

Psychometric Validation of the Chinese Version of the Diabetes Health Literacy Scale Among Adults with Type 2 Diabetes Mellitus: Application of Factor Analysis and Item Response Theory

Yao Li , Congzhi Wang, Liu Yang , Lin Zhang 

Department of Internal Medicine Nursing, School of Nursing, Wannan Medical College, Wuhu, Anhui, People's Republic of China

Correspondence: Lin Zhang, Department of Internal Medicine Nursing, School of Nursing, Wannan Medical College, 22 Wenchang West Road, Higher Education Park, Wuhu, Anhui, People's Republic of China, Email yaoran2008@163.com

Background: Diabetes health literacy is a critical determinant of a patient's ability to effectively self-manage the condition, and its accurate assessment is paramount for developing targeted interventions. This study aims to translate and validate the reliability and validity of the Chinese version of the Diabetes Health Literacy Scale (DHLS) among Chinese patients with type 2 diabetes mellitus based on factor analysis and item response theory (IRT) models.

Methods: The cross-sectional study was conducted among 251 adult patients with type 2 diabetes mellitus admitted to a tertiary hospital in Wuhu City. The 14-item DHLS was administered to all participants. Reliability was evaluated using Cronbach's alpha. The three-factor structure of the scale was assessed through both exploratory and confirmatory factor analysis (EFA/CFA). Furthermore, the Graded Response Model from IRT was employed to analyze item discrimination, difficulty parameters, and the scale information function.

Results: The Chinese version of the DHLS demonstrated excellent overall reliability (Cronbach's alpha = 0.938). Both EFA and CFA strongly supported the original three-factor structure (informational, computational, and communicative literacy). Convergent validity was satisfactory, with all subscales achieving the average variance extracted (AVE) > 0.50, and composite reliability values > 0.70. Discriminant validity was also confirmed, as the square roots of the AVE for each subscale exceeded the inter-construct correlations. IRT analysis confirmed that all 14 items had satisfactory discrimination parameters and exhibited monotonic increasing difficulty patterns. Importantly, the total information curve revealed that the scale provides the most precise measurement for individuals with moderate levels of health literacy.

Conclusion: The Chinese version of the DHLS is a valid and reliable instrument with robust psychometric properties. Its application can facilitate the development of targeted interventions, ultimately improving diabetes management outcomes.

Keywords: diabetes health literacy scale, type 2 diabetes mellitus, reliability, validity, item response theory

Introduction

Background

Between 1990 and 2022, the global prevalence of diabetes among adults (including both type 1 and 2) doubled, rising sharply from approximately 7% to about 14%, with the most significant increase occurring in low- and middle-income countries (LMICs). However, diabetes treatment rates in LMICs have stagnated at low levels, while residents in North America, Central and Western Europe, and the Pacific region have seen significant improvements in treatment access. This has led to a widening global disparity in diabetes care, leaving nearly 450 million people (59%) with diabetes worldwide untreated in 2022.¹ To address this, it is crucial to vigorously implement health programs in LMICs that facilitate diagnosis and effective management. China currently has the largest diabetic population in the world,

accounting for approximately one-quarter of all global cases. The focus is primarily on type 2 diabetes mellitus (T2DM), which represents over 90% of all diabetes cases in the country. As a highly preventable disease, its high incidence and persistently rising prevalence trend pose significant challenges for chronic disease management.²

Effective self-management is the foundation of diabetes control, and health literacy (HL) is widely recognized as a critical factor influencing self-management behaviors and health outcomes.^{3–6} The Consortium Health Literacy Project European defines HL as the capacity of individuals to access, understand, evaluate, and apply health information to make decisions regarding healthcare, disease prevention, and health promotion.⁷

Literature Survey

Adequate HL empowers patients with diabetes to navigate complex medical information, adhere to prescribed treatments, and engage in scientific dietary and exercise planning, all of which contribute to improved glycemic control and a reduced risk of complications.^{8–10} Currently, a variety of general HL assessment tools are available.^{11–15} Among them, the Rapid Estimate of Adult Literacy in Medicine and the Test of Functional Health Literacy in Medicine are the most frequently adopted in clinical research. The former assesses HL through a word recognition test, while the latter incorporates a numeracy component.¹⁶ However, these generic instruments often fail to comprehensively capture the specific knowledge and skills required for diabetes self-management. Subsequently, the 14-item Health Literacy Scale (HL-14) is a concise instrument specifically designed to assess the HL of diabetic patients in clinical settings.¹⁷ Its primary strength lies in evaluating three key dimensions—functional, communicative, and critical—that patients employ when confronted with health information, which has demonstrated good reliability and validity.^{18,19} Nevertheless, a notable limitation of the HL-14 is its inability to assess numeracy skills, which are essential for diabetes management tasks such as interpreting blood glucose levels and calculating carbohydrate intake based on food packaging labels. To address this gap, Lee et al developed the new comprehensive Diabetes Health Literacy Scale (DHLS).²⁰ This tool is designed to measure the integrated dimensions of informational, numeracy, and communicative HL. The DHLS has demonstrated robust psychometric properties, meeting four validity indicators (content, construct, convergent, and criterion-related validity) and two reliability indicators (internal consistency and test-retest reliability). The psychometric properties of the scale have been examined in both Persian and Malaysian versions, with studies reporting satisfactory reliability and validity.^{21,22} Furthermore, application of the DHLS by Shirooka et al demonstrated its value in reducing distress and burnout, enhancing self-management skills, fostering supportive networks, and ultimately improving patient quality of life.²³ Similar findings were reported in other studies.^{24,25} While the DHLS has been developed internationally, its validity and reliability among Chinese patients with type 2 diabetes remain unvalidated. Given the association between HL and metabolic outcomes in Chinese population,²⁶ developing culturally appropriate assessment tools is crucial.

Study Aim

Therefore, this study aimed to introduce the Chinese version of the DHLS and evaluated its psychometric properties for T2DM patients.

Method

Design and Sample

A cross-sectional study was conducted in Wuhu City from May 2023 to June 2024, involving hospitalized patients with type 2 diabetes from the First Affiliated Hospital of Wannan Medical College. The inclusion criteria were as follows: (1) age ≥ 45 years; (2) clinically confirmed diagnosis of T2DM according to the the International Classification of Diseases (10th Revision); (3) clear consciousness with intact mobility and cognitive ability; and (4) willingness to participate and complete the questionnaire. Patients were excluded if they met any of the following criteria: (1) presence of severe mental disorders or intellectual impairment; (2) severe diabetes-related complications or loss of self-care ability; (3) other critical illnesses (eg, severe cardiovascular diseases, serious infectious diseases, or cancer) or auditory/visual impairment due to T2DM complications; or (4) pregnancy or other specific types of diabetes.

The sample size was determined using the criterion proposed by Kendall, which recommends 10 to 20 times the number of variables.²⁷ Given 14-item questionnaire, we inflated the initial estimate by 20% to compensate for potential

non-response or invalid data, arriving at a minimum target of 168 participants. Ultimately, a total of 251 valid questionnaires were collected.

The Instrument

The DHLS comprises 14 items across three dimensions, assessing psychometric properties through questioning. Items measure composite aspects of informational (Q1:Q7), numeracy (Q8:Q11), and communicative HL (Q12:Q14). Responses are rated on a five-point Likert scale indicating the answer most consistent with the respondent's current state (0 = Strongly Disagree, 4 = Strongly Agree). The original scale demonstrated good reliability across all dimensions, with a Cronbach's alpha of 0.91 and an intraclass correlation coefficient of 0.89.

Translation Procedure

The cross-cultural adaptation was conducted with permission from the original DHLS developers and follows the Brislin model used in other studies.^{28,29} The process consisted of four stages: (1) Forward translation: The original scale was independently translated by a Ph.D. in Nursing and an Doctor of Medicine in Endocrinology, both native Chinese speakers proficient in English. The research team discussed and reconciled discrepancies to produce a consensus version. (2) Back-translation: This consensus version was independently back-translated into English by two medical English teachers who were blind to the original DHLS. This process was iterative, ensuring semantic and content equivalence with the original, resulting in the preliminary Chinese version of the DHLS. (3) Expert review: A multidisciplinary panel comprising three nursing professionals and two psychologists evaluated each item for cultural and linguistic appropriateness. (4) Pilot testing: The pre-final version was administered to 30 diabetic patients, and their feedback was used to refine the wording and finalize the Chinese DHLS. The final content of the DHLS is provided in [Supplementary Table 1](#).

The Stage of Pre-Survey

A pre-survey was conducted in May 2023 using a convenience sample of 30 patients with T2DM who met the inclusion criteria, recruited from the First Affiliated Hospital of Wannan Medical College. During the pre-survey, participant comprehension of each questionnaire item was assessed through interviews. The results indicated a total scale score ranging from 17 to 47, with a mean of 31.27 ± 8.29 . The time to complete the questionnaire ranged from 3 to 5 minutes, averaging 3.53 minutes.

Data Collection

This study was completed using a questionnaire that collected sociodemographic information, diabetes-related conditions, and DHLS. A multistage sampling approach was adopted. First, the First Affiliated Hospital of Wannan Medical College was randomly selected from all tertiary hospitals in Wuhu City, Anhui Province. Second, four departments were purposively chosen within the hospital: endocrinology, geriatrics, traditional Chinese medicine, and dermatology. Third, survey stations were set up in these departments to randomly recruit patients with T2DM. Data collection was carried out in two phases: a pre-survey and a formal survey. All questionnaires were distributed and collected on-site, with each participant completing the questionnaire only once. A total of 260 questionnaires were distributed, and 251 valid responses were returned, with a response rate of 96.54%. To ensure data quality, all investigators received standardized training prior to the survey, which included instruction on patient communication strategies and scale scoring criteria. After obtaining written informed consent, questionnaires were administered in a one-on-one setting and completed independently by participants.

Statistical Analysis

Descriptive statistics were used to summarize the demographic and clinical characteristics of the participants. Continuous variables were presented as mean and standard deviation for normally distributed continuous variables, whereas medians and interquartile ranges were utilized for non-normal continuous variables, while categorical variables were expressed as frequencies and percentages. Group comparisons of scale scores across gender and educational attainment were performed using independent samples *t*-tests. No missing data were encountered in the final sample of 251 participants.

The psychometric validation was conducted using a two-stage analytical strategy that integrates Classical Test Theory (CTT) and Item Response Theory (IRT). This combined approach advances beyond previous validations by evaluating both scale-level properties (reliability, factor structure) and item-level functioning (discrimination, difficulty, precision).

Reliability of the Chinese version of the DHLS was assessed through internal consistency, measured by Cronbach's alpha coefficient. An alpha value greater than 0.70 was considered acceptable. Item-total correlations were also examined, with a threshold of 0.50 used to indicate adequate correlation between each item and the total scale score. Validity was evaluated through multiple approaches: Construct validity was examined using both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The same sample was used for both EFA and CFA. The factorability of the data was confirmed by the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity before conducting EFA with principal axis factoring.^{30–32} EFA (the principal component analysis with Varimax rotation) was used to study the structural validity. With EFA, the criterion for the factor loading of each item was set at not less than 0.40 on its primary factor, and the cumulative contributing rate of the extracted common factors was required to be higher than 40%. Furthermore, the three-factor structure was examined using CFA, and multiple indices were employed to assess model fit: the chi-square to degrees of freedom ratio (χ^2/df) < 5 was considered acceptable, the comparative fit index (CFI) and Tucker-Lewis index (TLI) both > 0.90, the root mean square error of approximation (RMSEA) < 0.08, and the standardized root mean square residual (SRMR) < 0.08. Convergent validity was supported by calculating the average variance extracted (AVE) > 0.50 and composite reliability (CR) > 0.70 for each subscale. Discriminant validity was confirmed using the Fornell-Larcker criterion, whereby the square root of the AVE for each construct was required to exceed the correlations between that construct and other constructs in the model.

IRT analysis was conducted to provide a more nuanced examination of item-level characteristics.^{33–35} Prior to IRT analysis, we assessed whether the data met the assumption of unidimensionality. This was evaluated using the ratio of the first to second eigenvalues from exploratory factor analysis, with a ratio greater than 3:1 considered indicative of a dominant general factor sufficient for unidimensional IRT modeling. Both the Graded Response Model (GRM) and Generalized Partial Credit Model (GPCM) were fitted, with model selection based on the smaller Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The selected model was subsequently used to estimate item discrimination (α) and difficulty parameters (β). Model fit was assessed at both the scale level (M_2 , RMSEA, TLI, CFI, SRMSR) and item level using Orlando-Thissen $S-X^2$ statistics. Item information curves (IICs) and the total (scale) information curve were examined to evaluate the measurement precision across different levels of the latent trait.

All statistical analyses were performed using R software (version 4.5.1). The following packages were utilized for specific analyses: psych for EFA; lavaan for CFA; mirt for IRT modeling; and ggplot2 for scree plots. A two-sided P -value of less than 0.05 was considered statistically significant.

Result

Descriptive Statistics

This study included a total of 251 patients with diabetes mellitus, and the relevant characteristics are described in [Table 1](#). The overall mean age was 64.62 ± 10.37 years, comprising 142 males (56.57%) and 109 females (43.43%). The mean BMI was 24.21 ± 3.26 kg/m². Among the participants, 98.8% were married, 44.62% had an educational level of primary school or below, and 59.76% resided in urban areas. Regarding medical insurance, 43.82% had employee health insurance. In terms of diabetes-related characteristics, 49.00% of patients had complications, 15.94% had a disease duration exceeding 20 years, and 62.15% reported a family history of diabetes.

As shown in [Supplementary Table 2](#), the distribution of DHLS scores across gender and educational attainment reveals notable patterns. Males consistently outperformed females in all three subscales and in the overall composite score (all $P < 0.001$). Furthermore, patients with primary school education or lower consistently demonstrated poorer performance on all HL scales (all $P < 0.001$).

Reliability Analysis

The reliability analysis of the Chinese version of the DHLS demonstrated excellent internal consistency ([Supplementary Table 3](#)), with a Cronbach's alpha coefficient of 0.938. All items were retained in the final scale as deletion of any

Table 1 Frequency Distribution of Patient Characteristics (n=251)

Variables	Groups	Number	%
Demographic information			
Sex	Male	142	56.57
	Female	109	43.43
Educational attainment	Elementary school or below	112	44.62
	Junior high school	80	31.87
	High school or technical secondary school	36	14.34
	College or above	23	9.16
Marital status	Married	248	98.8
	Divorced/never married	3	1.2
Area of residence	Urban	150	59.76
	Town	10	3.98
	Rural	91	36.25
Medical insurance	Employee Medical Insurance	110	43.82
	Urban Resident Medical Insurance	119	47.41
	Out-of-Pocket	22	8.76
Diabetes-related conditions			
Complications	Yes	123	49.00
	No	128	51.00
Duration of illness (years)	< 5	66	26.29
	5–10	83	33.07
	11–20	62	24.70
	>20	40	15.94
Family history	Yes	156	62.15
	No	95	37.85

individual item did not improve the overall alpha coefficient, and all item-total correlations exceeded the recommended threshold of 0.5.

Validity Analysis

Exploratory Factor Analysis

The KMO measure was 0.90, and Bartlett's test of sphericity was significant ($\chi^2 = 3119.70$; $P < 0.001$), supporting the factorability of the data. The scree plot justified the extraction of three factors ([Figure 1](#)), which collectively explained 65% of the total variance, with individual contributions of 34%, 12%, and 19%. All items demonstrated strong loadings (0.49–0.99) on their respective factors, exceeding the 0.40 threshold with no cross-loadings ([Supplementary Table 4](#)), thus confirming the unmodified Chinese version of the scale.

Confirmatory Factor Analysis

CFA was initially conducted to test the original three-factor structure of the Chinese version of the DHLS. The initial model demonstrated an unacceptable fit to the data (see [Table 2](#) and [Figure 2a](#)). The model was subsequently refined using modification indices to enhance model fit. The revised three-factor model showed acceptable fit across multiple indices: $\chi^2/\text{df} = 3.522$, SRMR=0.051, RMSEA=0.079 (0.066, 0.093), TLI = 0.926, and CFI = 0.943 (see [Table 2](#) and [Figure 2b](#)).

The AVE values for the three subscales were 0.760, 0.575, and 0.524, respectively, all exceeding the recommended threshold of 0.50. The corresponding CR values were 0.899, 0.764, and 0.705, all above the 0.70 benchmark, thus supporting satisfactory convergent validity. Discriminant validity was also established. The square roots of the AVE values for each subscale (0.872, 0.758, 0.724) were greater than the correlations between any pair of subscales ($r_{12} = 0.744$, $r_{13} = 0.714$, $r_{23} = 0.692$), confirming that the constructs are distinct from one another.

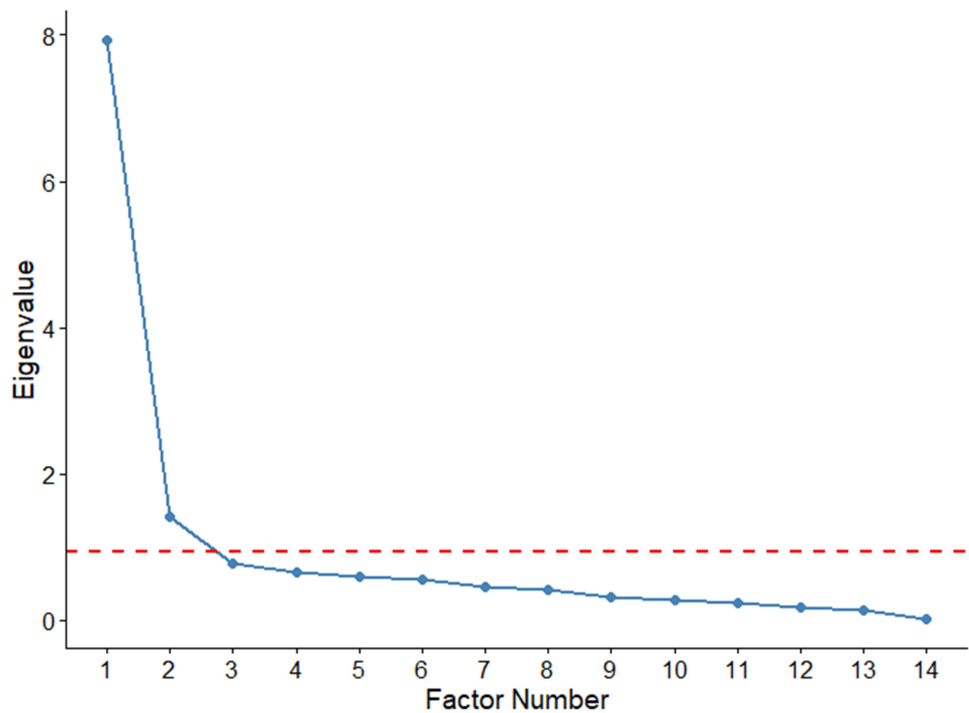


Figure 1 Scree Plot of the Exploratory Factor Analysis.

Item Response Theory Models

The assumption of unidimensionality was supported by a first-to-second eigenvalue ratio of 5.59, indicating a strong dominant factor underlying the scale. Model fit comparison between the GRM and the GPCM supported the selection of the GRM, which showed lower values on both the AIC and the BIC ($AIC_{GRM}=5374.689$, $BIC_{GRM}=5593.268$; $AIC_{GPCM}=5381.763$, $BIC_{GPCM}=5600.341$).

The overall fit of the GRM was acceptable across multiple indices: $M_2=60$ ($P=0.043$), $SRMSR = 0.094$, $RMSEA = 0.077$ (0.060, 0.094), $TLI = 0.906$, and $CFI = 0.914$. While the $M2$ was statistically significant, this is commonly observed in models with moderate to large sample sizes and is interpreted alongside the approximate indices, which collectively suggest acceptable model-data fit. Item-level fit was assessed using the Orlando-Thiesen $S-X^2$ statistic. All 14 items demonstrated satisfactory fit (all $P>0.05$), with item-level $RMSEA$ values ranging from 0.000 to 0.073. Notably, all estimated discrimination parameters (α) fell within an acceptable range of 1.065 to 62.831 (Table 3). Furthermore, the difficulty parameters (β) for each item followed the expected monotonic decreasing pattern from threshold β_1 to β_4 (Table 3). The steep slopes of the Item characteristic curves confirm high discrimination, and the ordered thresholds (β_1 to β_4) are evidenced by the sequential rightward shift of the category response curves for each item (Figure 3a).

Analysis of the IICs revealed a multi-peaked distribution, with Items 1 and 2 providing the greatest informational value (Figure 3b). Collectively, the total (scale) information curve peaked within the ability range of -1 to 1 standard deviation, confirming that the scale offers optimal precision for assessing diabetic patients with moderate levels of the latent trait (Figure 3c).

Table 2 Evaluation of Fitness of Three-Factor Structural Model

Model	CMIN/DF	SRMR	RMSEA	TLI	CFI
Initial model	7.512	0.095	0.161 (0.149, 0.174)	0.810	0.845
Modified model	3.522	0.051	0.078 (0.067, 0.092)	0.926	0.943

Abbreviations: CMIN/DF, Chi-square to degrees of freedom ratio; SRMR, Standardized Root Mean Square Residual; RMSEA, Root Mean Square Error of Approximation; TLI, Tucker–Lewis Index; CFI, Comparative Fit Index.

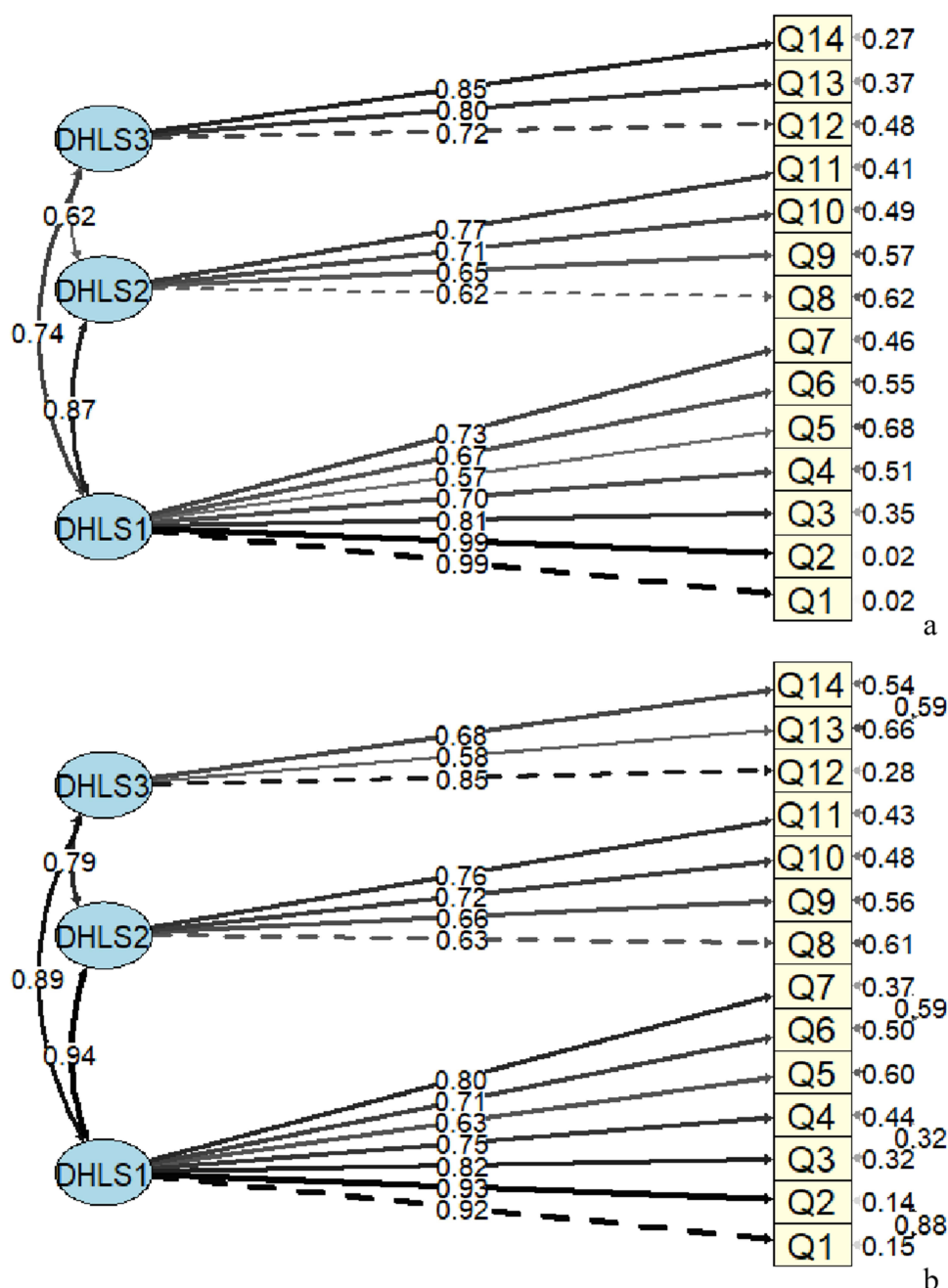


Figure 2 Standardized three-factor structural model of DHLS (n=251). (a) Initial model, (b) Modified model.

Discussion

This study introduced and localized the DHLS and evaluated its psychometric properties among Chinese patients with T2DM using a three-factor structural equation model and IRT modeling. The results demonstrated that the Chinese version of the DHLS exhibits good reliability and validity, with its three-factor structure validated, establishing it as a reliable and effective measurement tool.

In terms of scale reliability, both the total DHLS score and its three constituent subscales demonstrated great internal consistency, with Cronbach's alpha coefficients all exceeding 0.80. This finding aligns closely with the reliability coefficient from the original DHLS development study, which achieved a Cronbach's alpha of 0.91. While the excellent test-retest reliability was validated through an intraclass correlation coefficient of 0.89, it is important to note that this

Table 3 Estimates of Discrimination and Threshold Parameters for the Scale Under the Graded Response Model with the Graded Response Model (n=251)

Items	Threshold				Discrimination
	β_1	β_2	β_3	β_4	α_i
Q1	69.126	1.642	-27.742	-102.208	62.831
Q2	69.078	1.642	-35.658	-102.140	62.788
Q3	6.452	1.099	-6.083	-16.839	3.452
Q4	7.863	1.430	-4.591	-25.728	2.650
Q5	6.369	3.540	-1.005	-5.416	1.832
Q6	5.031	-0.056	-2.670	-7.557	2.201
Q7	5.458	0.556	-2.619	-8.777	2.861
Q8	4.609	2.274	-2.790	-4.618	1.623
Q9	2.822	0.093	-2.224	-4.388	1.327
Q10	6.450	1.171	-3.496	-9.924	1.817
Q11	6.422	1.765	-2.657	-7.540	2.251
Q12	7.746	3.314	-0.370	-7.365	2.134
Q13	6.150	3.305	-0.013	-4.971	1.065
Q14	3.276	0.024	-2.153	-4.324	1.445

study did not assess temporal stability.²⁰ Similarly, the Persian version and Malaysian version scales also demonstrated robust psychometric properties. The Persian version achieved a reliability coefficient of $\alpha=0.919$, with the computational literacy subscale at 0.879, the communication literacy subscale at 0.784, and the information literacy subscale at 0.865.²³ Furthermore, the Malaysian localized version achieved composite reliability indices of 0.962, 0.836, and 0.828 based on Cronbach's alpha coefficients for its three components,²² indicating that the scale maintains strong measurement stability across different cultural contexts.

The combined application of factor analysis and IRT in this study offers complementary validation perspectives and represents a methodological advancement over previous CTT-based approaches to the DHLS in the Chinese context. While factor analysis confirms structural validity at the scale level, IRT delivers essential insights into item-level functioning by revealing how effectively each item discriminates across the HL spectrum and pinpointing where measurement precision is greatest. This dual-method framework advances the field by shifting psychometric validation from confirming scale structure to precisely characterizing measurement performance, thereby establishing a more rigorous and clinically actionable standard for future HL instrument development.

The three-factor structure of the Chinese version was robustly supported, which is consistent with the theoretical model established during the original scale's development,²⁰ and similar to other linguistic adaptations.^{22,23} Both Exploratory and Confirmatory Factor Analyses confirmed the model, with EFA showing three well-defined factors (informational, numeracy, and communicative literacy) without cross-loadings, and CFA indicating an acceptable fit for the modified model. This collective evidence demonstrates that the three target HL constructs are distinct and structurally stable in the Chinese cultural context, aligning well with the core competencies required for diabetes self-management. A cross-cultural comparison of item retention further highlights the scale's adaptability. Both the present study and the Persian version validation retained all scale items without deleting any questions, with all factor loadings exceeding the 0.4 threshold.²³ However, in contrast to these findings, the Malaysian version required the deletion of item Q8 (Calculate the next time to take diabetes medication) due to a factor loading exceeding 1.0.²² The researchers hypothesized that this anomaly might reflect the specific context of long-term medication management among the studied population, where extensive prior education on medication timing could have rendered this item less relevant for measuring health literacy competencies,³⁶ suggesting a potential ceiling effect or cultural specificity in its application.³⁷ Furthermore, differences in DHLS scores across gender and educational groups may reflect sociodemographic disparities in HL access and self-

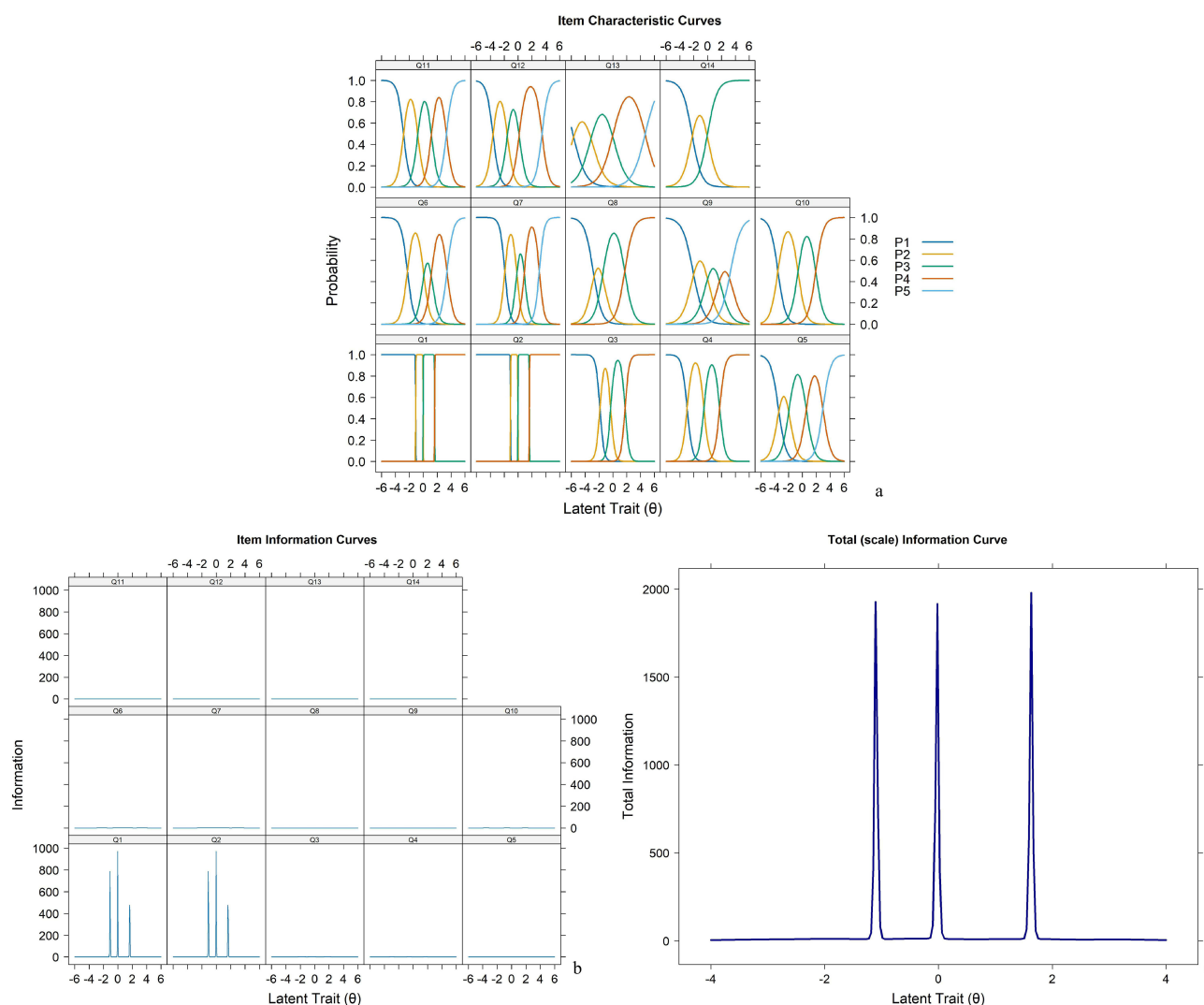


Figure 3 Item response theory model of DHLS. (a) Item characteristic curves for 14 items; (b) Item information curves for 14 items; (c) Total (scale) information curve.

management engagement. These findings underscore the importance of considering population characteristics when interpreting scale scores and designing targeted interventions.

Subsequently, this study applied IRT to conduct an in-depth, item-level evaluation of the scale's measurement properties. Previous validation studies of other language versions of the DHLS did not employ this psychometric approach. The GRM model demonstrated satisfactory discrimination parameters for all 14 items, indicating each item's strong ability to differentiate between respondents with different levels of the underlying HL trait. The difficulty parameters exhibited the expected monotonically increasing pattern across response categories, thereby confirming the validity of the item design and the logical ordering of response options. Analysis of the scale's information function revealed that the DHLS provides the most precise measurement for patients with moderate HL, which indicates the scale most effectively distinguishes and identifies the large patient population whose HL is neither extremely high nor low—yet precisely those most in need of targeted education and support. However, when evaluating extremist groups, scores should be interpreted with caution or supplemented by other assessment methods.

The validation of these measurement characteristics highlights the potential clinical value introduced by the DHLS. The Chinese version of the DHLS effectively fills a crucial gap in domestic tools for diabetes-specific HL assessment, particularly in assessing numeracy skills. Compared to more general HL instruments like the HL-14,³⁸ its design treats numeracy as an independent dimension, thereby enabling precise capture of patients' actual capabilities in core self-

management tasks such as interpreting blood glucose values and calculating carbohydrates. This specialized focus is crucial for comprehensively evaluating the multifaceted nature of health literacy among diabetes patients. From a practical implementation perspective, this scale could provide Chinese clinicians and public health researchers with a concise and effective screening tool. By evaluating scores across three dimensions, educators may be able to develop personalized intervention plans: providing clear science communication materials for patients with insufficient information literacy; offering medication and dietary calculation aids for patients with weak computational skills; and conducting role-play training in doctor-patient communication techniques for patients with poor communication literacy. This multidimensional assessment approach will help optimize the allocation of limited healthcare resources by directing specific interventions to address identified literacy gaps, which might in turn support better clinical outcomes and quality of life, as suggested by prior research.^{11,12} Particularly in a country like China, where the burden of diabetes is heavy and healthcare resources are unevenly distributed,^{39,40} the systematic application of this tool could facilitate the systematic identification of patients with insufficient HL in clinical practice and inform the design of targeted interventions. Therefore, future adoption of this scale may contribute to efforts in tiered diagnosis and treatment, enhance community-based diabetes management, and potentially improve the overall landscape of diabetes care in China.^{41,42}

The main strength of this study lies in its first combined application of the structural equation model and IRT modeling to assess the psychometric properties of the Chinese version of the DHLS. However, our study also presents certain limitations. First, the factor structure was derived and confirmed within the same sample, which may limit the external validity of the structural model. Future research should replicate the CFA in an independent sample. Second, the sample primarily originated from hospitals within the same region, potentially introducing selection bias. Future research should validate the generalizability of the DHLS using broader, more representative samples. Third, this cross-sectional design precludes longitudinal testing of the predictive validity of HL on long-term clinical outcomes (eg, HbA1c changes). Future longitudinal studies are warranted.

Conclusions

The Chinese version of the DHLS exhibits a robust three-dimensional structure, comprehensively assessing the three core competencies of information, calculation, and communication among adults with type 2 diabetes. Notably, the inclusion of calculation as a distinct dimension adds specific value in capturing patients' abilities to handle quantitative self-management tasks. This study employed both CTT and IRT as complementary validation approaches, confirming the scale's reliability, structural validity, and item-level functionality. Critically, IRT analysis indicated that the scale offers optimal measurement precision for individuals with moderate HL, making it particularly suitable for screening and tailoring interventions for this large patient subgroup. However, several limitations should be acknowledged. The use of a single-region hospital-based sample and a cross-sectional design may affect generalizability and preclude causal inferences regarding HL and health outcomes. This scale demonstrates potential for application in clinical screening, HL research, and intervention evaluation. Future multicenter longitudinal studies are recommended to establish its predictive validity and effectiveness in real world diabetes management.

Data Sharing Statement

The data that support the findings of this study are available from the corresponding author [LZ] upon reasonable request.

Human Ethics and Consent to Participate Declarations

Ethical approval for this study was conducted by the Medical Ethics Committee of Wannan Medical College (Approval No: 2021-3). This study was performed in line with the principles of the Declaration of Helsinki. All participants were informed about the purpose of the study, assured of confidentiality, and provided written consent prior to participation. Participation was voluntary, and respondents could withdraw at any time without consequence.

Acknowledgments

We express our great gratitude to all the respondents and all the authors of the scales that have been used in this study.

Author Contributions

Lin Zhang: Conceptualization, Funding acquisition, Validation, Writing – review & editing; Yao Li: Methodology, Data curation, Visualization, Resources, Writing – review & editing; Congzhi Wang: Investigation, Data curation, Methodology, Writing – original draft; Liu Yang: Investigation, Data curation, Writing – original draft. All authors gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the Internal Medicine Nursing Teaching Team (2021jxtd02), Doctoral Research Start-up Fund of Wannan Medical College (X600100191), and Natural Science Research Project of Anhui Provincial Department of Education for Universities and Colleges (2025AHGXZK40175).

Disclosure

The authors declare that there is no conflict of interest.

References

1. Collaboration NCDRF. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. *Lancet*. 2024;404(10467):2077–2093. doi:10.1016/S0140-6736(24)02317-1
2. Xu Y, Lu J, Li M, et al. Diabetes in China part 1: epidemiology and risk factors. *Lancet Public Health*. 2024;9(12):e1089–e1097. doi:10.1016/S2468-2667(24)00250-0
3. Xu Y, Lu J, Li M, et al. Diabetes in China part 2: prevention, challenges, and progress. *Lancet Public Health*. 2024;9(12):e1098–e1104. doi:10.1016/S2468-2667(24)00251-2
4. Ong-Artborirak P, Seangpraw K, Boonyathe S, Auttama N, Winaiprasert P. Health literacy, self-efficacy, self-care behaviors, and glycemic control among older adults with type 2 diabetes mellitus: a cross-sectional study in Thai communities. *BMC Geriatr*. 2023;23(1):297. doi:10.1186/s12877-023-04010-0
5. Chen J, Wen L, Fu G, et al. The relationship between health literacy and blood sugar control in rural areas among diabetes patients. *Front Endocrinol*. 2024;15:1334100. doi:10.3389/fendo.2024.1334100
6. Li J, Lin Y, Li X, et al. Mediating effects of health literacy between self-efficacy and quality of life in patients with type 2 diabetes mellitus. *Diabetes Metab Syndr Obes*. 2025;18:3625–3637. doi:10.2147/DMSO.S546698
7. Sorensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health*. 2012;12:80. doi:10.1186/1471-2458-12-80
8. Abdallah SMA, Ayoub AI, Makhlof MME, Ashour A. Diabetes knowledge, health literacy and diabetes self-care among older adults living with diabetes in Alexandria, Egypt. *BMC Public Health*. 2024;24(1):2848. doi:10.1186/s12889-024-20238-w
9. Butayeva J, Ratan ZA, Downie S, Hosseinzadeh H. The impact of health literacy interventions on glycemic control and self-management outcomes among type 2 diabetes mellitus: a systematic review. *J Diabetes*. 2023;15(9):724–735. doi:10.1111/1753-0407.13436
10. Seger M, Ryan CD, Januszewski AS, et al. Is it time to screen health literacy in diabetes clinical practice? *Diabet Res Clin Pract*. 2024;208:111117. doi:10.1016/j.diabres.2024.111117
11. Pourhabibi N, Mohebbi B, Sadeghi R, et al. Determinants of poor treatment adherence among patients with type 2 diabetes and limited health literacy: a scoping review. *J Diabetes Res*. 2022;2022:2980250. doi:10.1155/2022/2980250
12. Torris C, Nortvedt L. Health literacy and self-care among adult immigrants with type 2 diabetes: a scoping review. *BMC Public Health*. 2024;24(1):3248. doi:10.1186/s12889-024-20749-6
13. Tavousi M, Mohammadi S, Sadighi J, et al. Measuring health literacy: a systematic review and bibliometric analysis of instruments from 1993 to 2021. *PLoS One*. 2022;17(7):e0271524. doi:10.1371/journal.pone.0271524
14. Parmar J, El Masri A, MacMillan F, McCaffery K, Arora A. Health literacy and medication adherence in adults from ethnic minority backgrounds with type 2 diabetes mellitus: a systematic review. *BMC Public Health*. 2025;25(1):222. doi:10.1186/s12889-024-20734-z
15. Zhang L, Ding D, Fethney J, Neubeck L, Gallagher R. Tools to measure health literacy among Chinese speakers: a systematic review. *Patient Educ Couns*. 2020;103(5):888–897. doi:10.1016/j.pec.2019.11.028
16. Mantwill S, Schulz PJ. Health literacy in Mainland China: validation of a functional health literacy test in simplified Chinese. *Health Promot Int*. 2016;31(4):742–748. doi:10.1093/heapro/dav043
17. Ishikawa H, Takeuchi T, Yano E. Measuring functional, communicative, and critical health literacy among diabetic patients. *Diabetes Care*. 2008;31(5):874–879. doi:10.2337/dc07-1932
18. Wu J, Tao Z, Gao N, et al. The use of multidimensional nomial logistic model and structural equation model in the validation of the 14-item health-literacy scale in chinese patients living with type 2 diabetes. *Risk Manag Healthc Policy*. 2023;16:1567–1579. doi:10.2147/RMHP.S419879
19. Batista MJ, Marques ACP, Silva Junior MF, Alencar GP, Sousa M. Translation, cross-cultural adaptation and psychometric evaluation of Brazilian Portuguese version of the 14-item Health Literacy Scale. *Cien Saude Colet*. 2020;25(7):2847–2857. doi:10.1590/1413-81232020257.22282018
20. Lee EH, Lee YW, Lee KW, Nam M, Kim SH. A new comprehensive diabetes health literacy scale: development and psychometric evaluation. *Int J Nurs Stud*. 2018;88:1–8. doi:10.1016/j.ijnurstu.2018.08.002
21. Jafari A, Naddafi F, Gholian-Aval M, Tehrani H. Relationship between diabetes health literacy, distress, burnout, social support, complications, self-care behaviors, and quality of life among patients with type 2 diabetes: a path analysis study. *Diabetol Metab Syndr*. 2024;16(1):150. doi:10.1186/s13098-024-01391-z

22. Ng PL, Arifin WN, Mohamed W, Zakaria R, Kueh YC. Validation of the Malay version of the diabetes health literacy scale among Malaysian adults with type 2 diabetes mellitus. *PeerJ*. 2025;13:e19660. doi:10.7717/peerj.19660
23. Moshki M, Alami A, Zadehahmad Z, et al. Psychometric properties of Persian version of diabetes health literacy scale (DHLS) in patients with type 2 diabetes. *Diabetol Metab Syndr*. 2022;14(1):139. doi:10.1186/s13098-022-00910-0
24. Jafari A, Zadehahmad Z, Armanmehr V, Talebi M, Tehrani H. The evaluation of the role of diabetes health literacy and health locus of control on quality of life among type 2 diabetes using the Path analysis. *Sci Rep*. 2023;13(1):5447. doi:10.1038/s41598-023-32348-3
25. Lee EH, Lee YW, Chae D, et al. Pathways linking health literacy to self-management in people with type 2 diabetes. *Healthcare*. 2021;9(12):1734. doi:10.3390/healthcare9121734
26. Chen Y, Ran X, Chen Y, Jiang K. Effects of health literacy intervention on health literacy level and glucolipid metabolism of diabetic patients in Mainland China: a systematic review and meta-analysis. *J Diabetes Res*. 2021;2021:1503446. doi:10.1155/2021/1503446
27. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. *Educ Psychol Meas*. 2013;76(6):913–934. doi:10.1177/0013164413495237
28. Mei Y, Yang X, Li Y, et al. Validity and reliability of the chinese version of the Diabetes Treatment Burden Questionnaire (DTBQ) among patients with type 2 diabetes. *Diabetes Metab Syndr Obes*. 2023;16:663–675. doi:10.2147/DMSO.S398438
29. Gao Z, Zhang L, Ma J, et al. Reliability and validity of the Chinese version of the self-directed learning instrument in Chinese nursing students. *BMC Nurs*. 2023;22(1):51. doi:10.1186/s12912-023-01201-3
30. Petric G, Atanasova S. Validation of the extended e-health literacy scale: structural validity, construct validity and measurement invariance. *BMC Public Health*. 2024;24(1):1991. doi:10.1186/s12889-024-19431-8
31. Akter J, Islam RM, Chowdhury HA, et al. Psychometric validation of diabetes distress scale in Bangladeshi population. *Sci Rep*. 2022;12(1):562. doi:10.1038/s41598-021-04671-0
32. Wang RH, Kao CC, Su YC, et al. Character strengths use in diabetes self-management scale: development and psychometric testing. *J Adv Nurs*. 2023;79(10):4034–4043. doi:10.1111/jan.15723
33. Egede LE, Walker RJ, Linde S, Williams JS. Identifying individuals with highest social risk in adults with type 2 diabetes using item response theory. *J Gen Intern Med*. 2024;39(9):1642–1648. doi:10.1007/s11606-024-08742-6
34. Garcia-Lugo JP, Carrillo Algarra AJ, Sierra-Matamoros FA, Hernandez-Zambrano SM, Henao-Carrillo DC. Validation using item response theory, and assessment of sensitivity to change of the Hypoglycemia Awareness Questionnaire in patients with type 2 diabetes treated with insulin. *J Diabetes Investig*. 2025;16(5):852–862. doi:10.1111/jdi.14406
35. Haugstvedt A, Nilsen RM, Breivik K, et al. Comparison of information provided on diabetes distress when using PAID-20, PAID-11 or PAID-5. A registry-based study applying IRT analyses among 10,190 adults with type 1 diabetes in Norway. *Diabet Res Clin Pract*. 2025;225:112267. doi:10.1016/j.diabres.2025.112267
36. Lee EH, Lee YW, Lee KW, Hong S, Kim SH. A new objective health numeracy test for patients with type 2 diabetes: development and evaluation of psychometric properties. *Asian Nurs Res*. 2020;14(2):66–72. doi:10.1016/j.anr.2020.01.006
37. Sousa VD, Rojjanasirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract*. 2011;17(2):268–274. doi:10.1111/j.1365-2753.2010.01434.x
38. Levic M, Bogavac-Stanojevic N, Krajnovic D. The instruments used to assess health literacy and pharmacotherapy literacy of diabetes mellitus type 2 patients: a scoping review. *Front Public Health*. 2021;9:747807. doi:10.3389/fpubh.2021.747807
39. Zhou YC, Liu JM, Zhao ZP, Zhou MG, Ng M. The national and provincial prevalence and non-fatal burdens of diabetes in China from 2005 to 2023 with projections of prevalence to 2050. *Mil Med Res*. 2025;12(1):28. doi:10.1186/s40779-025-00615-1
40. Deng W, Zhao L, Chen C, et al. National burden and risk factors of diabetes mellitus in China from 1990 to 2021: results from the Global Burden of Disease study 2021. *J Diabetes*. 2024;16(10):e70012. doi:10.1111/1753-0407.70012
41. Hu H, Liang H, Wang H. Longitudinal study of the earliest pilot of tiered healthcare system reforms in China: will the new type of chronic disease management be effective? *Soc Sci Med*. 2021;285:114284. doi:10.1016/j.socscimed.2021.114284
42. Jia W, Chan JC, Wong TY, Fisher EB. Diabetes in China: epidemiology, pathophysiology and multi-omics. *Nat Metab*. 2025;7(1):16–34. doi:10.1038/s42255-024-01190-w

Diabetes, Metabolic Syndrome and Obesity

Publish your work in this journal

Diabetes, Metabolic Syndrome and Obesity is an international, peer-reviewed open-access journal committed to the rapid publication of the latest laboratory and clinical findings in the fields of diabetes, metabolic syndrome and obesity research. Original research, review, case reports, hypothesis formation, expert opinion and commentaries are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/diabetes-metabolic-syndrome-and-obesity-journal>

Dovepress
Taylor & Francis Group