Analysis of Psychometric Properties of the Patient-Reported Outcome Measure for Inflammatory Bowel Disease: Postprint

Authors: Luo Na, Ruan Yanqin, Lei Pingguang, Wan Chonghua, Wan Keyan, Song Ying, Chen Ying, Chen Ying

Date: 2023-07-18T00:00:00+00:00

Abstract

Background: The level of patient-reported outcomes in inflammatory bowel disease (IBD) has received attention. Currently, there are few mature IBD patient-reported outcome scales with Chinese cultural characteristics, and developed scales require rigorous evaluation.

Objective: To analyze and evaluate the measurement properties of the Inflammatory Bowel Disease Patient-Reported Outcome Measurement Scale [PROISCD-IBD (V1.0)], providing a basis for scientific assessment of IBD patient-reported outcomes.

Methods: PROISCD-IBD (V1.0) was administered to 274 IBD patients who visited the gastroenterology outpatient and inpatient departments of the First Affiliated Hospital of Kunming Medical University and the Affiliated Hospital of Guangdong Medical University between October 2020 and January 2022. PROISCD-IBD (V1.0) comprises one generic module and one IBD-specific module. The generic module contains 30 items, which are further divided into four domains: Physical Health (PHD), Mental Health (MHD), Social Health (SHD), and Spiritual/Belief Health (SBD). The IBD-specific module covers four facets: Digestive System Symptoms (DSS), Extra-intestinal Symptoms (EXS), Special Psychological aspects (SPP), and Treatment Side Effects (TSE). The full scale contains 44 items. Cronbach' s α coefficient and split-half coefficient were used to assess reliability; correlation coefficient method, exploratory factor analysis, and structural equation modeling were used to analyze structural validity; t-test was used to analyze clinical validity of each domain.

Results: The Cronbach' s α coefficients for the five domains of PROISCD-IBD (V1.0)—PHD, MHD, SHD, SBD, and the specific module—were 0.732, 0.838,

0.781, 0.673, and 0.884, respectively, with a total scale Cronbach' s α coefficient of 0.932. The split-half coefficients for the five domains were 0.669, 0.859, 0.610, 0.494, and 0.795, respectively, with a total scale split-half reliability of 0.879. Correlation analysis revealed that the correlation coefficients between PHD, MHD, SHD, and SBD domain scores and the generic module score were all >0.6 (P<0.05). Exploratory factor analysis extracted three principal components with a cumulative variance contribution rate of 58.05%. The structural equation model showed 2/df=2.568, root mean square error of approximation (RMSEA)=0.076, normed fit index (NFI)=0.677, non-normed fit index (NNFI)=0.774, comparative fit index (CFI)=0.772, incremental fit index (IFI)=0.774, and standardized root mean square residual (SRMR)=0.1031. IBD patients were divided into active phase (n=90) and remission phase (n=184) according to clinical staging; patients in remission scored higher than those in active phase in all domains, the generic module, the specific module, and total scale score (P<0.05).

Conclusion: PROISCD-IBD (V1.0) demonstrates good reliability and validity and can be used for measuring patient-reported outcomes in IBD patients.

Full Text

Analysis of the Measurement Characteristics of Inflammatory Bowel Disease Patient-Reported Outcomes Measurement Scale

LUO Na[1], RUAN Yanqin[1], LEI Pingguang[2], WAN Chonghua[3], WANG Keyan[2], SONG Ying[1], CHEN Ying[1]*

[1]School of Public Health, Kunming Medical University, Kunming 650500, China
[2]Songgang People' s Hospital, Baoan District, Shenzhen 518105, China
[3]Center for Quality of Life and Applied Psychology, Guangdong Medical University, Dongguan 523808, China
*Corresponding author: CHEN Ying, Professor, Doctoral supervisor; E-mail: chenying2128@126.com

Abstract

Background The level of patient-reported outcomes in inflammatory bowel disease (IBD) has garnered increasing attention. Currently, there are few mature IBD patient-reported outcome scales with Chinese cultural characteristics, and developed scales require rigorous evaluation. **Objective** To analyze and evaluate the measurement properties of the Inflammatory Bowel Disease Patient-Reported Outcome Measurement Scale (PROISCD-IBD [V1.0]), providing a basis for scientific assessment of patient-reported outcomes in IBD patients. **Methods** PROISCD-IBD (V1.0) was administered to 274 IBD patients treated in the gastroenterology outpatient and inpatient departments of the First Affiliated Hospital of Kunming Medical University and the Affil-

iated Hospital of Guangdong Medical University between October 2020 and January 2022. PROISCD-IBD (V1.0) comprises one common module and one IBD-specific module. The common module contains 30 items divided into four domains: physical health (PHD), mental health (MHD), social health (SHD), and spiritual/belief health (SBD). The IBD-specific module (TIBD) covers 44 items. Reliability was assessed using Cronbach' s α coefficient and split-half coefficient; construct validity was analyzed using correlation coefficient methods, exploratory factor analysis, and structural equation modeling; and clinical validity of each domain was analyzed using t-tests. **Results** The Cronbach' s α coefficients for the PHD, MHD, SHD, SBD, and specific module domains of PROISCD-IBD (V1.0) were 0.732, 0.838, 0.781, 0.673, and 0.884, respectively, with a total scale Cronbach's α coefficient of 0.932. The split-half coefficients for these five domains were 0.669, 0.859, 0.610, 0.494, and 0.795, respectively, with a total scale split-half reliability of 0.879. Correlation analysis revealed that correlation coefficients between domain scores (PHD, MHD, SHD, SBD) and the common module score were all >0.6 (P<0.05). Exploratory factor analysis extracted three principal components with a cumulative variance contribution rate of 58.05%. The structural equation model showed ^{2}/df=2.568, root-mean-square error of approximation (RMSEA)=0.076, normed fit index (NFI)=0.677, non-normed fit index (NNFI)=0.774, comparative fit index (CFI)=0.772, incremental fit index (IFI)=0.774, and standardized root-mean-square residual (SRMR)=0.1031. IBD patients were divided by clinical stage into active (n=90) and remission (n=184) phases; patients in remission scored higher than those in the active phase across all domains, common modules, specific modules, and total scale scores (P<0.05). **Conclusion** PROISCD-IBD (V1.0) demonstrates good reliability and validity for measuring patient-reported outcomes in IBD patients.

Keywords: Inflammatory bowel diseases; Patient reported outcome measures; Reliability; Validity; Structural equation model

---

Inflammatory bowel disease (IBD), which generally includes ulcerative colitis (UC) and Crohn' s disease (CD), is a chronic, non-specific, recurrent inflammatory disease primarily affecting the gastrointestinal tract. Its pathogenesis remains unclear. In recent years, the incidence of IBD has shown an increasing trend, not only in Europe, America, and Western countries but also in the United States, Asia, Africa, and Eastern Europe. Current epidemiological data and statistical figures in China indicate that both the prevalence and incidence of IBD are rising, with the number of IBD patients in China expected to reach 1.5 million by 2025. IBD often begins in adolescence, with recurrent episodes that gradually worsen. Most patients require expensive treatments and lifelong medical care, severely impacting their studies, work, and daily life while reducing quality of life for both patients and their families.

Patient-reported outcomes (PRO) serve as an effective supplement to quality of life assessment, reflecting patients' physical health status, functional status, and

treatment experience without interpretation from medical staff or other personnel, thus directly reflecting patients' quality of life levels. Currently, over ten scales have been developed abroad to measure IBD patient-reported outcomes from perspectives such as symptom scoring and quality of life assessment. However, domestically, there is a lack of recognized disease-specific scales, with most studies using translated versions of the IBD-Q scale for quality of life measurement in IBD patients. Due to cultural differences, these may not accurately estimate the reported outcome levels of Chinese IBD patients.

In recent years, Professor Wan Chonghua and his research team have developed and promoted the Quality of Life Instrument for Chronic Diseases (QLICD) system and, building upon QLICD, developed the Patient-Reported Outcome Measurement System for Chronic Diseases (PROISCD), which includes a common module and over ten disease-specific scales for chronic conditions. This study primarily re-evaluates the reliability and validity of the PROISCD-IBD (V1.0) common module, develops and analyzes the measurement characteristics of the specific module, thereby constructing a complete PROISCD-IBD (V1.0) scale system.

1.1 Study Subjects

This study selected IBD patients treated in the gastroenterology outpatient and inpatient departments of the First Affiliated Hospital of Kunming Medical University and the Affiliated Hospital of Guangdong Medical University between October 2020 and January 2022 as the primary study population. Inclusion criteria were: (1) meeting the diagnostic criteria for UC and CD recommended by the Gastroenterology Branch of the Chinese Medical Association; (2) primary school education or higher; and (3) signed informed consent. Exclusion criteria were: (1) patients with mental disorders; and (2) patients with severe conditions such as malignant tumors. Investigators, after unified training and in their capacity as physicians, provided a brief explanation of the PROISCD-IBD (V1.0) scale and, upon obtaining patient consent, had patients complete the scale independently. Completed scales were collected and checked, with missing items filled in on-site, and invalid scales were excluded. All patients completed the scale on their first day of hospitalization or during their outpatient visit. This study protocol was approved by the Medical Ethics Committee of Kunming Medical University (KMMU2021MEC031).

1.2 Survey Instrument

PROISCD-IBD (V1.0) was collaboratively developed by Professor Wan Chonghua and his research team, comprising one common module and one IBD-specific module (TIBD), demonstrating excellent systematicity. The common module contains 30 items divided into four domains: physical health (PHD), mental health (MHD), social health (SHD), and spiritual/belief health (SBD). TIBD covers four facets: digestive system symptoms (DSS),

extraintestinal symptoms (EXS), special psychological symptoms (SPP), and treatment side effects (TSE). The full scale contains 44 items.

This scale employs a 5-point Likert method. Scale items are divided into positively and negatively worded items; scores for positively worded items require no conversion, while scores for negatively worded items are calculated as 6 minus the original item score. Standardized conversion is also required for facet and domain scores calculated from adjusted items.

1.3 Statistical Methods

SPSS 27.0 and AMOS 24.0 were used for statistical analysis. Cronbach' s α coefficient was calculated for each facet, domain, and the total scale. Items within each domain and the total scale were divided into two halves according to odd and even item numbers, with negatively worded items reverse-scored before calculating correlations between the two halves using the Spearman-Brown formula. Construct validity was analyzed using Spearman correlation coefficient methods, factor analysis, and structural equation modeling. Clinical validity of each domain was analyzed using t-tests. $P<0.05$ was considered statistically significant.

2.2 Reliability

The Cronbach' s α coefficients for the PHD, MHD, SHD, SBD, and TIBD domains of PROISCD-IBD (V1.0) were 0.732, 0.838, 0.781, 0.673, and 0.884, respectively, with a total scale Cronbach' s α coefficient of 0.932. Except for the SBD domain, which was slightly lower, all other domains demonstrated good reliability. From the facet perspective, except for four facets (cognition, positive emotion, social support, and social role) where Cronbach' s α was relatively low, all other facets exceeded 0.6. The split-half coefficients for PHD, MHD, SHD, SBD, and TIBD were 0.669, 0.859, 0.610, 0.494, and 0.795, respectively, with a total scale split-half reliability of 0.879 .

2.3 Content Validity

This study constructed the TIBD theoretical framework through reference to existing foreign IBD scales, expert consultation, and patient group discussions. More than ten experts in gastroenterology, statistics, and PRO scale research discussed and determined the preliminary specific module scale. Subsequently, 42 gastroenterology medical staff completed item importance assessments and 69 patients participated in qualitative interviews. Data analysis revealed an expert response rate of 100%, with full-score ratios for PRO-IBD item importance ratings ranging from 20.0% to 61.9%. Experts demonstrated high consensus on item operability, with coordination degrees ranging from 88.1% to 100.0%, indicating high credibility. Incorporating patient feedback, a second expert discussion was convened, resulting in the deletion of eight unreasonable items

and a final IBD TIBD scale containing 14 items. PROISCD-IBD (V1.0) thus demonstrates adequate content validity.

2.4 Construct Validity

2.4.1 Item-Domain Correlation Analysis Item-domain correlation analysis revealed that, except for MHD3 (life pleasure) and MHD8 (positive attitude toward disease), correlations between items and their respective domains were significantly stronger than correlations with domains outside their own . Correlations between PHD, MHD, SHD, and SBD domain scores and the common module score were all >0.6 (0.775, 0.893, 0.875, and 0.688, respectively).

2.4.2 Exploratory Factor Analysis The KMO value was 0.88, indicating strong correlations among factors, and Bartlett' s test was 1,514.531 (P<0.001), confirming the suitability of factor analysis. Factor analysis of the specific module extracted three principal components with a cumulative variance contribution rate of 58.05%. The three components were: the first principal component included seven items primarily reflecting gastrointestinal symptoms, with a variance contribution rate of 25.76%; the second principal component included four items reflecting treatment side effects and disease-related psychological effects, with a variance contribution rate of 19.34%; and the third principal component included two items reflecting extraintestinal symptoms, with a variance contribution rate of 12.96% .

2.4.3 Structural Equation Model The revised structural equation model yielded the following results: ^{2}=2,046.469 (P<0.001), df=797, ^{2}/df=2.568 (<3.000). The root-mean-square error of approximation (RMSEA) was 0.076, with a 90% CI of 0.072-0.080. The normed fit index (NFI), non-normed fit index (NNFI), comparative fit index (CFI), and incremental fit index (IFI) were 0.677, 0.774, 0.772, and 0.774, respectively. The standardized root-mean-square residual (SRMR) was 0.1031, with standardized factor loadings >0.5 for most items. These results indicate adequate model fit [Figure 1: see original paper].

2.5 Discriminant Validity (Clinical Validity)

IBD patients were divided by clinical stage into active (n=90) and remission (n=184) phases. Patients in remission scored significantly higher than those in the active phase across all domains, common modules, specific modules, and total scale scores (P<0.05) .

Discussion

Reliability is a key indicator of measurement tool stability, reflecting the impact of random error on measurement results. This study assessed reliability using internal consistency and split-half reliability methods. The most common

reliability coefficient is Cronbach' s α, which ranges from 0 to 1 and is generally considered acceptable when $ $0.7. Split-half reliability coefficients should generally be >0.7. In this study, PROISCD-IBD (V1.0) demonstrated a Cronbach' s α coefficient of 0.932 and split-half coefficient of 0.879. Except for the SBD domain, which had slightly lower reliability, all other domains exceeded 0.7. Similarly, except for the SBD domain, all other domains approached or exceeded 0.7 for split-half coefficients. The lower reliability of the SBD domain may be attributed to some patients not providing authentic evaluations of satisfaction and spiritual/belief content: although this is a self-reported scale, some patients may not have fully authentic evaluations of medical care and insurance satisfaction (items SBD1 and SBD2) due to being in a hospital setting. Additionally, regarding whether personal beliefs, religion, or worship provide strength to overcome difficulties and face life and death (items SBD4 and SBD5), some patients lacked religious beliefs or did not understand the item content, resulting in poor inter-item correlations and lower reliability.

Validity, also known as authenticity, was evaluated from three perspectives: content validity, construct validity, and discriminant validity. By drawing upon existing foreign IBD scales and incorporating input from medical staff, patients, and experts during item pool development, the scale reflects the true level of IBD patient-reported outcomes and aligns with Chinese cultural characteristics, thus demonstrating good content validity. Construct validity was assessed using correlation analysis, exploratory factor analysis, and structural equation modeling. Most items showed significantly stronger correlations with their own domains than with other domains. Factor analysis of the specific module extracted three common factors with a cumulative variance contribution rate of 58.05%. The items within these three factors reflected gastrointestinal symptoms, treatment side effects and disease-related psychological effects, and extraintestinal symptoms, respectively. Item IBD7 (fever) did not load on any common factor, while the PROISCD-IBD (V1.0) specific module is structured into four facets: gastrointestinal symptoms, extraintestinal symptoms, special psychological symptoms, and treatment side effects. Although the facet structure differs slightly from the exploratory factor analysis results, the item content classification is largely similar. In clinical practice, some IBD patients do experience fever symptoms, and based on expert medical opinion, the model results and item-domain relationships require further adjustment. Regarding structural equation modeling, ^{2}/df should generally be $ $3 and RMSEA $ $0.08. The revised model met these criteria, with NFI, NNFI, CFI, and IFI approaching acceptable standards and standardized factor loadings >0.5 for most items, indicating that the relationships between scale items and domains generally match the theoretical construct. Discriminant validity is a primary indicator of item and scale quality, often referred to as clinical validity in medical scale development. This study divided patients into active and remission groups by clinical stage; paired t-test results showed statistically significant differences across all domains, common modules, specific modules, and total scale scores, with higher scores in the remission group, as expected. PROISCD-IBD (V1.0) can effectively distinguish

patient-reported outcome levels across different disease phases, indicating good discriminant validity.

Conclusion

In summary, results from classical test theory and structural equation modeling indicate that PROISCD-IBD (V1.0) has good reliability and validity. However, some items and domains remain difficult for patients to understand or do not align with actual patient conditions, potentially failing to authentically reflect patient-reported outcome levels. Future research should focus on revising item wording and adjusting content structure to capture patient-reported outcomes with fewer items. Additionally, due to time and practical constraints, this study did not evaluate test-retest reliability, criterion validity, or responsiveness of the scale, which should be addressed in subsequent research.

Author Contributions: LUO Na was responsible for proposing the main research objectives, conceptualizing and designing the study, and drafting the manuscript. RUAN Yanqin and LEI Pingguang were responsible for data collection and management. WAN Chonghua was responsible for determining the research topic and providing guidance. WANG Keyan and SONG Ying were responsible for data verification and manuscript revision. CHEN Ying was responsible for quality control and review of the article, overall responsibility, and supervision. All authors confirmed the final manuscript.

Conflicts of Interest: None declared.

References

[1] Inflammatory Bowel Disease Professional Committee of China Medical Education Association. Chinese expert recommendations on biological agents for inflammatory bowel disease (trial version) [J]. Chinese Journal of Digestive Diseases and Imaging (Electronic Edition), 2021, 11(6): 10.12104/j.issn.1674-4748.2022.22.009.

[2] YU Bing, LI Xin. Research progress on the pathogenesis of inflammatory bowel disease [J]. Journal of Snake, 2020, 32(1): 114-116. DOI: 10.3969/j.issn.1001-5639.2020.01.038.

[3] NG S C, SHI H Y, HAMIDI N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies [J]. Lancet, 2017, 390(10114): 2769-2778. DOI: 10.1016/S0140-6736(17)32448-0.

[4] HE Qiong, LI Jiandong. Research progress on the epidemiology of inflammatory bowel disease [J]. Journal of Practical Medicine, 2019, 35(18): 2962-2966. DOI: 10.3969/j.issn.1006-5725.2019.18.001.

[5] KAPLAN G G. The global burden of IBD: from 2015 to 2025 [J]. Nat Rev Gastroenterol Hepatol, 2015, 12(12): 720-727. DOI: 10.1038/nrgastro.2015.150.

[6] LI Hui, LI Mingsong. Challenges and opportunities for inflammatory bowel disease in China [J]. Modern Digestion & Intervention, 2019, 24(6): 569-572, 582. DOI: 10.3969/j.issn.1672-2159.2019.06.001.

[7] LEE J W. Depressive symptoms and quality of life in the patients of inflammatory bowel disease [J]. Gut Liver, 2017, 11(4): 449-450. DOI: 10.5009/gnl17227.

[8] Department of Health and Human Services FDA Center for Drug Evaluation and Research U S, Department of Health and Human Services FDA Center for Biologics Evaluation and Research U S, Department of Health and Human Services FDA Center for Devices and Radiological Health U S. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance [J]. Health Qual Life Outcomes, 2006, 4: 79. DOI: 10.1186/1477-7525-4-79.

[9] LU Yi, GE Changchang, SHEN Hong, et al. Research progress on patient-reported outcome scales for inflammatory bowel disease [J]. Gastroenterology, 2022, 27(2): 119-123. DOI: 10.3969/j.issn.1008-7125.2022.02.012.

[10] ZHAI Yunxiao, HUANG Chan. Research progress on self-reported outcome assessment tools for inflammatory bowel disease patients [J]. Chinese General Practice Nursing, 2022, 20(22): 3061-3064. DOI: 10.12104/j.issn.1674-4748.2022.22.009.

[11] ZHOU Yunxian. Clinical application study of the Chinese version of the Inflammatory Bowel Disease Quality of Life Questionnaire [D]. Hangzhou: Zhejiang University, 2006.

[12] ZHU Danling, ZHOU Yunxian, XU Hongwen. Translation and reliability and validity testing of the Brief Health Scale for inflammatory bowel disease [J]. Chinese Nursing Management, 2018, 18(12): 1630-1634. DOI: 10.3969/j.issn.1672-1756.2018.12.010.

[13] WAN Chonghua, GAO Li, LI Xiaomei, et al. Development methodology for the generic module of the Quality of Life Instrument for Chronic Diseases (QLICD) system (Part 1): item selection and formation of the generic module [J]. Chinese Mental Health Journal, 2005, 19(11): 723-726. DOI: 10.3321/j.issn:1000-6729.2005.11.002.

[14] ZHANG Chuanmeng, HUANG Zhiwen, WAN Chonghua, et al. Development and responsiveness evaluation of the coronary heart disease scale in the patient-reported outcome measurement system for chronic diseases [J]. Chinese Journal of Modern Nursing, 2018, 24(17): 1991-1995. DOI: 10.3760/cma.j.issn.1674-2907.2018.17.002.

[15] WAN Chonghua, WU Xiaoyu, LIU Yuxi, et al. Research and application status of the second version of the Quality of Life Instrument for Chronic Diseases (QLICD [V2.0]) [J]. Journal of Guangdong Medical University, 2022, 40(3): 243-249.

[16] WU Kaichun, LIANG Jie, RAN Zhihua, et al. Consensus on diagnosis and treatment of inflammatory bowel disease (2018 · Beijing) [J]. Chinese Journal of Practical Internal Medicine, 2018, 38(9): 796-813. DOI: 10.19538/j.nk2018090106.

[17] WU Yongze, WANG Wenjuan. Effects of different response categories on Likert scale characteristics [J]. Chinese Journal of Prevention and Control of Chronic Diseases, 2010, 18(2): 215-217. DOI: 10.16386/j.cjpccd.issn.1004-6194.2010.02.003.

[18] VAN DER STEEG A F W, DE VRIES J, ROUKEMA J A. Quality of life and health status in breast carcinoma [J]. Eur J Surg Oncol EJSO, 2004, 30(10): 1051-1057. DOI: 10.1016/j.ejso.2004.07.003.

[19] OU Jinqing, LEI Pingguang, REN Dandan, et al. Evaluation of measurement properties of the quality of life scale for patients with fatty liver disease [J]. Soft Science of Health, 2020, 34(8): 36-40. DOI: 10.3969/j.issn.1003-2800.2020.08.008.

[20] LIU Baoyan. Measurement of patient-reported outcomes: principles, methods, and applications [M]. Beijing: People' s Medical Publishing House, 2011: 168.

(Received: February 25, 2023; Revised: June 29, 2023)

Note: Figure translations are in progress. See original paper for figures.