

# Path2Prevention an Explainable Web Platform for Personalized Type 2 Diabetes Risk Prevention

## Problem Driven Design, Transparent Risk Scoring, and Rule Based Guidance for Preventive Self Management

Anupamkumar Kushwaha, Riya Yadav, Asmita Chauhan, Arushi Singh, Prof. Vaishali Rane

Department of Computer Engineering,

Thakur Shyamnarayan Engineering College, Mumbai, India

**Abstract** - Type 2 diabetes prevention needs digital systems that do more than collect isolated health logs. Many existing tools provide generic advice, limited transparency, and weak linkage between family history, behaviour, and next actions. This paper presents Path2Prevention, an explainable web platform engineered to operationalize prevention evidence in a user facing workflow. The implemented system combines guided onboarding, daily lifestyle logging, a progressive web application frontend, a Node.js and Express backend, MongoDB persistence, and a deterministic pipeline that transforms recent logs, baseline profile data, family history, and optional laboratory values into a visible 0 to 100 risk score. The method computes rolling metrics, applies South Asian sensitive anthropometric thresholds, weights family history, evaluates rule based recommendations, triggers safety overrides for diabetes range laboratory values, estimates behaviour correlations, projects short horizon risk trajectory, and updates dynamic goals. The paper contributes a clearly specified full stack architecture, an interpretable risk and recommendation engine, and scenario based functional validation showing how onboarding estimates, daily logs, auto resolution, and referral alerts work end to end. Rather than claiming clinical efficacy, the study provides an engineering foundation for pilot deployment and future longitudinal validation of transparent digital diabetes prevention.

**Keywords** - Type 2 diabetes prevention; explainable risk scoring; digital health; personalized recommendations; web application; prediabetes.

### I. INTRODUCTION

Diabetes and prediabetes continue to expand at a scale that makes prevention a systems problem as much as a clinical one. The IDF Diabetes Atlas reports 589 million adults aged 20 to 79 years living with diabetes in 2024, while the ICMR INDIAB national study estimated roughly 101 million people with diabetes and 136 million with prediabetes in India [15], [14]. Prediabetes is commonly identified through impaired fasting glucose, impaired glucose tolerance, or HbA1c in the 5.7 to 6.4 percent range [9]. These figures make scalable prevention support a public health and engineering priority.

The case for prevention is already strong. The Diabetes Prevention Program showed that lifestyle intervention built around at least 150 min per week of moderate activity and approximately 7 percent weight loss reduced diabetes incidence by 58 percent and outperformed metformin in adults at high risk

[1]. Similar prevention effects were reported in the Finnish Diabetes Prevention Study and the Da Qing trial, establishing that structured behaviour change can delay or prevent progression to type 2 diabetes [2], [3].

Digital delivery has attempted to bring these targets into everyday life through apps, web programs, messaging, and blended interventions. Trials report improvements in participation, steps, and weight change, yet review evidence remains mixed for long term glycaemic outcomes and often notes limited personalization and limited explanation of why a user is receiving a specific message [4]-[8]. This is especially problematic for prevention, where trust depends on whether users can connect their own habits and family history to understandable next steps.

Path2Prevention was conceived to bridge that gap. The platform combines guided onboarding, daily logging, a transparent risk score, prioritized recommendations, insights, and export features inside one explainable website. The contribution of this paper is not a claim of clinical efficacy. Instead, the paper presents the design rationale, architecture, risk and recommendation method, and functional validation logic of the implemented system so that the engineering approach can be inspected, critiqued, and extended.

### II. RELATED WORK AND RESEARCH GAP

#### A. Lifestyle prevention evidence and digital translation

Landmark prevention trials established the behavioural targets that still underpin most diabetes prevention programs. In the DPP, participants were coached toward weekly moderate activity and clinically meaningful weight loss; comparable findings in Finland and China reinforced the idea that early lifestyle intervention can outperform delayed treatment [1]-[3]. These studies provide the clinical backbone for any digital prevention platform that claims to be evidence based.

Digital and mobile translation has focused on portability, engagement, and self monitoring. Fukuoka et al. demonstrated that a mobile app intervention could support short term changes in weight and step count among adults at risk [4]. Later digital diabetes prevention work such as Transform and real world implementation studies among older adults further showed that structured prevention content can be adapted into web or app mediated formats [5], [6].

**B. Persistent gaps in existing digital prevention tools**

Review studies provide a more cautious picture. Smartphone and mobile health interventions are promising, but the strength of evidence is inconsistent across weight, glycaemic, and adherence outcomes, particularly in prediabetes and older cohorts [7], [8]. Many published systems also emphasize behaviour change outcomes without describing the underlying software and decision logic in a way that another engineering team can inspect or reproduce.

From an implementation perspective, four gaps recur. First, many systems collect behaviour but do not decompose risk into understandable contributing factors. Second, personalization is often limited to reminders or generic education rather than active rule based prioritization. Third, family history is captured but seldom used as an operational modifier of urgency. Fourth, culturally relevant adiposity thresholds and safety handling for diabetes range laboratory values are not consistently embedded in the user facing workflow. Path2Prevention was designed directly against these gaps.

**Path2Prevention Problem Framing and Design Response**

How the project moved from a practical prevention gap to an explainable engineering approach

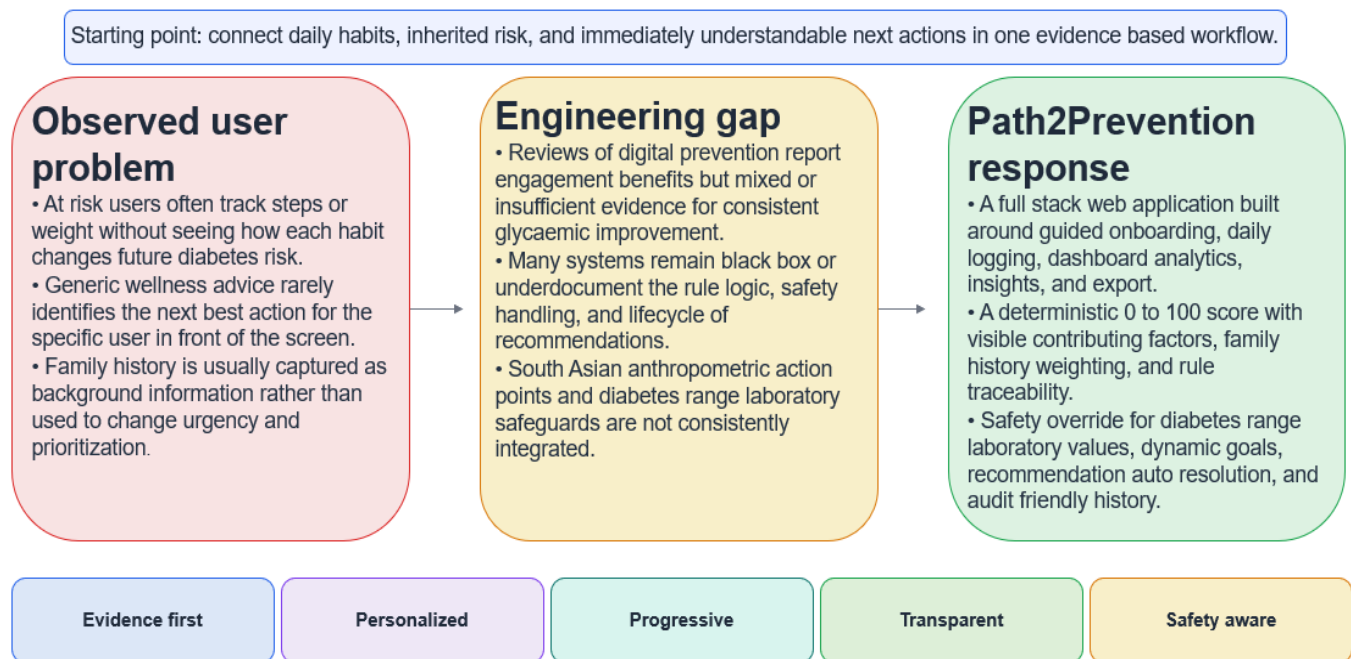


Fig. 1. Problem framing, engineering gap, and project response that motivated the development of Path2Prevention.

**III. PROJECT STARTING POINT, DESIGN OBJECTIVES, AND CONTRIBUTIONS**

**A. Why the project was started**

The project began from a practical observation visible across generic wellness products: users can count steps, sleep, or weight, but they still struggle to understand whether their overall prevention status is improving and what they should do next. The internal project documentation frames this gap in concrete terms: people do not know their risk level, daily habits are not linked to long term risk, family history creates anxiety without guidance, and generic advice produces a weak feedback loop between data and action. Path2Prevention therefore started as an attempt to convert prevention evidence into a transparent, daily use web workflow rather than another passive tracking tool.

The intended use case is preventive self management rather than diagnosis. The system is aimed at adults who want to

understand modifiable type 2 diabetes risk, with particular attention to users whose risk profile is influenced by family history, rising body mass index, sedentary behaviour, diet quality, or optional laboratory values. This starting point shaped both the user experience and the backend architecture.

**B. Design objectives and engineering requirements**

Five design principles guided implementation: evidence first, personalized not generic, progressive, non judgmental, and transparent. In engineering terms this translated into several concrete requirements: the score had to be explainable, recommendations had to change with individual data, safety handling had to override routine coaching when laboratory values entered diabetes range, and the system had to preserve sufficient history for auditability and later evaluation. Another non trivial requirement was immediate usefulness for new users, which led to onboarding estimates based on lifestyle snapshot data before longitudinal logging exists.

Requirement	Why it matters	Implemented response
Explainability	Users need to understand why risk changed and why a recommendation appeared.	Visible factor by factor score breakdown, interpretable bands, and explicit rule based recommendation reasons.
Personalization	Prevention advice should reflect inherited risk, behaviour patterns, and current context.	Family history weighting, category priorities, dynamic goals, and one active recommendation per category.
Safety	A prevention app must not hide evidence suggestive of established diabetes.	Fasting glucose and HbA1c override logic that suppresses routine coaching and issues a medical review alert.
Traceability	Engineering evaluation requires auditable state changes over time.	Append only risk score history, locked daily logs, and versioned rules that can be updated without rewriting the core engine.
Immediate usability	New users should not face an empty dashboard before enough logs are collected.	Lifestyle snapshot based onboarding estimate that initializes score, risk band, and priority cards from day one.

TABLE I. KEY ENGINEERING REQUIREMENTS AND SYSTEM RESPONSE

### C. Contributions of this paper

This paper makes four contributions. First, it documents a complete full stack prevention architecture rather than only a user interface or behaviour study. Second, it specifies a deterministic risk and recommendation method that remains inspectable at the level of metrics, factor caps, and rule evaluation. Third, it integrates culturally relevant South Asian adiposity thresholds, family history weighting, and laboratory safety override behaviour inside one coherent workflow. Fourth, it presents scenario based functional validation that demonstrates how onboarding estimates, daily recomputation, auto resolution, and dashboard synthesis operate in the implemented system.

## IV. SYSTEM ARCHITECTURE AND IMPLEMENTATION

### A. Full stack architecture

Path2Prevention is implemented as a browser based progressive web application with a Node.js and Express backend and MongoDB Atlas persistence. The client layer supports registration, onboarding, daily logging, dashboard viewing, insights, and export. On the backend, route modules manage authentication, profile and settings updates, logs, recommendations, dashboard assembly, insights, engagement, content delivery, and report generation. Requests pass through JWT authentication and validation before interacting with the computation layer.

The architecture was chosen to support an end to end prevention journey rather than a single use screen. A new user can register, provide consent, complete onboarding, receive a first estimate, return for daily logging, inspect trend changes, simulate behaviour changes, and export reports without leaving the same application context. This continuity is important

because prevention depends on repeated interactions rather than one time screening.

### B. Data model, persistence, and security decisions

The documented schema spans 21 collections grouped into user data, computed data, and seeded content. User facing collections store accounts, profiles, settings, daily logs, weekly measures, streaks, and related personal records. Computed collections store append only risk scores, recommendation states, goals, trajectory records, and correlation snapshots. Seeded collections store evidence sources, tips, educational content, food items, activity guides, and rule versions. This separation keeps mutable personal data distinct from generated analytics and reusable content.

Several persistence decisions directly support traceability. Daily logs are locked after saving so that historical score computation is not silently altered later. Risk scores are stored in append only form instead of being overwritten, which preserves the basis for first score versus current score comparisons and trajectory analysis. Recommendation state tracks active, snoozed, and resolved conditions. The service layer is protected with helmet headers, CORS controls, rate limiting, mongo sanitize middleware, JWT guards, and server side input validation.

### C. Operational workflow from onboarding to insights

After account creation, the user completes a guided onboarding flow covering demographics, body metrics, family history, a lifestyle snapshot, optional laboratory values, and consent. Once onboarding is complete, the application can generate an initial estimate even when no daily logs yet exist. Later, each saved daily log updates rolling metrics, risk score, recommendations, goals, and the data assembled for dashboard and insights views. Export services transform the same history into CSV and PDF outputs, while the what if simulator reuses the risk logic without committing new records.

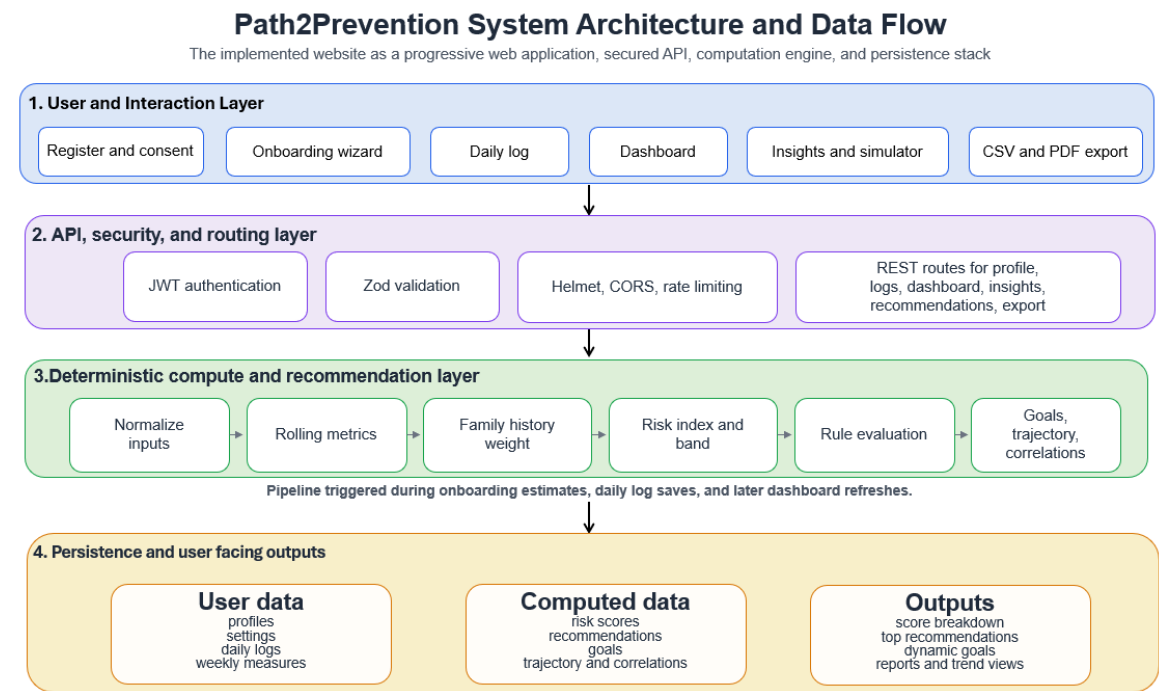


Fig. 2. Full stack architecture and end to end data flow across the user interface, secured API layer, deterministic compute engine, persistence modules, and user facing outputs.

## V. EXPLAINABLE RISK AND RECOMMENDATION METHOD

### A. Pipeline execution stages

The core computation engine is a deterministic pipeline that executes during onboarding estimation, daily log submission, and later dashboard refresh. In the current implementation the sequence can be summarized as ten stages: input normalization, rolling metric computation, family history weighting, risk index construction, band mapping, recommendation generation, safety evaluation, correlation analysis, trajectory estimation, and persistence with goal updates. This design avoids stale batch computation and ensures that recommendations reflect the latest available state.

Input normalization combines several sources rather than a single daily form. The pipeline may read profile data, lifestyle

snapshot values, recent daily logs, weekly body measures, optional laboratory values, and active recommendation state. As a result, the score is not merely a reflection of one day of behaviour. It is a structured summary of recent patterns, background risk, and context.

### B. Rolling metrics and computation windows

Metric extraction mirrors how prevention goals are described in practice. Activity is summarized through moderate equivalent minutes over a seven day window because the underlying target is weekly rather than daily. Steps, sleep, stress, sedentary time, and diet signals are aggregated over recent windows, while weight trend, waist circumference, and latest laboratory values capture slower moving risk information. These windows allow the system to react to behaviour change without overreacting to one isolated entry.

Metric	Interpretation	Window or note
avgSteps7d	Average daily steps used to assess ambulatory activity.	7 days
moderateEqMin7d	Weekly moderate equivalent activity volume derived from logged sessions.	7 days
avgSleepHours7d	Average non zero sleep duration.	7 days
sleepStdDev7d	Sleep consistency and variability.	7 days
sugaryDrinks7d	Sugar sweetened drink count.	7 days
fastFood7d	Fast food exposure.	7 days
avgStressScore7d	Self reported stress burden.	7 days
avgSedentaryHours7d	Average daily sedentary time.	7 days
daysLogged7d	Data completeness gate for rule evaluation.	7 days

Metric	Interpretation	Window or note
weightTrend28dPct	Recent body weight direction and magnitude.	28 days
waistCm	Central adiposity indicator.	Latest available value
latestHbA1c or latestFastingGlucose	Laboratory risk signal that can dominate the score when values are in prediabetes or diabetes range.	Latest available value

TABLE II. RISK INPUTS AND COMPUTATION WINDOWS

### C. Weighted score construction and band mapping

The user facing score is normalized to a 0 to 100 scale but is built from bounded raw contributions. Family history can contribute up to 20 points, activity deficit up to 25, BMI up to 25, steps up to 20, sugary drinks up to 20, and sleep, fast food, sedentary exposure, and waist circumference lower but still meaningful penalties. Laboratory values can contribute far more strongly because diabetes range findings should not be handled like ordinary lifestyle friction. The system also recognizes compound patterns such as poor sleep with high stress or low activity with high dietary burden.

Band mapping converts the normalized score into four interpretable levels: Low for 0 to 24, Medium for 25 to 49, High for 50 to 74, and Very High for 75 to 100. The weighting logic also incorporates South Asian relevant anthropometric action points, including BMI attention from 23 and lower waist cut offs than common Western defaults [10]-[13]. This calibration is important because the intended user context includes populations in whom metabolic risk appears at lower adiposity thresholds.

### D. Rule evaluation, safety override, and personalized outputs

Recommendation generation is rule driven rather than opaque. The documented active configuration contains 14 base rules in the current production rule version, while a broader importable ruleset contains more than 35 rules and modifiers. Each rule defines trigger logic, resolve logic, category, base priority, and user facing action text. Candidate recommendations are scored, family history can add a priority boost, duplicate categories are collapsed so that the user sees a balanced set of cards, and only the highest priority items are surfaced together.

Safety logic sits above ordinary coaching. When fasting glucose reaches 7.0 mmol/L or HbA1c reaches 6.5 percent, routine prevention cards are replaced by a medical review alert. Recommendation state can also move from active to resolved when later behaviour satisfies the relevant resolve condition. Beyond recommendation cards, the pipeline computes dynamic goals, short horizon trajectory, and Pearson style correlations among selected behaviour pairs when sufficient history exists. A what if simulator reuses the same score logic without persisting hypothetical data.

## Explainable Risk Scoring and Recommendation Pipeline

How recent logs, baseline profile data, family history, and optional laboratory values become a user facing score and action set

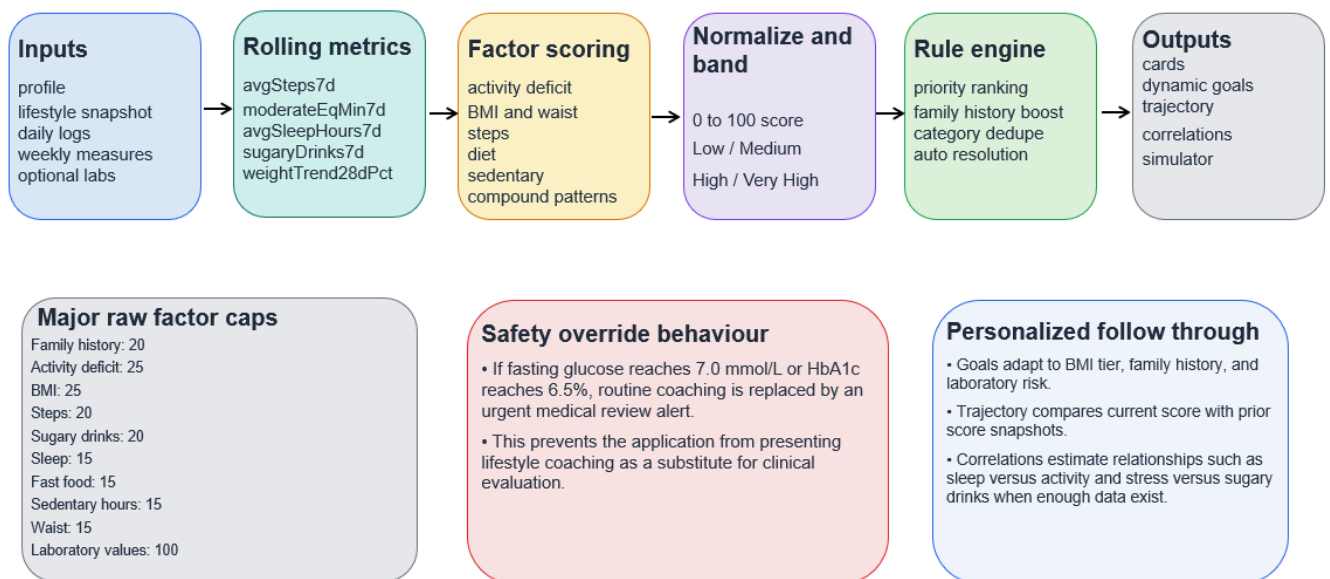


Fig. 3. Explainable risk pipeline from inputs and rolling metrics to factor scoring, recommendation generation, safety override behaviour, and personalized follow through.

## VI. PROTOTYPE WORKFLOWS AND FUNCTIONAL VALIDATION

### A. Validation strategy

The present study reports functional validation rather than prospective clinical outcomes. The question is whether the implemented system behaves consistently and transparently when exposed to representative prevention scenarios. This is appropriate for an engineering paper because the main

contribution is the design and operation of the platform, not a completed longitudinal intervention trial.

Validation therefore focuses on workflow correctness: whether onboarding can produce a first estimate without logs, whether daily logs trigger consistent recomputation, whether diabetes range laboratory values invoke the safety pathway, whether recommendations can auto resolve, and whether the dashboard can assemble the latest score, trajectory, goals, and recommendation state into one user facing response.

Scenario	Trigger condition	Expected system response
Onboarding estimate	Profile completed before sufficient longitudinal logs exist.	Lifestyle snapshot fields bootstrap the first risk score, band, and recommendation set so the dashboard is informative from the first session.
Daily recomputation	A new daily log records activity, steps, sleep, diet, stress, or body measures.	Rolling metrics update, a new append only score record is created, and recommendations and goals are recalculated.
Safety override	Fasting glucose is at least 7.0 mmol/L or HbA1c is at least 6.5 percent.	Routine prevention cards are suppressed and an urgent medical review alert is presented.
Auto resolution	A previously active problem later satisfies the relevant resolve condition.	The recommendation transitions out of the active state and no longer competes for dashboard priority.
Dashboard synthesis	The user opens dashboard or insights after prior interactions.	The backend assembles current score, trajectory, goals, recommendation state, recent history, and available correlation summaries in one response.

TABLE III. FUNCTIONAL VALIDATION SCENARIOS AND EXPECTED SYSTEM BEHAVIOUR

### B. Representative medium risk example

A documented representative case helps illustrate how the score remains interpretable at the factor level. In this scenario the user has parental type 2 diabetes history, BMI 24.8 kg per m squared, activity below the weekly target, fewer than 8000 steps per day, one sugary drink over the recent period, 7.2 h average sleep, and acceptable sedentary exposure. The resulting score is 38 out of 100, which maps to the Medium band. Family history contributes 15 raw points, BMI 10, activity deficit 5, sugary

drinks 3, and step deficit 1.8, while sleep and sedentary time do not add further burden.

The value of this example is not the absolute number alone but the traceability behind it. The user can see which factors drove the score, which categories became recommendation candidates, and why activity receives emphasis when family history is present. This level of transparency is difficult to achieve in black box wellness scoring, but it is central to user trust in prevention settings.

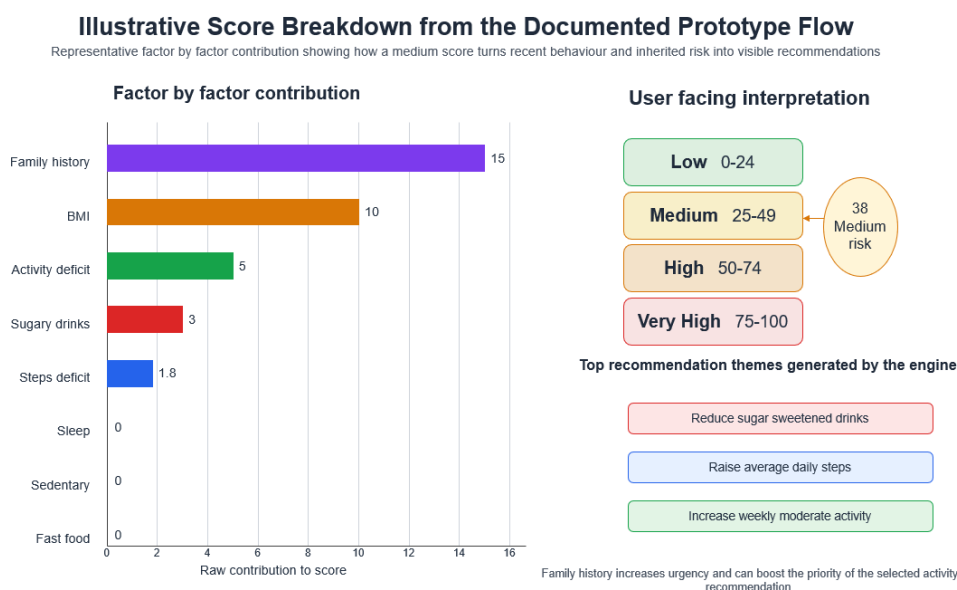


Fig. 4. Illustrative score breakdown from a representative prototype scenario, including factor contributions, resulting risk band, and the dominant recommendation themes.

### C. What the prototype output demonstrates

Taken together, the scenarios and representative case show that Path2Prevention operates as a closed prevention loop: onboarding creates an initial estimate, daily logs update the state, safety conditions can supersede routine coaching, recommendation lifecycle state is preserved, and the latest view can be rendered through dashboard, insights, and report services. These are engineering outcomes rather than clinical endpoints, but they matter because they determine whether the platform can support later pilot deployment and evaluation without redesigning the decision logic.

## VII. DISCUSSION

Path2Prevention contributes value precisely because it treats diabetes prevention as an explainability problem as well as a behaviour change problem. Generic trackers can store activity or diet entries, but they seldom show how family history, adiposity, short sleep, sedentary time, or sugar exposure are translated into an overall prevention view. In contrast, the deterministic method used here makes the risk score inspectable down to the level of windows, factors, and rule outcomes. This improves both user understanding and engineering auditability.

The platform also offers a culturally relevant design choice by embedding South Asian anthropometric action points rather than relying only on thresholds derived from Western populations<sup>[13]</sup>. Combined with family history weighting and safety override behaviour, this makes the system more context aware than many generic wellness experiences. Because the recommendations are deduplicated by category and can resolve when conditions improve, the interface can move from warning to reinforcement instead of repeatedly showing the same generic instruction.

## VIII. LIMITATIONS AND FUTURE WORK

Several limitations remain. The current paper describes architecture, logic, and prototype behaviour rather than prospective reduction in diabetes incidence or glycaemic progression. Much of the behavioural input is self reported, which introduces recall bias and missingness. The platform does not yet integrate wearables, clinician review workflows, or laboratory verification streams. The normalization denominator used for the displayed score is a practical interface design decision and should be revisited after longitudinal calibration.

Future work should proceed in three linked directions. First, usability and adherence studies are needed to evaluate whether users understand the score breakdown, recommendation phrasing, and dynamic goals. Second, prospective validation should test whether risk bands and factor weights correspond to observed change in weight, fasting glucose, or HbA1c over time. Third, the platform can be extended with wearable ingestion, clinician review tools, multilingual delivery, adaptive notification timing, and controlled comparison against more complex machine learning assisted personalization while preserving interpretability as a baseline.

## IX. CONCLUSION

Path2Prevention demonstrates how a full stack website can translate established diabetes prevention evidence into an explainable digital workflow. The implemented system integrates guided onboarding, daily logging, a secure backend, append only analytics, a deterministic risk score, rule based recommendations, safety aware laboratory logic, dynamic goals,

and user facing export tools. Its main contribution is a transparent and extensible engineering foundation for personalized type 2 diabetes prevention. That foundation is suitable for journal reporting now and for pilot deployment and outcome validation in the next stage of the project.

## REFERENCES

- [1] Diabetes Prevention Program Research Group, "Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin," *New England Journal of Medicine*, vol. 346, no. 6, pp. 393-403, 2002.
- [2] J. Tuomilehto, J. Lindstrom, J. G. Eriksson, T. T. Valle, H. Hamalainen, P. Ilanne-Parikka, and M. Uusitupa, "Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance," *New England Journal of Medicine*, vol. 344, no. 18, pp. 1343-1350, 2001.
- [3] X. R. Pan, G. W. Li, Y. H. Hu, J. X. Wang, W. Y. Yang, Z. X. An, and B. V. Howard, "Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: The Da Qing IGT and Diabetes Study," *Diabetes Care*, vol. 20, no. 4, pp. 537-544, 1997.
- [4] Y. Fukuoka, C. L. Gay, K. L. Joiner, and E. Vittinghoff, "A novel diabetes prevention intervention using a mobile app: A randomized controlled trial with overweight adults at risk," *American Journal of Preventive Medicine*, vol. 49, no. 2, pp. 223-237, 2015.
- [5] M. F. Alwashmi, G. Mugford, W. Abu-Ashour, and M. Nuccio, "A digital diabetes prevention program Transform for adults with prediabetes: Secondary analysis," *JMIR Diabetes*, vol. 4, no. 3, e13904, 2019.
- [6] S. L. Fitzpatrick, M. Mayhew, A. M. Rawlings, N. Smith, D. B. Nyongesa, W. M. Vollmer, V. J. Stevens, S. K. Grall, and S. P. Fortmann, "Evaluating the implementation of a digital diabetes prevention program in an integrated health care delivery system among older adults: Results of a natural experiment," *Clinical Diabetes*, vol. 40, no. 3, pp. 345-353, 2022.
- [7] E. Jahan, R. Almansour, K. Ijaz, S. Baptista, L. B. Giordan, R. Ronto, S. Zaman, E. O'Hagan, and L. Laranjo, "Smartphone applications to prevent type 2 diabetes: A systematic review and meta analysis," *American Journal of Preventive Medicine*, vol. 66, no. 6, pp. 1060-1070, 2024.
- [8] Y. A. Jeem, R. N. Andriani, R. Nabila, D. D. Emelia, L. Lazuardi, and H. Koesnanto, "The use of mobile health interventions for outcomes among middle aged and elderly patients with prediabetes: A systematic review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 20, art. 13638, 2022.
- [9] J. B. Echouffo-Tcheugui, L. Perreault, L. Ji, and S. Dagogo-Jack, "Diagnosis and management of prediabetes: A review," *JAMA*, vol. 329, no. 14, pp. 1206-1216, 2023.
- [10] World Health Organization, WHO Guidelines on Physical Activity and Sedentary Behaviour. Geneva: World Health Organization, 2020.
- [11] N. F. Watson, M. S. Badr, G. Belenky, D. L. Bliwise, O. M. Buxton, D. J. Buysse, D. F. Dinges, J. Gangwisch, M. A. Grandner, C. Kushida, R. K. Malhotra, J. L. Martin, S. R. Patel, S. F. Quan, and E. Tasali, "Recommended amount of sleep for a healthy adult: A joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society," *Sleep*, vol. 38, no. 6, pp. 843-844, 2015.
- [12] R. K. Johnson, L. J. Appel, M. Brands, B. V. Howard, M. Lefevre, R. H. Lustig, F. Sacks, L. M. Steffen, and J. Wylie-Rosett, "Dietary sugars intake and cardiovascular health: A scientific statement from the American Heart Association," *Circulation*, vol. 120, no. 11, pp. 1011-1020, 2009.
- [13] A. Misra, P. Chowbey, B. M. Makkar, N. K. Vikram, J. S. Wasir, D. Chadha, S. R. Joshi, S. Sadikot, R. Gupta, S. Gulati, and Y. P. Munjal, "Consensus statement for diagnosis of obesity, abdominal obesity and the metabolic syndrome for Asian Indians and recommendations for physical activity, medical and surgical management," *Journal of the Association of Physicians of India*, vol. 57, pp. 163-170, 2009.
- [14] R. M. Anjana, R. Unnikrishnan, M. Deepa, R. Pradeepa, N. Tandon, A. K. Das, and V. Mohan, "Metabolic non communicable disease health report of India: the ICMR INDIAB national cross sectional study (ICMR INDIAB 17)," *Lancet Diabetes and Endocrinology*, vol. 11, no. 7, pp. 474-489, 2023.
- [15] D. J. Magliano and E. J. Boyko, Eds., *Diabetes Atlas*, 11th ed. Brussels: International Diabetes Federation, 2025.