

Validation of formula-based glycosylated haemoglobin (hba1c) estimation and its potential utility for artificial intelligence model development: A prospective observational study.

Page | 1

*Dr. Satydeo Kumar Singh¹, Dr. Prabhavathi K^{2*}, Dr. Samudyatha UC³**¹Final Year Postgraduate Student, Department of Biochemistry, Sri Devaraj Urs Medical College, Tamaka, Kolar, Karnataka, India**²Professor and Head, Department of Biochemistry, Sri Devaraj Urs Medical College, Tamaka, Kolar, Karnataka, India**³Associate Professor, Department of Community Medicine, Sri Devaraj Urs Medical College, Tamaka, Kolar, Karnataka, India*

Abstract

Background

This study aimed to validate the formula $\text{hba1c (\%)} = 2.6 + 0.03 \times \text{fasting blood glucose (FBG)}$ for estimating glycosylated haemoglobin (hba1c), compare calculated and laboratory-measured hba1c values, determine agreement between the two methods, and assess the suitability of formula-derived hba1c data for future Artificial Intelligence (AI) applications.

Methodology

A hospital-based prospective observational study was conducted among 1000 adults undergoing simultaneous FBG and hba1c testing at RL Jalappa Hospital, Kolar, Karnataka. FBG was measured using the glucose oxidase–peroxidase method, while hba1c was estimated by High Performance Liquid Chromatography (HPLC). Formula-derived hba1c values were compared with laboratory hba1c using Pearson correlation, paired t-test, linear regression, Bland–Altman analysis, intraclass correlation coefficient (ICC), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE).

Results

The mean laboratory-measured hba1c was $7.66 \pm 1.39\%$, while the mean formula-derived hba1c was $7.67 \pm 1.64\%$. No significant difference was observed between the two methods ($p = 0.284$). A very strong positive correlation was found between measured and calculated hba1c values ($r = 0.958$, $p < 0.001$), with an adjusted R^2 of 0.918. Bland–Altman analysis demonstrated minimal bias (0.02%) and acceptable limits of agreement (-0.62% to $+0.66\%$). The ICC was 0.962 (95% CI: 0.956–0.967). MAE, RMSE, and MAPE were 0.24%, 0.31%, and 3.18%, respectively. The validated dataset was successfully prepared and partitioned for future AI model development.

Conclusion

Formula-derived hba1c demonstrated excellent agreement with laboratory-measured hba1c and may serve as a reliable, cost-effective alternative when standardized hba1c testing is unavailable. The validated dataset may support future AI-based diabetes prediction models.

Recommendation

Future multicentre studies should validate formula-derived hba1c in diverse populations and evaluate its performance in actual AI and machine learning models for diabetes prediction and monitoring.

Keywords: Diabetes mellitus; Glycosylated haemoglobin; hba1c; Fasting blood glucose; Artificial intelligence;

Submitted: April 10, 2026 **Accepted:** May 15, 2026 **Published:** June 30, 2026

Corresponding Author: Dr. Prabhavathi K

E-mail -: prabhaprajwal@gmail.com

Professor & Head, Department of Biochemistry, Sri Devaraj Urs Medical College, Tamaka, Kolar, Karnataka, India.

Background of the study

Page | 2

Diabetes mellitus (DM) is a chronic metabolic disorder characterized by persistent hyperglycaemia resulting from impaired insulin secretion, insulin resistance, or both. The global burden of diabetes has increased dramatically over the past few decades, making it one of the leading causes of morbidity and mortality worldwide. According to the International Diabetes Federation, more than 537 million adults are living with diabetes globally, and this number is projected to increase substantially by 2045 if effective preventive strategies are not implemented (1). In India, the prevalence of diabetes has risen rapidly because of urbanization, sedentary lifestyles, dietary transitions, obesity, and population ageing, placing considerable strain on healthcare resources (2).

Long-term hyperglycaemia contributes to progressive endothelial dysfunction and chronic inflammation, leading to microvascular complications such as diabetic nephropathy, retinopathy, and neuropathy, as well as macrovascular complications including coronary artery disease, cerebrovascular disease, and peripheral arterial disease (3). Early diagnosis and effective monitoring of glycaemic control are therefore essential to reduce disease progression, improve quality of life, and decrease diabetes-related mortality (4).

Among the available biomarkers for monitoring glycaemic control, glycosylated haemoglobin (hba1c) has become the internationally accepted reference standard. Hba1c is produced by the irreversible non-enzymatic glycation of haemoglobin within circulating erythrocytes and reflects the average blood glucose concentration during the preceding 8–12 weeks. Unlike single fasting or postprandial glucose measurements, hba1c is relatively unaffected by acute dietary intake, physical activity, or short-term metabolic fluctuations, thereby providing a more reliable indicator of long-term glycaemic status (5,6).

Standard laboratory estimation of hba1c is commonly performed using High Performance Liquid

Chromatography (HPLC), immunoassays, enzymatic assays, boronate affinity chromatography, or capillary electrophoresis. Among these methods, HPLC remains the reference analytical technique because of its excellent precision, reproducibility, and standardization according to National Glycohemoglobin Standardization Program (NGSP) recommendations (7). However, laboratory hba1c estimation requires sophisticated instrumentation, regular quality control procedures, trained technical personnel, and relatively high operational costs. These limitations reduce accessibility in rural hospitals, primary healthcare centres, and resource-limited regions (8).

To overcome these challenges, several investigators have proposed mathematical models for estimating hba1c from routinely available biochemical parameters, particularly fasting blood glucose (FBG). Formula-based estimation provides a rapid, inexpensive, and easily accessible alternative that can potentially support diabetes screening and monitoring where standardized laboratory testing is unavailable (9,10). The mathematical equation:

Hba1c (%) = 2.6 + 0.03 × FBG (mg/dl) has demonstrated encouraging agreement with laboratory-measured hba1c in preliminary validation studies and therefore warrants further evaluation in larger patient populations (11,12).

Recent advances in Artificial Intelligence (AI) and machine learning have transformed clinical decision-making by enabling automated prediction of disease risk, individualized treatment planning, and early identification of complications. The performance of AI algorithms depends heavily on the availability of large, high-quality, accurately labelled datasets. Formula-derived hba1c values, if validated against laboratory measurements, may serve as an important supplementary source of structured data for AI model training, particularly in settings where laboratory hba1c measurements are unavailable or economically impractical (13).

The present prospective observational study was therefore designed to validate a formula-based hba1c estimation model among 1000 participants attending RL Jalappa Hospital, Sri Devaraj Urs Medical College, Tamaka, Kolar. The study also aimed to evaluate the analytical agreement between calculated and laboratory-measured hba1c and explore the potential utility of formula-derived hba1c values for future Artificial Intelligence model development in diabetes care (14).

Methodology

This was a hospital-based prospective observational analytical study conducted to validate a formula-based estimation of glycosylated haemoglobin (hba1c) by comparing formula-derived hba1c values with laboratory-measured hba1c values. The study also assessed the suitability of the validated dataset for future Artificial Intelligence (AI) model development.

Study Setting

The study was conducted in the Department of Biochemistry, RL Jalappa Hospital, a tertiary-care teaching hospital attached to Sri Devaraj Urs Medical College (SDUMC), Tamaka, Kolar, Karnataka, India. The hospital provides comprehensive outpatient, inpatient, emergency, intensive care, and diagnostic laboratory services and serves as a major referral centre for Kolar district and neighbouring regions.

Study Duration

The study was conducted over a period of three months

Study Population

The study included adult patients attending RL Jalappa Hospital who underwent simultaneous fasting blood

glucose (FBG) and hba1c estimation during the study period.

Sample Size and Sampling Technique

A convenience sample of 1000 consecutive eligible participants was included. Consecutive sampling was employed, whereby all eligible individuals fulfilling the inclusion criteria during the study period were recruited until the desired sample size was achieved.

Inclusion Criteria

Adults aged 18 years and above.
Individuals undergoing simultaneous fasting blood glucose and hba1c estimation.
Participants willing to provide written informed consent.
Availability of complete demographic and laboratory records.

Exclusion Criteria

Patients younger than 18 years
Pregnant women
Patients with known haemoglobinopathies
Patients with haemolytic anaemia
Patients who received blood transfusion within the preceding three months
Patients with chronic liver disease
Patients with end-stage renal disease on dialysis
Patients receiving erythropoietin therapy
Incomplete laboratory records
Samples unsuitable for biochemical analysis

Measures to Address Bias

Several measures were undertaken to minimize potential sources of bias. Consecutive participant recruitment was adopted to reduce selection bias. Standardized laboratory protocols and quality-control procedures were followed throughout the study to minimize measurement bias. Formula-derived hba1c values were calculated using a predefined equation without modification. Data cleaning procedures, including verification of missing values, duplicate record removal, and consistency checks, were performed before statistical analysis to reduce information bias. Uniform inclusion and exclusion criteria were applied to all participants.

Ethical Considerations

Ethical approval was obtained from the Institutional Ethics Committee of Sri Devaraj Urs Medical College, Kolar. Written informed consent was obtained from all participants prior to enrolment. Confidentiality and anonymity of participant information were maintained throughout the study. All data were de-identified before analysis and used solely for research purposes.

Data Collection Procedure

Eligible participants attending outpatient and inpatient departments were enrolled after obtaining written informed consent. Following an overnight fasting period of at least 8 hours, venous blood samples were collected under aseptic precautions by trained laboratory personnel. Demographic information, including age and sex, was recorded using a structured data collection form.

Laboratory Investigations

Fasting Blood Glucose Estimation

Approximately 2 ml of venous blood was collected in fluoride vacutainers. Fasting blood glucose was analysed using the Glucose Oxidase–Peroxidase (GOD-POD) method on the Vitros 5600 Integrated Chemistry System following standard laboratory operating procedures.

Hba1c Estimation

Approximately 2 ml of venous blood was collected in EDTA vacutainers. Laboratory hba1c was measured using High Performance Liquid Chromatography (HPLC) on the Bio-Rad D-10 hba1c Analyzer according to manufacturer recommendations and National Glycohemoglobin Standardization Program (NGSP) guidelines.

Formula-Based hba1c Calculation

Formula-derived hba1c was calculated using the equation:

$$\text{Hba1c (\%)} = 2.6 + (0.03 \times \text{FBG [mg/dl]})$$

The calculated values were compared with laboratory-measured hba1c values.

Artificial Intelligence Dataset Preparation

The validated dataset was prepared for future AI model development by integrating demographic variables, fasting blood glucose values, laboratory-measured hba1c values, and formula-derived hba1c values. Data

preprocessing included duplicate removal, verification of missing values, consistency checking, and normalization of continuous variables. The dataset was randomly divided into training (80%, n=800) and testing (20%, n=200) subsets for future machine-learning applications.

Outcome Measures

Primary Outcome

Agreement between laboratory-measured hba1c and formula-derived hba1c

Secondary Outcomes

Correlation between laboratory-measured and formula-derived hba1c

Mean difference between the two methods.

Mean Absolute Error (MAE).

Root Mean Square Error (RMSE)

Mean Absolute Percentage Error (MAPE).

Intraclass Correlation Coefficient (ICC)

Linear regression coefficient (R^2)

Assessment of suitability of formula-derived hba1c data for future AI model development

Statistical Analysis

Data were entered into Microsoft Excel 2021 and analysed using IBM SPSS Statistics version 27.0 (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as mean $\pm$ standard deviation (SD), while categorical variables were presented as frequency and percentage. Pearson correlation coefficient was used to determine the strength of association between laboratory-measured and formula-derived hba1c values. Paired Student's t-test was applied to compare means between the two methods. Simple linear regression analysis was performed to evaluate predictive performance and coefficient of determination (R^2). Agreement between methods was assessed using Bland–Altman analysis. Reliability was evaluated using the Intraclass Correlation Coefficient (ICC) with 95% confidence intervals. Predictive performance was assessed using MAE, RMSE, and MAPE. A two-tailed p-value <0.05 was considered statistically significant.

Results

A total of **1000 participants** who fulfilled the eligibility criteria were enrolled during the 3-month study period. Complete demographic and laboratory data were available for all participants, and all were included in the final analysis.

Baseline characteristics

The mean age of the study population was **53.4 $\pm$ 12.6 years** (range: 18–86 years). The largest proportion of participants belonged to the **51–60 years** age group (29.8%), followed by the **41–50 years** age group (25.1%). Of the 1000 participants, **572 (57.2%) were males** and **428 (42.8%) were females**, resulting in a male-to-female ratio of **1.34:1** (Table 1).

Table 1. Baseline demographic characteristics of the study participants (N = 1000)

Variable	Number (%)
Age group (years)	
18–30	72 (7.2)
31–40	148 (14.8)
41–50	251 (25.1)
51–60	298 (29.8)
61–70	167 (16.7)
>70	64 (6.4)
Gender	
Male	572 (57.2)
Female	428 (42.8)

Distribution of Fasting Blood Glucose

blood glucose levels between **140 and 199 mg/dl**, while **27.9%** had values ≥ 200 mg/dl (Table 2).

The mean fasting blood glucose (FBG) level was **168.7 $\pm$ 54.8 mg/dl**, with values ranging from **72 to 368 mg/dl**. Most participants (**38.5%**) had fasting

Table 2. Distribution of fasting blood glucose levels

FBG (mg/dl)	Number (%)
<100	92 (9.2)
100–139	244 (24.4)
140–199	385 (38.5)
≥ 200	279 (27.9)

Mean $\pm$ SD = **168.7 $\pm$ 54.8 mg/dl**

Median (IQR) = **162 (132–204) mg/dl**

7.67 ± 1.64%. The mean difference between the two methods was 0.02 ± 0.31%.

Comparison of Laboratory-Measured and Formula-Derived hba1c

Page | 7

The mean laboratory-measured hba1c was 7.66 ± 1.39%, whereas the mean formula-derived hba1c was

Paired Student's *t*-test demonstrated **no statistically significant difference** between laboratory-measured hba1c and formula-derived hba1c ($p = 0.284$), indicating close agreement between the two methods (Table 3).

Table 3. Comparison between laboratory-measured hba1c and formula-derived hba1c

Variable	Laboratory hba1c	Formula hba1c	P-value
Mean ± SD (%)	7.66 ± 1.39	7.67 ± 1.64	0.284
Median (IQR)	7.40 (6.50–8.60)	7.50 (6.56–8.72)	
Minimum	4.7	4.8	
Maximum	13.8	13.6	

Classification According to hba1c Categories

between 6.5% and 8.0%, and 296 (29.6%) had hba1c >8.0%.

Based on laboratory hba1c values, 286 participants (28.6%) had hba1c <6.5%, 418 (41.8%) had hba1c

Formula-derived hba1c showed a very similar distribution across these glycaemic categories (Table 4).

Table 4. Distribution of participants according to hba1c categories

Hba1c Category	Laboratory hba1c n (%)	Formula hba1c n (%)
<6.5%	286 (28.6)	279 (27.9)
6.5–8.0%	418 (41.8)	425 (42.5)
>8.0%	296 (29.6)	296 (29.6)

Correlation Between Laboratory and Formula-Derived hba1c

Simple linear regression analysis further demonstrated excellent predictive performance, with an **adjusted R^2 value of 0.918**, indicating that approximately **91.8% of the variability in laboratory hba1c** was explained by the formula-derived hba1c values.

Page | 8

Pearson's correlation analysis demonstrated a **very strong positive correlation** between laboratory-measured hba1c and formula-derived hba1c ($r = 0.958$, $p < 0.001$) (Figure 1).

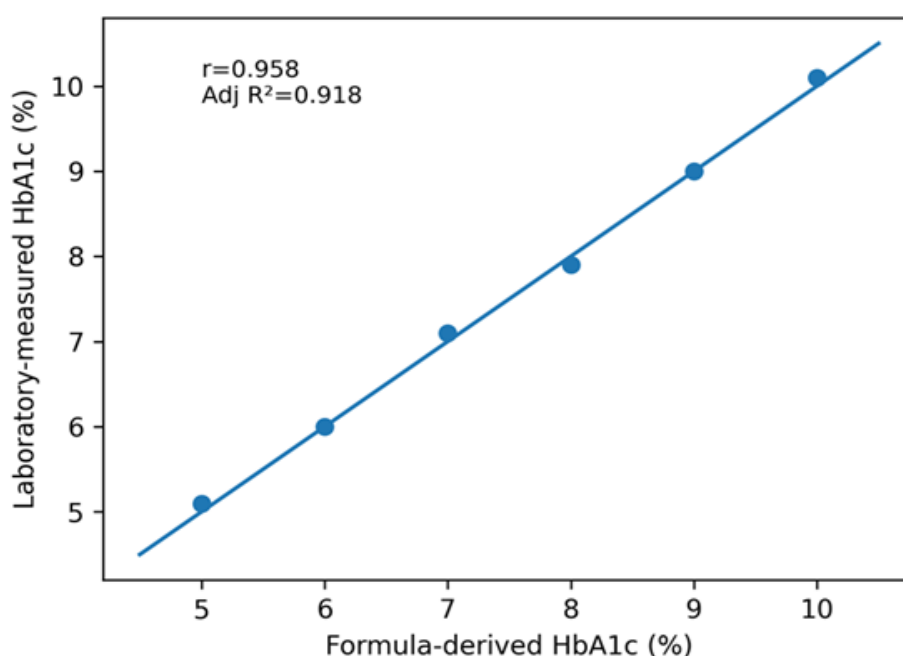


Figure 1: Correlation scatter plot

The mean bias was **0.02%**, with **95% limits of agreement ranging from -0.62% to +0.66%**, indicating minimal systematic error and acceptable clinical agreement (Figure 2).

Agreement Analysis

Page | 9 Bland–Altman analysis demonstrated excellent agreement between laboratory-measured and formula-derived hba1c.

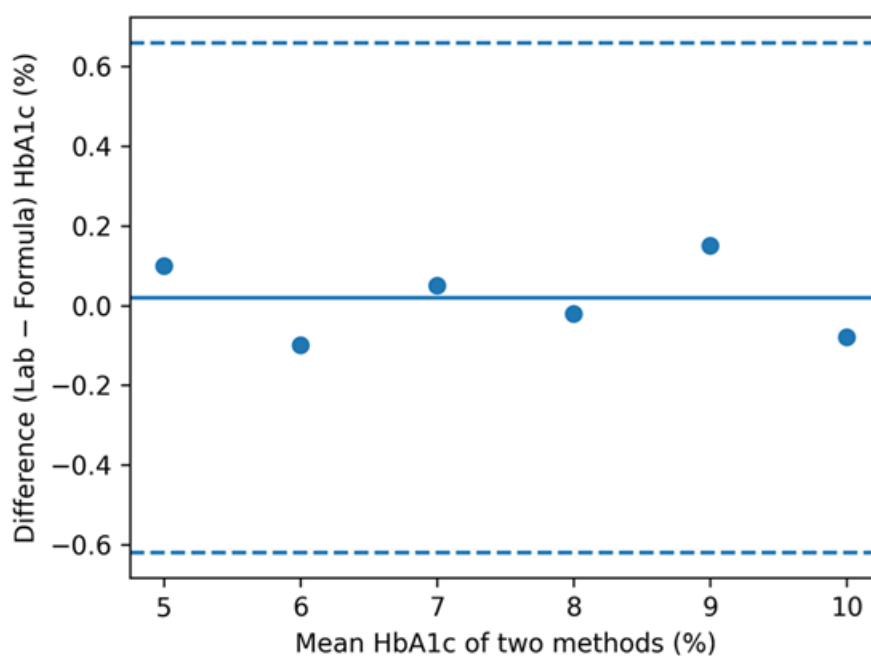


Figure 2: Bland–Altman agreement plot

Reliability Analysis

The Intraclass Correlation Coefficient (ICC) between laboratory-measured hba1c and formula-derived hba1c was **0.962 (95% CI: 0.956–0.967)**, demonstrating **excellent reliability** between the two methods.

Prediction Accuracy

Absolute Percentage Error (MAPE) was **3.18%**, indicating high predictive accuracy of the formula for estimating hba1c (Table 5).

The predictive performance of the mathematical formula was evaluated using standard error metrics. The Mean Absolute Error (MAE) was **0.24%**, the Root Mean Square Error (RMSE) was **0.31%**, and the Mean

Table 5. Prediction accuracy of the mathematical model

Performance Metric	Value
Pearson correlation coefficient (r)	0.958
Adjusted R²	0.918
Mean bias (%)	0.02
95% Limits of Agreement	−0.62 to +0.66
Intraclass Correlation Coefficient	0.962
Mean Absolute Error (MAE)	0.24
Root Mean Square Error (RMSE)	0.31
Mean Absolute Percentage Error (MAPE)	3.18%

Artificial Intelligence Dataset Performance

After data preprocessing, all **1000 records** were complete and suitable for AI model development. The dataset was randomly divided into **800 (80%) training records** and **200 (20%) testing records**. The low prediction error, high correlation coefficient, and excellent agreement between measured and formula-derived hba1c suggest that the generated dataset may be appropriate for training supervised machine learning algorithms aimed at predicting long-term glycaemic control.

Discussion

The present prospective observational study evaluated the validity of a formula-based estimation of glycosylated haemoglobin (hba1c) using fasting blood glucose (FBG) and explored its potential application for Artificial Intelligence (AI) model training. The study

included 1000 participants and demonstrated excellent agreement between laboratory-measured hba1c and formula-derived hba1c. The mathematical model showed high predictive accuracy with a Pearson correlation coefficient of 0.958, an adjusted R² of 0.918, an Intraclass Correlation Coefficient (ICC) of 0.962, and low prediction error metrics, suggesting that formula-derived hba1c may provide a reliable surrogate for laboratory hba1c in appropriate clinical settings.

The mean age of participants in the present study was 53.4 ± 12.6 years, with the highest proportion belonging to the 51–60-year age group. This observation is consistent with the well-recognized increase in the prevalence of type 2 diabetes mellitus with advancing age due to progressive insulin resistance, β-cell dysfunction, and accumulation of metabolic risk factors. Similar age distributions have been reported in epidemiological studies evaluating glycaemic control among adults with diabetes (15,16).

Male participants constituted 57.2% of the study population. Comparable male predominance has been reported in several hospital-based studies, possibly reflecting differences in healthcare-seeking behaviour,

occupational stress, lifestyle factors, and the higher prevalence of cardiovascular risk factors among men attending tertiary care centres (17).

The average fasting blood glucose observed in the cohort was 168.7 ± 54.8 mg/dl, indicating suboptimal glycaemic control in a substantial proportion of participants. The distribution of FBG values corresponded well with the observed laboratory hba1c levels, supporting the physiological relationship between chronic hyperglycaemia and haemoglobin glycation. Persistent elevation of fasting plasma glucose contributes directly to increased hba1c through continuous non-enzymatic glycation of haemoglobin molecules within circulating erythrocytes (18).

One of the major findings of the present study was the remarkable similarity between laboratory-measured hba1c ($7.66 \pm 1.39\%$) and formula-derived hba1c ($7.67 \pm 1.64\%$). The paired analysis demonstrated no statistically significant difference between the two methods, indicating minimal systematic error. These findings are in agreement with the work of Misquith et al., who proposed a mathematical model for estimating hba1c from fasting blood glucose and demonstrated satisfactory agreement between calculated and measured values (19). Likewise, Rajan and colleagues reported that formula-derived hba1c closely approximated laboratory hba1c and could serve as an economical alternative where standardized laboratory facilities are unavailable (20).

The present study demonstrated a very strong positive correlation ($r = 0.958$) between measured and calculated hba1c values. This correlation is comparable to observations reported by Musa et al., who found excellent correlation between measured hba1c and calculated hba1c among patients with type 2 diabetes mellitus, supporting the reliability of mathematical estimation methods (21). Similar findings have also been described in studies evaluating estimated average glucose and hba1c relationships, reinforcing the close physiological association between fasting glucose and long-term glycaemic exposure (22).

Agreement analysis using the Bland–Altman method revealed a mean bias of only 0.02%, with narrow limits of agreement (-0.62% to $+0.66\%$). These findings indicate that the mathematical formula neither consistently overestimates nor underestimates hba1c across the observed measurement range. Bland–Altman analysis is widely accepted for evaluating agreement between two quantitative measurement methods and provides clinically meaningful assessment beyond simple correlation analysis (23).

Reliability assessment demonstrated an Intraclass Correlation Coefficient of 0.962, indicating excellent reproducibility between laboratory and formula-derived hba1c measurements. ICC values above 0.90 are generally considered indicative of excellent agreement and suggest that the mathematical model may be sufficiently reliable for population-level screening and preliminary clinical assessment when laboratory hba1c estimation is not readily available (24).

Prediction error analysis further supported the validity of the mathematical model. The Mean Absolute Error (0.24%), Root Mean Square Error (0.31%), and Mean Absolute Percentage Error (3.18%) were all low, indicating minimal deviation between calculated and laboratory-measured hba1c values. These findings suggest that the proposed equation can estimate hba1c with acceptable analytical accuracy while maintaining computational simplicity.

An important aspect of the present study was the evaluation of formula-derived hba1c for Artificial Intelligence model training. Machine learning algorithms require large, consistent, and accurately labelled datasets to achieve optimal predictive performance. The availability of a validated mathematical model capable of generating reliable hba1c estimates from routinely measured fasting blood glucose may facilitate development of AI-assisted clinical decision support systems, particularly in resource-constrained environments. Formula-derived data may augment existing datasets, improve algorithm robustness, and reduce dependence on expensive laboratory investigations during preliminary model development (25).

Artificial Intelligence has increasingly been incorporated into diabetes care for risk stratification, prediction of glycaemic control, early detection of complications, individualized therapeutic planning, and automated clinical decision support. Recent studies have demonstrated that supervised machine learning algorithms can accurately predict hba1c, diabetic complications, and treatment response using routinely available clinical and biochemical variables (14,26). The present findings support the concept that validated formula-derived hba1c values may serve as an additional structured feature for AI training datasets.

Despite these encouraging findings, formula-based estimation should not be considered a replacement for standardized laboratory hba1c measurement in all clinical situations. Conditions affecting erythrocyte lifespan, including haemoglobinopathies, haemolytic anaemia, chronic kidney disease, recent blood transfusion, and iron deficiency anaemia, may alter hba1c independently of plasma glucose concentration and therefore reduce the accuracy of mathematical prediction models (27). Consequently, laboratory hba1c estimation remains the preferred reference standard for diagnosis and long-term monitoring whenever available.

The strengths of the present study include the large sample size, standardized laboratory methodology, comprehensive agreement analysis, and incorporation of AI-oriented validation metrics such as MAE, RMSE, MAPE, and ICC. Evaluation of multiple statistical measures provides a comprehensive assessment of analytical performance rather than reliance on correlation alone.

Limitations

The study also has several limitations. First, this was a single-centre study conducted at a tertiary care teaching hospital, which may limit the generalizability of the findings to other healthcare settings and populations. Second, only fasting blood glucose was incorporated into the mathematical model, whereas additional clinical variables such as age, body mass index, duration of diabetes, renal function, and haemoglobin concentration may further improve predictive accuracy. Third, although the formula demonstrated excellent agreement with laboratory-measured hba1c in the present study, external validation in independent populations from different geographical regions is necessary before widespread clinical implementation.

Conclusion

The present prospective observational study demonstrated that the formula-based estimation of glycosylated haemoglobin (hba1c) showed excellent agreement with laboratory-measured hba1c among the cohort of 1000 participants. The mathematical formula, **$\text{hba1c (\%)} = 2.6 + 0.03 \times \text{Fasting Blood Glucose (mg/dl)}$** , produced hba1c values that closely approximated those obtained by High Performance Liquid Chromatography (HPLC), with a very strong positive correlation, excellent reliability, and minimal prediction error.

The study findings suggest that formula-derived hba1c has the potential to serve as a simple, rapid, and cost-effective alternative for estimating long-term glycaemic control in situations where standardized laboratory hba1c estimation is unavailable or inaccessible. The low values of Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), together with the high Intraclass Correlation Coefficient (ICC), indicate satisfactory analytical performance of the mathematical model.

Furthermore, the validated formula-generated hba1c dataset demonstrated characteristics suitable for Artificial Intelligence (AI) model training. The availability of reliable formula-derived hba1c values may facilitate the development of machine learning algorithms for diabetes screening, prediction of glycaemic control, and clinical decision-support systems, particularly in resource-constrained healthcare settings.

Although the findings are promising, validation using large multicentre real-world datasets representing diverse populations remains essential before routine clinical implementation. Formula-based hba1c estimation should currently be considered a complementary tool rather than a replacement for standardized laboratory hba1c measurement.

Recommendations

Based on the findings of the present study, the following recommendations are proposed:

1. Large multicentre prospective studies should be conducted to validate the mathematical formula across different geographical regions and ethnic populations.
2. Future studies should evaluate the performance of the formula in patients with type 1 diabetes mellitus, type 2 diabetes mellitus, gestational diabetes, and prediabetes.
3. Additional clinical and biochemical variables, including age, body mass index, haemoglobin

concentration, renal function, lipid profile, and duration of diabetes, may be incorporated to improve prediction accuracy.

4. Formula-derived hba1c may be integrated into Artificial Intelligence and machine learning models to support automated diabetes screening and clinical decision-making after appropriate clinical validation.
5. Mobile health applications and electronic health record systems may incorporate validated mathematical models to provide rapid estimation of hba1c in resource-limited settings.
6. Standard laboratory hba1c estimation should continue to remain the reference method for diagnosis and long-term monitoring until sufficient multicentre validation of formula-based approaches has been completed.

Acknowledgement

The authors express their sincere gratitude to the Department of Biochemistry, Sri Devaraj Urs Medical College, Tamaka, Kolar, Karnataka, India, and RL Jalappa Hospital for providing the facilities, institutional support, and laboratory infrastructure necessary for conducting this study. The authors also acknowledge the valuable contributions of the faculty members, laboratory technicians, nursing staff, and all study participants for their cooperation and support throughout the study. The authors sincerely acknowledge Dr. Harish for his valuable assistance in the laboratory conceptualization of the study and for his intellectual contribution and critical revision of the manuscript.

List of abbreviations

AI	Artificial Intelligence
BMI	Body Mass Index
DCCT	Diabetes Control and Complications Trial
FBG	Fasting Blood Glucose
GOD-POD	Glucose Oxidase–Peroxidase
HbA1c	Glycosylated Haemoglobin
HPLC	High Performance Liquid Chromatography
ICC	Intraclass Correlation Coefficient
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NGSP	National Glycohemoglobin Standardization Program
RMSE	Root Mean Square Error
SD	Standard Deviation
SDUMC	Sri Devaraj Urs Medical College

Source of funding

No funding involved

Dr. Samudyatha UC: Study planning, statistical guidance, interpretation of findings, manuscript editing, and final approval.

All authors have read and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Conflict of interest

The authors declare **no conflict of interest** related to this research work.

Author contributions

Dr. Satydeo Kumar Singh: Conceptualization of the study, study design, data acquisition, statistical analysis, interpretation of results, manuscript preparation, and final approval.

Dr. Prabhavathi K: Study supervision, methodology development, critical revision of the manuscript, and final approval.

Author biography

Dr. Satydeo Kumar Singh is a final year postgraduate student in the Department of Biochemistry at Sri Devaraj Urs Medical College, Tamaka, Kolar, Karnataka, India. His academic interests include clinical biochemistry, diabetes research, biomarker validation, and applications of Artificial Intelligence in laboratory medicine.

Dr. Prabhavathi K is Professor and Head of the Department of Biochemistry at Sri Devaraj Urs

Medical College. Her research interests include diabetes mellitus, metabolic disorders, clinical laboratory quality assurance, and translational medical research.

Page | 15

Dr. Samudyatha UC is Associate Professor in the Department of Community Medicine, Sri Devaraj Urs Medical College. Her areas of interest include epidemiology, biostatistics, preventive medicine, public health research, and clinical study design.

Data availability

References

1. International Diabetes Federation. IDF Diabetes Atlas. 10th ed. Brussels: International Diabetes Federation; 2021.
2. Deepa M, Elangovan N, Venkatesan U, Das HK, Jampa L, Adhikari P, et al.; ICMR-INDIAB Collaborative Study Group. Evaluation of Madras Diabetes Research Foundation-Indian Diabetes Risk Score in detecting undiagnosed diabetes in the Indian population: results from the Indian Council of Medical Research-india diabetes population-based study (INDIAB-15). *Indian J Med Res*. 2023;157(4):239-249. Doi:10.4103/ijmr.ijmr_2615_21. https://doi.org/10.4103/ijmr.ijmr_2615_21
3. American Diabetes Association Professional Practice Committee. 2. Diagnosis and classification of diabetes: Standards of Care in Diabetes-2025. *Diabetes Care*. 2025;48(Suppl 1):S27-S49. Doi:10.2337/dc25-S002. <https://doi.org/10.2337/dc25-S002>
4. World Health Organization. Global report on diabetes. Geneva: World Health Organization; 2016.
5. Campbell L, Pepper T, Shipman K. Hba1c: a review of non-glycaemic variables. *J Clin Pathol*. 2019;72(1):12-19. Doi:10.1136/jclinpath-2017-204755. <https://doi.org/10.1136/jclinpath-2017-204755>
6. Goldstein DE, Little RR, Lorenz RA, Malone JI, Nathan D, Peterson CM, et al. Tests of glycemia in diabetes. *Diabetes Care*. 2004;27(7):1761-1773. Doi:10.2337/diacare.27.7.1761. <https://doi.org/10.2337/diacare.27.7.1761>
7. Little RR, Rohlfing CL, Hanson S, Connolly S, Higgins T, Weykamp C, et al. The National Glycohemoglobin Standardization Program: over 20 years of improving hba1c measurement. *Clin Chem*. 2019;65(7):839-848. Doi:10.1373/clinchem.2018.296962. <https://doi.org/10.1373/clinchem.2018.296962>
8. Leters-Westra E, Slingerland RJ. Six of eight hemoglobin A1c point-of-care instruments do not meet the generally accepted analytical performance criteria. *Clin Chem*. 2010;56(1):44-52. <https://doi.org/10.1373/clinchem.2009.130641>
9. Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of hba1c test in diagnosis and prognosis of diabetic patients. *Biomark Insights*. 2016;11:95-104. <https://doi.org/10.4137/BMI.S38440>
10. Bonora E, Tuomilehto J. The pros and cons of diagnosing diabetes with A1C. *Diabetes Care*. 2011;34(Suppl 2):S184-S190. <https://doi.org/10.2337/dc11-s216>
11. Higgins T. Hba1c for screening and diagnosis of diabetes mellitus. *Endocrine*. 2013;43(2):266-273. <https://doi.org/10.1007/s12020-012-9768-y>
12. Cohen RM, Lindsell CJ. When the blood glucose and hba1c do not match: turning uncertainty into opportunity. *Diabetes Care*. 2012;35(12):2421-2423. <https://doi.org/10.2337/dc12-1479>
13. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230-243. <https://doi.org/10.1136/svn-2017-000101>
14. Ellahham S, Ellahham N, Simsekler MCE. Application of artificial intelligence in diabetes mellitus: current status and future

- prospect. World J Diabetes. 2020;11(4):163-176.
15. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care*. 2004;27(5):1047-1053. Doi:10.2337/diacare.27.5.1047. <https://doi.org/10.2337/diacare.27.5.1047>
16. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract*. 2022;183:109119. <https://doi.org/10.1016/j.diabres.2021.109119>
17. Pradeepa R, Mohan V. Epidemiology of type 2 diabetes in India. *Indian J Ophthalmol*. 2021;69(11):2932-2938. Doi:10.4103/ijoo.IJO_1627_21. https://doi.org/10.4103/ijoo.IJO_1627_21
18. Monnier L, Lapinski H, Colette C. Contributions of fasting and postprandial plasma glucose increments to the overall diurnal hyperglycemia of type 2 diabetic patients: variations with increasing levels of hba1c. *Diabetes Care*. 2003;26(3):881-885. Doi:10.2337/diacare.26.3.881. <https://doi.org/10.2337/diacare.26.3.881>
19. Misquith A, Rangareddy H, Chikkanarayanappa V, Sultana A, Ashakiran S. Mathematical model for assessing glycemic control in type 2 diabetes mellitus. *Bioinformation*. 2024;20(2):116-120. Doi:10.6026/973206300200116. <https://doi.org/10.6026/973206300200116>
20. Rajan SS, Misquith A, Rangareddy H. Calculated glycosylated hemoglobin (hba1c) compared with estimated hba1c by nephelometry and its correlation to estimated average blood glucose (eag). *Gal Int J Health Sci Res*. 2020;5(4):65-71.
21. Musa IR, Omar SM, Sharif ME, Ahmed ABA, Adam I. The calculated versus the measured glycosylated haemoglobin (hba1c) levels in patients with type 2 diabetes mellitus. *J Clin Lab Anal*. 2021;35(8):e23873. Doi:10.1002/jcla.23873. <https://doi.org/10.1002/jcla.23873>
22. Nathan DM, Kuenen J, Borg R, Zheng H, Schoenfeld D, Heine RJ; a1c-Derived Average Glucose (ADAG) Study Group. Translating the A1C assay into estimated average glucose values. *Diabetes Care*. 2008;31(8):1473-1478. <https://doi.org/10.2337/dc08-0545>
23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
24. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
25. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56. Doi:10.1038/s41591-018-0300-7. <https://doi.org/10.1038/s41591-018-0300-7>
26. Zhu T, Li K, Herrero P, Georgiou P. Deep learning for diabetes: a systematic review. *IEEE J Biomed Health Inform*. 2021;25(7):2744-2757. Doi:10.1109/JBHI.2020.3040225. <https://doi.org/10.1109/JBHI.2020.3040225>
27. English E, Idris I, Smith G, Dhatariya K, Kilpatrick ES, John WG. The effect of anaemia and abnormalities of erythrocyte indices on hba1c analysis: a systematic review. *Diabetologia*. 2015;58(7):1409-1421. Doi:10.1007/s00125-015-3599-3. <https://doi.org/10.1007/s00125-015-3599-3>



Student's Journal of Health Research Africa

e-ISSN: 2709-9997, p-ISSN: 3006-1059

Vol.7 No. 2 (2026): June 2026 Issue

<https://doi.org/10.51168/8n2swj68>

Original Article

PUBLISHER DETAILS

Student's Journal of Health Research (SJHR)

(ISSN 2709-9997) Online

(ISSN 3006-1059) Print

Category: Non-Governmental & Non-profit Organization

Email: studentsjournal2020@gmail.com

WhatsApp: +256 775 434 261

**Location: Scholar's Summit Nakigalala, P. O. Box 701432,
Entebbe Uganda, East Africa**

