate shorter SCARED versions (e.g., SCARED-5) for rapid screening without loss of diagnostic utility

---

Digital Repository Format
Measure: Screen for Child Anxiety Related Emotional Disorders (SCARED)
Type: Symptom Screening (Child Anxiety)
Languages: Multilingual (English, Spanish, Korean, French, Persian, Sinhala, Swahili, Dutch, etc.)
Validated Populations: U.S., Korean, European, South Asian, Middle Eastern, African descent, clinical and community youth
Cutoffs:

- Total score ≥ 25 = Clinically significant anxiety
- Subscale-specific thresholds vary by version
  Psychometrics: Cronbach's α = 0.74–0.96; strong construct and discriminant validity
  Setting: Primary care, integrated behavioral health, schools, pediatric outpatient clinics
  Use Case: Screening and monitoring of anxiety in children and adolescents across diverse systems of care

Social Responsive Scale (SRS)[256-265]

1. Objective

To evaluate the psychometric properties, cross-cultural utility, and clinical relevance of the Social Responsiveness Scale (SRS) as a quantitative tool for assessing autistic traits across developmental stages, diverse populations, and cultural contexts.

---

2. Measure Evaluated

Social Responsiveness Scale (SRS) – A 65-item parent, teacher, or self-report questionnaire designed to measure the severity of social impairment associated with Autism Spectrum Disorders (ASD). Each item is rated on a 4-point Likert scale (0 = Not True to 3 = Almost Always True), providing a total score reflecting the degree of social reciprocity difficulty.

---

3. PAPERS Framework Evaluation
Psychometric Properties
Internal Consistency:

- Cronbach's α ranges from 0.81–0.97 across U.S., German, Japanese, Korean, Chinese, and Vietnamese samples.

- German SRS: α = 0.91–0.97 for general and clinical samples.

Test-Retest Reliability:

- ICC ranges from 0.81–0.96 across international validations.
- German sample: ICC = 0.84–0.97.

Inter-Rater Reliability:

- Parent–teacher ICC = 0.75–0.91 in U.S. and German samples.

Construct Validity:

- Correlations with ADOS, ADI-R, and SCQ: r = 0.35–0.70.
- Consistent associations with CBCL, VABS, and temperament inventories support validity.

Discriminant Validity:

- Differentiates ASD from ADHD, mood disorders, and typical development in both Western and Asian contexts.

Factor Structure:

- Typically supports a one-factor model across multiple studies and cultures.

Pragmatic Properties

Acceptability:

- Widely used in clinics, schools, and research for early identification of ASD traits.

Feasibility:

- 15–20 minutes to complete; easily scored by hand or software.

Interpretability:

- U.S. cut-off: T-score ≥ 60 = clinically significant social impairment.
- Country-specific cut-offs:
    - Korean children: T > 52
    - Japanese adults: T > 65
    - Chinese children: T > 56.5

Equity:

- Validated in Germany, Japan, Korea, China, Vietnam, France, the U.K., and the U.S.
- Cultural adaptation through translation, back-translation, and item-level modification supported comparability.

Sustainability:

- Embedded in many autism intervention programs and screening protocols globally.

---

4. Application to Integrated Primary Care

The SRS enables routine screening for autism traits in pediatric and adolescent primary care settings. Its sensitivity to both subclinical and clinical ASD presentations supports early referral to behavioral health and developmental services. Its versions for adults also allow integration

into neurodevelopmental and psychiatric screening within adult care systems. Cultural adaptations increase appropriateness in global health and multilingual clinical environments.

---

5. Strengths and Limitations

Strengths:

- High reliability and validity across cultures and age groups
- Sensitive to a continuum of autism traits, not just diagnostic thresholds
- Translated and validated in over 10 countries
- Versions available for preschoolers, school-age children, and adults

Limitations:

- May conflate social anxiety or social communication disorders with ASD
- Norms vary by culture, requiring local validation
- Less effective in detecting non-autistic social impairments

---

6. Recommendations for Practice and Research

Practice:

- Incorporate SRS into primary care and school-based developmental screenings
- Adjust scoring thresholds to align with local validation studies
- Use in conjunction with diagnostic tools (e.g., ADOS, ADI-R) for comprehensive assessment

Research:

- Further examine SRS's specificity in distinguishing ASD from related conditions
- Longitudinal studies to assess SRS predictive value for developmental outcomes
- Expand validation in underrepresented populations (e.g., rural, non-Western, and minority communities)

Digital Repository Format

Measure: Social Responsiveness Scale (SRS)

Type: Symptom Screening (Autism Spectrum Disorder Traits)

Languages: English, German, Japanese, Korean, Chinese, Vietnamese, French, etc.

Validated Populations: U.S., Germany, Japan, Korea, China, Vietnam, and others

Cutoffs:

- U.S. T-score ≥ 60 = clinically significant
- Korean children: >52
- Japanese adults: >65
- Chinese children: >56.5

Psychometrics: Internal consistency α = 0.81–0.97; ICC = 0.81–0.96; Validated across

multiple cultures

Setting: Primary care, psychiatry, pediatrics, school health, developmental clinics

Use Case: Screening and monitoring for ASD traits in children, adolescents, and adults across diverse clinical and cultural contexts.

St. Louis University Mental Status Examination (SLUMS)[266-275]

1. Objective

To evaluate the reliability, diagnostic accuracy, and longitudinal utility of the SLUMS in detecting both dementia and mild neurocognitive disorder (MNCD), particularly in older adults and U.S. veteran populations, and to assess its equity across racially diverse populations and its predictive capacity for institutionalization and mortality.

---

2. Measure Evaluated

Saint Louis University Mental Status Examination (SLUMS) – An 11-item, 30-point cognitive screening tool developed to detect mild neurocognitive disorder and dementia. SLUMS assesses orientation, short- and long-term memory, executive function, attention, language, and reasoning. It is freely available and provides education-adjusted cutoff scores for identifying MNCD and dementia, distinguishing it from tools like the MMSE and MoCA. Adaptations and longitudinal analyses have further validated its use.

---

3. PAPERS Framework Evaluation

Psychometric Properties:

- Internal consistency: α = .71–.85 (across White, Black, and veteran samples)
- Test–retest reliability: ρ = .72 (1-year interval)
- Convergent validity: Strong correlations with MMSE, MoCA, and comprehensive neuropsychological batteries
- Diagnostic accuracy:
  - MNCD (education-adjusted cutoffs): Sensitivity = 92%, Specificity = 81%
  - Dementia detection: AUC = 0.927 (SLUMS) vs. 0.671 (MMSE)
- Factor Structure: Best supported as unidimensional (single-factor), though some items contribute uniquely to executive function or memory domains

Pragmatic Properties:

- **Acceptability:** Brief (~7–10 minutes), free to use, paper-based administration
- **Feasibility:** Routinely used in VA hospitals and long-term care; widely adopted in geriatrics
- **Interpretability:**
  - ≥27 = Normal (with high school), 21–26 = MNCD, ≤20 = Dementia

– Adjusted for education level (cutoffs lowered for < high school)

• **Equity:** Studies demonstrate racial disparities in classification (e.g., Black veterans 2x more likely to score in dementia range), underscoring the need for culturally responsive interpretation

• **Sustainability:** Longitudinal studies confirm SLUMS predicts institutionalization, cognitive decline, and mortality

---

4. Application to Integrated Primary Care

SLUMS is a practical tool for primary care, geriatrics, and community health settings, especially within the VA and long-term care systems. It supports early detection of MNCD, helps differentiate between cognitive and mood-related impairment, and informs treatment planning. Integration with EHRs and alignment with cognitive wellness protocols makes SLUMS useful for dementia care pathways, care coordination, and preventative screenings.

---

5. Strengths and Limitations

Strengths:

• Free and accessible, unlike MMSE and MoCA
• Detects both MNCD and dementia with high sensitivity
• Education-adjusted cutoffs improve equity
• Predictive of institutionalization and mortality
• Reliable over time and across diverse clinical samples

Limitations:

• Less psychometric research than MoCA/MMSE
• Racial disparities in score interpretation not fully addressed
• Underutilized in non-veteran and multicultural populations
• May not distinguish well between MNCD and depression without supplemental tools

---

6. Recommendations for Practice and Research

Practice:

• Use SLUMS in aging veterans, long-term care, and primary care patients at risk of cognitive impairment
• Apply education-adjusted scoring to reduce misclassification
• Combine with functional assessments and depression screening for improved diagnostic clarity
• Provide cultural competence training to reduce racial/ethnic misinterpretation

Research:

• Further validate SLUMS in non-veteran, multilingual, and racially diverse populations
• Compare performance with MoCA and MMSE in integrated care and community-based aging populations

• Develop digital or EHR-integrated SLUMS for scalable implementation
• Investigate item-level disparities contributing to race-based misclassification

---

Digital Repository Format
Measure: Saint Louis University Mental Status Examination (SLUMS)
Type: Cognitive Screening
Languages: English; validated in VA settings; adaptable for multilingual and low-literacy contexts
Validated Populations: Older adults, U.S. veterans, racially diverse cohorts, long-term care residents
Cutoffs: High school+: 27–30 (Normal), 21–26 (MNCD), ≤20 (Dementia); <High school: 25–30 (Normal), 20–24 (MNCD), ≤19 (Dementia)
Psychometrics: Sensitivity 92%, Specificity 81%, Reliability α = .71–.85, Test-retest ρ = .72
Setting: Primary care, geriatrics, VA hospitals, long-term care, community health clinics
Use Case: Early detection of mild cognitive impairment and dementia, risk stratification, outcome prediction

Vanderbilt ADHD Diagnostic Rating Scales (VADPRS/VADTRS)[276-285]

1. Objective
To evaluate the psychometric validity, clinical utility, and cross-cultural adaptability of the Vanderbilt ADHD Diagnostic Parent Rating Scale (VADPRS) and Teacher Rating Scale (VADTRS) for assessing ADHD and common comorbidities in children across community, clinical, and educational settings. These tools aim to facilitate evidence-based diagnosis, treatment planning, and monitoring in pediatric primary care.

---

2. Measure Evaluated
Vanderbilt ADHD Diagnostic Rating Scales – These DSM-IV-based tools include 18 core ADHD items (inattention, hyperactivity/impulsivity), 8 ODD items, 4 CD items, and 7 anxiety/depression items. Both parent and teacher versions include performance impairment items. Designed for use in children aged 6–12, the scales are structured, standardized, and freely available.

---

3. PAPERS Framework Evaluation
Psychometric Properties:
• Internal consistency:
  – ADHD Inattention: α = .89 (parent), .93 (teacher)
  – Hyperactivity/Impulsivity: α = .90–.94

– ODD: α = .91 (parent), .89 (teacher)
– Anxiety/Depression: α = .77–.84
• Test–retest reliability: ICC = .80–.91 across scales
• Construct validity: Confirmed 4-factor structure (ADHD-I, ADHD-HI, ODD, Anxiety/Depression) via CFA
• Criterion validity: Sensitivity = .80, Specificity = .75, NPV = .98 vs. structured clinical interview
• Predictive validity: Performance scores correlate with learning disorders; recommended cutoff of 7.5 rules out reading/spelling LDs

Pragmatic Properties:

• **Acceptability:** Widely used by pediatricians and schools; high clinician satisfaction
• **Feasibility:** Can be completed quickly (5–10 min); incorporated in AAP/NICHQ toolkit; widely adopted in EHRs
• **Interpretability:** DSM-IV-based scoring; clinical cutoffs established for ADHD types and comorbidities
• **Equity:** Validated in U.S., Czech, Greek, and Spanish-speaking populations; useful in both community and clinical samples
• **Sustainability:** Recommended in AAP ADHD Guidelines and CDC surveillance protocols; used in school-linked care pathways

---

4. Application to Integrated Primary Care

The Vanderbilt scales are central to the AAP's model for ADHD diagnosis in pediatric care. They help clinicians assess symptoms from multiple informants (parent and teacher), identify coexisting conditions (ODD, CD, anxiety, depression), and evaluate academic impairment. Their integration into EHRs supports collaborative care, referral coordination, and treatment monitoring in behavioral health integration models.

---

5. Strengths and Limitations

Strengths:

• High reliability and diagnostic utility across informants
• Measures comorbid conditions and functional impairment
• Widely accepted and accessible; supports AAP guidelines
• Supported by U.S. normative data and adapted for multiple languages

Limitations:

• Less sensitive for ruling in learning disorders (math LD) without additional testing
• Cutoffs for comorbid screens may lack specificity in diverse populations
• Performance items require contextual interpretation with academic data
• Age and symptom severity may influence factor loadings differently across cultures

---

6. Recommendations for Practice and Research

Practice:

- Use both parent and teacher versions to meet diagnostic criteria for ADHD
- Include performance items in learning disorder screenings
- Embed within ADHD care pathways and multidisciplinary evaluation teams
- Interpret comorbidity screens cautiously; use interviews for confirmation

Research:

- Continue validating versions for non-English-speaking and low-resource populations
- Examine dimensional scoring over categorical cutoffs for improved specificity
- Assess longitudinal sensitivity to treatment change in real-world pediatric settings
- Explore EHR-based decision support tools using Vanderbilt data

---

Digital Repository Format

Measure: Vanderbilt ADHD Diagnostic Rating Scales (VADPRS / VADTRS)

Type: Behavioral Health Symptom and Comorbidity Screening

Languages: English, Czech, Greek, Spanish, others under adaptation

Validated Populations: U.S. clinical/community samples, ADHD referrals, diverse ethnic and language groups

Cutoffs:

- ADHD: ≥6 symptoms in a domain + functional impairment
- ODD: ≥4 symptoms
- Anxiety/Depression: Item-specific and sum-score criteria explored

Psychometrics: Internal consistency α = .77–.94; ICC = .80–.91; Sensitivity = .80; Specificity = .75; NPV = .98

Setting: Pediatric primary care, school-linked mental health, developmental-behavioral pediatrics

Use Case: ADHD diagnosis, comorbidity screening, performance monitoring, support for school referrals

WHO Disability Assessment Schedule (WHODAS)[286-295]

1. Objective

To evaluate the psychometric properties, cross-cultural utility, and clinical relevance of the WHODAS 2.0 as a generic tool for assessing health and disability across physical, mental, and emotional conditions, in both clinical and general populations, globally and in integrated care contexts.

---

2. Measure Evaluated

WHODAS 2.0 – A standardized 12-, 36-, or 36+ item self-report instrument developed by the World Health Organization to assess functioning in six domains: cognition, mobility, self-care, getting along, life activities, and participation. Response options are based on a 5-point Likert scale from "none" to "extreme or cannot do."

---

3. PAPERS Framework Evaluation

Psychometric Properties:

Internal Consistency:

- **36-item version:** Cronbach's α = 0.94 in general population samples (e.g., Sweden, Nigeria, Turkey, Spain)
- **12-item version:** α ranges from 0.78–0.89 across cultural groups and patient samples (e.g., Turkish psychiatric patients, Igbo-speaking Nigerian pain patients, Spanish primary care)

Construct Validity:

- Strong correlations with related disability and mental health measures (e.g., PHQ-9, GAF, SF-36)
- Factor analysis supports the six-domain model across versions and languages

Test-Retest Reliability:

- ICC > 0.85 for 36-item version over 2–4 weeks in psychiatric and primary care populations

Discriminant Validity:

- Accurately differentiates between clinical vs. non-clinical groups and between varying severities of disability

---

Pragmatic Properties:

Acceptability: Widely accepted in primary care, mental health, public health, and international health monitoring.

Feasibility:

- Completion time: 5–20 minutes, depending on version
- Available in interviewer-administered, self-report, and proxy formats

Interpretability:

- Domain and total scores converted to a 0–100 scale; higher scores = greater disability
- Cutoffs are not universally standardized but are often tailored to context-specific needs (e.g., pain, depression, mobility)

Equity:

- Validated in >30 languages, including Turkish, Spanish, Igbo, Swedish, Hindi, Mandarin Chinese
- Applicable across age, literacy, and socioeconomic levels with culturally sensitive adaptations

Sustainability:

- Used globally in WHO health surveys, research studies, and integrated care settings. Free to use and updatable through WHO resources.

---

4. Application to Integrated Primary Care

WHODAS 2.0 allows providers to comprehensively assess patients' functional impairments across mental and physical domains, supporting biopsychosocial assessments. Its brevity, broad scope, and cultural adaptability make it ideal for team-based care models, disability evaluations, and tracking of recovery outcomes over time.

---

5. Strengths and Limitations

Strengths:

- Cross-culturally validated in multiple languages and settings
- Broadly applicable across conditions, including pain, depression, psychosis, and disability
- High reliability and strong correlations with functional and clinical measures
- Free, open-access tool backed by WHO

Limitations:

- Some items may require clarification in populations with low literacy or limited insight
- Time burden may be high for the 36-item version in busy clinical settings
- Cut-off points for clinical decision-making are not universally standardized

---

6. Recommendations for Practice and Research

Practice:

- Integrate WHODAS 2.0 into behavioral health screening in primary care, especially for chronic illness, disability, or psychiatric comorbidities
- Use the 12-item version for rapid screening, and the 36-item version for in-depth assessment or research

Research:

- Further validation needed in youth and non-Western populations
- Evaluate responsiveness to change in integrated behavioral healthcare interventions
- Explore the use of domain-specific scoring in functional goal-setting for care planning

---

Digital Repository Format

Measure: World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0)

Type: Functional Status Assessment
Languages: 30+ (Spanish, Turkish, Igbo, Swedish, Hindi, Mandarin, etc.)
Validated Populations: General population, psychiatric, pain, primary care, disability, global health
Cutoffs:

- Not standardized globally; typically use relative score percentiles or clinical judgment
  Psychometrics:
- 36-item α = 0.94; 12-item α = 0.78–0.89; excellent test-retest reliability
  Setting: Primary care, mental health, pain clinics, rehabilitation, international health surveys
  Use Case: Measuring health-related functioning in clinical care, research, disability assessment, or integrated care monitoring

Young Mania Rating Scale (YMRS)[296-301]

1. Objective
To evaluate the reliability, validity, and cross-cultural utility of the Young Mania Rating Scale (YMRS) for detecting and quantifying the severity of manic symptoms in adults and children across diverse populations, including low-resource and multicultural settings. Emphasis is placed on distinguishing bipolar disorder (BD) from other neuropsychiatric conditions such as ADHD and depression.

---

2. Measure Evaluated
Young Mania Rating Scale (YMRS) – An 11-item clinician-rated semi-structured tool originally developed to assess the severity of manic symptoms in adults. It incorporates both patient self-report and observational inputs. Certain items (e.g., elevated mood, sleep, thought content) are double-weighted to reflect their clinical significance. Versions validated for children (e.g., CMRS) and various languages and regions (e.g., Portuguese, Spanish, Kinyarwanda, Korean) extend the scale's applicability.

---

3. PAPERS Framework Evaluation
Psychometric Properties:
• Internal consistency: α = .67–.91 across versions (Portuguese, Spanish, Korean, Rwandese).
• Inter-rater reliability: Excellent (ICC = .93 for original; >0.80 for Portuguese and Rwandese versions).
• Validity: Strong concurrent validity with other mania/depression scales (e.g., BPRS, CGI-S, CMRS).

• Unidimensional factor structure supported across most studies; Rasch modeling affirmed construct validity.
• Pediatric and adult versions differentiated mania from ADHD, especially in controlled clinical samples.

Pragmatic Properties:

• **Acceptability:** Widely used and accepted in clinical trials and practice; suitable for inpatient/outpatient settings.
• **Feasibility:** Administered in 10–20 minutes by trained clinicians; adaptations with structured interviews improve usability.
• **Interpretability:** Scores range from 0–60; higher scores = greater mania severity; double-weighted items enhance specificity.
• **Equity:** Validated across diverse populations—Portuguese, Spanish, Korean, Rwandese, and pediatric samples.
• **Sustainability:** Endorsed globally; standard in psychiatry research; used in academic, government, and hospital systems.

---

4. Application to Integrated Primary Care

While the YMRS is primarily used in psychiatric care, it can be integrated into collaborative care and stepped care models in primary care to assess manic symptoms and monitor treatment response. It is particularly relevant for pediatric populations in integrated behavioral health systems. The scale's structured format enables standardization across multi-provider teams, and cultural adaptations make it usable in low-resource or multilingual clinical environments.

---

5. Strengths and Limitations

Strengths:

• Highly reliable and valid across diverse settings and populations.
• Supported by decades of clinical trial usage.
• Adaptable for pediatric and global use.
• Strong discriminative power for core manic symptoms (e.g., sleep, grandiosity, activity).

Limitations:

• Requires trained clinician for accurate administration.
• Double-weighting can skew interpretations if not used carefully.
• Less effective as a screening tool compared to structured diagnostic interviews.
• Some items (e.g., insight, irritability) show lower factor loadings or internal consistency.

---

6. Recommendations for Practice and Research

Practice:

• Use YMRS for structured assessment of manic episodes, particularly in psychiatric and

collaborative care settings.
• Train primary care clinicians or behavioral health providers in YMRS scoring protocols.
• Use culturally adapted versions (e.g., Rwandese, Portuguese, Korean) in relevant populations to ensure fidelity.

Research:

• Continue validation studies for low-literacy and non-English-speaking populations.
• Explore simplified formats or digital adaptations for integrated care.
• Examine the responsiveness of YMRS to treatment changes in diverse clinical trials and global mental health programs.

Digital Repository Format

Measure: Young Mania Rating Scale (YMRS)
Type: Symptom Severity Rating
Languages: Multilingual (English, Spanish, Portuguese, Korean, Kinyarwanda, etc.)
Validated Populations: Adults with BD, Children/adolescents with ADHD/BD, African, Asian, Latin American populations
Cutoffs: No universal cutoff; higher total scores (out of 60) indicate greater mania severity; items 5, 6, 8, and 9 are double-weighted
Psychometrics: Internal consistency α = .67–.91; Inter-rater reliability ICC > 0.90; AUC = 0.90 (child ADHD vs. mania)
Setting: Psychiatric care, pediatric psychiatry, integrated primary care, global health clinics
Use Case: Assessing and tracking manic symptoms in individuals with BD, distinguishing BD from ADHD, and measuring treatment response.

Youth Top Problems Assessment (YTPA)[302]

Objective

To evaluate the psychometric properties, cultural adaptability, and clinical applicability of the Youth Top Problems Assessment (YTP) as an idiographic, youth-centered mental health screening and progress monitoring tool, particularly in low-resource and global mental health settings.

---

2. Measure Evaluated

Youth Top Problems Assessment (YTP) – A flexible, open-ended tool in which children and adolescents identify and rate their top three psychological or functional concerns on a scale (typically 0–10) regarding their severity or impairment. Designed for use across diverse clinical and research contexts, the YTP emphasizes youth voice and individualized tracking of treatment response.

3. PAPERS Framework Evaluation

Psychometric Properties:

- Internal Consistency: Not applicable due to idiographic and non-standardized item content across respondents.
- Test–Retest Reliability: Moderate short-term reliability in repeated administrations (ICCs ~0.63–0.70 in Kenyan adolescents) despite individualized content.
- Construct Validity:
    - Converges with standardized distress and functioning measures (e.g., PHQ-9, SDQ, WHODAS).
    - Sensitively tracks changes in youth-identified priorities during treatment.
- Discriminant Validity: Effectively distinguishes between youth with clinically significant vs. subclinical mental health concerns.

Pragmatic Properties:

- **Acceptability:**
    - High among youth and providers, especially in LMIC and cross-cultural settings.
    - Recognized for enhancing therapeutic alliance and engagement.
- **Feasibility:**
    - Easy to administer, low cost, minimal training required.
    - Adaptable for group settings and task-shifting models.
- **Interpretability:**
    - Clinically meaningful changes were detected by tracking individual concern severity ratings over time.
- **Equity:**
    - Demonstrated cross-cultural utility (e.g., validated in Kenya).
    - Elicits culturally relevant stressors not captured in standardized tools.
- **Sustainability:**
    - Integrated into mental health interventions across schools, NGOs, and clinics in LMICs.
    - Recommended in global mental health toolkits (e.g., WHO and UNICEF resources).

---

4. Application to Integrated Primary Care

The YTP is well-suited for integrated behavioral health models within primary care, especially when tailoring services to youth from marginalized or culturally diverse backgrounds. Its idiographic nature supports shared decision-making and patient-centered care, making it ideal for brief interventions, school-based health centers, and mobile mental health teams.

---

5. Strengths and Limitations

Strengths:

- Prioritizes youth voice and agency in treatment planning
- Sensitive to change and responsive to culturally diverse expressions of distress
- Low-cost, low-literacy, and scalable
- Promotes patient engagement and individualized care

Limitations:

- Lack of standardization limits use for diagnostic comparisons
- Not suitable for population-level prevalence estimates
- Requires some clinician interpretation to guide structured treatment

---

6. Recommendations for Practice and Research

Practice:

- Integrate into intake assessments to guide personalized care
- Use longitudinally to monitor change in client-identified outcomes
- Apply in community and school-based behavioral health services

Research:

- Further validate in diverse linguistic and clinical populations
- Examine use alongside standardized screening tools for complementary insight
- Explore the impact on engagement, retention, and treatment outcomes

---

Digital Repository Format

Measure: Youth Top Problems Assessment (YTP)

Type: Idiographic Assessment (Self-Reported Priority Concerns)

Languages: Multilingual; adaptable for cultural and regional use (e.g., English, Swahili)

Validated Populations: Youth in the U.S., Kenya, and other LMICs; school-based and clinical samples

Cutoffs: No standardized cutoff; severity scores (0–10 per problem) interpreted individually and over time

Psychometrics: Good test–retest reliability; strong convergent validity; culturally relevant content

Setting: Primary care, schools, community mental health, LMIC global health interventions

Use Case: Youth-centered screening, treatment planning, and progress monitoring across settings and cultures

References

1. Ballmann C, Kölle MA, Bekavac-Günther I, et al. Evaluation of the German Version of the Adult Attention-Deficit/Hyperactivity Disorder Self-Report Screening Scale for DSM-5 as a Screening Tool for Adult Attention-Deficit/Hyperactivity Disorder in Primary Care. *Frontiers in Psychology*. 2022;13. doi:https://doi.org/10.3389/fpsyg.2022.858147
2. Hines JL, King TS, Curry WJ. The Adult ADHD Self-Report Scale for Screening for Adult Attention Deficit-Hyperactivity Disorder (ADHD). *The Journal of the American Board of Family Medicine*. 2012;25(6):847-853. doi:https://doi.org/10.3122/jabfm.2012.06.120065
3. Regnart J, Truter I, Zingela Z, Meyer A. A pilot study: Use of the Adult Attention-Deficit/Hyperactivity Disorder Self-Report Scale in a South African patient population. *South African Journal of Psychiatry*. 2019;25. doi:https://doi.org/10.4102/sajpsychiatry.v25i0.1326
4. Ganzenmüller JL, Ballmann C, von Nessen-Lapp RMW, et al. Screening tools for adult ADHD patients in primary care. *Journal of Affective Disorders Reports*. 2024;17:100800. doi:https://doi.org/10.1016/j.jadr.2024.100800
5. Kessler RC, Adler LA, Gruber MJ, Sarawate CA, Spencer T, Van Brunt DL. Validity of the World Health Organization Adult ADHD Self-Report Scale (ASRS) Screener in a representative sample of health plan members. *International Journal of Methods in Psychiatric Research*. 2007;16(2):52-65. doi:https://doi.org/10.1002/mpr.208
6. Ustun B, Adler LA, Rudin C, et al. The World Health Organization Adult Attention-Deficit/Hyperactivity Disorder Self-Report Screening Scale for DSM-5. *JAMA Psychiatry*. 2017;74(5):520. doi:https://doi.org/10.1001/jamapsychiatry.2017.0298
7. Kessler RC, Adler L, Ames M, et al. The World Health Organization Adult ADHD self-report Scale (ASRS): a Short Screening Scale for Use in the General Population. *Psychological Medicine*. 2005;35(2):245-256. doi:https://doi.org/10.1017/s0033291704002892
8. Evren C, Umut G, Bozkurt M, Teksin-Unal G, Agachanli R, Evren B. Psychometric properties of the Turkish version of the Adult ADHD Self-Report Scale (ASRS-v1.1) in a sample of Inpatients with alcohol use disorder. *Dusunen Adam: The Journal of Psychiatry and Neurological Sciences*. Published online June 28, 2016:109-119. doi:https://doi.org/10.5350/dajpn2016290202
9. Ramos Quiroga JA, Daigre Blanco C, Valero Ventura S, et al. Validación al español de la escala de cribado del trastorno por déficit de atención/hiperactividad en adultos (ASRS v. 1.1): una nueva estrategia de puntuación. *Revista de Neurología*. 2009;48(09):449. doi:https://doi.org/10.33588/rn.4809.2008677

10. Simon CB, McCabe CJ, Matson TE, Oliver M, Bradley KA, Hallgren KA. High test-retest reliability of the Alcohol Use Disorder Identification Test-Consumption version (AUDIT-C) completed by primary care patients in routine care. *Alcohol Clinical and Experimental Research*. 2023;48(2). doi:https://doi.org/10.1111/acer.15245
11. Frank D, DeBenedetti AF, Volk RJ, Williams EC, Kivlahan DR, Bradley KA. Effectiveness of the AUDIT-C as a Screening Test for Alcohol Misuse in Three Race/Ethnic Groups. *Journal of General Internal Medicine*. 2008;23(6):781-787. doi:https://doi.org/10.1007/s11606-008-0594-0
12. Bradley KA, DeBenedetti AF, Volk RJ, Williams EC, Frank D, Kivlahan DR. AUDIT-C as a brief screen for alcohol misuse in primary care. *Alcoholism, Clinical and Experimental Research*. 2007;31(7):1208-1217. doi:https://doi.org/10.1111/j.1530-0277.2007.00403.x
13. Johnson JA, Lee A, Vinson D, Seale JP. Use of AUDIT-Based Measures to Identify Unhealthy Alcohol Use and Alcohol Dependence in Primary Care: A Validation Study. *Alcoholism: Clinical and Experimental Research*. 2012;37:E253-E259. doi:https://doi.org/10.1111/j.1530-0277.2012.01898.x
14. Osaki Y, Ino A, Matsushita S, Higuchi S, Kondo Y, Kinjo A. Reliability and Validity of the Alcohol Use Disorders Identification Test - Consumption in Screening for Adults with Alcohol Use Disorders and Risky Drinking In Japan. *Asian Pacific Journal of Cancer Prevention*. 2014;15(16):6571-6574. doi:https://doi.org/10.7314/apjcp.2014.15.16.6571
15. Atkins DL, Cumbe VFJ, Muanido A, et al. Validity and item response theory properties of the Alcohol Use Disorders Identification Test for primary care alcohol use screening in Mozambique (AUDIT-MZ). *Journal of Substance Abuse Treatment*. 2021;127:108441. doi:https://doi.org/10.1016/j.jsat.2021.108441
16. Kriston L. Meta-analysis: Are 3 Questions Enough to Detect Unhealthy Alcohol Use? *Annals of Internal Medicine*. 2008;149(12):879. doi:https://doi.org/10.7326/0003-4819-149-12-200812160-00007
17. Knopf A. AUDIT-C shows reliability when used online in patient portals. *Alcoholism & Drug Abuse Weekly*. 2024;36(41):8-8. doi:https://doi.org/10.1002/adaw.34301
18. Meneses-Gaya C, Zuardi AW, Loureiro SR, et al. Is the Full Version of the AUDIT Really Necessary? Study of the Validity and Internal Construct of Its Abbreviated Versions. *Alcoholism: Clinical and Experimental Research*. Published online June 25, 2010:no-no. doi:https://doi.org/10.1111/j.1530-0277.2010.01225.x
19. Hallgren KA, Matson TE, Oliver M, et al. Practical Assessment of Alcohol Use Disorder in Routine Primary Care: Performance of an Alcohol Symptom Checklist. *Journal of General Internal Medicine*. 2021;37. doi:https://doi.org/10.1007/s11606-021-07038-3
20. Bryan CJ, Blount T, Kanzler KA, et al. Reliability and normative data for the Behavioral Health Measure (BHM) in primary care behavioral health settings. *Families, Systems, & Health*. 2014;32(1):89-100. doi:https://doi.org/10.1037/fsh0000014

21. Bryan CJ, Morrow C, Appolonio KK. Impact of behavioral health consultant interventions on patient symptoms and functioning in an integrated family medicine Clinic. *Journal of Clinical Psychology*. 2009;65(3):281-293. doi:https://doi.org/10.1002/jclp.20539
22. Ray-Sannerud BN, Dolan DC, Morrow CE, et al. Longitudinal outcomes after brief behavioral health intervention in an integrated primary care clinic. *Families, Systems, & Health*. 2012;30(1):60-71. doi:https://doi.org/10.1037/a0027029
23. Nelson KG, Young K, Chapman H. Examining the performance of the brief addiction monitor. *Journal of Substance Abuse Treatment*. 2014;46(4):472-481. doi:https://doi.org/10.1016/j.jsat.2013.07.002
24. Schumm J, Wong C, Okrant E, Tharp JA, Embree J, Lester N. Factor Structure of the Brief Addiction Monitor in a Non-Veteran Substance Use Disorder Outpatient Treatment Sample. *Drug and Alcohol Dependence Reports*. Published online November 2022:100125. doi:https://doi.org/10.1016/j.dadr.2022.100125
25. Gaddy MA, Casner HG, Rosinski J. Factor structure and measurement invariance of the Brief Addiction Monitor. *Journal of Substance Abuse Treatment*. 2018;90:29-37. doi:https://doi.org/10.1016/j.jsat.2018.04.010
26. Blanchard BE, Lynch KG, Malte CA, et al. Towards shortening the Brief Addiction Monitor-Revised (BAM-R). *Drug and Alcohol Dependence Reports*. 2023;8:100183. doi:https://doi.org/10.1016/j.dadr.2023.100183
27. Hallinan S, Gaddy M, Ghosh A, Burgen E. Factor structure and measurement invariance of the Revised Brief Addiction Monitor. *Psychological Assessment*. 2021;33(3):273-278. doi:https://doi.org/10.1037/pas0000973
28. Dowdy-Hazlett T, Killian M, Woods M. Measurement of traumatic experiences of children within survey and intervention research: A systematic review of the child and adolescent trauma screen. *Children and Youth Services Review*. 2021;131. doi:https://doi.org/10.1016/j.childyouth.2021.106259
29. Nilsson D, Dävelid I, Ledin S, Svedin CG. Psychometric properties of the Child and Adolescent Trauma Screen (CATS) in a sample of Swedish children. *Nordic Journal of Psychiatry*. 2020;75(4):247-256. doi:https://doi.org/10.1080/08039488.2020.1840628
30. Sachser C, Berliner L, Holt T, et al. International development and psychometric properties of the Child and Adolescent Trauma Screen (CATS). *Journal of Affective Disorders*. 2017;210:189-195. doi:https://doi.org/10.1016/j.jad.2016.12.040
31. Danyliuk I, Malysheva K, Oleksandra Loshenko. Adapting the Child and Adolescent Trauma Screen (CATS) methodology for the Ukrainian-speaking population. *TECHNOLOGIES OF INTELLECT DEVELOPMENT*. 2024;8(1). doi:https://doi.org/10.31108/3.2024.8.1.7
32. Sachser C, Berliner L, Risch E, et al. The child and Adolescent Trauma Screen 2 (CATS-2) - validation of an instrument to measure DSM-5 and ICD-11 PTSD and complex PTSD in

children and adolescents. *European Journal of Psychotraumatology*. 2022;13(2):2105580. doi:https://doi.org/10.1080/20008066.2022.2105580
33. Redican E, Sachser C, Pfeiffer E, et al. Validation of the Ukrainian caregiver-report version of the Child and Adolescent Trauma Screen (CATS) in children and adolescents in Ukraine. *Psychological Trauma: Theory, Research, Practice, and Policy*. Published online August 24, 2023. doi:https://doi.org/10.1037/tra0001570
34. Akkuş PZ, Serdaroğlu E, Kömürlüoğlu A, et al. Screening traumatic life events in preschool aged children: cultural adaptation of child and adolescent trauma screen (cats) caregiver-report 3-6 years version. *The Turkish Journal of Pediatrics*. 2021;63(1):95. doi:https://doi.org/10.24953/turkjped.2021.01.011
35. Müller LRF, Unterhitzenberger J, Wintersohl S, Rosner R, König J. Screening for Posttraumatic Stress Symptoms in Young Refugees: Comparison of Questionnaire Data with and without Involvement of an Interpreter. *International Journal of Environmental Research and Public Health*. 2021;18(13):6803. doi:https://doi.org/10.3390/ijerph18136803
36. Kim BK, Sim KS, Lee NB, Kim DH, Joo HS. A Validation of the Korean Version of the Child and Adolescent Trauma Screen 2(K-CATS-2). *The Association of Korea Counseling Psychology Education Welfare*. 2024;11(3):111-137. doi:https://doi.org/10.20496/cpew.2024.11.3.111
37. Casey P, Patalay P, Deighton J, Miller SD, Wolpert M. The Child Outcome Rating Scale: validating a four-item measure of psychosocial functioning in community and clinic samples of children aged 10–15. *European Child & Adolescent Psychiatry*. 2019;29(8):1089-1102. doi:https://doi.org/10.1007/s00787-019-01423-4
38. Duncan B, Sparks J, Bohanske RT, Claud DA. Giving youth a voice: A preliminary study of the reliability and validity of a brief outcome measure for children, adolescents, and caretakers. *Psychology*. Published online 2008.
39. Corona CD, Gutierrez PM, Wagner BM, Jobes DA. The psychometric properties of the Collaborative Assessment and Management of Suicidality rating scale. *Journal of Clinical Psychology*. 2018;75(1):190-201. doi:https://doi.org/10.1002/jclp.22699
40. Nielsen AC, Alberdi F, Rosenbaum B. Collaborative assessment and management of suicidality method shows effect. *PubMed*. 2011;58(8):A4300-A4300.
41. Comtois KA, Jobes DA, S. O'Connor S, et al. Collaborative assessment and management of suicidality (CAMS): feasibility trial for next-day appointment services. *Depression and Anxiety*. 2011;28(11):963-972. doi:https://doi.org/10.1002/da.20895
42. Ryberg W, Zahl PH, Diep LM, Landrø NI, Fosse R. Managing suicidality within specialized care: A randomized controlled trial. *Journal of Affective Disorders*. 2019;249:112-120. doi:https://doi.org/10.1016/j.jad.2019.02.022

43. Comtois KA, Hendricks KE, DeCou CR, et al. Reducing short-term suicide risk after hospitalization: A randomized controlled trial of the Collaborative Assessment and Management of Suicidality. *Journal of Affective Disorders*. 2023;320:656-666. doi:https://doi.org/10.1016/j.jad.2022.09.042
44. Swift JK, Trusty WT, Penix EA. The effectiveness of the collaborative assessment and management of suicidality (CAMS) compared to alternative treatment conditions: A meta-analysis. *Suicide and Life-Threatening Behavior*. 2021;51(5). doi:https://doi.org/10.1111/sltb.12765
45. Jobes DA, Comtois KA, Gutierrez PM, et al. A Randomized Controlled Trial of the Collaborative Assessment and Management of Suicidality versus Enhanced Care as Usual With Suicidal Soldiers. *Psychiatry*. 2017;80(4):339-356. doi:https://doi.org/10.1080/00332747.2017.1354607
46. Arkov K, Rosenbaum B, Christiansen L, Jønsson H, Münchow M. [Treatment of suicidal patients: The Collaborative Assessment and Management of Suicidality]. *PubMed*. 2008;170(3):149-153.
47. Huh D, Jobes DA, Comtois KA, et al. The collaborative assessment and management of suicidality (CAMS) versus enhanced care as usual (E-CAU) with suicidal soldiers: Moderator analyses from a randomized controlled trial. *Military Psychology*. 2018;30(6):495-506. doi:https://doi.org/10.1080/08995605.2018.1503001
48. Corona CD, Gutierrez PM, Wagner BM, Jobes DA. Assessing the Reliability of the CAMS Rating Scale Using a Generalizability Study. *Crisis*. 2019;40(4):273-279. doi:https://doi.org/10.1027/0227-5910/a000565
49. Posner K, Brown GK, Stanley B, et al. The Columbia–Suicide Severity Rating Scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry*. 2011;168(12):1266-1277. doi:https://doi.org/10.1176/appi.ajp.2011.10111704
50. Austria-Corrales F, Jiménez-Tapia A, Astudillo-García CI, et al. The Columbia-suicide severity rating scale: Validity and psychometric properties of an online Spanish-language version in a Mexican population sample. *Frontiers in Public Health*. 2023;11. doi:https://doi.org/10.3389/fpubh.2023.1157581
51. Kilincaslan A, Gunes A, Eskin M, Madan A. Linguistic adaptation and psychometric properties of the Columbia-Suicide Severity Rating Scale among a heterogeneous sample of adolescents in Turkey. *The International Journal of Psychiatry in Medicine*. 2018;54(2):115-132. doi:https://doi.org/10.1177/0091217418791454
52. Matarazzo BB, Brown GK, Stanley B, et al. Predictive Validity of the Columbia-Suicide Severity Rating Scale among a Cohort of At-risk Veterans. *Suicide and Life-Threatening Behavior*. 2018;49(5):1255-1265. doi:https://doi.org/10.1111/sltb.12515

53. Serrani Azcurra D. Psychometric validation of the Columbia-Suicide Severity rating scale in Spanish-speaking adolescents. *Colombia médica*. 2017;48(4):174-182. doi:https://doi.org/10.25100/cm.v48i4.2294
54. Katz I, Barry CN, Cooper SA, Kasprow WJ, Hoff RA. Use of the columbia-suicide severity rating scale (C- SSRS) in a large sample of veterans receiving mental health services in the veterans health administration. *Suicide and Life-Threatening Behavior*. 2019;50(1). doi:https://doi.org/10.1111/sltb.12584
55. Al-Halabí S, Sáiz PA, Burón P, et al. Validation of a Spanish version of the Columbia-Suicide Severity Rating Scale (C-SSRS). *Revista de Psiquiatría y Salud Mental (English Edition)*. 2016;9(3):134-142. doi:https://doi.org/10.1016/j.rpsmen.2016.06.004
56. Kerr DCR, Gibson B, Leve LD, DeGarmo DS. Young Adult Follow-up of Adolescent Girls in Juvenile Justice Using the Columbia Suicide Severity Rating Scale. *Suicide and Life-Threatening Behavior*. 2014;44(2):113-129. doi:https://doi.org/10.1111/sltb.12072
57. Ji Y, Liu X, Zheng S, et al. Validation and application of the Chinese version of the Columbia-Suicide Severity Rating Scale: Suicidality and cognitive deficits in patients with major depressive disorder. *Journal of Affective Disorders*. 2023;342:139-147. doi:https://doi.org/10.1016/j.jad.2023.09.014
58. Dhalla S, D. Zumbo B, Poole G. A Review of the Psychometric Properties of the CRAFFT Instrument: 1999-2010. *Current Drug Abuse Reviews*. 2011;4(1):57-64. doi:https://doi.org/10.2174/1874473711104010057
59. Rial A, Kim-Harris S, Knight JR, et al. Validación empírica del CRAFFT Abuse Screening Test en una muestra de adolescentes españoles. *Adicciones*. 2018;31(2):160. doi:https://doi.org/10.20882/adicciones.1105
60. Knight JR, Sherritt L, Shrier LA, Harris SK, Chang G. Validity of the CRAFFT Substance Abuse Screening Test Among Adolescent Clinic Patients. *Archives of Pediatrics & Adolescent Medicine*. 2002;156(6):607. doi:https://doi.org/10.1001/archpedi.156.6.607
61. Lee CA, Smith DC, Lanker A, Clary KL. Sensitivity and specificity of the CRAFFT to identify heavy cannabis use: Evidence from a large statewide adolescent sample. *Addictive Behaviors*. 2021;122:107006. doi:https://doi.org/10.1016/j.addbeh.2021.107006
62. Bertini M, Busaniche J, Baquero F, et al. Adaptación transcultural y validación del test CRAFFT como prueba de pesquisa para consumo problemático, abuso y dependencia de alcohol y otras sustancias en un grupo de adolescentes argentinos. *Archivos Argentinos de Pediatria*. 2015;113(2). doi:https://doi.org/10.5546/aap.2015.114
63. Harris SK, Knight, Jr. JR, Van Hook S, et al. Adolescent substance use screening in primary care: Validity of computer self-administered versus clinician-administered screening. *Substance Abuse*. 2016;37(1):197-203. doi:https://doi.org/10.1080/08897077.2015.1014615

64. Cummins LH, Chan KK, Burns KM, Blume AW, Larimer M, Marlatt GA. Validity of the CRAFFT in American-Indian and Alaska-Native adolescents: screening for drug and alcohol risk. *Journal of Studies on Alcohol*. 2003;64(5):727-732. doi:https://doi.org/10.15288/jsa.2003.64.727
65. Mitchell SG, Kelly SM, Gryczynski J, et al. The CRAFFT Cut-Points and DSM-5 Criteria for Alcohol and Other Drugs: A Reevaluation and Reexamination. *Substance Abuse*. 2014;35(4):376-380. doi:https://doi.org/10.1080/08897077.2014.936992
66. Knight JR, Harris SK, Sherritt L, et al. Prevalence of Positive Substance Abuse Screen Results Among Adolescent Primary Care Patients. *Archives of Pediatrics & Adolescent Medicine*. 2007;161(11):1035. doi:https://doi.org/10.1001/archpedi.161.11.1035
67. Levy S, Sherritt L, Harris SK, et al. Test-Retest Reliability of Adolescents??? Self-Report of Substance Use. *Alcoholism: Clinical & Experimental Research*. 2004;28(8):1236-1241. doi:https://doi.org/10.1097/01.alc.0000134216.22162.a5
68. Asghari A. Psychometric properties of the Depression Anxiety Stress Scales-21 (DASS-21) in a non-clinical Iranian sample. *Psychology*. Published online 2010.
69. Zolotareva AA. Systematic review of the psychometric properties of the Depression Anxiety and Stress Scale-21 (DASS-21). *VM BEKHTEREV REVIEW OF PSYCHIATRY AND MEDICAL PSYCHOLOGY*. 2020;(2):26-37. doi:https://doi.org/10.31363/2313-7053-2020-2-26-37
70. Gloster AT, Rhoades HM, Novy D, et al. Psychometric properties of the Depression Anxiety and Stress Scale-21 in older primary care patients. *Journal of Affective Disorders*. 2008;110(3):248-259. doi:https://doi.org/10.1016/j.jad.2008.01.023
71. Sinclair SJ, Siefert CJ, Slavin-Mulford JM, Stein MB, Renna M, Blais MA. Psychometric Evaluation and Normative Data for the Depression, Anxiety, and Stress Scales-21 (DASS-21) in a Nonclinical Sample of U.S. Adults. *Evaluation & the Health Professions*. 2011;35(3):259-279. doi:https://doi.org/10.1177/0163278711424282
72. Tonsing KN. Psychometric properties and validation of Nepali version of the Depression Anxiety Stress Scales (DASS-21). *Asian Journal of Psychiatry*. 2014;8:63-66. doi:https://doi.org/10.1016/j.ajp.2013.11.001
73. Alfonsson S, Wallin E, Maathz P. Factor structure and validity of the Depression, Anxiety and Stress Scale-21 in Swedish translation. *Journal of Psychiatric and Mental Health Nursing*. 2017;24(2-3):154-162. doi:https://doi.org/10.1111/jpm.12363
74. Wang K, Shi HS, Geng FL, et al. Cross-cultural validation of the Depression Anxiety Stress Scale-21 in China. *Psychological Assessment*. 2016;28(5):e88-e100. doi:https://doi.org/10.1037/pas0000207
75. Henry JD, Crawford JR. The short-form Version of the Depression Anxiety Stress Scales (DASS-21): Construct Validity and Normative Data in a Large non-clinical Sample. *The*

British journal of clinical psychology. 2005;44(2):227-239. doi:https://doi.org/10.1348/014466505X29657
76. Crawford JR, Henry JD. The Depression Anxiety Stress Scales (DASS): Normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology*. 2003;42(2):111-131. doi:https://doi.org/10.1348/014466503321903544
77. Kakemam E, Navvabi E, Albelbeisi AH, Saeedikia F, Rouhi A, Majidi S. Psychometric properties of the Persian version of Depression Anxiety Stress Scale-21 Items (DASS-21) in a sample of health professionals: a cross-sectional study. *BMC Health Services Research*. 2022;22(1). doi:https://doi.org/10.1186/s12913-022-07514-4
78. Evren C, Can Y, Yilmaz A, et al. Psychometric properties of the Drug Abuse Screening Test (DAST-10) in heroin dependent adults and adolescents with drug use disorder. *Dusunen Adam: The Journal of Psychiatry and Neurological Sciences*. Published online December 15, 2013:351-359. doi:https://doi.org/10.5350/dajpn2013260404
79. Sashiprabha Dulanjalee Nawaratne, Sashiprabha Dulanjalee Nawaratne. Psychometric properties of the modified Drug Abuse Screening Test Sinhala version (DAST-SL): evaluation of reliability and validity in Sri Lanka. *BMC public health*. 2024;24(1). doi:https://doi.org/10.1186/s12889-024-19288-x
80. Pérez Gálvez B, García Fernández L, De Vicente Manzanaro MP, Oliveras Valenzuela MA, Lahoz Lafuente M. Spanish Validation of the Drug Abuse Screening Test (DAST-20 y DAST-10) / Validación española del Drug Abuse Screening Test (DAST-20 y DAST-10). *Health and Addictions/Salud y Drogas*. 2010;10(1). doi:https://doi.org/10.21134/haaj.v10i1.35
81. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. A Single-Question Screening Test for Drug Use in Primary Care. *Archives of Internal Medicine*. 2010;170(13). doi:https://doi.org/10.1001/archinternmed.2010.140
82. Gavin DR, Ross HE, Skinner HA. Diagnostic Validity of the Drug Abuse Screening Test in the Assessment of DSM-III Drug Disorders. *Addiction*. 1989;84(3):301-307. doi:https://doi.org/10.1111/j.1360-0443.1989.tb03463.x
83. Cassidy CM, Schmitz N, Malla A. Validation of the Alcohol Use Disorders Identification Test and the Drug Abuse Screening Test in First Episode Psychosis. *The Canadian Journal of Psychiatry*. 2008;53(1):26-33. doi:https://doi.org/10.1177/070674370805300105
84. Shirinbayan P, Salavati M, Soleimani F, et al. The Psychometric Properties of the Drug Abuse Screening Test. *Addiction & Health*. 2020;12(1):25-33. doi:https://doi.org/10.22122/ahj.v12i1.256
85. Fatemi S, Soleimani R, Yazdanipour MA, Novin MH, Abdollahi E. Psychometric Properties of 20-Item and 10-Item Persian Versions of Drug Abuse Screening Test. *Journal of Holistic Nursing And Midwifery*. 2022;32(3):234-241. doi:https://doi.org/10.32598/jhnm.32.3.2366

86. Murad HAS, AlHarthi NA, Bakarman MA, Gazzaz ZJ. Development and Validation of an Arabic Version of the Drug Abuse Screening Test-10 (DAST-10) among Saudi Drug Abusers. *Journal of Psychoactive Drugs*. 2021;54(5):1-8. doi:https://doi.org/10.1080/02791072.2021.2013580
87. S. McCann, Tracy L. Simpson, Richar B. Reliability and Validity of Screening Instruments for Drug and Alcohol Abuse in Adults Seeking Evaluation for Attention-Deficit/ Hyperactivity Disorder. *American Journal on Addictions*. 2000;9(1):1-9. doi:https://doi.org/10.1080/10550490050172173
88. Parkerson GR, Broadhead WE, Tse CKJ. The Duke Health Profile. *Medical Care*. 1990;28(11):1056-1072. doi:https://doi.org/10.1097/00005650-199011000-00007
89. Parkerson GR, Gehlbach SH, Wagner EH, James SA, Clapp NE, Muhlbaier LH. The Duke-UNC Health Profile: An Adult Health Status Instrument for Primary Care. *Medical Care*. 1981;19(8):806-828. doi:https://doi.org/10.1097/00005650-198108000-00002
90. Novin MH, Farzadfar F, Pashaei T, Razaghi E. Cross-Culture Adaptation and Psychometric Properties of the Persian Version of Duke Health Profile. *Iranian Journal of Psychiatry and Behavioral Sciences*. 2020;14(4). doi:https://doi.org/10.5812/ijpbs.102765
91. Rl B, Ta V, Sc Z, Ee B. Evaluation of a health status measure in adults with high psychosocial risk. *PubMed*. 1986;5(3):158-166.
92. Guillemin F, A Paul-Dauphin, Virion JM, Bouchet C, S Briançon. [The DUKE health profile: a generic instrument to measure the quality of life tied to health]. *PubMed*. 1997;9(1):35-44.
93. Rapin A, Drame M, Jolly D, et al. Psychometric properties of the Duke Health Profile in a neuromuscular disease population. *European Journal of Physical and Rehabilitation Medicine*. 2016;52(1):57-64.
94. Parkerson GR, Willke RJ, Hays RD. An International Comparison of the Reliability and Responsiveness of the Duke Health Profile for Measuring Health-Related Quality of Life of Patients Treated With Alprostadil for Erectile Dysfunction. *Medical Care*. 1999;37(1):56-67. doi:https://doi.org/10.1097/00005650-199901000-00009
95. Schuntermann MF. Duke Health Profile (DUKE). *The Sage Dictionary of Health and Society*. Published online January 1, 2006. doi:https://doi.org/10.4135/9781446215159.n274
96. Parkerson GR. Duke Health Profile (DUKE). *Springer eBooks*. Published online January 1, 2023:1892-1895. doi:https://doi.org/10.1007/978-3-031-17299-1_785
97. Shrestha SD, Pradhan R, Tran TD, Gualano RC, Fisher JRW. Reliability and validity of the Edinburgh Postnatal Depression Scale (EPDS) for detecting perinatal common mental disorders (PCMDs) among women in low-and lower-middle-income countries: a systematic review. *BMC Pregnancy and Childbirth*. 2016;16(1). doi:https://doi.org/10.1186/s12884-016-0859-2

98. Gholamreza Kheirabadi, Mohammad Reza Maracy, Sahar Akbaripour, Nasrin Masaeli. Psychometric properties and diagnostic accuracy of the edinburgh postnatal depression scale in a sample of Iranian women. *PubMed*. Published online March 1, 2012.
99. Gyimah L, Agyepong IA, Owiredu D, et al. Tools for screening maternal mental health conditions in primary care settings in sub-Saharan Africa: systematic review. *Frontiers in Public Health*. 2024;12. doi:https://doi.org/10.3389/fpubh.2024.1321689
100. Fellmeth G, Harrison S, Opondo C, Nair M, Kurinczuk JJ, Alderdice F. Validated screening tools to identify common mental disorders in perinatal and postpartum women in India: a systematic review and meta-analysis. *BMC Psychiatry*. 2021;21(1). doi:https://doi.org/10.1186/s12888-021-03190-6
101. Hartley CM, Barroso N, Rey Y, Pettit JW, Bagner DM. Factor Structure and Psychometric Properties of English and Spanish Versions of the Edinburgh Postnatal Depression Scale Among Hispanic Women in a Primary Care Setting. *Journal of Clinical Psychology*. 2014;70(12):1240-1250. doi:https://doi.org/10.1002/jclp.22101
102. Smith-Nielsen J, Matthey S, Lange T, Væver MS. Validation of the Edinburgh Postnatal Depression Scale against Both DSM-5 and ICD-10 Diagnostic Criteria for Depression. *BMC Psychiatry*. 2018;18(1). doi:https://doi.org/10.1186/s12888-018-1965-7
103. Mohd W, Amir Awang, Mohamed MN. Revalidation of the Malay Version of the Edinburgh Postnatal Depression Scale (EPDS) Among Malay Postpartum Women Attending the Bakar Bata Health Center in Alor Setar, Kedah, Northwest Of Peninsular Malaysia. *PubMed*. 2003;10(2):71-75.
104. Wang Y, Guo X, Lau Y, Chan KS, Yin L, Chen J. Psychometric evaluation of the Mainland Chinese version of the Edinburgh Postnatal Depression Scale. *International Journal of Nursing Studies*. 2009;46(6):813-823. doi:https://doi.org/10.1016/j.ijnurstu.2009.01.010
105. Tesfaye M, Hanlon C, Wondimagegn D, Alem A. Detecting postnatal common mental disorders in Addis Ababa, Ethiopia: Validation of the Edinburgh Postnatal Depression Scale and Kessler Scales. *Journal of Affective Disorders*. 2010;122(1-2):102-108. doi:https://doi.org/10.1016/j.jad.2009.06.020
106. Green EP, Tuli H, Kwobah E, Menya D, Chesire I, Schmidt C. Developing and validating a perinatal depression screening tool in Kenya blending Western criteria with local idioms: A mixed methods study. *Journal of Affective Disorders*. 2018;228:49-59. doi:https://doi.org/10.1016/j.jad.2017.11.027
107. Maria Teresa Gonçalves, Malafaia S, Santos, Roth T, Daniel Ruivo Marques. Epworth sleepiness scale: A meta-analytic study on the internal consistency. *Sleep Medicine*. 2023;109:261-269. doi:https://doi.org/10.1016/j.sleep.2023.07.008
108. Bloch KE, Schoch OD, Zhang JN, Russi EW. German Version of the Epworth Sleepiness Scale. *Respiration*. 1999;66(5):440-447. doi:https://doi.org/10.1159/000029408

109. Baumgartel KL, Terhorst L, Conley YP, Roberts JM. Psychometric evaluation of the Epworth Sleepiness Scale in an obstetric population. *Sleep Medicine*. 2013;14(1):116-121. doi:https://doi.org/10.1016/j.sleep.2012.10.007
110. Sandoval-Rincón M, Alcalá-Lozano R, Iván Herrera-Jiménez, Jiménez-Genchi A. [Validation of the Epworth sleepiness scale in Mexican population]. *PubMed*. 2013;149(4):409-416.
111. Bertolazi AN, Fagondes SC, Hoff LS, Pedro VD, Menna Barreto SS, Johns MW. Portuguese-language version of the Epworth sleepiness scale: validation for use in Brazil. *Jornal Brasileiro de Pneumologia*. 2009;35(9):877-883. doi:https://doi.org/10.1590/s1806-37132009000900009
112. Beaudreau SA, Spira AP, Stewart A, et al. Validation of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Scale in older black and white women. *Sleep Medicine*. 2012;13(1):36-42. doi:https://doi.org/10.1016/j.sleep.2011.04.005
113. Elbiaze M. Reliability and Validity of a Moroccan Arabic Dialect Version of the Epworth Sleepiness Scale. *Acta Scientific Medical Sciences*. 2020;4(11):55-59. doi:https://doi.org/10.31080/asms.2020.04.0770
114. Shukla G, Bajpai G, Pandey R, et al. Validation of a modified Hindi version of the Epworth Sleepiness Scale among a North Indian population. *Annals of Indian Academy of Neurology*. 2016;19(4):499. doi:https://doi.org/10.4103/0972-2327.194427
115. Bajpai G, Shukla G, Pandey R, Gupta A, Goyal V, Behari AM. Validation of a modified hindi version of the Epworth Sleepiness Scale in a north Indian population. *Sleep Medicine*. 2013;14:e68-e69. doi:https://doi.org/10.1016/j.sleep.2013.11.133
116. Jimenez-Correa U, Haro R, Poblano A n, et al. Mexican Version of the Epworth Sleepiness Scale. *The Open Sleep Journal*. 2009;2(1):6-10. doi:https://doi.org/10.2174/1874620900902010006
117. Campo-Arias A, Herazo E, Pedrozo-Pupo JC. Validity and reliability of the Spanish version of the Epworth Sleepiness Scale. *Sleep and Breathing*. 2025;29(3). doi:https://doi.org/10.1007/s11325-025-03339-7
118. Wu S, Wang R, Ma X, Zhao Y, Yan X, He J. Excessive daytime sleepiness assessed by the Epworth Sleepiness Scale and its association with health related quality of life: a population-based study in China. *BMC Public Health*. 2012;12(1). doi:https://doi.org/10.1186/1471-2458-12-849
119. Ahn JK, Kim Y, Choi KH. The psychometric properties and clinical utility of the Korean version of GAD-7 and GAD-2. *Frontiers in Psychiatry*. 2019;10(1). doi:https://doi.org/10.3389/fpsyt.2019.00127
120. García-Campayo J, Zamorano E, Ruiz MA, Pérez-Páramo M, López-Gómez V, Rejas J. The assessment of generalized anxiety disorder: psychometric validation of the Spanish version of the self-administered GAD-2 scale in daily medical practice. *Health and*

Quality of Life Outcomes. 2012;10(1):114. doi:https://doi.org/10.1186/1477-7525-10-114

121. Vrublevska J, Renemane L, Kivite-Urtane A, Rancans E. Validation of the generalized anxiety disorder scales (GAD-7 and GAD-2) in primary care settings in Latvia. *Frontiers in Psychiatry*. 2022;13. doi:https://doi.org/10.3389/fpsyt.2022.972628
122. Mills SD, Fox RS, Malcarne VL, Roesch SC, Champagne BR, Sadler GR. The psychometric properties of the Generalized Anxiety Disorder-7 Scale in Hispanic Americans with English or Spanish language preference. *Cultural Diversity and Ethnic Minority Psychology*. 2014;20(3):463-468. doi:https://doi.org/10.1037/a0036523
123. Müller F, Hansen A, Kube M, Arnetz JE, Holman HT. Translation, cultural adaptation, and validation of the PHQ-9 and GAD-7 in Kinyarwanda for primary care in the United States. *PLOS ONE*. 2024;19(10). doi:https://doi.org/10.1371/journal.pone.0302953
124. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety disorder: the GAD-7. *Archives of Internal Medicine*. 2006;166(10):1092-1097. doi:https://doi.org/10.1001/archinte.166.10.1092
125. Luo Z, Li Y, Hou Y, et al. Adaptation of the two-item generalized anxiety disorder scale (GAD-2) to Chinese rural population: A validation study and meta-analysis. *General Hospital Psychiatry*. 2019;60:50-56. doi:https://doi.org/10.1016/j.genhosppsych.2019.07.008
126. Rósa Ingólfsdóttir. Psychometric Properties of the Icelandic Version of the Generalized Anxiety Disorder-7. *Psychology*. Published online May 1, 2014.
127. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine*. 2007;146(5):317. doi:https://doi.org/10.7326/0003-4819-146-5-200703060-00004
128. Gagnon C, Bélanger L, Ivers H, Morin CM. Validation of the Insomnia Severity Index in Primary Care. *The Journal of the American Board of Family Medicine*. 2013;26(6):701-710. doi:https://doi.org/10.3122/jabfm.2013.06.130064
129. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*. 2011;34(5):601-608. doi:https://doi.org/10.1093/sleep/34.5.601
130. Horacio Balam Álvarez-García, Isaías Vicente Lugo-González, Fabiola González Betanzos. Psychometric properties of the Insomnia Severity Index (ISI) in Mexican adults. *Interacciones*. Published online May 19, 2023:e311-e311. doi:https://doi.org/10.24016/2023.v9.311
131. Yu DSF. Insomnia Severity Index: psychometric properties with Chinese community-dwelling older people. *Journal of Advanced Nursing*. 2010;66(10):2350-2359. doi:https://doi.org/10.1111/j.1365-2648.2010.05394.x

132. Ba S, Gang Ck. Reliability and Validity of the Chinese Translation of Insomnia Severity Index and Comparison with Pittsburgh Sleep Quality Index. ˜*The Malaysian journal of psychiatry*. 2013;22(2):3-9.
133. Clemente V, Ruivo Marques D, Miller-Mendes M, Morin C, Serra J, Allen Gomes A. The European Portuguese version of the Insomnia Severity Index. *Journal of Sleep Research*. Published online September 30, 2020. doi:https://doi.org/10.1111/jsr.13198
134. Kaufmann CN, Orff HJ, Moore RC, Delano-Wood L, Depp CA, Schiehser DM. Psychometric Characteristics of the Insomnia Severity Index in Veterans With History of Traumatic Brain Injury. *Behavioural Sleep Medicine*. 2017;17(1):12-18. doi:https://doi.org/10.1080/15402002.2016.1266490
135. Suleiman KH, Yates BC. Translating the Insomnia Severity Index Into Arabic. *Journal of Nursing Scholarship*. 2011;43(1):49-53. doi:https://doi.org/10.1111/j.1547-5069.2010.01374.x
136. Michaud AL, Zhou ES, Chang G, Recklitis CJ. Validation of the Insomnia Severity Index (ISI) for identifying insomnia in young adult cancer survivors: comparison with a structured clinical diagnostic interview of the DSM-5 (SCID-5). *Sleep Medicine*. 2021;81:80-85. doi:https://doi.org/10.1016/j.sleep.2021.01.045
137. Thakral M, Von Korff M, McCurry SM, Morin CM, Vitiello MV. ISI-3: Evaluation of a brief screening tool for insomnia. *Sleep Medicine*. Published online August 2020. doi:https://doi.org/10.1016/j.sleep.2020.08.027
138. Hu J, Zhou W, Hu S, et al. Cross-cultural difference and validation of the Chinese version of Montreal Cognitive Assessment in older adults residing in Eastern China: Preliminary findings. *Archives of Gerontology and Geriatrics*. 2013;56(1):38-43. doi:https://doi.org/10.1016/j.archger.2012.05.008
139. Suchy-Dicey AM, Vo TT, Oziel K, Buchwald DS, Rhoads K, French BF. Psychometric Reliability, Validity, and Generalizability of MoCA in American Indian Adults: The Strong Heart Study. *Assessment*. Published online July 24, 2024. doi:https://doi.org/10.1177/10731911241261436
140. Daniel B, Agenagnew L, Workicho A, Abera M. Psychometric properties of the montreal cognitive assessment (moca) to detect major neurocognitive disorder among older people in ethiopia: A validation study. *Neuropsychiatric Disease and Treatment*. 2022;18:1789-1798. doi:https://doi.org/10.2147/ndt.s377430
141. D Agnani P, R Bhise A. Translation, cultural adaptation and validation of Gujarti version of Montreal Cognitive Assessment (MOCA) in older adults. *International Journal of Scientific Research.* Published online February 1, 2023:80-82. doi:https://doi.org/10.36106/ijsr/8702792
142. Freitas S, Prieto G, Simões MR, Santana I. Psychometric Properties of the Montreal Cognitive Assessment (MoCA): An Analysis Using the Rasch Model. *The Clinical*

Neuropsychologist. 2014;28(1):65-83. doi:https://doi.org/10.1080/13854046.2013.870231
143. Goldstein FC, Ashley AV, Miller E, Alexeeva O, Zanders L, King V. Validity of the Montreal Cognitive Assessment as a Screen for Mild Cognitive Impairment and Dementia in African Americans. *Journal of Geriatric Psychiatry and Neurology*. 2014;27(3):199-203. doi:https://doi.org/10.1177/0891988714524630
144. Pedraza OL, Salazar AM, Sierra FA, et al. Reliability, criterion and discriminant validity of the Montreal Cognitive Assessment Test (MoCA) in a group of adults from Bogotá. *Medicine*. 2016;41(4):221-228.
145. Lee JY, Dong Woo Lee, Cho SJ, et al. Brief Screening for Mild Cognitive Impairment in Elderly Outpatient Clinic: Validation of the Korean Version of the Montreal Cognitive Assessment. *Journal of Geriatric Psychiatry and Neurology*. 2008;21(2):104-110. doi:https://doi.org/10.1177/0891988708316855
146. Aguilar-Navarro SG, Mimenza-Alvarado AJ, Palacios-García AA, Samudio-Cruz A, Gutiérrez-Gutiérrez LA, Ávila-Funes JA. Validez y confiabilidad del MoCA (Montreal Cognitive Assessment) para el tamizaje del deterioro cognoscitivo en méxico. *Revista Colombiana de Psiquiatría*. 2018;47(4):237-243. doi:https://doi.org/10.1016/j.rcp.2017.05.003
147. Pinto T, Machado L, Costa Maria Lúcia G, et al. Accuracy and Psychometric Properties of the Brazilian Version of the Montreal Cognitive Assessment as a Brief Screening Tool for Mild Cognitive Impairment and Alzheimer's Disease in the Initial Stages in the Elderly. *Dementia and Geriatric Cognitive Disorders*. 2019;47(4-6):366-374. doi:https://doi.org/10.1159/000501308
148. Graves RE, Alim TN, Aigbogun N, et al. Diagnosing bipolar disorder in trauma exposed primary care patients. *Bipolar Disorders*. 2007;9(4):318-323. doi:https://doi.org/10.1111/j.1399-5618.2007.00449.x
149. Lin CJ, Shiah I-Shin, Chu H, et al. Reliability and Validity of the Chinese Version of the Mood Disorder Questionnaire. *Archives of Psychiatric Nursing*. 2011;25(1):53-62. doi:https://doi.org/10.1016/j.apnu.2010.03.003
150. Ouali U, Jouini L, Zgueb Y, et al. The Factor Structure of the Mood Disorder Questionnaire in Tunisian Patients. *Clinical practice and epidemiology in mental health: CP & EMH*. 2020;16(Suppl-1):82-92. doi:https://doi.org/10.2174/1745017902016010082
151. Zimmerman M, Galione JN. Screening for Bipolar Disorder with the Mood Disorders Questionnaire: A Review. *Harvard Review of Psychiatry*. 2011;19(5):219-228. doi:https://doi.org/10.3109/10673229.2011.614101
152. Jon DI, Hong N, Yoon BH, et al. Validity and reliability of the Korean version of the Mood Disorder Questionnaire. *Comprehensive Psychiatry*. 2009;50(3):286-291. doi:https://doi.org/10.1016/j.comppsych.2008.07.008

153. Numan Konuk, Kiran S, Lut Tamam, Elif Karaahmet, Aydin H, Atik L. [Validation of the Turkish version of the mood disorder questionnaire for screening bipolar disorders]. *PubMed*. 2007;18(2):147-154.
154. de Sousa Gurgel W, Rebouças DB, Negreiros de Matos KJ, Carneiro AHS, Gomes de Matos e Souza F. Brazilian Portuguese validation of Mood Disorder Questionnaire. *Comprehensive Psychiatry*. 2012;53(3):308-312. doi:https://doi.org/10.1016/j.comppsych.2011.04.059
155. Hirschfeld RMA, Holzer C, Calabrese JR, et al. Validity of the Mood Disorder Questionnaire: A General Population Study. *American Journal of Psychiatry*. 2003;160(1):178-180. doi:https://doi.org/10.1176/appi.ajp.160.1.178
156. González A, Arias A, Mata S, Lima L. [Validation of the Spanish version of the Mood Disorder Questionnaire to detect bipolar disorder type II in patients with major depression disorder]. *PubMed*. 2009;50(2):163-171.
157. Cheatle MD, Compton PA, Dhingra L, Wasser TE, O'Brien CP. Development of the Revised Opioid Risk Tool to Predict Opioid Use Disorder in Patients with Chronic Nonmalignant Pain. *The Journal of Pain*. 2019;20(7):842-851. doi:https://doi.org/10.1016/j.jpain.2019.01.011
158. Ma JD, Horton JM, Hwang M, Atayee RS, Roeland EJ. A Single-Center, Retrospective Analysis Evaluating the Utilization of the Opioid Risk Tool in Opioid-Treated Cancer Patients. *Journal of Pain & Palliative Care Pharmacotherapy*. 2014;28(1):4-9. doi:https://doi.org/10.3109/15360288.2013.869647
159. Chamoun K, Mouawad J, Salameh P, et al. Opioid use disorder in two samples of the Lebanese population: scale validation and correlation with sleep and mood disorders. *BMC Psychiatry*. 2023;23(1). doi:https://doi.org/10.1186/s12888-023-05304-8
160. Strand MA, Eukel HN, Frenzel O, Skoy E, Steig J, Werremeyer A. Opioid risk stratification in the community pharmacy: The utility of the Opioid Risk Tool. *Research in Social and Administrative Pharmacy*. Published online July 2022. doi:https://doi.org/10.1016/j.sapharm.2022.07.009
161. Esteve R, Reyes-Pérez Á, Ramírez-Maestre C, et al. Diagnostic and Predictive Capacity of the Spanish Versions of the Opioid Risk Tool and the Screener and Opioid Assessment for Patients with Pain—Revised: A Preliminary Investigation in a Sample of People with Noncancer Chronic Pain. *Pain and Therapy*. 2022;11(2):493-510. doi:https://doi.org/10.1007/s40122-022-00356-2
162. Teulings L, Broglio K. Opioid Misuse Risk: Implementing Screening Protocols in an Ambulatory Oncology Clinic. *Clinical Journal of Oncology Nursing*. 2020;24(1):11-14. doi:https://doi.org/10.1188/20.cjon.11-14
163. Perry JS, Stoll KE, Allen AD, Hahn JC, Ostrum RF. The Opioid Risk Tool Correlates With Increased Postsurgical Opioid Use Among Patients With Orthopedic Trauma.

Orthopedics. Published online February 10, 2023:1-4. doi:https://doi.org/10.3928/01477447-20230207-04

164. Haller IV, Renier CM, Juusola M, et al. Enhancing Risk Assessment in Patients Receiving Chronic Opioid Analgesic Therapy Using Natural Language Processing. *Pain Medicine*. Published online December 29, 2016:pnw283. doi:https://doi.org/10.1093/pm/pnw283
165. Bringhurst, D. L., Watson, C. W., Miller, S. D., & Duncan, B. L. (2006). The reliability and validity of the Outcome Rating Scale: A replication study of a brief clinical measure. *Journal of brief Therapy*, *5*(1), 23-30.
166. Moggia D, Niño-Robles N, Miller SD, Feixas G. Psychometric Properties of the Outcome Rating Scale (ORS) in a Spanish Clinical Sample. *The Spanish Journal of Psychology*. 2018;21. doi:https://doi.org/10.1017/sjp.2018.32
167. Campbell A, Hemsley S. Outcome Rating Scale and Session Rating Scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist*. 2009;13(1):1-9. doi:https://doi.org/10.1080/13284200802676391
168. Hafkenscheid A. De Outcome rating scale (ORS) en de Session rating scale (SRS). *Tijdschrift voor Psychotherapie*. 2010;36(6):394-403. doi:https://doi.org/10.1007/s12485-010-0173-9
169. Meier ST. Construct Validity of Outcome Rating Scale (ORS) Scores in Clinical Samples: Extension of Harris, Murphy, and Rakes' (2019) Narrative Review. *Journal of Evidence-Based Social Work*. 2020;17(6):648-661. doi:https://doi.org/10.1080/26408066.2020.1784345
170. Seryjová Juhová D, Řiháček T, Cígler H, et al. Czech version of the outcome rating scale: Selected psychometric properties. *Ceskoslovenska psychologie*. 2021;65(4):353-368. doi:https://doi.org/10.51561/cspsych.65.4.353
171. Janse P, Boezen-Hilberdink L, van Dijk MK, Verbraak MJPM, Hutschemaekers GJM. Measuring Feedback From Clients. *European Journal of Psychological Assessment*. 2014;30(2):86-92. doi:https://doi.org/10.1027/1015-5759/a000172
172. Andrade-González N, Rodrigo-Holgado I, Fernández-Rozas J, et al. Spanish Versions of the Outcome Rating Scale and the Session Rating Scale: Normative Data, Reliability, and Validity. *Frontiers in Psychology*. 2021;12. doi:https://doi.org/10.3389/fpsyg.2021.663791
173. Munro L, Rodwell J. Validation of an Australian Sign Language Instrument of Outcome Measurement for Adults in Mental Health Settings. *Australian & New Zealand Journal of Psychiatry*. 2009;43(4):332-339. doi:https://doi.org/10.1080/00048670902721111
174. Kapos FP, Hancock C, Viviana Guerrero Torres, Gonzalez MI, Do A, Jensen MP. Validation of the PEG Scale in Spanish (PEG-S) Among Adults Receiving Care for Pain in US Primary Care. *The Journal of Pain*. 2023;24(11):1897-1904. doi:https://doi.org/10.1016/j.jpain.2023.06.005

175. Krebs EE, Bair MJ, Damush TM, Tu W, Wu J, Kroenke K. Comparative Responsiveness of Pain Outcome Measures Among Primary Care Patients With Musculoskeletal Pain. *Medical Care*. 2010;48(11):1007-1014. doi:https://doi.org/10.1097/mlr.0b013e3181eaf835
176. Margot, Wendy, Achterberg WP, Jacobijn Gussekloo, Blom JW. A Postal Screener for Pain and Need for Treatment in Older Persons in Primary Care. *Journal of the American Geriatrics Society*. 2014;62(10):1832-1837. doi:https://doi.org/10.1111/jgs.13064
177. Wuyek LA, Antony MM, McCabe RE. Psychometric properties of the panic disorder severity scale: clinician-administered and self-report versions. *Clinical Psychology & Psychotherapy*. 2010;18(3):234-243. doi:https://doi.org/10.1002/cpp.703
178. Roberge P, Provencher M, Norton P, et al. Panic Disorder Severity Scale self-report: transcultural validation and sensitivity to change of the French-Canadian adaptation. *European Psychiatry*. 2022;65(S1):S388-S389. doi:https://doi.org/10.1192/j.eurpsy.2022.982
179. Svensson M, Nilsson T, Johansson H, Viborg G, Perrin S, Sandell R. Psychometric analysis of the Swedish panic disorder severity scale and its self-report version. *Nordic Journal of Psychiatry*. 2019;73(1):58-63. doi:https://doi.org/10.1080/08039488.2018.1554699
180. Fuste G, Gil MÁ, López-Solà C, et al. Psychometric Properties of the Spanish Version of the Panic Disorder Severity Scale. *The Spanish Journal of Psychology*. 2018;21. doi:https://doi.org/10.1017/sjp.2018.6
181. Roberge P, Marx P, Couture J, et al. French adaptation and validation of the Panic Disorder Severity Scale—self-report. *BMC Psychiatry*. 2022;22(1). doi:https://doi.org/10.1186/s12888-022-03989-x
182. Santacana M, Fullana MA, Bonillo A, et al. Psychometric properties of the Spanish self-report version of the Panic Disorder Severity Scale. *Comprehensive Psychiatry*. 2014;55(6):1467-1472. doi:https://doi.org/10.1016/j.comppsych.2014.04.007
183. Lee EH, Kim JH, Yu BH. Reliability and validity of the self-report version of the Panic Disorder Severity Scale in Korea. *Depression and Anxiety*. 2009;26(8):E120-E123. doi:https://doi.org/10.1002/da.20461
184. Liu X, Xu T, Chen D, et al. Reliability, validity and cut-off score of the Chinese version of the panic disorder severity scale self-report form in patients with panic disorder. *BMC Psychiatry*. 2020;20(1). doi:https://doi.org/10.1186/s12888-020-02560-w
185. Lim YJ, Yu BH, Kim JH. Korean panic disorder severity scale: construct validity by confirmatory factor analysis. *Depression and Anxiety*. 2007;24(2):95-102. doi:https://doi.org/10.1002/da.20206

186. He Y, Zeng Q, Wei J, et al. Reliability and validity of the Chinese version of Panic Disorder Severity Scale and Panic-Associated Symptom Scale. *Chin J Psychiatry*. 2013;46(04):217-221. doi:https://doi.org/10.3760/cma.j.issn.1006-7884.2013.04.006
187. Arrieta J, Aguerrebere M, Raviola G, et al. Validity and Utility of the Patient Health Questionnaire (PHQ)-2 and PHQ-9 for Screening and Diagnosis of Depression in Rural Chiapas, Mexico: A Cross-Sectional Study. *Journal of Clinical Psychology*. 2017;73(9):1076-1090. doi:https://doi.org/10.1002/jclp.22390
188. El-Den S, Chen TF, Gan YL, Wong E, O'Reilly CL. The psychometric properties of depression screening tools in primary healthcare settings: A systematic review. *Journal of Affective Disorders*. 2018;225:503-522. doi:https://doi.org/10.1016/j.jad.2017.08.060
189. Rancans E, Trapencieris M, Ivanovs R, Vrublevska J. Validity of the PHQ-9 and PHQ-2 to screen for depression in nationwide primary care population in Latvia. *Annals of General Psychiatry*. 2018;17(1). doi:https://doi.org/10.1186/s12991-018-0203-5
190. Aslan J, Cova F, Saldivia S, et al. Psychometric Properties of the Patient Health Questionnaire-9 in Elderly Chilean Primary Care Users. *Frontiers in Psychiatry*. 2020;11. doi:https://doi.org/10.3389/fpsyt.2020.555011
191. Vrublevska J, Trapencieris M, Rancans E. Adaptation and validation of the Patient Health Questionnaire-9 to evaluate major depression in a primary care sample in Latvia. *Nordic Journal of Psychiatry*. 2017;72(2):112-118. doi:https://doi.org/10.1080/08039488.2017.1397191
192. Cassiani-Miranda CA, Cuadros-Cruz AK, Torres-Pinzón H, et al. Validity of the Patient Health Questionnaire-9 (PHQ-9) for depression screening in adult primary care users in Bucaramanga, Colombia. *Revista Colombiana De Psiquiatria*. 2021;50(1):11-21. doi:https://doi.org/10.1016/j.rcp.2019.09.001
193. Molebatsi K, Motlhatlhedi K, Wambua GN. The validity and reliability of the Patient Health Questionnaire-9 for screening depression in primary health care patients in Botswana. *BMC Psychiatry*. 2020;20(1). doi:https://doi.org/10.1186/s12888-020-02719-5
194. Ferreira T, Sousa M, Salgado J. Brief assessment of depression: Psychometric properties of the Portuguese version of the Patient Health Questionnaire (PHQ-9). *The Psychologist: Practice & Research Journal*. 2019;1(2):1. doi:https://doi.org/10.33525/pprj.v1i2.36
195. Gelaye B, Williams MA, Lemma S, et al. Validity of the patient health questionnaire-9 for depression screening and diagnosis in East Africa. *Psychiatry Research*. 2013;210(2):653-661. doi:https://doi.org/10.1016/j.psychres.2013.07.015
196. Al-Amer R, Maneze D, Ramjan L, Villarosa AR, Darwish R, Salamonson Y. Psychometric testing of the Arabic version of the Patient Health Questionnaire among adolescent

refugees living in Jordan. *International Journal of Mental Health Nursing*. 2020;29(4):685-692. doi:https://doi.org/10.1111/inm.12702
197. Luitel NP, Rimal D, Eleftheriou G, et al. Translation, cultural adaptation and validation of Patient Health Questionnaire and generalized anxiety disorder among adolescents in Nepal. *Child and Adolescent Psychiatry and Mental Health*. 2024;18(1). doi:https://doi.org/10.1186/s13034-024-00763-7
198. Chen YQ, Huang XJ, Yang F, et al. A Chinese adaptation of the Patient Health Questionnaire for Adolescents (PHQ-A): factor structure and psychometric properties. *BMC psychiatry*. 2024;24(1). doi:https://doi.org/10.1186/s12888-024-05783-3
199. Tele AK, Carvajal-Velez L, Nyongesa V, et al. Validation of the English and Swahili Adaptation of the Patient Health Questionnaire–9 for Use Among Adolescents in Kenya. *Journal of Adolescent Health*. 2023;72(1, Supplement):S61-S70. doi:https://doi.org/10.1016/j.jadohealth.2022.10.003
200. Zachar-Tirado R, Chen Z, Donders J. Clinical Utility of the Patient Health Questionnaire–Adolescent in Adolescents With Traumatic Brain Injury. *Journal of Head Trauma Rehabilitation*. 2020;Publish Ahead of Print. doi:https://doi.org/10.1097/htr.0000000000000610
201. Richardson LP, McCauley E, Grossman DC, et al. Evaluation of the Patient Health Questionnaire-9 Item for Detecting Major Depression Among Adolescents. *PEDIATRICS*. 2010;126(6):1117-1123. doi:https://doi.org/10.1542/peds.2010-0852
202. Borghero F, Martínez V, Zitko P, et al. Tamizaje de episodio depresivo en adolescentes. Validación del instrumento PHQ-9. *Revista médica de Chile*. 2018;146(4):479-486. doi:https://doi.org/10.4067/s0034-98872018000400479
203. Naveed S, Waqas A, Memon AR, Jabeen M, Sheikh MH. Cross-cultural validation of the Urdu translation of the Patient Health Questionnaire for Adolescents among children and adolescents at a Pakistani school. *Public Health*. 2019;168:59-66. doi:https://doi.org/10.1016/j.puhe.2018.11.022
204. Pitts BH, Sheeder J, Sigel E, Love-Osborne K, Woods J. Informing Use of the Patient Health Questionnaire-2 to Detect Moderate or Greater Depression Symptoms in Adolescents and Young Adults in Outpatient Primary Care. *Journal of Adolescent Health*. Published online April 2023. doi:https://doi.org/10.1016/j.jadohealth.2023.02.043
205. Cerimele JM, Russo J, Bauer AM, et al. The Patient Mania Questionnaire (PMQ-9): a Brief Scale for Assessing and Monitoring Manic Symptoms. *Journal of General Internal Medicine*. Published online June 18, 2021. doi:https://doi.org/10.1007/s11606-021-06947-7
206. Cerimele JM, Fortney JC. Bipolar disorder assessment and monitoring measures in clinical care: Updates from a large randomized controlled trial in primary care. *Bipolar Disorders*. Published online November 3, 2023. doi:https://doi.org/10.1111/bdi.13382

207. Kushalnagar P, Paludneviciene R, Kallen M, Atcherson S, Cella D. PROMIS-deaf profile measure: cultural adaptation and psychometric validation in American sign language. *Journal of Patient-Reported Outcomes*. 2020;4(1). doi:https://doi.org/10.1186/s41687-020-00208-7
208. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*. 2010;63(11):1179-1194. doi:https://doi.org/10.1016/j.jclinepi.2010.04.011
209. Rimehaug SA, Kaat AJ, Nordvik JE, Klokkerud M, Robinson HS. Psychometric properties of the PROMIS-57 questionnaire, Norwegian version. *Quality of Life Research*. Published online June 18, 2021. doi:https://doi.org/10.1007/s11136-021-02906-1
210. Abma IL, Butje BJD, ten Klooster PM, van der Wees PJ. Measurement properties of the Dutch–Flemish patient-reported outcomes measurement information system (PROMIS) physical function item bank and instruments: a systematic review. *Health and Quality of Life Outcomes*. 2021;19(1). doi:https://doi.org/10.1186/s12955-020-01647-y
211. Jensen RE, Potosky AL, Reeve BB, et al. Validation of the PROMIS physical function measures in a diverse US population-based cohort of cancer patients. *Quality of Life Research*. 2015;24(10):2333-2344. doi:https://doi.org/10.1007/s11136-015-0992-9
212. Teresi JA, Katja Ocepek-Welikson, Kleinman M, Ramirez M, Kim G. Psychometric Properties and Performance of the Patient Reported Outcomes Measurement Information System® (PROMIS®) Depression Short Forms in Ethnically Diverse Groups. *PubMed*. 2016;58(1):141-181.
213. Rose AJ, Bayliss E, Huang W, et al. Evaluating the PROMIS-29 v2.0 for Use Among Older Adults with Multiple Chronic Conditions. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 2018;27(11):2935-2944. doi:https://doi.org/10.1007/s11136-018-1958-5
214. Edwards TC, Fredericksen RJ, Crane HM, et al. Content validity of Patient-Reported Outcomes Measurement Information System (PROMIS) items in the context of HIV clinical care. *Quality of Life Research*. 2015;25(2):293-302. doi:https://doi.org/10.1007/s11136-015-1096-2
215. Kroenke K, Yu Z, Wu J, Kean J, Monahan PO. Operating Characteristics of PROMIS Four-Item Depression and Anxiety Scales in Primary Care Patients with Chronic Pain. *Pain Medicine*. 2014;15(11):1892-1901. doi:https://doi.org/10.1111/pme.12537
216. Meinert AC, Mire SS, Kim HJ, Shellman AB, Keller-Margulis MA, Curtis DF. A Study of the Psychometric Properties of the Pediatric Symptom Checklist-17 for Children With Developmental Delays and Disorders. *Clinical Pediatrics*. Published online September 28, 2024. doi:https://doi.org/10.1177/00099228241284095

217. Kostanecka A, Power T, Clarke A, Watkins M, Hausman CL, Blum NJ. Behavioral Health Screening in Urban Primary Care Settings: Construct Validity of the PSC-17. *Journal of Developmental & Behavioral Pediatrics*. 2008;29(2):124-128. doi:https://doi.org/10.1097/dbp.0b013e31816a0d9e
218. Gardner W, Lucas A, Kolko DJ, Campo JV. Comparison of the PSC-17 and Alternative Mental Health Screens in an At-Risk Primary Care Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2007;46(5):611-618. doi:https://doi.org/10.1097/chi.0b013e318032384b
219. Murphy JM, Bergmann P, Chiang C, et al. The PSC-17: Subscale Scores, Reliability, and Factor Structure in a New National Sample. *PEDIATRICS*. 2016;138(3):e20160038-e20160038. doi:https://doi.org/10.1542/peds.2016-0038
220. Erdogan S, Ozturk M. Psychometric evaluation of the Turkish version of the Pediatric Symptom Checklist-17 for detecting psychosocial problems in low-income children. *Journal of Clinical Nursing*. 2011;20(17-18):2591-2599. doi:https://doi.org/10.1111/j.1365-2702.2010.03537.x
221. Piqueras JA, Vidal-Arenas V, Falcó R, et al. Short Form of the Pediatric Symptom Checklist-Youth Self-Report (PSC-17-Y): Spanish Validation Study. *Journal of Medical Internet Research*. 2021;23(12):e31127. doi:https://doi.org/10.2196/31127
222. Murphy JM, Reede J, Jellinek MS, Bishop SJ. Screening for Psychosocial Dysfunction in Inner-City Children: Further Validation of the Pediatric Symptom Checklist. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1992;31(6):1105-1111. doi:https://doi.org/10.1097/00004583-199211000-00019
223. Jutte DP, Burgos A, Mendoza F, Ford CB, Huffman LC. Use of the Pediatric Symptom Checklist in a Low-Income, Mexican American Population. *Archives of Pediatrics & Adolescent Medicine*. 2003;157(12):1169. doi:https://doi.org/10.1001/archpedi.157.12.1169
224. Han DH, Woo J, Jeong JH, Hwang S, Chung US. The Korean Version of the Pediatric Symptom Checklist: Psychometric Properties in Korean School-aged Children. *Journal of Korean Medical Science*. 2015;30(8):1167. doi:https://doi.org/10.3346/jkms.2015.30.8.1167
225. Leiva L, Rojas R, Peña F, Vargas B, Scquicciarini A. Detectando las Dificultades Emocionales y Conductuales en la Escuela: Validación de PSC-17. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*. 2019;50(4). doi:https://doi.org/10.21865/ridep50.1.08
226. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. *Journal of traumatic stress*. 2015;28(6):489-498. doi:https://doi.org/10.1002/jts.22059

227. Ibrahim H, Ertl V, Catani C, Ismail AA, Neuner F. The validity of Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5) as screening instrument with Kurdish and Arab displaced populations living in the Kurdistan region of Iraq. *BMC Psychiatry*. 2018;18(1). doi:https://doi.org/10.1186/s12888-018-1839-z
228. Forkus SR, Raudales AM, Rafiuddin HS, Weiss NH, Messman BA, Contractor AA. The posttraumatic stress disorder (PTSD) checklist for DSM–5: A systematic review of existing psychometric evidence. *Clinical Psychology: Science and Practice*. 2022;30(1):110-121. doi:https://doi.org/10.1037/cps0000111
229. Verhey R, Chibanda D, Gibson L, Brakarsh J, Seedat S. Validation of the posttraumatic stress disorder checklist – 5 (PCL-5) in a primary care population with high HIV prevalence in Zimbabwe. *BMC Psychiatry*. 2018;18(1). doi:https://doi.org/10.1186/s12888-018-1688-9
230. Lathan EC, Petri JM, Haynes T, et al. Evaluating the Performance of the Primary Care Posttraumatic Stress Disorder Screen for DSM-5 (PC-PTSD-5) in a Trauma-Exposed, Socioeconomically Vulnerable Patient Population. *Journal of Clinical Psychology in Medical Settings*. Published online January 30, 2023. doi:https://doi.org/10.1007/s10880-023-09941-9
231. Blanchard BE, Johnson M, Campbell SB, et al. Minimal important difference metrics and test–retest reliability of the PTSD Checklist for *DSM*-5 with a primary care sample. *Journal of Traumatic Stress*. 2023;36(6):1102-1114. doi:https://doi.org/10.1002/jts.22975
232. Pereira-Lima K, Loureiro SR, Bolsoni LM, Apolinario da Silva TD, Osório FL. Psychometric properties and diagnostic utility of a Brazilian version of the PCL-5 (complete and abbreviated versions). *European Journal of Psychotraumatology*. 2019;10(1):1581020. doi:https://doi.org/10.1080/20008198.2019.1581020
233. Hansen M, Vaegter HB, Ravn SL, Andersen TE. Validation of the Danish PTSD Checklist for DSM-5 in trauma-exposed chronic pain patients using the Clinician-Administered PTSD Scale for DSM-5. *European Journal of Psychotraumatology*. 2023;14(1). doi:https://doi.org/10.1080/20008066.2023.2179801
234. Sveen J, Bondjers K, Willebrand M. Psychometric properties of the PTSD Checklist for DSM-5: a pilot study. *European Journal of Psychotraumatology*. 2016;7(1):30165. doi:https://doi.org/10.3402/ejpt.v7.30165
235. Ferrie O, Richardson T, Smart T, Ellis-Nee C. A validation of the PCL–5 questionnaire for PTSD in primary and secondary care. *Psychological Trauma: Theory, Research, Practice, and Policy*. 2022;15(5). doi:https://doi.org/10.1037/tra0001354
236. Thapa A, Rayens MK, Chung ML, et al. Psychometric Testing of the Protocol for Responding to and Assessing Patient Assets, Risks, and Experiences (PRAPARE) in

Patients With Heart Failure and Coronary Heart Disease. *Research in Nursing & Health*. Published online January 7, 2025. doi:https://doi.org/10.1002/nur.22440

237. Drake C, Batchelder H, Lian T, et al. Implementation of social needs screening in primary care: a qualitative study using the health equity implementation framework. *BMC Health Services Research*. 2021;21(1). doi:https://doi.org/10.1186/s12913-021-06991-3

238. Wan W, Li V, Chin MH, et al. Development of PRAPARE Social Determinants of Health Clusters and Correlation with Diabetes and Hypertension Outcomes. *The Journal of the American Board of Family Medicine*. 2022;35(4):668-679. doi:https://doi.org/10.3122/jabfm.2022.04.200462

239. Weir RC, Proser M, Jester M, Li V, Hood-Ronick CM, Gurewich D. Collecting Social Determinants of Health Data in the Clinical Setting: Findings from National PRAPARE Implementation. *Journal of Health Care for the Poor and Underserved*. 2020;31(2):1018-1035. doi:https://doi.org/10.1353/hpu.2020.0075

240. Lynch S, Street NW, Sharpe L, Kellish A. Social Determinants of Health Screening: Primary Care PRAPARE Tool Implementation. *The Journal for Nurse Practitioners*. 2024;20(4):104955-104955. doi:https://doi.org/10.1016/j.nurpra.2024.104955

241. Zucoloto ML, Martinez EZ. General aspects of quality of life in heterogeneous populations: notes on Flanagan's Quality of Life Scale (QoLS). *Trends in Psychiatry and Psychotherapy*. 2019;41(3):268-275. doi:https://doi.org/10.1590/2237-6089-2018-0077

242. Burckhardt CS, Anderson KL, Archenholtz B, Hägg O. The Flanagan Quality of Life Scale: Evidence of Construct Validity. *Health and Quality of Life Outcomes*. 2003;1(1):59. doi:https://doi.org/10.1186/1477-7525-1-59

243. Burckhardt CS, Archenholtz B, Bjelle A. Measuring the Quality of Life of Women with Rheumatoid Arthritis or Systemic Lupus Erythematosus: A Swedish Version of the Quality of Life Scale (QOLS). *Scandinavian Journal of Rheumatology*. 1992;21(4):190-195. doi:https://doi.org/10.3109/03009749209099220

244. Wahl A, Burckhardt C, Wiklund I, Hanestad BR. The Norwegian Version of the Quality of Life Scale (QOLS-N). *Scandinavian Journal of Caring Sciences*. 1998;12(4):215-222. doi:https://doi.org/10.1111/j.1471-6712.1998.tb00500.x

245. Novelli MMPC, Rovere HHD, Nitrini R, Caramelli P. Cross-cultural adaptation of the quality of life assessment scale on Alzheimer disease. *Arquivos de Neuro-Psiquiatria*. 2005;63(2a):201-206. doi:https://doi.org/10.1590/s0004-282x2005000200002

246. Danao LL, Padilla GV, Johnson DA. An English and Spanish quality of life measure for rheumatoid arthritis. *Arthritis and Rheumatism*. 2001;45(2):167-173. doi:https://doi.org/10.1002/1529-0131(200104)45:2%3C167::AID-ANR170%3E3.0.CO;2-X

247. Reeves AJ, Baker RT, Casanova MP, Cheatham SW, Pickering MA. Examining the factorial validity of the Quality of Life Scale. *Health and Quality of Life Outcomes*. 2020;18(1). doi:https://doi.org/10.1186/s12955-020-01292-5
248. Gómez-Gallego M, Gómez-Amor J, Gómez-García J. Validación de la versión española de la escala QoL-AD en pacientes con enfermedad de Alzheimer, cuidadores y profesionales sanitarios. *Neurología*. 2012;27(1):4-10. doi:https://doi.org/10.1016/j.nrl.2011.03.006
249. Akinpelu AO, Odetunde MO, Odole AC. Cross-cultural adaptation and initial validation of the Stroke-Specific Quality of Life Scale into the Yoruba language. *International Journal of Rehabilitation Research*. 2012;35(4):339-344. doi:https://doi.org/10.1097/mrr.0b013e328355dd54
250. Runyon K, Chesnut SR, Burley H. Screening for childhood anxiety: A meta-analysis of the screen for child anxiety related emotional disorders. *Journal of Affective Disorders*. 2018;240:220-229. doi:https://doi.org/10.1016/j.jad.2018.07.049
251. Birmaher B, Brent DA, Chiapetta L, Bridge J, Monga S, Baugher M. Psychometric Properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): A Replication Study. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1999;38(10):1230-1236. doi:https://doi.org/10.1097/00004583-199910000-00011
252. Muris P, Steerneman P. The Revised version of the Screen for Child Anxiety Related Emotional Disorders (SCARED-R): First evidence for its reliability and validity in a clinical sample. *British Journal of Clinical Psychology*. 2001;40(1):35-44. doi:https://doi.org/10.1348/014466501163463
253. Wren FJ, Berg EA, Heiden LA, et al. Childhood Anxiety in a Diverse Primary Care Population. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2007;46(3):332-340. doi:https://doi.org/10.1097/chi.0b013e31802f1267
254. Wren FJ, Bridge JA, Birmaher B. Screening for Childhood Anxiety Symptoms in Primary Care: Integrating Child and Parent Reports. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2004;43(11):1364-1371. doi:https://doi.org/10.1097/01.chi.0000138350.60487.d3
255. Gonzalez A, Weersing VR, Warnick E, Scahill L, Woolston J. Cross-Ethnic Measurement Equivalence of the SCARED in an Outpatient Sample of African American and Non-Hispanic White Youths and Parents. *Journal of Clinical Child & Adolescent Psychology*. 2012;41(3):361-369. doi:https://doi.org/10.1080/15374416.2012.654462
256. Bölte S, Poustka F, Constantino JN. Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Research*. 2008;1(6):354-363. doi:https://doi.org/10.1002/aur.49
257. Nguyen PH, Ocansey ME, Miller M, Le DTK, Schmidt RJ, Prado EL. The reliability and validity of the social responsiveness scale to measure autism symptomology in

Vietnamese children. *Autism Research*. 2019;12(11):1706-1718. doi:https://doi.org/10.1002/aur.2179
258. Cen CQ, Liang YY, Chen QR, et al. Investigating the validation of the Chinese Mandarin version of the Social Responsiveness Scale in a Mainland China child population. *BMC Psychiatry*. 2017;17(1). doi:https://doi.org/10.1186/s12888-016-1185-y
259. Cheon KA, Park JI, Koh YJ, et al. The social responsiveness scale in relation to DSM IV and DSM5 ASD in Korean children. *Autism Research*. 2016;9(9):970-980. doi:https://doi.org/10.1002/aur.1671
260. Stordeur C, Boele A, Peyre H, Delorme R, Acquaviva E. Psychometric properties of the French Version of the Social Responsiveness Scale in autism spectrum disorder with or without attention deficit hyperactivity disorder. *L'Encéphale*. 2019;45(4):285-289. doi:https://doi.org/10.1016/j.encep.2018.08.004
261. Wigham S, McConachie H, Tandos J, Le Couteur AS. The reliability and validity of the Social Responsiveness Scale in a UK general child population. *Research in Developmental Disabilities*. 2012;33(3):944-950. doi:https://doi.org/10.1016/j.ridd.2011.12.017
262. Chun J, Bong G, Han JH, Oh M, Yoo HJ. Validation of Social Responsiveness Scale for Korean Preschool Children With Autism. *Psychiatry Investigation*. 2021;18(9):831-840. doi:https://doi.org/10.30773/pi.2021.0182
263. Kaiser S, Bergquist KÅ, Halvorsen MB. The psychometric properties of the Norwegian version of the social responsiveness scale in a neuropediatric sample. *Research in Autism Spectrum Disorders*. 2022;95:101973. doi:https://doi.org/10.1016/j.rasd.2022.101973
264. Lyall K, Rando J, Toroni B, et al. Examining shortened versions of the Social Responsiveness Scale for use in autism spectrum disorder prediction and as a quantitative trait measure: Results from a validation study of 3–5 year old children. *JCPP Advances*. 2022;2(4). doi:https://doi.org/10.1002/jcv2.12106
265. Takei R, Matsuo J, Takahashi H, Uchiyama T, Kunugi H, Kamio Y. Verification of the utility of the social responsiveness scale for adults in non-clinical and clinical adult populations in Japan. *BMC Psychiatry*. 2014;14(1). doi:https://doi.org/10.1186/s12888-014-0302-z
266. Noyes ET, Major S, Wilson AM, Campbell EB, Ratcliffe LN, Spencer RJ. Reliability and Factor Structure of the Saint Louis University Mental Status (SLUMS) Examination. *Clinical Gerontologist*. Published online September 6, 2022:1-7. doi:https://doi.org/10.1080/07317115.2022.2120446
267. Shwartz SK, Morris RD, Penna S. Psychometric properties of the Saint Louis University Mental Status Examination. *Applied Neuropsychology: Adult*. 2017;26(2):101-110. doi:https://doi.org/10.1080/23279095.2017.1362407

268. Yang YP, Huang YC, Chen CS, Yang YC, Wang JJ. Sensitivity and Specificity of the Saint Louis University Mental Status Examination to Detect Mild Cognitive Impairment and Dementia in Chinese Population. *Gerontology*. Published online January 21, 2021:1-8. doi:https://doi.org/10.1159/000511904
269. Merz ZC, Lace JW. Clinical utility of the Saint Louis University Mental Status Examination (SLUMS) in a mixed neurological sample: Proposed revised cutoff scores for normal cognition, mild cognitive impairment, and dementia. *Applied Neuropsychology: Adult*. Published online August 5, 2022:1-8. doi:https://doi.org/10.1080/23279095.2022.2106572
270. Cummings-Vaughn L, Cruz-Oliver D, Malmstrom T, Tumosa N, Morley J. The Veterans Affairs Medical Center Saint Louis University Mental Status Examination Comparison Study. *Alzheimer's & Dementia*. 2012;8(4):P485. doi:https://doi.org/10.1016/j.jalz.2012.05.1313
271. Tariq SH, Tumosa N, Chibnall JT, Perry MH, Morley JE. Comparison of the Saint Louis University Mental Status Examination and the Mini-Mental State Examination for Detecting Dementia and Mild Neurocognitive Disorder—A Pilot Study. *The American Journal of Geriatric Psychiatry*. 2006;14(11):900-910. doi:https://doi.org/10.1097/01.jgp.0000221510.33817.86
272. Cruz-Oliver DM, Malmstrom TK, Allen CM, Tumosa N, Morley JE. The veterans affairs Saint Louis University mental status exam (slums exam) and the mini-mental status exam as predictors of mortality and institutionalization. *The journal of nutrition, health & aging*. 2012;16(7):636-641. doi:https://doi.org/10.1007/s12603-012-0098-9
273. Cruz-Oliver DM, Malmstrom TK, Roegner M, Tumosa N, Grossberg GT. Cognitive Deficit Reversal as Shown by Changes in the Veterans Affairs Saint Louis University Mental Status (SLUMS) Examination Scores 7.5 Years Later. *Journal of the American Medical Directors Association*. 2014;15(9):687.e5-687.e10. doi:https://doi.org/10.1016/j.jamda.2014.05.004
274. Roberg BL, Anzalone C, Nicholson JD, Peruggia PE, Buckley TR. Performance Comparisons on the Saint Louis University Mental Status Examination Between Black and White Veterans and Education Classification in a Large Outpatient Sample from the Southern United States. *Archives of Clinical Neuropsychology*. 2022;38(4):633-643. doi:https://doi.org/10.1093/arclin/acac090
275. Cahn-Hidalgo D, Estes PW, Benabou R. Validity, reliability, and psychometric properties of a computerized, cognitive assessment test (Cognivue®). *World Journal of Psychiatry*. 2020;10(1):1-11. doi:https://doi.org/10.5498/wjp.v10.i1.1
276. Wolraich ML. Psychometric Properties of the Vanderbilt ADHD Diagnostic Parent Rating Scale in a Referred Population. *Journal of Pediatric Psychology*. 2003;28(8):559-568. doi:https://doi.org/10.1093/jpepsy/jsg046

277. Bard DE, Wolraich ML, Neas B, Doffing M, Beck L. The Psychometric Properties of the Vanderbilt Attention-Deficit Hyperactivity Disorder Diagnostic Parent Rating Scale in a Community Population. *Journal of Developmental & Behavioral Pediatrics*. 2013;34(2):72-82. doi:https://doi.org/10.1097/dbp.0b013e31827a3a22
278. Wolraich ML, Bard DE, Neas B, Doffing M, Beck L. The Psychometric Properties of the Vanderbilt Attention-Deficit Hyperactivity Disorder Diagnostic Teacher Rating Scale in a Community Population. *Journal of Developmental & Behavioral Pediatrics*. 2013;34(2):83-93. doi:https://doi.org/10.1097/dbp.0b013e31827d55c3
279. Kapogiannis A, Makris G, Darviri C, et al. The Greek Version of the Vanderbilt ADHD Diagnostic Parent Rating Scale for Follow-up Assessment in Prepubertal Children with ADHD. *International Journal of Disability, Development and Education*. 2020;69(5):1-10. doi:https://doi.org/10.1080/1034912x.2020.1802647
280. Xiao ZH, Wang QH, Luo TT, Zhong L. [Diagnostic value of Vanderbilt ADHD Parent Rating Scale in attention deficit hyperactivity disorder]. *PubMed*. 2013;15(5):348-352.
281. Becker SP, Langberg JM, Vaughn AJ, Epstein JN. Clinical Utility of the Vanderbilt ADHD Diagnostic Parent Rating Scale Comorbidity Screening Scales. *Journal of Developmental & Behavioral Pediatrics*. 2012;33(3):221-228. doi:https://doi.org/10.1097/dbp.0b013e318245615b
282. Vňuková M, Sebalo I, Anders M, Ptacek R, Surman C. Psychometric Properties of the Czech Version of the Vanderbilt ADHD Diagnostic Parent Rating Scale. *Journal of Attention Disorders*. 2023;27(10):1075-1080. doi:https://doi.org/10.1177/10870547231167567
283. Langberg JM, Vaughn AJ, Brinkman WB, Froehlich T, Epstein JN. Clinical Utility of the Vanderbilt ADHD Rating Scale for Ruling Out Comorbid Learning Disorders. *PEDIATRICS*. 2010;126(5):e1033-e1038. doi:https://doi.org/10.1542/peds.2010-1267
284. Leslie LK. Implementing the American Academy of Pediatrics Attention-Deficit/Hyperactivity Disorder Diagnostic Guidelines in Primary Care Settings. *PEDIATRICS*. 2004;114(1):129-140. doi:https://doi.org/10.1542/peds.114.1.129
285. Mohammed, Al M, Nawal Mohammed Alsharef, et al. Psychometric Properties of the Arabic Vanderbilt Children's ADHD Diagnostic Rating Scale (VADRS-A) in a Saudi Population Sample. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*. 2024;12(1):72-83. doi:https://doi.org/10.2478/sjcapp-2024-0008
286. Federici S, Bracalenti M, Meloni F, Luciano JV. World Health Organization disability assessment schedule 2.0: An international systematic review. *Disability and Rehabilitation*. 2016;39(23):2347-2380. doi:https://doi.org/10.1080/09638288.2016.1223177
287. Saltychev M, Katajapuu N, Bärlund E, Laimi K. Psychometric properties of 12-item self-administered World Health Organization disability assessment schedule 2.0 (WHODAS

2.0) among general population and people with non-acute physical causes of disability – systematic review. *Disability and Rehabilitation*. 2019;43(6):1-6. doi:https://doi.org/10.1080/09638288.2019.1643416

288. Üstün TB, Chatterji S, Kostanjsek N, et al. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bulletin of the World Health Organization*. 2010;88(11):815-823. doi:https://doi.org/10.2471/blt.09.067231
289. Luciano JV, Ayuso-Mateos JL, Fernández A, Serrano-Blanco A, Roca M, Haro JM. Psychometric properties of the twelve item World Health Organization Disability Assessment Schedule II (WHO-DAS II) in Spanish primary care patients with a first major depressive episode. *Journal of Affective Disorders*. 2010;121(1-2):52-58. doi:https://doi.org/10.1016/j.jad.2009.05.008
290. Igwesi-Chidobe CN, Kitchen S, Sorinola IO, Godfrey EL. World Health Organisation Disability Assessment Schedule (WHODAS 2.0): development and validation of the Nigerian Igbo version in patients with chronic low back pain. *BMC Musculoskeletal Disorders*. 2020;21(1). doi:https://doi.org/10.1186/s12891-020-03763-8
291. Habtamu K, Alem A, Medhin G, et al. Validation of the World Health Organization Disability Assessment Schedule in people with severe mental disorders in rural Ethiopia. *Health and Quality of Life Outcomes*. 2017;15(1). doi:https://doi.org/10.1186/s12955-017-0647-3
292. Kunt DA, Dereboy F. [Validity and Reliability of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) in Turkish Psychiatry Patients and Healthy Controls]. *Turkish Journal of Psychiatry*. Published online 2018.
293. Igwesi-Chidobe CN, Kitchen S, Sorinola IO, Godfrey EL. World Health Organisation Disability Assessment Schedule (WHODAS 2.0): development and validation of the Igbo version in patients with chronic low back pain. *Research Square (Research Square)*. Published online May 29, 2020. doi:https://doi.org/10.21203/rs.3.rs-27535/v1
294. Cheung MKT, Hung ATF, Poon PKK, et al. Validation of the World Health Organization Assessment Schedule II Chinese Traditional Version (WHODAS II CT) in persons with disabilities and chronic illnesses for Chinese population. *Disability and Rehabilitation*. 2014;37(20):1902-1907. doi:https://doi.org/10.3109/09638288.2014.989336
295. Buist-Bouwman MA, Ormel J, De Graaf R, et al. Psychometric properties of the World Health Organization Disability Assessment Schedule used in the European Study of the Epidemiology of Mental Disorders. *International Journal of Methods in Psychiatric Research*. 2008;17(4):185-197. doi:https://doi.org/10.1002/mpr.261
296. Jacek Wciórka, Schaeffer E, Switaj P, et al. [Bech-Rafaelsen Mania Scale and Young Mania Rating Scale--comparison of psychometric properties of the two instruments for rating a manic syndrome]. *Psychiatria Polska*. 2011;45(1):61-78.

297. Ronnachai Kongsakon, Daochompu Bhatanaprabhabhan. Validity and Reliability of the Young Mania Rating Scale: Thai version. *PubMed*. 2005;88(11):1598-1604.
298. E. Musoni-Rwililiza, Arnbjerg CJ, Rurangwa NU, et al. Adaption and validation of the Rwandese version of the Young Mania Rating Scale to measure the severity of a manic or hypomanic episode. *BMC Psychiatry*. 2024;24(1). doi:https://doi.org/10.1186/s12888-024-05890-1
299. Moritz Mühlbacher, Egger C, Kaplan P, et al. [Reliability and Concordance Validity of a German Version of the Young Mania Rating Scale (YMRS-D)]. *PubMed*. 2011;25(1):16-25.
300. Serrano E, Ezpeleta L, Alda JA, Matalí JL, San L. Psychometric Properties of the Young Mania Rating Scale for the Identification of Mania Symptoms in Spanish Children and Adolescents with Attention Deficit/Hyperactivity Disorder. *Psychopathology*. 2011;44(2):125-132. doi:https://doi.org/10.1159/000320893
301. Vilela JAA, Crippa JAS, Del-Ben CM, Loureiro SR. Reliability and validity of a Portuguese version of the Young Mania Rating Scale. *Brazilian Journal of Medical and Biological Research*. 2005;38(9):1429-1439. doi:https://doi.org/10.1590/s0100-879x2005000900019
302. Wasil AR, Venturo-Conerly KE, Gillespie S, Osborn TL, Weisz JR. In Their Own Words: Using Open-Ended Assessment to Identify Culturally Relevant Concerns among Kenyan Adolescents. *Culture, Medicine, and Psychiatry*. Published online February 2, 2021. doi:https://doi.org/10.1007/s11013-020-09706-1