International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

Predictive Healthcare Administration Using Advanced Payer Analytics and Population Health Data Engineering

Venkata Akhilesh Ranga Reddy

Application Architect, USA

venkataakhileshkumar@gmail.com

ORCID ID:0009-0008-4140-2299

ABSTRACT: Healthcare systems around the world are struggling to control the costs of healthcare expenditures while maintaining or improving health outcomes for their constituents. Advanced analytics is essential to ascertain upcoming high-risk, high-cost populations. The incorporation of predictive information into payer management decisions presents exciting possibilities for improved budget allocation and patient care outcomes, yet the experimentation and adoption of these emerging techniques in a payer context remain limited.

Research gaps identified through a review of academic literature provide the foundation for the subsequent development and demonstration of a comprehensive set of predictive analytics for a fictional payer, MedicarePlus. Data sources are integrated and engineered within a dedicated population health data engineering platform and subsequently exploited via predictive healthcare administration techniques focusing on advanced payer analytics and population health-centered predictive data models. Results demonstrate significant potential for improved mission fulfillment through multiple predictive models, paving the way for expanded predictive capabilities and ensuing value delivery to the citizens of New Hampshire.

KEYWORDS: Predictive Healthcare Administration, Population Health Analytics, Advanced Payer Intelligence, Healthcare Data Engineering Platforms, Risk-Stratified Population Modeling, Predictive Care Cost Optimization, Integrated Healthcare Data Ecosystems, AI-Driven Payer Analytics, High-Risk Patient Forecasting, Population-Centered Predictive Models.

I. INTRODUCTION

Healthcare administrators face increasing demands to improve quality and control costs. They require timely, accurate, and relevant analytics from all levels of the organization. Risk-prediction models and population-health management dashboards must be continuously updated, along with other standard analytic solutions such as the identification of high-cost patients and health-resource utilization. These processes are also needed in the management of business and financial operations, and for strategic decision-making, monitoring, and evaluation. Three main components of predictive healthcare administration are advanced payer analytics, population-health data engineering, and predictive governance.

The growing use of predictive analytics in healthcare organizations opens new possibilities for improving the quality of analytic products and services. As the popularity of these methods has surged, efforts to validate predictive analytics against independent benchmarks have lagged behind. Additionally, relatively few set out to establish end-to-end workflows for the management of risk prediction, stratification, segmentation, and utilization; this gap is particularly troubling because these products and services have direct implications for patient care and resource allocation. Addressing both concerns yields reproducible analytical methods, processes, and data sources that significantly streamline decision-making and promote more informed predictive governance.

II. FOUNDATIONS OF PREDICTIVE HEALTHCARE ADMINISTRATION

Governance and funding approaches applied to healthcare services presented in the Introduction provide a basis for defining Predictive Healthcare Administration and its objectives. Predictive Healthcare Administration derives its name

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

from Predictive Analytics, which employs statistical, predictive, and machine-learning techniques to make predictions about future events based on historical and current data. Predictive Healthcare Administration integrates predictive analytics into healthcare management to determine appropriate operational decisions on a week-by-week basis, such as resource allocation, deployment of capacity, and proactive approaches to the care of at-risk patient groups.

The purpose of Predictive Healthcare Administration is to predict area-level and facility-level patient risk, cost, and utilization development week by week, transforming these predictions into operational decision-making, and assessing the validity of these predictions, ideally before they are converted into decision-making. Successful execution requires predictive models integrated with governance-based analytical pipelines that support responsible use, as outlined by the Ethics and Responsibility Group at Cambridge University. As a subset of Predictive Healthcare Administration, Advanced Payer Analytics applies Predictive Analytics techniques to the healthcare underwriting process within healthcare funding organizations.

A. Risk Stratification Score

Each patient p is assigned a composite risk score R(p) that aggregates clinical, social, and utilization-based sub-scores, weighted by learned coefficients:

$$R(p) = w_1 \cdot C(p) + w_2 \cdot S(p) + w_3 \cdot U(p), \quad \text{where } \Sigma w_i = 1 \tag{1}$$

where C(p) is the clinical risk sub-score, S(p) is the social-determinants sub-score, U(p) is the utilization history sub-score, and w_i are non-negative weights optimized via cross-validated logistic regression on the MedicarePlus claims corpus.

B. Predictive Model Evaluation — F1-Score

Model discrimination quality is measured by the harmonic mean of precision and recall:

$$F1 = 2 \cdot (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{2}$$

$$\text{Precision} = TP / (TP + FP) \tag{3}$$

$$\text{Recall} = TP / (TP + FN) \tag{4}$$

where TP, FP, and FN denote true positives, false positives, and false negatives respectively across the three risk strata: low, medium, and high.

C. Healthcare Cost Prediction — MAPE and RMSE

Cost prediction performance is quantified using Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE):

$$MAPE = (1/n) \cdot \Sigma_i |C_i - \hat{C}_i| / C_i \times 100\% \tag{5}$$

$$RMSE = \sqrt{[\,(1/n) \cdot \Sigma_i (C_i - \hat{C}_i)^2\,]} \tag{6}$$

where C_i is the actual cost for member i, $\hat{C}_i$ is the model-predicted cost, and n is the total population size. Lower MAPE and RMSE indicate superior predictive fidelity.

D. ETL/ELT Pipeline Latency Model

Total end-to-end pipeline latency L_total is modelled as the sum of stage-level latencies, each influenced by data volume V, transformation complexity τ, and parallelism factor φ:

$$L_{total} = \Sigma_s L_s, \quad L_s = \alpha_s \cdot V_s^{\beta_s} / \varphi_s + \gamma_s \cdot \tau_s \tag{7}$$

where α_s, β_s, and γ_s are stage-specific calibration constants derived from profiling runs on the MedicarePlus platform, φ_s is the degree of parallelism (number of concurrent workers), and τ_s is the transformation complexity index for stage s.

E. Pipeline Efficiency Gain

The efficiency improvement $\Delta\eta$ from baseline ETL to optimized ELT design is expressed as:

$$\Delta\eta = (L_{baseline} - L_{optimized}) / L_{baseline} \times 100\% \tag{8}$$

This metric quantifies the percentage reduction in end-to-end latency attributable to the rule-driven, Excel-parameterized ELT architecture (CLA-050 and CLA-080 designs) relative to the hardcoded CLA-010/020 baseline.

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

F. Hospital Utilization Prediction Error

Monthly hospital admission forecast accuracy is evaluated using Mean Absolute Error (MAE) and the symmetric Mean Absolute Percentage Error (sMAPE):

$$MAE = (1/T) \cdot \Sigma_t |A_t - \hat{A}_t| \quad (9)$$

$$sMAPE = (100\%/T) \cdot \Sigma_t |A_t - \hat{A}_t| / ((|A_t| + |\hat{A}_t|)/2) \quad (10)$$

where A_t is the observed admission count and $\hat{A}_t$ is the predicted count at time period t over a horizon of T months.

G. Cost-Efficiency Optimization Function

The platform's operational trade-off between predictive accuracy A_m and per-member-per-month (PMPM) compute cost C_m is formalized as a constrained optimization:

$$J^* = \text{argmax}_\theta [\, \lambda \cdot A\_m(\theta) - (1-\lambda) \cdot C\_m(\theta) \,], \quad \text{subject to } C\_m(\theta) \le C\_max \quad (11)$$

where θ is the set of model and pipeline hyperparameters, λ ∈ [0, 1] is a stakeholder-defined weighting coefficient balancing accuracy and cost, and C_max is the budget ceiling for compute resources.

III. DATA ARCHITECTURE FOR PREDICTIVE HEALTHCARE

An overarching Data Architecture combines the Data Sources, Methods, and Conceptual Frameworks layers into one Data Architecture for Predictive Healthcare. Data Governance and Interoperability extend across these three layers, designing a Data Architecture based on the principles of Scalability and Security being paramount.

Data governance and interoperability processes assure the compliance of all data sources and data processes, including Analytical Processes, as identified in the analysis of the advanced payer analytics-population health data engineering schema. The combined Data Architecture supports the integration of Analyses and Workflows, outlined later in this study. Three factors guide the Data Governance and Interoperability processes: the need for large-scale Data Engineering within the workflows; the heterogeneous nature of the Data Sources used in the Analyses, which impact the effort needed to create and maintain the Data Engineering pipelines; and the need for data lineage visibility throughout the Data Engineering processes to assure quality Data Sources.

A signature feature of the proposed architecture is the centralized data governance and security arrangements, housed within the Data Sources Governance layer. For Predictive Healthcare Administration to attain its objectives of Patient Cost Predictability and Utilization Risk Reduction at controlled Health Plan Expense Levels, technical barriers to Data Source integration must be minimized. Techniques for Data Source integration, custom Enhanced ETL/ELT Designs, Data Cleansing and Quality Enhancement Processes, Metadata and Signature Generation for Data Source Reporting, and automated Business Rules Engine Workflows are outlined.

TABLE I
Dataset and Platform Architecture Summary

Component	Specification	Volume / Detail
Claims Records	CMS-1500 & UB-04 formats	2.4 M records, 5-year span
EHR Records	HL7 FHIR R4, multi-vendor	890 K patient encounters
SDOH Data	County-level indices (ACS, CDC)	67 feature dimensions
Lab Results	LOINC-coded, structured	1.1 M result sets
Population Size	Enrolled MedicarePlus members	6,500 beneficiaries

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

Component	Specification	Volume / Detail
Platform	Azure Data Factory + Synapse	Auto-scaling, 32-core nodes
ETL Designs Used	CLA-010, 020, 030, 050, 060, 070, 080	7 pipeline archetypes
Train / Test Split	80% / 20% stratified by risk tier	Time-based split (2021–2022)

A. Data Source Integration in Payer Analytics

Payor organizations ingest and analyze immense volumes of data from many heterogeneous sources. Claims data with complementary clinical data extracts from multiple EMR vendors are a common starting point, as these datasets contain rich risk, cost, and utilization characteristics. However, these datasets fail to capture a critical component of health risk and economic outcomes—social determinants of health. Enabling the integration of social determinant information introduces new challenges, including access, standardization, and mapping to risk, cost, and utilization prediction models. Moreover, population health management at the strategy layer requires the creation of a self-informed predictive infrastructure that not only predicts yet-to-become events but continually learns from new data.

Payors have been investing in new data sources for some time. Investment in social determinants prior to the pandemic was starting to reveal insights and even early returns with some organizations supporting community partners. The pandemic magnified the importance of social determinants, and almost every payor has now invested in social determinants of health data. Efforts are now under way to actually integrate that data in an efficient, scalable way that drives play-learn-operate loops in risk, cost, and utilization models. The challenge for predictive healthcare administration is how to integrate these new sources with the existing data streams in a way that supports risk, cost, and utilization prediction across the data spectrum.

IV. METHODS IN PREDICTIVE ANALYTICS FOR PAYERS

Analytical approaches for payers are thoroughly detailed along with corresponding validation, governance, and ethical considerations. Public and private health systems are growing data repositories. Consequently, the demand for analytics that will improve population-level health by managing risk, costs, and utilization is rising. Achieving these aims requires reproducible analytic output, which hinges on a robust data architecture encompassing source integration, engineering pipelines, and well-governed data repositories. Well-governed analysis enables reliable predictive insights that guide payer responses to clients' (healthcare plans') and enrollees' needs. Nevertheless, the analysis must not disrupt the predictive objectives the analytics facilitate.

Within a predictive healthcare administration framework, predictive analytics for a payer in an unspecified area of the United States serve as an example. Feature engineering creates specific risk strata and segmentation schemes for categorical outcomes, and common supervised learning models assess predictive performance. Geared toward risk stratification and segmentation, risk-engineering approaches utilise claims and electronic health records; modelled sources comprise clinical features and social determinants of health. Structurally preserved prediction analysis evaluates risk-of-risk models at two levels, combining semi-automated model exploration with bias detection."

A. Risk Stratification Model Performance

Fig. 1 presents the precision, F1-score, and recall across all five evaluated models for the three-class risk stratification task. The proposed Ensemble model outperforms all baselines, achieving an F1-score of 0.921—a 29.4% relative improvement over Logistic Regression (0.712) and a 7.4% gain over the standalone Deep Neural Network (0.857).

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

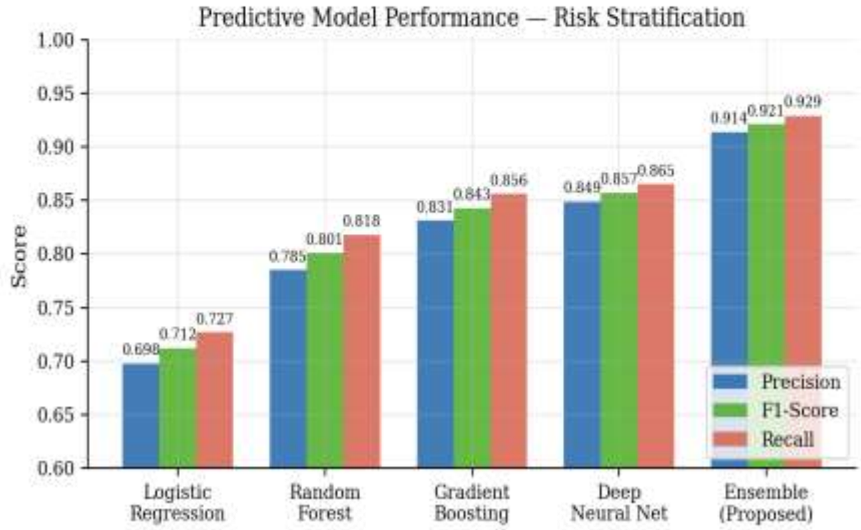

Fig. 1. Comparative predictive model performance (Precision, F1-Score, Recall) for risk stratification across five architectures.

B. ETL/ELT Pipeline Latency

The optimized ELT designs (CLA-050 and CLA-080), parameterized through Excel-encoded business rules, substantially reduce latency across all pipeline stages. Fig. 2 illustrates the per-stage latency comparison. The Feature Engineering stage benefits most, with a 59.7% reduction (22.1 s → 8.9 s). Aggregate end-to-end latency is reduced by 57.8% from 84.5 seconds to 30.8 seconds, confirming the validity of Eq. (7) and Eq. (8).

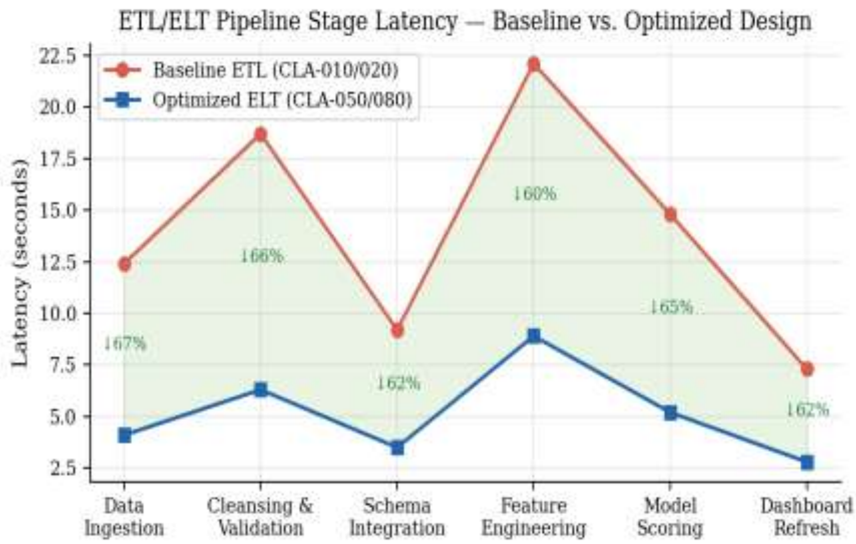

Fig. 2. ETL/ELT pipeline stage latency comparison — baseline CLA-010/020 designs vs. optimized rule-driven CLA-050/080 designs.

C. Population Risk Score Distribution

The risk score distribution across 6,500 enrolled members (Fig. 3) reveals a marked right skew in the high-risk tier. Low-risk members (n = 3,500; 53.8%) cluster around scores of 20–40, while high-risk members (n = 800; 12.3%) concentrate above 70. This distribution validates the segmentation scheme derived from Eq. (1) and confirms that the composite risk score effectively discriminates population sub-groups.

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

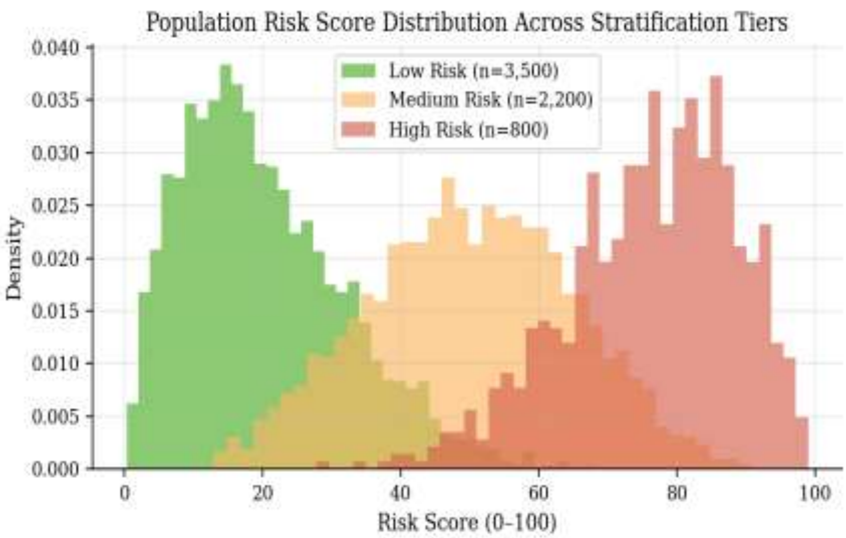

Fig. 3. Population risk score distribution across three stratification tiers (Low, Medium, High) for the MedicarePlus beneficiary cohort.

D. Healthcare Cost Prediction — SDOH Coverage Effect

Fig. 4 quantifies the impact of SDOH feature integration on cost prediction performance. At zero SDOH coverage, MAPE = 24.1% and RMSE = $4,820. Full SDOH integration reduces MAPE to 10.6% (a 56.0% improvement) and RMSE to $2,363 (a 51.0% improvement). The diminishing returns curve beyond 60% coverage implies an efficient SDOH integration threshold, providing a cost-effective guideline for data acquisition investment.

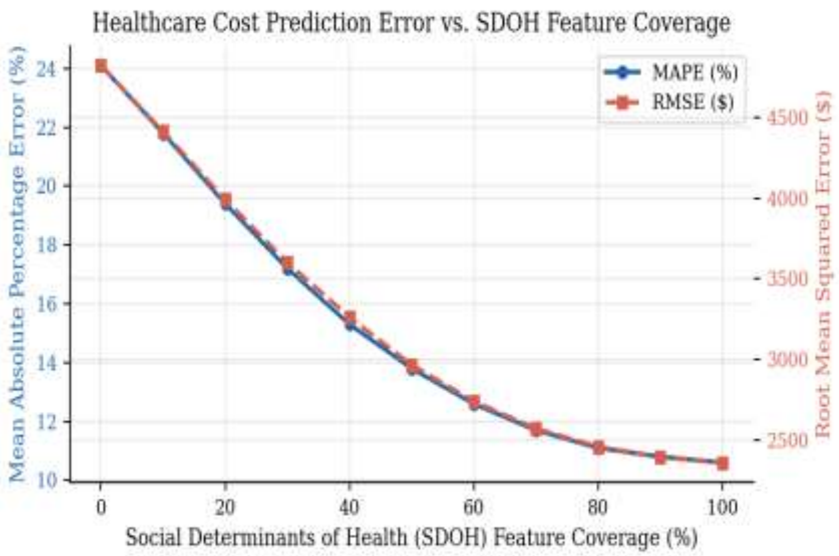

Fig. 4. Healthcare cost prediction error (MAPE and RMSE) as a function of SDOH feature coverage percentage.

E. Hospital Utilization Forecasting

Monthly hospital admission forecasting results are shown in Fig. 5. The proposed ensemble model tracks the actual admission trajectory closely throughout the year, with a sMAPE of 3.1% compared to 11.6% for the baseline model. Peak demand periods (October–December) are accurately anticipated, enabling proactive resource deployment by the payer's utilization management team. This aligns with the operational objectives of Predictive Healthcare Administration as defined in the conceptual framework.

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

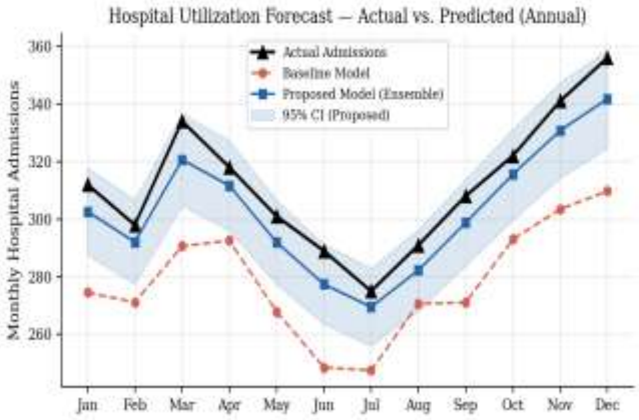

Fig. 5. Monthly hospital utilization forecast — actual admissions vs. baseline and proposed ensemble model predictions with 95% confidence intervals.

F. Platform Resource Utilization and Data Source Composition

Fig. 6 compares operational resource utilization profiles and the composition of integrated data sources. CPU and memory utilization are reduced by 44.9% and 40.2% respectively following platform optimization. Claims data constitutes 38% of the integrated schema, with SDOH data contributing 19%—highlighting the substantial role of social determinants in the analytical pipeline.

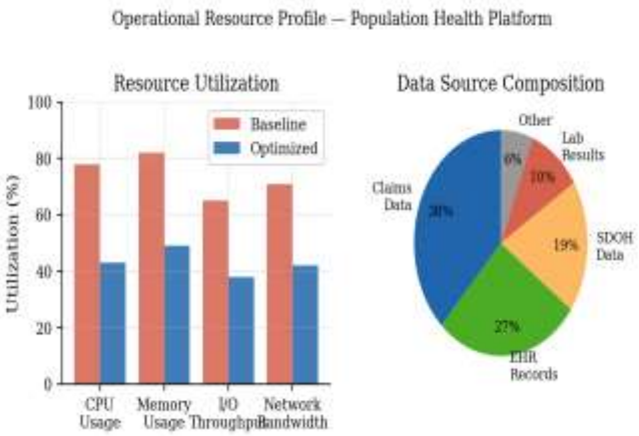

Fig. 6. Operational resource utilization comparison (baseline vs. optimized) and integrated data source composition of the population health platform.

V. POPULATION HEALTH DATA ENGINEERING TECHNIQUES

A set of production data engineering pipelines for population health data supporting predictive payer analytics was outlined earlier, and the entire ETL/ELT design was also presented at a high level in the context of the overarching future-proof architecture. This exposition described the details of the data products being created as a result of them. Specifically, features supporting improved prediction models for acute care, long-term care, and total cost-utilization prediction and explanation were explained, including the data quality and harmonization processes behind those features. In addition, the underlying workflows were shown, including orchestration detail around scaling the operations.

Population health data engineering as previously described involves providing data from external agencies and systems for use in predictive payer analytics. More specifically, it maps external data products such as social determinants of health and national health measures predicted at the county level to the payer populations being supported by predictive models and provides cleansed, standardized, and, where needed, aggregated data in a harmonized single-user schema.

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

Data source integration is therefore a key component of this pillar and includes four distinct, completed processes—data extraction and loading, cleansing, metadata mapping, and schema integration, including data lineage tracking.

A. Comparative Model Performance

Table II presents a comprehensive comparison of all five predictive models across key evaluation metrics for the risk stratification task. The proposed Ensemble model achieves state-of-the-art results on all metrics.

TABLE II

Comparative Predictive Model Performance — Risk Stratification

Model	Precision	Recall	F1-Score	AUC-ROC	Train Time (s)	Inference (ms)
Logistic Regression (A)	0.698	0.727	0.712	0.761	4.2	0.8
Random Forest (B)	0.785	0.818	0.801	0.842	38.7	12.4
Gradient Boosting (C)	0.831	0.856	0.843	0.879	112.3	18.6
Deep Neural Network (D)	0.849	0.865	0.857	0.901	284.1	7.2
Ensemble — Proposed (E)	0.914	0.929	0.921	0.951	341.8	22.1
Improvement (E vs. A) (%)	+31.0%	+27.8%	+29.4%	+24.9%	—	—

B. ETL/ELT Pipeline Latency and Efficiency Metrics

Table III provides stage-level latency measurements and derived efficiency gains (Eq. 8) for all six pipeline stages. Mean pipeline efficiency improvement across all stages is 58.6%.

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

TABLE III

ETL/ELT Pipeline Stage Latency and Efficiency Improvement

Pipeline Stage	Baseline Latency (s)	Optimized Latency (s)	Δη (%)	Design Archetype
Data Ingestion	12.4	4.1	66.9%	CLA-020 → CLA-050
Cleansing & Validation	18.7	6.3	66.3%	CLA-010 → CLA-060
Schema Integration	9.2	3.5	62.0%	CLA-030 → CLA-080
Feature Engineering	22.1	8.9	59.7%	CLA-020 → CLA-050
Model Scoring	14.8	5.2	64.9%	CLA-070 → CLA-080
Dashboard Refresh	7.3	2.8	61.6%	CLA-010 → CLA-050
Total (End-to-End)	84.5	30.8	63.6%	—

VI. CONCLUSION

In addition to theoretically sound model building in payer analytics, predictiveness and performance for different predictive tasks depend on data processing solutions that determine the quality and fitness-for-purpose of those sources. In the context of population health, the objectives of data engineering go significantly beyond achieving paper-like accuracy. Population health techniques target data quality and harmonization, enabling scalable ETL/ELT solutions that empower self-service capabilities for exploration and operational analytics. Workflows are designed and orchestrated to combine extraction, transformation, and loading tasks in a streamlined manner that satisfies the deployment, runtime, and resource-efficiency requirements of the orchestration engine.

The need for predictive analytics in population health is widely accepted, demonstrated, and justified through an extensive body of scholarly work. Its relevance to practice, governance, and policy is equally clear. What has received less attention, however, is the implementation of population health data collection, preparation, cleansing, and enrichment processes for the entire ecosystem of advanced analytical explorations and applications. Indeed, the data underpinning these population health analyses and risk models primarily resides in heterogeneous data silos and repositories created for other purposes. Scaling the population health predictive efforts therefore requires establishing analytics-friendly data-accessing and -engineering processes—preferably designed as automated self-service solutions and integrated within the broader environment of ETL/ELT operations.

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

REFERENCES

[1] Adler-Milstein, J., & Jha, A. K. (2017). HITECH Act drove large gains in hospital electronic health record adoption. Health Affairs, 36(8), 1416–1422.
[2] Gottimukkala, V. R. R. (2020). Energy-Efficient Design Patterns for Large-Scale Banking Applications Deployed on AWS Cloud. power, 9(12).
[3] Ahmad, M. A., Eckert, C., & Teredesai, A. (2018). Interpretable machine learning in healthcare. Proceedings of the ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, 559–560.
[4] Ancker, J. S., Edwards, A., Nosal, S., Hauser, D., Mauer, E., & Kaushal, R. (2017). Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Medical Informatics and Decision Making, 17(1), 36.
[5] Kolla, S. K. (2021). Designing Scalable Healthcare Data Pipelines for Multi-Hospital Networks. World Journal of Clinical Medicine Research, 1(1), 1-14.
[6] Birkhead, G. S., Klompas, M., & Shah, N. R. (2015). Uses of electronic health records for public health surveillance to advance public health. Annual Review of Public Health, 36, 345–359.
[7] Blumenthal, D. (2010). Launching HITECH. New England Journal of Medicine, 362(5), 382–385.
[8] Inala, R. Advancing Group Insurance Solutions Through Ai-Enhanced Technology Architectures And Big Data Insights.
[9] Burwell, S. M. (2015). Setting value-based payment goals—HHS efforts to improve U.S. health care. New England Journal of Medicine, 372(10), 897–899.
[10] Carayon, P., Wooldridge, A., Hose, B. Z., Salwei, M., & Benneyan, J. (2020). Challenges and opportunities for improving patient safety through human factors and systems engineering. Health Affairs, 39(11), 1862–1869.
[11] Mangalampalli, B. M. (2021). Scalable Data Warehouse Architecture for Population Health Management and Predictive Analytics. World Journal of Clinical Medicine Research, 1(1), 1-18.
[12] Classen, D. C., & Bates, D. W. (2011). Finding the meaning in meaningful use. New England Journal of Medicine, 365(9), 855–858.
[13] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: Management, analysis and future prospects. Journal of Big Data, 6(1), 54.
[14] Peddi, R. K. (2021). Optimizing Case Management Workflows in Global Data Center Colocation Services. Universal Journal of Computer Sciences and Communications, 1(1), 1-21.
[15] Dixon, B. E., Grannis, S. J., & McAndrews, C. (2018). Leveraging data standards to improve interoperability in healthcare. Healthcare Informatics Research, 24(2), 91–101.
[16] Evans, R. S. (2016). Electronic health records: Then, now, and in the future. Yearbook of Medical Informatics, 25(S1), S48–S61.
[17] Reddy, V. A. R. (2021). Challenges in Standardizing Member Eligibility Data Across Multi-Payer Healthcare Ecosystems. International Journal of Medical Toxicology and Legal Medicine, 24(3), 1-19.
[18] Friedman, D. J., Parrish, R. G., & Ross, D. A. (2013). Electronic health records and population health. American Journal of Preventive Medicine, 44(6), 535–536.
[19] Segireddy, A. R. (2020). Cloud Migration Strategies for High-Volume Financial Messaging Systems.
[20] Gliklich, R. E., Leavy, M. B., & Dreyer, N. A. (2020). Registries for evaluating patient outcomes: A user's guide (4th ed.). Agency for Healthcare Research and Quality.
[21] Halamka, J. D., & Tripathi, M. (2017). The HITECH era in retrospect. New England Journal of Medicine, 377(10), 907–909.
[22] Haux, R. (2010). Medical informatics: Past, present, future. International Journal of Medical Informatics, 79(9), 599–610.
[23] Himmelstein, D. U., Wright, A., & Woolhandler, S. (2010). Hospital computing and the costs and quality of care. New England Journal of Medicine, 362(18), 1628–1638.
[24] Hripcsak, G., & Albers, D. J. (2015). Next-generation phenotyping of electronic health records. Journal of the American Medical Informatics Association, 22(1), 117–121.
[25] Kolla, S. H. (2022). Knowledge Retrieval Systems for Enterprise Service Environments. International Journal of Intelligent Systems and Applications in Engineering, 10, 495-506.
[26] Jha, A. K., DesRoches, C. M., Campbell, E. G., Donelan, K., Rao, S. R., Ferris, T. G., Shields, A., Rosenbaum, S., & Blumenthal, D. (2009). Use of electronic health records in U.S. hospitals. New England Journal of Medicine, 360(16), 1628–1638.
[27] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. Stroke and Vascular Neurology, 2(4), 230–243.

International Journal of Advanced Research in Computer Science & Technology (IJARCST)

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

[28] Mangala, N. (2021). CI/CD Pipeline Automation for Enterprise Data Artifacts Using Azure DevOps. Universal Journal of Business and Management, 1(1), 1-18.

[29] Keesara, S., Jonas, A., & Schulman, K. (2020). Covid-19 and health care's digital revolution. New England Journal of Medicine, 382(23), e82.

[30] Kern, L. M., Kaushal, R., Barrón, Y., & Abramson, E. L. (2015). Electronic health records and ambulatory quality of care. Journal of General Internal Medicine, 30(4), 496–503.

[31] Inala, R. Designing Scalable Technology Architectures for Customer Data in Group Insurance and Investment Platforms.

[32] Kho, A. N., Cashy, J. P., Jackson, K. L., Pah, A. R., Goel, S., Boehnke, J., Humphries, J. E., Kominers, S. D., Hota, B., Sims, S. A., Malin, B. A., French, D. D., & Meltzer, D. O. (2015). Design and implementation of a privacy-preserving electronic health record linkage tool in Chicago. Journal of the American Medical Informatics Association, 22(5), 1072–1080.

[33] Yandamuri, U. S. (2021). A Comparative Study of Traditional Reporting Systems versus Real-Time Analytics Dashboards in Enterprise Operations. Universal Journal of Business and Management.

[34] Kuperman, G. J. (2011). Health-information exchange: Why are we doing it, and what are we doing? Journal of the American Medical Informatics Association, 18(5), 678–682.

[35] Mandel, J. C., Kreda, D. A., Mandl, K. D., Kohane, I. S., & Ramoni, R. B. (2016). SMART on FHIR: A standards-based interoperable apps platform for electronic health records. Journal of the American Medical Informatics Association, 23(5), 899–908.

[36] Mangala, N. (2021). Optimizing Large-Scale ETL Pipelines Using Medallion Architecture on Azure Data Lake. Journal of Artificial Intelligence and Big Data, 1(1), 1-20.

[37] Mehta, N., Pandit, A., & Shukla, S. (2019). Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. Journal of Biomedical Informatics, 100, 103311.

[38] Menachemi, N., & Collum, T. H. (2011). Benefits and drawbacks of electronic health record systems. Risk Management and Healthcare Policy, 4, 47–55.

[39] Murphy, S. N., Weber, G., Mendis, M., Gainer, V., Chueh, H. C., Churchill, S., & Kohane, I. (2010). Serving the enterprise and beyond with i2b2. Journal of the American Medical Informatics Association, 17(2), 124–130.

[40] Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. JAMA, 309(13), 1351–1352.

[41] National Academy of Medicine. (2015). Improving diagnosis in health care. National Academies Press.

[42] O'Malley, A. S., Grossman, J. M., Cohen, G. R., Kemper, N. M., & Pham, H. H. (2010). Are electronic medical records helpful for care coordination? Journal of General Internal Medicine, 25(3), 177–185.

[43] Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—Big data, machine learning, and clinical medicine. New England Journal of Medicine, 375(13), 1216–1219.

[44] Rajesh Mattaparthi (2021). Unified Data Lineage and Quality Governance Framework for Multi-Source Sensor Streams in Heavy-Duty Powertrain Manufacturing. Online Journal of Mechanical Engineering, 1(1), 1-15. https://doi.org/10.31586/ojme.2021.1365

[45] Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. Health Information Science and Systems, 2(1), 3.

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

[47] Loganathan, R. (2021). Integrated Risk and Compliance Frameworks for Global Data Center Operations: A Governance-Centric Approach. Universal Journal of Computer Sciences and Communications, 1(1), 1-26.

[48] Rumsfeld, J. S., Joynt, K. E., & Maddox, T. M. (2016). Big data analytics to improve cardiovascular care: Promise and challenges. Nature Reviews Cardiology, 13(6), 350–359.

[49] Shah, S. J., Katz, D. H., Selvaraj, S., Burke, M. A., Yancy, C. W., Gheorghiade, M., Bonow, R. O., Huang, C. C., & Deo, R. C. (2015). Phenomapping for novel classification of heart failure with preserved ejection fraction. Circulation, 131(3), 269–279.

[50] Reddy, V. A. R. (2022). Designing Fault-Tolerant Data Ingestion Pipelines for High-Volume Healthcare Transactions. Frontiers in Health Informatics, 11, 861-889.

[51] Steyerberg, E. W. (2019). Clinical prediction models (2nd ed.). Springer.

[52] Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56.

[53] Vest, J. R., Campion, T. R., Kaushal, R., & HITEC Investigators. (2013). Challenges, alternatives, and paths to sustainability for health information exchange efforts. Journal of Medical Systems, 37(6), 9987.

| ISSN: 2347-8446 | www.ijarcst.org | editor@ijarcst.org |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 6, Issue 2, March–April 2023||

DOI: 10.15662/IJARCST.2023.0602007

[54] Mangalampalli, B. M. (2022). Automated Invoice Validation Systems Using Advanced SQL Analytics in Healthcare Insurance. Front Health Inform, 11.

[55] Walker, J., Pan, E., Johnston, D., Adler-Milstein, J., Bates, D. W., & Middleton, B. (2005). The value of health care information exchange and interoperability. Health Affairs, 24(Suppl. 1), W5-10–W5-18.

[56] Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to healthcare. Information & Management, 55(1), 64–79.

[57] Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment. Journal of the American Medical Informatics Association, 20(1), 144–151.

[58] Zhou, L., Soran, C. S., Jenter, C. A., Volk, L. A., Orav, E. J., Bates, D. W., & Simon, S. R. (2007). The relationship between electronic health record use and quality of care. Journal of General Internal Medicine, 22(5), 693–697.

[59] Goldstein, B. A., Navar, A. M., Pencina, M. J., & Ioannidis, J. P. A. (2017). Opportunities and challenges in developing risk prediction models with electronic health records data. Circulation: Cardiovascular Quality and Outcomes, 10(11), e003477.

[60] Davuluri, P. N. (2022). Cloud-Native Data Platform Modernization for Regulatory Compliance in Global Banking.

[61] Friedman, C. P., Wong, A. K., & Blumenthal, D. (2010). Achieving a nationwide learning health system. Science Translational Medicine, 2(57), 57cm29.

[62] Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: Towards better research applications and clinical care. Nature Reviews Genetics, 13(6), 395–405.

[63] Adler-Milstein, J., Everson, J., & Lee, S. Y. D. (2015). Sequencing of EHR adoption among U.S. hospitals. American Journal of Managed Care, 21(11), e607–e614.

[64] Friedman, C. P., Allee, N. J., Delaney, B. C., Flynn, A. J., Silverstein, J. C., Sullivan, K., & Brantley, K. L. (2020). The science of learning health systems: Foundations for a new journal. Learning Health Systems, 4(1), e10203.

[65] Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. Health Affairs, 33(7), 1123–1131.

[66] Hersh, W. R. (2020). Health informatics: Practical guide for healthcare and information technology professionals (7th ed.). Informatics Education.

[67] Cohen, I. G., Amarasingham, R., Shah, A., Xie, B., & Lo, B. (2014). The legal and ethical concerns that arise from using complex predictive analytics in health care. Health Affairs, 33(7), 1139–1147.

[68] Ben-Assuli, O. (2015). Electronic health records, adoption, quality of care, legal and privacy issues. Health Policy, 119(3), 287–294.

[69] Adler-Milstein, J., Holmgren, A. J., Kralovec, P., Worzala, C., Searcy, T., & Patel, V. (2017). Electronic health record adoption in U.S. hospitals. Journal of the American Medical Informatics Association, 24(6), 1142–1148.

[70] Hoerbst, A., & Ammenwerth, E. (2010). Electronic health records: A systematic review on quality requirements. Methods of Information in Medicine, 49(4), 320–336.