Development and Psychometric Evaluation of the QLICD-NS (V2.0) Quality of Life Instrument for Nephrotic Syndrome Using Classical Test Theory: A Post-print

Authors: Zhang Xingshan, Lin Yue, Wan Chonghua, Song Aifeng, Pan Haiyan

Date: 2022-10-28T00:00:00+00:00

Abstract

Objective: To develop a specific module for a quality of life measurement scale for patients with nephrotic syndrome, combine it with the second edition generic module of the Quality of Life Instrument for Chronic Diseases series to constitute a Quality of Life Scale for Nephrotic Syndrome Patients, and evaluate its psychometric properties. Methods: A total of 203 nephrotic syndrome patients attending the Department of Nephrology at the Affiliated Hospital of Guangdong Medical University in 2021 were selected for the assessment. Face-to-face interviews and questionnaire surveys were conducted. Based on the questionnaire results, item selection was performed using the variation method, correlation coefficient method, Cronbach' s alpha method, factor analysis method, and importance ratings by doctors and patients, combined with expert discussion. Finally, evaluation methods based on classical test theory were applied. Results: The specific module of the Quality of Life Scale for Nephrotic Syndrome Patients contains 15 items. The Cronbach' s alpha coefficients and split-half reliability coefficients for all domains and the total scale were greater than 0.7. The item content covers three aspects: clinical symptoms, side effects, and psychological impact. Using SF-36 as the criterion standard, the correlation between each domain and its similar domains was higher than with other domains. The standardized response means for all domains and the total scale were greater than 0.80. Conclusion: The specific module of the Quality of Life Scale for Nephrotic Syndrome Patients developed in this study includes 15 items, and the scale demonstrates good reliability, validity, and responsiveness.

Full Text

Development and Psychometric Evaluation of the Nephrotic Syndrome Quality of Life Instrument QLICD-NS (V2.0) Using Classical Test Theory

Authors: Zhang Xingshan[1], Lin Yue[1], Wan Chonghua[2], Song Aifeng[1], Pan Haiyan[1]*

[1]Department of Epidemiology and Health Statistics, Guangdong Medical University, Dongguan 523808, China

[2]School of Humanities and Management, Guangdong Medical University, Dongguan 523808, China

*Corresponding author: Pan Haiyan, Professor, Master' s Supervisor; E-mail: panhaiyan@gdmu.edu.cn

Abstract

Objective: To develop a disease-specific module for measuring quality of life (QOL) in patients with nephrotic syndrome, integrate it with the second edition common module of the Quality of Life Instruments for Chronic Diseases (QLICD) series, and evaluate its psychometric properties. **Methods:** A total of 203 nephrotic syndrome patients from the Department of Nephrology at the Affiliated Hospital of Guangdong Medical University were enrolled in 2021. On-site interviews and questionnaire surveys were conducted. Item selection was performed using variation degree method, correlation coefficient method, Cronbach' s alpha method, factor analysis, and importance ratings from both physicians and patients, supplemented by expert discussions. Classical Test Theory (CTT) was applied for final psychometric evaluation. **Results:** The specific module contained 15 items covering clinical symptoms, side effects, and psychological impact. Cronbach' s alpha and split-half reliability for all domains and the total scale exceeded 0.7. Using SF-36 as the criterion standard, correlations between corresponding domains were higher than those with non-corresponding domains. The standardized response mean for all domains and the total scale was greater than 0.80. **Conclusion:** The nephrotic syndrome-specific QOL module developed in this study comprises 15 items and demonstrates satisfactory reliability, validity, and responsiveness.

Keywords: Quality of life; Nephrotic syndrome; Classical Test Theory

Introduction

Global disease burden studies estimated that approximately 1.4 million deaths worldwide were attributed to kidney disease in 2019, representing a 20% increase from 2010 [1]. Nephrotic syndrome (NS) has emerged as a common clinical diagnosis among chronic kidney diseases, accounting for approximately 40% of renal biopsy cases [2-3]. NS is characterized by treatment difficulties, high recurrence rates, and significant treatment-related adverse effects. The clinical

management of NS often overlooks patients' subjective experiences, leading to detrimental impacts on psychological well-being, physical health, personality, social activities, and work/study performance [4-7]. Consequently, attention to and improvement of patients' quality of life represents a critical need.

Currently, generic QOL scales and some kidney disease-specific instruments are available for evaluating NS patients, yet no NS-specific scale has been identified in domestic or international literature [8-10]. While generic scales are broadly applicable to healthy and diseased populations, they fail to capture disease-specific characteristics and are less sensitive to changes in patients' QOL. Additionally, many existing scales were developed in Western countries, and cultural, economic, and educational differences may affect their validity in Chinese populations, as patients may avoid certain questions, compromising result accuracy. Therefore, developing a culturally appropriate QOL scale for Chinese NS patients is warranted.

The first edition of the Quality of Life Instruments for Chronic Diseases (QLICD V1.0) was developed by Professor Wan Chonghua and colleagues, comprising a common module and disease-specific modules for eight chronic conditions [11]. The common module could be used across multiple chronic diseases, providing a foundation for subsequent QOL research. However, as research progressed, limitations of QLICD V1.0 became apparent. First, with only eight chronic diseases included, many conditions lacked specific modules, requiring analysis using only the common module. Second, advances in medical technology have altered disease-specific manifestations, necessitating revision of specific module items. Consequently, the research team initiated development of the second edition. To date, the common module QLICD-GM (V2.0) and 26 disease-specific modules have been completed. The common module encompasses three domains, nine facets, and 28 items, while six disease-specific modules remain under development. This second edition series adopts the structure of combining common and specific modules to create culturally appropriate QOL instruments for Chinese chronic disease patients [12].

1.1 Research Objectives

This study aimed to develop a nephrotic syndrome-specific QOL module, integrate it with the second edition common module to form the test version QLICD-NS (V2.0), and evaluate its psychometric properties using Classical Test Theory. This work provides a measurement tool for assessing QOL in NS patients and further enriches the chronic disease QOL instrument series.

1.2 Research Design

The research team conducted literature searches from 2017 to 2022 across CNKI, Wanfang, Chinese Biomedical Literature Database, PubMed, and Web of Science. Drawing upon commonly used generic scales, an initial item pool reflecting NS patients' QOL was established. Two research groups performed preliminary

item screening, pre-survey analysis, and subsequent item refinement to construct the specific module.

Convenience sampling was employed to recruit 20 clinicians from urology, respiratory, gastroenterology, and dermatology departments, along with 30 patients for on-site interviews. For patients, investigators explained the survey purpose and completion methods, obtained informed consent, and administered the questionnaire (for minor patients, guardians assisted after providing consent). For clinicians, investigators discussed item modifications after questionnaire completion.

Statistical analysis of pre-survey data was conducted, followed by another round of expert and scholar discussions regarding item selection. Survey instruments included: (a) a basic demographic information questionnaire; (b) the Nephrotic Syndrome Quality of Life Scale QLICD-NS (V2.0); and (c) the Chinese version of the SF-36 Health Survey.

1.3 Study Subjects

NS patients from the Department of Nephrology at the Affiliated Hospital of Guangdong Medical University between March 2021 and November 2021 were selected as study participants. Patients completed the first measurement on admission day, a subset completed a second measurement during hospitalization, and a third measurement was completed on discharge day. **Inclusion criteria:** (1) confirmed NS diagnosis according to diagnostic criteria; (2) primary school education or above, with reading and comprehension ability to complete the questionnaire independently or with assistance (for minor patients, guardians could assist after providing consent); (3) non-critical condition, without other severe comorbidities, and provided informed consent. **Exclusion criteria:** (1) patients with psychiatric disorders or confused consciousness; (2) illiterate patients; (3) patients refusing to cooperate. This study complied with informed consent requirements and obtained ethical approval from the Affiliated Hospital of Guangdong Medical University Ethics Committee (Approval No.: PJ2014040).

1.4 Statistical Methods

Negative items (where higher scores indicate worse conditions) were first reverse-coded. Six methods were applied for item selection, retaining items meeting four or more criteria: (1) Variation degree method: calculating standard deviation for each item (with consistent basic units, variation reflected by standard deviation), deleting items with SD < 1.0; (2) Correlation coefficient method: calculating correlation between each item and its domain score, deleting items with $r < 0.4$; (3) Cronbach' s alpha method: calculating domain α, then α_1 after deleting the item, deleting if $\alpha < \alpha_1$; (4) Factor analysis: extracting common factors with eigenvalues > 1, deleting items with loading coefficients < 0.6 after varimax rotation; (5) Physician importance rating: deleting items with mean score < 80;

(6) Patient importance rating: deleting items with mean score < 70 [13-14].

A 5-point scoring system was applied to each item. Positive items were scored directly as 1, 2, 3, 4, 5, while negative items were reverse-scored as 5, 4, 3, 2, 1. Facet scores were calculated by summing relevant items, and total scale score by summing all items. To enable comparison across facets, raw scores were transformed to standard scores using the range method: S = (X - Min) × 100 / R, where S is the standard score, X the raw score, Min the minimum possible score for the facet or total scale, and R the range (maximum minus minimum score).

SPSS 26.0 was used for statistical analysis. Measurement data were expressed as mean ± standard deviation, and count data as percentages. For factor analysis, common factors were extracted based on eigenvalues > 1 [15]. Correlation analysis considered $P < 0.05$ statistically significant.

1.5 Psychometric Evaluation of the Scale

Classical Test Theory (CTT) provides a macro-level evaluation of measurement tools with low sample size requirements and straightforward implementation [16]. CTT comprehensively assesses measurement properties including reliability, validity, and responsiveness. Reliability encompasses test-retest reliability, split-half reliability, and internal consistency. Validity includes content validity, construct validity, and criterion validity. Reliability reflects scale precision (consistency and stability), validity reflects scale accuracy (utility and correctness), and responsiveness reflects sensitivity to change through pre-post treatment measurements [17], as shown in Table 1 .

Table 1 Contents and methods of CTT assessment

Evaluation Aspect	Evaluation Indicator	Method
Reliability	Test-retest reliability	Compare first and second measurement domain and total scores
	Split-half reliability	Split items into odd/even groups, calculate correlation (Pearson)

Evaluation Aspect	Evaluation Indicator	Method
	Internal consistency	Calculate Cronbach's alpha for each domain
Validity	Content validity	Based on literature and expert analysis
	Construct validity	Calculate item-domain correlations, conduct exploratory factor analysis
	Criterion validity	Use SF-36 as criterion, apply Pearson correlation analysis
Responsiveness	Standardized response mean	Use paired t-test for pre-post treatment scores, calculate SRM

Results

2.1 Study Subjects

A total of 203 NS patients were surveyed, with three repeated measurements conducted. The second survey yielded 162 valid scales, and the third survey yielded 201 valid scales. The sample included 132 males and 71 females, with a mean age of 41.42 ± 19.42 years. All participants were Han Chinese. The most common occupation category was "other" (primarily students and retired elderly). Primary and junior high school education accounted for 72.9% of the sample. Demographic details are presented in Table 2 .

Table 2 Demographic information

Characteristic	n (%)	Characteristic	n (%)
Gender		**Education**	
Male	132 (65.02)	Primary school	78 (38.42)
Female	71 (34.98)	Junior high school	70 (34.48)
Age (years)		High school/technical school	34 (16.75)
<18	29 (14.29)	College degree or above	14 (6.90)
18-44	100 (49.26)	**Occupation**	
45-59	24 (11.82)	Employed	48 (23.65)
$ $60	30 (14.78)	Unemployed	19 (9.36)
Insurance		Student	1 (0.49)
Social medical insurance	138 (68.00)	Retired	2 (0.99)
Self-pay	65 (32.00)	Other	124 (61.08)

2.2 Item Screening Results for the Specific Module

The research team proposed an initial item pool covering urinary, respiratory, gastrointestinal, and dermatological systems, as well as medication side effects. After core group review, discussion, modification, and consolidation, a final item pool of 10 items was formed. Screening results are shown in Table 3 .

Table 3 Item screening results for specific modules

Item	Physician Rating	Patient Rating	SD	Item-Domain Correlation	Cronbach' s α if Deleted
TNS1: Infection symptoms?	89.40*	90.75*	1.402*	0.380*	0.92
TNS2: Edema?	94.90*	92.83*	1.299*	0.413*	0.91
TNS3: Walking difficulty/falls?	81.50*	81.37*	1.642*	0.430*	0.90
TNS4: Decreased urine output?	94.45*	86.50*	1.217*	0.340*	0.91

Item	Physician Rating	Patient Rating	SD	Item-Domain Correlation	Cronbach' s α if Deleted
TNS5: Foamy urine?	83.65*	73.87*	0.781*	0.484*	0.92
TNS6: Dizziness/headache?	73.90*	76.97*	0.821*	0.468*	0.91
TNS7: Back/lower limb pain?	90.15*	73.30*	1.245*	0.460*	0.91
TNS8: Skin itching?	83.65*	70.47*	0.657*	0.689*	0.90
TNS9: Hematuria?	73.90*	73.87*	1.456*	0.484*	0.91
TNS10 Shortness of breath/chest pain?	83.50*	72.00*	1.006*	0.488*	0.92

*Note: Items meeting retention criteria

2.3 Formation of the Specific Module Test Version

Based on recommendations from clinical staff and statistical experts, the following modifications were made:

1. Item "NS1: Do you have infection symptoms (respiratory, skin, urinary, intestinal)?" was subdivided into more specific items for better patient comprehension: "Do you have fever, cough, sputum, or sore throat?" , "Do you have abdominal pain, diarrhea, or bloating?" , "Do you have skin damage or rash?" , and "Do you have frequent, urgent, or painful urination?"
2. Item "NS3: Do you have difficulty walking or fall easily?" was deleted, as most hospitalized patients could walk and clinicians considered this symptom uncommon.
3. Item "NS8: Do you have skin itching?" was deleted, as clinicians noted few patients experienced itching; some had rashes without itching.

4. Item "NS10: Do you have shortness of breath or chest pain?" was revised to "Do you have chest tightness, chest pain, or shortness of breath?" for better specificity.
5. Given that NS treatment primarily involves glucocorticoids or immunosuppressants with various side effects (gastrointestinal dysfunction, endocrine disorders, blood pressure elevation), three items were added: "Do you experience hair loss, acne, hirsutism, or stretch marks after medication?" , "Do you have stomach discomfort?" , and "Do you have palpitations?"
6. Considering the chronic and relapsing nature of NS, a psychological item was added: "Do you worry about disease recurrence?"

The final specific module contained 15 items, which were reorganized into facets and renumbered. The final items are presented in Table 4 .

Table 4 Specific module items of QLICD-NS (V2.0)

Item	Content
TNS1	Do you have fever, cough, sputum, or sore throat?
TNS2	Do you have abdominal pain, diarrhea, or bloating?
TNS3	Do you have skin damage or rash?
TNS4	Do you have edema?
TNS5	Do you have frequent, urgent, or painful urination?
TNS6	Do you have decreased urine output?
TNS7	Do you have foamy urine?
TNS8	Do you have hematuria?
TNS9	Do you have dizziness or headache?
TNS10	Do you have chest tightness, chest pain, or shortness of breath?
TNS11	Do you have palpitations?
TNS12	Do you have back or lower limb pain?
TNS13	Do you have stomach discomfort?
TNS14	Do you experience hair loss, acne, hirsutism, or stretch marks after medication?
TNS15	Do you worry about disease recurrence?

2.4.1 Scale Reliability Evaluation

Test-retest reliability: Paired t-tests comparing first and second measurements showed no statistically significant differences in domain or total scores ($P > 0.05$). Correlation coefficients between first and second measurements exceeded 0.70 for all domains ($P < 0.001$), with intraclass correlation coefficients (ICC) all greater than 0.75, indicating good scale stability (Table 5).

Table 5 Test-retest reliability assessment

Domain	First Measurement (Mean ± SD)	Second Measurement (Mean ± SD)	Paired t-test	P-value	r	P-value	ICC (95% CI)
Physical Function (PHD)	62.96 ± 13.70	63.75 ± 14.10	1.917	0.057	0.940	<0.001	0.939 (0.918, 0.955)
Psychological Function (PSD)	70.16 ± 13.64	69.81 ± 12.25	0.085	0.932	0.795	<0.001	0.791 (0.725, 0.842)
Social Function (SOD)	68.37 ± 16.48	69.00 ± 12.72	0.771	0.524	0.762	<0.001	0.757 (0.682, 0.816)
Common Module (CGD)	67.27 ± 11.39	67.63 ± 10.54	0.121	0.835	0.901	<0.001	0.899 (0.864, 0.925)
Specific Module (SPD)	59.08 ± 13.67	58.18 ± 13.69	0.519	0.604	0.938	<0.001	0.938 (0.916, 0.954)
Total Scale (TOT)	64.43 ± 10.87	64.33 ± 10.41	1.614	0.108	0.940	<0.001	0.940 (0.919, 0.956)

Internal consistency: Cronbach' s alpha and split-half reliability were 0.878 and 0.821 for the common module, 0.774 and 0.708 for the specific module, and 0.884 and 0.782 for the total scale, respectively. All domain coefficients exceeded 0.700, indicating good internal consistency (Table 6).

Table 6 Internal consistency assessment

Domain	Number of Items	Cronbach' s α	Split-half Reliability
Physical Function (PHD)	9	0.823	0.801
Psychological Function (PSD)	8	0.851	0.792

Domain	Number of Items	Cronbach' s α	Split-half Reliability
Social Function (SOD)	11	0.762	0.741
Common Module (CGD)	28	0.878	0.821
Specific Module (SPD)	15	0.774	0.708
Total Scale (TOT)	43	0.884	0.782

2.4.2 Scale Validity Evaluation

Content validity: The QLICD-NS (V2.0) common module is the second edition of the chronic disease QOL scale, which has undergone specialized evaluation demonstrating good content validity. The entire scale development process involved clinical physicians, epidemiologists, and psychologists. The specific module covers clinical symptoms, side effects, and psychological impact, comprehensively reflecting NS patients' QOL. Therefore, QLICD-NS (V2.0) can be considered to have good content validity.

Construct validity: Item-domain correlations were calculated to verify that each item correlated most strongly with its intended domain. Results showed low correlation with GPH3 in the physical function domain; GPS2 ("Has your disease caused memory decline?") correlated weakly with psychological function (r = 0.382); GSO6 ("Has illness and treatment caused family financial difficulties?") correlated weakly with social function (r = 0.376); and TNS3 ("Do you have skin damage or rash?") and TNS8 ("Do you have hematuria?") showed correlations of 0.240 and 0.184, respectively. All other item-domain correlations exceeded 0.40, indicating satisfactory construct validity (Table 7).

Table 7 Pearson correlation analysis of QLICD-NS (V2.0) items with domains

Item	Physical Function	Psychological Function	Social Function	Specific Module
GPH1	0.624**	0.546**	0.180*	0.601**
GPH2	0.522**	0.616**	0.554**	0.623**
GPH3	0.532**	0.403**	0.161*	0.423**
GPS1	0.292**	0.216**	0.284**	0.140*
GPS2	0.323**	0.382**	0.152*	0.255**
GSO1	0.351**	0.280**	0.416**	0.694**
GSO2	0.380**	0.550**	0.576**	0.166*

Item	Physical Function	Psychological Function	Social Function	Specific Module
GSO3	0.146*	0.222**	0.210**	0.205**
GSO4	0.530**	0.219**	0.433**	0.178*
GSO5	0.318**	0.323**	0.409**	0.289**
GSO6	0.221**	0.363**	0.376**	0.207**
TNS1	0.380**	0.369**	0.217**	0.267**
TNS2	0.280**	0.210**	0.302**	0.229**
TNS3	0.345**	0.286**	0.279**	0.240**
TNS4	0.155*	0.251**	0.204**	0.326**
TNS5	0.203**	0.399**	0.331**	0.319**
TNS6	0.218**	0.344**	0.289**	0.438**
TNS7	0.478**	0.523**	0.606**	0.704**
TNS8	0.689**	0.779**	0.728**	0.184**
TNS9	0.578**	0.164*	0.362**	0.406**
TNS10	0.350**	0.163*	0.331**	0.326**
TNS11	0.324**	0.379**	0.294**	0.317**
TNS12	0.567**	0.189**	0.366**	0.304**
TNS13	0.506**	0.330**	0.619**	0.498**
TNS14	0.466**	0.670**	0.544**	0.376**
TNS15	0.530**	0.486**	0.179*	0.178*

Note: $P < 0.05$, $^{**}P < 0.01^*$

Factor analysis: Exploratory factor analysis of the specific module yielded KMO = 0.677 (> 0.600) [18] and Bartlett's sphericity test $^2 = 605.474$, $P < 0.01$, indicating data suitability for factor analysis. Principal component analysis extracted common factors with eigenvalues > 1, and varimax rotation yielded three factors explaining 51.491% of variance. Factor 1 (items TNS1-TNS8) reflected disease symptoms (25.167% variance); Factor 2 (items TNS9-TNS14) reflected treatment side effects (18.565% variance); Factor 3 (item TNS15) reflected psychological impact (7.759% variance). This three-factor structure aligned with the intended facet structure, supporting good construct validity (Table 8).

Table 8 Exploratory factor analysis of QLICD-NS (V2.0) specific module

Item	Factor 1	Factor 2	Factor 3
Fever/cough/sputum/sore throat	0.406	0.301	0.102
Abdominal pain/diarrhea/bloating	0.521	0.289	0.201
Skin damage/rash	0.612	0.198	0.156
Edema	0.698	0.201	0.089
Frequent/urgent/painful urination	0.589	0.234	0.167
Decreased urine output	0.634	0.189	0.145
Foamy urine	0.701	0.156	0.098

Item	Factor 1	Factor 2	Factor 3
Hematuria	0.578	0.234	0.201
Dizziness/headache	0.234	0.456	0.189
Chest tightness/pain/shortness of breath	0.189	0.523	0.234
Palpitations	0.156	0.489	0.267
Back/lower limb pain	0.201	0.567	0.189
Stomach discomfort	0.145	0.634	0.156
Medication side effects	0.123	0.701	0.234
Worry about recurrence	0.089	0.156	0.823

Criterion validity: Using the Chinese SF-36 as the criterion, correlations between corresponding domains were examined. Physical functioning correlated > 0.60 with the physical domain; role-physical showed low correlations with QLICD-NS (V2.0) domains; mental health and role-emotional correlated > 0.50 with psychological function; general health and health transition correlated 0.619 and 0.633 with total scale score, respectively. Overall, SF-36 demonstrated acceptable criterion validity for QLICD-NS (V2.0) (Table 9).

Table 9 Correlation analysis between QLICD-NS (V2.0) and SF-36 domain scores

SF-36 Domain	Physical Function	Psychological Function	Social Function	Common Module	Specific Module	Total Score
Physical Functioning	0.632**	0.561**	0.386**	0.343**	0.240**	0.457**
Role-Physical	0.466**	0.636**	0.175**	0.714**	0.619**	0.470**
Bodily Pain	0.539**	0.481**	0.502**	0.633**	0.503**	0.539**
General Health	0.423**	0.270**	0.314**	0.147*	0.318**	0.507**
Vitality	0.607**	0.194**	0.698**	0.634**	0.514**	0.586**
Social Functioning	0.463**	0.530**	0.600**	0.501**	0.193**	0.603**
Role-Emotional	0.569**	0.465**	0.633**	0.540**	0.567**	0.554**
Mental Health	0.316**	0.483**	0.359**	0.358**	0.379**	0.283**
Health Transition	0.215**	0.391**	0.283**	0.215**	0.391**	0.283**

2.4.3 Scale Responsiveness Evaluation

Paired t-tests analyzed pre-post treatment scores for QLICD-NS (V2.0) domains/facets, with standardized response means (SRM) calculated as (post-score - pre-score) / standard deviation of difference. Results showed no statistical difference in psychological function ($P > 0.05$), though the cognition and will/personality facets were significant ($P < 0.05$). All other domains/facets showed significant differences ($P < 0.05$). Except for cognition and social support facets, all SRM values exceeded 0.8, indicating good responsiveness (Table 10).

Table 10 Responsiveness assessment of QLICD-NS (V2.0)

Domain/Facet	Pre-treatment (Mean ± SD)	Post-treatment (Mean ± SD)	t-value	P-value	SRM
Physical Function	62.76 ± 13.48	69.35 ± 9.66	8.83	<0.001	1.22
Basic physiological function	60.11 ± 15.89	65.08 ± 11.61	11.13	<0.001	1.21
Cognition	74.25 ± 19.21	81.43 ± 18.70	11.24	<0.001	1.02
Will and personality	50.81 ± 20.08	59.76 ± 11.12	15.67	<0.001	1.00
Psychological Function	70.10 ± 13.61	71.42 ± 11.00	8.19	0.020	0.15
Emotion	77.24 ± 16.43	81.78 ± 28.53	12.13	<0.001	1.02
Social Function	67.77 ± 14.28	67.54 ± 10.77	9.08	0.771	0.11
Social support	71.14 ± 19.74	74.63 ± 16.20	13.31	<0.001	0.73
Common Module	68.27 ± 16.53	72.17 ± 10.02	11.74	<0.001	1.15
Specific Module	69.65 ± 16.73	73.05 ± 12.64	14.84	<0.001	1.17
Total Scale	80.18 ± 27.83	80.22 ± 12.71	13.31	<0.001	1.73
Symptom facet	48.32 ± 22.45	58.77 ± 16.59	18.66	<0.001	1.48
Side effect facet	67.22 ± 11.36	70.97 ± 8.05	7.11	<0.001	1.17
Psychological impact facet	64.35 ± 10.85	71.94 ± 7.83	15.82	<0.001	1.48

Discussion

Classical Test Theory employs relatively simple mathematical models that are easy to understand, calculate, and widely applicable [19]. This study used CTT to analyze the scale' s reliability, validity, and responsiveness.

Reliability includes test-retest reliability, split-half reliability, and internal consistency. Test-retest reliability showed no significant differences between first and second measurements ($P > 0.05$), with all domain correlations exceeding 0.70 ($P < 0.05$), indicating satisfactory test-retest reliability. Internal consistency, measuring homogeneity among domain items, was assessed using Cronbach' s α. Values between 0.70 and 0.95 indicate good internal consistency [20]. QLICD-NS (V2.0) achieved $\alpha > 0.70$ for both common and specific modules, with total scale $\alpha = 0.884$, demonstrating excellent internal consistency.

Validity comprises content validity, construct validity, and criterion validity. Content validity, a prerequisite for scale development, reflects coverage of intended content and is typically assessed through expert review [21]. QLICD-NS (V2.0) includes common and specific modules covering physical, psychological, and social function domains, with broad coverage. Developed through rigorous procedures with clear, comprehensible items, the scale comprehensively reflects patients' QOL, supporting good content validity.

Construct validity analysis revealed low correlation between physical function domain and GPH3 (sexual function), likely due to cultural differences in discussing sexual topics, potentially causing underreporting. GPS2 correlated weakly with psychological function ($r = 0.382$), possibly because 38.92% of patients were aged >55 years, and younger patients have better memory than elderly patients. GSO6 correlated weakly with social function ($r = 0.376$), as only 14.29% of patients were self-pay, with most having social medical insurance that substantially reduced financial burden. TNS3 and TNS8 showed low correlations (0.240 and 0.184) due to low prevalence of skin symptoms and hematuria, though these are genuine disease manifestations requiring larger samples for definitive assessment. All other item-domain correlations exceeded 0.40, indicating acceptable construct validity. Exploratory factor analysis yielded three factors matching the intended structure, confirming good construct validity.

Using SF-36 as the criterion, role-physical showed low correlations with QLICD-NS (V2.0) domains, likely because hospitalized patients had restricted work/activity, resulting in uniform responses. Overall, SF-36 demonstrated acceptable criterion validity.

Responsiveness refers to a scale' s ability to detect clinically meaningful changes over time [22]. Comparing first and third (discharge day) measurements, cognition and will/personality facets showed significant differences ($P < 0.05$), though the emotion facet in psychological function did not, possibly due to reduced daily activities during hospitalization resulting in consistently low emotion scores. All other domains/facets showed significant differences ($P < 0.05$). Except for cogni-

tion and social support facets, all SRM values exceeded 0.8. Generally, SRM > 0.5 indicates moderate responsiveness and SRM > 0.8 indicates high responsiveness [23]. Combined results support good responsiveness of QLICD-NS (V2.0).

The NS-specific module was developed using CTT, generalizability theory, and item response theory, comprising 15 items integrated with QLICD-GM (V2.0) to form QLICD-NS (V2.0). CTT evaluation demonstrated satisfactory reliability, validity, and responsiveness. Study limitations include: (1) all patients were from a single hospital, limiting generalizability; (2) all participants were Han Chinese. Future work should expand the sample to multiple hospitals and communities, and estimate the minimal clinically important difference (MCID) after scale revision.

Author Contributions: Author 1 collected and analyzed data and drafted the manuscript; Author 2 verified case data and revised the manuscript; Author 3 determined the research topic and provided guidance; Author 4 checked data and revised the manuscript; Author 5 supervised data collection and controlled manuscript quality. All authors approved the final version.

Conflict of Interest: The authors declare no conflicts of interest.

References

[1] Basch E, Deal AM, Dueck AC, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment[J]. JAMA, 2017, 318(2). DOI:10.1001/jama.2017.7156.

[2] Haraldstad K, Wahl A, Andenæs R, et al. A systematic review of quality of life research in medicine and health sciences[J]. Qual Life Res, 2019, 28(10): 2641-2650. DOI: 10.1007/s11136-019-02214-9.

[3] ZIDAROV D, VISCA R, AHMED S. Type of clinical outcomes used by healthcare professionals to evaluate health-related quality of life domains to inform clinical decision making for chronic pain management[J]. Quality of Life Research, 2019, 28(10): 2761-2771. DOI: 10.1007/s11136-019-02215-8.

[4] LI Q W. Effect of targeted nursing based on goal management theory on psychological state, compliance behavior and prognosis of children with nephrotic syndrome[J]. Heilongjiang Medicine, 2020, 44(04): 567-568. DOI: 10.3969/j.issn.1004-5775.2020.04.056.

[5] ZHAO X P. Influence of empathy nursing on negative emotion of patients with nephrotic syndrome[J]. Rural Health in China, 2020, 12(06): 66. DOI: 10.3969/j.issn.1674-361X.2020.06.066.

[6] WANG Y J, WANG Y Y. Preventive effect of predictive nursing intervention on complications of patients with nephrotic syndrome[J]. World Latest Medicine Information, 2019, 19(A0): 269. DOI: 10.19613/j.cnki.1671-3141.2019.100.145.

[7] Troost JP, Gipson DS, Carlozzi NE, et al. Using PROMIS to create clinically meaningful profiles of nephrotic syndrome patients[J]. Health Psychology

Research, 2019, 38(5): 410-421. DOI:10.1037/hea0000679.

[8] Selewski DT, Troost JP, Massengill SF, et al. The impact of disease duration on quality of life in children with nephrotic syndrome: a Midwest Pediatric Nephrology Consortium study[J]. Pediatric Nephrology, 2015, 30(9): 1467-1476. DOI:10.1007/s00467-015-3074-x.

[9] WANG C L, YANG X J, ZHAO Y. Effect of evidence-based nursing intervention on treatment compliance and quality of life of patients with nephrotic syndrome[J]. Journal of Community Medicine, 2018, 16(05): 76-77.

[10] Yoshiko Shutto, Hideaki Yamabe, Michiko Shimada, et al. Quality of Life in Patients with Minimal Change Nephrotic Syndrome[J]. The Scientific World Journal, 2013, 124315. DOI:10.1155/2013/124315.

[11] PAN H Y, YANG Z, WAN C H, et al. Secondary evaluation of common modules of the chronic disease quality of life scale[J]. Chinese General Practice, 2012, 15(10): 1101-1103, 1107. DOI: 10.3969/j.issn.1007-9572.2012.10.011.

[12] Thomas ML. Advances in applications of item response theory to clinical assessment[J]. Psychol Assess. 2019, 31(12): 1442-1455. DOI: 10.1037/pas0000597.

[13] WAN C H, GAO L, LI X M, et al. Development of common modules of the chronic disease patients' quality of life scale system (I)——Item screening and formation of common modules[J]. Chinese Journal of Mental Health, 2005(11): 8-11.

[14] WAN C H. Quality of Life Measurement and Evaluation Methods[M]. Kunming: Yunnan University Press, 1999.

[15] LIU Z D, ZHANG Q Q, TIAN G L. Summary of factor analysis comprehensive evaluation[J]. Statistics and Decision-making, 2019, 35(19): 68-73. DOI:10.13546/j.cnki.tjyjc.2019.19.015.

[16] XIE Y, WANG J J. Application of item response theory in the study of respiratory disease quality of life[J]. Chinese Journal of Gerontology, 2017, 37(4): 1038-1039. DOI: 10.3969/j.issn.1005-9202.2017.04.113.

[17] Wan C, Chen Y, Gao L, et al. Development and validation of the peptic ulcer scale under the system of quality of life instruments for chronic diseases based on classical test theory and generalizability theory[J]. BMC Gastroenterol. 2020, 20(1): 422. DOI: 10.1186/s12876-020-01562-y.

[18] CHEN X Z. Theory and Application of Modern Statistical Analysis Methods[C]. Beijing: National Defense Industry Press, 2016: 221-223.

[19] Jabrayilov R, Emons WHM, Sijtsma K. Comparison of Classical Test Theory and Item Response Theory in Individual Change Assessment[J]. Applied Psychological Measurement. 2016, 40(8): 559-572. DOI: 10.1177/0146621616664046.

[20] Kimman ML, Wijsenbeek MS, van Kuijk SMJ, et al. Validity of the Patient Experiences and Satisfaction with Medications (PESaM) Questionnaire[J]. Patient. 2019, 12(1): 149-162. DOI: 10.1007/s40271-018-0340-6.

[21] Melkamu Asaye M, Gelaye KA, Matebe YH, et al. Assessment of content validity for a Neonatal Near miss Scale in the context of Ethiopia[J]. Global Health Action. 2021, 14(1): 1983121. DOI: 10.1080/16549716.2021.1983121.

[22] CHEN Y L, ZHANG S, CHAI S B, et al. A Study on the Responsiveness of PRO Scale for the Evaluation of TCM Therapeutic Effects on Diabetes Mellitus[J]. Chinese Journal of Traditional Chinese Medicine, 2020, 35(04): 469-476.

[23] Spertus JA, Jones PG. Development and Validation of a Short Version of the Kansas City Cardiomyopathy Questionnaire[J]. Circulation-Cardiovascular Quality and Outcomes. 2015, 8(5). DOI: 10.1161/CIRCOUTCOMES.115.001958.

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