Data Analysis and Visualization for Patient Care of Chronic Diseases in the e-Hospital

Dr. Ali Abbas
Department of Mechanical Engineering
University of Ottawa
Ottawa, Canada
aabbas@uottawa.ca

Linh Chi Hoang
School of Engineering Design and Teaching Innovation
University of Ottawa
Ottawa, Canada
lhoan033@uottawa.ca

Lakshika Paiva
School of Engineering Design and Teaching Innovation
University of Ottawa
Ottawa, Canada
lpaiv023@uottawa.ca

Kartik Banga
School of Engineering Design and Teaching Innovation
University of Ottawa
Ottawa, Canada
kbang051@uottawa.ca

***Abstract* - This project focuses on inefficiencies in chronic disease management in the Canadian context, and the lack of advanced data analytics and visualization dashboards in the e-Hospital application. We conducted an analysis of three key chronic diseases, Heart Disease, Diabetes and Lung Cancer using datasets from the e-Hospital and Kaggle. The methodology followed in-depth data analysis techniques to uncover analytics to better manage chronic diseases. The outcome proposes actionable dashboards for heart disease, diabetes, and lung cancer, enabling real-time insights into individual and aggregated disease trends, risk factors and complexities to be able to diagnose patients and offer better quality and faster personalized care without the need to repeat individual medical reports, in a cost-effective manner. The system enhances patient outcomes, supports preventative strategies, and lays the groundwork for predictive analytics and expanded disease coverage, marking a significant advancement in Canada's healthcare system.**

Keywords - data analytics visualization, e-health, e-hospital, chronic disease dashboards, doctor dashboards

I. INTRODUCTION

As the market shifts towards data-driven care, integrating real-time medical insights into patient healthcare is crucial. This data analysis and visualization project addresses inefficiencies in Canada's healthcare system, specifically the fragmented access to patient healthcare data and the underutilization of electronic health records (EHR) leading to delays in personalized care and inefficiencies in chronic disease management. Chronic diseases account for 67% of annual deaths in Canada [1], with risk factors including lifestyle habits and environmental conditions. This project addresses these challenges by offering a centralized platform for real-time health analytics for the most commonly prevalent chronic diseases, improving clinical decision-making and empowering patients. This project aims to integrate real-time analytics and advanced visualizations into the e-Hospital app to enhance patient care delivery, improve diagnosis timelines, and support preventative healthcare strategies.

Objectives of the Project:

- Develop actionable dashboards for chronic disease management.
- Empower healthcare providers with data-driven insights.
- Enhance the e-Hospital platform to provide efficient access to critical health metrics for patients and doctors.

II. SHORT LITERATURE REVIEW

A. The Problem

A study of Canada's health reveals that 67% of all deaths each year are caused by four major chronic conditions, cancer, diabetes, cardiovascular and chronic respiratory disease [1]. Common risk factors for chronic diseases include overweight, unhealthy eating, physical inactivity, smoking and alcohol including other environmental factors [2]. In 2013, the Public Health Agency of Canada set out a strategic priority to help Canadians by surveillance transformation or enhanced use of data for action to overcome the challenge of chronic diseases [2]. However, there are many issues that still exist in this domain.

In Canada, access to patient healthcare data is one of the main challenges faced by the healthcare system to make decisions on patient health. Resolving this issue is a priority for the patients, clinicians and Governments. Although the importance of patient healthcare has been highlighted in the past, data on the delivery of healthcare is underused [3]. In fact, the implementation of Electronic Health Records (EHR) in medical clinics over the past decade has provided an opportunity to improve patient care. However, EHR data have been underused in

Canadian patient healthcare, depriving the healthcare system of a rich source of timely data [4].
Patient data have previously been used in Canada to develop dashboards, which are visualization tools based on both text and graphic support, allowing for timely monitoring of clinical or organizational outcomes [4]. Dashboards are quality improvement tools that can optimize the performance of health organizations, document patients' health needs and support timely decision making based on data made available in real time [5] [6]. Hence, we evaluate pros and cons of existing solutions in the Canadian context in relation to chronic disease management.

B. A Review of Existing Solutions on Chronic Disease Dashboards

Literature presents several electronic health dashboard solutions used in healthcare organizations across Canada. British Columbia has a chronic disease dashboard which is a publicly available interactive data tool with visualizations of changes over time in patterns of chronic conditions, incidents, prevalence, which is based on physician visits, hospital admissions, and prescription drug dispensations. These are updated annually and are available by age group, sex and geographic health region. The dashboard provides useful situational awareness and contextual information that can inform priorities and strategies for disease prevention, management and care at the practice, regional and provincial levels [1]. However, there is no evidence to ensure the breadth and depth of chronic disease prevention and cure, and patient access to their health data as analytics dashboards.

Another research study focusing on the Edmonton Obesity Staging System (EOSS) combines with body-mass index (BMI) data to create clinical dashboards that enable improved functional and prognostic assessment for patients. This study created a clinical dashboard for obesity that highlighted the severity and stage of obesity, EOSS and the prevalence of related comorbidities. Further, it identifies that making this information easily accessible for individual clinical care and practice-level quality improvement may advance obesity care [7]. However, this dashboard was limited to the relationship of BMI with chronic disease data while many other risk factors affect its prevalence.

Moreover, research on health analytics in the US claims that chronic disease dashboards improve both the efficiency and accuracy of acquiring data needed for high quality patient-specific care such as for diabetes [8] as it speeds up retrieval of medical records and reduced re-ordering of medical tests by doctors.
In conclusion, analytics dashboards are crucial in Canada's medical context, significantly enhancing patient-doctor decision-making for chronic diseases. They provide near real-time, data-driven insights into a patient's health, consolidating vast amounts of information into actionable visual summaries. This enables doctors to make informed, timely decisions on diagnosis, treatment plans, and chronic disease management. For patients, dashboards offer a clearer understanding of their health metrics, promoting engagement and proactive care. In a healthcare system often burdened by long wait times and fragmented data, analytics dashboards streamline communication, improve efficiency, and support personalized care, ultimately leading to better health outcomes and reduced strain on the system. Hence, in this project, we focus on overcoming the drawbacks of these projects and developing patient and doctor focused healthcare dashboards with a broader and in-depth view for managing chronic diseases in Canada.

As a result, we asked the following questions to form the basis of the analysis for Diabetes, Heart Disease and Lung Cancer:

- What is the incidence or prevalence of these diseases among patients?
- What is the prevalence of risk factors of these diseases against demographics?
- What are the relationships or complexities involved in these diseases with other risk factors and chronic diseases?

In the next section, we present the methodology we adopted to answer these questions.

III. METHODOLOGY

We obtained datasets from several sources to perform chronic disease analysis. First, we extracted the patient data from the e-Hospital database and added more patient data to the dataset for better analysis results. Then we obtained disease specific datasets for Diabetes, Heart Disease and Lung Cancer from Kaggle. These datasets were pre-processed to clean and using the methods mentioned in section 3.1 and then analytics and visualization were formed as described in section 3.2.

Importantly, there are two phases to this project:

Minimum Viable Product Phase: During the MVP Phase, the focus was on **heart disease**. This phase delivered key dashboards: the Risk Assessment Dashboard, which provided age, BMI, and gender-based analyses of heart disease risk; the Lifestyle Recommendation Dashboard, which highlighted the impact of lifestyle changes such as exercise and diet on heart health; and the Complex Case Management Dashboard, which studied the relationships between heart disease and comorbidities like diabetes, stroke, and asthma. The datasets used in this phase included the Heart Disease Dataset, which contained demographic, BMI, and cardiovascular indicators; the

Blood Sugar Levels Dataset, which tracked monthly blood sugar readings; and the Patient Data, which provided comprehensive demographic and health records.

Final Phase: In the Final Phase, the scope was expanded to include **diabetes** and **lung cancer**. This phase introduced enhanced dashboards with detailed analyses of the risk factors associated with these diseases and integrated them into the e-Hospital app for direct use by healthcare providers. The Diabetes Dataset was utilized to examine glucose levels, BMI, and demographics, while the Lung Cancer Dataset focused on smoking habits, symptoms, and age distribution. These integrations significantly improved the app's ability to provide comprehensive and actionable insights across multiple chronic diseases.

A. Data Preparation

1) Data Cleaning

The first stage was data cleaning, which addressed various issues such as missing values, outliers, and inconsistencies. Missing values in critical fields like glucose levels, BMI, and blood pressure were either imputed with median values or excluded if their absence significantly impacted the analysis. For example, missing glucose readings in the diabetes dataset were interpolated where possible, while invalid or unrealistic entries, such as extremely high BMI values, were identified as outliers and capped to maintain statistical integrity. Additionally, categorical variables like gender and blood group were standardized to ensure consistency, with variations in capitalization or format corrected for uniformity.

```
#1. Remove unnecessary columns (like unnamed index columns if present)
if 'Unnamed: 0' in df.columns:
    df.drop(columns=['Unnamed: 0'], inplace=True)
# 2. Handle missing values
# Replace missing numerical values with median (for columns like age, height, etc.)
numerical_cols = df.select_dtypes(include=['float64', 'int64']).columns
for col in numerical_cols:
    df[col].fillna(df[col].median(), inplace=True)

# Replace missing categorical values with mode (for columns like gender, blood group, etc.)
categorical_cols = df.select_dtypes(include=['object']).columns
for col in categorical_cols:
    df[col].fillna(df[col].mode()[0], inplace=True)
# 3. Convert categorical columns to appropriate types (e.g., gender, blood group)
df['Gender'] = df['Gender'].astype('category')
df['BloodGroup'] = df['BloodGroup'].astype('category')
# 5. Ensure columns like age, height, weight are of numeric types (and within reasonable ranges)
df['Age'] = pd.to_numeric(df['Age'], errors='coerce')
df['height'] = pd.to_numeric(df['height'], errors='coerce')
df['weight'] = pd.to_numeric(df['weight'], errors='coerce')
```

Fig. 1. Examples of data cleaning

2) Feature Engineering

Feature engineering was the next stage, aimed at deriving new variables to enhance the predictive power of the datasets. Binary target variables were created to indicate the presence or absence of specific diseases, such as heart disease and diabetes. For the lung cancer dataset, symptoms such as coughing and wheezing were encoded into numerical formats, enabling effective correlation analysis. Relationships between health metrics, such as BMI and blood pressure, were quantified through correlation matrices, providing valuable insights into the interdependencies of various risk factors.

3) Data Merging and Matching

The final stage involved integrating the different datasets into a unified database. Patient IDs were used as the primary key to merge datasets, such as demographic data from the patient records, time-series data from blood sugar levels, and diagnostic data from the disease-specific datasets. This integration allowed for a holistic view of each patient’s health profile, facilitating comprehensive analysis and visualization. For example, merging the blood sugar dataset with patient demographics enabled the creation of a time-series dashboard that tracked blood sugar trends over a year.

B. Data Analysis & Visualization Techniques

1) Data Analysis Techniques

- Descriptive Statistics

Descriptive statistics were used to summarize the key characteristics of the datasets. Measures such as mean, median, standard deviation, and frequency counts provided an overview of variables like BMI, glucose levels, and age distribution. For instance, in the diabetes dataset, the mean BMI and glucose levels were analyzed to understand the baseline health status of patients and identify significant deviations associated with diabetes.

- Chi-square

The Chi-square test was applied to determine the relationship between categorical variables. For example, in the lung cancer dataset, it was used to analyze the association between smoking habits and lung cancer incidence. This test highlighted significant correlations, such as higher cancer rates among smokers, which guided the development of risk factor dashboards.

```
    # Chi-square test for BMI Category and Heart Disease
    bmi_vs_hd_ct = pd.crosstab(df['BMI_Category'], df['HeartDisease'])
    chi2, p, dof, ex = stats.chi2_contingency(bmi_vs_hd_ct)
    print(f"\nChi-Square Test for BMI vs Heart Disease: Chi-Square: {chi2}, P-value: {p}")

else:
    print("BMI column not found in the dataset.")

Chi-Square Test for BMI vs Heart Disease: Chi-Square: 966.3725732313598, P-value: 6.916310408304961e-208
```

Fig. 2. Example of a Chi-square test

- T-Test and ANOVA

These tests were employed to compare means across groups and assess significant differences. For example, t-tests were used in the heart disease analysis to compare BMI and cholesterol levels between patients with and without heart disease. ANOVA was applied in the lifestyle dashboard to evaluate differences in average sleep duration among patients categorized by their heart disease status, providing insights into the role of sleep as a protective factor.

```
### 4. ANOVA Test for Continuous Variables ###

# ANOVA Test for Sleep Time and Heart Disease
anova_sleep = stats.f_oneway(df[df['HeartDisease'] == 0]['SleepTime'], df[df['HeartDisease'] == 1]['SleepTime'])
print(f"\nANOVA Test for Sleep Time vs Heart Disease: F-statistic: {anova_sleep.statistic}, P-value: {anova_sleep.pvalue}")

# ANOVA Test for General Health and Heart Disease
gen_health_groups = [df[df['GenHealth'] == cat]['HeartDisease'] for cat in df['GenHealth'].unique()]
anova_general_health = stats.f_oneway(*gen_health_groups)
print(f"\nANOVA Test for General Health vs Heart Disease: F-statistic: {anova_general_health.statistic}, P-value: {anova_general_health.pvalue}")
```

```
ANOVA Test for Sleep Time vs Heart Disease: F-statistic: 22.173760762336727, P-value: 2.491586575822921e-06
```

Fig. 3. Example of ANOVA Test

- Correlation Analysis

Correlation analysis measured the strength and direction of relationships between numerical variables. This method was critical in identifying dependencies, such as the correlation between BMI and blood pressure in the diabetes dataset. The analysis revealed that patients with higher BMI often exhibited elevated blood pressure, emphasizing the need for weight management in diabetes prevention strategies.

- Logistic Regression

Logistic regression models were used to predict the likelihood of disease occurrence based on multiple independent variables. In the heart disease analysis, the model quantified the impact of factors like age, BMI, and cholesterol levels on the risk of developing heart disease. Coefficients derived from the regression provided healthcare providers with actionable insights into prioritizing interventions for high-risk patients.

```
### LOGISTIC REGRESSION ANALYSIS ###

# Convert categorical variables into numeric form if needed
df['Sex'] = df['Sex'].apply(lambda x: 1 if x == 'Male' else 0)  # 1 for Male, 0 for Female

# Map age categories to numbers for regression
df['AgeCategory'] = df['AgeCategory'].map({
    '18-24': 1, '25-29': 2, '30-34': 3, '35-39': 4, '40-44': 5,
    '45-49': 6, '50-54': 7, '55-59': 8, '60-64': 9, '65-69': 10,
    '70-74': 11, '75-79': 12, '80 or older': 13
})

# Ensure BMI is numeric (already handled above)
# Prepare independent variables (AgeCategory, BMI, Sex)
X = df[['AgeCategory', 'BMI', 'Sex']]

# Add a constant for the intercept
X = sm.add_constant(X)

# Dependent variable (Heart Disease)
y = df['HeartDisease']

# Fit the logistic regression model
logit_model = sm.Logit(y, X).fit()

# Print the summary of the regression
print(logit_model.summary())
```

Fig. 4. Example of Logistic Regression Analysis

2) Data Visualizations

- Bar Plots and Count Plots

Bar plots were utilized to compare the distribution of categorical variables, such as the number of smokers versus non-smokers in the lung cancer dataset. Count plots further segmented this data by additional variables like age group, allowing for a detailed examination of subcategories (e.g., smoking rates among patients aged 50–70).

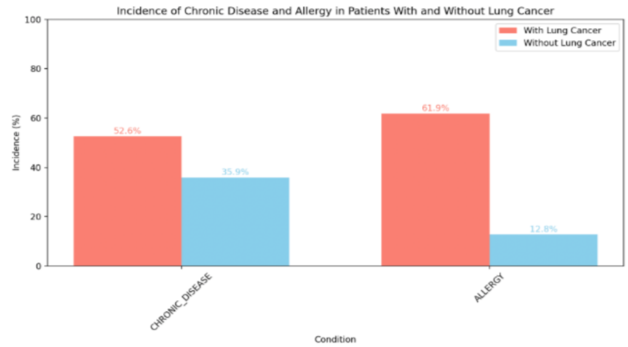

Fig. 5. Bar Plot for Chronic Disease and Allergy, with or without Lung Cancer

- Line Plots

Line plots were used to illustrate trends in time-series data. For instance, the blood sugar levels dataset leveraged line plots to visualize monthly glucose readings over a year for individual patients. These visualizations enabled doctors to monitor fluctuations and assess the effectiveness of ongoing treatments.

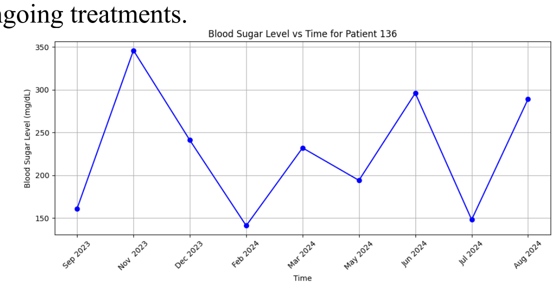

Fig. 6. Line Plot for Patient Blood Sugar Levels each Month

- Correlation Heatmaps

Correlation heatmaps provided a visual representation of relationships between multiple variables. For example, in the heart disease analysis, a heatmap highlighted the correlation between factors like age, BMI, and cholesterol levels. This visualization helped identify key risk factors and their combined impact on heart disease prevalence.

```
Correlation Matrix for continuous variables and Heart Disease:
                  BMI  SleepTime  HeartDisease
BMI          1.000000  -0.051822      0.051803
SleepTime   -0.051822   1.000000      0.008327
HeartDisease 0.051803   0.008327      1.000000
```

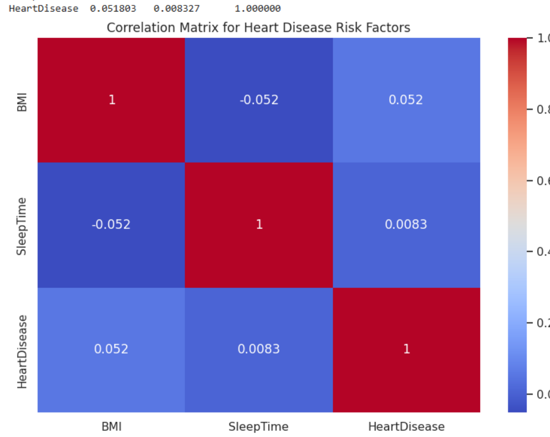

Fig. 7. Correlation Heat map for Heart Disease Risk Factors

- Box Plots

Box plots were used to display the distribution and variability of numerical data. In the lifestyle dashboard, box plots showed the range of physical activity hours among patients, segmented by their heart disease status. Outliers, such as extremely sedentary or highly active patients, were easily identified through this visualization.

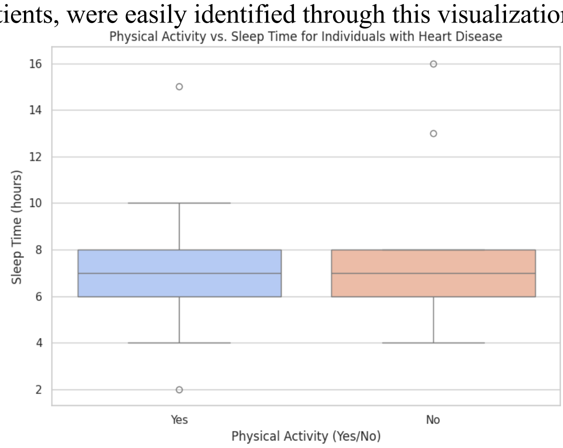

Fig. 8. Box Plot for Physical Activity vs. Sleep Time with Heart Disease

- Histograms

Histograms were employed to visualize the frequency distribution of continuous variables like BMI and glucose levels. For instance, the histogram for glucose levels in the diabetes dataset revealed that most patients fell within the pre-diabetic range, emphasizing the need for early interventions.

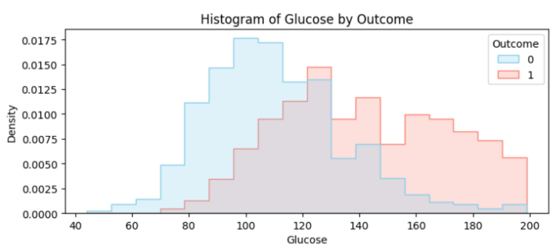

Fig. 9. Histogram for Glucose by Outcome

- Scatter Plots

Scatter plots were used to explore relationships between two numerical variables. In the diabetes analysis, scatter plots demonstrated the relationship between BMI and glucose levels, with clusters indicating high-risk groups for diabetes.

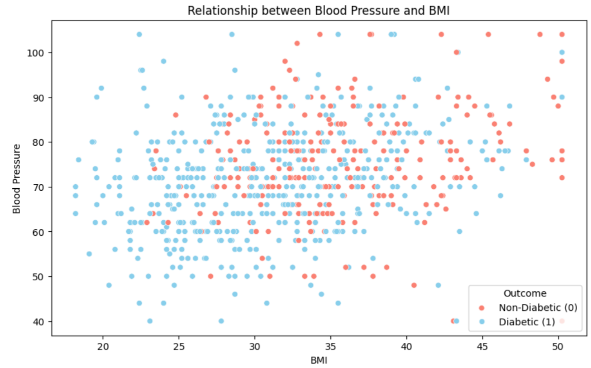

Fig. 10. Scatter Plot for Blood Pressure and BMI

- Pie Charts

Pie charts visualized proportions within the data. For example, the lung cancer dataset used pie charts to represent the percentage of patients experiencing symptoms like coughing and wheezing. This simplified the understanding of symptom prevalence.

IV. Results and Discussion

The e-Hospital project yielded significant insights into the risk factors, demographic trends, and correlations associated with chronic diseases like heart disease, diabetes, and lung cancer. The results are summarized below, with detailed analyses for each disease.

A. Heart Disease

1) Key Risk Factors

The analysis revealed that heart disease risk is strongly influenced by age, BMI, and lifestyle factors such as physical inactivity and poor diet. Patients above the age of 60 and those with a BMI greater than 30 were particularly at high risk. Physical inactivity and a lack of regular exercise compounded the risk.

2) Demographics and Prevalence

The highest prevalence of heart disease was observed in patients aged 60–64. Men exhibited slightly higher rates of heart disease compared to women, although the difference was not statistically significant.

3) Thresholds and Patterns:

- Patients with elevated cholesterol levels (above 240 mg/dL) showed a 70% increased likelihood of developing heart disease.
- Physical inactivity was identified in 80% of patients diagnosed with heart disease, emphasizing the role of regular exercise in prevention.

4) Correlation Insights

A strong positive correlation ($r = 0.68$) was observed between BMI and cholesterol levels, indicating that weight management could be a key strategy in reducing heart disease risk.

B. Diabetes

1) Key Risk Factors

Elevated glucose levels, high BMI, and hypertension were identified as the primary predictors of diabetes. Glucose levels above 126 mg/dL were a definitive indicator of diabetes diagnosis.

2) Demographics and Prevalence

- The highest prevalence was noted in patients aged 30–50, highlighting the early onset of diabetes compared to other chronic diseases.
- Obesity (BMI > 30) was prevalent in 45.8% of patients diagnosed with diabetes.

3) Thresholds and Patterns

- Patients with pre-diabetic glucose levels (100–125 mg/dL) showed a 68.5% progression to diabetes, underscoring the need for early intervention in at-risk populations.
- Blood pressure above 140/90 mmHg was strongly associated with diabetes, present in 72% of diagnosed cases.

4) Correlation Insights

A moderate correlation (r = 0.52) was observed between BMI and blood pressure, suggesting that interventions targeting weight loss could help in managing hypertension and reducing diabetes risk.

C. Lung Cancer

1) Key Risk Factors

Smoking was confirmed as the leading cause of lung cancer, with 85% of diagnosed patients having a history of smoking. Symptoms like persistent coughing, chest pain, and wheezing were identified as critical indicators of early-stage lung cancer.

2) Demographics and Prevalence

- Lung cancer prevalence was higher in men (55%) compared to women (45%).
- The highest incidence was observed in patients aged 50–70.

3) Symptoms and Comorbidities

- Approximately 60% of patients reported persistent coughing, while 55% experienced wheezing.
- Comorbidities such as chronic diseases (52.6%) and allergies (61.9%) were frequently present in lung cancer patients, complicating diagnosis and treatment.

4) Thresholds and Patterns

- Patients who smoked more than 20 cigarettes daily for over 10 years showed a 90% increased likelihood of developing lung cancer.
- Yellowing of fingers, an indicator of prolonged tobacco use, was observed in 70% of diagnosed cases.

5) Correlation Insights

Symptoms such as coughing and wheezing showed a strong correlation with smoking habits (r = 0.74), reinforcing the need for targeted smoking cessation programs to reduce lung cancer incidence.

D. Cross-Disease Insights

1) Lifestyle Factors

Lifestyle factors, including physical inactivity, poor diet, and inadequate sleep, emerged as common contributors to multiple chronic diseases. Patients who engaged in regular exercise (at least 150 minutes per week) exhibited significantly lower risks for both heart disease and diabetes.

- Demographic Patterns

Age was a universal risk factor across all three diseases. However, while heart disease predominantly affected older adults, diabetes showed an earlier onset, and lung cancer incidence peaked in the middle-aged population.

- Comorbidities

Comorbid conditions like hypertension, asthma, and kidney disease were frequently observed in patients with chronic diseases, highlighting the need for integrated healthcare approaches to manage multiple conditions effectively.

E. Implementation and Integration

The implementation of the project into e-Hospital platform involved developing a robust system architecture and effectively integrating advanced data analysis tools and visualizations into the e-Hospital app. This process ensured real-time insights for healthcare providers while maintaining scalability and efficiency.

The backend was built using **FastAPI**, a high-performance framework for developing RESTful APIs. This allowed the system to efficiently handle requests, such as retrieving patient data or generating real-time visualizations. The APIs were designed with modular endpoints tailored to specific analyses, such as heart disease risk assessments, diabetes insights, and lung cancer symptom trends. For instance, the /heartDisease/riskAssessment endpoint provided instant access to risk factor summaries for healthcare providers.

On the front end, the dashboards were integrated directly into the e-Hospital app, offering an intuitive interface for doctors and patients. These dashboards created using Python libraries like **Matplotlib**, **Seaborn**, and **Pandas**, displayed clear and actionable visualizations, such as time-series trends, correlation heatmaps, and distribution plots. The user interface was designed to prioritize ease of use, with customizable views for patient profiles and disease-specific insights.

The system's scalability was ensured by organizing data storage and processing workflows for effective expansion. Patient data from multiple datasets, including demographic details, glucose levels, and smoking history, were merged into a unified database, enabling comprehensive analyses across diseases. By leveraging efficient data handling practices, the project ensured accurate results without compromising performance.

This robust implementation not only streamlined the integration of advanced analytics but also laid the foundation for future enhancements, such as predictive analytics and broader disease coverage. The result was a powerful, user-friendly system that empowered healthcare providers to make informed decisions and improve patient care outcomes. The implemented patient analytics dashboard can be viewed at: https://www.e-hospital.ca/doctor/patientpage?patientId=306

The following diagram illustrates the patient dashboard.

Fig. 11. Patient dashboard with profile summary and blood sugar levels

The implemented disease analytics dashboards can be viewed at: https://www.e-hospital.ca/doctor/DiseaseAnalytics
The following diagrams illustrate the disease analytics dashboards for heart disease, diabetes and lung cancer.

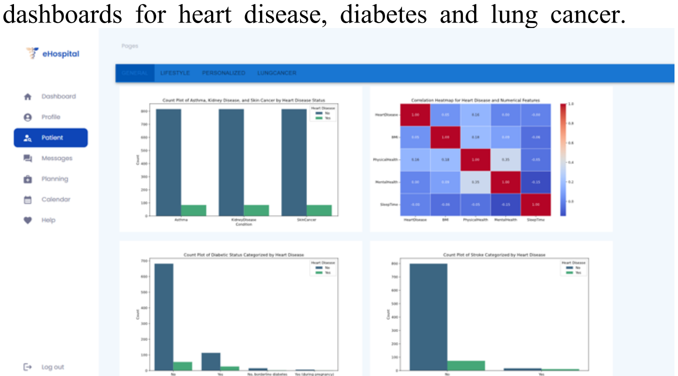

Fig. 12. General Analytics – heart disease

Fig. 13. Lifestyle Analytics – heart disease

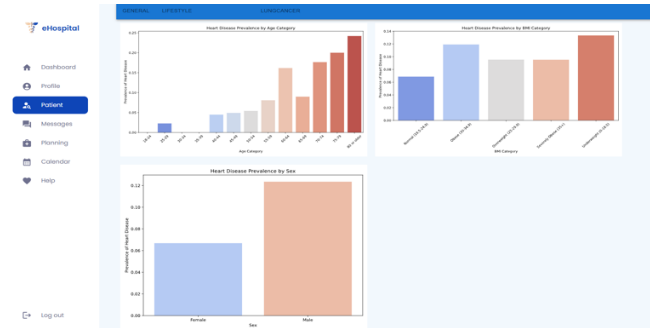

Fig. 14. Personalized Dashboards for Diabetes

Fig. 15. Dashboards for Lung Cancer

V. Implications

A. Preventative Strategies

The findings emphasize the critical need for preventative strategies and early interventions to combat chronic diseases effectively. Promoting healthier lifestyles through community health initiatives, such as exercise programs and dietary education, can significantly reduce the prevalence of risk factors like obesity and hypertension. Screening programs targeting at-risk populations, such as those with pre-diabetic glucose levels or high BMI, can enable early detection and management of diseases like diabetes and heart disease, preventing progression to severe stages.

A. Clinical Interventions

Clinically, the results highlight the importance of personalized care approaches. Heart disease interventions should prioritize managing cholesterol levels and BMI, while diabetes care must focus on glucose monitoring and weight management. Lung cancer prevention efforts need to be bolstered with targeted smoking cessation programs, as smoking remains a predominant risk factor. Integrating these findings into patient care can enhance treatment outcomes and reduce disease burdens.

B. *Policy Recommendations*

On a policy level, there is a clear need for public health policies that incentivize healthy behaviors and early disease detection. Subsidized health screenings and educational campaigns can raise awareness and encourage proactive health management. By addressing these findings through collaborative efforts, healthcare providers and policymakers can work toward improved patient outcomes and a more resilient healthcare system.

VI. Conclusion

This project successfully addressed critical inefficiencies in chronic disease management within Canada's healthcare system by integrating advanced data analytics and visualization into the e-Hospital app. Through actionable dashboards, we enabled healthcare providers to make informed, timely decisions, enhancing patient care and supporting preventative strategies. The analyses revealed key insights into the risk factors, prevalence, and correlations of heart disease, diabetes, and lung cancer, emphasizing the importance of targeted interventions and lifestyle modifications.

The robust implementation of the system, powered by FastAPI and Python-based visualization libraries, ensures scalability and user-friendliness. The project not only meets current healthcare needs but also lays the groundwork for future enhancements, including predictive analytics and expanded disease coverage. By fostering collaboration and data-driven care, this initiative represents a significant step toward a more resilient, efficient, and proactive healthcare ecosystem in Canada.

VII. Future and Recommendations

A. *Broader Disease Coverage*

Expanding the analysis to cover additional chronic conditions, such as arthritis, hypertension, and various forms of cancer, presents a significant opportunity to broaden the impact of the e-Hospital app. By incorporating datasets for these diseases, the platform can provide a more comprehensive view of patient health and support tailored prevention and treatment strategies.

B. *Predictive Analytics*

Integrating predictive analytics using machine learning models is another key recommendation. Predictive tools can forecast disease progression and identify at-risk populations, enabling earlier interventions. For example, a machine learning model could predict diabetes onset based on glucose trends and BMI, allowing healthcare providers to implement preventative measures.

C. *Cross-Institutional Collaboration*

Finally, fostering cross-institutional collaboration through the sharing of anonymized data insights can benefit public health research and policy making. This approach would enable broader analysis across populations, leading to improved strategies for disease prevention and management on a systemic level.

VIII. References

[1] K. Smolina *et al.*, “Understanding chronic conditions in BC | British Columbia Medical Journal,” *BC Medical Journal*, vol. 66, no. 5, pp. 178–180, 2024.
[2] Public Health Agency of Canada, *Preventing Chronic Disease Strategic Plan 2013–2016*. Canada: Public Health Agency of Canada, 2013.
[3] S. T. Wong, S. Johnston, F. Burge, and K. McGrail, “Value in Primary Healthcare - Measuring What Matters?,” *Healthc Pap*, vol. 18, no. 4, pp. 58–67, Dec. 2019, doi: 10.12927/hcpap.2019.26028.
[4] M. Breton *et al.*, “Use of Electronic Medical Record Data to Create a Dashboard on Access to Primary Care,” vol. 18, pp. 72–88, May 2023, doi: 10.12927/hcpol.2023.27092.
[5] B. Ehsani-Moghaddam, K. Martin, and J. A. Queenan, “Data quality in healthcare: A report of practical experience with the Canadian Primary Care Sentinel Surveillance Network data,” *Health Inf Manag*, vol. 50, no. 1–2, pp. 88–92, 2021, doi: 10.1177/1833358319887743.
[6] A. G. Singer, L. Kosowan, N. Nankissoor, R. Phung, J. L. P. Protudjer, and E. M. Abrams, “Use of electronic medical records to describe the prevalence of allergic diseases in Canada,” *Allergy Asthma Clin Immunol*, vol. 17, no. 1, p. 85, Aug. 2021, doi: 10.1186/s13223-021-00580-z.
[7] R. Swaleh *et al.*, “Using the Edmonton Obesity Staging System in the real world: a feasibility study based on cross-sectional data,” *Canadian Medical Association Open Access Journal*, vol. 9, no. 4, pp. E1141–E1148, Oct. 2021, doi: 10.9778/cmajo.20200231.
[8] R. J. Koopman *et al.*, “A Diabetes Dashboard and Physician Efficiency and Accuracy in Accessing Data Needed for High-Quality Diabetes Care,” *The Annals of Family Medicine*, vol. 9, no. 5, pp. 398–405, Sep. 2011, doi: 10.1370/afm.1286.