

Research Article

Development of a Composite Performance Score Tool for Diabetes Mellitus Management: Features and Framework

Argwings Kodhek¹ , Alice Achieng' Ojwang¹ , Francis Oguya¹ ,
Fredrick Otieno² , Isaac Okeyo^{1,*} 

¹Department of Health and Biomedical Sciences, Technical University of Kenya, Nairobi, Kenya

²Faculty of Health Sciences, University of Nairobi, Nairobi, Kenya

Abstract

Background: Type 2 diabetes mellitus (T2DM) remains a critical public health challenge in low- and middle-income countries (LMICs), where existing assessment tools often fail to integrate the full spectrum of clinical, behavioral, and contextual determinants of care. This study developed and validated the Management Performance Composite Scoring Tool (MPCST), a multidimensional instrument designed for primary healthcare (PHC) settings, grounded in the Chronic Care Model (CCM). **Methods:** An exploratory sequential mixed-methods design was conducted over four months across five public PHC facilities in Nairobi, Kenya. The qualitative phase involved six focus group discussions (n=52) and ten key informant interviews to identify domains relevant to T2DM management. In the quantitative phase, 181 adults with stable T2DM were selected via systematic random sampling. Data collection employed validated instruments alongside standardized clinical and biochemical assessments. A composite scoring algorithm was developed using statistical machine learning and deployed in a Streamlit-based application for real-time clinical utility. Psychometric validation included content validity (Lawshe's method), generalizability theory, and convergent and concurrent validity analysis. **Results:** Among 181 participants (63.5% female; mean age 55.9 years), glycemic control was poor (mean HbA1c: 9.39 mmol/L), and lifestyle performance was suboptimal. Tool development incorporated clinical (e.g., HbA1c, BMI), behavioral (diet, activity, distress), and contextual (social determinants, healthcare access) domains. Ordinal regression yielded excellent model fit ($pseudo R^2 = 0.949$, AIC = 69.821). Geometric mean aggregation improved behavioral sensitivity, while minimax strategies optimized clinical indicator selection. The MPCST demonstrated high reliability ($G = 0.8351$), strong convergent ($r = 0.74$) and concurrent ($r = 0.565$) validity. **Conclusion:** The MPCST is a rigorously validated, contextually relevant, and scalable tool for evaluating T2DM care quality in LMIC PHC systems. Its dual-format (manual and app-based) design supports routine use, longitudinal tracking, and integration into health systems, thereby enhancing clinical decision-making and patient-centered care.

Keywords

Type 2 Diabetes Mellitus, Primary Healthcare, Psychometrics, Chronic Care Model

*Corresponding author: isaacokeyo@tukenya.ac.ke (Isaac Okeyo)

Received: 12 October 2025; **Accepted:** 12 November 2025; **Published:** 9 December 2025



Copyright: © The Author(s), 2025. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

1. Introduction

T2DM represents a significant and growing global public health challenge, particularly in LMICs, where the burden continues to escalate alongside rapid urbanization and lifestyle changes [1]. T2DM is characterized by chronic hyperglycemia, leading to substantial morbidity and mortality due to its associated microvascular and macrovascular complications [2]. Effective management of T2DM requires comprehensive monitoring of clinical, behavioral, and contextual factors to optimize glycemic control and prevent adverse outcomes [3].

Existing research underscores that beyond biomedical parameters such as HbA1c and fasting glucose, behavioral components including medication adherence, diet, and physical activity play crucial roles in diabetes control [4]. Furthermore, social determinants of health (SDoH) such as healthcare access, socioeconomic status, and environmental factors have been increasingly recognized as independent predictors of diabetes outcomes [5, 6]. Despite this, many current clinical assessment tools predominantly emphasize isolated clinical indicators or knowledge-attitude-practice (KAP) metrics, often neglecting the multidimensional nature of diabetes management in real-world settings, especially within resource-limited PHC environments [7, 8].

Several gaps remain unaddressed. First, there is a lack of integrated, validated composite tools that simultaneously assess clinical status, patient behaviors, and contextual factors, key domains identified by theoretical models such as the CCM [9]. Second, many tools lack adaptability for use in diverse healthcare settings, limiting scalability and real-time clinical utility [10]. Additionally, most existing instruments have been developed in high-income contexts and lack rigorous psychometric validation using advanced methodologies, including generalizability theory, which can account for multiple sources of measurement error [11]. Finally, longitudinal assessment frameworks to monitor changes in diabetes management performance over time remain scarce, impeding dynamic quality improvement efforts [12].

Given these limitations, there is a compelling need for a robust, multidimensional performance scoring tool tailored to PHC settings in LMICs that incorporates clinical, behavioral, and contextual domains, validated with rigorous statistical methods and designed for both manual and digital administration. Such a tool would enable comprehensive, actionable evaluation of T2DM management, support clinical decision-making, facilitate resource allocation, and promote patient empowerment through personalized feedback.

The current study aims to develop and validate the MPCST, an innovative instrument integrating these critical domains into a practical framework for assessing T2DM management performance. Specifically, the study seeks to: identify key clinical, behavioral, and contextual indicators relevant to PHC settings; employ advanced analytic techniques to construct a composite scoring algorithm; validate the tool's psychometric

properties including content validity, generalizability, and convergent and concurrent validity; and demonstrate feasibility for implementation through both manual and app-based modalities. This work addresses a vital gap in diabetes care assessment and aligns with global health priorities to improve chronic disease management in resource-constrained environments [1, 13].

2. Methods

2.1. Study Design, Period and Setting

An exploratory sequential mixed-methods design was employed, with initial qualitative data informing the development of the subsequent quantitative phase. The study was conducted from June to September 2024 in selected public PHC facilities within Nairobi County, Kenya. These urban PHC facilities serve a socioeconomically diverse population and are representative of diabetes care delivery in high-volume, resource-constrained public sector settings.

2.2. Study Population and Eligibility Criteria

The study included adults (≥ 30 years) with confirmed, stable T2DM and healthcare providers involved in T2DM care, all recruited from selected PHC facilities in Nairobi County. Eligible patients had no recent hospitalizations or acute complications within the past three months, and were able to consent and participate in interviews and clinical assessments. Providers were eligible if directly engaged in T2DM management and consented to participate. Exclusion criteria included non-T2DM diagnoses, unstable medical or psychiatric conditions, or refusal to provide informed consent.

2.3. Sample Size and Recruitment

Five PHC facilities were selected via stratified purposive sampling within a multistage design. Six focus group discussions (FGDs) and ten key informant interviews (KII) were conducted until thematic saturation. FGD participants were randomly drawn from T2DM patient rosters, while key informants were purposively selected for their expertise. For the quantitative phase, 181 individuals with T2DM were recruited through systematic random sampling, proportionate to facility caseloads. This sample satisfies simulation-based recommendations for confirmatory factor analysis ($\geq 80\%$ power, $\alpha = 0.05$, CFI > 0.95) and meets the 10:1 participant-to-item ratio for psychometric validation [14].

2.4. Data Collection Instruments and Procedures

This mixed-methods study employed both qualitative and

quantitative approaches. Qualitative data were gathered through Focus Group Discussions (FGDs) and Key Informant Interviews (KIIs) using semi-structured guides to explore patient and healthcare provider perspectives. Quantitative data were collected via face-to-face interviews using digitized, validated instruments on the Open Data Kit (ODK) platform. Key assessments included: Dietary adherence (Perceived Dietary Adherence Questionnaire, a seven-day recall), physical activity (International Physical Activity Questionnaire-Short Form, quantifying MET-minutes/week over seven days), diabetes knowledge (structured Yes/No/Don't Know questions), attitudes (five-point Likert scale), and self-care practices (five-point frequency scale). Additional tools included the Short Form-8 Health Survey (SF-8) for general health, the Diabetes Quality of Care Summary Score, the Medication Adherence Questionnaire, and the adapted Modified Kuppuswami Scale for socioeconomic status.

Weight (kg) was measured using a *portable digital scale* with participants standing upright in light clothing; height (m) was recorded using a *wall-mounted stadiometer* on a firm, level surface; and blood pressure was measured using a *validated automated sphygmomanometer*, with all procedures adhering to *WHO-recommended protocols*. *Capillary blood glucose* was measured using *oxidase-based glucometer strips*, while *HbA1c*, *lipid profiles*, and *microalbuminuria* were analyzed in a centralized facility using the *Fuji NX500i dry chemistry analyzer*, recognized for its accuracy in chronic disease biomarker analysis. Biological samples were stored in *temperature-controlled containers* and transported within two hours of collection, maintaining a *cold chain between 2-8 °C* to ensure sample integrity. Upon completion of data collection, a *composite decision-support and predictive analytics tool* was developed using the *Python programming language*, incorporating *statistical machine learning algorithms* to compute individualized diabetes care quality scores. The tool was implemented via the *Streamlit framework*, allowing for *real-time user interaction*, *dynamic visualization*, and *integration of clinical algorithms*, thus enhancing its usability and relevance in resource-constrained primary care settings.

2.5. Data Quality Control

To ensure methodological rigor, all research assistants, holding relevant bachelor's degrees and prior field experience, completed a two-day training on the study protocol, research ethics, and Open Data Kit (ODK) use. Questionnaires were pretested on a demographically comparable population, and ambiguous items were revised based on pilot feedback. The Principal Investigator provided direct oversight for adherence to the study protocol, including recruitment, consent, questionnaire administration, and proper biological sample handling. In the laboratory, internal quality control employed Levey-Jennings charts for assay performance and Westgard rules (1-2s warning, 1-3s rejection) to guide interpretation, with failed analytical runs repeated to ensure biochemical result accuracy.

2.6. Statistical Analysis

Descriptive statistics (means, standard deviations, frequencies, proportions) summarized sociodemographic and clinical characteristics. Group differences were assessed using parametric tests (e.g., t-tests, Welch's ANOVA) with Games-Howell tests for post-hoc comparisons. Qualitative data underwent thematic analysis, involving transcription, coding, theme identification, review, and synthesis. Quantitative data were characterized using coefficient of variation, skewness, and Shapiro-Wilk tests. Given non-normal predictor distributions, geometric means were employed. A minimax strategy justified the non-compensatory nature of T2DM components for enhanced clinical relevance. Tool modeling utilized Ordinal logistic regression and Generalized Linear Models to examine associations, with a backward elimination procedure guided by p-value thresholds, Akaike Information Criterion (AIC), and Pseudo R-squared for optimization. Tool validation included: content validity using Lawshe's Content Validity Ratio (CVR) to assess expert agreement; reliability evaluated via the generalizability coefficient (G-coefficient), accounting for diverse measurement error sources; and construct validation through convergent and concurrent validity testing using Pearson correlation coefficient, assessing alignment with related theoretical constructs.

2.7. Ethical Approval and Consent to Participate

Ethical approval was obtained from Kenyatta National Hospital and the National Commission for Science, Technology, and Innovation (NACOSTI). Access to PHC facilities was authorized by the Nairobi County Government and facility managements. All T2DM-P were fully informed of the study's aims and procedures, and written informed consent was secured before enrollment.

3. Results

3.1. Sociodemographic and Clinical Characteristics

The mean age of participants in the Focus Group Discussions (FGDs) was 51.79 years (95% CI: 49.48–54.09), while Key Informant Interview (KII) participants had a mean age of 47.10 years (95% CI: 40.13–54.07), as detailed in [Tables 1 and 2](#). Gender distribution did not significantly differ within the FGDs ($\chi^2(1)=2.28$, $p=0.167$), indicating no overrepresentation of either sex. Similarly, gender distribution in the KIIs was not significantly different ($\chi^2(1)=0.40$, $p=0.527$). Furthermore, across the six FGD groups, there were no significant differences in either gender or age distribution (gender: $\chi^2(5)=1.43$, $p=0.921$; age: $F(5,47)=0.73$, $p=0.605$), with

Levene's test confirming homogeneity of variances for age. These findings collectively indicate overall homogeneity in

age and gender across all groups involved in the qualitative interviews.

Table 1. Sociodemographic Characteristics of T2DM-P in the FGDs.

Item	FGD1	FGD2	FGD3	FGD4	FGD5	FGD6	Total
Male	3	3	5	3	4	3	21
Female	4	6	4	5	6	6	31
Total	7	9	9	8	10	9	52
Mean age (SD)	47.3 (7.158)	56 (13.693)	54.4 (11.823)	52.1 (12.005)	54.5 (11.712)	49.1 (11.163)	52.5 (11.415)
Female were not significantly more than males, $\chi^2 (1) = 2.283$, $p = 0.167$							
Grand Mean for age is 51.79 (CI 49.48-54.09)							

Table 2. Sociodemographic Characteristics of Participants in the KII.

Expert	Doctors	Diabetic Nurses / nutritionist	Community Health Worker	Expert Patient	Community Organization Manager	Based	Total
Number	4	3	1	1	1		
						Male	4
						Female	6
Mean age 47.1 (SD 9.746); Minimum Age: 33 years; Maximum Age: 59 years							

As shown in Table 3, female participants significantly outnumbered males ($\chi^2(1) = 13.27$, $p < 0.001$). Among individuals with T2DM, a significantly higher proportion were married ($\chi^2(1) = 10.22$, $p = 0.001$) and employed ($\chi^2(1) = 6.02$, $p = 0.014$). The overall mean HbA1c was elevated at 9.39 mmol/L (SD = 2.43), reflecting poor glycemic control. Welch's ANOVA indicated a significant difference in HbA1c across categorized ranges: (6.0-7.0, 7.1-7.6, 7.7-8.2, 8.3-8.8, and >8.8 mmol/L), ($F(4,176) = 113.83$, $p < 0.001$), with post hoc Games-Howell tests confirming significant differences

between all category pairs. Mean lifestyle performance scores were average, with dietary habit scores of 46.43% (SD = 9.508%) and physical activity scores of 470.64 MET/week (SD = 9.508). Mean LDL cholesterol was borderline high at 2.85 mmol/L (SD = 0.74), and mean BMI indicated overweight status at 28.66 kg/m² (SD = 5.87). Overall, the quantitative phase identified a predominance of females, suboptimal glycemic control, and inadequate lifestyle factor performance among participants.

Key: c Chi Square test; a Analysis of Variance

Table 3. Sociodemographic and Clinical Characteristics of T2DM-P in Quantitative Study.

Males (%)	66 (36.5%)	Mean Age in years	55.92	Married (%)	120 (66.3)	Employed (%)	107 (59.1)
Females (%)	115 (63.5%)	SD	12.041	Unmarried (%)	61 (33.7)	Unemployed (%)	74 (40.9)
p value	0.000 ^c				0.000 ^c		0.014 ^c
HbA1c categories mmol/lit	6-7	7.1-7.6	7.7-8.2	8.3-8.8	>8.8	Overall Mean 9.39 (SD 2.430). Mean HbA1c between categories was significantly different $p = < 0.001^a$	
Number	27	24	26	17	87		
Mean (SD)	6.49 (1.906)	7.44 (0.212)	7.94 (0.202)	8.60 (0.296)	11.416 (0.465)		
						$\eta^2 = 0.697$ (0.619-0.744)	

Variable	Systolic Blood Pressure (mmHg)	BMI	LDL (mmol/lit)	Dietary Scores (%)	Knowledge Scores (%)	Physical Activity (met/wk)	Distress scores (%)
Mean	137.41	28.655	2.849	46.43	58.90	470.64	41.14
SD	20.872	5.868	0.744	9.508	15.33	178.126	11.867

3.2. Item Selection Process

An exploratory mixed-methods approach informed the iterative selection of items for the MPCST, balancing contextual relevance with statistical rigor.

3.3. Qualitative Findings: Determinants of Diabetes Care

Item generation was initially guided by thematic analysis of FGDs with T2DM patients, who identified key determinants of care: biomedical parameters (e.g., blood sugar, blood pressure, cholesterol, weight), comorbidities, lifestyle behaviors (carbohydrate and water intake, physical activity), psychosocial factors (stress, diabetes knowledge), and structural barriers (e.g., healthcare access, insurance). These insights were triangulated and refined through KIIs with healthcare providers and policymakers, who emphasized three overarching domains: clinical outcomes (biomarkers, comorbidities), behavioral and psychosocial factors, and contextual determinants (health system capacity and SDoH).

3.4. Considerations in T2DM Item Combinations

Self-Care Behavioral Measures and Their Multiplicative Interactions

To derive an average of the behavioral measures, a logistic regression model was conducted with five HbA1c categories (very high, high, elevated, normal, and optimal) as the outcome variable. The behavioral predictors and their interaction terms were included as independent variables. The overall model was statistically significant, (χ^2 (76) = 102.404, p = 0.023).

The model took the following form:

$$\log\left(\frac{P(Y=j)}{P(Y=\text{reference})}\right) = \beta_{0j} + \beta_{1j}D + \beta_{2j}P + \beta_{3j}S + \beta_{4j}(D \times P \times S) + \varepsilon$$

Where;

Y = HbA1C categories

P(Y=j) is the probability of being in category j

β_{0j} is the intercept

D, P and S are dietary habits, Physical activity and diabetes distress respectively

β are the coefficients of regression

β_{4j} is the multiplicative interaction term for category j.

The interaction term (D×P×S) was significant (p = 0.034; OR = 1.725, 95% CI: 1.276-2.330), indicating a multiplicative effect of diet, physical activity, and diabetes distress on glycemic control. The geometric mean of behavioral scores showed stronger sensitivity to variability (r = -0.672, p < 0.001) than the arithmetic mean (r = -0.172, p = 0.048), and was less affected by outliers (e.g., diet z = -3.247; activity z = 2.857) and skewness. Skewness values and Shapiro-Wilk test results further confirmed non-normality in key variables: physical activity (skew = 0.841, W = 0.946, p < 0.001), HbA1c (skew = 0.653, W = 0.928, p < 0.001), and LDL (skew = -0.316, W = 0.984, p < 0.001). Accordingly, behavioral factors were best modeled using the geometric mean, which preserved interaction effects and provided a more robust, clinically meaningful summary.

3.5. Minimax Strategy

Acknowledging the non-compensatory nature of diabetes care domains, a minimax strategy was used to retain the most deranged lipid and most burdensome comorbidity, capturing peak clinical risk. Triglycerides showed the greatest deviation among lipids (mean = 1.782, SD = 0.856; t (180) = 46.573, p < 0.001), while peripheral neuropathy was the most burdensome comorbidity (mean = 1.56 out of 5), supporting their inclusion in the final model.

3.6. Management Performance Composite Score (MPCS) and Related Indices

The MPCS is an annual index of scaled scores from 1 [poor] to 5 [excellent] of component raw values using geometric mean.

$$\text{Variable-MPCS} = (\text{HbA1c}_1 \times \text{HbA1c}_2 \times \text{HbA1c}_3)^{1/3}$$

$$\text{Annual-MPCS} = (\text{var. MPCS}_{\text{HbA1c}} \times \text{var. MPCS}_{\text{SBP}} \times \text{var. MPCS}_{\text{BMI}} \times \text{var. MPCS}_{\text{Lipid}} \times \text{var. MPCS}_{\text{Comorbidity}} \times \text{var. MPCS}_{\text{Behavioral}} \times \text{var. MPCS}_{\text{SDoH}} \times \text{var. MPCS}_{\text{Healthcare}})^{1/8}$$

The geometric mean was chosen to minimize the influence of extreme values and to respect the multiplicative and non-compensatory nature of the included domains.

Variable Concordance Value (VCV): the slope of the regression line of scaled scores over time; negative indicates

decline, zero indicates no change, and positive indicates improvement in care.

Aggregate Concordance Value (ACV): the slope of Annual-MPCS over time, reflecting long-term care performance trends.

3.7. Modelling of MPCS

The optimal model for Annual-MPCS was an ordinal regression ($\chi^2(14) = 272.10$, $p < 0.001$; pseudo $R^2 = 0.949$; AIC = 69.82), outperforming a generalized linear model ($F(9,171) = 47.55$, $p < 0.001$; $R^2 = 0.699$; AIC = 114.90). Inclusion of additional predictors reduced model performance: KAP (AIC = 341.92), disaggregated lipids (pseudo $R^2 = 0.873$; AIC = 120.53), individual comorbidities (pseudo $R^2 = 0.804$; AIC = 149.63), and arithmetic versus geometric means for self-care (pseudo $R^2 = 0.886$; AIC = 689.00). Systolic BP was retained over diastolic BP due to a higher pseudo $R^2(0.029$ vs. $0.024)$ and lower p -value (0.117 vs. 0.167). These findings underscore the utility of geometric aggregation and the minimax strategy in enhancing model parsimony and predictive accuracy.

3.8. Content Validity

To ensure content validity, 10 experts reviewed all items using Lawshe's CVR method. All exceeded the 0.80 threshold except "healthcare system" (CVR = 0.60), which was revised to "healthcare access and processes" to enhance conceptual relevance.

3.9. Reliability Analysis

Generalizability Theory (G-Theory) using linear mixed-effects models estimated variance from persons and measurement facets, yielding a generalizability coefficient (G) to assess the MPCST's reliability across heterogeneous components.

$$G \text{ Coefficient} = \frac{\sigma_p^2}{\sigma_p^2 + \frac{\sigma_{pi,e}^2}{n_i}}$$

This can also be rewritten as:

$$G = \frac{\text{Pearson Variance}}{\text{Pearson Variance} + \frac{\text{Residual Error Variance}}{\text{Number of Items}}}$$

Where:

- 1) σ_p^2 = variance between patients
- 2) $\sigma_{pi,e}^2$ = residual (person $\times$ item + error)
- 3) n_i = number of items

Components of the tool: Clinical outcomes, Behavioral measures and contextual factors.

Using SPSS Mixed Model Analysis,

$$\sigma_p^2 = 0.085739$$

$$\sigma_{pi,e}^2 = 0.118479$$

$$n_i = 7$$

$$G = 0.8351.$$

A G-coefficient of 0.8351 indicates strong reliability, confirming that MPCST composite scores are consistent across items and suitable for monitoring T2DM.

3.10. Validity Analysis

To evaluate convergent validity, the SF-8 Health Survey was administered as a reference measure of overall health status among individuals with T2DM. The mean SF-8 score was 16.20 (95% CI: 15.58-16.82; SD = 4.226), representing 40.5% of the maximum score. The mean MPCST score was 2.273 (95% CI: 2.206-2.340; SD = 0.457), corresponding to 45.46% of the total scale. A strong positive correlation was observed between the MPCST and SF-8 scores ($r=0.74$, $p < 0.001$), indicating good convergent validity.

Concurrent validity was assessed using the Q Score as a criterion measure of diabetes care quality. The mean Q Score was 25.05 (95% CI: 23.69-26.42; SD = 9.305), equivalent to 41.8% of the maximum score. The MPCST showed a moderate positive correlation with the Q Score ($r=0.565$, $p < 0.001$), supporting its concurrent validity. Additionally, a strong positive correlation was found between the MPCST and HbA1c levels ($r=0.752$, $p < 0.001$), demonstrating the tool's clinical relevance in reflecting glycemic control.

4. Discussion

4.1. Demographic and Clinical Profile of Study Participants

This exploratory mixed methods study provides important insights into the demographic, clinical, behavioral, and contextual dimensions of T2DM management among patients attending public PHC facilities. The mean participant age was 55.92 years (SD = 12.041), consistent with global evidence showing T2DM prevalence peaks in the mid-50s [15], and aligns with data from the Congo [16] where T2DM patient ages ranged from 40-65 years. A gender disparity was evident, with more female participants, reflecting trends in Sub-Saharan Africa where T2DM prevalence is higher in females (5.9%) than males (5.5%) [17]. This disparity is further supported by greater metabolic syndrome (a known risk factor) rates among women than men in sub-Saharan Africa [18]. This dimorphism has been attributed to hormonal changes, fat distribution, and psychosocial stressors increasing women's susceptibility to T2DM [19].

Behaviorally, participants showed moderate adherence to lifestyle recommendations, with dietary scores averaging 46.43% and physical activity measuring 470.64 MET-min/week—just below the recommended threshold of

≥ 500 MET-min/week. Such low levels of physical activity and suboptimal nutrition are common, particularly in African settings [20], and closely align with the observed poor glycemic control. Indeed, glycemic management was notably inadequate, with a mean HbA1c of 9.39 mmol/L and 48.07% of participants exceeding 8.8 mmol/L. These results reflect broader global trends, where 69.1% of individuals with T2DM in low- and middle-income countries experience poor glycemic control [6], consistent with regional data from Kenya [8].

4.2. Integrating Clinical, Behavioral, and Contextual Domains

Qualitative data revealed three foundational domains: biomedical, psychosocial, and structural factors, later refined through key informant interviews into clinical outcomes, behavioral/psychosocial metrics, and contextual factors such as healthcare access and SDoH. The inclusion of contextual variables is strongly supported by evidence demonstrating that social and environmental determinants independently influence diabetes control and complication risk beyond biomedical parameters [21, 22].

This multidimensional perspective aligns with the CCM, which emphasizes productive interactions between well-informed patients and proactive healthcare teams, supported by community resources and organized health systems [9]. The MPCST operationalizes the CCM by integrating these three domains into a measurable framework for chronic disease management in PHC settings.

4.3. Methodological Innovations and Model Performance

Quantitatively, ordinal logistic regression demonstrated excellent model fit ($\chi^2(14) = 272.096$, $p < 0.001$, $\text{pseudo } R^2 = 0.949$, $\text{AIC} = 69.821$), with the final model prioritizing adherence to care over traditional KAP metrics due to its stronger predictive validity [23]. The MPCST aggregates behavioral indicators using the geometric mean, a statistically robust approach that handles skewed data, controls outliers and rewards consistency, avoiding compensatory masking effects [24, 25]. This approach aligns with literature supporting multiplicative interactions among behavioral factors [26]. Clinical variable selection employed the minimax strategy, emphasizing vulnerability to enhance fairness and resilience in decision-making [27].

4.4. Validation and Practical Utility of the MPCST

The MPCST classifies T2DM management into five tiers—from Excellent [5] to Poor (1)—based on clinically relevant thresholds. Validation results showed strong psychometric properties: content validity ratio (CVR) of 0.8,

generalizability coefficient (G) of 0.8351, convergent validity (Pearson's $r = 0.74$, $p < 0.001$), and concurrent validity ($r = 0.565$, $p < 0.001$). Crucially, the tool incorporates contextual factors like healthcare access and SDoH, recognized independent predictors of poor glycemic control and complications [21].

The MPCST features the Annual-MPCS, enabling longitudinal evaluation of care quality. Regression of Annual-MPCS produces the Annual Concordance Value (ACV), indicating whether patient management is improving, stable, or deteriorating—a novel contribution aligned with prior evidence underscoring the value of trend analysis in chronic care monitoring [28, 29]. Together, these components establish a comprehensive framework for benchmarking and dynamic assessment of T2DM care.

4.5. Feasibility, Strengths, Limitations, and Future Directions

Designed for adaptability, the MPCST can be administered manually via standardized paper forms—ideal for resource-constrained settings—or through a smartphone/tablet app enabling automatic scoring and data visualization. This dual modality supports implementation across diverse PHC environments, including those with limited digital infrastructure. The tool emphasizes simplicity, rapid administration, and user-friendly interfaces to encourage uptake by clinicians, healthcare managers, and potentially patients. Localization through language adaptation and integration with existing electronic medical records is a promising pathway to scale.

A major strength of this study is its methodological pluralism, integrating qualitative insights with robust quantitative validation to ensure both contextual relevance and analytical rigor. The use of Generalizability Theory in place of conventional reliability metrics represents a methodological advance, enabling the decomposition of multiple sources of measurement error [30]. The adoption of a Minimax approach introduces a novel optimization strategy, enhancing statistical parsimony in combining clinical and behavioral variables. Additionally, the use of the geometric mean—rather than the arithmetic mean—strengthens the composite score by reducing the masking of poor performance and mitigating the impact of outliers. This is especially relevant given the high interindividual variability in T2DM care metrics, often driven by disparities in glycemic control, adherence, and healthcare access [31, 32]. By incorporating adherence and contextual factors into the Annual-MPCS (a composite index used as an evaluation metric) and ACV - Trajectory based quality index, the MPCST captures the multifactorial nature of diabetes care and is well-suited for both cross-sectional evaluation and longitudinal monitoring.

Limitations include exclusive focus on public PHC facilities in Nairobi, which serves ~63% of Kenyans relying on

public care [33], potentially missing perspectives from private care or self-managed patients [34]. Nairobi's relatively well-resourced urban context [5] may limit generalizability to rural or less-resourced settings. Additionally, validation was conducted on a single sample without external replication, raising concerns of overfitting. White coat adherence—temporary behavioral improvements before clinic visits—may have biased clinical metrics, partially addressed by KAP inclusion but not eliminated.

Future research should prioritize external validation across diverse geographic and healthcare contexts, including rural and private sectors, to refine thresholds and assess generalizability. Implementation studies are needed to evaluate real-world usability, acceptability, and clinical impact when integrated into PHC workflows and digital health platforms. Longitudinal studies assessing the predictive validity of Annual-MPCS and ACV against complication rates and healthcare utilization would further demonstrate clinical utility. Expanding the tool for patient self-management and personalized feedback holds potential to enhance empowerment and chronic disease control.

5. Conclusion: Implications for Practice and Policy

This study introduces the MPCST, a rigorously developed and validated instrument integrating clinical, behavioral, and contextual factors, including SDoH, to provide a comprehensive, scalable framework for T2DM management assessment in resource-limited primary care settings. Grounded in established theoretical models and supported by robust psychometric evidence, the MPCST advances chronic disease evaluation by emphasizing adherence metrics and employing novel analytic methods such as generalizability theory and geometric mean aggregation. The use of machine learning for care modeling and Streamlet for interface prototyping highlights the tool's analytical strength and practical utility. Its flexible design, enabling both manual and digital administration, enhances feasibility across diverse healthcare contexts. While further external validation is warranted, the MPCST represents a significant advancement bridging theoretical frameworks and real-world application, offering clinicians, health systems, and policymakers a practical tool to improve diabetes care quality and outcomes in LMICs. Integration with digital health strategies aligns with Kenya's National e-Health Policy 2016-2030 [13], facilitating scalability and supporting patient empowerment through personalized self-monitoring.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] IDF. International Diabetes Federation. IDF Diabetes Atlas. 11th ed. Brussels: International Diabetes Federation; 2023 (cited 2025 Jun 28). Available from: <https://www.idf.org/component/attachments/attachments.html?id=1938&task=downloadreddit.com+3diabetesatlas.org+3en.wikipedia.org+3>
- [2] Młynarska E, Czarnik W, Dzieża N, Jędraszak W, Majchrowicz G, Prusinowski F, et al. Type 2 Diabetes Mellitus: New Pathogenetic Mechanisms, Treatment and the Most Important Complications. *Int J Mol Sci*. 2025 Jan; 26(3): 1094.
- [3] WHO WH. *WHO guideline on self-care interventions for health and well-being, 2022 revision*. Geneva: World Health Organization; 2022. 186 p.
- [4] Sugandh F, Chandio M, Raveena F, Kumar L, Karishma F, Khuwaja S, et al. Advances in the Management of Diabetes Mellitus: A Focus on Personalized Medicine. *Cureus*. 2023 Aug 18 (cited 2025 Jul 1). Available from: <https://www.cureus.com/articles/177570-advances-in-the-management-of-diabetes-mellitus-a-focus-on-personalized-medicine>
- [5] Ayuya OI. Ethnicity, social connectedness, and the rural-urban food continuum: Food security among urban informal settlement dwellers in Kenya. *Heliyon*. 2024 May 15 (cited 2025 Jun 28); 10(9). Available from: [https://www.cell.com/heliyon/abstract/S2405-8440\(24\)06512-5](https://www.cell.com/heliyon/abstract/S2405-8440(24)06512-5)
- [6] Azagew T, et al. Poor glycemic control and its predictors among people living with diabetes in low- and middle-income countries: a systematic review and meta-analysis. *BMC Public Health*. (cited 2025 Jun 28). Available from: <https://link.springer.com/article/10.1186/s12889-025-21828-y>
- [7] Aje AA, Fakeye TO. Factors associated with disease knowledge and attitude among ambulatory patients with type 2 diabetes - a multicenter study. *BMC Endocr Disord*. 2024 Aug 26; 24(1): 158.
- [8] Onteri SN, Kariuki J, Mathu D, Wangui AM, Magige L, Mutai J, et al. Diabetes health care specific services readiness and availability in Kenya: Implications for Universal Health Coverage. *PLOS Glob Public Health*. 2023 Sep 27; 3(9): e0002292.
- [9] Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving Chronic Illness Care: Translating Evidence Into Action. *Health Aff (Millwood)*. 2001 Nov; 20(6): 64-78.
- [10] Sivakumar A, Pan RY, Wang A, Choi D, Charif AB, Kastner M, et al. Assessing the sustainability and scalability of a diabetes eHealth innovation: a mixed-methods study. *BMC Health Serv Res*. 2023 Jun 14; 23(1): 630.
- [11] Del Castanhel F, Fonseca FR, Bonnassis Burg L, Maia Nogueira L, Rodrigues de Oliveira Filho G, Grosseman S. Applying the Generalizability Theory to Identify the Sources of Validity Evidence for the Quality of Communication Questionnaire. *Am J Hosp Palliat Med*. 2024 Jun 1; 41(7): 792-9.

- [12] Awiphan R, Kowatcharakul W, Ruengorn C, Noppakun K, Thavorn K, Nochaiwong S. Temporal trend in quality indicators of diabetes care performance among persons with type 2 diabetes in primary care practice: a serial cross-sectional analytical study, 2013/14 to 2021/22. *Front Endocrinol*. 2024 Oct 23 (cited 2025 Jul 1); 15. Available from: <https://www.frontiersin.org/articles/10.3389/fendo.2024.1359904/full>
- [13] Ministry of Health, Kenya. *Kenya National e-Health Policy 2016-2030*: KEN_Kenya_National-eHealth-Policy_2016-2030
- [14] Colin RM, Hollins CJ. Minimum Sample Size Requirements for a Validation Study of the Birth Satisfaction Scale-Revised (BSS-R). *J Nurs Pract*. 2017 Sep 15 (cited 2025 Jul 1); 1(1). Available from: <https://scholars.direct/Articles/nursing/jnp-1-005.php?jid=nursing>
- [15] Abdul Basith Khan M, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of Type 2 Diabetes — Global Burden of Disease and Forecasted Trends. *J Epidemiol Glob Health*. 2020 Mar 1; 10(1): 107-11.
- [16] Sagastume D, Mertens E, Sibongwere DK, Dimbelolo JC, Kabundi JCK, de Man J, et al. A retrospective database study of the demographic features and glycemic control of patients with type 2 diabetes in Kinshasa, Democratic Republic of the Congo. *BMC Med*. 2022 Aug 19; 20(1): 258.
- [17] Diawara A, et al. Type 2 diabetes prevalence, awareness, and risk factors in rural Mali: a cross-sectional study. *Sci Rep*. (cited 2025 Jun 28). Available from: <https://www.nature.com/articles/s41598-023-29743-1>
- [18] Jaspers FW-W, Kengne AP, Meeks KAC, Agyemang C. Prevalence of metabolic syndrome in sub-Saharan Africa: A systematic review and meta-analysis. *Nutr Metab Cardiovasc Dis*. 2020 Apr 12; 30(4): 547-65.
- [19] Clark JR, Chen G. Current understanding of gender dimorphism regarding type 2 diabetes. *Arch Nutr Food Sci*. 2020 Jul 23; 1(2): 20-9.
- [20] Muhammad EA, Shewarega ES, Baffa LD, Fentie EA, Belew AK, Mengistu B, et al. Dietary practice and associated factors among patients with type 2 diabetes mellitus in the Horn of Africa: a systematic review and meta-analysis. *BMC Nutr*. 2025 Apr 21; 11(1): 81.
- [21] Walker AF, Graham S, Maple-Brown L, Egede LE, Campbell JA, Walker RJ, et al. Interventions to address global inequity in diabetes: international progress. *Lancet*. 2023 Jul 15; 402(10397): 250-64.
- [22] Solar O, Irwin A. *A Conceptual Framework for Action on the Social Determinants of Health*. WHO Document Production Services; 2010 (cited 2025 Jun 27). Available from: <http://hdl.handle.net/1903/23135>
- [23] Choe JH, Xuan S, Goldenberg A, Matian J, McCombs J, Kim RE. Medication Persistence and Its Impact on Type 2 Diabetes. *Am J Manag Care*. 2024 Apr 1; 30(4): e124-34.
- [24] Eregata GT, Hailu A, Memirie ST, Norheim OF. Measuring progress towards universal health coverage: national and sub-national analysis in Ethiopia. *BMJ Glob Health*. 2019 Nov 1 (cited 2025 Jun 28); 4(6). Available from: <https://gh.bmj.com/content/4/6/e001843>
- [25] Mchenga M, Manthulu G, Chingwanda A, Chirwa E. Developing Malawi's Universal Health Coverage Index. *Front Health Serv*. 2022 Feb 10 (cited 2025 Jun 28); 1. Available from: <https://www.frontiersin.org/articles/10.3389/frhs.2021.786186/full>
- [26] Mehta N, Preston SH. Are major behavioral and sociodemographic risk factors for mortality additive or multiplicative in their effects? *Soc Sci Med*. 2016 Apr 1; 154: 93-9.
- [27] Vanitha G, Kasthuri M. Medical data classification using enhanced correlation attribute. *The Scientific Temper (2025) Vol. 16 (2): 3736-3746*.
- [28] Calderon-Margalit R, Cohen-Dadi M, Opas D, Jaffe DH, Levine J, Ben-Yehuda A, et al. Trends in the performance of quality indicators for diabetes care in the community and in diabetes-related health status: an Israeli ecological study. *Isr J Health Policy Res*. 2018 Jan 17; 7(1): 10.
- [29] Thompson W, Wang H, Xie M, Kolassa J, Rajan M, Tseng CL, et al. Assessing Quality of Diabetes Care by Measuring Longitudinal Changes in Hemoglobin A1c in the Veterans Health Administration. *Health Serv Res*. 2005; 40(6 Pt 1): 1818-35.
- [30] Lertsakulbunlue S, Kantiwong P. Evaluating the dependability of peer assessment in project-based learning for pre-clinical students: a generalizability theory approach. *BMC Med Educ*. (cited 2025 Jun 28); Available from: <https://link.springer.com/article/10.1186/s12909-025-06772-0>
- [31] Wang X, et al. Biological variation of cardiovascular biochemical markers in patients with Type 2 Diabetes Mellitus. *Clin Chim Acta*. 2022 Sep 1; 534: 161-6.
- [32] Vural Keskinler M, Takir M, Caklili OT, Oguz A. The Frequency and Determinants of HbA1c Variability in Type 2 Diabetic Patients. *Metab Syndr Relat Disord*. 2021 Sep; 19(7): 372-7.
- [33] Kariuki S. World Health Day: Universal Health Coverage - Everyone, Everywhere - Celebrating Kenya's journey towards universal health coverage. Nairobi: Ministry of Health; 2019 (cited 2025 Jun 28). Available from: <https://www.health.go.ke/wp-content/uploads/2019/04/WORLD-HEALTH-DAY-SUPPORT-07-04-2019.pdf>
- [34] Hearn C, et al. Self-management of non-communicable diseases in low- and middle-income countries: A scoping review. *PLoS One*. (cited 2025 Jun 28). Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0219141>