Patient Preference and Adherence

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Cross-Cultural Adaptation and Psychometric Validation of a Diabetes-Specific Health Literacy Scale in Mainland China

Ru Ma[1], Xia Yu[2], Li Zhang[2], Zihan Hu[1], Yi Zhang[1], Qiangqiang Luo[3], Jiahui Huang[4], Caihong Li[2], Jing Li[1]

[1]School of Nursing, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, 100144, People's Republic of China; [2]Department of Nursing, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, Beijing, 102218, People's Republic of China; [3]Hospital for Skin Diseases, Institute of Dermatology, Chinese Academy of Medical Sciences, Peking Union Medical College, Nanjing, 210042, People's Republic of China; [4]Department of Nursing, Peking Union Medical College Hospital, Beijing, 100730, People's Republic of China

Correspondence: Jing Li, Department of Nursing, School of Nursing, Chinese Academy of Medical Sciences & Peking Union Medical College, 33 Ba Da Chu Road, Beijing, People's Republic of China, Tel +86 10 8877 1124, Email lijing@nursing.pumc.edu.cn; Caihong Li, Department of Nursing, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, Beijing, 102218, People's Republic of China, Tel +86 15 32153 0986, Email iris870716@hotmail.com

Purpose: The study aimed to cross-culturally adapt and psychometrically test the Mandarin version of the Chinese Health Literacy Scale for Diabetes (M-CHLSD).

Patients and Methods: A cross-sectional study was conducted in Beijing with 221 adults diagnosed with type 2 diabetes. The adaptation followed a rigorous five-step process: forward translation, synthesis, back translation, expert review (utilizing a two-round Delphi method), and cognitive interviews using the Questionnaire Appraisal System. Psychometric properties of the M-CHLSD were evaluated using reliability, content validity, and construct validity. Criterion validity was assessed by correlating the scale with glycemic measures and conducting group comparisons using the Mann–Whitney *U*-test.

Results: The Delphi review highlighted the need for linguistic refinement, clinical accuracy, and contextual adaptation, leading to revisions aligned with Mainland China's healthcare system. Pilot testing with 15 patients confirmed comprehensibility. Confirmatory Factor Analysis indicated a good model fit (Comparative Fit Index = 0.93, Tucker-Lewis Index = 0.91, Root Mean Square Error of Approximation = 0.064). The scale demonstrated excellent internal consistency (Cronbach's α=0.884) and test-retest reliability (Intraclass Correlation Coefficient=0.82–0.91). The M-CHLSD showed moderate correlations with fasting plasma glucose (ρ = −0.36) and 2-hour postprandial glucose (ρ = −0.39). The scale successfully distinguished between individuals with good versus poor glycemic control (Cohen's d = 0.54–0.56). A mild ceiling effect (15.1%) was observed.

Conclusion: The M-CHLSD is a robust, reliable, and clinically relevant tool for assessing health literacy in Mandarin-speakers with diabetes. Its rigorous methodology, including Delphi, cognitive interviews, and comprehensive psychometric testing, makes it superior to other adapted tools. The scale's strong clinical correlations and ability to stratify patients based on glycemic control ensure its utility in personalized education and follow-up care pathways.

Keywords: health literacy, type 2 diabetes mellitus, scale, psychometrics

Introduction

Diabetes is one of the most prevalent chronic diseases worldwide. Over the past three decades, China has experienced a significant rise in the prevalence of diabetes, with type 2 diabetes mellitus (T2DM) accounting for over 90% of cases.[1] Despite progress in public awareness, treatment rates, and glycemic control, these indicators remain suboptimal due to poor self-management.[2] Effective self-management is essential for improving glycemic outcomes, enhancing quality of life, and reducing healthcare costs.[3]

Don Nutbeam defines health literacy as the skills enabling individuals to access, understand, and apply health information.[4] Evidences show a positive link between health literacy and self-management, while low health literacy

https://doi.org/10.2147/PPA.S586903
Received: 8 December 2025
Accepted: 27 February 2026
Published: 10 March 2026

increases risks of complications and disability.[5] For instance, enhanced health literacy was correlated with higher adherence to the total of self-care activities, nonpharmacological therapy, and foot care.[6] Conversely, low health literacy is associated with poorer self-management practices and increased risks of diabetes-related complications and disability.[7]

In populations with diabetes, interventions based on health literacy have demonstrated substantial benefits in improving glycemic control and enhancing self-management behaviors. Notably, individuals with a disease duration of more than seven years show greater improvements in self-care following such interventions, underscoring that individuals with longstanding diabetes can still achieve substantial benefits from efforts aimed at improving health literacy.[8] However, about 70% of individuals with T2DM in China have inadequate health literacy, highlighting both clinical gaps and communication barriers between patients and providers.[9] Clinical tools that translate health literacy assessment into actionable insights are therefore needed—not only to assess diabetes-specific knowledge and skills but also to generate actionable insights that guide individualized education and clinical decision making.

The accurate assessment requires valid and reliable tools. However, existing instruments face two limitations. First, most measure general health literacy rather than diabetes-specific skills.[10] Instruments such as the Rapid Estimate of Adult Literacy in Medicine[11] and the Test of Functional Health Literacy in Adults[12] focus on reading comprehension and numeracy skills related to medical contexts. As a result, they fail to capture the unique knowledge, reasoning, and decision-making processes essential for managing diabetes-specific tasks such as insulin titration, carbohydrate counting, and blood glucose monitoring.[13] This lack of disease specificity limits the clinical relevance of such tools when applied to diabetes care.

Second, a notable gap persists between assessment and intervention.[10] Most existing instruments provide only summative scores, yielding a single global metric that fails to identify specific deficits in distinct self-management domains or underlying cognitive processes. Diabetes self-management is, in essence, a cognitive task that requires patients not merely to recall information but to interpret, apply, and adjust decisions in daily contexts. Bloom's taxonomy, which delineates cognitive functioning into six hierarchical levels, namely remembering, understanding, applying, analyzing, evaluating, and creating, offers a framework to map these very competencies.[14] A total score obscures critical variations along this hierarchy; for example, a patient may accurately recall diabetes-related terminology yet struggle to apply this knowledge in real-world scenarios, such as adjusting insulin doses based on blood glucose readings. Without a detailed breakdown of performance across cognitive domains, clinicians cannot pinpoint the precise cognitive level at which a deficit occurs, thereby hindering the development of targeted educational interventions and ultimately undermining the integration of health literacy into routine clinical care.[15]

The Chinese Health Literacy Scale for Diabetes (CHLSD) operationalizes Bloom's Taxonomy into a disease-specific health literacy instrument. Its cognitive level-stratified design enhances clinical relevance and facilitates targeted screening and stratified interventions.[14] However, it was developed and validated in Cantonese for Hong Kong. This poses challenges for application in mainland China. Mandarin is the official language in Mainland China, which differs significantly from Cantonese in lexical, phonological, and syntactic aspects, potentially compromising validity among Mandarin speakers. Furthermore, Hong Kong's healthcare system and educational practices differ from those in mainland China, particularly regarding clinical guidelines, care pathways, and community-based management. Directly applying the Cantonese version may introduce bias and limit its clinical utility.

Therefore, there is a pressing need for a culturally adapted Mandarin version of the CHLSD (M-CHLSD). Such an adaptation would ensure linguistic and contextual appropriateness, enable accurate assessment, and enhance patient-centered communication. The M-CHLSD would support healthcare professionals in delivering tailored education and counseling, while also generating evidence to inform clinical training, program evaluation, and policy development in diabetes care.

Materials and Methods

Design, Setting, and Sampling

A cross-sectional study was conducted in a general hospital and affiliated community health centers in Tiantongyuan, Beijing, to evaluate the validity of the M-CHLSD. Eligibility criteria included: adults aged ≥18 years with a clinical diagnosis of T2DM; cognitive competence (scoring ≥8 on the Short Portable Mental Status Questionnaire); proficiency in Mandarin Chinese for verbal communication.

The CHLSD consists of 34 items. Following statistical guidelines,[16] the required sample was 5–10 times the number of items. With a 15% buffer, the minimum sample was 196. From January to April 2025, 221 patients with T2DM were enrolled, exceeding the requirement.[17]

Written permission for CHLSD adaptation was obtained. This study has received approval from the Ethics Committee of the School of Nursing, Peking Union Medical College (PUMCSON-2025-01) and was conducted in accordance with the principles of the Declaration of Helsinki. All participants provided informed consent before participation. Trained research assistants conducted face-to-face interviews to ensure data quality and confidentiality.

Measurements

Short Portable Mental Status Questionnaire (SPMSQ)

The SPMSQ was used to screen for cognitive competence among participants with T2DM, which consists of 10 items with a binary response scale, where the interviewer assigns a score of 0 for wrong answers and a score of 1 for correct answers. Eligibility required a score≥8, signifying sufficient cognitive function to participate in the study.[18] The scale showed excellent test-retest reliability (0.93, $p < 0.001$), strong convergent validity (0.74, $p < 0.001$), acceptable discriminant validity (0.23, $p < 0.001$), and high internal consistency ($\alpha = 0.82$).[19]

Demographics Questionnaire

A structured questionnaire was used to collect sociodemographic information such as age, gender, educational attainment, occupation, and income, as well as lifestyle factors and diabetes-related clinical parameters.

The Chinese Health Literacy Scale for Diabetes

The CHLSD incorporates disease-specific vocabulary, medication labels, consultation notes, and nursing instructions to assess multiple dimensions of health literacy. The CHLSD has 34 items covering four domains of Bloom's Taxonomy: remembering, understanding, applying, and analyzing. The remembering subscale contains 18 items assessing diabetes vocabulary. Items were scored 2 for correct, 1 for hesitant-correct, and 0 for incorrect. The remaining 16 items include 7 for understanding, 5 for application, and 4 for analysis, reflecting decision-making in daily diabetes care. These items were scored 2 for correct and 0 for incorrect. The CHLSD total score ranges from 0 to 68, with a score of 48 or below denoting inadequate health literacy. The scale demonstrated strong internal consistency and good validity. The Cronbach's alpha for CHLSD and its four subscales were 0.884, 0.885, 0.667, 0.654, and 0.717.[14]

Chinese Version of the Health Literacy Scale

The Chinese Health Literacy Scale, adapted from the Functional Comprehensive Chronic Illness Health Literacy Scale,[20] contains 14 items across functional (5 items), interactive (5 items), and critical health literacy (4 items). Scoring uses a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree), with total scores ranging from 14 to 56. The Cronbach's alpha for the scale and its three subscales were 0.868, 0.839, 0.765, and 0.814.[21]

Translation and Cross-Cultural Adaptation Process

The translation and cultural adaptation of the instrument followed the five-stage cross-cultural validation framework proposed by Koopmans et al, which includes forward translation, synthesis, back translation, expert committee review, and pilot testing.[22]

Forward Translation

Two bilingual experts independently translated the scale into Mandarin. One had a nursing research background and knew the study objectives; the other, with medical training, was blinded to the study purpose to ensure neutrality.

Synthesis

A research investigator convened a consensus meeting with both translators to resolve discrepancies through iterative discussion, culminating in a unified preliminary Mandarin version.

 https://doi.org/10.2147/PPA.S586903

Back Translation

Two additional bilingual medical professionals (a research fellow and a PhD student) independently back-translated the Mandarin version into Cantonese to verify conceptual equivalence with the original scale. Neither translator had prior exposure to the CHLSD, ensuring unbiased reverse translation.

Expert Committee Review

A Delphi consultation was conducted with experts in clinical medicine, nursing, public health, and psychology. The expert review focused on two aspects: (1) semantic accuracy and conceptual consistency of the translation, and (2) content validity and cultural adaptability. The Delphi process comprised two rounds, with 17 experts participating in the first round and 10 in the second. All experts had at least 10 years of professional experience and held intermediate or senior professional credentials. With psychologists required to have this experience specifically in diabetes-related psychological counseling.

Pilot Testing

The preliminary M-CHLSD was pilot-tested among participants recruited through convenience sampling. Cognitive interviews, utilizing think-aloud and probing techniques, evaluated comprehensibility, comprehensiveness, and cultural relevance of the scale. Content analysis was performed using the Questionnaire Appraisal System (QAS),[23] a structured tool designed to identify potential issues in survey questions. Each item was assessed through eight steps: reading, instructions, clarity, assumptions, knowledge/memory, sensitivity/bias, response categories, and other considerations. For each step, potential problems were documented, with reasons specified for any identified issues.

Data Analysis

Statistical analyses were performed using IBM SPSS Statistics 27 and AMOS 24. Descriptive statistics summarized baseline demographic and clinical characteristics. Continuous variables were expressed as mean ± standard deviation (SD), and categorical variables as frequencies and percentages.

Ceiling and Floor Effect

A floor effect was defined as 15% or more of individuals from the sample achieving the score's worst level, and the ceiling effect was defined as when 15% or more of individuals had the best score.[24]

Reliability

Internal consistency was assessed using Cronbach's α, with values ≥0.7 considered acceptable.[25] Test–retest reliability was evaluated using the Intraclass Correlation Coefficient (ICC) and Spearman'ρ over a two-week interval. The ICC was computed using a two-way mixed-effects model for absolute agreement of single measures, denoted as ICC (3,1). ICC values between 0.75 and 0.9 indicate good reliability, and >0.90 indicate excellent.[26] Spearman's ρ was classified as weak (0.10–0.29), moderate (0.30–0.49), or strong (0.50–1.0).[27]

Content Validity

Content validity was examined to ensure the scale comprehensively measured the intended constructs. Ten experts in clinical medicine, nursing, public health, and psychology evaluated each item using a 4-point Likert scale, considering aspects such as grammar, wording, necessity, importance, item placement, and scoring. Content validity was quantified by calculating the content validity index (CVI), probability of chance agreement, and the adjusted Kappa (K*) statistic. Acceptable content validity was defined as an item-level CVI (I-CVI) ≥0.78 and a scale-level CVI (S-CVI) ≥0.90. An adjusted K*>0.74 was considered excellent content validity.[28]

Construct Validity

Construct validity was evaluated by confirmatory factor analysis (CFA). Before CFA, the Kaiser-Meyer-Olkin (KMO) statistic and Bartlett's test of sphericity were used to assess the appropriateness of factor analysis. A KMO value above 0.5 and a significant Bartlett's test ($p<0.05$) indicated adequate sampling adequacy.[29] Model fit criteria were Comparative fit index (CFI), Tucker-Lewis Index (TLI), and Incremental Fit Index (IFI)>0.9; Root Mean Square Error of

Approximation (RMSEA)≤0.08; and Chi-Square/Degrees of Freedom (CMIN/df)<3.[30] Convergent validity required average variance extracted (AVE)>0.50 and composite reliability (CR)>0.70.

Criterion-Related Validity

Criterion-related validity was evaluated by calculating Spearman's rank-order correlations between M-CHLSD scores and three clinical benchmarks: (1) Chinese Health Literacy Scale scores, (2) fasting plasma glucose (FPG) levels, and (3) 2-hour postprandial glucose (2-h PG) levels. Correlation strengths were classified as weak (|rs|<0.3), moderate (0.3≤|rs|≤0.5), and strong (|rs|>0.50).[27] Participants were stratified into adequate and inadequate glycemic control groups according to Chinese Diabetes Society criteria (FPG≤7.0 mmol/L and 2-h PG ≤10.0 mmol/L for adequacy). Differences in M-CHLSD scores between groups were assessed using the Mann–Whitney *U*-test, with Cohen's d calculated to assess effect size: small (0.2≤|d|<0.5), moderate (0.5≤|d|<0.8), and large (|d|≥0.8).[31] Due to the study's design limitations and recruitment period, HbA1c values were not available for all participants, which prevented their inclusion in the criterion validity analysis.

Results

Cross-Cultural Adaptation

Expert Committee Review

After forward translation, synthesis, and back translation, an expert committee review was conducted using the Delphi method. The first round included 17 experts (six clinicians, nine nurses, one public health expert, one psychologist), with 100% response rate and an authority coefficient of 0.9. The second round included 10 experts (four clinicians, six nurses), with 100% response and a coefficient of 0.89. This ensured rigorous multidisciplinary validation.

The first round identified three areas for revision: linguistic adaptation, clinical content accuracy, and contextual applicability. Terms were adjusted to fit mainland usage. For example, "Comprehensive Social Security Scheme" was substituted with "critical illness medical insurance system" to align with local healthcare policies. After these revisions (detailed in Table 1), the second round confirmed linguistic suitability and semantic equivalence, validating cultural adaptation for mainland China.

Pilot Testing

A pilot study using cognitive interviews was carried out with 15 patients diagnosed with T2DM. The first round included 10 participants (four males, six females; mean age 53.7 years) with disease durations ranging from newly diagnosed to 20 years and a mean M-CHLSD score of 60.8. The second round included five patients (three males, two females; mean age 51 years) with disease durations up to 10 years and M-CHLSD scores ranging from 45 to 68 (mean=61), demonstrating consistency across iterations.

QAS coding identified issues in Instructions and Clarity. For example, the Amoxicillin labeling (a penicillin antibiotic) was supplemented with Mainland Chinese terminology; the query method for item B16 (the serial number of the question in M-CHLSD) was clarified to improve semantic transparency; and the rosiglitazone labeling (a medication for T2DM) was simplified to enhance contextual coherence (see Table 1). No further issues were raised in round two, confirming improved comprehensibility and successful adaptation.

Psychometric Testing

Participants' Characteristics

A total of 206 individuals were recruited for the main study using convenience sampling. To assess test-retest reliability, a subsample was selected through purposive sampling based on clinical stability, willingness to participate in a follow-up assessment, and accessibility.[26,32] Of the 40 eligible participants invited to repeat the survey after a two-week interval, 33 completed it. The mean age of participants was 54.71 years (SD = 13.21). The average height and weight were 167.73 cm (SD = 14.44) and 76.53 kg (SD = 16.00), respectively. The mean duration of diabetes was 7.09 years (SD = 8.44). The sample was mostly male (64.6%). The majority of participants were employed (51.9%), lived with family members (94.2%), and were married (93.7%) (see Table 2).

https://doi.org/10.2147/PPA.S586903

Table 1 Cross-Cultural Adaptation Specifications Based on Delphi and Cognitive Interviews

Source	Type	Specific Problem	Solution
First-round Delphi	Linguistic Adaptation	Inadequate linguistic politeness markers Not in common use	Politeness enhancement: "you(ni)"was changed to "you(nin)"[a] Colloquial adaptation: "drink(yin)" was changed to "drink(he)" [a] "physician" was changed to "doctor" "calorie" was changed to "kilocalorie" "diabetes girl"[b] was changed to "diabetes specialist nurse"
	Clinical Content Accuracy	Inaccurate insulin administration instructions Imprecise medication and educational labeling	Defined insulin type, injection timing, and usage protocols in Tip 2. Refined misleading terminology in Tips 5–7 to provide clinically precise descriptions.
	Contextual Applicability	Healthcare system and clinical documentation discrepancies	The Comprehensive Social Security Scheme reference was replaced with the critical illness medical insurance system in Tip 8. Physician identification fields were incorporated (eg, name, credentials) in Tip 4.
First-round Cognitive Interviews	Step 2: Instructions	Low semantic transparency: Ambiguous medication instructions for Amoxicillin	A specified 7-day treatment course was added to the usage instructions to emphasize time-bound antimicrobial efficacy.
	Step 3: Clarity	Poor discourse coherence: Overly complex wording in Rosiglitazone labels Imprecise information mapping: B16 was used without explicit specification	Contextual descriptors in Tip 6 were streamlined to enhance therapeutic coherence. The questioning approach regarding B16 was revised to remove extraneous details and enhance clarity.

Notes: [a]The terms in parentheses are Pinyin (phonetic transcriptions). Although both terms function as second-person pronouns with identical denotations, the latter form is distinguished by its heightened politeness, colloquial register, and prevalence in daily communication. [b]In Hong Kong, diabetes specialist nurses are affectionately referred to as "Diabetes Girls" in local parlance.

Table 2 Participants' Sociodemographic Characteristics (n=206)

Variable	Mean	SD
Age (years)	54.71	13.21
Height (cm)	167.73	14.44
Weight (kg)	76.53	16.00
Duration(years)	7.09	8.44
	Frequency	**Percentage**
Gender, n (%)		
Male	133	64.6
Female	73	35.4
Education, n (%)		
Bachelor's degree and above	74	35.9
High school/Technical school/Vocational school	39	18.9
Associate degree	50	24.3
Junior high school and below	43	20.9
Employment status, n (%)		
Working	107	51.9
Not working (retired or unemployed)	99	48.1
Living status, n (%)		
Live with family	194	94.2
Live alone	12	5.8

(*Continued*)

 https://doi.org/10.2147/PPA.S586903

Table 2 (Continued).

Per capita monthly income of the household		
>5000	171	83
2000–5000	26	12.6
<2000	9	4.4
Marital Status		
Married	193	93.7
Single	7	3.4
Divorced or Widowed	6	2.9

Ceiling and Floor Effect

Among the 206 participants, 15.1% attained the maximum score of 68, indicating a ceiling effect. In contrast, no floor effect was observed.

Reliability

As detailed in Table 3, the internal consistency of the M-CHLSD subscales was excellent, with Cronbach's α coefficients of 0.95 for Remembering, 0.90 for Understanding, 0.82 for Applying, and 0.87 for Analyzing. The scale demonstrated excellent test–retest reliability. The ICC was 0.92 for the total score and ranged from 0.795 to 0.986 across subscales. In addition, Spearman's ρ coefficients ranged from 0.760 to 0.946 ($p < 0.001$), supporting strong temporal stability across all domains.

Content Validity

All 34 items showed an I-CVI of≥0.90 and a S-CVI of>0.90. The probability of chance agreement was negligible (Pc≤0.01), and the adjusted K* values for all items ranged from 0.899 to 0.972, confirming excellent inter-rater consistency and strong content validity.

Construct Validity

The KMO measure was highly satisfactory at 0.89, and Bartlett's test of sphericity reached statistical significance ($p<0.001$), confirming the appropriateness of factor analysis. CFA demonstrated an optimal model fit (IFI=0.91, TLI=0.90, CFI=0.91, RMSEA=0.06, and CMIN/df=1.78). CR values ranged from 0.827 to 0.951 (all>0.7), and the AVE was≥0.50 in all domains. All factor loadings were≥0.45, meeting the acceptability criterion (see Table 4).

Criterion-Related Validity

Spearman correlation analysis demonstrated a moderate positive association between the M-CHLSD and the Chinese Health Literacy Scale (ρ=0.399, $p<0.001$), supporting convergent validity. Regarding glycemic outcomes, the scale was

Table 3 Indicators of Internal Consistency and Test-Retest Reliability

Domain	Coefficient Type	Coefficients	95% CI	p values	Cronbach's α
M-CHLSD	ICC(3,1)	0.938	0.874–0.969	<0.001	0.86
	Spearman's ρ	0.866	0.739–0.934	<0.001	
Remembering	ICC(3,1)	0.795	0.579–0.900	<0.001	0.95
	Spearman's ρ	0.760	0.557–0.878	<0.001	
Understanding	ICC(3,1)	0.912	0.822–0.957	<0.001	0.90
	Spearman's ρ	0.837	0.688–0.919	<0.001	
Applying	ICC(3,1)	0.963	0.926–0.982	<0.001	0.82
	Spearman's ρ	0.920	0.839–0.961	<0.001	
Analyzing	ICC(3,1)	0.986	0.972–0.993	<0.001	0.87
	Spearman's ρ	0.946	0.890–0.974	<0.001	

Note: ICC (3,1) indicates a two-way mixed-effects model for absolute agreement. All p-values reflect two-tailed significance.
Abbreviations: ICC, Intraclass Correlation Coefficient; CI, Confidence Interval.

https://doi.org/10.2147/PPA.S586903

Table 4 Factor Loadings and Convergent Validity

Domains	Items	Factor Loading	AVE	CR
Remembering	A1	0.67	0.52	0.95
	A2	0.63		
	A3	0.81		
	A4	0.77		
	A5	0.69		
	A6	0.75		
	A7	0.75		
	A8	0.53		
	A9	0.70		
	A10	0.87		
	A11	0.70		
	A12	0.65		
	A13	0.74		
	A14	0.57		
	A15	0.78		
	A16	0.75		
	A17	0.78		
	A18	0.78		
Understanding	B3	0.77	0.57	0.90
	B8	0.87		
	B9	0.88		
	B10	0.77		
	B11	0.63		
	B12	0.70		
	B13	0.59		
Applying	B4	0.73	0.50	0.83
	B5	0.67		
	B14	0.79		
	B15	0.83		
	B16	0.46		
Analyzing	B1	0.86	0.68	0.89
	B2	0.96		
	B6	0.86		
	B7	0.59		

moderately negatively correlated with both fasting plasma glucose ($\rho=-0.340$, $p<0.001$) and 2-hour postprandial glucose ($\rho=-0.324$, $p<0.001$), highlighting its utility in assessing health literacy relevant to diabetes management. Furthermore, Mann–Whitney *U*-tests showed that participants with adequate glycemic control had significantly higher M-CHLSD scores (fasting glucose: 59.80 vs.54.08, Cohen's d=0.56; 2-h postprandial glucose: 59.45 vs.53.86, d=0.54), demonstrating good discriminative validity across clinical benchmarks (Table 5).

Discussion

In clinical practice, adequate health literacy is essential for the effective management of chronic conditions such as type 2 diabetes. Validated assessment tools allow healthcare providers to identify patients with limited health literacy and to design targeted educational interventions, which are critical for implementing evidence-based self-management strategies. The CHLSD has been translated and applied in various international settings. To our knowledge, this is the first study to culturally adapt and validate a Mandarin version of the CHLSD for use in mainland China. By addressing both linguistic and healthcare system differences from the original Cantonese version, the M-CHLSD fills an important gap in diabetes-specific health literacy assessment among Mandarin-speaking populations.

Table 5 Group Comparisons of M-CHLSD Scores by Glycemic Control Status

Glycemic Indicator	Inadequate Control (n = 125)	Adequate Control (n = 81)	U	Z	p-values	Cohen's d [95% CI]
Fasting plasma glucose ((FPG > 7 mmol/L))	M = 54.08 SD = 11.01	M = 59.80 SD = 9.01	3315	−4.20	<0.001	0.56 [0.27, 0.84]
2-h Postprandial glucose (2hPG ≥ 10 mmol/L)	M = 53.86 SD = 11.05	M = 59.45 SD = 9.20	3427	−4.27	<0.001	0.54 [0.26, 0.82]

Note: All p-values reflect two-tailed significance.
Abbreviation: CI, Confidence Interval.

Compared with the Iranian version,[33] which followed the WHO forward-backward translation method with minor terminology adjustments, our study adopted a more comprehensive and structured adaptation approach. This included two rounds of Delphi consultation with experts, cognitive interviews with patients, and systematic item appraisal using the QAS-99. This process enabled precise revisions to address linguistic ambiguities, inconsistencies related to health policy, and culturally inappropriate medical terminology. The resulting instrument demonstrated superior psychometric performance relative to the Cantonese and Persian versions, with higher internal consistency, stronger test-retest reliability, and improved model fit indices. In contrast to the Turkish adaptation, which emphasized organizational health literacy and communication strategies within its own healthcare context, the M-CHLSD was specifically tailored to the linguistic environment and healthcare structures of mainland China, including its national insurance policies, community health models, and patient education pathways. These contextual adaptations enhance the tool's relevance and practical utility in clinical settings.[34]

Beyond its measurement properties, the M-CHLSD holds strong clinical and educational value. In our study, higher scores were significantly associated with better glycemic control. Patients with inadequate glycemic control (FPG > 7 mmol/L or 2hPG ≥ 10 mmol/L) had significantly lower M-CHLSD total scores (FPG >7 mmol/L: 54.08; 2hPG ≥ 10 mmol/L: 53.86) compared to those with adequate control (59.80 and 59.45, respectively), indicating good discriminative validity by clinical outcome. This clinical sensitivity supports patient stratification and enables tailored intervention strategies. Grounded in Bloom's Taxonomy, the M-CHLSD aligns with the cognitive progression from remembering and understanding to application and analysis. For patients with low literacy, the scale can help clinicians recognize the need for simplified language, visual materials, and repetition. For those with higher literacy, education can focus on applied self-management strategies, problem-solving, and shared decision-making. The tool therefore serves not only as a diagnostic instrument, but also as a foundation for planning communication, designing personalized education pathways, and informing clinical training curricula.[35] Although the mean scores observed in the inadequate-control group (approximately 53–54) suggest an elevated risk of poor glycemic outcomes, a clinically applicable cut-off requires formal validation, such as through ROC curve analysis.

At the healthcare system level, the M-CHLSD supports the development of health-literate care environments by highlighting the importance of matching educational content to patients' cognitive readiness. Its application can inform institutional policies and professional training programs aimed at improving provider–patient communication,[36] especially within chronic disease care. From an international perspective, this study contributes a rigorous, multi-phase methodology for the cross-cultural adaptation of health literacy instruments. The approach, which includes Delphi consultation, cognitive interviewing, and psychometric validation, can serve as a valuable reference for similar adaptations in other languages, diseases, or health systems.

This study also has limitations. First, data were collected from a single city, which may limit generalizability to broader populations across China. Second, the cross-sectional design prevents assessment of the scale's predictive validity for long-term clinical outcomes. Third, the presence of a moderate ceiling effect indicates that the current scale may not fully capture distinctions among individuals with higher levels of health literacy. Fourth, the sample size for test-retest reliability assessment (n = 33) fell short of the COSMIN guideline recommendation of at least 50 participants, potentially leading to an imprecise or overestimated intraclass correlation coefficient. Future research should explore the scale's predictive performance in longitudinal studies and assess the impact of tailored interventions guided by M-CHLSD results. Expanding the scale to include domains such as dietary behaviors and physical activity may also enhance its discriminative power across the full spectrum of health literacy.

Conclusion

This study completed a rigorous cross-cultural adaptation and comprehensive psychometric evaluation of the CHLSD for use in mainland China, resulting in the Mandarin version (M-CHLSD). Through a structured process involving forward–backward translation, expert committee review via a two-round Delphi method, and cognitive debriefing with patients, linguistic equivalence, clinical accuracy, and contextual relevance were rigorously established. The M-CHLSD demonstrated strong content validity, internal consistency, and test-retest reliability, with construct validity confirmed through confirmatory factor analysis. Its significant associations with glycemic outcomes further support the instrument's discriminative and criterion-related validity, underscoring its clinical sensitivity. These measurement properties establish the M-CHLSD as a valid, reliable, and culturally appropriate tool for assessing diabetes-specific health literacy among Mandarin-speaking adults in mainland China.

The M-CHLSD is a culturally adapted, diabetes-specific health literacy tool that enables clinicians to accurately identify patients with limited health literacy and tailor education to their cognitive needs. By linking assessment results to specific domains such as application or understanding, it supports targeted interventions, improves communication, and enhances self-management. Its association with glycemic outcomes underscores its clinical utility in risk stratification and personalized care planning within routine diabetes management.

Abbreviations

T2DM, Type 2 Diabetes Mellitus; REALM, The Rapid Estimate of Adult Literacy in Medicine; TOFHLA, The Test of Functional Health Literacy in Adults; CHLSD, The Chinese Health Literacy Scale for Diabetes; M-CHLSD, Mandarin version of the Chinese Health Literacy Scale for Diabetes; FCCHL, The Functional Comprehensive Chronic Illness Health Literacy Scale; QAS, Questionnaire Appraisal System; ICC, Intraclass Correlation Coefficient; CFA, Confirmatory Factor Analysis; KMO, Kaiser-Meyer-Olkin; CVI, Content Validity Index; I-CVI, Item-level CVI; S-CVI, Scale-level CVI; CFI, Comparative Fit Index; TLI, Tucker-Lewis Index; IFI, Incremental Fit Index; RMSEA, Root Mean Square Error of Approximation; CMIN/df, Chi-Square/Degrees of Freedom; AVE, Average Variance Extracted; CR, Composite Reliability; FPG, Fasting Plasma Glucose; 2-h PG, 2-hour Postprandial Glucose.

Data Sharing Statement

The datasets used and/or analyzed during the current study are available from the corresponding author (Jing Li) on reasonable request.

Ethics Approval and Informed Consent

This study has received approval from the Ethics Committee of the School of Nursing, Peking Union Medical College (PUMCSON-2025-01).

Acknowledgments

We gratefully acknowledge Professor Angela YM Leung for authorizing the cultural adaptation of the Chinese Health Literacy Scale for Diabetes and providing expert guidance during the process.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; 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 work was supported by the School of Nursing, Peking Union Medical College (Research number: PUMCSON202510).

Disclosure

The authors declare that they have no competing interests.

References

1. 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
2. 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
3. Powers MA, Bardsley JK, Cypress M, et al. Diabetes self-management education and support in adults with type 2 diabetes: a consensus report of the American Diabetes Association, the Association of Diabetes Care and Education Specialists, the Academy of Nutrition and Dietetics, the American Academy of Family Physicians, the American Academy of PAs, the American Association of Nurse Practitioners, and the American Pharmacists Association. *J Acad Nutr Diet*. 2021;121(4):773–788.e9. doi:10.1016/j.jand.2020.04.020
4. Nutbeam D. Health literacy as a public health goal: 25 years on. *Health Promot Int*. 2025;40(4). doi:10.1093/heapro/daaf119
5. Ong-Artborirak P, Seangpraw K, Boonyathee 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
6. Rodrigues S, Isabel Patrício A, Cristina C, et al. Health literacy and adherence to therapy in type 2 diabetes: a cross-sectional study in portugal. *Health Lit Res Pract*. 2024;8(4):e194–e203. doi:10.3928/24748307-20240625-01
7. Sriklo M, Sermsinsiri K. Effects of a health literacy program on self-management behaviors of patients with uncontrolled type 2 diabetes in Thailand: a quasi-experimental study. *Belitung Nurs J*. 2026;12(1):97–105. doi:10.33546/bnj.4156
8. 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
9. Meilan J, Chunyu L, Xiaoshuang Z, Caifu L. Analysis on the health literacy status and influencing factors in type 2 diabetic patients. *J Med Sci Yanbian Univ*. 2011;34(2):3.
10. Ru M, Wumei H, Jing L. A scoping review of health literacy interventions for type 2 diabetes patients. *Chin J Nurs*. 2024;59(19):2420–2426. doi:10.3761/j.issn.0254-1769.2024.19.017
11. Davis TC, Crouch MA, Long SW, et al. Rapid assessment of literacy levels of adult primary care patients. *Fam Med*. 1991;23(6):433–435.
12. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *J Gen Intern Med*. 1995;10(10):537–541. doi:10.1007/bf02640361
13. 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
14. Leung AY, Lou VW, Cheung MK, Chan SS, Chi I. Development and validation of Chinese health literacy scale for diabetes. *J Clin Nurs*. 2013;22(15–16):2090–2099. doi:10.1111/jocn.12018
15. Paasche-Orlow MK, Wolf MS. Addressing health literacy. *JAMA*. 2025;333(20):1826–1827. doi:10.1001/jama.2025.2198
16. Kline R. *Principles and Practice of Structural Equation Modeling*. 4th ed. Guilford publications; 2016.
17. Hair J, Hult GTM, Ringle C, Sarstedt M. A primer on Partial Least Squares Structural Equation Modeling (PLS-SEM); 2022.
18. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *J Am Geriatr Soc*. 1975;23(10):433–441. doi:10.1111/j.1532-5415.1975.tb00927.x
19. Martínez de la Iglesia J, Dueñas Herrero R, Onís Vilches MC, Aguado Taberné C, Albert Colomer C, Luque Luque R. Adaptación y validación al castellano del cuestionario de Pfeiffer (SPMSQ) para detectar la existencia de deterioro cognitivo en personas mayores de 65 años [Spanish language adaptation and validation of the Pfeiffer's questionnaire (SPMSQ) to detect cognitive deterioration in people over 65 years of age]. *Med Clin*. 2001;117(4):129–134. Spanish. doi:10.1016/s0025-7753(01)72040-4
20. 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
21. Zhao X, Li Y, Zeng Y, Chen L. Evaluation of the reliability and validity of Health Literacy Scale-Chinese. *Chin J Nurs Educ*. 2021;18(3):266–269. doi:10.3761/j.issn.1672-9234.2021.03.015
22. Koopmans L, Bernaards CM, Hildebrandt VH, Lerner D, de Vet HC, van der Beek AJ. Cross-cultural adaptation of the Individual Work Performance Questionnaire. *Work*. 2015;53(3):609–619. doi:10.3233/wor-152237
23. Jang MK, Kim S, Collins EG, Quinn LT, Park CG, Ferrans CE. Enriching the quality of cross-cultural instrument development through cognitive interviewing: implications for nursing research. *Jpn J Nurs Sci*. 2020;17(2):e12301. doi:10.1111/jjns.12301
24. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4(4):293–307. doi:10.1007/bf01593882
25. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42. doi:10.1016/j.jclinepi.2006.03.012
26. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. doi:10.1016/j.jcm.2016.02.012
27. Khamis H. Measures of Association: how to choose? *J Diagnos Med Sonogr*. 2008;24(3):155–162. doi:10.1177/8756479308317006
28. Lynn MR. Determination and quantification of content validity. *Nurs Res*. 1986;35(6):382–385. doi:10.1097/00006199-198611000-00017
29. Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas*. 1960;20(1):141–151. doi:10.1177/001316446002000116

30. McDonald RP, Ho MH. Principles and practice in reporting structural equation analyses. *Psychol Methods*. 2002;7(1):64–82. doi:10.1037/1082-989x.7.1.64
31. Glaros AG. Statistical significance, clinical importance and effect sizes: enhancing understanding of a study's results. *J Oral Rehabil*. 2025;52(8):1185–1192. doi:10.1111/joor.13759
32. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;17(1):101–110. doi:10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E
33. Ahmadi A, Niknami S, Ghaffari M. Type 2 diabetes health literacy assessment tool: translation and psychometric evaluation of the Iranian version. *Int J Endocrinol Metab*. 2022;20(2):e116983. doi:10.5812/ijem-116983
34. Krosnick JA. Improving question design to maximize reliability and validity. In: *The Palgrave Handbook of Survey Research*. Springer; 2017:95–101.
35. Ogbeide SA, Johnson-Esparza Y, George D. Applying bloom's taxonomy in primary care behavioural health training. *Clin Teach*. 2025;22(3):e70099. doi:10.1111/tct.70099
36. Mor-Anavy S, Lev-Ari S, Levin-Zamir D. Health literacy, primary care health care providers, and communication. *Health Literacy Res Pract*. 2021;5(3):e194–e200. doi:10.3928/24748307-20210529-01