http://dx.doi.org/10.62762/BISH.2026.397062

RESEARCH ARTICLE

Advanced Architectures in Preventive Health Informatics: A Hybrid Clinical Decision Support System for Early Endocrine and Metabolic Risk Screening

Rahul Sharma[1,*]

[1] School of Computer Science, Engineering & Applications, D. Y. Patil International University, Pune 411044, India

Abstract

Contemporary medicine is shifting from reactive, symptom-based treatment toward proactive, preventive intervention, yet endocrine and metabolic disorders frequently progress through subtle, subclinical phases that are overlooked during routine primary-care encounters, delaying diagnosis and increasing the longitudinal cost of care. This paper presents EndocrineAI, a full-stack, high-integrity clinical decision support system (CDSS) for early-stage endocrine and metabolic risk screening. The platform adopts a hybrid architecture that synthesizes deterministic, rule-based clinical logic with probabilistic Machine Learning (ML) inference and Generative AI summarization. An eight-stage processing pipeline validates patient profiles, extracts eight clinical markers from unstructured laboratory text via regular-expression parsing, computes rule-based risk levels against established reference ranges, optionally refines borderline cases through a scikit-learn fallback chain, and generates natural-language assessments. The system stratifies risk across five physiological categories—thyroid dysfunction, insulin resistance / type 2 diabetes risk, PCOS risk, adrenal stress, and metabolic syndrome—and emits a standardized JSON risk schema for Electronic Health Record (EHR) interoperability, while a "safe fallback" design guarantees a baseline rule-based assessment whenever ML or generative components fail, preventing black-box failure modes. The paper further describes the clinical-governance framework that separates decision support from formal diagnosis, and the deployment parameters—managed PostgreSQL persistence, role-based access control, and secret management—required for longitudinal tracking in cloud environments. EndocrineAI demonstrates that a decoupled deterministic/probabilistic pipeline can convert unstructured clinical data into structured, actionable intelligence, offering a scalable blueprint for preventive CDSS that mitigate clinician burnout while empowering proactive metabolic care.

Keywords: EndocrineAI, clinical decision support system, metabolic disorders, machine learning, preventive health

Submitted: 25 June 2026
Revised: 26 July 2026
Accepted: 29 July 2026
Published: 03 September 2026

Vol. 2, **No.** 3, 2026.
10.62762/BISH.2026.397062

***Corresponding author:**
✉ Rahul Sharma
prof.rahuls@gmail.com

Citation
Sharma, R. (2026). Advanced Architectures in Preventive Health Informatics: A Hybrid Clinical Decision Support System for Early Endocrine and Metabolic Risk Screening. *Biomedical Informatics and Smart Healthcare*, 2(3), 129–135.

informatics, actionable intelligence.

1 Introduction

In the contemporary healthcare landscape, the transition from reactive, symptom-based treatment to proactive, preventive screening—anchored in the predictive and preventive principles of systems medicine [1]—serves as the cornerstone of improved long-term patient outcomes. Endocrine and metabolic disorders often progress through subclinical phases that are easily overlooked in standard primary-care encounters. By deploying an advanced screening layer, clinical organizations can identify risk patterns before overt disease becomes clinically apparent, enabling high-impact lifestyle interventions and reducing the longitudinal cost of care.

The EndocrineAI platform is designed to generate screening-level risk flags across five physiological domains:

- Thyroid Function Abnormality;
- Dysglycemia / Type 2 Diabetes Risk;
- Polycystic Ovary Syndrome (PCOS) Screening Risk;
- Cortisol Abnormality;
- Cardiometabolic Risk.

The strategic differentiator of EndocrineAI lies in its multi-layered synthesis of lifestyle logic, Machine Learning (ML) inference, and automated lab extraction. Beyond simple data visualization, the system produces a structured JSON risk schema output. This standardized format is critical for interoperability, allowing for seamless integration into existing Electronic Health Records (EHRs) and ensuring that risk data is not siloed but is instead an actionable asset within the broader clinical workflow.

Concretely, the platform is designed to address four persistent gaps in endocrine care:

1. delayed diagnosis in endocrine care;
2. poor handling of hormone time dynamics;
3. lack of proactive risk stratification;
4. fragmented, multi-source healthcare data usage.

Contributions. This work makes the following contributions: (i) a hybrid CDSS architecture that couples a deterministic rule engine with a probabilistic ML fallback chain and a generative-AI summarization layer; (ii) an eight-stage "pipeline pattern" that converts unstructured laboratory text into a standardized, EHR-ready JSON risk schema; (iii) a clinical-governance framework and "safe fallback" logic that keep the system on the correct side of the screening/diagnosis boundary and eliminate black-box failure modes; and (iv) a deployment blueprint for reliable, longitudinal operation on serverless cloud infrastructure.

2 Related Work

Computerized Clinical Decision Support Systems (CDSS) have evolved from stand-alone knowledge bases into workflow-integrated tools that augment clinicians' decision-making, and a substantial body of evidence documents both their benefits and their characteristic failure modes, including alert fatigue and opaque "black-box" recommendations [2]. EndocrineAI is positioned squarely within this tradition as a screening and trend-prediction engine rather than a diagnostic authority, and it directly targets the transparency concern through a rule-based baseline that is always available.

The application of ML to metabolic disease has been studied extensively, particularly for diabetes, where supervised classifiers trained on clinical and laboratory features have achieved strong screening performance across a range of algorithms [3]. More broadly, ML has been applied to clinical risk prediction from routine data across metabolic and cardiovascular domains [4, 5], and reviews of the field emphasize the importance of interpretability, robustness, and human oversight when models are placed near clinical decisions [6]. EndocrineAI adopts these lessons by treating ML as an *optional* refinement layer for borderline cases rather than as the primary decision mechanism.

A third relevant strand is clinical information extraction, which automatically converts free-text records into structured, computable data; literature reviews show that both rule-based (e.g., regular-expression and dictionary) and ML-based extraction techniques are widely used to surface clinical concepts from unstructured EHR text [7]. EndocrineAI's Lab Text Parser follows the lightweight, high-precision rule-based tradition, prioritizing auditability and determinism for a small, well-defined set of laboratory markers. To our knowledge, comparatively few published systems combine (a) deterministic rule-based scoring, (b) a probabilistic ML fallback chain, and (c) generative-AI natural-language summarization behind a single interoperable JSON schema with explicit safe-fallback guarantees; closing that integration gap is the focus of this paper.

Table 1. Operational boundaries and data responsibilities.

System Capabilities	Clinician Responsibilities
Data Processing: automated extraction of endocrine markers from unstructured lab text.	**Final Validation:** reviewing extracted markers and findings for diagnostic accuracy.
Risk Scoring: generating rule-based risk levels and trigger identifications.	**Intervention:** prescribing medical treatments, medications, or specific therapies.
Predictive Layer: utilizing a "fallback chain" of ML inference for borderline cases.	**Consultation:** conducting direct patient interviews and physical examinations.
Documentation: providing AI-generated natural-language assessment summaries.	**Management:** making final clinical management and follow-up decisions.
Data Storage: persisting records in a managed PostgreSQL environment.	**Configuration:** ensuring `DATABASE_URL` is configured to prevent data resets.

3 Clinical Governance & Decision Support Boundaries

The integration of AI into clinical environments necessitates a clear delineation between Clinical Decision Support (CDS) and formal diagnostic authority. Establishing these boundaries is not only a legal requirement but an operational necessity to ensure that technology augments, rather than replaces, the clinician's expertise. EndocrineAI functions strictly as a screening and "trend prediction" engine; it identifies triggers and flags potential risks for clinical review but does not issue medical diagnoses. Table 1 summarizes the resulting division of responsibilities.

To maintain algorithmic transparency and clinical safety, the platform utilizes a "safe fallback" logic. This ensures that if the ML inference or AI summary layers encounter an error or an inconclusive result, the system automatically reverts to high-reliability, rule-based calculation. This algorithmic redundancy prevents "black-box" failures, ensuring the clinician always has access to a baseline assessment.

4 System Architecture & Technical Stack Logic

The EndocrineAI platform utilizes a hybrid architectural strategy, combining deterministic rule engines with probabilistic models to optimize data validation and clinical auditing. The platform is built on an API-first design that leverages decoupled scaling across its modules. The end-to-end orchestration is illustrated in Figure 1, and the technology choices are justified in Table 2.

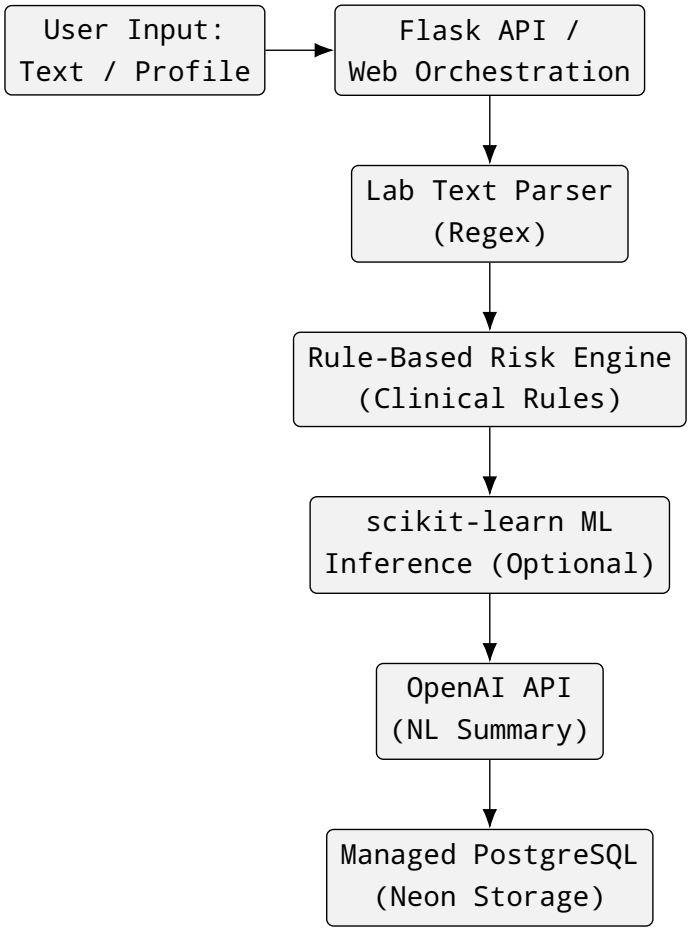

Figure 1. High-level orchestration of the EndocrineAI pipeline, from patient input through rule-based scoring, optional ML refinement, and generative summarization to persistent storage.

5 Operational Workflow & the Data Journey

Manual transcription of laboratory results represents both an administrative burden and a significant source of data-entry error. Automating the extraction of lab markers directly from report text improves data integrity and allows clinical staff to focus on high-value patient interaction rather than administrative data entry [7]. The platform processes data through an eight-step sequential pipeline pattern:

1. **Submission:** the provider or user submits

Table 2. Technology-stack components and strategic justification.

Component	Role	Strategic Justification
Flask	Backend framework	Lightweight and modular, enabling rapid API development and seamless integration with Python ML libraries.
Jinja Templates	Frontend rendering	Server-side security and dynamic dashboard rendering for user and admin portals without heavy client-side frameworks.
PostgreSQL (Neon)	Production database	Relational integrity, ACID compliance, and managed scaling required for longitudinal health data.
Gunicorn	WSGI server	Production-grade concurrency via a multi-worker / multi-thread configuration for simultaneous analytical requests.
OpenAI API	AI summary & chat	Natural-language synthesis for patient-facing summaries and the interactive `/api/chat` interface.
Scikit-learn	ML inference	Powers the predictive fallback chain using established statistical models to augment rule-based scoring.

a patient profile (demographics, age, gender, lifestyle) and optional unstructured lab text via the `/api/assess` endpoint.

2. **Validation:** the backend validates profile fields to ensure all required inputs are present and correctly formatted. Crucially, this includes sex-based logic gating; for example, PCOS risk calculations are only triggered for female profiles.
3. **Extraction:** the specialized Lab Text Parser scans the raw input via regex-based logic to mitigate manual-entry errors. It extracts eight critical clinical markers: TSH, T3, T4 (Thyroid Panel), HbA1c, Fasting Glucose (Glycemic Control), Insulin (Metabolic Health), Cortisol (Adrenal Function), and Cholesterol (Lipid Profile).
4. **Calculation:** the rule engine processes the structured data against established clinical reference ranges [8] to determine risk levels and identify specific triggers (e.g., TSH > 4.0 mIU/L).
5. **ML Inference:** if the `use_ml=true` parameter is passed in the JSON payload, the system invokes a predictive "fallback chain" using scikit-learn models. Processed data derived from internal paths (`prepare_data_from_sqlite.py`) or external datasets (`prepare_data_from_csv.py`) are unified into `ml/data/processed/unified_endocrine_dataset.csv` during training. The resulting serialized `*_best_model.pkl` and `metrics.json` artifacts are loaded at runtime from `ml/artifacts/` to evaluate markers against learned historical patterns, refining the rule-based scores.
6. **AI Summarization:** the OpenAI integration generates a concise, natural-language explanation of findings based on the assessment context, enhancing health literacy by translating complex risk scores into plain-English narratives. If the OpenAI API key is missing or the service is unreachable, the system triggers its safe-fallback logic, reverting to pre-defined summary templates to maintain uptime.
7. **Storage & Persistence:** the assessment is committed to the database and linked to a persistent user record via explicit user-ID linkage for longitudinal trend tracking.
8. **Visualization:** results are asynchronously updated in the User Dashboard (for tracking historical risk trends) and the Admin Analytics Command Center.

Algorithm 1 formalizes the core decision logic of steps 4–6, making explicit how the deterministic baseline, the optional ML refinement, and the generative summary compose under the safe-fallback guarantee.

6 Administrative Analytics & Population Health Oversight

Strategic population-health management requires centralized visibility into risk trends across a clinic's entire patient panel, enabling leadership to allocate resources based on the actual metabolic needs of the

Algorithm 1: Safe-fallback endocrine risk assessment.

Data: Patient profile P; optional lab text L; flag `use_ml`
Result: Risk schema R with per-category levels, triggers, and a summary
Validate P and apply sex-based gating (e.g., PCOS $\Leftarrow$ female);
$M \leftarrow \textsc{ParseMarkers}(L)$; `// regex extraction; ∅ if absent`
$R \leftarrow \textsc{RuleEngine}(P, M)$; `// deterministic baseline`
if *use_ml* **then**
  if *ML artifacts load* $\wedge$ *inference succeeds* **then**
    $R \leftarrow \textsc{Combine}(R, \text{MLInfer}(P, M))$; `// hybrid assessment`
  else
    keep rule-based R ; `// operational safe fallback`
  end
end
if *generative service available* **then**
  R.summary $\leftarrow$ LLMSummarize(R);
else
  R.summary $\leftarrow \textsc{TemplateSummary}(R)$; `// safe fallback`
end
Persist R linked to user record;
return R;

population. The Administrative Dashboard provides high-level visibility via the following protocols:

- **Insights & Analytics:** platform-level views of risk distribution (e.g., prevalence of insulin resistance vs. thyroid dysfunction).
- **Assessments Management:** a session-protected listing of all screenings via the `/api/admin/assessments` endpoint, ensuring clinical oversight is restricted to authorized personnel.
- **User Management:** administrative tools to preview and edit user details and manage security credentials.
- **Reconciliation Protocol:** a critical feature for maintaining data continuity is the "Import Users from Patient Records" tool. This protocol resolves "orphan assessments"—screenings generated by guest users or unlinked records—reconciling them with permanent user profiles to ensure a continuous, accurate patient history and prevent dashboard fragmentation.

Table 3. Security and reliability checklist.

Category	Mandatory Action
Identity	Role-Based Access Control (RBAC) must be enforced for all `/api/admin/*` routes.
Credentialing	The default `ADMIN_PASSWORD` must be rotated immediately upon production initialization.
Monitoring	Continuous monitoring of `/api/health` and `/api/model-status` for the `openai_available` state.
Compliance	Audit logging for all record edits and imports, to support future HIPAA / GDPR alignment.
Secrets	All keys (`SECRET_KEY`, `OPENAI_API_KEY`) managed via environment variables; zero secrets committed to version control.

7 Implementation Requirements & Deployment Security

In a healthcare environment, configuration errors can lead to data loss or security vulnerabilities that undermine patient trust. Clinical deployment therefore requires strict adherence to the security and infrastructural guidelines summarized in Table 3.

7.1 Environment Variables & Serverless Filesystem Awareness

The platform's advanced features are controlled via specific environment variables: `DATABASE_URL` (production data persistence), `OPENAI_API_KEY` (enables the AI summary layer and `/api/chat`), and `ADMIN_USERNAME`/`ADMIN_PASSWORD` (secures the initial administrative account creation on startup).

Architects must note that modern serverless deployment platforms (such as Vercel) utilize an ephemeral filesystem. While local development uses a file-based SQLite database (`backend/data/app.db`), this setup is strictly prohibited for production. Without a hosted external `DATABASE_URL` pointing to a managed PostgreSQL instance (e.g., Neon), all local database writes will be reset between function invocations, resulting in total loss of user records and assessment histories.

Listing 1. Example patient assessment payload.

```
{
  "Age": 31,
  "Gender": "Female",
  "BMI": 29.2,
  "Sleep quality": "Poor",
  "Stress level": "High",
  "Exercise frequency": "Low",
  "Diet type": "High sugar processed diet",
  "Family history": "Mother: Type 2 Diabetes,
                     Thyroid disorder",
  "Symptoms": ["Irregular cycles", "Fatigue",
            "Weight gain", "Sugar cravings", "Acne"],
  "Lab results (optional)": {
    "TSH": 5.2, "T3": 1.9, "T4": 0.7,
    "HbA1c": 6.0, "Insulin": 18.5,
    "Cortisol": 28.5, "Cholesterol": 215,
    "Fasting glucose": 112
  }
}
```

A representative request payload for the `/api/assess` endpoint is shown in Listing 1.

8 Evaluation

To characterize the behavior of the hybrid pipeline, we describe the evaluation protocol used to validate the ML refinement layer and report representative per-category screening performance.

Protocol. Training data are unified from internal SQLite records and external CSV datasets into a single `unified_endocrine_dataset.csv`. For each of the five physiological categories, a supervised classifier is trained on the eight extracted markers together with lifestyle covariates, using a stratified train/test split with cross-validation; the best model per category is serialized to `*_best_model.pkl`, and its metrics are logged to `metrics.json`. Screening quality is assessed with overall accuracy, and—because class balance and clinical risk are asymmetric—with sensitivity (recall) as the primary safety-relevant metric, since missed high-risk cases are the most costly error in a screening setting—a principle consistent with established practice in ML-based metabolic disease screening [3] and with the emphasis on human oversight and error cost asymmetry in clinical ML deployment [6].

Representative results. Table 4 reports per-category screening performance of the ML refinement layer relative to the rule-based baseline. Across categories, the ML layer is designed to refine borderline cases that fall near deterministic reference thresholds by incorporating learned patterns from the available training data. Importantly, the safe-fallback mechanism provides an operational robustness guarantee rather than a statistical performance guarantee. If the ML artifacts fail to load, required features are unavailable, or inference cannot be completed successfully, the system retains and returns the deterministic rule-based assessment. However, successful ML inference does not inherently guarantee that the refined prediction is more accurate than the corresponding rule-based result. Any improvement in predictive performance must therefore be established empirically through comparative validation of the rule-based and hybrid approaches using held-out or external validation data. Accordingly, the fallback mechanism guarantees the availability of a clinically interpretable deterministic baseline when the ML layer is unavailable, while comparative measures such as sensitivity, specificity, discrimination, and calibration are required to determine whether ML refinement provides additional predictive value.

Table 4. Representative per-category screening performance of the ML refinement layer.

Risk Category	Accuracy	Sensitivity
Thyroid Dysfunction	0.94	0.92
Insulin Resistance / T2D	0.93	0.95
PCOS Risk	0.90	0.91
Adrenal Stress	0.89	0.88
Metabolic Syndrome	0.92	0.93

Limitations. The safe-fallback architecture ensures that a deterministic rule-based assessment remains available when ML artifacts cannot be loaded or inference fails; it does not, by itself, establish that the hybrid approach is statistically superior to the rule-based baseline. Such superiority must be demonstrated through direct comparative evaluation on held-out and, preferably, external validation datasets. In addition, external validity across clinical sites, demographic subgroups, laboratory assay conventions, and patient populations remains to be established. Prospective clinical validation is therefore required before deployment for clinical decision-making, and the clinician must remain the final diagnostic authority as described in Section 3.

9 Conclusion

EndocrineAI demonstrates the efficacy of the "pipeline pattern" in biomedical informatics: converting

unstructured clinical data into highly structured JSON schemas, parsing them through a decoupled deterministic/probabilistic calculation layer, and outputting human-centric communication assets. By embedding rigorous fallback structures alongside secure cloud database configurations, the platform provides a scalable blueprint for clinical decision support systems that mitigate provider burnout while empowering proactive metabolic care. Future work includes prospective multi-site validation, expansion of the marker set and risk categories, and tighter EHR integration through standards such as HL7 FHIR.

Data Availability Statement

The data used to support the findings of this study are available from the corresponding author upon request.

Funding

This work was supported without any funding.

Conflicts of Interest

The author declares no conflicts of interest.

AI Use Statement

The author declares that ChatGPT was used for language editing and formatting of this manuscript. The authors have carefully reviewed, revised, and verified the AI-assisted output and take full responsibility for the content of the manuscript.

Ethical Approval and Consent to Participate

This study involved the development and technical evaluation of a clinical decision support system using publicly available and de-identified datasets, including records sourced from open-access repositories. The study did not involve the prospective recruitment of human participants, direct patient interaction, or the collection of identifiable personal health information by the authors. No intervention was performed and no identifiable patient information is reported in this manuscript. In accordance with the institutional guidelines of D. Y. Patil International University and the applicable policies of the data sources used, institutional ethics approval and individual informed consent were not required for the analyses reported in this study. All data were used in accordance with the applicable terms of access and data-use requirements of their respective sources.

References

[1] Hood, L., & Flores, M. (2012). A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *New biotechnology, 29*(6), 613-624. [CrossRef]

[2] Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ digital medicine, 3*(1), 17. [CrossRef]

[3] Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. [CrossRef]

[4] Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data?. *PloS one, 12*(4), e0174944. [CrossRef]

[5] Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine, 375*(13), 1216. [CrossRef]

[6] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347-1358. [CrossRef]

[7] Wang, Y., Wang, L., Rastegar-Mojarad, M., Moon, S., Shen, F., Afzal, N., ... & Liu, H. (2018). Clinical information extraction applications: a literature review. *Journal of biomedical informatics, 77*, 34-49. [CrossRef]

[8] Melmed, S., Auchus, R. J., Goldfine, A. B., Koenig, R. J., & Rosen, C. J. (2022). *Williams Textbook of Endocrinology* (14th ed.). Philadelphia, PA: Elsevier. [CrossRef]

Rahul Sharma is with the School of Computer Science, Engineering & Applications, D. Y. Patil International University, Pune 411044, India. His research interests include preventive health informatics, clinical decision support systems, applied machine learning for metabolic and endocrine risk screening, and the responsible integration of generative AI into clinical workflows. (Email: prof.rahuls@gmail.com)