



The **Strawberry** Hub
Empowering Patient Engagement 

Comprehensive Assessment Protocol for Gut-Brain Axis Evaluation

A Fully Free, Evidence-Based Battery for Clinical and Educational Settings

Version 1.0
Date: January 2026

Table of Contents

1. [Introduction](#)
 2. [Protocol Overview](#)
 3. [Assessment Tools](#)
 - A. [PROMIS Gastrointestinal Symptom Scales](#)
 - B. [Patient Health Questionnaire-9 \(PHQ-9\)](#)
 - C. [Generalized Anxiety Disorder-7 \(GAD-7\)](#)
 - D. [General Well-Being Schedule \(GWBS\)](#)
 - E. [Bristol Stool Form Scale \(BSFS\)](#)
 4. [Administration Guidelines](#)
 5. [Interpretation Framework](#)
 6. [References](#)
-

Introduction

The gut-brain axis represents a bidirectional communication system between the gastrointestinal tract and the central nervous system. Comprehensive assessment requires evaluation across multiple domains: gastrointestinal symptoms, mental health status, overall well-being, and objective bowel function indicators.

This protocol presents a fully free, evidence-based assessment battery consisting of five validated patient-reported outcome measures (PROMs) that collectively capture the multidimensional nature of gut-brain axis function. All instruments are in the public domain or freely available with proper attribution.

Rationale for Multi-Domain Assessment

Single-domain assessment fails to capture the complex interplay between physical symptoms, psychological distress, and quality of life that characterizes disorders of gut-brain interaction (DGBIs). This comprehensive battery addresses:

- **Gastrointestinal symptoms** (PROMIS GI domains)
 - **Depression** (PHQ-9)
 - **Anxiety** (GAD-7)
 - **Overall well-being** (GWBS)
 - **Stool consistency** (Bristol Stool Form Scale)
-

Protocol Overview

Assessment Timeline

Baseline Assessment (Week 0) –

- All five instruments

Follow-up Assessments (Weeks 1, 2, 3, 5, etc.) –

- PROMIS GI symptom scales (primary outcomes)
- PHQ-9 and GAD-7 (mental health monitoring) –
- GWBS (overall well-being)
- Bristol Stool Form Scale (daily patient diary, reviewed at each assessment)

Administration Method

- **Setting:** Clinical, research, or educational
 - **Mode:** Self-administered electronic format
 - **Time required:** Approximately 15-20 minutes for complete battery
 - **Language:** Available in multiple validated translations
-

Assessment Tools

1. PROMIS Gastrointestinal Symptom Scales

Description

The Patient-Reported Outcomes Measurement Information System (PROMIS) Gastrointestinal Symptom Scales comprise eight distinct symptom domains developed through rigorous NIH-funded methodology using Item Response Theory (IRT) and extensive patient input.

Eight Symptom Domains

1. **Gastroesophageal Reflux** (13 items)
2. **Disrupted Swallowing** (7 items)
3. **Diarrhea** (5 items)
4. **Bowel Incontinence/Soilage** (4 items)
5. **Nausea and Vomiting** (4 items)
6. **Constipation** (9 items)
7. **Belly Pain** (6 items)
8. **Gas/Bloating/Flatulence** (12 items)

Total: 60 items across all domains. Just administer those domains that are relevant to that patient.

Psychometric Properties

Reliability: - Excellent internal consistency across all scales - Test-retest reliability: 0.70-0.85 - IRT-based scoring ensures precision across symptom severity spectrum

Validity: - Developed using systematic literature review and patient focus groups - Excellent convergent validity with legacy GI instruments (correlations 0.40-0.80) - Construct validity confirmed across multiple GI conditions - Discriminates between healthy individuals and patients with GI disorders ($p < 0.001$) - Successfully validated in systemic sclerosis, inflammatory bowel disease, irritable bowel syndrome

Structural Validity: - Confirmatory factor analysis supports unidimensionality within each domain - Items satisfy IRT assumptions of monotonicity - Three domains (Bowel Incontinence, Gas/Bloating, Belly Pain) show strongest psychometric properties

Appropriate Use

Population: - Adults (18+ years) with diverse gastrointestinal conditions - General population (normative data available) - Validated for use in: - Irritable bowel syndrome - Inflammatory bowel disease - Functional dyspepsia - Gastroesophageal reflux disease - Systemic sclerosis with GI involvement - Research and clinical settings

Purpose: - Quantify GI symptom severity - Track symptom changes over time - Evaluate treatment response - Compare symptom burden across different GI conditions - Support clinical decision-making

Not Appropriate For: - Children under 18 years (pediatric versions needed) - Diagnostic purposes (should complement, not replace, clinical evaluation)

Scoring

Scoring Method: T-score metric

T-Score Interpretation: - Mean: 50 (U.S. general population) - Standard Deviation: 10 points - Higher scores = MORE symptoms (worse status)

T-Score Categories: - **T-score 40:** One SD better than average (fewer symptoms) - **T-score 50:** Average symptom level - **T-score 60:** One SD worse than average (more symptoms) - **T-score 70:** Two SD worse than average (severe symptoms)

Scoring Resources: - Manual scoring tables available (response pattern scoring preferred) - Online automated scoring: www.healthmeasures.net - Scoring manual: <https://www.healthmeasures.net/promis>

Important Scoring Notes: 1. Domains include skip patterns (if participant answers “never” to initial items, subsequent items may be skipped) 2. Response pattern scoring is preferred over sum score tables 3. Standard Error (SE) is provided with each T-score 4. Minimum one item response required per domain

Response Options: Most items use 5-point frequency scale: - Never (0) - Rarely (1) - Sometimes (2) - Often (3) - Always (4)

Access and Licensing

Cost: FREE - Public domain **Access:** www.healthmeasures.net **Attribution Required:** Yes - cite PROMIS and HealthMeasures **Modifications:** Not permitted without authorization

2. Patient Health Questionnaire-9 (PHQ-9)

Description

The PHQ-9 is a 9-item depression severity measure based on DSM-5 criteria for major depressive disorder. It is one of the most widely validated and commonly used depression screening and severity instruments globally.

Psychometric Properties

Reliability: - Internal consistency (Cronbach's α): 0.81-0.91 across diverse populations - Test-retest reliability: 0.81-0.85 - MacDonal'd's ω : >0.70 across populations - Consistently demonstrates excellent reliability in clinical and community samples

Validity: - Sensitivity: 71-95% (varies by cutoff and population) - Specificity: 61-88% (varies by cutoff and population) - Area Under Curve (AUC): 0.85-0.95 for detecting major depression - Convergent validity with clinician-rated scales (HAM-D): $r = 0.32-0.75$ - Strong correlation with other depression measures ($r = 0.70-0.85$) - Criterion validity confirmed against structured diagnostic interviews

Optimal Cutoff Scores by Population: - General population: ≥ 10 (sensitivity 88%, specificity 88%) - Clinical samples with mood disorders: ≥ 11 (sensitivity 72%, specificity 72%) - Adolescents (South Africa): ≥ 10 (sensitivity 91%, specificity 76%) - Primary care: ≥ 10 provides best balance

Factor Structure: - Predominantly one-factor (depression) - Some studies support two-factor model: - Cognitive/affective symptoms - Somatic symptoms

Appropriate Use

Population: - Adults (18+ years) - Adolescents (10-19 years) - validated in multiple countries - Successfully validated across diverse cultural contexts: - South African (isiXhosa, urban populations) - Lithuanian, Iranian, Japanese populations - African American, Mexican American women - HIV-affected populations - Medical patients (heart failure, inflammatory bowel disease)

Purpose: - Initial depression screening - Monitoring depression severity over time - Evaluating treatment response - Research outcomes assessment - **NOT a substitute for diagnostic interview**

Clinical Utility: - Excellent as initial screening tool - Positive screens should be followed by comprehensive clinical assessment - More sensitive than specific (higher false positive rates) - Can be used for criterion-based DSM-5 diagnosis using algorithm

Scoring

Scoring Method: Sum of 9 items

Response Options (past 2 weeks): - Not at all (0) - Several days (1) - More than half the days (2) - Nearly every day (3)

Total Score Range: 0-27

Severity Categories: - **0-4:** Minimal/none depression - **5-9:** Mild depression - **10-14:** Moderate depression - **15-19:** Moderately severe depression - **20-27:** Severe depression

Diagnostic Algorithm (Alternative Scoring): Can be used to make provisional DSM-5 diagnosis: - At least 5 of 9 symptoms present “more than half the days” - Must include either depressed mood OR anhedonia - Excludes item 9 (suicidal ideation) unless clearly related to depression

Clinical Significance: - Score ≥ 10 : Consider active treatment - Score ≥ 15 : Warrants active treatment with pharmacotherapy and/or psychotherapy - Score ≥ 20 : Immediate treatment; consider referral to mental health specialist

Item 9 (Suicidal Ideation) Protocol: Any score >0 on item 9 requires immediate clinical follow-up and risk assessment.

Access and Licensing

Cost: FREE - Public domain **Access:** www.phqscreeners.com **Available in:** 80+ language translations **Attribution:** Developed by Drs. Robert L. Spitzer, Janet B.W. Williams, Kurt Kroenke **Modifications:** Not permitted without permission

3. Generalized Anxiety Disorder-7 (GAD-7)

Description

The GAD-7 is a 7-item anxiety severity measure initially developed to screen for generalized anxiety disorder but validated as a broader measure of anxiety symptoms across various anxiety disorders.

Psychometric Properties

Reliability: - Internal consistency (Cronbach's α): 0.82-0.94 across populations - Test-retest reliability: 0.80-0.85 - MacDonal's ω : >0.80 consistently - Excellent reliability across clinical and community settings

Validity: - Sensitivity: 73-89% (varies by cutoff and population) - Specificity: 54-88% (varies by cutoff and population) - AUC: 0.82-0.92 for detecting anxiety disorders - Generally more sensitive than specific (higher false positives than PHQ-9) - Convergent validity with HAM-A: $r = 0.31-0.75$ - Performs well across different anxiety disorder types

Optimal Cutoff Scores: - ≥ 5 : Mild anxiety - ≥ 7 : Commonly used cutoff (sensitivity 73%, specificity 54-67%) - ≥ 8 : Alternative cutoff for better specificity (Peruvian hospitals) - ≥ 9 : Used in student populations (sensitivity 73%, specificity 70%) - ≥ 10 : Moderate anxiety (sensitivity 89%, specificity 82%)

Factor Structure: - Predominantly one-factor (anxiety) - Some studies support two-factor model: - Cognitive/affective symptoms - Somatic symptoms - Structure stable across gender and age groups (with minor variations)

Performance Notes: - Anxiety scales generally perform less well than depression scales - Clinical utility as specific GAD diagnostic tool is limited - Best used as general anxiety symptom measure rather than GAD-specific tool - Excellent for detecting individuals at risk requiring further assessment

Appropriate Use

Population: - Adults (18+ years) - Adolescents (validated in multiple countries) - Successfully validated across diverse populations: - South African populations - Lithuanian, Lebanese, Indonesian, Filipino populations - Kenyan healthcare workers - Inflammatory bowel disease patients - HIV-affected populations - General medical and psychiatric settings

Purpose: - Initial anxiety screening - Monitoring anxiety severity - Assessing treatment response - Broad anxiety symptom assessment (not GAD-specific) - Research outcomes - **NOT a substitute for diagnostic interview**

Clinical Utility: - Effective initial screening across anxiety disorder types: - Generalized anxiety disorder - Panic disorder - Social anxiety disorder - Post-traumatic stress disorder - All positive screens require comprehensive clinical assessment - Less specific than PHQ-9; expect more false positives - Useful in heterogeneous psychiatric populations

Scoring

Scoring Method: Sum of 7 items

Response Options (past 2 weeks): - Not at all (0) - Several days (1) - More than half the days (2) - Nearly every day (3)

Total Score Range: 0-21

Severity Categories: - **0-4:** Minimal anxiety - **5-9:** Mild anxiety - **10-14:** Moderate anxiety - **15-21:** Severe anxiety

Clinical Decision Points: - **Score 5-9:** Watchful waiting; consider psychoeducation - **Score 10-14:** Probable clinically significant anxiety; requires clinical assessment - **Score ≥ 15 :** Active treatment indicated; consider referral

Interpretation Cautions: - GAD-7 measures anxiety symptoms broadly, not specific disorders - High scores indicate need for further diagnostic evaluation - May capture general distress rather than specific anxiety pathology - Cultural and demographic factors may influence optimal cutoffs

Access and Licensing

Cost: FREE - Public domain **Access:** www.phqscreeners.com **Available in:** 80+ language translations (same website as PHQ-9) **Attribution:** Developed by Drs. Robert L. Spitzer, Kurt Kroenke, Janet B.W. Williams, Bernd Löwe **Modifications:** Not permitted without permission

4. General Well-Being Schedule (GWBS)

Description

The General Well-Being Schedule is a multidimensional self-report measure of subjective feelings of psychological well-being and distress. Originally developed in the 1970s by H.J. Dupuy for the U.S. National Center for Health Statistics, it assesses “inner personal state” across both positive and negative dimensions.

Versions Available

18-item version (most widely used): - 14 items on 6-point scale - 4 items on 10-point (0-10) scale - Completion time: 5-7 minutes

Six Subscale Domains

1. **Anxiety**
2. **Depression (Depressed Mood)**
3. **Positive Well-Being**
4. **Self-Control**
5. **Vitality (Energy)**
6. **General Health**

Psychometric Properties

Reliability: - Internal consistency (Cronbach's α): 0.90-0.94 (18-item version) - Test-retest reliability (Pearson's r): 0.81-0.85 - Intraclass correlation coefficient: 0.85 - Subscale reliability: 0.70-0.88 - Consistently excellent across diverse populations

Validity: - Concurrent validity with established measures: - General Health Questionnaire: $r = -0.75$ - State-Trait Anxiety Inventory: $r = -0.68$ - Center for Epidemiologic Studies Depression Scale: $r = -0.63$ - Profile of Mood States: $r = -0.53$ - Self-Rating Depression Scale: $r = -0.59$ - Construct validity: Strong negative correlations with distress measures - Discriminates well between clinical and non-clinical populations

Factor Structure: - Original six-factor structure supported - Alternative models validated: - Three-factor model (general mood/affect; life satisfaction/vitality; somatic complaints) - Four-factor model - Rasch analysis confirms fit when local dependencies addressed - Some cultural variations in factor structure

Responsiveness: - Sensitive to changes in clinical status - Successfully used to evaluate intervention outcomes - Demonstrates change across treatment conditions

Appropriate Use

Population: - Adults (general population and clinical samples) - Successfully validated across: - Japanese, Iranian populations - Mexican American, African American women - University students (non-clinical) - Patients with stress-related exhaustion/burnout - Individuals with chronic medical conditions - Yoga/meditation practitioners

Purpose: - Comprehensive assessment of psychological well-being - Monitoring overall mental health status - Evaluating intervention effects on well-being - Population health surveys - Screening for psychological distress - Complement to specific symptom measures - Assesses positive states, not just absence of pathology

Advantages: - Measures both positive well-being AND distress - Brief yet comprehensive - Suitable for repeated administration - Complements condition-specific measures - Free and widely available - Easy to complete (80% of respondents find it easy)

Scoring

Scoring Method: Sum of all items after reverse scoring negative items

18-Item Version Scoring:

Total Score Range: - Traditional scoring (1-6 scale): 22-132 - Alternative scoring (0-5 scale): 0-110

Interpretation Categories (Traditional Scoring): - **0-60:** Severe distress - **61-72:** Moderate distress - **73-110:** Positive well-being

Higher scores = Better well-being (opposite direction from PHQ-9/GAD-7)

Subscale Scoring: Each of six subscales can be scored separately: - Anxiety (5 items) - Depressed mood (3 items) - Positive well-being (4 items) - Self-control (2 items) - Vitality (4 items) - General health (1 or more items, depending on version)

Response Options: Mixed format: - Six-point scales for most items (ranging from descriptive anchors like “All of the time” to “None of the time”) - 0-10 rating scales for select items (e.g., “How have you been feeling in general?” from “The worst I have ever felt” to “The best I have ever felt”)

Rasch Conversion Available: For the 22-item PGWBI, interval-scale metric conversions are available for more precise measurement.

Access and Licensing

Cost: FREE - Public domain **Access:** Available through NIH/NHANES and various research repositories **Repository:** <https://cde.nlm.nih.gov/> (search “General Well-Being Schedule”) **Translations:** Japanese, Iranian validated versions available **Attribution:** Developed by H.J. Dupuy for National Center for Health Statistics **Modifications:** Various validated short forms exist

5. Bristol Stool Form Scale (BSFS)

Description

The Bristol Stool Form Scale is a visual analog scale that classifies human feces into seven categories based on form and consistency. Developed in 1997 by K.W. Heaton and S.J. Lewis at the University of Bristol, it provides a standardized, patient-friendly method for describing stool consistency that correlates with intestinal transit time.

Seven Stool Types

- Type 1:** Separate hard lumps, like nuts (hard to pass)
- Type 2:** Sausage-shaped but lumpy
- Type 3:** Like a sausage but with cracks on surface
- Type 4:** Like a sausage or snake, smooth and soft
- Type 5:** Soft blobs with clear-cut edges (passed easily)
- Type 6:** Fluffy pieces with ragged edges, mushy stool
- Type 7:** Watery, no solid pieces, entirely liquid

Clinical Categories

- Constipation:** Types 1-2
- Normal:** Types 3-5 (Type 4 often considered ideal)
- Diarrhea:** Types 6-7

Psychometric Properties

Reliability: - Inter-rater reliability (among experts): ICC = 0.88-0.89 (excellent) - Intra-rater reliability: ICC = 0.86-0.91; κ = 0.72-0.89 - General agreement within one category type: 95.1% of ratings - Modified versions (for children): - Inter-rater ICC: 0.71-0.79 - Intra-rater κ : 0.71-0.81

Reliability Challenges: - Absolute agreement across all 7 types is limited (only extreme types 1 and 7 have perfect agreement) - Reliability decreases when categorizing into Rome III clinical categories (constipation/normal/diarrhea) - Difficulties particularly at boundaries: Type 2 vs. 3 and Type 5 vs. 6 - Patient self-ratings may differ from expert ratings (patients tend to score lower/harder than experts)

Validity: - Validated as measure of colonic transit time (original purpose) - Correlates with stool water content (Spearman's ρ = 0.49, $p < 0.001$) - Discriminates between healthy volunteers and IBS-D patients ($p < 0.001$) - Overall accuracy in classifying stool models: 81% (κ = 0.78 - substantial) - Construct validity supported across multiple studies - Content validity established through systematic development

Responsiveness: - Successfully used to evaluate treatment response across numerous conditions - Recommended by Rome Foundation for IBS subtyping - FDA-recommended for clinical trials of IBS-D - Sensitive to changes in bowel function

Appropriate Use

Population: - Adults (original version) - Children and adolescents (validated with modifications) - Toilet-trained and non-toilet-trained toddlers (modified versions with diaper pictures) - Patients with: - Irritable bowel syndrome - Inflammatory bowel disease - Functional constipation - Chronic diarrhea - Faecal incontinence - General population (epidemiological studies)

Purpose: - Objective assessment of stool consistency - Classification of IBS subtypes (IBS-C, IBS-D, IBS-M) - Monitoring treatment response - Daily symptom tracking - Estimating gut transit time - Clinical communication tool between patients and providers - Research outcome measure

Validated Uses: - Diagnostic criteria for functional bowel disorders (Rome IV) - Evaluating response to: - Probiotics - Laxatives - Diet modifications - Pharmacological interventions - Dose titration for medications affecting bowel function - Longitudinal monitoring in GI disorders

Advantages: - Simple, visual, patient-friendly - Quick to use (<1 minute) - No complex scoring required - Free and widely available - Translated into multiple languages (Spanish, Portuguese, Polish, Persian, Dutch, etc.) - Suitable for repeated daily assessment

Limitations: - Reliability decreases at clinical decision boundaries - Patient-reported may differ from objective assessment - Limited psychometric evaluation for construct validity - Best used for identifying changes of ≥ 2 stool types rather than single-type differences - Some age-related bias (younger children less reliable)

Scoring

Scoring Method: Visual classification

Administration: 1. Present patient with BSFS chart showing all 7 types 2. Patient selects type that best matches their stool 3. Can be used for: - Single assessments - Daily diaries - Predominant pattern over time period

Interpretation:

Individual Stool Assessment: - Type 1-2: Indicates slow transit (constipation) - Type 3-4-5: Normal transit - Type 6-7: Indicates fast transit (diarrhea)

Rome IV IBS Subtyping: Based on proportion of bowel movements in each category when NOT on medication:

- **IBS-C** (IBS with constipation): $\geq 25\%$ Type 1-2, $< 25\%$ Type 6-7
- **IBS-D** (IBS with diarrhea): $\geq 25\%$ Type 6-7, $< 25\%$ Type 1-2
- **IBS-M** (IBS mixed): $\geq 25\%$ Type 1-2 AND $\geq 25\%$ Type 6-7
- **IBS-U** (IBS unclassified): Insufficient abnormal stools to meet criteria for C, D, or M

Daily Diary Recording: Patients record predominant stool type for each bowel movement. Summary can be expressed as: - Average type over time period - Percentage of movements in each category - Predominant type - Range of types

Clinical Significance: - Change of ≥ 2 types generally considered clinically meaningful - Shift from Type 1-2 to Type 3-5 indicates successful constipation treatment - Shift from Type 6-7 to Type 3-5 indicates successful diarrhea treatment

Training Recommendations: - Brief training improves reliability, especially for Types 2, 3, 5, 6 - Having descriptive text read aloud improves pediatric reliability - Showing examples of stools in diapers helpful for caregivers of young children

Access and Licensing

Cost: FREE - Widely available public resource **Access:** Available in published literature and numerous online sources **Original Publication:** Scandinavian Journal of Gastroenterology 1997;32(9):920-924 **Translations:** Spanish, Brazilian Portuguese, Polish, Persian, Dutch, validated versions **Modifications:** Modified versions for children validated **Usage:** No permission required for clinical or research use **Citation:** Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol. 1997;32(9):920-924.

Administration Guidelines

Recommended Assessment Sequence

1. Initial Session (Baseline - Week 0)

Order of administration: 1. Bristol Stool Form Scale (explain daily diary) 2. PROMIS GI symptom scales (select relevant domains) 3. PHQ-9 4. GAD-7 5. GWBS

Rationale for sequence: - Begin with simplest, least intrusive measure (BSFS) - Progress from physical symptoms to psychological - End with comprehensive well-being measure

Follow-Up Assessments

Weekly/Biweekly Monitoring: - PROMIS GI domains (primary outcomes) - PHQ-9 (if baseline score ≥ 5) - GAD-7 (if baseline score ≥ 5) - GWBS - Bristol Stool Scale diary review

Frequency Recommendations: - PROMIS GI: Every assessment - PHQ-9/GAD-7: Every 2-4 weeks (more frequently if elevated) - GWBS: Every assessment - BSFS: Daily patient diary, reviewed at each assessment

Administration Format

Paper-Based: - Advantages: No technology barriers, accessible to all - Considerations: Requires manual scoring, data entry - Best for: Limited-resource settings, older adults

Electronic/Computer-Based: - Advantages: Automated scoring, immediate results, data storage - Considerations: Requires technology access, computer literacy - Best for: Research settings, tech-savvy populations - PROMIS offers computer-adaptive testing (CAT) online

Practical Considerations

Time Burden:

Complete battery: 15-20 minutes

PROMIS (select domains): 5-10 minutes

PHQ-9: 2-3 minutes

GAD-7: 2-3 minutes

GWBS: 5-7 minutes

BSFS: <1 minute

Literacy Requirements: - All instruments require approximately 6th-8th grade reading level - Visual aids (BSFS) reduce literacy barriers - Having questions read aloud improves comprehension and reliability - Translated versions available for non-English speakers

Patient Burden Mitigation: - Use only relevant PROMIS domains (don't administer all 8 if not applicable) - Consider alternating PHQ-9/GAD-7 administration if burden is concern - BSFS diary can be simplified to once-daily "predominant type" - Electronic administration reduces completion time

Interpretation Framework

Individual Instrument Interpretation

See individual tool sections above for: - Scoring ranges - Severity categories - Clinical significance thresholds

Integrated Interpretation Approach

The comprehensive battery allows identification of distinct clinical patterns:

Pattern 1: High GI Symptoms + High Psychological Distress

- **PROMIS GI:** T-scores >60 across multiple domains
- **PHQ-9:** ≥10
- **GAD-7:** ≥10
- **GWBS:** <73
- **Interpretation:** Classic gut-brain axis dysregulation; psychological factors likely exacerbating GI symptoms
- **Implications:** Multidisciplinary treatment (GI + mental health) indicated

Pattern 2: High GI Symptoms + Low Psychological Distress

- **PROMIS GI:** T-scores >60
- **PHQ-9:** <5
- **GAD-7:** <5
- **GWBS:** >73
- **Interpretation:** Primary GI pathology with good psychological resilience
- **Implications:** Focus on GI-specific treatments; monitor for psychological sequelae

Pattern 3: Moderate GI Symptoms + High Psychological Distress

- **PROMIS GI:** T-scores 50-60
- **PHQ-9:** ≥10 and/or **GAD-7** ≥10
- **GWBS:** <73
- **Interpretation:** Psychological factors may be primary; somatic manifestation of distress
- **Implications:** Address mental health as priority; GI symptoms may improve secondarily

Pattern 4: Improvement in GI + Persistent Psychological Symptoms

- **PROMIS GI:** Decreased by ≥5 T-score points
- **PHQ-9/GAD-7:** Unchanged or worsening
- **Interpretation:** Physical improvement not translating to psychological recovery
- **Implications:** Add or intensify psychological intervention

Clinical Significance Thresholds

Minimal Clinically Important Difference (MCID):

- **PROMIS GI scales:** Change of 5 T-score points (0.5 SD)
- **PHQ-9:** Change of 5 points
- **GAD-7:** Change of 4-5 points
- **GWBS:** Change of 10-15 points (18-item version)
- **BSFS:** Change of ≥2 types

Response to Treatment Criteria

Favorable Response: - PROMIS GI: Decrease ≥5 T-score points in primary symptom domain - PHQ-9/GAD-7: Decrease ≥30-50% from baseline - GWBS: Increase ≥10 points - BSFS: Movement toward Types 3-5 (normal range)

Excellent Response: - PROMIS GI: Return to near-normal range (T-score <55) - PHQ-9: Decrease to <5 (remission) - GAD-7: Decrease to <5 (remission) - GWBS: Increase to >73 (positive well-being range) - BSFS: Consistent Types 3-5

Red Flags Requiring Immediate Action

1. **PHQ-9 item 9 (suicidal ideation) >0:** Immediate risk assessment
2. **PHQ-9 \geq 20 OR GAD-7 \geq 15:** Consider specialist referral
3. **GWBS <60:** Severe distress; requires clinical evaluation
4. **Worsening scores despite treatment:** Reassess diagnosis and treatment plan
5. **BSFS Type 7 with alarm features:** Rule out infectious/inflammatory causes

Documentation and Communication

Essential Documentation: - Date of assessment - Scores for each instrument - Interpretation category - Clinical action taken - Patient's subjective response to questionnaire burden

Communication with Patients: - Explain that questionnaires help track progress objectively - Normalize fluctuations in scores - Emphasize collaborative nature of monitoring - Provide context for score meanings (avoid over-medicalizing normal variation)

Communication with Referring Providers: - Summary of baseline and current scores - Interpretation of clinical pattern - Treatment response trajectory - Specific recommendations based on findings

References

PROMIS GI Scales

1. Spiegel BMR, Hays RD, Bolus R, et al. Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) Gastrointestinal Symptom Scales. *Am J Gastroenterol*. 2014;109(11):1804-1814.
2. van der Ende-van Loon MCM, Klomberg RCW, Jansen EJS, et al. Dutch-Flemish translation and validation of the gastrointestinal symptom scales from the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Patient Rep Outcomes*. 2023;7(1):119.
3. Almazar AE, Chock NK, Wood RK, et al. Construct validity of the PROMIS GI symptom scales in systemic sclerosis. *Arthritis Care Res*. 2014;66(12):1800-1808.
4. HealthMeasures. PROMIS Gastrointestinal Symptoms Scoring Manual. Updated February 2024. Available at: www.healthmeasures.net

PHQ-9

5. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-613.
6. Gyasi RM, Agyemang-Duah W, Mensah CM, Arthur F. Examining the psychometric properties of the Patient Health Questionnaire-9 and Generalized Anxiety Disorder-7 among young urban South African women. *J Affect Disord.* 2024;363:133-141.
7. Wittkamp KA, Naeije L, Schene AH, et al. Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review. *Gen Hosp Psychiatry.* 2007;29(5):388-395.
8. Gelaye B, Tadesse MG, Williams MA, et al. Detecting depression and anxiety among adolescents in South Africa: validity of the isiXhosa Patient Health Questionnaire-9 and Generalized Anxiety Disorder-7. *J Adolesc Health.* 2023;72(5):784-791.

GAD-7

9. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-1097.
10. Kroenke K, Spitzer RL, Williams JB, et al. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med.* 2007;146(5):317-325.
11. Beard C, Björgvinsson T. Beyond generalized anxiety disorder: psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *J Anxiety Disord.* 2014;28(6):547-552.
12. Richardson T, Wrightman M, Yeebo M, et al. Reliability and score ranges of the PHQ-9 and GAD-7 in a primary and secondary care mental health service. *J Psychosoc Rehabil Ment Health.* 2017;4:237-240.

GWBS

13. Dupuy HJ. The Psychological General Well-Being (PGWB) Index. In: Wenger NK, Mattson ME, Furberg CF, Elinson J, eds. *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies.* New York: Le Jacq Publishing; 1984:170-183.

14. Nakayama T, Toyoda H, Ohno K, et al. Validity, reliability and acceptability of the Japanese version of the General Well-Being Schedule (GWBS). *Qual Life Res.* 2000;9(5):529-539.
15. Alagheband M, Ahmadabadi NM, Abdollahi A. Psychometric characteristics of the General Well-Being Schedule (GWB) in an Iranian sample. *Int J Body Mind Culture.* 2013;1(1):69-76.
16. McDowell I. The General Well-Being Schedule. In: McDowell I. *Measuring Health: A Guide to Rating Scales and Questionnaires.* 3rd ed. Oxford University Press; 2006.
17. Lundgren-Nilsson A, Jonsdottir IH, Pallant J, Ahlborg G. Construct validity of the Psychological General Well Being Index (PGWBI) in a sample of patients undergoing treatment for stress-related exhaustion: a Rasch analysis. *Health Qual Life Outcomes.* 2013;11:2.

Bristol Stool Form Scale

18. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol.* 1997;32(9):920-924.
19. Blake MR, Raker JM, Whelan K. Validity and reliability of the Bristol Stool Form Scale in healthy adults and patients with diarrhoea-predominant irritable bowel syndrome. *Aliment Pharmacol Ther.* 2016;44(7):693-703.
20. Chumpitazi BP, Self MM, Czyzewski DI, et al. Bristol Stool Form Scale reliability and agreement decreases when determining Rome III stool form designations. *Neurogastroenterol Motil.* 2016;28(3):443-448.
21. Lane MM, Czyzewski DI, Chumpitazi BP, Shulman RJ. Reliability and validity of a modified Bristol Stool Form Scale for children. *J Pediatr.* 2011;159(3):437-441.
22. Wegh CAM, Schoterman MHC, Vaughan EE, Belzer C, Benninga MA. The modified Bristol Stool Form Scale: a reliable and valid tool to score stool consistency in Dutch (non) toilet trained toddlers. *J Pediatr Gastroenterol Nutr.* 2021;73(2):186-191.

Gut-Brain Axis and Assessment

23. Drossman DA, Hasler WL. Rome IV-Functional GI Disorders: Disorders of Gut-Brain Interaction. *Gastroenterology.* 2016;150(6):1257-1261.
24. Simrén M, Törnblom H, Palsson OS, et al. Cumulative effects of psychologic distress, visceral hypersensitivity, and abnormal transit on patient-reported outcomes in irritable bowel syndrome. *Gastroenterology.* 2019;157(2):391-402.

25. Palssson OS, Whitehead WE, van Tilburg MAL, et al. Development and validation of the Rome IV diagnostic questionnaire for adults. *Gastroenterology*. 2016;150(6):1481-1491.
-

Quick Reference Score Summary

Instrument	Score Range	Clinical Significance	Access
PROMIS GI	T-score: 20-80	≥60 = 1 SD worse than average	www.healthmeasures.net
PHQ-9	0-27	≥10 = Moderate depression	www.phqscreeners.com
GAD-7	0-21	≥10 = Moderate anxiety	www.phqscreeners.com
GWBS	22-132 (or 0-110)	<73 = Distress; >73 = Positive well-being	NIH repositories
BSFS	Types 1-7	1-2 = Constipation; 6-7 = Diarrhea	Public domain

Recommended Assessment Schedule Template

Baseline (Week 0): All 5 instruments + daily BSFS diary

Week 1: PROMIS GI + BSFS review

Week 2: PROMIS GI + PHQ-9/GAD-7 + GWBS + BSFS review

Week 3: PROMIS GI + BSFS review

Week 5: All 5 instruments (reassessment)

Monthly thereafter: PROMIS GI + PHQ-9/GAD-7 + GWBS + BSFS review

Adaptation for Different Settings

Clinical Practice: - Use full battery at baseline and every 4-8 weeks - Weekly PROMIS GI + BSFS for symptom tracking - PHQ-9/GAD-7 if baseline elevated or symptoms emerge

Research Studies: - Full battery at all designated timepoints - Consider PROMIS computer-adaptive testing for efficiency - Daily BSFS diaries for granular data

Educational Settings: - Use as teaching cases with de-identified data - Demonstrate multidimensional assessment approach - Practice interpretation using clinical vignettes

Resource-Limited Settings: - Prioritize: PROMIS GI + GWBS + BSFS (core triad) - Add PHQ-9 or GAD-7 based on clinical presentation - Paper-based administration acceptable

Document End

Citation for This Protocol: Comprehensive Assessment Protocol for Gut-Brain Axis Evaluation: A Fully Free, Evidence-Based Battery for Clinical and Educational Settings. Version 1.0. January 2026.

For Questions or Updates: This protocol represents a synthesis of publicly available research and validated assessment tools. Users should consult primary literature for the most current psychometric data and clinical guidelines.

Acknowledgments: This protocol was developed based on extensive review of published psychometric literature and represents best practices as of January 2026. All instruments described are freely available or in the public domain.

Appendices

The following appendices contain the complete assessment instruments for clinical use.

Appendix A: Patient Health Questionnaire-9 (PHQ-9)

Instructions: Over the last 2 weeks, how often have you been bothered by any of the following problems?

Item	Not at all (0)	Several days (1)	More than half the days (2)	Nearly every day (3)
1. Little interest or pleasure in doing things	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Feeling down, depressed, or hopeless	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Trouble falling or staying asleep, or sleeping too much	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Feeling tired or having little energy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Poor appetite or overeating	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Feeling bad about yourself - or that you are a failure or have let yourself or your family down	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Trouble concentrating on things, such as reading the newspaper or watching television	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. Moving or speaking so slowly that other people could have noticed. Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual

9. Thoughts that you would be better off dead, or of hurting yourself in some way

Scoring: Sum all items (0-27). Categories: 0-4 minimal, 5-9 mild, 10-14 moderate, 15-19 moderately severe, 20-27 severe depression.

Source: Public domain. Available at www.phqscreeners.com

Appendix B: Generalized Anxiety Disorder-7 (GAD-7)

Instructions: Over the last 2 weeks, how often have you been bothered by the following problems?

Item	Not at all (0)	Several days (1)	More than half the days (2)	Nearly every day (3)
1. Feeling nervous, anxious, or on edge	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Not being able to stop or control worrying	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Worrying too much about different things	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Trouble relaxing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Being so restless that it is hard to sit still	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Becoming easily annoyed or irritable	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Feeling afraid, as if something awful might happen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Scoring: Sum all items (0-21). Categories: 0-4 minimal, 5-9 mild, 10-14 moderate, 15-21 severe anxiety.

Source: Public domain. Available at www.phqscreeners.com

Appendix C: Bristol Stool Form Scale (BSFS)

Instructions: Please select the type that best describes your stool consistency.

Type	Description
Type 1	Separate hard lumps, like nuts (hard to pass) - CONSTIPATION
Type 2	Sausage-shaped but lumpy - CONSTIPATION
Type 3	Like a sausage but with cracks on surface - NORMAL
Type 4	Like a sausage or snake, smooth and soft - NORMAL (IDEAL)
Type 5	Soft blobs with clear-cut edges (passed easily) - NORMAL
Type 6	Fluffy pieces with ragged edges, mushy stool - DIARRHEA
Type 7	Watery, no solid pieces, entirely liquid - DIARRHEA

Clinical Use: Types 1-2 indicate constipation, Types 3-5 are normal (Type 4 ideal), Types 6-7 indicate diarrhea.

Source: Public domain. Original publication: Scand J Gastroenterol 1997;32(9):920-924

Appendix D: General Well-Being Schedule (GWBS) - 18 Item

Instructions: This section contains questions about how you feel and how things have been going with you. For each question, mark the answer that best applies to you during the past month.

Questions 1-14: 6-point response scale

1. How have you been feeling in general?

In excellent spirits (5) | In very good spirits (4) | In good spirits mostly (3) | I have been up and down in spirits a lot (2) | In low spirits mostly (1) | In very low spirits (0)

2. Have you been bothered by nervousness or your nerves?

Extremely so - to the point where I could not work or take care of things (0) | Very much so (1) | Quite a bit (2) | Some - enough to bother me (3) | A little (4) | Not at all (5)

3. Have you been in firm control of your behavior, thoughts, emotions, or feelings?

Yes, definitely so (5) | Yes, for the most part (4) | Generally so (3) | Not too well (2) | No, and I am somewhat disturbed (1) | No, and I am very disturbed (0)

4. Have you felt so sad, discouraged, hopeless, or had so many problems that you wondered if anything was worthwhile?

Extremely so - to the point that I have just about given up (0) | Very much so (1) | Quite a bit (2) | Some - enough to bother me (3) | A little bit (4) | Not at all (5)

5. Have you been under or felt you were under any strain, stress, or pressure?

Yes - almost more than I could bear (0) | Yes - quite a bit of pressure (1) | Yes - some, more than usual (2) | Yes - some, but about usual (3) | Yes - a little (4) | Not at all (5)

6. How happy, satisfied, or pleased have you been with your personal life?

Extremely happy - could not have been more satisfied or pleased (5) | Very happy (4) | Fairly happy (3) | Satisfied - pleased (2) | Somewhat dissatisfied (1) | Very dissatisfied (0)

7. Have you had any reason to wonder if you were losing your mind, or losing control over the way you act, talk, think, feel, or of your memory?

Not at all (5) | Only a little (4) | Some - but not enough to be concerned (3) | Some - and I have been a little concerned (2) | Some - and I am quite concerned (1) | Yes, very much so - and I am very concerned (0)

8. Have you been anxious, worried, or upset?

Extremely so - to the point of being sick or almost sick (0) | Very much so (1) | Quite a bit (2) | Some - enough to bother me (3) | A little bit (4) | Not at all (5)

9. Have you been waking up fresh and rested?

Every day (5) | Most every day (4) | Fairly often (3) | Less than half the time (2) | Rarely (1) | None of the time (0)

10. Have you been bothered by any illness, bodily disorder, aches, or pains?

All the time (0) | Most of the time (1) | A good bit of the time (2) | Some of the time (3) | A little of the time (4) | None of the time (5)

11. Has your daily life been full of things that were interesting to you?

All the time (5) | Most of the time (4) | A good bit of the time (3) | Some of the time (2) | A little of the time (1) | None of the time (0)

12. Have you felt downhearted and blue?

All the time (0) | Most of the time (1) | A good bit of the time (2) | Some of the time (3) | A little of the time (4) | None of the time (5)

13. Have you been feeling emotionally stable and sure of yourself?

All the time (5) | Most of the time (4) | A good bit of the time (3) | Some of the time (2) | A little of the time (1) | None of the time (0)

14. Have you felt tired, worn out, used up, or exhausted?

All the time (0) | Most of the time (1) | A good bit of the time (2) | Some of the time (3) | A little of the time (4) | None of the time (5)

Questions 15-18: 0-10 rating scale

15. How concerned or worried about your health have you been?

Rate from 0 (Not concerned at all) to 10 (Very concerned)

16. How relaxed or tense have you been?

Rate from 0 (Very tense) to 10 (Very relaxed)

17. How much energy, pep, or vitality have you felt?

Rate from 0 (No energy at all, listless) to 10 (Very energetic, dynamic)

18. How depressed or cheerful have you been?

Rate from 0 (Very depressed) to 10 (Very cheerful)

Scoring: Sum all items. Traditional range: 22-132 (or 0-110 alternative). Lower scores indicate distress, higher scores indicate well-being. 0-60: severe distress, 61-72: moderate distress, 73-110: positive well-being.

Source: Public domain. Available through NIH/NHANES repositories.

Appendix E: PROMIS Gastrointestinal Symptom Scales

Note: *The complete PROMIS GI assessment contains 60 items across 8 domains. Due to space constraints, domain descriptions are provided below. Full item banks are freely available at www.healthmeasures.net*

Eight Symptom Domains:

1. Gastroesophageal Reflux (13 items)

Assesses heartburn, acid regurgitation, and related symptoms over the past 7 days.

2. Disrupted Swallowing (7 items)

Evaluates difficulty swallowing solids and liquids.

3. Diarrhea (5 items)

Measures frequency and urgency of loose/watery stools.

4. Bowel Incontinence/Soilage (4 items)

Assesses involuntary loss of stool.

5. Nausea and Vomiting (4 items)

Evaluates nausea severity and vomiting episodes.

6. Constipation (9 items)

Measures stool frequency, straining, and completeness of evacuation.

7. Belly Pain (6 items)

Assesses abdominal pain severity and impact.

8. Gas/Bloating/Flatulence (12 items)

Evaluates gas, bloating, and abdominal distension.

Response Format: Most items use a 5-point frequency scale (Never, Rarely, Sometimes, Often, Always) referring to the past 7 days.

Scoring: T-score metric (Mean=50, SD=10 in U.S. general population). Higher scores indicate MORE symptoms (worse status). T-score ≥ 60 indicates 1 SD worse than average.

Access: FREE at www.healthmeasures.net. Registration required. Attribution: "PROMIS® was funded with cooperation from the NIH Science and Technology Research Network (STTR) and the NIH Roadmap for Medical Research."

Implementation Notes

All instruments in this protocol are completely free for clinical, educational, and research use. No licensing fees or permissions are required. Clinicians may photocopy these questionnaires for patient use. Electronic administration is also permitted.

Citation: *Comprehensive Assessment Protocol for Gut-Brain Axis Evaluation: A Fully Free, Evidence-Based Battery for Clinical and Educational Settings. Version 1.0. January 2026.*