



# Predicting Health-Related Quality of Life Among Patients With Type 2 Diabetes Using Random Forest and SHAP-Based Feature Importance Analysis

Karl. Pöder<sup>1</sup>, Ana Lucía. Martínez<sup>2\*</sup>

<sup>1</sup> Department of Cognitive Psychology, University of Tartu, Tartu, Estonia

<sup>2</sup> Department of Psychology, National Autonomous University of Mexico (UNAM), Mexico City, Mexico

\* Corresponding author email address: ana.martinez@unam.mx

## Article Info

### Article type:

Original Research

### How to cite this article:

Pöder, K., & Martínez, A. L. (2025). Predicting Health-Related Quality of Life Among Patients With Type 2 Diabetes Using Random Forest and SHAP-Based Feature Importance Analysis. *Quality of Life and Health Sciences*, 1(2) 1-16. <http://dx.doi.org/10.61838/kman.qlhs.5784>



© 2025 the authors. Published by KMAN Publication Inc. (KMANPUB), Ontario, Canada. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

## ABSTRACT

**Objective:** This study aimed to predict health-related quality of life among patients with type 2 diabetes in Mexico using a Random Forest model and to identify the most influential clinical, behavioral, psychological, and sociodemographic predictors through SHAP-based feature importance analysis.

**Methods and Materials:** This cross-sectional predictive modeling study was conducted among 462 adult patients with type 2 diabetes receiving outpatient care in Mexico City, Guadalajara, and Monterrey. Health-related quality of life was assessed using the 36-Item Short Form Health Survey. Clinical, behavioral, psychological, and sociodemographic data were collected using medical records and standardized self-report instruments, including measures of diabetes self-care and psychological distress. The dataset was divided into training and testing subsets using an 80:20 split. A Random Forest regression model was developed to predict total health-related quality of life scores. Model performance was evaluated using the coefficient of determination, root mean square error, mean absolute error, and mean squared error. SHAP analysis was applied to interpret global feature importance and the direction of predictor effects.

**Findings:** The Random Forest model demonstrated strong predictive performance, explaining 73.5% of the variance in health-related quality of life in the independent testing set. The model achieved an RMSE of 8.34, MAE of 6.49, and MSE of 69.56. Cross-validation and out-of-bag estimation showed comparable performance, with  $R^2$  values of 0.721 and 0.704, respectively. SHAP analysis identified depression score as the strongest predictor of lower health-related quality of life, followed by HbA1c, number of diabetes-related complications, physical activity self-care, duration of diabetes, body mass index, stress score, medication adherence, sleep duration, age, hypertension, anxiety score, foot care behavior, household income, and insulin use.

**Conclusion:** The findings indicate that health-related quality of life among patients with type 2 diabetes can be predicted with acceptable accuracy using Random Forest modeling, while SHAP analysis provides clinically interpretable evidence that psychological distress, glycemic control, complication burden, and self-care behaviors are central determinants of patient-perceived health.

**Keywords:** Type 2 diabetes; health-related quality of life; Random Forest; SHAP; machine learning; self-care; depression; glycemic control.

## 1. Introduction

Type 2 diabetes mellitus is a complex, chronic, and progressive metabolic disorder that extends far beyond abnormalities in blood glucose regulation and imposes a multidimensional burden on physical functioning, psychological well-being, social participation, and daily self-management. As the disease progresses, patients are required to maintain long-term adherence to medication, dietary regulation, physical activity, glucose monitoring, regular medical follow-up, and prevention of microvascular and macrovascular complications. These demands make type 2 diabetes not only a biomedical condition but also a lived experience that continuously affects perceived health, autonomy, emotional stability, and quality of life. In clinical diabetes care, the traditional emphasis on biochemical indicators such as glycated hemoglobin, fasting blood glucose, lipid profile, renal markers, and body mass index remains essential; however, these indicators do not fully capture how patients experience their illness in everyday life. Health-related quality of life has therefore become a central outcome in diabetes research because it reflects the patient's subjective perception of physical capacity, emotional well-being, social functioning, symptom burden, and overall health status. In patients with type 2 diabetes, health-related quality of life is influenced by disease duration, treatment complexity, self-care behaviors, psychological distress, sleep quality, comorbidities, socioeconomic resources, and diabetes-related complications, making it a suitable outcome for integrative predictive modeling (Ali et al., 2024; Mansour et al., 2023).

The clinical course of type 2 diabetes is often shaped by the cumulative effect of metabolic dysregulation and chronic complications. Poor glycemic control has been associated with adverse clinical outcomes, particularly among patients with type 2 diabetes and cardiovascular comorbidity, highlighting the importance of HbA1c not only as a marker of metabolic control but also as a predictor of broader health risks (Tang et al., 2024). Cardiovascular complications are especially important because they directly affect physical functioning, treatment burden, and perceived vulnerability, while molecular and biomarker studies continue to clarify the mechanisms through which diabetes contributes to vascular and systemic deterioration (Elshafae et al., 2020; Mouriño-Álvarez et al., 2021; Rabizadeh et al., 2021). In addition, kidney impairment is one of the most consequential diabetes-related complications because it may remain asymptomatic during early stages but later impose

substantial restrictions on physical health and daily life. Recent studies have emphasized the potential role of renal biomarkers, including cystatin C, beta trace protein, and related markers, for earlier detection of kidney injury and diabetic nephropathy in patients with type 2 diabetes (Ali et al., 2024; Mansour et al., 2023; Ramadevi & Sunitha, 2025). Such findings suggest that quality of life in diabetes should not be interpreted as a simple psychological outcome but as a multidimensional construct affected by biological risk, complication burden, treatment history, and functional decline.

Beyond renal and cardiovascular complications, type 2 diabetes is associated with a broad range of health problems that may influence health-related quality of life through pain, disability, sensory impairment, sexual health problems, oral health difficulties, musculoskeletal deterioration, and reduced independence. Diabetic retinopathy and ocular changes may impair vision and interfere with mobility, self-care, and social participation (Deepa et al., 2021). Sarcopenia, osteoporosis, and changes in body composition may also contribute to frailty and reduced physical capacity among patients with type 2 diabetes, making markers such as the creatinine-to-cystatin C ratio potentially relevant for identifying musculoskeletal vulnerability (Dai & Xu, 2022). Erectile dysfunction has been identified as a significant diabetes-related problem in African populations, indicating that quality of life is also affected by sexual health and intimate functioning, domains that are often underreported in routine diabetes care (Shiferaw et al., 2020). Other diabetes-related conditions, such as tooth wear and oral complications, may be affected by treatment type and metabolic status and may further reduce comfort, nutrition, and social confidence (Younes & Hamam, 2022). These diverse complications show that type 2 diabetes can disturb multiple bodily systems, and this systemic burden must be reflected in predictive approaches that move beyond single-risk-factor analysis.

Psychological and behavioral factors are equally central to understanding quality of life among patients with type 2 diabetes. Depression is common among patients with type 2 diabetes and can reduce motivation, interfere with self-care, worsen perceived health, and amplify the subjective burden of symptoms (Algaows et al., 2021). Psychological distress can also interact with treatment adherence, sleep, physical activity, and dietary behavior, producing a cycle in which emotional symptoms and metabolic burden reinforce each other. Self-care behavior is a major determinant of diabetes outcomes because patients must translate medical

recommendations into repeated daily actions. Evidence on self-care practices, treatment satisfaction, and quality of life among patients with type 2 diabetes indicates that behavioral engagement and satisfaction with care are closely connected with perceived well-being (Banstola et al., 2022). Health literacy and self-care behavior have also been shown to influence health-related quality of life in patients with type 2 diabetes, suggesting that patients' ability to understand, access, and apply health information is an important resource for disease management (Gaffari-fam et al., 2020). Similarly, structured education in diabetes management has been emphasized as a key strategy for improving self-management and supporting long-term disease control (Chowdhury, 2022). These findings indicate that quality of life in diabetes is not determined exclusively by disease severity but also by patients' knowledge, skills, behavioral routines, and confidence in managing their condition.

Sleep and family support are additional factors that may meaningfully influence health-related quality of life in type 2 diabetes. Sleep deprivation in adults with type 2 diabetes has been associated with adverse consequences for metabolic regulation, emotional functioning, and daily performance, making sleep duration and sleep quality important candidate predictors of health-related quality of life (Woods et al., 2023). Inadequate sleep may reduce energy, impair cognitive functioning, increase stress, worsen glycemic control, and lower the patient's ability to engage in physical activity and dietary regulation. Family support is also important because diabetes management frequently occurs within household routines involving meals, finances, emotional encouragement, transportation to care, and medication support. Perceived family support has been associated with diabetes control among adults with type 2 diabetes, indicating that interpersonal resources may contribute to both clinical outcomes and subjective well-being (Emeka et al., 2022). In this respect, health-related quality of life should be viewed as the result of interaction between clinical burden, psychological distress, lifestyle behavior, sleep, social support, and access to health resources.

The increasing complexity of diabetes care has encouraged the use of predictive analytics and machine-learning methods to identify high-risk patients and support individualized decision-making. Traditional statistical models are useful for testing linear associations and estimating average effects, but they may be limited when predictors are numerous, nonlinear, interdependent, or differently important across patient subgroups. Machine-

learning approaches can model complex interactions among clinical, psychological, behavioral, and demographic variables and may therefore improve the prediction of patient-centered outcomes such as health-related quality of life. Studies using predictive supervised learning and data analysis have demonstrated the potential of machine learning for classification and prediction in type 2 diabetes and related diabetic conditions (Prabakar et al., 2022; Roobini & Lakshmi, 2021). Ensemble approaches have also been applied for diabetes prediction and classification, supporting the value of multistage and combined learning strategies in metabolic disease research (Simaiya et al., 2022). Furthermore, artificial intelligence has been proposed for predicting cardiovascular complications in patients with type 2 diabetes, indicating that predictive modeling can be extended from disease detection to complication forecasting and risk stratification (Ferdosian et al., 2021). These developments provide a strong rationale for applying machine-learning models to health-related quality of life, particularly because quality of life is influenced by a combination of variables that may not operate in a purely linear or additive manner.

Among machine-learning methods, Random Forest is especially suitable for health-related quality of life prediction because it can manage high-dimensional predictor sets, detect nonlinear relationships, model interactions, and reduce overfitting through ensemble decision trees. However, one limitation of advanced machine-learning models is that they are often perceived as "black box" methods because they may produce accurate predictions without clearly explaining how each variable contributes to the final output. This limitation is particularly important in health sciences, where interpretability is essential for clinical translation, patient trust, and intervention planning. Explainable artificial intelligence methods such as SHAP-based feature importance analysis help address this limitation by estimating the contribution of each predictor to model predictions. SHAP analysis allows researchers to identify which variables have the greatest influence on predicted quality of life and whether high or low values of a variable increase or decrease the predicted outcome. This is important because clinicians and policymakers need not only to know whether a model predicts health-related quality of life accurately, but also to understand which factors are most actionable, such as depression, physical activity, medication adherence, sleep duration, glycemic control, and complication burden.

The growing use of digital health technologies further strengthens the relevance of predictive and explainable models in diabetes care. Technology use among patients with diabetes is influenced by underlying motives for adoption, suggesting that digital health behaviors cannot be understood only as access to devices but should also be examined in relation to patient motivation, perceived usefulness, and behavioral engagement (Csuka et al., 2025). Digital tools, remote monitoring, automated risk prediction, and explainable dashboards may help clinicians identify patients at risk of low quality of life and provide targeted support before deterioration becomes severe. Broader advances in precision medicine, biomedical innovation, and algorithmic decision-making also illustrate the movement toward individualized care across medical fields, including chronic inflammatory disease management, outcome measurement, gene-based treatment strategies, and prognostic modeling (Fokkens, Corso, et al., 2024; Fokkens, Sedaghat, et al., 2024; Suwito et al., 2023; Wu et al., 2021). Although these fields differ from diabetes care, they collectively demonstrate the increasing importance of algorithms, biomarkers, outcome measures, and personalized risk assessment in contemporary health research. Similarly, predictive studies outside diabetes, including research on hemorrhagic infarction after thrombolysis, show how multicenter clinical data can be used to identify predictors of adverse outcomes and improve risk-informed clinical decision-making (Zeinhom et al., 2024).

Health-related quality of life has also gained importance because chronic disease outcomes are shaped by population-level disruptions and broader health contexts. The COVID-19 pandemic had measurable effects on health-related quality of life in general populations, emphasizing that perceived health is sensitive not only to individual illness but also to social restrictions, psychological stress, access to care, and public health disruption (Nshimirimana et al., 2023). For patients with type 2 diabetes, such disruptions may be particularly consequential because delayed appointments, reduced physical activity, financial stress, and limited access to routine monitoring may worsen both metabolic status and subjective well-being. In addition, research on gestational diabetes screening highlights that diabetes-related risk identification depends on accurate assessment strategies and that reliance on limited risk-factor screening may be insufficient in some contexts (Olumodeji et al., 2021). Genetic, inflammatory, and metabolic research has also pointed to the biological heterogeneity of diabetes

and related chronic conditions, including associations between genetic polymorphism, body mass index, inflammatory biomarkers, and immune pathways in different disease contexts (Fujita & Inagaki, 2021; Maes et al., 2020; Oishi et al., 2020; Sabry et al., 2020). Although some of these studies address conditions outside type 2 diabetes or focus on related biological mechanisms, they collectively reinforce the need for integrative models capable of incorporating diverse predictors rather than relying on isolated clinical indicators.

Despite increasing attention to type 2 diabetes outcomes, several gaps remain in the prediction of health-related quality of life. First, many studies focus on biomedical markers or self-care behavior separately, whereas quality of life is shaped by the combined effects of metabolic control, complications, psychological distress, sleep, lifestyle, treatment burden, and social context. Second, conventional analyses often estimate average relationships but may fail to capture nonlinear or interaction effects among predictors. Third, machine-learning studies in diabetes have frequently prioritized diagnosis, classification, or complication prediction rather than patient-centered outcomes such as health-related quality of life. Fourth, even when predictive models are used, their clinical usefulness may be limited if the model does not provide interpretable explanations of feature importance. Finally, there remains a need for evidence from different sociocultural and healthcare settings, including Mexico, where diabetes care is affected by urbanization, lifestyle transitions, unequal access to care, and varying levels of health literacy, social support, and treatment continuity. A Random Forest model combined with SHAP-based feature importance analysis can address these gaps by producing both predictive accuracy and clinically interpretable insight into the most influential determinants of health-related quality of life.

The aim of this study was to predict health-related quality of life among patients with type 2 diabetes in Mexico using a Random Forest model and to identify the most influential sociodemographic, clinical, behavioral, and psychological predictors through SHAP-based feature importance analysis.

## 2. Methods and Materials

### 2.1. Study Design and Participants

This study was designed as a cross-sectional predictive modeling study aimed at estimating health-related quality of life among patients with type 2 diabetes and identifying the most influential sociodemographic, clinical, behavioral, and

psychological predictors through a Random Forest algorithm and SHAP-based feature importance analysis. The study population consisted of adult patients diagnosed with type 2 diabetes who were receiving outpatient care at public and private endocrinology and internal medicine clinics in Mexico. Data were collected from patients attending diabetes care centers in Mexico City, Guadalajara, and Monterrey between March 2024 and December 2024. A total of 462 patients with type 2 diabetes participated in the study. Participants were eligible for inclusion if they were 30 years of age or older, had a confirmed medical diagnosis of type 2 diabetes for at least one year, were able to read and understand Spanish, and provided informed consent to participate. Patients were excluded if they had type 1 diabetes, gestational diabetes, severe cognitive impairment, acute psychiatric symptoms, advanced renal failure requiring dialysis, active cancer treatment, or incomplete responses on the main outcome measure. The final analytical sample included 462 participants after screening for eligibility and completeness of data. Participation was voluntary, and all participants were informed about the objectives of the study, confidentiality of responses, and their right to withdraw at any stage without any effect on their medical care.

## 2.2. Measures

Health-related quality of life was assessed using the 36-Item Short Form Health Survey, which is one of the most widely used standardized instruments for evaluating perceived physical and mental health status in clinical and community populations. The questionnaire includes 36 items covering eight dimensions of health-related quality of life: physical functioning, role limitations due to physical health problems, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health. Scores for each domain were transformed to a scale ranging from 0 to 100, with higher scores indicating better perceived health-related quality of life. In the present study, the total health-related quality of life score was used as the primary outcome variable for predictive modeling, while the physical and mental component-related domains were also examined descriptively. The Spanish version of the instrument has been used extensively in chronic disease populations, including patients with diabetes, and its validity and reliability have been confirmed in previous studies. In the present study, internal consistency was examined using

Cronbach's alpha before inclusion of the scale score in the predictive model.

Diabetes-related clinical information was collected using a structured clinical information form developed for this study based on routine diabetes care indicators. This form included duration of diabetes, age at diagnosis, type of pharmacological treatment, insulin use, number of diabetes-related medications, presence of hypertension, dyslipidemia, obesity, cardiovascular disease, neuropathy, retinopathy, nephropathy, history of hospitalization due to diabetes complications, and frequency of medical follow-up. Recent glycated hemoglobin values were extracted from medical records when available, and when the most recent laboratory result was older than three months, patients were asked to provide updated laboratory information from their treating clinic. Body mass index was calculated using measured or clinically recorded weight and height. These clinical variables were included because health-related quality of life in type 2 diabetes is influenced not only by glycemic control but also by disease duration, treatment burden, comorbid conditions, and diabetes-related complications.

Diabetes self-management behaviors were assessed using the Summary of Diabetes Self-Care Activities measure. This instrument evaluates the frequency of self-care behaviors during the previous seven days and includes items related to general diet, specific diet, physical activity, blood glucose monitoring, foot care, and medication adherence. Responses are typically scored according to the number of days per week on which the patient performed each self-care activity, with higher scores indicating greater engagement in diabetes self-management. For the purposes of the present study, domain scores were calculated for dietary behavior, physical activity, glucose monitoring, foot care, and medication adherence. The Spanish version of the instrument has demonstrated acceptable psychometric properties in diabetes populations, and its validity and reliability have been supported in previous research. In this study, self-care domain scores were treated as candidate predictors in the machine-learning model.

Psychological distress was measured using the Depression, Anxiety, and Stress Scale-21. This self-report instrument consists of 21 items distributed across three subscales: depression, anxiety, and stress. Each item is rated on a four-point scale based on the extent to which the respondent experienced the symptom during the previous week. Higher scores indicate greater psychological distress. Because emotional distress may reduce perceived quality of life and interfere with diabetes self-management,

depression, anxiety, and stress scores were included as psychological predictors in the model. The Spanish version of the scale has shown acceptable construct validity and internal consistency in adult and clinical populations. In the present study, subscale scores were calculated separately and entered into the analysis as continuous variables.

Sociodemographic and lifestyle characteristics were measured using a researcher-developed demographic questionnaire. This questionnaire included age, sex, marital status, educational level, employment status, monthly household income, place of residence, health insurance status, smoking status, alcohol use, sleep duration, perceived social support, and regular physical activity. These variables were selected because health-related quality of life among patients with type 2 diabetes may be affected by social position, access to care, lifestyle conditions, and everyday health behaviors. All questionnaires were administered in Spanish by trained research assistants in a private setting at the participating clinics. When necessary, items were read aloud to participants with limited literacy, while care was taken to avoid influencing responses.

### 2.3. Data Analysis

Data analysis was conducted in several stages. First, the dataset was inspected for missing values, outliers, inconsistent responses, and distributional characteristics. Participants with missing values on the main health-related quality of life outcome were excluded from the analysis. For predictor variables, missing values were handled using median imputation for continuous variables and mode imputation for categorical variables when the proportion of missingness was low. Continuous predictors were examined using means, standard deviations, medians, and ranges, while categorical predictors were summarized using frequencies and percentages. Although Random Forest models do not require normal distribution of predictors, descriptive statistics were used to understand the sample profile and identify implausible values before model development. Categorical variables were dummy coded where necessary, and all variables were prepared in a single analytic dataset.

The primary predictive analysis was performed using a Random Forest regression model, with the total health-related quality of life score as the dependent variable. Candidate predictors included sociodemographic characteristics, diabetes duration, glycosylated hemoglobin, body mass index, treatment type, insulin use, diabetes-

related complications, comorbidities, self-care behaviors, depression, anxiety, stress, sleep duration, smoking status, and physical activity. The dataset was randomly divided into training and testing subsets, with 80% of the sample used for model training and 20% used for model evaluation. To reduce overfitting and improve model generalizability, hyperparameters were tuned using repeated k-fold cross-validation within the training set. The main hyperparameters considered during tuning included the number of trees, maximum tree depth, minimum number of samples required to split a node, minimum number of samples required at a leaf node, and the number of predictors randomly selected at each split. Model performance was evaluated in the testing dataset using the coefficient of determination, root mean square error, mean absolute error, and mean squared error.

To interpret the Random Forest model, SHAP analysis was applied to estimate the contribution of each predictor to the model's predictions. SHAP values were calculated to determine both global feature importance and the direction of each variable's influence on predicted health-related quality of life. Global importance was assessed by ranking predictors according to their mean absolute SHAP values, indicating the average magnitude of each feature's contribution to the prediction. SHAP summary plots and dependence patterns were used analytically to examine whether higher or lower values of key predictors were associated with higher or lower predicted quality of life. This approach was selected because Random Forest models are highly flexible and can capture nonlinear relationships and interactions, but they are less directly interpretable than conventional regression models. SHAP analysis therefore provided an interpretable explanation of the model by showing how clinical, behavioral, psychological, and demographic variables contributed to individual and overall predictions. All analyses were conducted using standard statistical and machine-learning procedures, and statistical reporting focused on predictive accuracy, ranking of influential features, and interpretation of the most important predictors of health-related quality of life among patients with type 2 diabetes.

### 3. Findings and Results

A total of 462 patients with type 2 diabetes were included in the final analysis. The mean age of the participants was 58.36 years with a standard deviation of 10.42 years, indicating that the sample mainly consisted of middle-aged and older adults receiving routine diabetes care. Of the total

participants, 255 patients were female, representing 55.2% of the sample, and 207 patients were male, representing 44.8%. Regarding marital status, 298 participants, equal to 64.5%, were married or living with a partner, while 164 participants, equal to 35.5%, were single, widowed, separated, or divorced. In terms of education, 107 participants, equal to 23.2%, had completed primary education or less, 158 participants, equal to 34.2%, had completed secondary education, 87 participants, equal to 18.8%, had completed technical or vocational education, and 110 participants, equal to 23.8%, had university-level education. With respect to employment status, 178 participants, equal to 38.5%, were employed, 153 participants, equal to 33.1%, were retired, 91 participants, equal to 19.7%, were homemakers, and 40 participants, equal to 8.7%, were unemployed at the time of data

collection. Most participants reported having some form of health insurance coverage, with 399 patients, equal to 86.4%, reporting public or private health insurance and 63 patients, equal to 13.6%, reporting no regular insurance coverage. Participants were recruited from three major urban clinical settings in Mexico, including 187 patients from Mexico City, 145 patients from Guadalajara, and 130 patients from Monterrey, corresponding to 40.5%, 31.4%, and 28.1% of the sample, respectively. Overall, the demographic profile of the sample showed adequate variability in age, sex, education, employment, and place of recruitment, which provided a suitable basis for examining the predictive contribution of sociodemographic, clinical, behavioral, and psychological variables to health-related quality of life.

**Table 1**

*Clinical, Behavioral, Psychological, and Health-Related Quality of Life Characteristics of the Participants*

Variable	Mean ± SD or n (%)	Minimum–Maximum
Duration of type 2 diabetes, years	9.28 ± 6.31	1–32
Age at diagnosis, years	49.08 ± 9.87	28–74
Glycated hemoglobin, HbA1c (%)	8.14 ± 1.72	5.60–13.90
Body mass index, kg/m <sup>2</sup>	29.84 ± 5.18	20.10–44.70
Number of diabetes-related medications	2.04 ± 0.91	1–5
Number of comorbid conditions	1.63 ± 1.07	0–5
Number of diabetes-related complications	1.08 ± 1.12	0–5
Insulin use	160 (34.6%)	—
Oral antidiabetic medication use	361 (78.1%)	—
Hypertension	263 (56.9%)	—
Dyslipidemia	228 (49.4%)	—
Diabetic neuropathy	131 (28.4%)	—
Diabetic retinopathy	80 (17.3%)	—
Diabetic nephropathy	56 (12.1%)	—
History of diabetes-related hospitalization	72 (15.6%)	—
General diet self-care, days per week	4.42 ± 1.61	0–7
Physical activity self-care, days per week	3.11 ± 1.94	0–7
Blood glucose monitoring, days per week	3.86 ± 2.17	0–7
Foot care behavior, days per week	3.29 ± 2.08	0–7
Medication adherence, days per week	5.92 ± 1.49	0–7
Depression score	13.76 ± 8.34	0–39
Anxiety score	10.92 ± 7.51	0–34
Stress score	15.44 ± 8.09	0–40
Sleep duration, hours per night	6.38 ± 1.31	3–10
Total health-related quality of life score	57.84 ± 16.22	18.40–94.70
Physical component-related quality of life	55.18 ± 17.36	14.20–96.10
Mental component-related quality of life	60.23 ± 16.91	16.80–97.40

As shown in Table 1, the participants had lived with type 2 diabetes for an average of 9.28 years, suggesting that a considerable proportion of the sample had experienced diabetes as a long-term chronic condition. The mean HbA1c

value was 8.14%, indicating that, on average, glycemic control was above the commonly recommended clinical target for many adults with type 2 diabetes. The mean body mass index was 29.84 kg/m<sup>2</sup>, showing that the average

participant was close to the obesity threshold, which is clinically relevant because excess body weight may intensify diabetes-related burden and reduce physical aspects of quality of life. More than half of the participants had hypertension, and nearly half had dyslipidemia, indicating a high burden of cardiometabolic comorbidity. Among diabetes-related complications, neuropathy was the most frequently reported complication, followed by retinopathy and nephropathy. In terms of self-care behaviors, medication adherence showed the highest mean score, whereas physical activity and foot care behaviors had comparatively lower mean scores. This pattern suggests that patients were more consistent in pharmacological self-management than in

lifestyle-related or preventive self-care activities. Psychological distress scores also showed considerable variability, with depression, anxiety, and stress symptoms present at different levels across the sample. The mean total health-related quality of life score was 57.84 out of 100, reflecting a moderate level of perceived quality of life, with the physical component score slightly lower than the mental component score. This finding indicates that physical limitations, treatment burden, diabetes complications, and comorbid conditions may have had a stronger adverse effect on perceived functioning than emotional and social domains for many participants.

**Table 2**

*Predictive Performance of the Final Random Forest Regression Model*

Model Evaluation Stage	R <sup>2</sup>	RMSE	MAE	MSE
Training set	0.941	3.89	2.85	15.13
Repeated k-fold cross-validation	0.721	8.55	6.68	73.10
Out-of-bag estimation	0.704	8.83	6.91	77.97
Independent testing set	0.735	8.34	6.49	69.56

Table 2 presents the predictive performance of the final Random Forest regression model for estimating health-related quality of life among patients with type 2 diabetes. The dataset was divided into a training set and an independent testing set, with 370 participants included in model training and 92 participants retained for final model evaluation. The model demonstrated very strong fit in the training set, with an R<sup>2</sup> value of 0.941, indicating that the Random Forest algorithm captured a large proportion of the variance in health-related quality of life during model development. However, because training performance alone may overestimate model accuracy, cross-validation, out-of-bag estimation, and independent testing performance were used to evaluate generalizability. The repeated k-fold cross-validation results showed an R<sup>2</sup> value of 0.721, while the out-of-bag estimate produced an R<sup>2</sup> value of 0.704. These results suggest that the model retained stable predictive

performance when evaluated on observations not directly used for each tree-building process. In the independent testing set, the model achieved an R<sup>2</sup> value of 0.735, an RMSE of 8.34, an MAE of 6.49, and an MSE of 69.56. This means that the final model explained approximately 73.5% of the variance in health-related quality of life among unseen participants, with an average absolute prediction error of approximately 6.49 points on the 0–100 quality-of-life scale. The relatively close correspondence between cross-validation, out-of-bag, and testing results indicates that the model was not merely memorizing the training data and had acceptable generalization capacity. Although the difference between training and testing performance reflects the expected optimism of a flexible ensemble model, the testing results support the usefulness of Random Forest regression for predicting health-related quality of life in patients with type 2 diabetes.

**Table 3**

*Prediction Error of the Random Forest Model Across Observed Health-Related Quality of Life Levels in the Testing Set*

Observed HRQoL Level in Testing Set	n	Observed HRQoL Mean ± SD	Predicted HRQoL Mean ± SD	MAE	RMSE	Prediction Bias
Low HRQoL, <50	31	39.91 ± 7.22	42.76 ± 8.04	6.98	8.75	+2.85
Moderate HRQoL, 50–69	43	59.72 ± 5.58	58.94 ± 6.37	5.73	7.49	-0.78
High HRQoL, ≥70	18	77.64 ± 6.11	73.06 ± 7.28	7.41	9.34	-4.58
Total testing set	92	57.96 ± 15.83	57.41 ± 13.97	6.49	8.34	-0.55

Table 3 shows the performance of the Random Forest model across clinically meaningful levels of observed health-related quality of life in the independent testing set. The model performed most accurately among participants with moderate health-related quality of life, where the MAE was 5.73 and the RMSE was 7.49. This suggests that the model was especially effective in predicting quality-of-life scores for patients located near the center of the outcome distribution. Among participants with low health-related quality of life, the model produced a positive prediction bias of +2.85, indicating that it slightly overestimated quality of life for patients with poorer perceived health status. In contrast, among participants with high health-related quality of life, the model produced a negative prediction bias of -4.58, indicating that it slightly underestimated quality of life

for patients with better perceived health status. This pattern is common in regression-based machine-learning models, particularly when extreme outcome values are less frequent than moderate values, because predictions tend to move toward the central tendency of the training data. Nevertheless, the total testing-set bias was only -0.55, indicating that the model did not show substantial overall systematic underprediction or overprediction. The results demonstrate that while the model had acceptable predictive accuracy across all quality-of-life levels, its strongest performance was observed for participants with moderate health-related quality of life, and slightly larger errors occurred at the lower and upper ends of the outcome distribution.

**Table 4**

*SHAP-Based Global Feature Importance for the Final Random Forest Model*

Rank	Predictor	Mean Absolute SHAP Value	Direction of Association With Predicted HRQoL
1	Depression score	4.86	Higher depression predicted lower HRQoL
2	Glycated hemoglobin, HbA1c	3.94	Higher HbA1c predicted lower HRQoL
3	Number of diabetes-related complications	3.62	More complications predicted lower HRQoL
4	Physical activity self-care	3.28	More frequent physical activity predicted higher HRQoL
5	Duration of type 2 diabetes	2.74	Longer duration predicted lower HRQoL
6	Body mass index	2.36	Higher body mass index predicted lower HRQoL
7	Stress score	2.18	Higher stress predicted lower HRQoL
8	Medication adherence	1.96	Higher adherence predicted higher HRQoL
9	Sleep duration	1.84	Adequate sleep predicted higher HRQoL
10	Age	1.62	Older age generally predicted lower HRQoL
11	Hypertension	1.45	Presence of hypertension predicted lower HRQoL
12	Anxiety score	1.36	Higher anxiety predicted lower HRQoL
13	Foot care behavior	1.24	More frequent foot care predicted higher HRQoL
14	Monthly household income	1.12	Higher income predicted higher HRQoL
15	Insulin use	1.03	Insulin use predicted lower HRQoL

Table 4 presents the SHAP-based global feature importance results for the final Random Forest model. Among all predictors, depression score had the largest mean absolute SHAP value, indicating that depressive symptoms made the strongest contribution to model predictions of health-related quality of life. Higher depression scores consistently shifted predicted quality-of-life values downward, suggesting that psychological distress was one of the central determinants of poor perceived health status among patients with type 2 diabetes. HbA1c was the second most influential predictor, and higher glycemic values were associated with lower predicted quality of life, reflecting the clinical relevance of glycemic control in perceived physical and general health. The number of diabetes-related complications ranked third, showing that accumulated

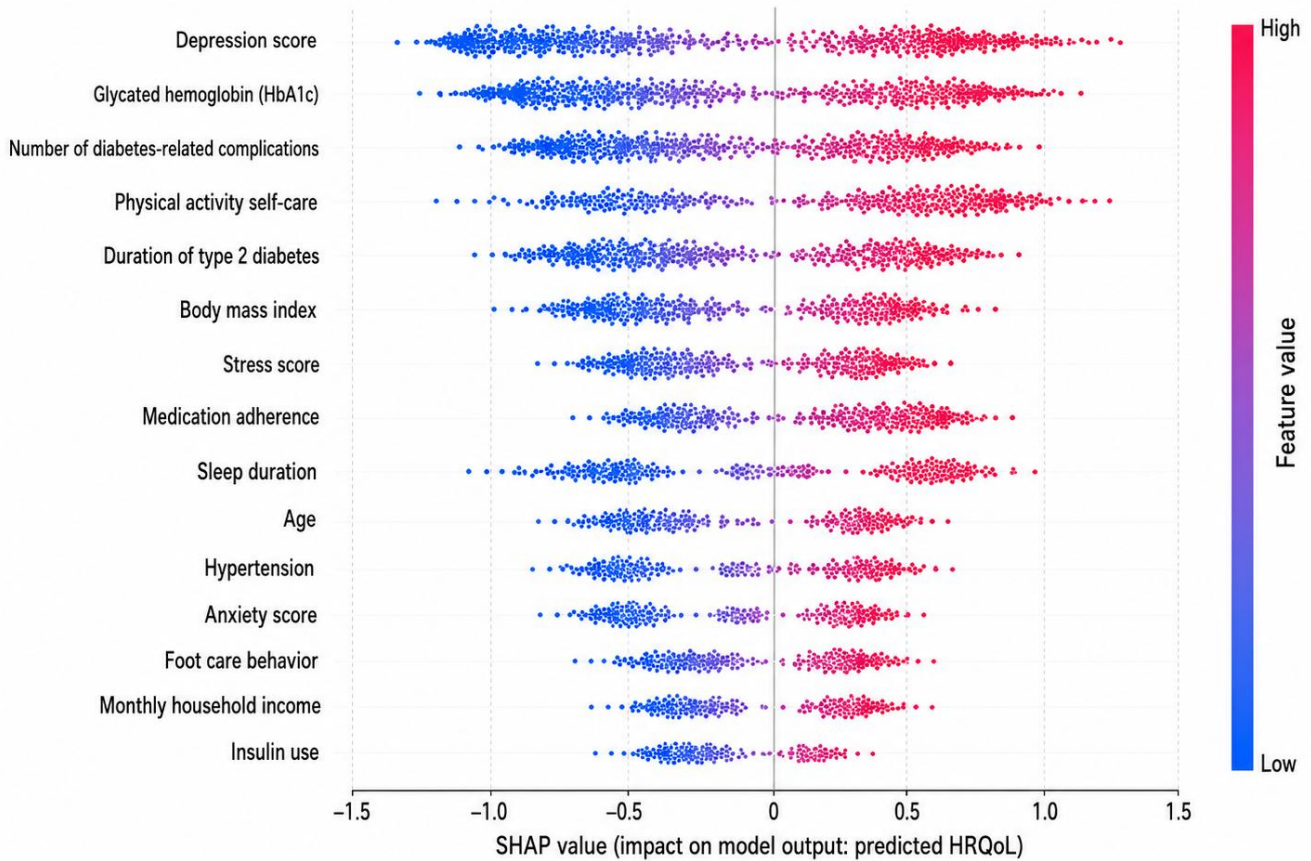
complication burden was strongly associated with reduced quality of life. Physical activity self-care ranked fourth and was the most influential positive behavioral predictor, indicating that participants who reported more frequent physical activity tended to receive higher predicted quality-of-life scores. Duration of diabetes, body mass index, stress, age, hypertension, anxiety, and insulin use were also associated with lower predicted health-related quality of life, whereas medication adherence, adequate sleep duration, foot care behavior, and higher household income were associated with higher predicted values. These results show that health-related quality of life among patients with type 2 diabetes was not predicted by a single clinical factor, but by a multidimensional combination of psychological distress, glycemic control, complication burden, self-care behavior,

cardiometabolic risk, and socioeconomic conditions. The SHAP findings also clarify that psychological variables, especially depression and stress, were not secondary

predictors; rather, they had a major role in the model's predictive structure alongside biomedical indicators.

**Figure 1**

*SHAP Summary Plot of the Final Random Forest Model for Predicting Health-Related Quality of Life Among Patients With Type 2 Diabetes*



The SHAP summary plot visually confirmed the ranking and direction of the most influential predictors identified in the global feature importance analysis. In the figure, predictors with wider horizontal dispersion contributed more strongly to variation in predicted health-related quality of life, whereas predictors with narrower distributions had smaller effects on model output. Depression score showed the widest distribution of SHAP values, confirming its dominant role in predicting health-related quality of life. High depression scores were concentrated on the negative side of the SHAP axis, showing that greater depressive symptoms reduced predicted quality-of-life values. Similarly, higher HbA1c levels and a greater number of diabetes-related complications were mainly located on the negative side of the plot, indicating that poor glycemic control and higher complication burden lowered predicted health-related quality of life. In contrast, higher levels of

physical activity self-care and medication adherence were generally located on the positive side of the SHAP axis, demonstrating their beneficial contribution to predicted quality-of-life scores. The plot also revealed nonlinear patterns for some variables. For example, sleep duration showed a positive contribution primarily when participants reported adequate sleep, whereas very short sleep duration contributed negatively to predicted health-related quality of life. Body mass index and diabetes duration also showed stronger negative SHAP values at higher levels, suggesting that the adverse effect of these variables became more pronounced as metabolic and disease-duration burden increased. Overall, the SHAP summary plot provided an interpretable explanation of the Random Forest model by showing how individual predictors shifted model predictions upward or downward and by demonstrating that reduced quality of life was most strongly associated with depressive

symptoms, poor glycemic control, accumulated diabetes complications, insufficient self-care behavior, and higher cardiometabolic burden.

#### 4. Discussion

The present study aimed to predict health-related quality of life among patients with type 2 diabetes in Mexico using a Random Forest model and SHAP-based feature importance analysis. The findings showed that the participants had a moderate level of health-related quality of life, with a mean total score of 57.84 out of 100, while physical component-related quality of life was slightly lower than mental component-related quality of life. Clinically, the sample demonstrated a considerable burden of diabetes-related risk, including elevated mean HbA1c, high body mass index, long disease duration, frequent hypertension and dyslipidemia, and a notable prevalence of neuropathy, retinopathy, and nephropathy. From a predictive perspective, the Random Forest model demonstrated acceptable and clinically meaningful performance, explaining 73.5% of the variance in health-related quality of life in the independent testing set. The relatively similar results obtained from repeated cross-validation, out-of-bag estimation, and testing-set evaluation indicate that the model had adequate generalizability and was not limited to memorizing the training data. The SHAP analysis further clarified the internal structure of the model and showed that depression score, HbA1c, number of diabetes-related complications, physical activity self-care, duration of diabetes, body mass index, stress score, medication adherence, sleep duration, age, hypertension, anxiety, foot care behavior, household income, and insulin use were the most influential predictors of health-related quality of life. This finding supports the view that quality of life in type 2 diabetes is not determined by one isolated clinical indicator but by the combined and nonlinear influence of psychological, metabolic, behavioral, complication-related, and socioeconomic variables.

The most important predictor in the SHAP analysis was depression score, indicating that higher depressive symptoms contributed strongly to lower predicted health-related quality of life. This result is theoretically and clinically coherent because depression can reduce energy, impair self-management motivation, intensify the perception of physical symptoms, disturb sleep, weaken treatment adherence, and increase social withdrawal. Previous evidence has emphasized the high prevalence of depression

among patients with type 2 diabetes and has shown that depressive symptoms are a major psychosocial burden in this population (Algaows et al., 2021). The present finding extends this line of evidence by showing that depression was not merely associated with health-related quality of life but was the strongest predictor in a nonlinear machine-learning model that simultaneously included clinical, behavioral, and demographic variables. This suggests that psychological distress may have a central role in explaining why patients with similar levels of glycemic control or disease duration experience different levels of perceived health. The importance of stress and anxiety in the SHAP ranking further confirms that emotional functioning is a core dimension of diabetes outcomes. In this regard, diabetes care should not be restricted to biomedical monitoring alone, because emotional symptoms may directly reduce quality of life and indirectly worsen self-care behavior, treatment continuity, and perceived disease control.

HbA1c was the second most influential predictor in the model, and higher HbA1c values were associated with lower predicted health-related quality of life. This finding is consistent with the clinical assumption that poor glycemic control increases symptom burden, accelerates complications, and raises the likelihood of adverse outcomes. Evidence from patients with coronary heart disease and type 2 diabetes has shown that baseline HbA1c has important implications for adverse outcomes, supporting the role of glycemic control as a central prognostic indicator (Tang et al., 2024). The present finding adds that HbA1c is also highly relevant to patient-centered outcomes, not only to biomedical endpoints. This is important because patients may experience poor glycemic control through fatigue, polyuria, sleep disturbance, worry about complications, dietary restriction, and the perceived failure of disease management. The model also identified the number of diabetes-related complications as the third most influential predictor. This result is strongly aligned with previous research on diabetic nephropathy, cardiovascular risk, retinopathy, sarcopenia-related vulnerability, and other complication pathways in type 2 diabetes (Ali et al., 2024; Dai & Xu, 2022; Deepa et al., 2021; Elshafae et al., 2020; Mansour et al., 2023; Ramadevi & Sunitha, 2025). Together, HbA1c and complication burden indicate that objective clinical deterioration is closely connected to subjective health status, especially when complications reduce mobility, vision, independence, sexual functioning, and physical comfort.

The role of comorbid and systemic complications in the present model also supports the interpretation that health-related quality of life is shaped by the whole disease burden rather than diabetes alone. Hypertension and body mass index were among the important predictors, and both were associated with lower predicted quality of life. This finding is compatible with evidence indicating that cardiometabolic risk factors and vascular complications are prominent concerns in type 2 diabetes (Mouriño-Álvarez et al., 2021; Rabizadeh et al., 2021). Similarly, studies on genetic susceptibility, insulin resistance, inflammatory pathways, and type 2 inflammatory biomarkers show that chronic diseases often involve complex biological mechanisms that may influence long-term functioning and perceived health (Fujita & Inagaki, 2021; Maes et al., 2020; Oishi et al., 2020; Sabry et al., 2020). Although these studies vary in disease focus and methodology, they support the broader principle that metabolic and inflammatory vulnerability may contribute to functional impairment and health deterioration. The inclusion of insulin use among the important predictors may also reflect disease severity and treatment complexity rather than the effect of insulin itself. Patients requiring insulin often have longer disease duration, poorer glycemic history, or more advanced metabolic impairment, which may explain why insulin use contributed to lower predicted quality of life in the model.

Physical activity self-care emerged as the strongest positive behavioral predictor of health-related quality of life. Participants who reported more frequent physical activity had higher predicted quality-of-life scores, suggesting that movement-related self-care may improve physical functioning, energy, weight control, psychological well-being, and perceived disease mastery. This result aligns with evidence showing that self-care practices, treatment satisfaction, and quality of life are interrelated among patients with type 2 diabetes (Banstola et al., 2022). Health literacy and self-care behaviors have also been shown to influence health-related quality of life, indicating that patients who are better able to understand and apply health information may experience better subjective health outcomes (Gaffari-fam et al., 2020). Structured diabetes education further supports self-management by improving patient knowledge, behavioral competence, and confidence in disease control (Chowdhury, 2022). In the present study, medication adherence and foot care behavior also contributed positively to predicted quality of life, although their relative importance was lower than physical activity. This pattern may indicate that lifestyle behaviors capable of

improving both physical and psychological functioning have broader quality-of-life effects than behaviors limited to treatment compliance alone. Nevertheless, medication adherence and preventive foot care remain clinically important because they reduce complication risk and support continuity of diabetes management.

Sleep duration was also identified as an important predictor in the SHAP analysis, with adequate sleep contributing positively to predicted quality of life and shorter sleep contributing negatively. This finding is consistent with evidence that sleep deprivation in adults with type 2 diabetes has negative consequences for metabolic regulation, emotional stability, and daily functioning (Woods et al., 2023). Sleep problems may worsen fatigue, appetite regulation, insulin resistance, mood symptoms, and self-care capacity, thereby influencing both clinical outcomes and subjective well-being. The nonlinear pattern observed for sleep duration is especially important because sleep may not function as a simple linear predictor; very short sleep may be harmful, while adequate sleep may be protective, and excessive sleep may reflect underlying illness or depression. The model also identified monthly household income as a positive predictor of quality of life, which suggests that socioeconomic resources may shape access to medication, nutritious food, transportation, follow-up care, glucose monitoring supplies, and opportunities for physical activity. In addition, perceived support and family involvement are likely to influence diabetes control and coping capacity, as previous work has shown that perceived family support is associated with diabetes control among adults with type 2 diabetes (Emeka et al., 2022). These findings reinforce the need to interpret diabetes outcomes within the patient's social and household context.

The predictive performance of the Random Forest model supports the value of machine learning for patient-centered outcomes in diabetes. Previous studies have demonstrated the usefulness of predictive supervised learning, data analysis, and ensemble approaches for diabetes classification and prediction (Prabakar et al., 2022; Roobini & Lakshmi, 2021; Simaiya et al., 2022). The present study extends this evidence by applying machine learning to health-related quality of life rather than only to diabetes diagnosis or classification. The use of SHAP analysis is particularly important because it provides transparency by identifying how variables contribute to model predictions. This is aligned with the growing application of artificial intelligence in diabetes-related complication prediction and risk stratification (Ferdosian et al., 2021). More broadly,

contemporary medical research increasingly emphasizes predictive modeling, personalized risk estimation, biomarkers, and algorithm-supported decision-making in different clinical domains (Wu et al., 2021; Zeinhom et al., 2024). The present results suggest that explainable machine learning can be used not only to identify patients at higher biomedical risk but also to identify those at risk for poor quality of life, which is a clinically meaningful and patient-centered outcome.

The findings should also be interpreted in the context of digital health and evolving models of chronic disease management. Technology use among patients with diabetes may depend on underlying motives for adoption, perceived usefulness, and health behavior patterns (Csuka et al., 2025). Therefore, predictive models could become more useful if integrated into digital health systems that provide individualized feedback, risk alerts, or self-management recommendations. The importance of physical activity, medication adherence, sleep, and psychological distress in the present model suggests that future digital interventions should not be limited to glucose tracking but should include behavioral and mental health monitoring. Similar trends toward algorithmic management, outcome measurement, and individualized care are visible across chronic disease research, including chronic rhinosinusitis management and recent advances in outcome assessment (Fokkens, Corso, et al., 2024; Fokkens, Sedaghat, et al., 2024). Although these fields are clinically different from diabetes, they show a common movement toward structured algorithms, measurable outcomes, and tailored intervention pathways. Likewise, advances in biomedical innovation, including gene-editing strategies and biomarker-driven medicine, illustrate the broader shift toward precision approaches in health sciences (Suwito et al., 2023). For diabetes care, this means that patient-centered predictive models should eventually combine clinical indicators, biomarkers, behavior, psychosocial factors, and digital data streams.

Several less direct references also help situate the present study within broader health-related quality-of-life and risk-prediction research. The COVID-19 pandemic demonstrated that health-related quality of life is vulnerable to social disruption, reduced access to care, psychological stress, and changes in daily routine (Nshimirimana et al., 2023). For patients with type 2 diabetes, such contextual disruptions may intensify metabolic risk and reduce perceived well-being. Research on gestational diabetes screening also highlights the importance of accurate risk identification and the limitations of relying only on simple risk-factor

approaches (Olumodeji et al., 2021). In the present study, the Random Forest model addressed this issue by integrating multiple predictors rather than relying on a single screening criterion. Additionally, studies on testosterone levels, oral health, and diabetes-related physiological complications emphasize that type 2 diabetes affects domains that may not always be central in routine glycemic monitoring but may still influence quality of life (Mahwi & Hassan, 2022; Younes & Hamam, 2022). Therefore, the present findings support a biopsychosocial interpretation of diabetes-related quality of life and demonstrate that predictive models are most useful when they account for clinical severity, psychological distress, health behavior, social resources, and treatment burden simultaneously.

## 5. Conclusion

Taken together, the findings indicate that the Random Forest model achieved acceptable predictive accuracy for estimating health-related quality of life among patients with type 2 diabetes. The model explained a substantial proportion of variance in the independent testing set and maintained stable performance across cross-validation and out-of-bag evaluation. The SHAP-based interpretation demonstrated that psychological distress, especially depression, was the strongest predictor of reduced health-related quality of life, followed by glycemic control, diabetes-related complications, physical activity self-care, duration of diabetes, body mass index, stress, medication adherence, sleep duration, and other clinical and socioeconomic variables. These findings suggest that health-related quality of life in type 2 diabetes is shaped by the combined influence of emotional, biomedical, behavioral, and social factors. The results also show that machine-learning approaches can provide clinically meaningful predictions while SHAP analysis can make the model interpretable by identifying the predictors that most strongly contribute to individual and overall quality-of-life estimates.

## 6. Limitations & Suggestions

This study had several limitations that should be considered when interpreting the findings. First, the cross-sectional design prevents causal inference, meaning that the model could identify important predictors of health-related quality of life but could not determine whether these factors directly caused changes in quality of life over time. Second, several variables, including self-care behaviors, sleep duration, psychological distress, and some

sociodemographic characteristics, were assessed through self-report instruments, which may be affected by recall bias, social desirability, or inaccurate reporting. Third, although the sample included patients from major urban clinical settings in Mexico, the findings may not fully represent rural populations, uninsured patients, Indigenous communities, or patients with limited access to specialized diabetes care. Fourth, some potentially relevant predictors, such as dietary quality, pain intensity, health literacy subdomains, social support quality, medication side effects, healthcare access barriers, and objective physical activity data, were not included in the final model. Fifth, although SHAP analysis improved interpretability, it does not replace clinical judgment and should be understood as an explanatory method for model behavior rather than proof of biological causality.

Future studies should use longitudinal designs to examine whether changes in depression, glycemic control, physical activity, sleep, medication adherence, and complication burden predict subsequent changes in health-related quality of life among patients with type 2 diabetes. Future research should also validate the present model in larger and more diverse samples, including patients from rural areas, different socioeconomic groups, primary care settings, and other Latin American countries. External validation is especially important before the model can be applied in clinical decision-making. Researchers should also compare Random Forest with other machine-learning approaches such as gradient boosting, XGBoost, support vector regression, neural networks, and regularized regression models to determine which methods provide the best balance between predictive accuracy and interpretability. In addition, future studies should include objective digital measures such as accelerometer-based activity, continuous glucose monitoring, sleep-tracking data, medication refill records, and electronic health record indicators. Qualitative studies may also be useful to understand how patients interpret quality of life, treatment burden, self-care expectations, and emotional distress in their own cultural and healthcare context.

The findings suggest that healthcare providers should assess health-related quality of life as a routine component of type 2 diabetes care rather than focusing exclusively on laboratory markers. Screening for depression, stress, anxiety, sleep problems, physical inactivity, low medication adherence, and diabetes-related complications should be integrated into regular follow-up visits. Patients identified as having low predicted quality of life should receive

multidisciplinary support, including psychological counseling, diabetes education, physical activity guidance, nutrition support, complication screening, and social assistance when needed. Clinicians should pay particular attention to patients with poor glycemic control, multiple complications, long disease duration, obesity, hypertension, and depressive symptoms, because these patients may be at higher risk of poor perceived health. The use of explainable machine-learning tools may help clinical teams prioritize patients who require more intensive follow-up and may support individualized care planning. However, such tools should be used as decision-support systems rather than replacements for professional judgment, patient dialogue, and shared decision-making.

### Acknowledgments

We would like to express our appreciation and gratitude to all those who cooperated in carrying out this study.

### Declaration of Interest

The authors of this article declared no conflict of interest.

### Ethical Considerations

The study protocol adhered to the principles outlined in the Helsinki Declaration, which provides guidelines for ethical research involving human participants.

### Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

### Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

### Authors' Contributions

All authors equally contributed to this article.

### References

- Algaows, F. S., Almisrea, S. N., Alshammari, F. K., Barakat, M., Alamri, A., Mesfer, A. A. S., Alahmadi, F. A., Altuwaijri, A. M., Naif, Z. A. N., Alawfi, R., Alqadeeb, M. M., & Zarnogi, D. N. (2021). Prevalence of Depression Among Type-2 Diabetic Patients-a Review. *Journal of Pharmaceutical*

- Research International*, 54-60. <https://doi.org/10.9734/jpri/2021/v33i40a32220>
- Ali, K. A. H., Adem, S., & El-Yassin, H. D. (2024). Role of Beta Trace Protein and Cystatin C as a Potential Biomarker for Early Detection of Type 2 Diabetic Nephropathy. *AL-Kindy College Medical Journal*, 20(1), 53-58. <https://doi.org/10.47723/a2qf5d64>
- Banstola, P., Yadav, D. K., & Pandey, R. (2022). Assessment of Self-Care Practices, Treatment Satisfaction and Quality of Life Among Diabetes Type2 Patients in Pokhara, Nepal. *Meds Alliance Journal of Medicine and Medical Sciences*, 2(3), 63-69. <https://doi.org/10.3126/mjmms.v2i3.47751>
- Chowdhury, S. R. (2022). Review the Evidence for Structured Education in Diabetes Management. *Dia&Endo*, 1(1). <https://doi.org/10.52803/is231>
- Csuka, S. I., Horvát, B., Csordás, G., Lakatos, C., & Martos, T. (2025). Technology Use and Health Behavior Among Patients With Diabetes: Do Underlying Motives for Technology Adoption Matter? *Frontiers in Digital Health*, 7. <https://doi.org/10.3389/fgdh.2025.1455261>
- Dai, H., & Xu, J. (2022). Creatinine-to-Cystatin C Ratio as a Marker of Sarcopenia for Identifying Osteoporosis in Male Patients With Type 2 Diabetes Mellitus. <https://doi.org/10.21203/rs.3.rs-1364770/v1>
- Deepa, K. P., Ghafoor, N., Siraj, N., Haque, M. Z., Islam, M. U., & Jubaida, M. R. (2021). Comparison of Ocular Lens Thickness of Type-2 Diabetic Retinopathy Patients With Non-Diabetic Control, Measured by High Resolution Ultrasonography. *Ibrahim Cardiac Medical Journal*, 10(1-2), 40-44. <https://doi.org/10.3329/icmj.v10i1-2.54002>
- Elshafae, M., Issa, H., moneam, A. m. A., Hafez, W. A. e., & Ahmed, E. (2020). Association Between Growth Differentiation Factor 15 and Cardiovascular Risk in Patients With Type 2 Diabetes Mellitus. *Benha Medical Journal*, 0(0), 0-0. <https://doi.org/10.21608/bmfj.2020.108486>
- Emeka, E. A., Echendu, S. T., Iloh, G. U. P., Umeadi, E. N., Adinnu, K. M., Ezeude, C. M., Onwuegbuna, A. A., Ibekwe, A. M., Okafor, F. N., Egbuniwe, M. C., & Eleje, G. U. (2022). Association Between Perceived Family Support and the Control of Adult Type 2 Diabetes Mellitus in a Nigerian Tertiary Hospital. *International Journal of Family Medicine & Healthcare*, 1(1). <https://doi.org/10.33425/2833-0382.1004>
- Ferdosian, H., Zamanian, H., Emami, S. A., Sedighi, E., Moridi, M., & Doustmehraban, M. (2021). Application of Artificial Intelligence in Prediction of Cardiovascular Complications in Patients With Diabetes Mellitus Type 2: A Protocol of Systematic Review. <https://doi.org/10.37766/inplasy2021.10.0076>
- Fokkens, W. J., Corso, E. D., Hellings, P. W., Harvey, R. J., Hopkins, C., & Reitsma, S. (2024). Algorithm for the Management of Primary Diffuse Chronic Rhinosinusitis. *Allergy*, 80(1), 359-361. <https://doi.org/10.1111/all.16417>
- Fokkens, W. J., Sedaghat, A. R., Soyka, M. B., & Reitsma, S. (2024). Recent Advances in Chronic Rhinosinusitis: Pathophysiology, Treatments, and Outcome Measures. *Rhinology Journal*, 0(0), 0-0. <https://doi.org/10.4193/rhin24.801>
- Fujita, Y., & Inagaki, N. (2021). Cell-autonomous Defects Contribute to Insulin Resistance in Skeletal Muscle. *Journal of Diabetes Investigation*, 12(7), 1136-1137. <https://doi.org/10.1111/jdi.13557>
- Gaffari-fam, S., Lotfi, Y., Daemi, A., Babazadeh, T., Sarbazi, E., Abbasabad, G. D., & Abri, H. (2020). Impact of Health Literacy and Self-Care Behaviors on Health-Related Quality of Life in Iranians With Type 2 Diabetes: A Cross-Sectional Study. <https://doi.org/10.21203/rs.3.rs-24026/v1>
- Maes, M., Vojdani, A., Sirivichayakul, S., Barbosa, D. S., & Kanchanatawan, B. (2020). Inflammatory and Oxidative Pathways Are New Drug Targets in Multiple Episode Schizophrenia and Leaky Gut, & Klebsiella Pneumoniae, and C1q Immune Complexes Are Additional Drug Targets in First episode Schizophrenia. <https://doi.org/10.20944/preprints202012.0457.v1>
- Mahwi, T., & Hassan, S. (2022). Prevalence of Low Free Testosterone Level in Men With Type 2 Diabetes Mellitus. *Journal of Sulaimani Medical College*, 12(3), 289-295. <https://doi.org/10.17656/jsmc.10371>
- Mansour, A., Abdelmoniem, R. O., Elbadawy, A., & Ibrahim, W. M. (2023). The Utility of Estimation of Glomerular Filtration Rate by Serum Cystatin C as a predictor of Diabetic Kidney Disease in Both Type I and Type II Diabetic Patients: A Single Center study. *The Egyptian Journal of Internal Medicine*, 35(1). <https://doi.org/10.1186/s43162-023-00243-y>
- Mouriño-Álvarez, L., Corbacho-Alonso, N., Sastre-Oliva, T., Corros-Vicente, C., Solís, J., Tejerina, T., Padial, L. R., & Barderas, M. G. (2021). Diabetes Mellitus and Its Implications in Aortic Stenosis Patients. *International Journal of Molecular Sciences*, 22(12), 6212. <https://doi.org/10.3390/ijms22126212>
- Nshimirimana, D. A., Kokonya, D., Gitaka, J., Wesonga, B., Nzioki, J. M., & Rukanikigitero, J. M. V. (2023). Impact of COVID-19 on Health-Related Quality of Life in the General Population: A Systematic Review and Meta-Analysis. *Plos Global Public Health*, 3(10), e0002137. <https://doi.org/10.1371/journal.pgph.0002137>
- Oishi, K., Matsunaga, K., Shirai, T., Hirai, K., & Gon, Y. (2020). Role of Type2 Inflammatory Biomarkers in Chronic Obstructive Pulmonary Disease. *Journal of clinical medicine*, 9(8), 2670. <https://doi.org/10.3390/jcm9082670>
- Olumodeji, A. M., Adebara, I., Adewara, O. E., Adeniyi, A. A., Awoyinka, S. B., Adeyemo, O. T., Ayankunle, O. M., Ajani, G. O., Ghazali, S. M., Olumodeji, U. O., & Okere, R. A. (2021). Gestational Diabetes Mellitus: Assessing the Reliability of Risk Factor-Based Screening in a Rural Nigerian Hospital. *Integrative Journal of Medical Sciences*. <https://doi.org/10.15342/ijms.2021.449>
- Prabakar, T. N., Shadaksharappa, B., & Ramkumar, P. (2022). Prediction of Diabetic Diseases Using Data Analysis. *International journal of health sciences*, 10013-10021. <https://doi.org/10.53730/ijhs.v6ns1.7364>
- Rabizadeh, S., Rajab, A., Mechanick, J. I., Moosaie, F., Rahimi, Y., Nakhjavani, M., & Esteghamati, A. (2021). LDL/apo B Ratio Predict Coronary Heart Disease in Type 2 Diabetes Independent of ASCVD Risk Score: A Case-Cohort Study. *Nutrition Metabolism and Cardiovascular Diseases*, 31(5), 1477-1485. <https://doi.org/10.1016/j.numecd.2021.01.013>
- Ramadevi, M., & Sunitha, S. (2025). Biomarkers for Early Detection of Kidney Injury Among Type 2 Diabetes Mellitus Patients. <https://doi.org/10.21203/rs.3.rs-6975415/v1>
- Roobini, M. S., & Lakshmi, M. R. (2021). Predictive Supervised Learning Model for Classification of Type 2 Diabetes Mellitus. <https://doi.org/10.21203/rs.3.rs-1009663/v1>
- Sabry, J. H., Ibrahim, M., Fadeel, S. A. E., & Elfallah, A. A. (2020). Study of FTO Gene Polymorphism Association With Type 2 Diabetes and Increasing BMI in Egyptian Patients. *Benha Journal of Applied Sciences*, 5(5), 1-9. <https://doi.org/10.21608/bjas.2020.136647>
- Shiferaw, W. S., Akalu, T. Y., Petrucka, P., Areri, H. A., & Aynalem, Y. A. (2020). Risk Factors of Erectile Dysfunction Among Diabetes Patients in Africa: A Systematic Review and Meta-Analysis. *Journal of Clinical & Translational*

- Endocrinology*, 21, 100232.  
<https://doi.org/10.1016/j.jcte.2020.100232>
- Simaiya, S., Kaur, R., Sandhu, J. K., Alsafyani, M., Alroobaea, R., Alsekait, D. M., Margala, M., & Chakrabarti, P. (2022). A Novel Multistage Ensemble Approach for Prediction and Classification of Diabetes. *Frontiers in Physiology*, 13. <https://doi.org/10.3389/fphys.2022.1085240>
- Suwito, B. E., Adji, A. S., Widjaja, J. S., Angel, S. C. S., Aufar Zimamuz Zaman Al, H., Salmay, N. F. W., & Choitrotussanijjah, C. (2023). A Review of CRISPR Cas9 for SCA: Treatment Strategies and Could Target B-Globin Gene and BCL11A Gene Using CRISPR Cas9 Prevent the Patient From Sickle Cell Anemia? *Open Access Macedonian Journal of Medical Sciences*, 11(F), 1-12. <https://doi.org/10.3889/oamjms.2023.11435>
- Tang, X., Li, Q., Han, Y., Wang, X., Song, Y., Zhang, Z., Jingjing, X., Liu, Z., Chen, Y., Zhang, Y. Z., Zhu, P., Guo, X., Jiang, L., Wang, Z., Liu, R., Wang, Q., Yao, Y., Feng, Y., Zhao, X., & Yuan, J. (2024). Implications of Baseline Glycemic Control by Plasma Glycated Hemoglobin A1c on Adverse Outcomes in Patients With Coronary Heart Disease and Type 2 Diabetes Mellitus: Results From the PROMISE Study. *Heliyon*, 10(22), e39748. <https://doi.org/10.1016/j.heliyon.2024.e39748>
- Woods, N. P., Tangpukdee, J., Thepa, T., & Methakanchanasak, N. (2023). Consequences of Sleep Deprivation in Adult Diabetes Mellitus Type 2 Patients: An Integrative Review. *Open Access Macedonian Journal of Medical Sciences*, 11(F), 1-10. <https://doi.org/10.3889/oamjms.2023.10029>
- Wu, R., Guo, S., Lai, S., Pan, G., Zhang, L., & Liu, H. (2021). A Stable Gene Set for Prediction of Prognosis and Efficacy of Chemotherapy in Gastric Cancer. <https://doi.org/10.21203/rs.3.rs-156785/v1>
- Younes, A., & Hamam, A. (2022). The Effect of Insulin and Metformin on Prevalence and Severity of Tooth Wear in Type2 Diabetic Patients. *Egyptian Dental Journal*, 68(3), 2659-2667. <https://doi.org/10.21608/edj.2022.133026.2070>
- Zeinhom, M. G., Ahmed, S. R., Kohail, A., Kamel, I. F. M., Abdelrahman, A., Nozha, O. M. A., Almoataz, M. A., Youssif, T. Y. O., Daabis, A. M. A., Refat, H. M., Ahmed Ahmed Mohamed Kamal, E., Elbassiouny, A., Akl, A. Z. O., Shuaib, A., Ismaiel, M., & Khalil, M. F. E. (2024). A Multicenter Trial on the Predictors of Different Subtypes of Hemorrhagic Infarction After Thrombolysis. *Scientific reports*, 14(1). <https://doi.org/10.1038/s41598-024-76189-0>