FHIMA, Vol 1, Issue 1, July-Sept- 2025

Contents lists available at Curevita Journals

Frontiers of Health Innovations and Medical Advances

journal homepage: www.curevitajournals.com

Data Intelligence Tools in Diabetes Mellitus: Applications, Methods, and Future Directions

Kailash Prasad Jaiswal

Department of Biotechnology, Bhabha University, Bhopal, MP, India

Articalinfo

Article history: Received 22 June 2025, Revised 18 Aug 2025, Accepted 20 Aug 2025, Published Sept 2025

Keywords: Diabetes mellitus; machine learning; artificial intelligence; digital health; continuous glucose monitoring; clinical decision support; population health.

Corresponding Author: Dr KP Jaiswal
Email:kailashjaiswal6@gmail.com

Citation: Jaiswal Kailash Prasad. 2025.Data Intelligence Tools in Diabetes Mellitus: Applications, Methods, and Future Directions.Curevita Innovation of BioData Intelligence 1,1,13-23..

Publisher: Curevita Research Pvt Ltd

Abstract

Diabetes mellitus (DM) is a heterogeneous metabolic disorder characterized by chronic hyperglycemia resulting from defects in insulin secretion, insulin action, or both. Globally, the prevalence and economic burden of DM continue to rise, necessitating new approaches for prevention, diagnosis, clinical management, and health-system planning. Data intelligence (DI)—the integrated use of data engineering, analytics, machine learning (ML), and artificial intelligence (AI) within robust socio-technical systems—has emerged as a transformative enabler across the diabetes continuum of care. We synthesize evidence on key application domains, highlight clinical and operational outcomes reported to date, and analyze barriers related to data quality, algorithmic bias, privacy, interoperability, and real-world implementation. We propose a pragmatic evaluation framework and a research roadmap focused on explainability, causal inference, hybrid mechanistic–ML models, and equitable deployment at scale.

Introduction

Diabetes mellitus (DM) encompasses a spectrum of disorders—including type 1 diabetes (T1D), type 2 diabetes (T2D), gestational diabetes, and monogenic forms—defined by dysregulated glucose homeostasis. Despite advances in pharmacotherapy and technology (e.g., CGM, insulin pumps), many individuals fail to achieve glycemic targets, and health systems face mounting costs due to complications. Concurrently, the digitization of health care has produced unprecedented volumes of multi-modal data: EHRs, claims, laboratory and imaging repositories, pharmacy data, wearable and sensor streams, patient-reported outcomes, and social determinants of health. Data intelligence (DI) leverages these assets to generate insights, guide decisions, and automate or augment clinical workflows.

This paper critically reviews DI tools across the diabetes care continuum, spanning prevention, diagnosis, acute and chronic management, and system-level optimization. We categorize tools by data source and analytic approach, map them to clinical/operational outcomes, and propose methods to evaluate safety, efficacy, and equity. Finally, we outline a research agenda for the next generation of trustworthy, human-centered AI in diabetes.

Data Ecosystem for Diabetes Intelligence

Data Sources

- **Clinical records:** Structured (diagnoses, vitals, medications, labs) and unstructured (progress notes, discharge summaries) EHR data.

- **Laboratory and imaging:** HbA1c, lipid profiles, renal function; retinal fundus images, OCT, CT angiography.
- **Devices and wearables:** CGM time-series, insulin pump logs, smart pens, fitness trackers, smart scales, blood pressure cuffs.
- **Patient-reported outcomes:** Symptom diaries, food logs, mood scores, pain/fatigue scales, QoL inventories.
- **Administrative and payer data:** Claims, authorizations, cost/utilization.
- **Genomic, proteomic, and metabolomic data:** For subtype discovery, pharmacogenomics, and precision nutrition.
- **Social determinants and environmental data:** Neighborhood deprivation indices, food deserts, walkability, air pollution, weather/seasonality.
- **Public health registries:** Diabetes registries, mortality data, vaccination status.

Data Engineering and Governance

- **Interoperability:** HL7 FHIR, OMOP CDM, DICOM for imaging, IEEE 11073 for device data.
- **Data quality:** Missingness handling, outlier detection, sensor drift correction, unit harmonization.
- **Identity resolution:** Patient matching across systems; privacy-preserving record linkage.
- **Streaming infrastructure:** Time-series ingestion from CGM/pumps via MQTT/Kafka;

edge analytics on mobile devices.

- **Security and privacy:** Role-based access, differential privacy, secure enclaves, de-identification, audit trails, and data use agreements.
- **Ethics and governance:** Data sharing governance, consent management, Indigenous data sovereignty, algorithmic impact assessments.

Analytic Paradigms and Model Classes

Supervised Learning

- **Risk prediction:** Incident diabetes, progression to insulin, hospitalization, hypoglycemia, ketoacidosis, and complications (retinopathy, nephropathy, neuropathy, CVD).
- **Outcome modeling:** HbA1c improvement, time-in-range (TIR), weight loss, medication adherence. Algorithms include regularized regression, tree ensembles (RF/XGBoost), and neural networks.

Unsupervised and Self-Supervised Learning

- **Phenotyping and subtyping:** Clustering of T2D into pathophysiologic subgroups; anomaly detection for rare events.
- **Representation learning:** Self-supervised encoders for CGM sequences or fundus images.

Deep Learning for Images and Signals

- **Computer vision:** DL models for diabetic retinopathy grading, macular edema detection, foot ulcer identification and segmentation.

- **Signal modeling:** Temporal CNNs, LSTMs, and Transformers for CGM, insulin, meals, and activity sequences.

Causal Inference and Uplift Modeling

- **Questions of comparative effectiveness:** What would happen under metformin vs. SGLT2 inhibitor initiation?
- **Methods:** Propensity scores, doubly robust estimators, causal forests, target trial emulation, instrumental variables.

Reinforcement Learning and Control

- **Insulin dosing and closed-loop control:** RL policies for basal-bolus optimization and AP systems; safety layers (model predictive control, constraints).
- **Behavioral nudging:** Context-aware RL for timing and content of digital coaching messages.

Hybrid and Mechanistic–ML Models

- **Physiological simulators + ML:** Use of minimal models of glucose–insulin kinetics coupled with ML for parameter personalization; digital twins for scenario testing.

Privacy-Preserving ML

- **Federated learning (FL):** Cross-institutional training without centralizing data.

- **Secure aggregation and differential privacy:** Mitigate leakage and membership inference attacks.

Applications Across the Diabetes Continuum

Prevention and Early Detection

- **Screening models** for undiagnosed diabetes using routine labs and vitals; opportunistic case-finding in dental/ophthalmology clinics.
- **Pre-diabetes progression prediction** and personalized lifestyle intervention targeting.
- **Risk calculators** embedded in primary care EHRs with actionable care pathways.

Diagnosis and Classification

- **Automated phenotyping** to distinguish T1D vs. T2D vs. LADA using labs (C-peptide, autoantibodies), age, BMI, and genetics.
- **Gestational diabetes** early risk prediction from first-trimester data.
- **Monogenic diabetes detection** via variant prioritization and clinical rules.

Glycemic Management

- **Decision support for insulin titration:** Algorithms recommending basal/bolus adjustments based on SMBG/CGM, carbohydrate intake, and activity.
- **Closed-loop/Artificial Pancreas (AP):** ML-enhanced control to improve TIR and reduce hypoglycemia.

- **Medication optimization:** Recommendations for antihyperglycemics considering comorbidities, eGFR, CVD risk, cost, and patient preference.
- **Nutrition analytics:** Food image recognition, macronutrient estimation, and glycemic impact prediction; personalized meal planning.
- **Behavioral and adherence analytics:** Predicting disengagement, micro-interventions, and digital therapeutics that adapt to context.

Complication Surveillance and Triage

- **Retinopathy:** DL for automated grading of fundus images and triage (refer vs. monitor).
- **Nephropathy:** Risk models for eGFR decline and albuminuria progression.
- **Neuropathy and foot ulcers:** Computer vision for ulcer detection, thermal imaging analytics, and remote monitoring.
- **Cardiovascular risk:** Integrated models combining labs, vitals, ECG signals, and imaging-derived features.

Acute Event Prediction and Safety

- **Hypo-/hyperglycemia forecasting** from CGM streams (minutes to hours ahead) to trigger alarms and proactive interventions.
- **DKA risk prediction** in T1D, especially in youth and during illness or pump failure.
- **Hospital care:** Predicting insulin requirements, steroid-induced hyperglycemia, and perioperative glycemic instability.

Population Health and Health-System Operations

- **Registry analytics:** Gap-in-care detection (overdue HbA1c, retinal exam).
- **Resource optimization:** Clinic capacity planning, remote monitoring thresholds, outreach prioritization.
- **Equity dashboards:** Stratified performance (TIR, HbA1c control) by demographics and SDoH to guide targeted interventions.
- **Cost and value analytics:** Budget impact, return on investment, and value-based care metrics.

Natural Language Processing (NLP)

- **Info extraction:** Automatic capture of complications, lifestyle factors, and adverse events from clinical notes.
- **Conversational agents:** Triage, education, and self-management support; escalation to clinicians when needed.

India-Focused Perspectives: Epidemiology, Digital Health Initiatives, and LMIC Deployment

Epidemiology in India

India is home to more than 101 million people living with diabetes, representing one of the largest affected populations globally. Rapid urbanization, lifestyle transitions, and genetic predisposition contribute to the increasing prevalence. Rural–urban disparities persist, with underdiagnosis in rural areas and a high burden of complications in underserved communities.

Gestational diabetes is also increasing, posing intergenerational risks.

Digital Health Initiatives in India

- **National Digital Health Mission (NDHM):** Envisions longitudinal electronic health records and interoperable data exchange, creating an enabling ecosystem for DI-driven diabetes care.
- **Ayushman Bharat Digital Mission (ABDM):** Links health facilities, insurance, and patient records, facilitating real-world analytics and population health management.
- **mHealth and telemedicine platforms:** Government-supported eSanjeevani and private apps enable remote consultations, CGM data sharing, and adherence tracking.
- **Public–private partnerships:** Deployment of AI-enabled retinal screening tools in primary health centers and mobile vans to expand coverage in semi-urban and rural regions.

LMIC Deployment Considerations

- **Infrastructure constraints:** Limited internet connectivity and device penetration require lightweight, on-device intelligence and offline-first solutions.
- **Affordability:** Low-cost CGM, SMS-based reminders, and AI

chatbots in regional languages improve accessibility.

- **Equity:** Prioritizing inclusive datasets, accounting for dietary diversity, socioeconomic determinants, and cultural practices.
- **Capacity building:** Training healthcare workers in rural areas to use DI dashboards and mobile applications.
- **Policy and regulation:** India's evolving SaMD regulatory framework and data protection bill will shape responsible DI deployment.

Conclusion

Data intelligence tools are reshaping diabetes care worldwide, and India exemplifies both the challenges and opportunities in large, diverse populations. By leveraging national digital health initiatives, cost-sensitive innovations, and equitable AI design, DI tools can bridge gaps in diagnosis, management, and complication prevention in low- and middle-income country contexts. To maximize impact, stakeholders must prioritize trustworthy, inclusive, and scalable DI solutions integrated into existing health systems. As the field matures, collaborative efforts between clinicians, researchers, policymakers, and technology developers will be essential to ensure that DI transforms diabetes care into a more precise, proactive, and patient-centered discipline.

Reference:

Wu, Y.-Y., Xiao, E. & Graves, D. T. (2015).Diabetes mellitus related bone metabolism and periodontal disease. Int. J. Oral Sci. 7, 63–72.

Vashist, P., Singh, S., Gupta, N. & Saxena, R. (2011).Role of early screening for diabetic retinopathy in patients with diabetes mellitus: An overview. Indian J. Community Med. 36, 247 .

Zakir, M. et al. (2023)Cardiovascular complications of diabetes: From microvascular to macrovascular pathways. Cureus 15, e45835

Cho, N. H. et al. IDF diabetes atlas: (2018).Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res. Clin. Pract. 138, 271–281

Kyrou, I. et al. (2020). Sociodemographic and lifestyle-related risk factors for identifying vulnerable groups for type 2 diabetes: A narrative review with emphasis on data from Europe. BMC Endocr. Disord. 20, 1–13

Charbuty, B. & Abdulazeez, A. (2021).Classification based on decision tree algorithm for machine learning. J. Appl. Sci. Technol. Trends 2, 20–28

Kavakiotis, I. et al. (2017). Machine learning and data mining methods in diabetes research. Comput. Struct. Biotechnol. J. 15, 104–116 .