Strategic Patient Risk Stratification & Readmission Predictive Modeling for Vitality Health Network (VHN)

Prepared for: Executive Leadership Board, Chief Medical Information Officer (CMIO)
Date: 11 January 2026
Subject: Implementation of a Strategic Patient Risk Stratification Framework using the Vitality Complexity Index (VCI)

---

1. Executive Summary

The healthcare industry is undergoing a profound transformation, moving from traditional fee-for-service models toward value-based care, where financial incentives are directly linked to patient outcomes. Vitality Health Network (VHN) currently faces a critical challenge, with its 30-day readmission rate for diabetic patients reaching **18%**, significantly exceeding national benchmarks. Such underperformance triggers substantial penalties under the **Hospital Readmissions Reduction Program (HRRP)**, impacting both financial performance and patient care continuity.

To address this, we developed the **Vitality Complexity Index (VCI)**, a comprehensive risk-scoring algorithm that identifies high-risk patients prior to discharge. Utilizing a robust dataset encompassing over **100,000 patient encounters** from 130 U.S. hospitals, our analysis highlights key readmission drivers, including clinical severity, operational factors, and medication management.

By implementing **proactive risk stratification** and integrating actionable insights into nursing workflows, VHN can mitigate readmission risks, optimize resource allocation, and improve patient outcomes. Strategic recommendations include targeted follow-up for high-acuity admissions, medication stabilization protocols, and VCI-guided staffing strategies.

---

2. Business Context and Problem Statement

2.1 The Hospital Readmission Crisis

Hospital readmissions are increasingly recognized as indicators of fragmented or suboptimal care. Under the HRRP, hospitals face substantial financial penalties for excessive readmission rates, particularly in chronic conditions such as diabetes. VHN's current 30-day readmission rate of 18% highlights a **continuity of care gap**, signaling operational inefficiencies and potential risks for patients with complex chronic diseases.

2.2 Pedagogical and Industry Convergence

Healthcare data is inherently complex: high-dimensional, sparse, and semantically rich. Patient records include ICD-9 coded diagnoses, intricate medication histories, and demographic variables requiring careful interpretation. The **CMIO's mandate** emphasized moving beyond descriptive statistics toward diagnostic and prescriptive analytics: "Stop reporting on what happened. Tell us why it happened, and identify who is at risk next."

---

3. Data Methodology and Enrichment

3.1 Data Source and Scope

Our analysis leveraged the **"Diabetes 130-US Hospitals" dataset**, spanning 1999–2008 and covering 130 hospitals. This dataset provides **encounter-level details**, including patient demographics, admission types, laboratory tests, procedures, medication regimens, and discharge outcomes. It represents a real-world, high-dimensional dataset with **over 100,000 patient encounters**, making it ideal for simulating clinical decision support analytics.

3.2 Clinical Sanitation (Phase 1)

Clinical data is prone to errors and inconsistencies due to multiple EHR systems and human factors. To ensure reliable analysis, we implemented rigorous **data sanitation** procedures:

- **Standardizing Null Values:** Columns with missing entries, such as `race`, `payer_code`, and `medical_specialty`, used "?" to indicate absence. These were converted to standard **NumPy NaN** values for analytical consistency.
- **Filtering Deceased Patients:** Patients with discharge codes indicating death (e.g., 11, 19, 20, 21) were removed, as they cannot be readmitted. Inclusion would introduce **noise into readmission predictions**.
- **Feature Removal:** The `weight` column showed >90% missingness and was excluded. This limitation was documented as a **data quality constraint**.
- **Deduplication:** Exact duplicate records were removed to prevent skewing analysis.

3.3 Data Enrichment via Web Scraping (Phase 2)

ICD-9 codes in the dataset (e.g., "250.02" or "428") are cryptic to administrators. To **translate numeric codes into actionable insights**, we used web scraping:

- **Technique:** The Python libraries `requests` and `BeautifulSoup` were used to fetch human-readable disease descriptions from a public ICD-9 repository.
- **Integration:** Descriptions for the **top 20 most frequent diagnoses** were mapped into the dataset. Codes outside this range were labeled as "Other or Not in Top 20."

- **Outcome:** This enrichment provided clarity for decision-making and facilitated a **data-driven discussion with clinical leadership**.

---

4. Exploratory Data Analysis: Clinical & Operational Insights

4.1 Demographic Disparities

Analyzing age, race, and gender distributions revealed:

- Diabetic readmissions predominantly affect the **elderly population** (ages 60–80).
- Intersectional analysis by race and gender highlighted **disparities**, suggesting unequal care outcomes. For example, certain minority groups exhibited **higher readmission probabilities**, underlining the need for targeted interventions.

4.2 Medication Efficacy and Protocol

Medication usage and dosage adjustments were significant predictors of readmission:

- **Insulin vs Oral Antidiabetics:** Patients on Insulin were generally at **higher clinical risk** and exhibited higher readmission rates compared to those on oral medications such as Metformin.
- **Dosage Changes:** Patients whose medications were changed during their stay ("change" = Yes) had **elevated readmission rates**, indicating instability during hospitalization.

4.3 Operational Metrics

Operational factors also influenced readmissions:

- **Length of Stay vs Laboratory Utilization:** A positive correlation was observed between `time_in_hospital` and `num_lab_procedures`. Patients with longer stays typically underwent more lab testing.
- **Discharge Disposition:** Patients discharged to **Skilled Nursing Facilities** had higher return rates than those discharged home, emphasizing the importance of **discharge planning and post-discharge support**.

---

5. The Vitality Complexity Index (VCI) Framework

5.1 Algorithm Development

To provide a **quantitative, actionable complexity score**, we implemented the **VCI**, a variant of the LACE index. The algorithm scores patients based on four dimensions:

1. **L – Length of Stay (LOS):**
    - <1 day: 0 points
    - 1–4 days: 1 point
    - 5–13 days: 4 points
    - ≥14 days: 7 points
2. **A – Acuity of Admission:**
    - Emergency (ID 1) or Trauma (ID 7) admissions: 3 points
    - Other types: 0 points
3. **C – Comorbidity Burden (Proxy):**
    - <4 diagnoses: 0 points
    - 4–7 diagnoses: 3 points
    - ≥8 diagnoses: 5 points
4. **E – Emergency Visit Intensity:**
    - 0 prior visits: 0 points
    - 1–4 visits: 3 points
    - 4 visits: 5 points

The total **VCI_Score = L + A + C + E**.

5.2 Stratification Results

Patients were categorized into **Low (<7), Medium (7–10), and High Risk (>10)** strata. Validation visualizations confirmed:

- The **High Risk** group consistently exhibited the **highest actual readmission rates (<30 days)**.
- Medium and Low risk groups had proportionally lower readmissions.

This stratification enables **nursing staff and discharge planners** to prioritize interventions efficiently.

---

6. Strategic Recommendations

Based on EDA and VCI insights, we propose the following **data-driven strategies**:

1. **Mandatory Follow-Up for High Acuity Patients:**
    - Patients admitted via the Emergency Room or Trauma should receive a **48-hour post-discharge call**, addressing potential early readmissions.
2. **Medication Stabilization Protocols:**
    - For patients undergoing **medication dosage changes**, extend observation periods to ensure metabolic and clinical stability before discharge.
3. **VCI-Based Staffing Dashboard:**

- Integrate the **VCI Score into a live nursing dashboard** to allocate resources dynamically toward **High Risk patients**, improving continuity of care and operational efficiency.

---

7. Conclusion

By combining robust data cleaning, **ICD-9 enrichment**, exploratory analysis, and the **VCI risk stratification**, VHN can **transition from reactive reporting to proactive care management**.

The **VCI provides a quantifiable, interpretable index**, enabling staff to focus on patients with the highest likelihood of readmission, optimize resource allocation, and improve overall patient outcomes. Implementing the recommended strategies will support both **clinical excellence** and **financial stability**, reducing penalties under HRRP and advancing VHN’s commitment to value-based care.