Quality of Life Research (2024) 33:1593–1603
https://doi.org/10.1007/s11136-024-03648-6

Construct validity of the 12-item Short Form Health Survey (SF-12) version 2 and the impact of lifestyle modifications on the health-related quality of life among Indian adults with prediabetes: results from the D-CLIP trial

Euisun Pyo[1] · Mary Beth Weber[1,3,7] · Jayalakshmi Sivaram[2] · Lisa R. Staimez[1,3,7] · Viswanathan Mohan[4] · Ranjit Mohan Anjana[4] · Regine Haardörfer[5] · Harish Ranjani[6]

Accepted: 8 March 2024 / Published online: 12 April 2024

Abstract

Purpose This study aimed to validate the factor structure of the 12-item Short-Form (SF-12) health-related quality of life (HRQOL) survey for Indian adults and assess the impact of lifestyle modification on the SF-12 of Indian adults with prediabetes.

Methods To validate the context-specific construct of the SF-12, two-factor confirmatory factor analysis (CFA) was performed using data from 1285 adults residing in Chennai, India, who screened for the Diabetes Community Lifestyle Improvement Program (D-CLIP). D-CLIP was a randomized controlled trial of 578 participants with prediabetes (283 treatment, 293 control), focusing on the effect of lifestyle modifications on the prevention of diabetes. Physical and mental component scores (PCS and MCS) were computed by using CFA standardized factor loadings. Multiple linear regression was subsequently conducted to estimate the effect of lifestyle modification on post-study changes of PCS and MCS among D-CLIP participants.

Results Cronbach's alpha and CFA fit indices demonstrated acceptable reliability and model fit of the SF-12 for Indian adults. The intervention group showed greater mean change in PCS after study participation compared to the controls (1.63 ± 0.82, $p = 0.046$); no significant difference was observed for MCS between two groups (1.00 ± 0.85, $p = 0.242$).

Conclusion The study confirmed that the SF-12 is suitable for assessing the physical and mental health dimensions of HRQOL for Indian adults. Our findings suggest that the benefits of diabetes prevention lifestyle modification strategies may primarily enhance the physical well-being of adults with prediabetes. Further studies validating the SF-12 in a broader Asian Indian population are needed.

Trial registration Clinicaltrials.gov, NCT01283308.

Keywords Health-related quality of life · SF-12 · Diabetes · India · CFA · Survey validation

✉ Euisun Pyo
pyo.euisun@gmail.com

1 Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta, GA, USA

2 Madras Diabetes Research Foundation, Chennai, India

3 Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

4 Department of Diabetology, Madras Diabetes Research Foundation and Dr. Mohan's Diabetes Specialities Centre, Chennai, India

5 Department of Behavioral, Social, and Health Education Sciences, Rollins School of Public Health, Emory University, Atlanta, GA, USA

6 Department of Preventive and Digital Health Research, Madras Diabetes Research Foundation, Chennai, India

7 Emory Global Diabetes Research Center, Woodruff Health Sciences Center, Emory University, Atlanta, GA 30322, USA

Introduction

India has an estimated 101 million adults with type 2 diabetes (T2DM) [1, 2]. Disease status alone is insufficient to understand and address the challenges faced by individuals with T2DM or comprehensively assess and monitor the impact of T2DM on an individual's well-being [3, 4]. Health-related quality of life (HRQOL) refers to experienced physical, social, and psychological health, and measurements of HRQOL offer greater insights into the lived experiences of various diseases [5, 6]. HRQOL surveys included two classes of measures: disease-specific descriptive measures of functional health status and generic preference-based measures of health and well-being [7–10]. While disease-specific HRQOL surveys include questionnaires targeted to the symptoms and functioning relevant to a certain disease, various disease applications and cross-contextual comparisons are possible with generic measures. Currently, numerous generic HRQOL instruments exist and have been extensively applied to interventions and policies to quantify their program efficacy and potential economic benefits. The commonly used examples are the WHO Quality of Life Questionnaire (WHOQOL) [11], the European Quality of Life Five Dimension—Visual Analogue Scale (EQ5D-VAS) [12], and the Short Form Health Survey 36 or 12 (SF-36, SF-12) [13, 14].

The SF-12 is a 12-item self-reported HRQOL instrument developed by QualityMetric Inc. as a modified version from their 36-item survey, the SF-36 [13]. The SF-12 has been extensively used in diverse research and program settings—particularly for large-scale interventions, national surveys, and populations with limited cognitive abilities—due to its low respondent burden [13–15]. The SF-12 is widely recognized for its reliability and responsiveness for various health outcomes [16, 17]. However, validating the survey is an especially important step when applying HRQOL surveys to non-Western countries since most of the existing instruments had been developed based on the US and European populations who might perceive social, cultural, and economical functioning differently than other populations.

Despite its wide applications in India [18–22], there are limited data on the validity of the SF-12 within Asian Indian populations. One study used Spearman correlations of the subdomains and reported Cronbach's alpha to estimate the internal reliability between the questionnaire items or subdomains of the SF-12 [21]. However, to apply the psychometric instrument to populations other than the USA, it is crucial to investigate beyond the internal reliability and assess the dimensions between the survey items and the latent factors that the instrument is designed to measure. To investigate the latent factors explained by a measurement, factor analysis such as principal component analysis (PCA), exploratory factor analysis (EFA), and confirmatory factor analysis (CFA) is commonly used [23]. While exploratory approaches such as PCA and EFA have been used in past studies when applying surveys to new population contexts, CFA is considered the appropriate method to validate the latent structures of a psychometric instrument since this approach tests relationships (or variance/co-variance structures) hypothesized a priori between the survey items (subdomains in the case of the SF-12) and the latent factor (or the SF-12 component scores) in one model [24, 25].

Studies in the USA have shown improvements in HRQOL among participants in lifestyle intervention studies designed to reduce the risk of diabetes [26, 27]. However, it is unclear how lifestyle modification programs will impact HRQOL in India, and limited research exists that explores the changes in HRQOL in the intervened participants. While the impact of a diabetes prevention program (DPP) on HRQOL has previously been assessed in the Indian population by the Kerela Diabetes Prevention Programme (KDPP) using the Short Form 6 Dimension (SF-6D), no other studies contribute to this evidence base, and using the SF-12 instead the SF-6D can provide more information on the multiple dimensions of HRQOL [28, 29].

The Diabetes Community Lifestyle Improvement Program (D-CLIP) implemented a group-based lifestyle DPP among adults with prediabetes in Chennai, India; a detailed protocol of the study can be found in a previous publication [30] and on Clinicaltrials.gov (NCT01283308). The trial found that the treatment group had a significantly lower rate of developing T2DM compared to the control, highlighting the effectiveness of lifestyle modifications against diabetes prevention [31]. This study aims to expand our understanding of the D-CLIP intervention outcomes by exploring the impact of lifestyle strategies on HRQOL of those participating in the intervention. Hence, the objectives of this study were (1) to first confirm the survey structure of the SF-12 for an Indian context using CFA and (2) assess the effect of D-CLIP on the SF-12 of Indian adults at risk of T2DM.

Methods

Study design and participants

D-CLIP was a randomized controlled translation trial (recruitment 2009–2012, $n = 578$ adults with prediabetes and overweight/obesity status) which aimed to prevent diabetes development among adults with prediabetes (impaired glucose tolerance [IGT], impaired fasting glucose [IFG] or IGT + IFG) by comparing expert recommendations for

diabetes prevention (lifestyle modification with metformin added as needed) to standard of care [30]. D-CLIP inclusion criteria were: adults aged 20–65 years; weight status categorized as overweight or obese based on the World Health Organization South Asian standards (BMI ≥ 23 kg/m^2 and/or waist circumference ≥ 90 cm for men and ≥ 80 cm for women) [32]; having prediabetes (baseline fasting plasma glucose, FPG, indicating IFG: 100–125 mg/dL and/or 2-h post-load glucose indicating IGT: 140–199 mg/dL); and not pregnant, breastfeeding, or reporting significant health issues.

Recruitment occurred at large-scale community screenings and referrals from the clinic databases in Chennai, India. Detailed recruitment procedure has been described previously [30]. After initial in-field screening, a non-random sample of 1285 adults (the "baseline screening sample") attended clinic-based testing, which served as a secondary screening step for D-CLIP participation and baseline testing for eligibility. Subsequently, 578 individuals were randomly assigned using a random number list to either the treatment or the control group. The control group ($n = 293$) received standard of care, including a health visit and diabetes prevention classes. The treatment group ($n = 283$) received six months of intervention including: 4 months of weekly group-based behavioral counseling for lifestyle modifications, 2 months of group-based maintenance classes, and for those at high risk of converting to diabetes after four or more months of participation (IFG + IGT or IFG + HbA1c ≥ 5.7%), twice daily prescription of up to 500 mg metformin. Due to the nature of the trial and the group-specific assessments, it was not feasible to maintain randomization blinding from the study participants, staff, and investigators. Follow-up assessments were conducted every 6 months for the duration of the trial. The primary health outcome of the D-CLIP trial was T2DM incidence which was diagnosed based on either annual oral glucose tolerance test (OGTT) or semiannual fasting blood glucose (FPG) [31]. D-CLIP participants who were diagnosed with T2DM during follow-up were excluded from further study activities and assessments. Sample size of 600 was calculated to detect a 35% group difference in T2DM incidence with 80% power, assuming $\alpha = 0.05$, 10% loss to follow-up, and 9% T2DM incidence per year for the control [30].

To validate the survey constructs of the SF-12 in Indian adults at risk of diabetes, we used the SF-12 data collected from the baseline screening sample with self-reported overweight and obesity but no previously known diagnosis of diabetes; four individuals (0.3% of the sample) were excluded from the analysis due to missing SF-12 data. For our secondary objective to assess the effect of D-CLIP on HRQOL, data from D-CLIP participants ($N = 578$) were used; only 3 participants had missing SF-12 data either at baseline. Sample selection is summarized in Fig. 1. Reporting of this study and the original randomized controlled trial was guided by the Consolidated Standards of Reporting Trials extension to randomized controlled trials (Supplementary Table S1) [33].

Measures

Descriptive characteristics of the recruited sample were collected by using administrator-assisted surveys. Body mass index (BMI, kg/m^2) was classified into four weight status categories according to the World Health Organization Asian-specific cut points [32]: BMI < 18.5 kg/m^2 for underweight, BMI 18.5 to < 23 kg/m^2 for normal weight, BMI 23 to < 27.5 kg/m^2 for overweight, BMI ≧ 27.5 kg/m^2 for obese. Mean minutes of exercise at baseline were estimated with a 2-item questionnaire and categorized by the physical activity goal (≧ 150 min/week) of the intervention [31]. Duration of study participation was calculated in the number of months between baseline and the last study visit with the participant's follow-up data (distribution shown in Supplementary Figure S1).

Short-Form 12 version 2 (SF-12) of 1-week recall was translated into Tamil, back-translated to English, and administered in either English or Tamil based on the participant's preference. The survey consists of twelve questions answered in 3- or 5-level Likert scales as originally developed by QualityMetric Inc. [13]. The questions were then grouped into eight subdomains: physical functioning (PF), role limitations due to physical health (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role limitations due to emotional health (RE), and mental health (MH). Z-scores of the subdomains were then aggregated into component scores, Physical Component Score (PCS) and Mental Component Score (MCS), as weighted sums (Supplementary Equation S1), which were standardized as mean of 50 and standard deviation of 10 for interpretability and cross-cultural comparisons.

Statistical analysis

Sample characteristics at clinical screening and baseline were compared by using one-way analysis of variance (ANOVA) or chi-squared test. We assessed the reliability of SF-12 and reported Cronbach's alpha (0.70 or greater was considered satisfactory internal reliability) [34] and Spearman correlations of the questionnaire items in relation to the subdomains and the two component scores (PCS, MCS) by using SAS software, Version 9.4. To confirm the latent constructs of PCS and MCS for the Indian context, a two-factor CFA was used, with subdomains PF, RP, BP, and GH as input variables for factor PCS, and variables VT, SF, RE, and MH for factor MCS. Using the lavaan package in R, the following fit indices

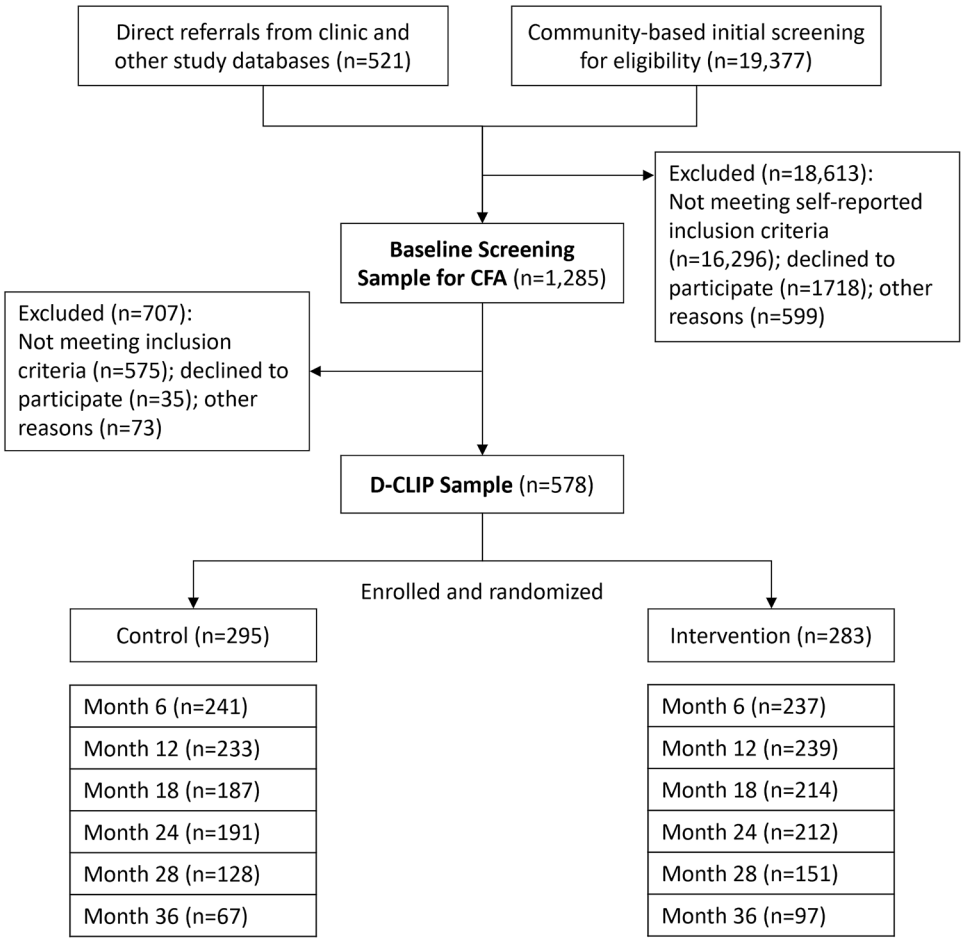

Fig. 1 Flow chart of the study

were estimated to determine the model fit of this factor structure [35]: chi-squared (χ^2, considered good fit if $p > 0.05$), square error of approximation (RMSEA, < 0.08), comparative fit index (CFI, > 0.90), Tucker-Lewis index (TLI, > 0.95), and standardized root mean squared residual (SRMR, < 0.08). Modification fit indices were also estimated to re-specify the model with appropriate error covariances that were theoretically justified and significantly improved overall model fit of the two-factor structure.

The subdomains were summarized into PCS and MCS by using a weighted sum of the corresponding subdomain z-scores with standardized factor loading estimates from CFA as their weights [13, 36]. The component scores were then standardized to a mean of 50 with standard deviation of 10 for each study visit. Mean PCS and MCS scores of the D-CLIP screening sample were compared between baseline sociodemographic characteristics, key lifestyle risk factors of chronic diseases, and prevalence of diagnosed diabetes by using t-test and analysis of variance (ANOVA). Mean SF-12 scores for the screening sample were also computed by using the original US-based norm scoring algorithm as described by the SF-12 manual in order to compare using t-test against the mean scores of the US population from which the instrument was developed [13].

To assess the effect of D-CLIP's lifestyle intervention on HRQOL, multiple linear regression was conducted to estimate the mean difference in the change of SF-12 component scores between intervention and control groups after study participation. The outcome of the regression model was determined by the changes in PCS and MCS scores from baseline to the final visit of each participant, with the random assignment to either control or intervention group being the main predictor. Since the time between baseline and the final study visit varied among the participants due to rolling enrollment, the model was adjusted for the number of months participating in the study. Another covariate was included to account for participant's T2DM diagnosis during the study period, as the development of diabetes during the study is associated with changes in the SF-12 scores, duration of study participation, and intervention exposure. Linear regression was conducted by using PROC GLM procedure in SAS 9.4; p values under 0.05 were considered statistically significant for all analyses.

Ethical approval

Emory University Institutional Review Board (IRB-00016503) and the Madras Diabetes Research Foundation Ethics Committee approved the study. All participants gave written consent before screening and randomization. Analysis and reporting are based on the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) reporting guidelines [37, 38]. No significant harm or adverse effects were reported during the trial period.

Results

Sample characteristics

Baseline characteristics of the study samples were previously published [31, 39]. To summarize, for the screening sample of 1285 adults, the mean age was 44.2 years and 63.7% were male. Mean BMI of the screening sample was 27.4 kg/m^2, with over 90% with overweight/obesity, which were consistent with the inclusion criteria of the initial community screening. For key lifestyle risk factors, 19.2% of the screened participants ever smoked and 78.3% reported exercising 150 min per week or less. Similarly, D-CLIP participants ($N=576$) were on average 44.4 years of age, 63.2% male, 94.3% with overweight or obesity (mean BMI 27.9 kg/m^2), 19.8% ever smoked, and 75.2% reported to exercise less than 150 min/week. Participant characteristics did not differ significantly by trial arm (Supplementary Table S2).

Construct validity and reliability of SF-12

When applying the standard loading values derived from the 1998 US sample population used for the development of SF-12, the mean of all subdomains and PCS were significantly different between the Indian and the US standard samples (Fig. 2). Mean scores of seven subdomains were significantly lower in the India sample than in the US standard; vitality was the only subdomain where the D-CLIP participants scored significantly higher. While the mean of MCS component scores were similar between the two populations, the mean PCS score was significantly higher for the US than the Indian sample.

All SF-12 items showed significant correlations to all corresponding subdomains and component scores, while showing moderate to low correlations (<|0.60|) with subdomains other than their own (Supplementary Table S3). All items had moderate to high correlations, with absolute values ranging from 0.51 to 0.78, to the same component scores as originally developed; PF items had a relatively low correlation of 0.51 to their appropriate component score, PCS. Internal consistency of the scale items to their respective component score was high with Cronbach's alpha of 0.80 overall, and 0.64 and 0.71 for PCS and MCS items, respectively.

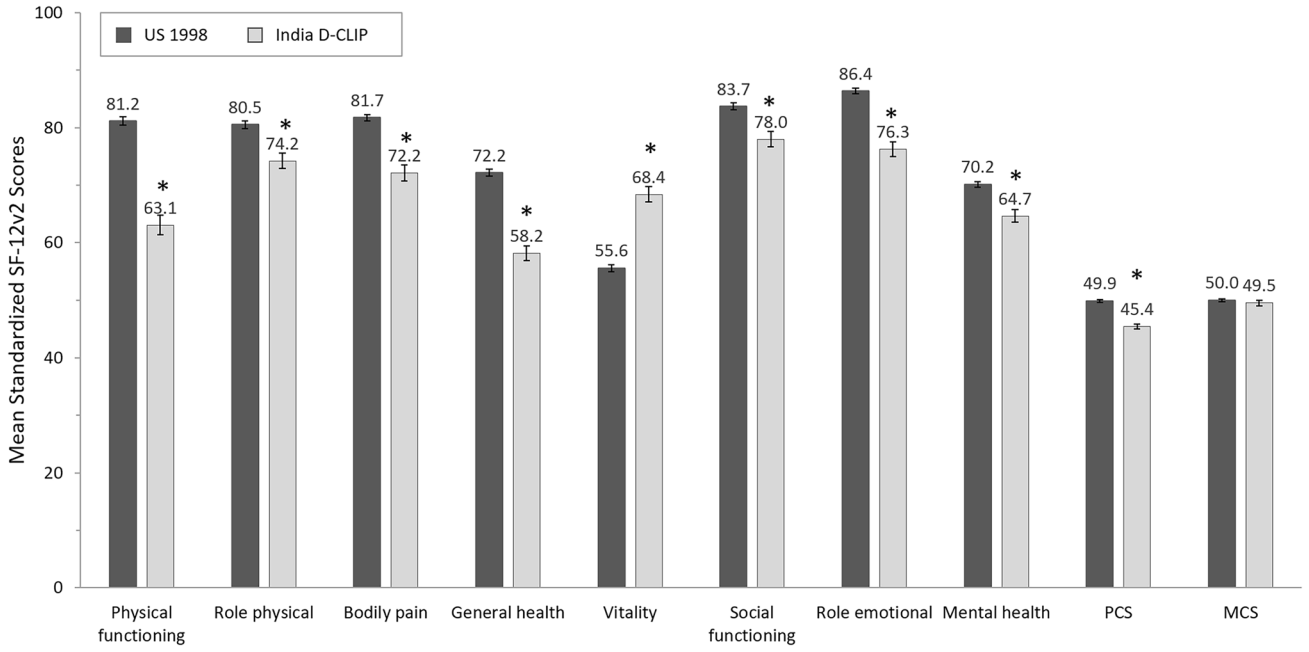

Fig. 2 Mean age- and sex-standardized scores for the 8 domains of SF-12 of the United States (1998 General US population) and the DCLIP Indian sample. Error bars indicate 95% confidence intervals. *$p<.001$

The CFA results showed that the SF-12 survey construct had an overall good fit for the screening sample: $\chi^2 = 110.832$ ($p < 0.0001$), RMSEA = 0.066, CFI = 0.962, TLI = 0.937, SRMR = 0.036. The two pairs of error covariances, RP-RE and VT-MH, were added to the model to optimize model fit since both pairs had two of the highest modification indices when added to the model (Supplementary Table S4). The languages used to frame the questions of RP and RE are also similar to each other, and VT and MH are within the same MCS factor (Supplementary Figure S2). The estimated factor loadings of all subdomains were also meaningfully high (|loading| ≥ 0.30) for their respective latent factors (Table 1). For PCS, PF was loaded the lowest of 0.398 while loadings of other subdomains, RP, BP, and GH, were 0.667, 0.666, and 0.516, respectively. The subdomains of MCS had a similar pattern where one subdomain (VT) had a substantially lower loading value of 0.478 while other domains had higher and similar loading estimations: SF with 0.617, RE 0.713, and MH 0.612.

Figure 3a and b shows the mean PCS and MCS scores by socio-demographic characteristics. Compared to men, women showed significantly lower scores in both PCS (mean ± 95% CI 46.7 ± 0.94 vs. 51.3 ± 0.64, $p < 0.0001$) and MCS (48.2 ± 0.97 vs. 51.0 ± 0.65, $p < 0.001$). While the mean PCS scores were significantly lower in older age groups compared to participants 30 years or younger ($p = 0.03$), mean MCS did not differ significantly ($p = 0.22$). Both mean PCS and MCS scores differed significantly by educational attainment: those with no formal education ($n = 25$) had the lowest mean PCS and MCS scores of 45.0 ± 4.23 and 46.3 ± 4.04, respectively, while those with high school or higher education had the highest (51.8 ± 0.67 and 51.0 ± 0.65, respectively). Mean PCS and MCS also differed significantly by weight status, with lower mean scores for individuals with obesity (48.6 ± 0.84 and 49.2 ± 0.88, respectively). Individuals who reported at least 150 weekly exercise minutes scored higher on both PCS and MCS (53.0 ± 1.10 and 51.2 ± 1.06, respectively) compared to those who did not (49.2 ± 0.62 and 49.4 ± 0.63, respectively). PCS and MCS scores did not differ significantly by smoking status or fasting plasma glucose diabetes category.

Table 1 Standardized factor loading estimates for the Short Form 12 (SF-12) health survey derived from the US standard population (principal component analysis performed by Ware et al. [13]) and D-CLIP screening sample of Indian adults (two-factor confirmatory factor analysis)

SF-12 domain	US standard		India D-CLIP	
	PCS-12	MCS-12	PCS-12	MCS-12
Physical functioning (PF)	0.424	−0.230	0.398	0
Role physical (RP)	0.351	−0.123	0.667	0
Bodily pain (BP)	0.318	−0.097	0.666	0
General health (GH)	0.250	−0.016	0.516	0
Vitality (VT)	0.029	0.235	0	0.478
Social functioning (SF)	−0.008	0.269	0	0.617
Role emotional (RE)	−0.192	0.434	0	0.713
Mental health (MH)	−0.221	0.486	0	0.612

Intervention effect of D-CLIP on HRQOL

Table 2 describes the adjusted mean difference between the intervention and control group regarding the longitudinal change in PCS and MCS scores and the raw scores of their subdomains. Results from the multiple linear regression indicated that compared to the control, the intervention group experienced a greater mean change in PCS scores after study participation. On average, the PCS scores of the intervention group resulted in an increase of 0.25 from baseline to the final study visit, while the control group experienced a mean decrease of 1.51 in their PCS scores (mean difference [95% CI] 1.63 [0.03, 3.24], $p = 0.046$). For MCS, no significant group difference was observed for the mean pre/post-study change (1.00 [0.67, 2.67], $p = 0.242$). Conclusions remained consistent across various standardization methods for computing PCS and MCS (Supplementary Table S5).

Discussion

This study is the first to report the construct validity of the SF-12 survey structure for Indian adults, and its application yielded important discussions on the effect of a lifestyle modification intervention on HRQOL of Indian adults with prediabetes. When using the US standard weights [13], the mean of all subdomains except vitality was overestimated, and subsequently the mean PCS score (Fig. 2). The discrepancy found between using the standard versus the Indian scoring weights highlighted the need to assess the suitability of the measurement structure and determine new scoring weights to compute the aggregated scores, PCS and MCS, that are specific to the context of our study population.

The subdomains of the Indian/Tamil version of the SF-12 used in our study demonstrated suitable internal reliability (Cronbach's alpha: 0.64 for PCS and 0.71 for MCS), suggesting strong consistency in measurement. This was consistent with Wind et al.'s findings (0.80 and 0.68, respectively) in a Northern Indian population exposed to recurring natural disasters [21]. The confirmatory factor analysis also demonstrated that all eight subdomains loaded sufficiently (≥ 0.3) onto their respective component scores with good model fit, suggesting that the SF-12 indicators were appropriate for the two-factor structure (i.e., PCS and MCS, each explained by four subdomains) for Indian adults. The CFA loadings also differed from those of the US sample used for survey

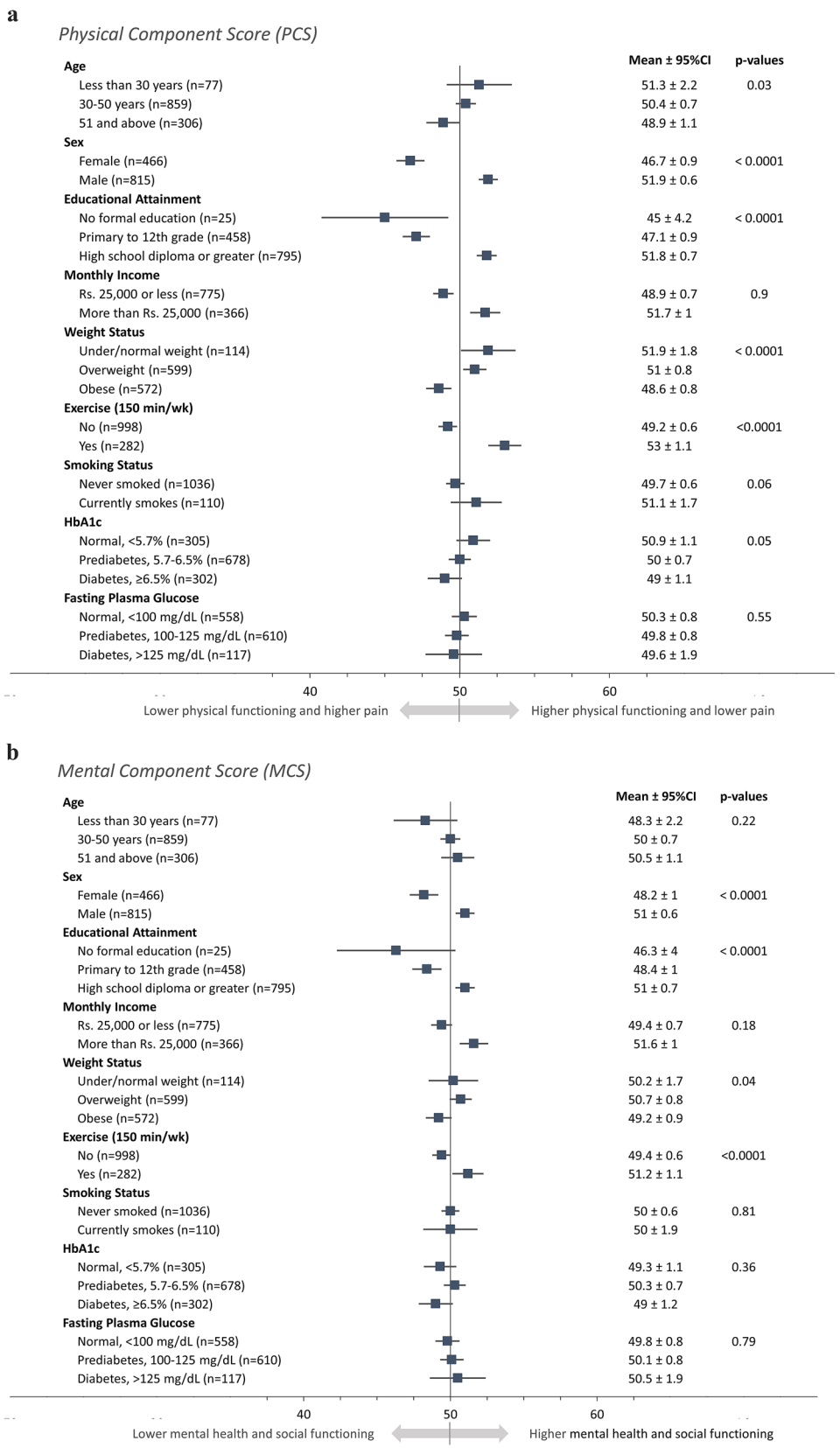

Fig. 3 Forest plot comparing mean SF-12 component scores, **a** physical component score and **b** mental component score, of the D-CLIP screening sample by sociodemographic characteristics

Table 2 Linear regression of the effect of D-CLIP intervention on the component scores of SF-12, physical component score (PCS), and mental component score (MCS)

	Unadjusted model			Adjusted model		
	Mean diff	±SE	*p* Value	Mean diff	±SE	*p* Value
PCS	1.76	±0.81	0.03*	1.63	±0.82	0.05*
Physical functioning (PF)	4.96	±2.96	0.09	4.40	±3.00	0.14
Role physical (RP)	0.81	±2.30	0.73	0.63	±2.32	0.79
Bodily pain (BP)	4.51	±2.38	0.06	4.21	±2.41	0.08
General health (GH)	2.97	±1.79	0.10	2.09	±1.80	0.25
MCS	1.00	±0.84	0.24	1.00	±0.85	0.24
Vitality (VT)	2.81	±2.20	0.20	2.00	±2.22	0.37
Social functioning (SF)	1.71	±2.49	0.49	1.47	±2.52	0.56
Role emotional (RE)	1.89	±2.12	0.37	1.81	±2.15	0.40
Mental health (MH)	3.28	±1.82	0.07	3.07	±1.84	0.10

Each parameter estimation refers to a mean difference of intervention to control groups of longitudinal change in SF-12 score of each participant. The change in SF-12 scores was measured by the difference in scores between baseline and participant's final follow-up visit before dropout or termination of the study

*$p < 0.05$

development (Table 1): the loading values were the lowest for GH, VT, and SF in the US standard sample, whereas in the D-CLIP sample, PF and VT loaded the least to their respective factors. Therefore, our study suggests using CFA loadings specific to the study sample, as opposed to using the standard weights [13], when computing PCS and MCS for the Indian context.

The mean PCS and MCS scores of the screening sample showed similar trends across multiple population characteristics: the scores did not differ significantly by income category or smoking status, while they were significantly higher among women versus men, and among those with high school diploma or higher. Past studies of other population contexts also found that men and adults with higher education scored higher in PCS and MCS, which may highlight health disparities among women and those without higher education [40–43]. Both scores were significantly lower for those reporting two of the major lifestyle risk factors for T2DM (i.e., overweight/obesity and low physical activity), although PCS and MCS did not differ by fasting blood glucose levels. Similar to our results, existing literature suggests that risk factors of metabolic disorders are negatively associated with the HRQOL [44–46].

Following our previous finding that the D-CLIP intervention significantly reduced diabetes incidence [31], we hypothesized its potential positive impact on the HRQOL. Compared to the control, the intervention group indeed resulted in a significantly greater increase in their PCS scores from baseline to study completion ($\beta = 1.63 \pm 0.82$, $p = 0.046$) and a greater improvement in MCS albeit no statistical significance ($\beta = 1.00 \pm 0.85$, $p = 0.24$). This finding may result from D-CLIP's greater focus on dietary changes and physical activity, which are integral to physical functioning. While physical and emotional wellbeing are closely interrelated, we were limited to assess whether the non-significant intervention effect on MCS is attributable to the behavioral focus of the intervention, the cultural perceptions toward negative responses to survey questions, or limited mental health awareness.

It is also essential to consider the clinical significance of the intervention effect. While the intervention group exhibited greater improvement in their physical functioning, the group difference on a scale of 0–100 may lack clinical relevance to prediabetes and diabetes. The KDPP study that implemented a similar lifestyle intervention in Indian adults with prediabetes found similar parameter estimates for the component scores (for their study, neither score showed statistically significance) [28]. Meanwhile, studies evaluating the effect of DPP in European countries and in the USA reported significant and clinically relevant improvements in HRQOL (measured by EuroQol, SF-36, and 15D-HRQOL) following program participation [47–51]. These interventions showed greater improvements in HRQOL compared to D-CLIP, possibly due to the different cultural context affecting how participants perceive and rate their physical and mental functioning levels. The mode of delivery was also different between studies; for instance, the Dwibedi study used self-paced digital education, whereas D-CLIP provided in-person group-based education. While we suspect heterogeneity of impact by intervention strategies [52]; further research is needed to understand which mode of delivery achieves greater and clinically meaningful impact on HRQOL.

Another plausible explanation of the similar mean change of the SF-12 scores between the intervention and control groups of the D-CLIP trial is due to the inherent design of the SF-12 to primarily measure a generalized physical and mental functioning rather than disease-specific symptoms.

Past studies mostly used disease-specific surveys to measure the HRQOL in relation to diabetes [44]. In some studies, generic measures such as the SF-12 and WHOQOL were also used to measure the HRQOL of patients with T2DM and found that diabetes-related risk factors, symptoms, and co-morbidities were correlated with the survey scores [51, 53–55]. While past studies found inverse correlation between the SF-12 and diabetes and its risk factors, more evidence is needed on whether the SF-12 measurements are sensitive enough to capture the changes of HRQOL among those who participate in DPPs.

In summary, this study contributes valuable initial evidence regarding the impact of lifestyle modifications on the HRQOL of Indian adults with prediabetes. Our findings emphasize the need for further intervention studies investigating various lifestyle modification tools that may enhance both the physical and mental components of HRQOL. This study was also the first to confirm that the two-factor structure of the SF-12 (i.e., PCS and MCS each explained by four subdomains) was suitable to measure the HRQOL of adults residing in Chennai, India. By using CFA, we have provided robust validation of the hypothesized latent structure of the instrument. The large sample size and minimal missingness (< 0.5%) of the SF-12 data also strengthened our analyses. However, the study was not without limitations. Despite the large sample size, the participants were recruited based on pre-determined eligibility from a single city in south India, limiting the generalizability of the SF-12 construct validity to the broader Indian population. Also, as D-CLIP was a prevention study, our study solely assessed the SF-12 changes before a T2DM diagnosis, unable to capture any post-diagnosis impact. Therefore, further study is needed to assess its construct validity in a broader Asian Indian population, and future research should include longitudinal HRQOL measures, spanning both pre- and post-T2DM diagnosis.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11136-024-03648-6.

Acknowledgements This project is supported by a BRiDGES grant from the International Diabetes Federation (LT07-115). BRiDGES, an International Diabetes Federation project, is supported by an educational grant from Lilly Diabetes. Additional support was provided by the Global Health Institute at Emory University. Additional support was provided through P30DK111024 from NIH/NIDDK. We also express our appreciation to Ms. Sharfaa Junaid who supported the preliminary analysis of this project.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Harish Ranjani, Jayalakshmi Sivaram, Mary Beth Weber, and Euisun Pyo. The first draft of the manuscript was written by Euisun Pyo, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This project is supported by a BRiDGES grant from the International Diabetes Federation (LT07-115). BRiDGES, an International Diabetes Federation project, is supported by an educational grant from Lilly Diabetes. Additional support was provided by the Global Health Institute at Emory University. Additional support was provided through P30DK111024 from NIH/NIDDK.

Data availability N/A, the study data can only be accessed by investigators registered in the IRB and their groups.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval Emory University Institutional Review Board (IRB-00016503) and the Madras Diabetes Research Foundation Ethics Committee approved the study.

Consent to participate All participants gave written informed consent before screening and randomization.

Consent to publish All participants gave written informed consent for the publication of their data for research purposes.

References

1. IDF. (2022). *Diabetes in South-East Asia—2021*. In *Factsheets*. International Diabetes Federation (IDF) Diabetes Atlas.
2. Anjana, R. M., et al. (2023). Metabolic non-communicable disease health report of India: The ICMR-INDIAB national cross-sectional study (ICMR-INDIAB-17). *The Lancet Diabetes and Endocrinology, 11*(7), 474–489.
3. Anjana, R. M., et al. (2017). Prevalence of diabetes and prediabetes in 15 states of India: Results from the ICMR-INDIAB population-based cross-sectional study. *The Lancet Diabetes and Endocrinology, 5*(8), 585–596.
4. Hills, A. P., et al. (2018). Epidemiology and determinants of type 2 diabetes in south Asia. *The Lancet Diabetes and Endocrinology, 6*(12), 966–978.
5. Haraldstad, K., et al. (2019). A systematic review of quality of life research in medicine and health sciences. *Quality of Life Research, 28*(10), 2641–2650.
6. Costa, D. S. J., et al. (2021). How is quality of life defined and assessed in published research? *Quality of Life Research, 30*(8), 2109–2121.
7. Farquhar, M. (1995). Definitions of quality of life: A taxonomy. *Journal of advanced nursing, 22*(3), 502–508.
8. Cieza, A., & Stucki, G. (2005). Content comparison of health-related quality of life (HRQOL) instruments based on the international classification of functioning, disability and health (ICF). *Quality of Life Research, 14*(5), 1225–1237.
9. Hand, C. (2016). Measuring health-related quality of life in adults with chronic conditions in primary care settings: Critical review of concepts and 3 tools. *Canadian Family Physician, 62*(7), e375–e383.
10. Karimi, M., & Brazier, J. (2016). Health, health-related quality of life, and quality of life: What is the difference? *PharmacoEconomics, 34*(7), 645–649.
11. The WHOQOL Group. (1998). The World Health Organization quality of life assessment (WHOQOL): Development and general psychometric properties. *Social science & medicine, 46*(12), 1569–1585.

12. The EuroQol Group. (1990). EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy, 16*(3), 199–208.
13. Ware, J., et al. (2002). *How to score version 2 of the SF-12 Health Survey (with a supplement documenting version 1).* QualityMetric Inc.
14. Ware, J., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220–233.
15. Jenkinson, C., et al. (1997). A shorter form health survey: Can the SF-12 replicate results from the SF-36 in longitudinal studies? *Journal of Public Health Medicine, 19*(2), 179–186.
16. Al Omari, O., et al. (2019). A review of the Short Form Health Survey-Version 2. *Journal of nursing measurement, 27*(1), 77–86.
17. Gandek, B., et al. (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*(11), 1171–1178.
18. Govindaraju, S. K., et al. (2021). Multiple high-energy open injuries in the same limb: Comparison of outcomes between salvage and amputation groups. *Indian Journal of Orthopaedics, 55*(3), 734–740.
19. Rangaswamy, U. K., & Ganachari, M. S. *Community pharmacist and clinical pharmacist interventions using MMAS-8, FACIT, SF-12 among pulmonary TB patients receiving anti-TB drugs from Urban Community Pharmacies, South India.*
20. Wiedemann, A. A., et al. (2018). Food addiction among men and women in India. *European Eating Disorders Review, 26*(6), 597–604.
21. Wind, T. R., et al. (2013). The impact of recurrent disasters on mental health: A study on seasonal floods in Northern India. *Prehospital and Disaster Medicine, 28*(3), 279–285.
22. Pal, S., et al. (2014). Long-term (1-year) safety and efficacy of a single 6-mL injection of hylan G-F 20 in Indian patients with symptomatic knee osteoarthritis. *The Open Rheumatology Journal, 8*, 54–68.
23. Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE: Life Sciences Education, 18*(1), rm1.
24. DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment, 23*, 225–241.
25. McArdle, J. J. (1996). Current directions in structural factor analysis. *Current Directions in Psychological Science, 5*(1), 11–18.
26. Florez, H., et al. (2012). Impact of lifestyle intervention and metformin on health-related quality of life: The diabetes prevention program randomized trial. *Journal of General Internal Medicine, 27*(12), 1594–1601.
27. Peña, A., et al. (2022). Effects of a diabetes prevention program on type 2 diabetes risk factors and quality of life among Latino Youths with prediabetes: A randomized clinical trial. *JAMA Network Open, 5*(9), e2231196–e2231196.
28. Thankappan, K. R., et al. (2018). A peer-support lifestyle intervention for preventing type 2 diabetes in India: A cluster-randomized controlled trial of the Kerala Diabetes Prevention Program. *PLoS Medicine, 15*(6), e1002575.
29. Cunillera, O., et al. (2010). Discriminative capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey. *Quality of Life Research, 19*(6), 853–864.
30. Weber, M. B., et al. (2012). A model of translational research for diabetes prevention in low and middle-income countries: The Diabetes Community Lifestyle Improvement Program (D-CLIP) trial. *Primary Care Diabetes, 6*(1), 3–9.
31. Weber, M. B., et al. (2016). The stepwise approach to diabetes prevention: Results from the D-CLIP randomized controlled trial. *Diabetes Care, 39*(10), 1760–1767.
32. WHO Expert Consultation. (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet, 363*(9403), 157–163.
33. Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ, 340*, c332.
34. Nunnally, J., & Bernstein, I. (1994). *Psychometric theory*, 3rd edn. MacGraw-Hill
35. Kline, R. B. (2023). Chapter 10: Model testing and indexing. In *Principles and practice of structural equation modeling*. Guilford publications.
36. McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods, 52*(6), 2287–2305.
37. Husereau, D., et al. (2022). Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: Updated reporting guidance for health economic evaluations. *BMC Medicine, 20*(1), 23.
38. Husereau, D., et al. (2013). Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *Value Health, 16*(2), e1-5.
39. Weber, M. B., et al. (2015). Behavioral and psychosocial correlates of adiposity and healthy lifestyle in Asian Indians. *Primary Care Diabetes, 9*(6), 418–425.
40. Montazeri, A., et al. (2009). The Iranian version of 12-item Short Form Health Survey (SF-12): Factor structure, internal consistency and construct validity. *BMC Public Health, 9*, 341.
41. Haddad, C., et al. (2021). Validation of the Arabic version of the "12-item short-form health survey" (SF-12) in a sample of Lebanese adults. *Arch Public Health, 79*(1), 56.
42. Galenkamp, H., et al. (2018). Measurement invariance of the SF-12 among different demographic groups: The HELIUS study. *PLoS ONE, 13*(9), e0203483.
43. Johnson, J. A., & Coons, S. J. (1998). Comparison of the EQ-5D and SF-12 in an adult US sample. *Quality of Life Research, 7*(2), 155–166.
44. Oluchi, S. E., et al. (2021). Health related quality of life measurements for diabetes: A systematic review. *International Journal of Environmental Research and Public Health*, *18*(17).
45. Kolotkin, R. L., & Andersen, J. R. (2017). A systematic review of reviews: Exploring the relationship between obesity, weight loss and health-related quality of life. *Clinical Obesity, 7*(5), 273–289.
46. Kolotkin, R. L., Meter, K., & Williams, G. R. (2001). Quality of life and obesity. *Obesity Reviews, 2*(4), 219–229.
47. Böhme, P., et al. (2020). Effectiveness of a type 2 diabetes prevention program combining FINDRISC scoring and telephone-based coaching in the French population of bakery/pastry employees. *European Journal of Clinical Nutrition, 74*(3), 409–418.
48. Karamanakos, G., et al. (2019). The effectiveness of a community-based, type 2 diabetes prevention programme on health-related quality of life. The DE-PLAN study. *PLoS ONE, 14*(10), e0221467.
49. Eaglehouse, Y. L., et al. (2016). Impact of a community-based lifestyle intervention program on health-related quality of life. *Quality of Life Research, 25*(8), 1903–1912.
50. Schmiedel, K., et al. (2020). Quality of life and satisfaction during the diabetes prevention program GLICEMIA: A cluster-randomized, controlled trial. *Gesundheitswesen, 82*(11), 844–851.
51. Dwibedi, C., Abrahamsson, B., & Rosengren, A. H. (2022). Effect of digital lifestyle management on metabolic control and quality of life in patients with well-controlled type 2 diabetes. *Diabetes Ther, 13*(3), 423–439.

52. Bian, R. R., et al. (2017). The effect of technology-mediated diabetes prevention interventions on weight: A meta-analysis. *Journal of Medical Internet Research, 19*(3), e76.
53. Achuko, O., et al. (2016). Pathways between discrimination and quality of life in patients with type 2 diabetes. *Diabetes Technology & Therapeutics, 18*(3), 151–158.
54. Jankowska, A., & Golicki, D. (2022). Self-reported diabetes and quality of life: Findings from a general population survey with the Short Form-12 (SF-12) Health Survey. *Archives of Medical Science, 18*(5), 1157–1168.
55. Singh, K., et al. (2017). Health-related quality of life variations by sociodemographic factors and chronic conditions in three metropolitan cities of South Asia: The CARRS study. *British Medical Journal Open, 7*(10), e018424.