CoralMD: A Multi-Modal Dashboard for Personalized Medicine

Tanner DeGrazia

Pomona College

Data Science Capstone Project

DS 190: Senior Seminar

November 2025

Abstract

Modern healthcare systems often remain reactive, addressing disease only after symptoms appear. CoralMD is a data-driven prototype dashboard that integrates genomic, wearable, and electronic health record (EHR) data to enable interpretable, proactive healthcare insights. By combining machine learning and ethical data practices, the system aims to make precision medicine transparent, equitable, and actionable. This paper outlines the motivation, data architecture, and proposed design for CoralMD, along with a review of current literature and ethical challenges surrounding personalized healthcare technologies.

1 Introduction

Modern healthcare is a system that is focused on reacting to disease. Cancer, Heart Disease, Neurodegenerative disease, and Diabetes are four of the leading causes of death in the United States. None of them are treated in a manner where prevention is the goal. Our current healthcare system is largely reactive, intervening only after illness occurs, despite an extreme growth in accessible biological and lifestyle data in recent years. The goal of personalized medicine is to reverse that trend by using data to predict and prevent disease. So while most current systems fail to bridge the gap between genomic insights, real-time physiological tracking like apple watches, and clinical health records, I propose a multi-modal system to act as a home base for the outlets to interact. CoralMD is an interpretable, multi-modal dashboard designed to bring these domains together and make personalized, proactive care accessible for both clinicians and patients.

While genomics, wearables, and electronic health records (EHRs) have individually transformed medical data collection, integrating them into a cohesive, interpretable model remains difficult. Research highlights issues in datasets like bias and lack of explainability in clinical AI systems. Existing machine learning approaches often look to prioritize accuracy at the expense of transparency, creating black-box models that clinicians cannot trust entirely as they are being told what to do. Another problem that exists is the lack of a dataset that encompasses all modes of data, making a related model across each avenue very hard. Thus, generalization is difficult when population-level genomic data overrepresents certain ethnicities.

CoralMD will address these challenges through a machine learning framework that integrates genomic (GRCh38, 1000 Genomes, ClinVar, gnomAD), wearables (Apple Watch, OhioT1DM), and Electronic Health Record (MIMIC-IV) datasets. We can bring these streams together in a model that works together with each outlet, despite not having one dataset that is all encompassing CoralMD's modular architecture will allow clinicians to see how genetic predispositions interact with real-time information and medical history to shape health risk.

The project must overcome several technical hurdles: aligning multi-modal data of differing time scales and formats, preserving privacy, and ensuring fairness in model predictions, all while delivering an accurate product. To validate the system, CoralMD will benchmark predictive models using metrics such as accuracy, SHAP-based visualizations are popular among other personalized medicine literature, so the aim is to use this to expose model reasoning. A prototype dashboard built in Streamlit will demonstrate how interpretability can create clinician trust.

If successful, CoralMD will illustrate a scalable pathway toward **Medicine 3.0**, where prediction, prevention, and patient participation become standard practice. Scalability will be achieved hopefully in the future, by switching it to AWS in order to make it a fully usable platform. The project's integration of ethical design, explainable AI, and multi-modal data could advance how clinicians and patients interact with health information. Future work will explore scaling the system with larger datasets and integrating real clinician feedback to refine both usability and fairness.

2 Literature Survey

2.1 AI and Data-Driven Personalized Medicine

Studies such as [1] and [2] explore frameworks combining genomic, wearable, and EHR data for individualized treatment planning. These works report high predictive performance (e.g., accuracy $\approx$ 94%, AUC $\approx$ 0.98) but often sacrifice interpretability. CoralMD builds upon these efforts by prioritizing model transparency and clinician usability by not directing the clinician, but actually letting them interpret the model by their knowledge.

2.2 Explainability and Trust in Clinical AI

Research by [3] and [4] emphasize the importance of interpretable AI and user-centered design for clinical decision-making. Their findings support CoralMD's design choice to integrate SHAP visualizations and transparent model logic to promote trust among healthcare providers.

2.3 Applications in Clinical Decision Support and Visualization

[5], [6], and [7] illustrate how visualization-driven systems enhance clinical interpretation. Their focus on user interface design and model explainability directly informs CoralMD's dashboard layer, which will visualize genomic, metabolic, and clinical data in real time.

2.4 Ethical and Generalization Challenges

Across studies, recurring concerns include fairness, data security, and equity in clinical AI. CoralMD addresses these by adopting fairness metrics, population-level normalization, and a privacy-first data handling strategy informed by [8] and [9].

3 Proposal for New Contribution

3.1 Motivation, Aims, and Objectives

- Build a prototype that integrates multi-modal health data.
- Provide interpretable ML outputs rather than black-box scores.
- Support clinician and patient decision-making ethically.
- Demonstrate the feasibility of equitable personalized medicine tools.

3.2 Research Design and Methods

Data Sources: Genomic: GRCh38, 1000 Genomes, ClinVar, gnomAD.
Wearable: Apple Watch/Fitbit dataset, OhioT1DM CGM data.
EHR: MIMIC-IV de-identified hospital dataset.

Technology Stack: Python (Pandas, PyTorch, Scikit-learn), PostgreSQL, Streamlit dashboard, AWS for scalability.

Modeling and Evaluation: Multi-modal feature integration, SHAP interpretability, metrics including Accuracy, Precision, Recall, F1, and Fairness indices.

Deliverables: Interactive CoralMD prototype, data integration pipeline, and visualization module.

4 Ethical Considerations

- **Bias and Representation:** Address dataset imbalances in genomic and wearable data.
- **Transparency and Trust:** Provide interpretable outputs to clinicians and patients.
- **Privacy and Consent:** Use public mock datasets and adopt privacy-by-design principles.
- **Equitable Access:** Consider which populations benefit from CoralMD and design for inclusivity.

References

[1] Z. Lotfi, R. Haji Hosseini, and M. Aminipour, "Artificial intelligence–driven approaches for prediction, management, and complication risk in type 2 diabetes: A systematic review," *InfoScience Trends*, vol. 2, no. 6, pp. 1–17, 2025.

[2] J. Ballard, Z. Wang, W. Li, *et al.*, "Deep learning-based approaches for multi-omics data integration and analysis," *BioData Mining*, vol. 17, no. 38, 2024.

[3] S. Sadhu, D. Solanki, L. Brick, N. Nugent, and K. Mankodiya, "Designing a clinician-centered wearable data dashboard (careportal): Participatory design study," *JMIR Formative Research*, vol. 7, p. e46866, 2023.

[4] L. H. Goetz and N. J. Schork, "Personalized medicine: Motivation, challenges, and progress," *Fertility and Sterility*, vol. 109, no. 6, pp. 952–963, 2018.

[5] M. K. Im, C. Rouphael, J. McMichael, N. Welch, and S. Dasarathy, "Challenges in and opportunities for electronic health record-based data analysis and interpretation," *Gut and Liver*, vol. 18, no. 2, pp. 201–208, 2024.

[6] E. J. Patterson, A. D. Bounds, S. K. Wagner, *et al.*, "Oculomics: A crusade against the four horsemen of chronic disease," *Ophthalmology and Therapy*, vol. 13, pp. 1427–1451, 2024.

[7] J. Wang, J. Luo, M. Ye, X. Wang, Y. Zhong, A. Chang, G. Huang, Z. Yin, C. Xiao, J. Sun, and F. Ma, "Recent advances in predictive modeling with electronic health records," in *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 8272–8280, 2024.

[8] C. M. Erickson, A. Wexler, and E. A. Largent, "Digital biomarkers for neurodegenerative disease," *JAMA Neurology*, vol. 82, no. 1, pp. 5–6, 2025.

[9] H. K. Brittain, R. Scott, and E. Thomas, "The rise of the genome and personalised medicine," *Clinical Medicine (London)*, vol. 17, no. 6, pp. 545–551, 2017.