Projecting Future Healthcare Needs for the State of Florida

Lucas Da Silva, Samuel Soares Walton, Bryan Lu, Gehan Velivitiya, Sruthi Soundarajan, Arjun Mohankumar, Gokul Ram, Gowtham Mani, Harish Suresh, Jeremy Bouhadana, Kaarthicsudhan Mani

1. INTRODUCTION

With the US population aging, demand for health services will continue to grow as the proportion of older populations increases [1]. Florida remains one of the states with the highest proportion of elderly individuals (65+), and this number is projected to grow in the coming years and decades [2] [3]. With this in mind, understanding the current and future need for healthcare services and facilities is essential to maximize efficiency, lower costs, and provide high quality critical care to the residents of the state [4]. In fact, some research suggests a close association between acceptable healthcare resource utilization and improving the overall patient experience and satisfaction [5]. While several attempts existed to predict hospitalizations and resource demand in Florida, especially during COVID-19 [6], there is a gap in county-level projections for some of the state's most prominent illnesses.

In this paper, the authors intend to use statistical and machine learning (ML) forecasting to project demand for colorectal cancer (CRC), chronic obstructive pulmonary disease (COPD), coronary heart disease (CHD), and mental illness for the state of Florida by county and report on counties that would have gaps in resources based on these projections. The aim is to create a comprehensive and cohesive insight into the current and future needs of the state down to the county level. This will allow governments and private entities to better prepare for these four illnesses. The selection of the four diseases is not arbitrary, but rather reflects populations of interest geared towards Florida's needs, as outlined in the below sections.

1.1 Colorectal Cancer (CRC)

CRC, a term referring to cancers of the colon and/or rectum [7], is the third most common cancer and the second most common cause of cancer death [8]. CRC also disproportionately impacts elderly populations. In fact, almost 60% of all cases are diagnosed in patients 70 years or older [9]. Considering the increased aging of Florida's population, CRC a one of the focal points of this analysis. Generally, CRC develops when epithelial cells acquire a certain abnormality that leads to uncontrolled, rapid growth [10].

The primary prevention tactic for CRC is appropriate screening for at-risk populations [11]. Generally, routine colorectal cancer screenings begin at age 45, and are intended to screen for tissue growths that are not yet cancerous [12]. There are several screening methods for colorectal cancer, including stool tests, sigmoidoscopies, and colonoscopies, but the latter two methods are common and may require prolonged hospital visits [13]. Treatments for CRC are based on the diagnosed stage of the cancer and the individual patient needs and may require surgery or a combination of surgery and chemotherapy [14].

1.2 Coronary Heart Disease (CHD)

CHD, characterized by reduced blood flow to the heart muscle due to plaque buildup in the coronary arteries, represents one of the most significant health challenges globally [15][16]. This condition can lead to severe outcomes such as heart attacks, heart failure, and death [17], making it the leading cause of mortality not only in Florida but across the United States [18]. Given that heart diseases and strokes continue to be major causes of disability and a significant contributor to increases in healthcare costs in the United States [18], CHD is yet another relevant disease studied in this research.

The approach to analyzing CHD cases in Florida centers around the importance of effective early detection and management to mitigate the disease's impact. The research strategy proposed focuses on the comprehensive tracking of critical diagnostic and therapeutic procedures crucial for CHD identification and treatment, including coronary angiographies [19], percutaneous coronary interventions (PCI) [20], and coronary artery bypass grafting (CABG) surgeries [21]. These procedures are instrumental in assessing coronary artery blockages and determining the most appropriate intervention to restore blood flow to the heart muscle with the goal of preventing severe cardiac events.

1.3 Chronic Obstructive Pulmonary Disease (COPD)

COPD is a significant public health issue [22], characterized by a persistent limitation of airflow into the lungs due to airway and/or alveolar abnormalities and is often attributed to exposure to harmful particles or gases [23]. This condition, primarily caused by smoking and environmental factors [24], requires long-term care and rehabilitation distinguishing it from acute respiratory conditions like pneumonia or influenza [25]. COPD is the most common type of chronic lower respiratory disease and ranks as one of the top five causes of death in Florida [26], highlighting the need for targeted healthcare strategies to effectively anticipate and deal with this disease’s future healthcare demands.

The cornerstone of COPD diagnosis in the U.S. healthcare system is spirometry, a pulmonary function test that measures lung capacity and confirms airflow limitations [27]. Additionally, oximetry, a non-invasive fingertip clip test provides a quick assessment of blood oxygen levels, which can be low in COPD patients [27]. To gain a more detailed picture of blood gas composition, an arterial blood gas (ABG) test might be employed [27]. Imaging tools like chest X-rays and CT scans are valuable for differential diagnosis but are not definitive for COPD [28]. Bronchoscopy and allergy skin testing are not routinely used for diagnosis but may be employed in specific cases [28].

Regardless of age, smoking cessation remains the single most critical intervention for COPD management [27]. Medications play a significant role, with bronchodilators (long-acting muscarinic antagonists [LAMAs] and long-acting beta2-agonists [LABAs]) that relax airways to ease breathing [28]. Inhaled corticosteroids may be prescribed for individuals with more severe COPD or have frequent flare-ups, particularly in cases associated with asthma [27]. Beyond medications, pulmonary rehabilitation includes combining exercise and education, which is crucial for improving symptoms, exercise capacity, and the overall well-being in patients of all ages [27]. Typically, in advanced stages of COPD, supplemental oxygen therapy is often employed in patients with consistently low blood oxygen levels [27][29].

In the most severe cases of COPD, surgery becomes an option. This includes lung transplants for replacing diseased lungs, lung volume reduction surgery (LVRS) to remove damaged tissue and improve remaining lung functions, and bullectomy to remove large air pockets (bullae) that can compress healthy lung tissue and worsen symptoms. Bullectomy can significantly improve breathing for suitable patients, but careful evaluation is needed as this procedure is not appropriate for everyone [30].

1.4 Mental Illness

Mental illness represents a wide range of mental health conditions that can affect an individual's mood, thinking, and behavior [31]. Examples of mental illness include depression, anxiety disorders, schizophrenia, eating disorders, and addictive behaviors [31]. Considering recent statistics from the State of Florida that highlight that 63.5% of adults with mental illness do not receive necessary care [32] and that suicide is the leading cause of death among those aged 5 through 54 years [33], mental health emerges as a critical area of focus within the state. In addition, Florida ranks 43rd nationally in access to mental health services, alongside one of the lowest per capita support for such services [32].

The complexity of diagnosing and treating mental illness arises from its multifaceted nature, which can be influenced by genetic, biological, environmental, and psychological factors [34]. Given this broad spectrum, it is often difficult to accurately diagnose and treat mental health disorders. Therefore, this analysis focuses on examining historical hospitalization data across all Florida counties and age groups to provide a tangible metric to gauge mental health needs and trends within specific age groups.

2. METHODS

This study utilized several forecasting methods to project Florida's population growth based on sex and age. The three methods include autoregressive integrated moving average (ARIMA), extreme gradient boosting (XGBoost), and random forest (RF). By using the model with the most accurate projections, projected rates on current demand from the four illnesses were used to better understand the needed changes in resources to service these diseases considering the changing population demographics. Since projections are based on county-level statistics, concerns for each county were isolated for each of the four studied diseases.

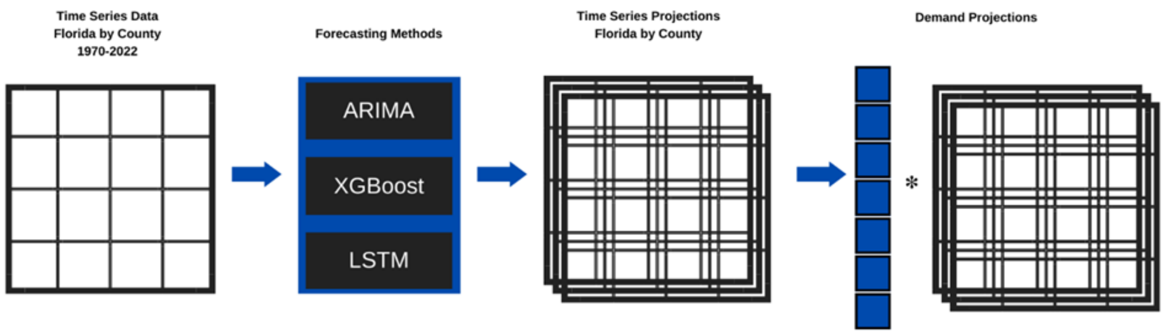


Figure 1. Overview of the forecasting workflow.

2.1 Data Acquisition

The primary data source for population growth within each county is based on the Florida Legislature's Office of Economic and Demographic Research's Midyear Population Estimates. These estimates were extracted and filtered from the Florida Health Community Health Assessment Resource Tool Set (CHARTS) [35]. This resource allowed for forecasting based on population data by county, which was filtered by specific age and sex groups. The age groups available included <1, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, and 85+ and the sex categorizations were male and female. The data encapsulated all 67 counties in Florida and was isolated for the years 1970 to 2022.

Data pertaining to the availability of hospital beds by county was also obtained through the CHARTS database, which provided a detailed perspective on the capacity for in-patient care spanning the years 2014 to 2023 [35]. This dataset helped facilitate real-world benchmark comparisons and enabled a high degree of precision in assessing the demands placed on healthcare facilities. Similarly, information on healthcare professionals by county, spanning the years from 2002 to 2023, was sourced from the same CHARTS database, which offered a comprehensive framework for evaluating the availability and distribution of specialized medical personnel needed for the treatment of the four diseases.

Leveraging these data resources assisted in deriving insights into Florida's healthcare system's readiness in managing the work-up (diagnoses) and treatments dependent to predict future demands, while measuring the state's capacity to address future needs.

For CRC, data on the numbers of colorectal nurses, anesthesiologists, and colorectal surgeons was procured; for COPD, counts of respiratory physical therapists and pulmonologists was acquired; and for mental health, the dataset encompassed of the number of psychiatrists, psychologists, licensed clinical social workers (LCSW), behavioral/mental health professionals, and licensed mental health counselors. This data was exclusively extracted from CHARTS [35].

For CHD, the US Health News was used to obtain the counts of thoracic surgeons and cardiologists by county within the state of Florida [36].

2.1.1 CRC Data Acquisition

For CRC, colonoscopy, sigmoidoscopy, and colorectal surgery data was obtained through the public dashboard on inpatient and outpatient data provided by the Florida Agency for Health Care Administration (AHCA) which provided information on facility treatment counts and patient demographics (i.e. age and sex) for the years 2018 and 2019 [37]. Colonoscopy and sigmoidoscopy data was obtained by isolating current procedural terminology (CPT) codes associated with both screening methods. For surgeries associated with CRC, Medicare Severity Diagnosis Related Groups (MS-DRG) counts were accumulated for specific procedures which were isolated by principal diagnoses of cancerous growth and respective treatment counts by group and county.

2.1.2 CHD Data Acquisition

For CHD, data was obtained on the official number of reports for diagnostics and treatments for CHD in the state of Florida by county from the Florida Healthcare and Administration Department (AHCA). This dataset covered the period from the third quarter of 2015 to the third quarter of 2023. MS-DRG codes were used to isolate the records related to PCI and CABG treatments.

2.1.3 COPD Data Acquisition

For COPD, the initial phase involved gathering COPD treatment data specifically for the years 2018 and 2019. This included information on bronchodilator, steroid therapy, and oxygen therapy for COPD patients, which was sourced from inpatient records using MS-DRG codes. Additionally, data on lung transplantation and bronchoscopy was collected using CPT codes from [37] for the same years. This data collection was conducted across counties, categorized by age groups and sex. Subsequently, a comprehensive dataset spanning from 2015 to 2023 was acquired [37], comprising records of over three million patients. This extended dataset was sourced from inpatient records and focused on extracting data based on primary diagnosis and other diagnosis columns, utilizing ICD-10-CM diagnosis codes. The data collection was conducted quarterly and categorized by age group, sex, and county.

2.1.4 Mental Illness Data Acquisition

For mental illness, data was obtained from the County Health Dashboard in the Florida Health CHARTS, which included hospitalizations from mental health disorders by age and hospitalizations from mood and depressive disorders, including schizophrenia, adjusted by age from 1994 to 2022 [38]. Data for hospitalizations from mental health disorder included the following age groups: 0-17, 18-21, 22-24, 25-44, 45-64, 65-74 and 75+ in all 67 counties from 1970 to 2022 and obtained from the Florida Health Charts Population estimate Query system [38]. Likewise, data of medical personnel: licensed clinical social workers (LCSW), psychologists, licensed mental health counselors and behavioral/mental health professionals was gathered from 2016-2017 to 2021-2022 from the County Health Dashboard - Health Resource Availability [39]. In addition, the count of psychiatrists was obtained from the Florida Department of Health - practitioner profile search [40].

2.2 Population Forecasting

Population forecasting used the data obtained from Florida Health CHARTS and was modeled using ARIMA, XGBoost, and Random Forest (RF) to produce a forecast spanning the 20 years past 2022, the latest year collected. ARIMA is a popular time-series forecasting technique that uses isolated past data (set by a lag parameter). Its transformations hold the mean and variance constant and error compensation to make predictions of future values based on historical data [37]. XGBoost is a Machine Learning (ML) tree-based forecasting technique that combines the results of multiple decision trees based on training data values, to perform forecasting. XGBoost also utilizes lag terms to isolate the amount of historical data used in prediction outputs [41]. RF is a classification and regression tool that utilizes combinations of decision trees based on a random subsection of the available training data [42]**.** These forecasting techniques were employed on each individual age group at a county level and trained on data from 1970 to 2017.

Testing was performed based on data from 2018 to 2022 to determine the best modeling technique to make predictions from 2023 to 2042. A set of metrics designed to comprehensively assess forecast accuracy and error rates were used in evaluating the performance of the models for population forecasting. Given the importance of accurate population forecasts for planning and resource allocation, metrics were employed that not only quantified the error in predictions but also provided insights into the models' predictive capabilities in a manner that was both interpretable and relevant to this study's objectives.

For errors, the Mean Absolute Percentage Error (MAPE) served as the primary metric as it offered an intuitive percentage-based measure of average error, which is particularly useful for communicating model forecast accuracy [43]. For instance, MAPE values allow for further understanding of the magnitude of prediction errors relative to actual population sizes and provide a clear, practical indication of the model's performance. In addition to MAPE, the Root Mean Squared Error (RMSE) was utilized to capture the variance in forecast errors; and, with its sensitivity to larger errors, it was a crucial measure for identifying models that might underestimate or overestimate significant population changes [44]. Furthermore, the Mean Absolute Error (MAE) was considered to evaluate the average magnitude of errors across the predictions, offering a straightforward metric in the same units as the population data [45]. Although less sensitive to outliers than RMSE, MAE provided an additional layer of insight into the models' consistency in prediction accuracy.

For the ARIMA model, parameter selection was automated through the 'auto_arima' function provided in the 'pmdarima' Python package, which iterates over multiple combinations of parameters to identify the most effective model configuration. By assessing potential models based on the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), 'auto_arima' automatically determined the order of differencing, the number of autoregressive terms, and the number of moving average terms [46]. Employing this automated approach ensured that the model was not only tailored to the idiosyncrasies of the dataset but also to prevent potential overfitting [47].

For the XGBoost model, lag variables were created capturing the population counts from the preceding one to three years for fine-tuning. Doing so enhanced the model's ability to effectively recognize and utilize historical trends [48]. Subsequently, all predictor variables were standardized to neutralize the scale differences amongst features. To refine the XGBoost model further, a randomized search for hyperparameter optimization utilizing 'RandomizedSearchCV' was implemented and a broad yet targeted search space for key parameters such as 'n_estimators', "max_depth", 'learning_rate', 'subsample', and 'colsample_bytree' was defined. RandomizedSearchCV efficiently explored this space by randomly sampling from predefined distributions and evaluating model configurations using 5-fold cross-validation [49]. This not only increased the efficiency of the search but also increased the model's generalizability by mitigating the risk of overfitting to the training data.

2.3 Demand Projections

2.3.1 CRC Demand Projections

Due to lack of historical data for CRC treatments, the 2018 and 2019 data was used to create a fixed baseline rate based on averages over total population for the same years. These rates were then projected onto the population forecasts to understand the increased need if rates were to be fixed based on available data. Similarly, average hospital length of stay (LOS) was calculated for CRC surgeries and the total LOS was calculated based on forecasted demand.

2.3.2 CHD Demand Projections

Originally, the same methodology was used as CRC for CHD but after obtaining additional data for CHD, a thorough exploratory data analysis was performed to identify trends and patterns for diagnostics and treatments. In addition to the regular analysis of the data to identify null values and imputations, trend analysis and clustering to groups of counties which had similar patterns was performed.

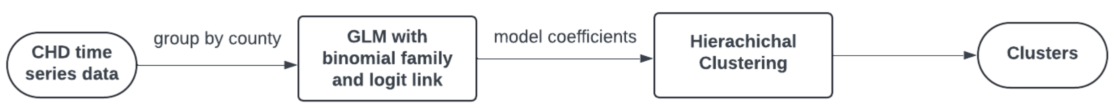


Figure 2. Grouping mechanism of the counties

Generalized linear models (GLM) for each county were fit and the model coefficients were vectorized. The model coefficient gave the direction of the fitted GLM line for the time series trend of each individual county. The resolution of the data points was the count of diagnostics and treatments per county by year from 2015 to 2023, however the scale of these counts was found to be drastically different. Therefore, before fitting the GLM, the counties were divided by their respective populations to obtain a standardized fraction of each year. Thereafter the years were encoded in a linear space between 0.1 and 0.99 with equal gaps. The GLM was then fit for the county fraction against the encoded year value and clustering was performed on the model coefficients. This was performed separately for both diagnostics and treatments. As 67 counties were analyzed with multiple procedures, the visual analysis itself was not robust and sufficient.

From the results shown in Figure 4, three groups of diagnostics trends were obtained. For treatment, all counties fell into one cluster. According to Figure 3, the trend for coronary artery bypass grafting (CABG) in Alachua county is different to that of Brevard county, however the PCI trends are similar. K-means was applied to see which coefficients prominently separated the counties.

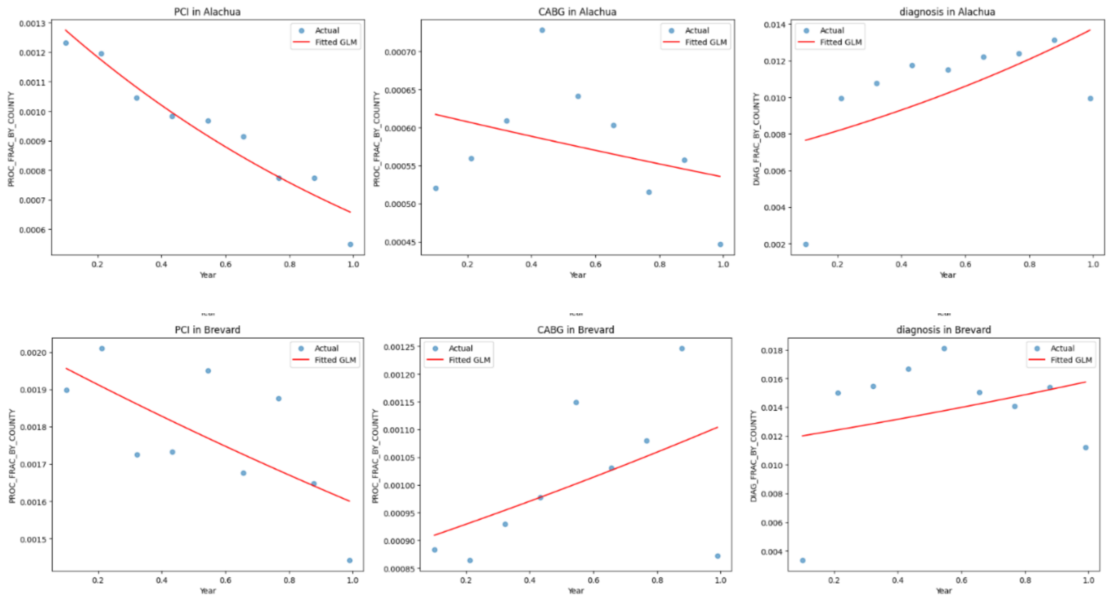


Figure 3. GLM fits

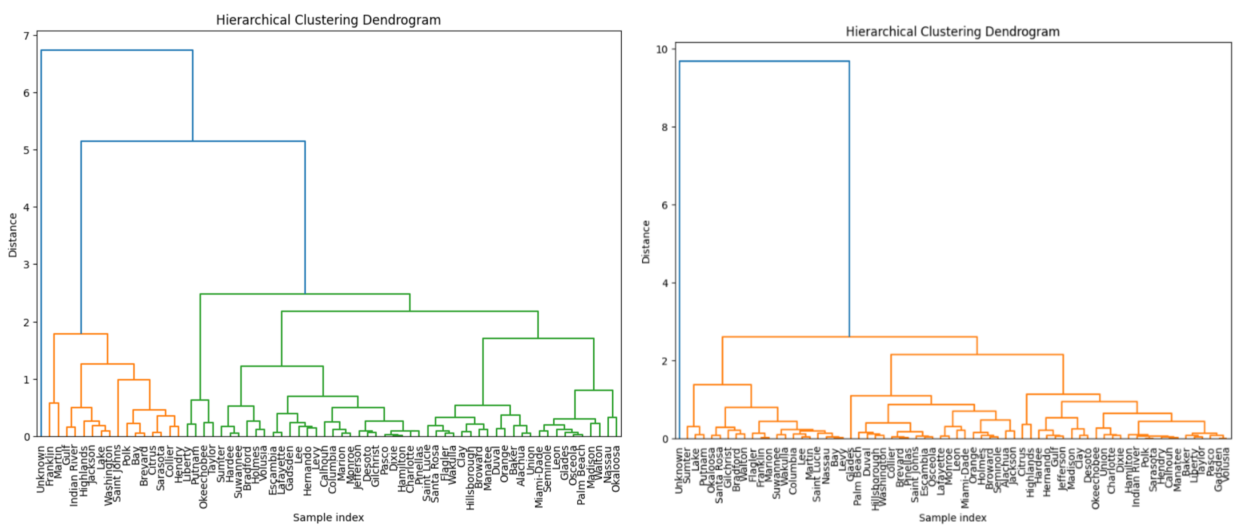


Figure 4. Dendrogram for CHD diagnostics (Left) & Dendrogram for CHD Treatment (Right)

Then, modeling for forecasting the diagnostics was performed with treatments for PCI and CABG done separately. The basic statistical model used was ARIMA, with the Augmented Dickey Fuller (ADF) Test performed to see if the time series were stationary. Based on the results, most of the time series were nonstationary, so the model needed to be fitted differently since ARIMA did not yield good results. Then, Seasonal ARIMA (SARIMA), an extension of the ARIMA model that includes seasonality by adding four additional seasonal parameters to the non-seasonal ARIMA parameters was employed [50]. This model showed more promising results for the majority of counties; however, a few counties were still not performing desirably. The Holt-Winters method, also known as the triple exponential smoothing algorithm, was used to address this. It extended the exponential smoothing to capture seasonality in addition to level and trend of the time series data. It was particularly well-suited for data with seasonal variations involving three types of smoothing layers [51].

2.3.3 COPD Demand Projections

Given the scarcity of historical treatment data, the initial approach to COPD analysis relied on information from 2018 and 2019. From this data, a constant baseline rate was established by calculating population-wide averages during those years. These rates were then projected onto population forecasts to gauge potential demand increases if rates were to remain static. Additionally, the average hospital length of stay (LOS) for COPD-related procedures was assessed and the total LOS was based on anticipated demand was extrapolated.

Following the initial analysis, additional data was received covering the period from 2015 to 2023. This dataset was structured into quarterly segments, focusing on various diagnosis codes pertinent to COPD. The quarterly diagnostic counts were utilized as inputs for the auto-arima model. This model, renowned for its automatic parameter selection, was configured with seasonality enabled parameters set at a frequency of four to accommodate the quarterly variations witnessed each year. By leveraging the auto-arima model, underlying patterns and trends were discerned within the COPD diagnostic data. Moreover, the projections extended over eight quarters, equivalent to a two-year horizon to anticipate future demand accurately.

Additionally, the exponential smoothing algorithm, a technique commonly used in time series forecasting, was employed. This algorithm assigned decreasing weights to past observations where more recent data points receiving higher weights, thus capturing short-term fluctuations while smoothing out noise and irregularities in the data [52]. Despite its intuitive simplicity, the exponential smoothing model exhibited higher Mean Absolute Percentage Error (MAPE) results compared to those obtained from the auto-arima model. Consequently, it was not pursued further for forecasting purposes.

2.3.4 Mental Health Demand Projections

Mental health disorder data was based on hospitalizations by age for mood and depressive disorders and schizophrenia over a period of 28 years. Forecasting of hospitalizations was done using five different ML algorithms that included: RF, XGBoost, long short-term memory (LSTM), light gradient boosting (Light GBM), and SARIMA. Light Gradient boosting was developed from gradient-boosting decision trees using a supervised machine-learning algorithm [53]. Decision tree regression integrated weak learners to achieve a more robust model through a serial training process [53]. Light GBM optimized memory usage and training time which improved it efficiency, scalability, and accuracy [53]. LSTM is an improved version of the Recurrent Neural Network (RNN) model. However, a traditional RNN finds it challenging to learn longer dependencies since it has a single hidden state, but an LSTM has a memory cell that can hold data longer to learn long-term dependencies [54]. For each county, the best model was chosen for forecasting hospitalizations based on the models' performance. Performance was calculated based on MAPE and MAE. Data between 1994 and 2016 was used for training, and data from 2017 to 2022 was used for testing. Then, for forecasting, the hospitalizations from the entire dataset from 1994 to 2022 were used. After forecasting the hospitalizations, the ratio of required medical professionals to population were calculated and applied to hospitalization counts in each county.

2.4 Sparse County Considerations

Throughout this project, an area of concern arose when compiling data for treatments and procedures from counties without hospitals capable of performing them. These counties had populations that had to receive these treatments and procedures in other neighboring Florida counties and leaving these populations out would have led to a large gap in the analysis of total resource usage. So, for counties that lacked such capabilities, the total populations by age and sex were added to the closest geographical counties. These exceptions are outlined in the Results Section for each individual illness when it occurred.

3. RESULTS

3.1 Population Forecasting Results

As previously mentioned, ARIMA, XGBoost, and RF were implemented in the forecasting of Florida county population by age and sex groups. The MAPE for all three models showed low error rates, going as low as 3.6% (Table 1). Compared to the tuned XGBoost and RF models, on average, the tuned ARIMA model outperformed both models in all categories. For adults ages 75 to 84 years, ARIMA did have the highest MAPE percentage, but the RMSE was lower for both the female and male groups. Figure 2 shows a comparison between the forecasted results from 2018 to 2022 and the true population counts. With these results and figures showing accurate trends, the study moved forward with ARIMA as the primary forecasting technique for population forecasting.

Table 1. MAPE, MAE, and RMSE for groups forecasted per ML Estimator Type. ARIMA = autoregressive integrated moving average; XGBoost= extreme gradient boosting; RF = random forest

	MAPE Results			MAE Results			RMSE Results		
Group	ARIMA	XGBoost	RF	ARIMA	XGBoost	RF	ARIMA	XGBoost	RF
Females:1 to 4	6.0%	8.0%	7.4%	238.46	301.01	295.18	517.36	675.51	656.04
Males:1 to 4	5.4%	7.9%	6.5%	233.77	357.15	296.65	571.74	956.40	672.76
Females:5 to 9	5.7%	7.5%	7.3%	291.53	336.42	366.02	622.29	645.27	768.40
Males:5 to 9	5.6%	7.3%	6.2%	316.07	364.87	322.04	671.18	745.55	655.32
Females:10 to 14	6.0%	6.0%	6.2%	421.79	468.76	398.88	823.17	1057.17	831.69
Males:10 to 14	5.6%	5.7%	5.9%	473.55	459.36	442.16	917.38	1048.79	911.89
Females:15 to 19	4.6%	5.9%	6.3%	261.98	327.00	293.85	506.20	771.02	568.73
Males:15 to 19	4.7%	6.3%	6.7%	266.15	379.86	343.51	556.51	836.99	675.59
Females:20 to 24	5.2%	7.8%	5.7%	344.77	317.89	415.68	711.16	716.94	1071.40
Males:20 to 24	7.1%	9.9%	8.5%	377.93	463.99	515.97	690.09	927.84	1171.32
Females:25 to 34	4.5%	6.9%	5.7%	878.46	1359.13	943.36	2539.98	4180.83	2522.90
Males:25 to 34	5.8%	8.5%	7.6%	856.43	1381.57	1098.04	2527.46	3778.48	2767.86
Females:35 to 44	6.2%	6.5%	6.0%	1012.84	784.70	879.20	1879.47	1581.49	1790.87
Males:35 to 44	8.3%	9.4%	7.6%	1213.07	1011.26	1002.98	2303.88	2168.57	2043.85
Females:45 to 54	3.6%	10.8%	8.6%	512.89	1378.13	1259.59	1692.11	3030.80	2589.70
Males:45 to 54	4.0%	12.3%	10.1%	525.82	1589.61	1360.34	1561.18	3939.26	2990.39
Females:55 to 64	3.8%	7.2%	6.8%	526.01	1622.55	1240.12	1091.83	5328.22	3904.12
Males:55 to 64	3.9%	7.5%	6.7%	447.69	1390.38	1209.07	811.60	4882.14	4067.67
Females:65 to 74	3.8%	6.4%	6.8%	743.60	1486.26	1330.14	1764.40	4412.98	3706.52
Males:65 to 74	4.4%	6.1%	6.2%	720.50	1208.36	1134.27	1678.50	3473.11	2834.02
Females:75 to 84	9.2%	7.8%	7.9%	1056.43	1102.79	994.87	2046.33	2685.91	2297.85
Males:75 to 84	9.0%	7.3%	6.6%	851.98	884.42	811.40	1559.38	2045.07	1834.12
Females:85+	8.1%	13.6%	13.7%	291.13	486.97	502.34	749.63	1138.83	1140.44
Males:85+	8.3%	11.1%	12.3%	185.06	254.18	300.97	379.59	471.32	618.28

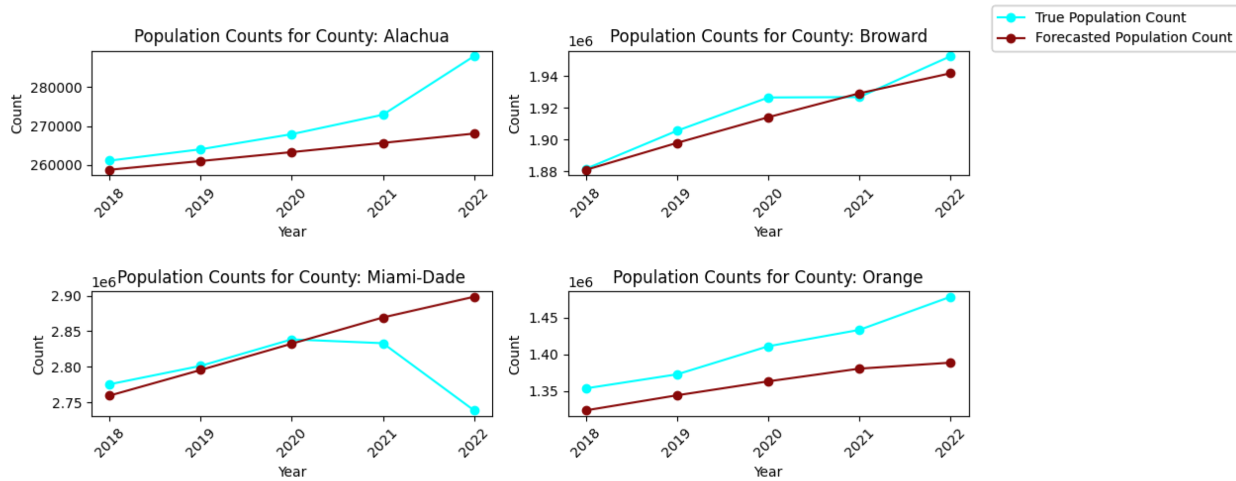


Figure 5. ARIMA Test Results for 4 Florida Counties (Alachua, Broward, Miami-Dade, Orange). Blue = True Population Count by County, Red = Forecasted Population Count by County

The tuned ARIMA model was used to create projections by county, age, and sex groups from 2023 to 2042, 20 years past the last recorded year in the dataset (2022). The same four counties as in Figure 2 can be shown projected out to 2042 in Figure 6. With these values in place, individual disease projections were performed to understand the impact of population change on resource demand.

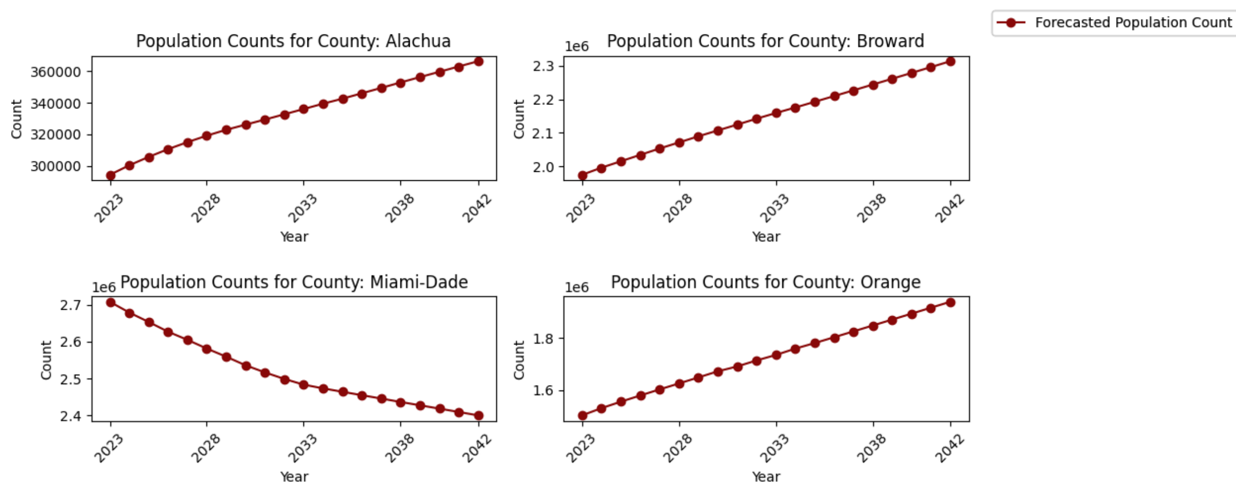


Figure 6. ARIMA Projections for 4 Florida Counties (Alachua, Broward, Miami-Dade, Orange)

3.2 Disease Results

3.2.1 CRC Results

The following counties were without reporting hospitals with capabilities to perform colonoscopies, sigmoidoscopies, and CRC surgery were aggregated, as shown in Table 2.

Table 2. Sparse Counties for CRC and their closest geographical neighbor.

** Hospital exists with capability but data not present		
Bordering County	Sparse County for Colonoscopy/Sigmoidoscopy	Sparse County for Colorectal Surgery
Alachua	Dixie, Gilchrist, Hamilton, Lafayette, Levy, Suwanee, Taylor, Union	Bradford, Dixie, Gilchrist, Hamilton, Lafayette, Levy, Suwanee, Taylor, Union
Bay	Calhoun, Holmes	Calhoun, Gulf, Holmes, Jackson, Washington
Clay	Putnam	Putnam
Duval	None	Baker
Highlands	Hardee	Hardee
Lee	Glades	Glades, Hendry
Leon	Franklin, Gadsden, Jefferson, Liberty, Madison, Wakulla	Franklin, Gadsden, Jefferson, Liberty, Madison, Wakulla
Sarasota	Desoto**	Desoto

When projecting demand for both screenings and surgery, the highest increase in demand is found in the 65–74 year age group in the ARIMA projections. For example, counties like Saint Johns had demand for colonoscopy and sigmoidoscopy increase by as much as 127% by 2042 compared to 2023 and an increase of 129 % for CRC surgery in the 65–74 year age group. In fact, in the same age group, Alachua, Broward, and Orange County are projected to have increased demand for both screenings and surgery by over 50% by 2042, with Orange County forecasted to have increased demand by over 93% in for the screenings and surgeries. Areas of concern were also present for ages 55-64 years, with projected increases of over 65% across all procedures in Orange County and 91% in Osceola County. Then, for ages 75-84 years, screening demand growth was projected as high as 115% in Saint Johns with increased surgery demand growth at over 123%. Several counties of interest can be seen in Figure 7 and Figure 8 with additional visualizations and information on individual counties found in the interactive dashboard.

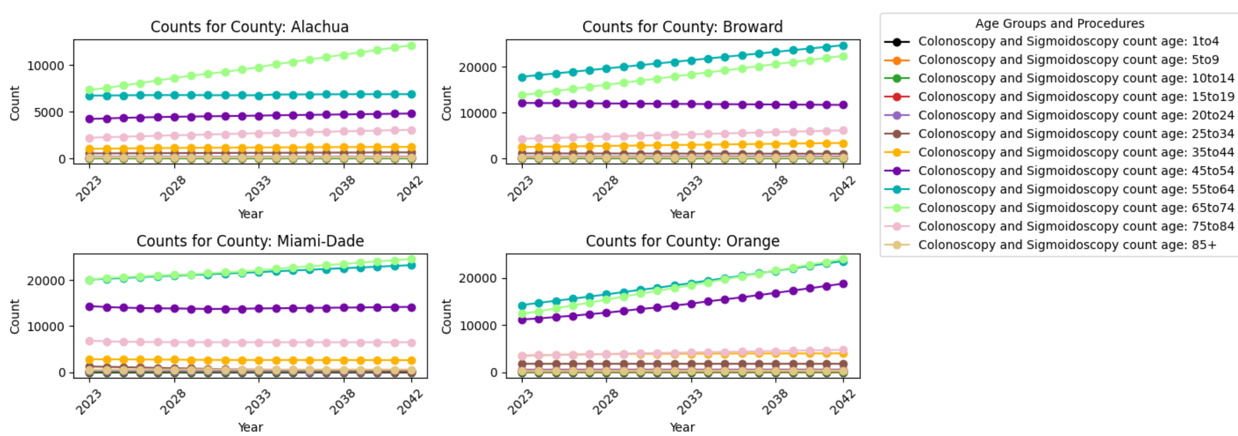


Figure 7. Colonoscopy/Sigmoidoscopy Projections for 4 Florida Counties (Alachua, Broward, Miami-Dade, Orange). Key on right provides color reference (also in supplement S3)

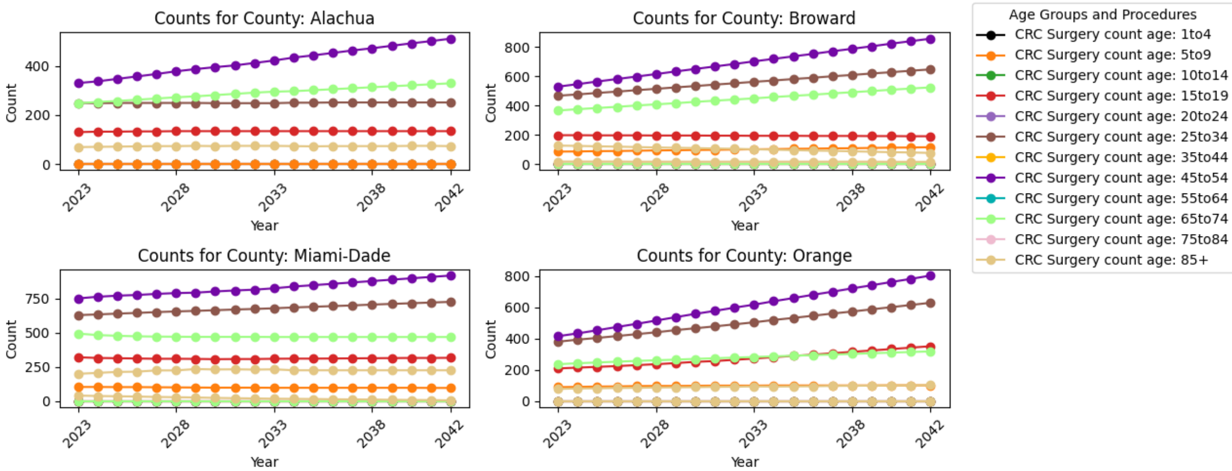


Figure 8. CRC Surgery Projections for 4 Florida Counties (Alachua, Broward, Miami-Dade, Orange). Key on right provides color reference (also in supplement S3)

The ideal personnel required to assist in colonoscopy and sigmoidoscopy procedures are two registered nurses (RNs) and an anesthesiologist, outside of the clinician specialist [56]. Using the data for current medical professionals available as described in section 2.1, areas of potential concern begin to appear for certain counties. With the ideal staff list, it is forecasted that by 2042, each RN will have to be involved in at least ten colonoscopies or sigmoidoscopies in Alachua, Leon, and Sarasota counties per year. With anesthesiologists, counties like Saint Johns would have a ratio upwards of 4000:1 by 2042. In addition, Clay, Highlands, Leon, Osceola, and Volusia Counties would all pass a ratio of 2000:1 in the same timeframe. For CRC surgeries, there also exists a potential for unsustainable ratios of surgeons. Osceola County is projected to have a ratio of 102:1 by 2042 and Leon and Saint Lucie are projected to pass a ratio of 50:1 in the same timeframe. More detailed information can be found in Figures 6, 7, and 8 and in the interactive dashboard.

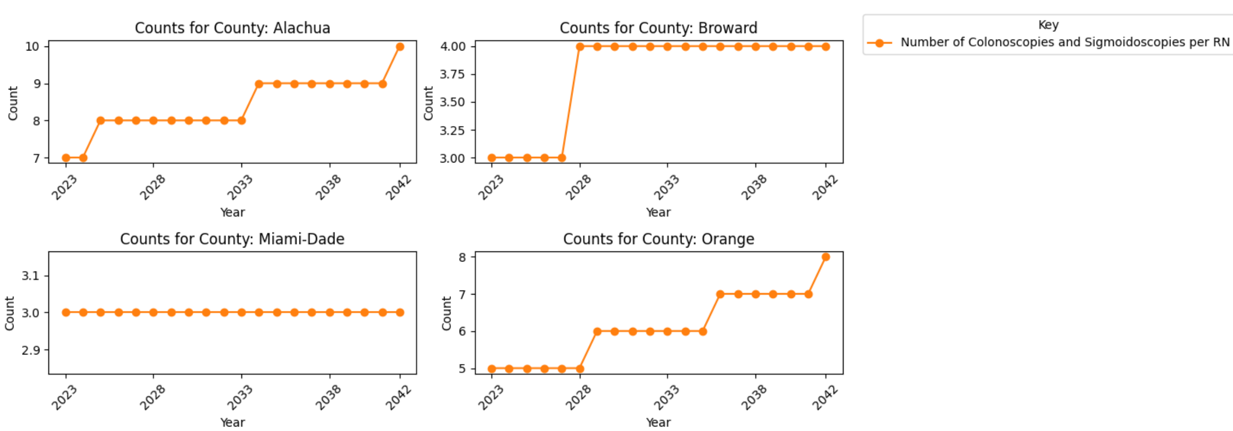


Figure 9. Number of Colonoscopy and Sigmoidoscopy Projections per RN for 4 Florida Counties (Alachua, Broward, Miami-Dade, Orange).

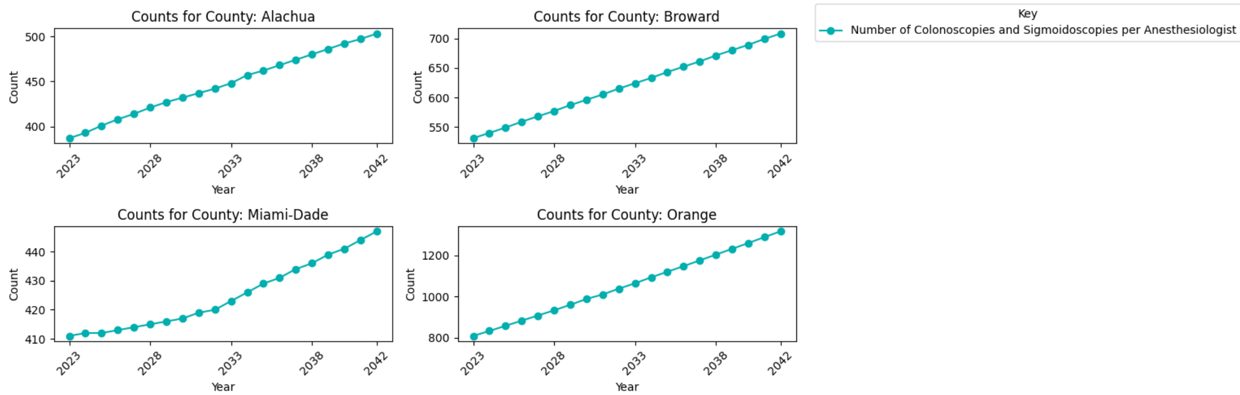


Figure 10. Number of Colonoscopy and Sigmoidoscopy Projections per Anesthesiologist for 4 Florida Counties (Alachua, Broward, Miami-Dade, Orange).

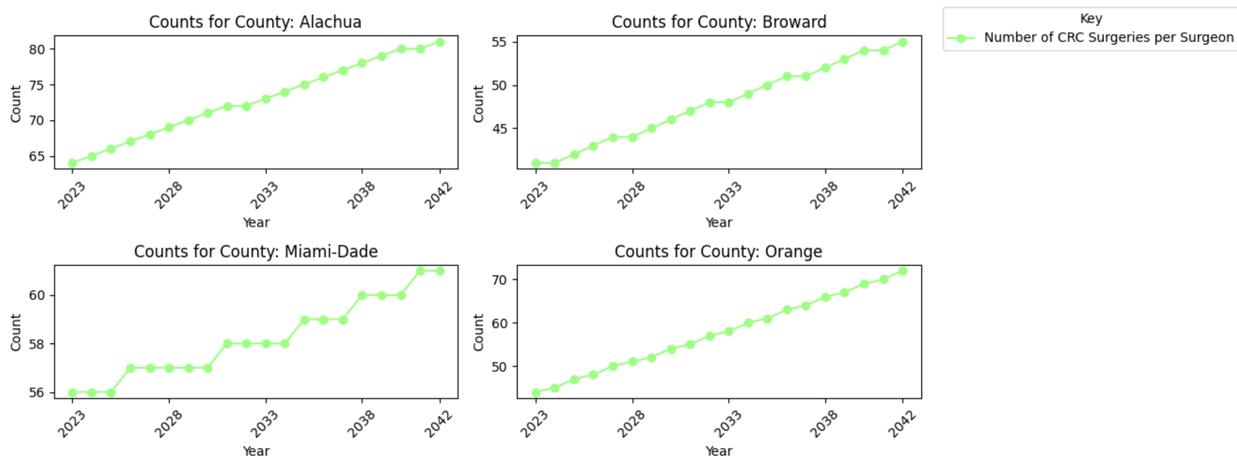


Figure 11. Number of CRC Surgery Projections per Surgeon for 4 Florida Counties (Alachua, Broward, Miami-Dade, Orange).

3.3.2 CHD Results

Two main datasets were taken into consideration, the outpatient emergency department data and the inpatient data. The first objective was to build statistical models for each county to forecast the expected number of patients diagnosed with CHD. Depending on the trend and seasonality of the time series in each county, either of the two models SARIMA and Exponential Smoothing fit best. Overall, SARIMA produced an average MAPE of 24.42% compared to Exponential Smoothing which produced 17.87%. Most of the time series data was nonstationary and a significant amount of data had white noise. Due to this, a damping in forecasting projections was observed after the year 2033. Thus, the model was trained with nine years of data (with quarters as the resolution for data points) and predicting 42 steps ahead. Forty-two quarters across nine years were aggregated to encapsulate the individual years.

Then, statistical models were built to forecast the PCI and CABG treatments using patient counts that living in any given county would expect to undergo. On average, SARIMA forecasting PCI resulted in an MAPE of 24.67% and using Exponential Smoothing had an MAPE of 18.73%. For CABG, SARIMA had an average MAPE of 25.44% and Exponential Smoothing had an average MAPE of 18.60%.

Before moving into the modeling of diseases, a thorough analysis of the data was performed to determine the percentage of patients from each county traveling to another county's facility for treatment (S3, S4) This allowed for better understanding of the rates at which a county's facility would need to expect patients from its neighboring counties. For example, in Bradford County, most of the patients needing PCI treatments traveled to Alachua (33.65%), or to Duval (27.88%) or Clay (17.31%) counties as Bradford does not have cardiac catheterization (Cath) lab facilities needed to conduct PCI [58]. Patients traveled to larger neighboring counties that had Cath lab facilities. In addition, Alachua and Duval appeared to be two of the most prominent counties providing PCI treatments. Alachua is the prominent hub for PCI, with geographically distant patients travelling to get treated there. However, looking at patients from Alachua, although most patients are getting treated in their home county, a number of patients tended to go to the neighboring Marion and Duval counties as well. This could be a reason for the large influx of patients to Alachua from other counties creating a shortage of infrastructure, doctors, and long wait times to conduct PCI. This is an indication that particularly in Alachua, the Cath Lab infrastructure required to conduct PCI may need to be expanded.

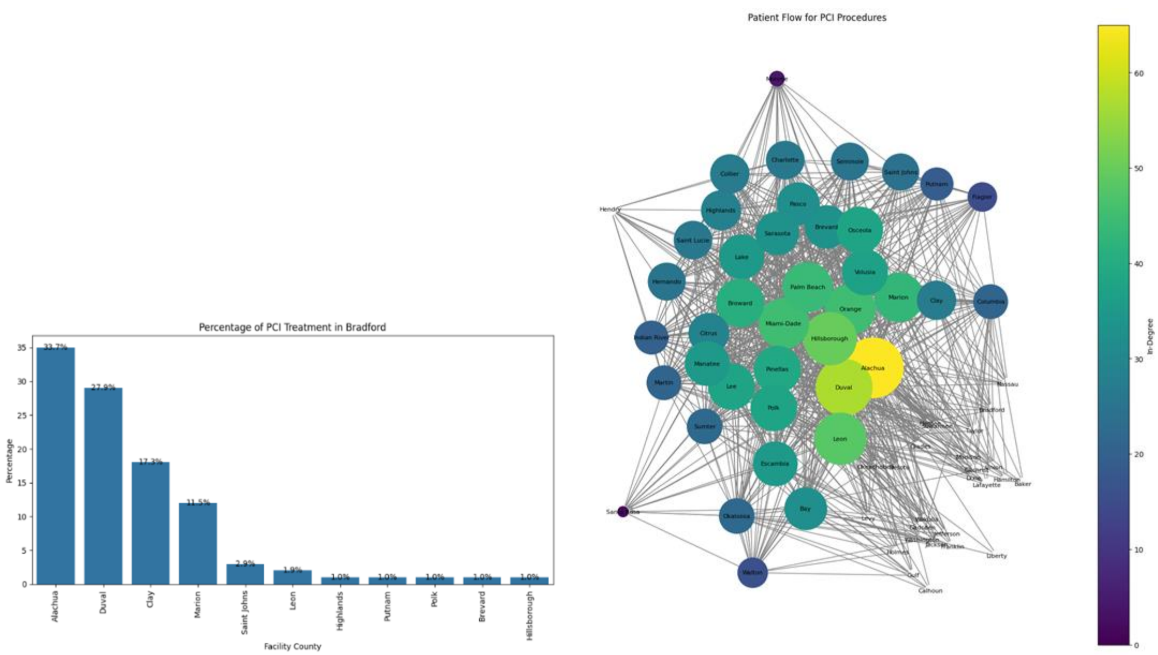


Figure 12. The percentage of PCI treatment cases from Bradford carried out in neighboring counties with facilities on left and PCI patients flow for Cath Lab facility counties to get treated on right.

According to Figure 13, Alachua is one of the counties that is rapidly growing in population at a fast rate. This population growth rate is around 5.5% annually and complements the staggering increase in CHD cases expected to be diagnosed in the upcoming years. Looking at the treatments, PCI count lowers across time while CABG count is rising gradually. The most effective treatment for CHD is CABG [59] and there is a high chance a patient who undergoes PCI will eventually need to be subjected to CABG treatment. This rising number of CABG treatments can be the reason for patients who previously underwent PCI having to undergo CABG in addition to new patients needing CABG in any particular year and an indication that patients in Alachua are prone to severe cases of CHD.

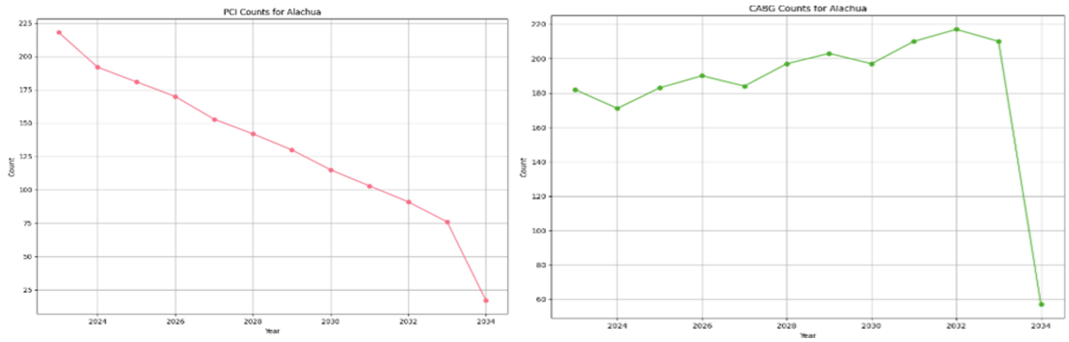


Figure 13. Diagnosis, PCI & CABG treatments for Alachua from 2023 to 2034

According to Figure 14, Citrus County tells a slightly different story. The diagnosis of patients appears to be consistent from the years 2024 to 2033 even though both PCI and CABG cases rise in number. This indicates that this county may potentially have even more CHD patients in critical condition than Alachua. This could be lifestyle factors could be one of the major factors influencing the number of patients in Citrus County in becoming severe CHD patients.

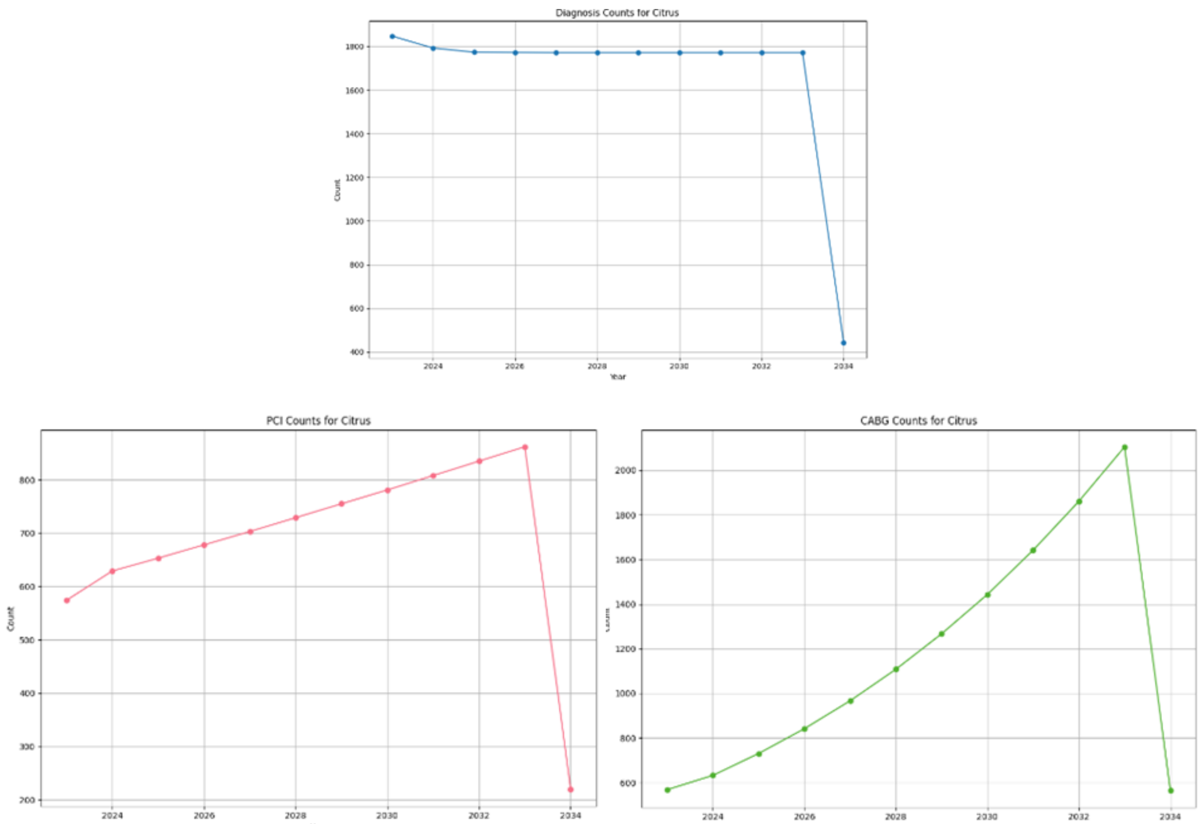


Figure 14. Diagnosis, PCI & CABG treatments for Citrus County from 2023 to 2034

The annual patient volume for PCI for interventional cardiologists is recommended to be 1:59 [60]. However, from the data gathered during the time span from 2015 to 2023, the operator to patient ratio is lower than the national average. The highest observed in 2018 in Florida was a 1:23 interventional cardiologists to patient ratio, still better than the national average. Though the recommended ratio is 1:59, the doctor’s efficacy, experience level, and attrition are factors that could impact performance. Results from the data show a 1:23 ratio would be an effective ratio to identify the number of doctors required to treat the growing PCI need.

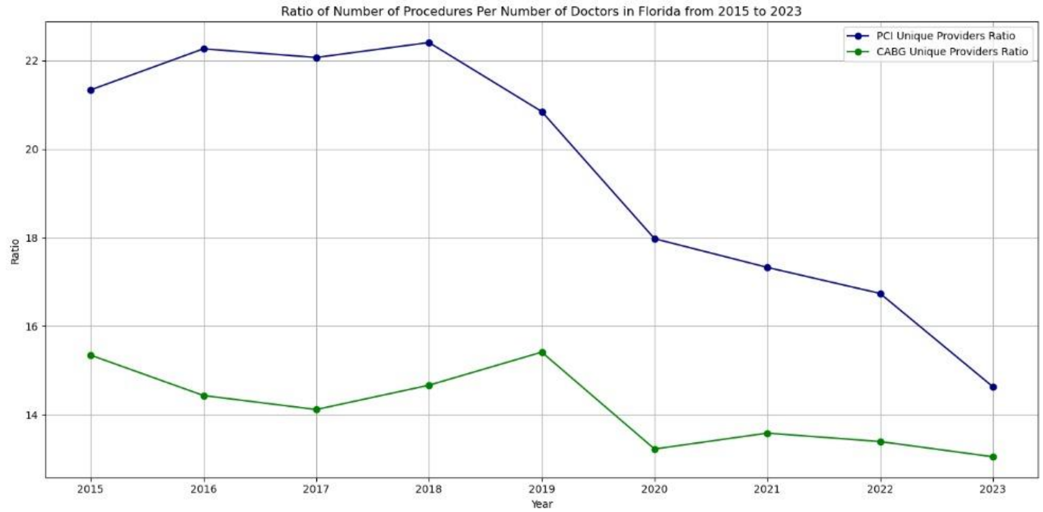
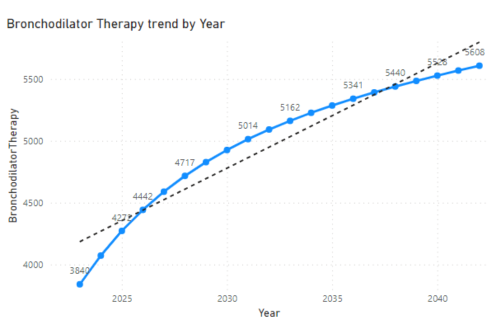
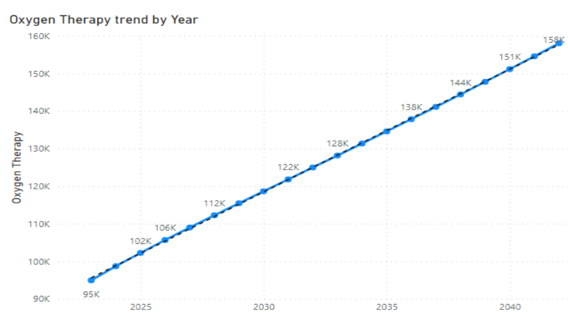


Figure 15. PCI & CABG procedures per doctor ratio from 2015 to 2023

3.2.3 COPD Results

Upon analyzing the data, discernible linear trends emerge for each treatment modality based on projected population demographics. The utilization of medical interventions displays a linear progression across all age groups with notable variations in treatment preferences. For instance, bronchoscopies demonstrate an upward linear trend as age advances, peaking in the age group of 65 to 74 years before tapering off in the oldest demographic. Similarly, bronchodilator therapy exhibits a steady increase with age, reaching its apex in the 55 to 64 age group before plateauing. Conversely, lung transplantation rates rise steadily with age until the 65 to 74 age group, where they reach their zenith before experiencing a decline in the oldest age group. Oxygen therapy also portrays a consistent linear trend, with usage escalating as age increases, indicating a greater reliance on this treatment modality among older populations. These linear trends underscore the evolving treatment needs across different age groups reflecting the dynamic interplay between age demographics and COPD treatment preferences.





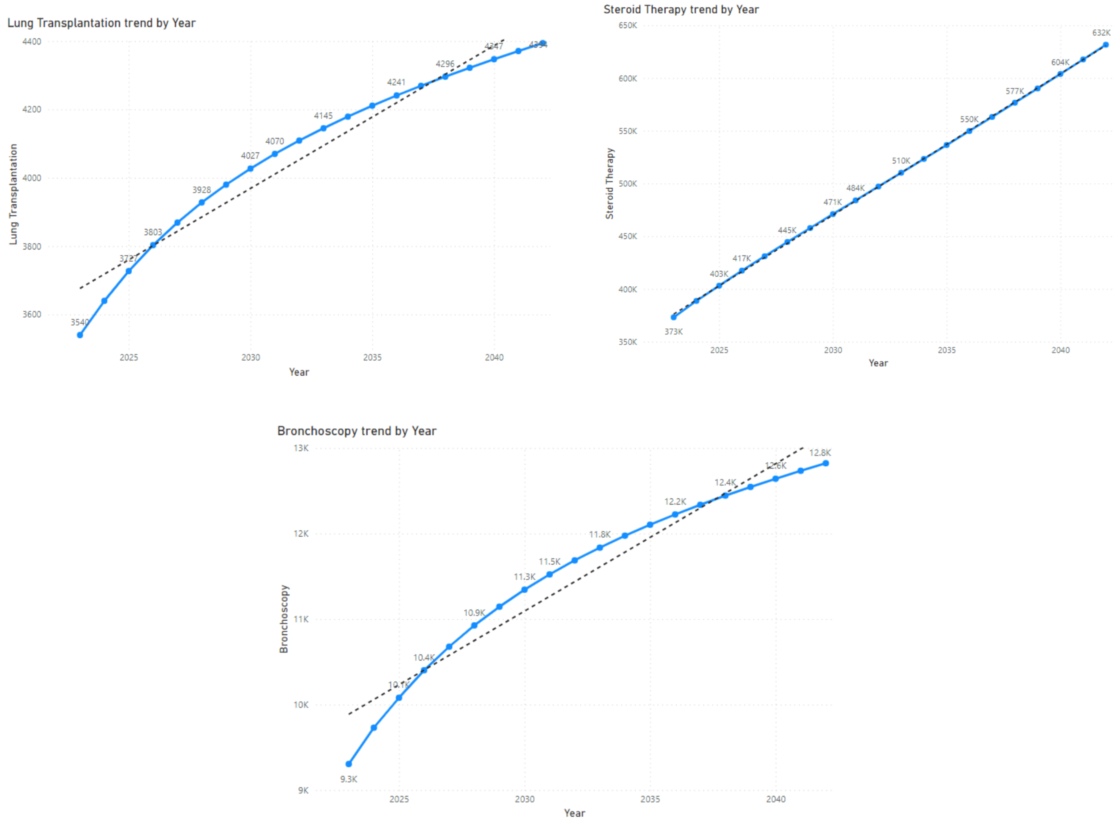


Figure 16. Illustrates a linear trend within the 35 to 44 age group, showcasing an upward trajectory for various treatments, indicative of an increasing trend.

With the more robust data, the analysis of MAPE results for each county using the auto-arima model revealed that out of 68 counties, 51 had a diagnosis count exceeding 40. Counties considered outliers were excluded from this analysis. Among the remaining 51 counties, 44 exhibited MAPE values ranging between 2% to 20%. Seminole County demonstrated the lowest error percentage, while Putnam County had the highest at 20%. Additionally, there were four counties with MAPE values ranging from 21% to 28%. Notably, the three counties of Taylor, Washington, and Calhoun displayed higher error rates and as such were omitted from further analysis. The charts in the supplement (S5) and Figure 17 illustrate the forecasting results generated by the auto-arima model, extending until the year 2025.

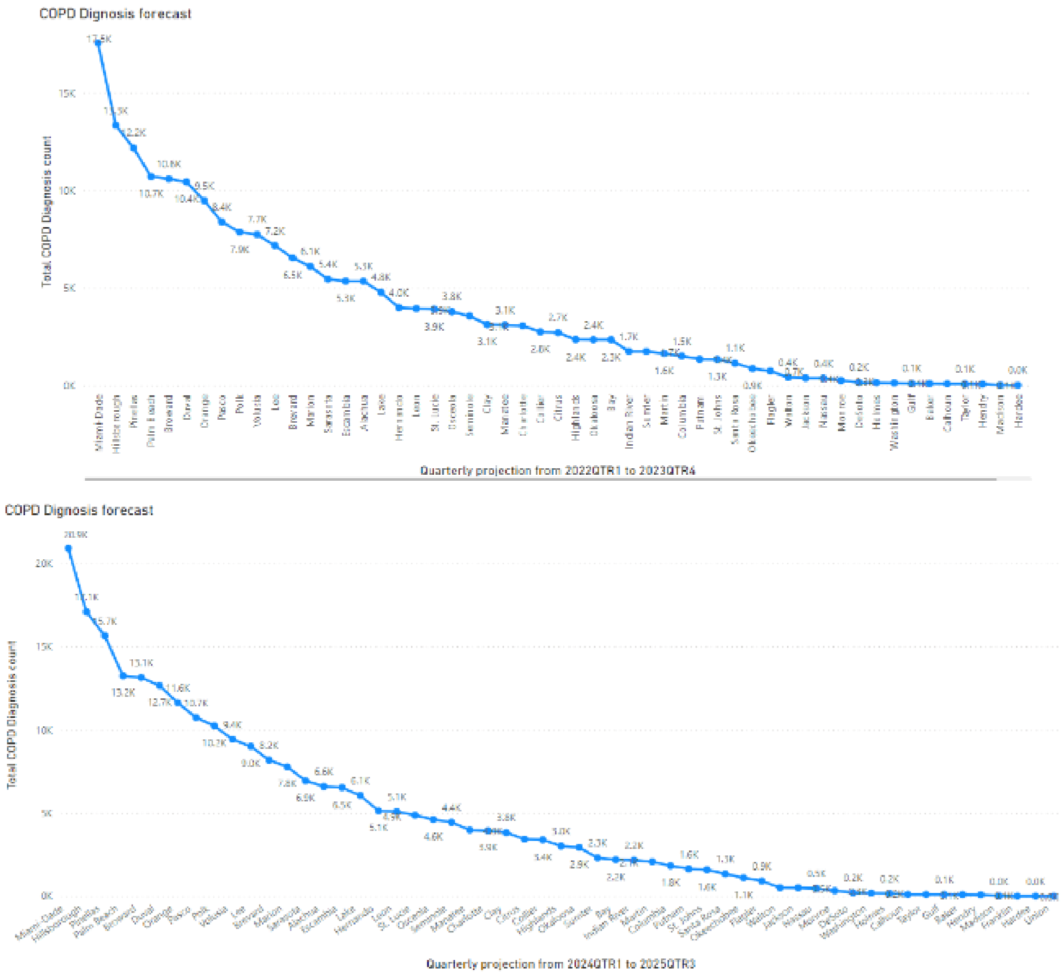


Figure 17. Counties exhibit an overall increasing trend in diagnosis count.

A substantial proportion, approximately 43%, of Florida's counties were analyzed to have shortages in crucial healthcare personnel, including nutritionists, MDs, physicians, and advanced medical nurses. The study's findings expose stark shortages of nutritionists in 43 counties, alongside concerning inadequacies of medical doctors, advanced medical nurses, and certified nurses in numerous areas. Of particular concern are Franklin, Glades, and Liberty counties, where shortages span all four categories of healthcare professionals. These findings emphasize the urgent need to bolster healthcare resources in these underserved regions, ensuring residents receive the necessary medical care.

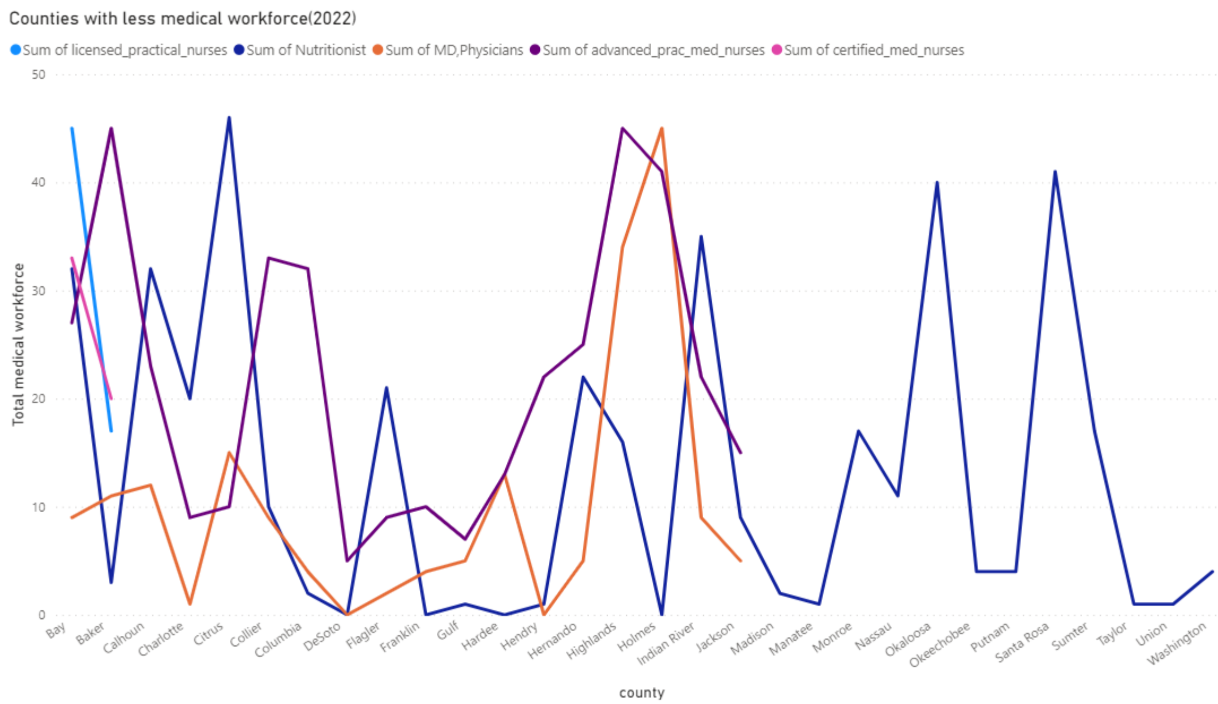


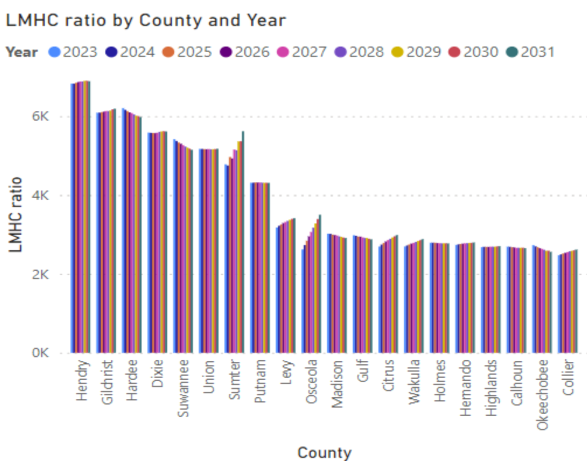
Figure 18. Chart of Medical Workforce by County

3.2.4 Mental Illness Results

As previously mentioned, mental health disorder data was based on hospitalizations by age for mood and depressive disorders and schizophrenia over a period of 28 years. Five different models were used for forecasting hospitalization count, as mentioned in 2.3.4. The MAPE for each age group can be seen in Table 3. These MAPE results were relatively low enough to continue analyzing the results.

Table 3. MAPE for Mental Health Hospitalization Forecasting	
Age Group	MAPE
0-17	25%
18-21	20%
22-24	19%
25-44	14%
45-64	14%
65-74	20%
75+	23%

Based on these projections, most hospitalizations from mental disorders are in the age groups 25-44 years and 45-64 years. Miami-Dade and Broward were projected to have the highest hospitalization counts for these age groups as well as for schizophrenia and mood and depressive disorders. Based on projections and current data on professionals, the professionals to population ratios in counties like Calhoun, Bradford, and Lake is zero which means that psychiatrists and marriage and family therapists are not currently available in those counties, although psychologists were present in low numbers. In Glades, psychologists, psychiatrists, and mental counselors are also not present. Population to mental health personnel is recommended to be in the ratio of 30000:1 and if the demand is high then the suggested ratio should be 20000:1 [57]. Additionally, the psychiatrists to population ratio is recommended to be around 10000:1 [61]. Based on the projections, Licensed Mental Health Counselors (LMHC) and Licensed Clinical Social Workers (LCSW) had acceptable proportions for almost all counties in Florida (Figure 19). For Licensed Marriage and Family Therapists (LMFT), Jackson, Charlotte and Wakulla Counties are already over the recommended ratio, 30000:1, whereas for the counties Polk, Citrus and Sumter, the ratio is currently around 26000:1. By the year 2031 the ratio will reach 30000:1 with the increase in county population if the same count of LMFT is maintained as in 2023 for the next nine years as shown in S7. For psychologists, the ratio for the counties Levy, Suwannee and Hendry already surpassed the desired ratio in the year 2023 and over the next nine years this upward trend would continue in the county Levy (Figure 20). In the case of psychiatrists to population ratio, the counties in Figure 28 will surpass the desired ratio of 10000:1 if the Psychiatrists count in 2023 is maintained same for the next nine years and those areas will demand in hiring of the personnels.



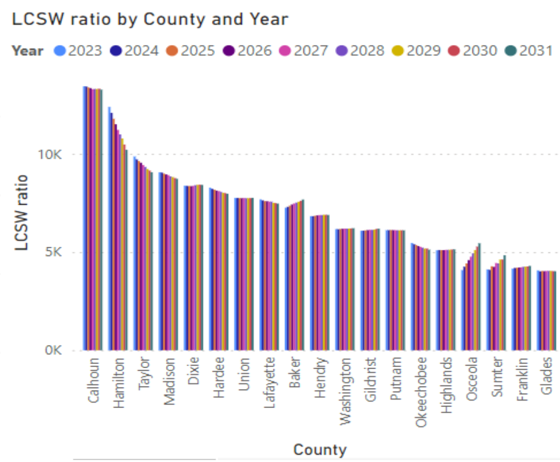


Figure 19. LMHC and LCSW ratios to the forecasted population

Figure 20. LMFT to population ratio and psychologists to population ratio

4. DISCUSSION

Based on the low error in population forecasting (MAPE 3.6% to 9.2%), the projection showed a significant increase in population size for the state of Florida from the years 2023 to 2042. Certain age groups, 45-74 years, had a larger growth rate than their younger counterparts, suggesting an increasingly aging population as years progress. With an ever-increasing Florida population, results also showed an increase in demand for procedures under each of the four diseases studied.

For CRC, procedure demand was projected to grow in several age groups, primarily 45 to 74 years. Specific counties like Saint Johns disproportionately had increased demand projections of over 100% by 2042 in several age groups. In turn, ratios of procedures to professionals could reach as high as 4000:1 for colonoscopies and sigmoidoscopies to anesthesiologists. Smaller counties with increased projected demand like Saint Johns and Osceola (over 90% demand growth by 2042 for ages 55-64), should be reviewed more closely to determine whether these counties can sustainably treat patients in the future.

For CHD, demand growth projections can be observed in several counties for treatments like CABG, including Alachua and Citrus. Fortunately, Florida tends to have ratios of procedures to doctors that are lower than the national threshold. It is important to note that for CHD, many professionals live in the same larger counties, which causes sparse county residents to potentially travel farther for treatment. For counties like Citrus, while diagnoses may be projected to be stagnant, CABG counts are projected to increase to over 2000 by 2042. Counties like Citrus should be cognizant of the potential for increased procedure demand. Moreover, larger counties like Alachua and Duval, should not only be aware of their own residents' demand but the demand of the state, due to the scarcity of specialized institutions, like Cath labs, needed to perform many key procedures.

For COPD, 43% of Florida counties were found to have potential shortages in healthcare personnel. In counties like Franklin, Glades, and Liberty, shortages exist for all four categories of healthcare professionals: nutritionists, MDs, physicians, and advanced medical nurses. For mental illness, growth is also projected with potential gaps in professionals equipped to handle the demand. In counties like Jackson, Charlotte and Wakulla, ratios of population to personnel exceed the recommended 30000:1. This underscores a potential insufficiency in the healthcare workforce, particularly in rural and underserved communities. Addressing these shortages may be vital to ensure equitable access to healthcare services and to improve health outcomes for Florida residents.

Overall, the projections show an increase in overall demand for procedures associated with each disease and potential gaps in professionals equipped to handle the demand. This increased demand should be further investigated to quantify the total gap between the current state of Florida healthcare and the future needs of the state's population.

5. CONCLUSION & FUTURE WORK

Based on the findings illustrated in this report, the authors see a need for further evaluation of the healthcare infrastructure, especially in several counties in Florida. Demand for treatments for all diseases in this report was projected to rise based on forecasts. Simultaneously, several potential areas of concern can be identified in ratios between procedures and the staff necessary to complete them. Counties with potentially unsustainable ratios may need to invest in human capital in the healthcare industry to address increasing demands.

The authors suggest future work to expand the analysis in future healthcare needs and preparedness in Florida. Data availability stemmed as a major issue in many facets of this study. With additional time, more detailed historical data on professionals and procedures could be procured from private and public organizations which would assist in more accurate and robust forecasting and analysis.

Future work to create an agent-based model that could simulate the impact of unexpected crises such as COVID-19 could be useful. Furthermore, an evolution of this model to a digital healthcare twin could be more revealing.

ACKNOWLEDGEMENTS

The assistance of Dr. Edwin Nassiff (UCF), Dr. Daniel Eilen (UCF), and KPMG is acknowledged in sponsoring and assisting with this project.

REFERENCES

[1] K. Christensen, G. Doblhammer, R. Rau, and J. W. Vaupel, "Ageing populations: the challenges ahead," *Lancet*, vol. 374, no. 9696, pp. 1196–1208, Oct. 2009, doi: 10.1016/S0140-6736(09)61460-4.

[2] D. Carr, G. Falchetta, and I. Sue Wing, "Population Aging and Heat Exposure in the 21st Century: Which U.S. Regions Are at Greatest Risk and Why?," *The Gerontologist*, vol. 64, no. 3, p. gnad050, Mar. 2024, doi: 10.1093/geront/gnad050.

[3] U. C. Bureau, "America Is Getting Older," Census.gov. Accessed: Apr. 03, 2024. [Online]. Available: https://www.census.gov/newsroom/press-releases/2023/population-estimates-characteristics.html

[4] C. J. van Oostveen, D. J. Gouma, P. J. Bakker, and D. T. Ubbink, "Quantifying the demand for hospital care services: a time and motion study," *BMC Health Serv. Res.*, vol. 15, p. 15, Jan. 2015, doi: 10.1186/s12913-014-0674-2.

[5] E. A. Biondi, M. Hall, M. S. Leonard, P. A. Pirraglia, and B. K. Alverson, "Association between resource utilization and patient satisfaction at a tertiary care medical center," *J. Hosp. Med.*, vol. 11, no. 11, pp. 785–791, 2016, doi: 10.1002/jhm.2621.

[6] T. Bhowmik and N. Eluru, “A comprehensive county level model to identify factors affecting hospital capacity and predict future hospital demand,” *Sci. Rep.*, vol. 11, no. 1, p. 23098, Nov. 2021, doi: 10.1038/s41598-021-02376-y.

[7] WHO, “Cancer,” World Health Organization, Feb. 03, 2022. https://www.who.int/news-room/fact-sheets/detail/cancer

[8] T. Sawicki, M. Ruszkowska, A. Danielewicz, E. Niedźwiedzka, T. Arłukowicz, and K. E. Przybyłowicz, “A Review of Colorectal Cancer in Terms of Epidemiology, Risk Factors, Development, Symptoms and Diagnosis,” *Cancers*, vol. 13, no. 9, Art. no. 9, Jan. 2021, doi: 10.3390/cancers13092025.

[9] M. Millan, S. Merino, A. Caro, F. Feliu, J. Escuder, and T. Francesch, “Treatment of colorectal cancer in the elderly,” *World J. Gastrointest. Oncol.*, vol. 7, no. 10, pp. 204–220, Oct. 2015, doi: 10.4251/wjgo.v7.i10.204.

[10] Md. S. Hossain *et al.*, “Colorectal Cancer: A Review of Carcinogenesis, Global Epidemiology, Current Challenges, Risk Factors, Preventive and Treatment Strategies,” *Cancers*, vol. 14, no. 7, p. 1732, Mar. 2022, doi: 10.3390/cancers14071732.

[11] B. K. Edwards *et al.*, “Annual Report to the Nation on the Status of Cancer, 1975–2006, Featuring Colorectal Trends and Impact of Interventions (Risk Factors, Screening, and Treatment) to Reduce Future Rates,” *Cancer*, vol. 116, no. 3, pp. 544–573, Feb. 2010, doi: 10.1002/cncr.24760.

[12] “Colorectal Cancer | Florida Department of Health,” www.floridahealth.gov. https://www.floridahealth.gov/diseases-and-conditions/cancer/colon-cancer/index.html

[13] National Cancer Institute, “Tests to Detect Colorectal Cancer and Polyps,” National Cancer Institute, 2016. https://www.cancer.gov/types/colorectal/screening-fact-sheet

[14] V. K. Morris *et al.*, “Treatment of Metastatic Colorectal Cancer: ASCO Guideline,” *J. Clin. Oncol.*, vol. 41, no. 3, pp. 678–700, Jan. 2023, doi: 10.1200/JCO.22.01690.

[15] American Heart Association, “Coronary Artery Disease - Coronary Heart Disease,” www.heart.org, Jul. 31, 2015. https://www.heart.org/en/health-topics/consumer-healthcare/what-is-cardiovascular-disease/coronary-artery-disease

[16] World Health Organization, “Cardiovascular Diseases,” World Health Organization, 2023. https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

[17] A. Kr. Malakar, D. Choudhury, B. Halder, P. Paul, A. Uddin, and S. Chakraborty, “A Review on Coronary Artery disease, Its Risk factors, and Therapeutics,” Journal of Cellular Physiology, vol. 234, no. 10, pp. 16812–16823, Feb. 2019, doi: https://doi.org/10.1002/jcp.28350.

[18] “Deaths from Coronary Heart Disease - Florida Health CHARTS - Florida Department of Health,” www.flhealthcharts.gov. https://www.flhealthcharts.gov/ChartsReports/rdPage.aspx?rdReport=Death.DataViewer&cid=0099

[19] E. L. Alderman et al., “Native coronary disease progression exceeds failed revascularization as cause of angina after five years in the bypass angioplasty revascularization investigation (BARI),” Journal of the American College of Cardiology, vol. 44, no. 4, pp. 766–774, Aug. 2004, doi: https://doi.org/10.1016/j.jacc.2004.05.041.

[20] M. Ahmad, P. Mehta, A. K. R. Reddivari, and S. Mungee, “Percutaneous coronary intervention,” PubMed, 2020. https://www.ncbi.nlm.nih.gov/books/NBK556123/

[21] V. Farooq et al., “The CABG SYNTAX Score - an angiographic tool to grade the complexity of coronary disease following coronary artery bypass graft surgery: from the SYNTAX Left Main Angiographic (SYNTAX-LE MANS) substudy,” EuroIntervention, vol. 8, no. 11, pp. 1277–1285, Mar. 2013, doi: https://doi.org/10.4244/eijv8i11a196.

[22] National Heart, Lung, and Blood Institute, “COPD - what is COPD?,” www.nhlbi.nih.gov, Oct. 25, 2022. https://www.nhlbi.nih.gov/health/copd

[23] Y. S. Jo, “Long-term outcome of chronic obstructive pulmonary disease: A review,” Tuberculosis and Respiratory Diseases, vol. 85, no. 4, Jul. 2022, doi: https://doi.org/10.4046/trd.2022.0074.

[24] American Lung Association, “COPD Causes and Risk Factors ,” American Lung Association, Jan. 24, 2024. https://www.lung.org/lung-health-diseases/lung-disease-lookup/copd/what-causes-copd

[25] American Lung Association, “Diagnosing and treating COPD,” www.lung.org, Mar. 05, 2021. https://www.lung.org/lung-health-diseases/lung-disease-lookup/copd/treating

[26] “Florida,” www.cdc.gov, Mar. 17, 2021. https://www.cdc.gov/nchs/pressroom/states/florida/fl3.htm

[27] “COPD: A National Health Issue Fact Sheet | NHLBI, NIH,” www.nhlbi.nih.gov, Sep. 02, 2022. https://www.nhlbi.nih.gov/resources/copd-national-health-issue-fact-sheet

[28] National heart, lung and blood institute, “COPD - diagnosis | NHLBI, NIH,” www.nhlbi.nih.gov, 2022. https://www.nhlbi.nih.gov/health/copd/diagnosis

[29] National Heart, Lung, and Blood Institute, “COPD - Treatment | NHLBI, NIH,” www.nhlbi.nih.gov, Mar. 24, 2022. https://www.nhlbi.nih.gov/health/copd/treatment

[30] N. C. for B. Information, U. S. N. L. of M. 8600 R. Pike, B. MD, and 20894 Usa, Chronic obstructive pulmonary disease (COPD): Surgical procedures for the treatment of COPD. Institute for Quality and Efficiency in Health Care (IQWiG), 2019. Available: https://www.ncbi.nlm.nih.gov/books/NBK543216/

[31] D. Bolton, “Conceptualisation of mental disorder and its personal meanings,” Journal of Mental Health, vol. 19, no. 4, pp. 328–336, Jul. 2010, doi: https://doi.org/10.3109/09638237.2010.492483.

[32] “Access to Care Data 2022,” Mental Health America. https://www.mhanational.org/issues/2022/mental-health-america-access-care-data

[33] SAMHSA, “2017 NSDUH Annual National Report | CBHSQ,” Samhsa.gov, 2017. https://www.samhsa.gov/data/report/2017-nsduh-annual-national-report [34] Zikmund, V. “Health--mental health--quality of life.” Bratislavske lekarske listy vol. 102,11 (2001): 527-9.

[35] “Population Query System - FL Health CHARTS - Florida Department of Health,” www.flhealthcharts.gov. https://www.flhealthcharts.gov/FLQUERY_New/Population/Count#

[36] “Health News Articles - US News Health,” @USNews, 2019. https://health.usnews.com

[37] “FloridaHealthFinder | Emergency Department Query Results,” quality.healthfinder.fl.gov. https://quality.healthfinder.fl.gov/QueryTool/QTResults?T=E# (accessed Apr. 16, 2024).

[38] “Social and Mental Health,” www.flhealthcharts.gov. https://www.flhealthcharts.gov/ChartsDashboards/rdPage.aspx?rdReport=CountyHealth.SocialAndMentalHealth&tabid=SocialAndMentalHealth

[39] “County Health Dashboard - Health Resource Availability,” www.flhealthcharts.gov. https://www.flhealthcharts.gov/ChartsDashboards/rdPage.aspx?rdReport=CountyHealth.HealthResourceAvailability&tabid=HealthResourceAvailability (accessed Apr. 16, 2024).

[40] “FL DOH MQA Search Portal | Practitioner Profile Search,” mqa-internet.doh.state.fl.us. https://mqa-internet.doh.state.fl.us/MQASearchServices/HealthCareProviders/PractitionerProfileSearch (accessed Apr. 16, 2024).

[41] H. Alabdulrazzaq, M. N. Alenezi, Y. Rawajfih, B. A. Alghannam, A. A. Al-Hassan, and F. S. Al-Anzi, “On the accuracy of ARIMA based prediction of COVID-19 spread,” *Results Phys.*, vol. 27, p. 104509, Aug. 2021, doi: 10.1016/j.rinp.2021.104509.

[42] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[43] S. Makridakis et al., “The accuracy of extrapolation (time series) methods: Results of a forecasting competition,” Journal of Forecasting, vol. 1, no. 2, pp. 111–153, Apr. 1982, doi: https://doi.org/10.1002/for.3980010202.

[44] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” International Journal of Forecasting, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: https://doi.org/10.1016/j.ijforecast.2006.03.001.

[45] C. J. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” Climate Research, vol. 30, no. 1, pp. 79–82, 2005, Available: https://www.jstor.org/stable/24869236

[46] R. J. Hyndman and Y. Khandakar, “Automatic Time Series Forecasting: the Forecast Package for R,” Journal of Statistical Software, vol. 27, no. 3, 2008, doi: https://doi.org/10.18637/jss.v027.i03.

[47] K. P. Burnham and D. R. Anderson, Eds., Model Selection and Multimodel Inference. New York, NY: Springer New York, 2004. doi: https://doi.org/10.1007/b97636.

[48] T. Chen and C. Guestrin, “XGBoost: a Scalable Tree Boosting System,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD ’16, pp. 785–794, 2016, doi: https://doi.org/10.1145/2939672.2939785.

[49] J. Bergstra, J. Ca, and Y. Ca, “Random Search for Hyper-Parameter Optimization Yoshua Bengio,” Journal of Machine Learning Research, vol. 13, pp. 281–305, 2012, Available: https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf

[50] A. E. Permanasari, I. Hidayah, and I. A. Bustoni, “SARIMA (Seasonal ARIMA) implementation on time series to forecast the number of Malaria incidence,” IEEE Xplore, Oct. 01, 2013. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6676239

[51] W. Jiang, X. Wu, Y. Gong, W. Yu, and X. Zhong, “Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption,” Energy, vol. 193, p. 116779, Feb. 2020, doi: https://doi.org/10.1016/j.energy.2019.116779.

[52] W. Junthopas and C. Wongoutong, “Setting the Initial Value for Single Exponential Smoothing and the Value of the Smoothing Constant for Forecasting Using Solver in Microsoft Excel,” Applied Sciences, vol. 13, no. 7, p. 4328, Mar. 2023, doi: https://doi.org/10.3390/app13074328.

[53] C. Ni et al., “Light Gradient Boosting Machine (LightGBM) to forecasting data and assisting the defrosting strategy design of refrigerators,” International journal of refrigeration, Jan. 2024, doi: https://doi.org/10.1016/j.ijrefrig.2024.01.025.

[54] C.-L. Hung, “Deep learning in biomedical informatics,” pp. 307–329, Jan. 2023, doi: https://doi.org/10.1016/b978-0-323-85796-3.00011-1.

[56] C. M. Stauffer and C. Pfeifer, “Colonoscopy,” PubMed, 2024. http://www.ncbi.nlm.nih.gov/books/NBK559274/ (accessed Apr. 16, 2024).

[57] KFF, “Mental Health Care Health Professional Shortage Areas (HPSAs),” The Henry J. Kaiser Family Foundation, Nov. 21, 2019. https://www.kff.org/other/state-indicator/mental-health-care-health-professional-shortage-areas-hpsas/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22

[58] “Hospital Search Results,” CathLab.com. https://cathlab.com/hospital-lab/hospital-search-results/?wpv_view_count=8292&wpv_post_search=&wpv-hospital-city=0&wpv-hospital-state=florida&wpv_filter_submit=Search (accessed Apr. 16, 2024).

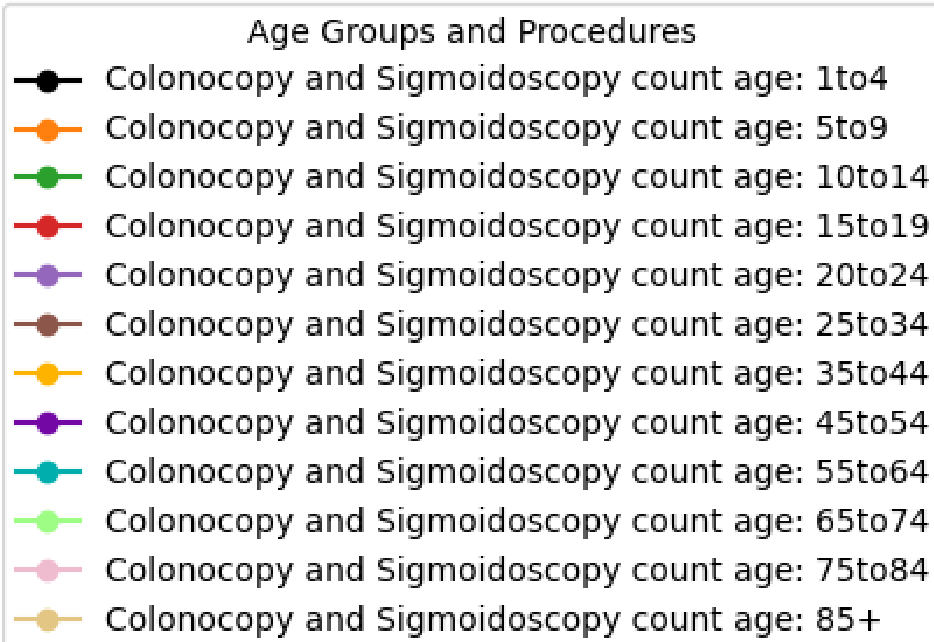
[59] R. Kumar et al., “Comparison of Outcomes of Percutaneous Coronary Intervention Versus Coronary Artery Bypass Grafting,” Cureus, vol. 12, no. 12, p. e12202, Dec. 2020, doi: https://doi.org/10.7759/cureus.12202.

[60] A. C. Fanaroff et al., “Outcomes of PCI in Relation to Procedural Characteristics and Operator Volumes in the United States,” Journal of the American College of Cardiology, vol. 69, no. 24, pp. 2913–2924, Jun. 2017, doi: https://doi.org/10.1016/j.jacc.2017.04.032.

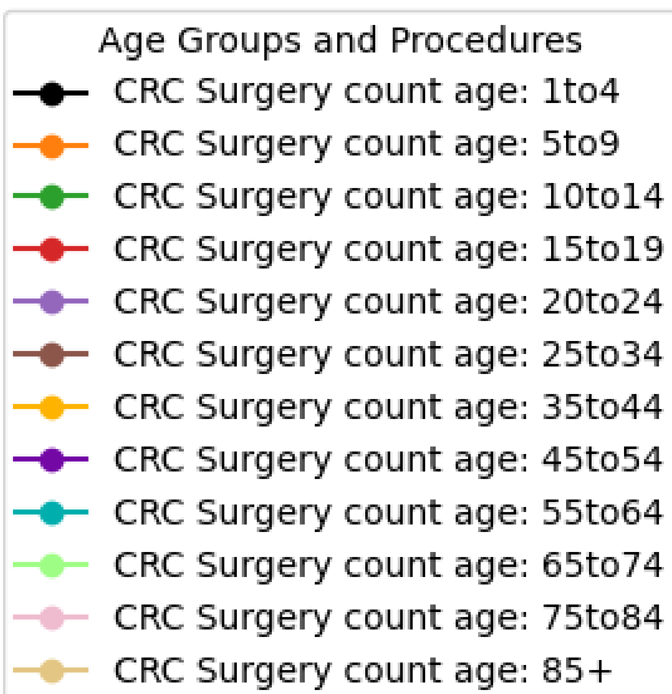
[61] P. W. Burvill, “Looking beyond the 1:10,000 ratio of psychiatrists to population,” The Australian and New Zealand Journal of Psychiatry, vol. 26, no. 2, pp. 265–269, Jun. 1992, Available: https://pubmed.ncbi.nlm.nih.gov/1642618/#:~:text=The%20Royal%20Australian%20and%20New

SUPPLEMENTARY MATERIALS

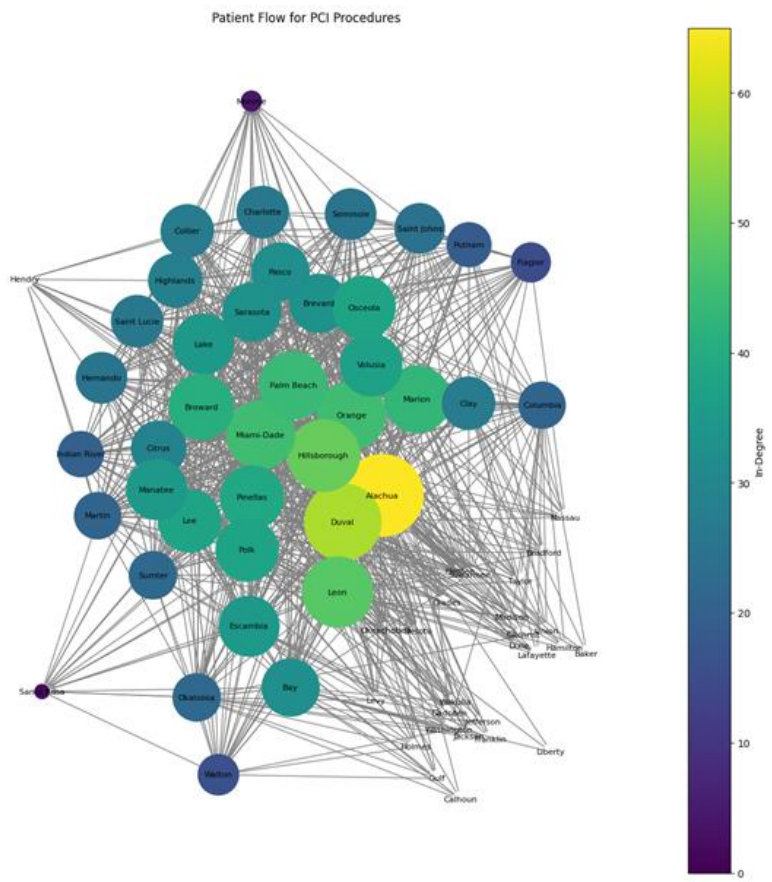
S1. Enlarged Key for Figure 4



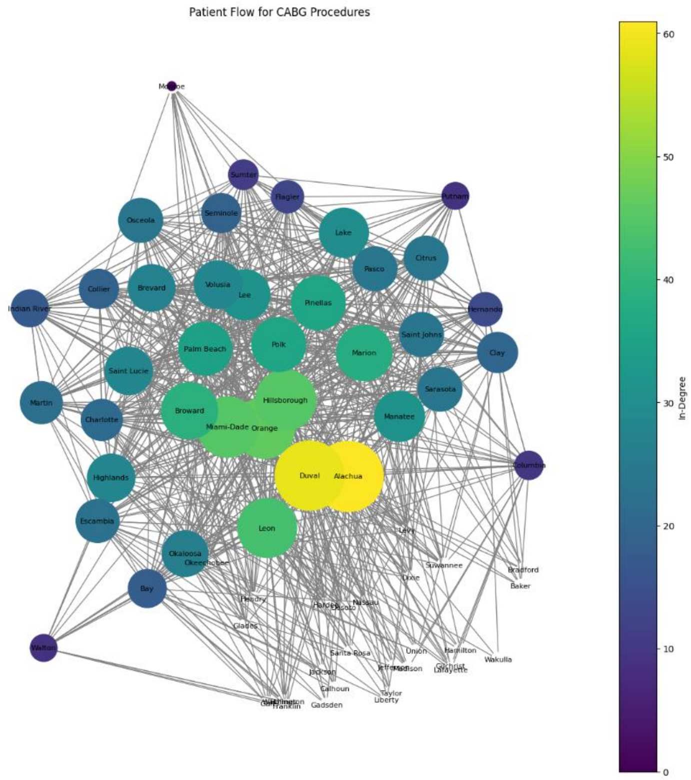
S2. Enlarged Key for Figure 5



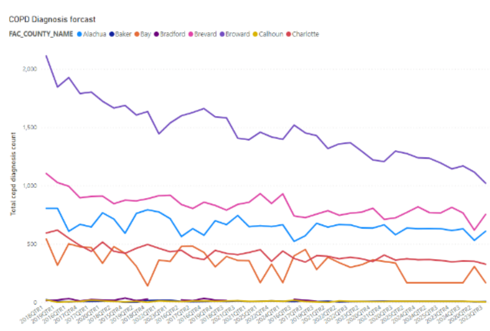
S3. Enlarged PCI patients flow for Cath Lab facility counties to get treated.

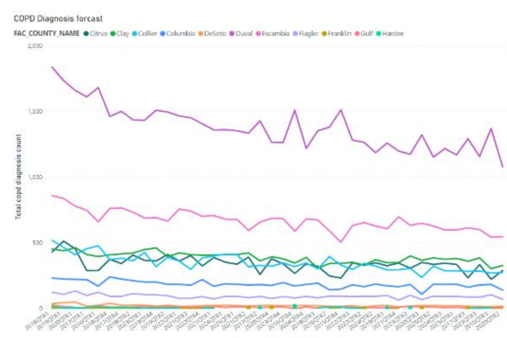


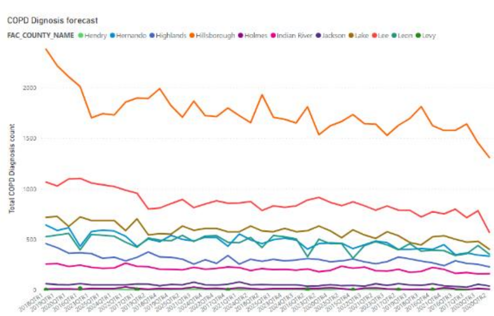
S4. Enlarged Figure CABG patients flow for Heart surgery facility counties to get treated.

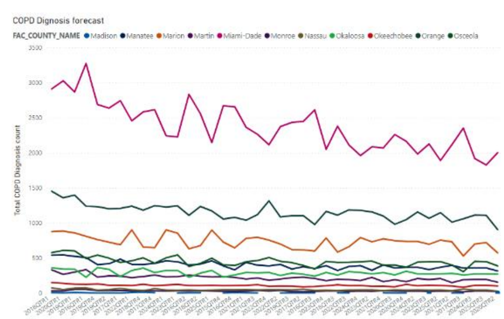


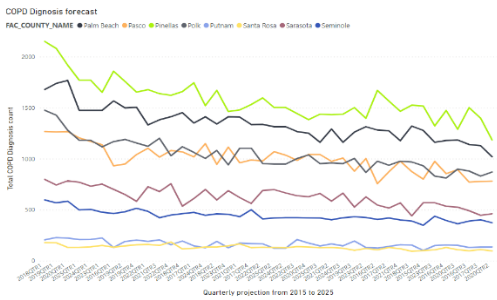
S5. Charts for COPD ARIMA results

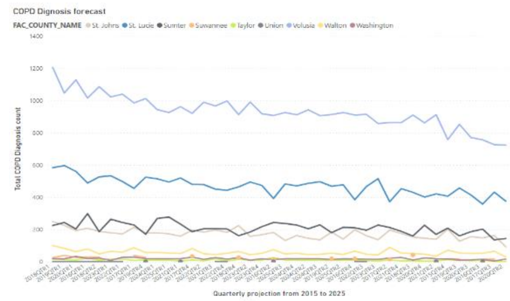




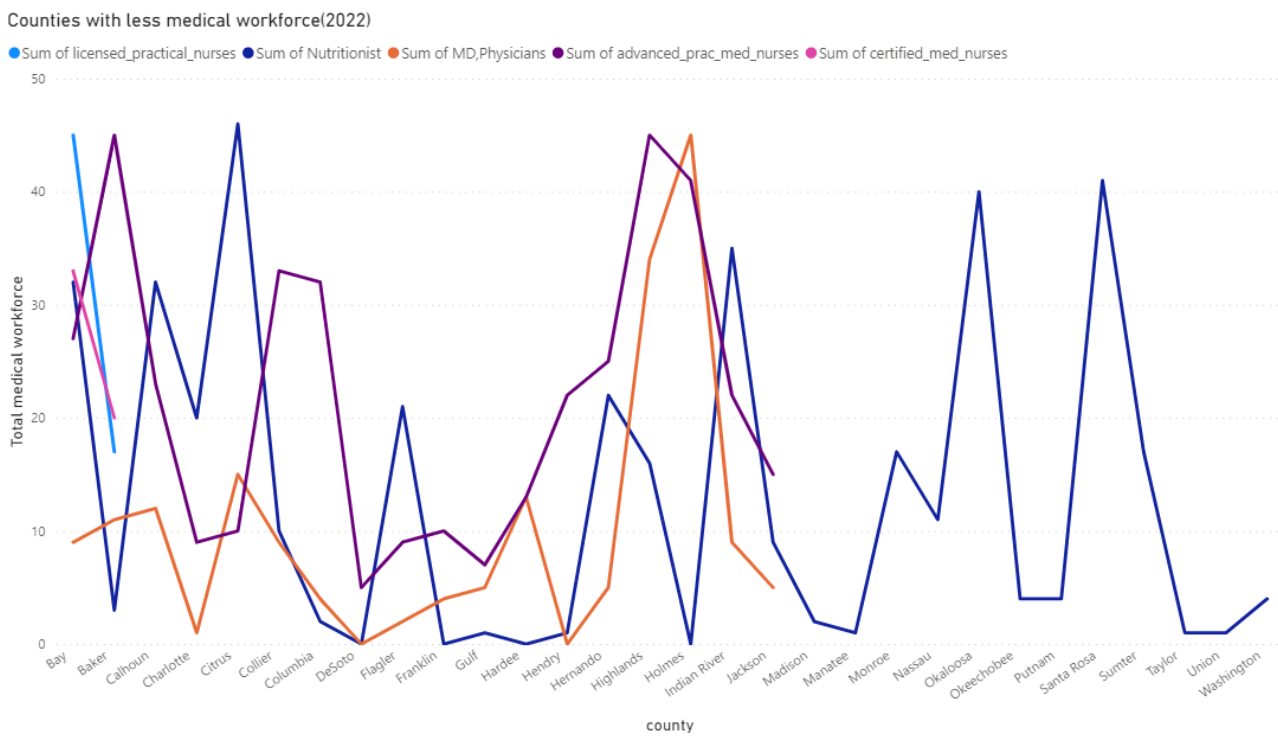


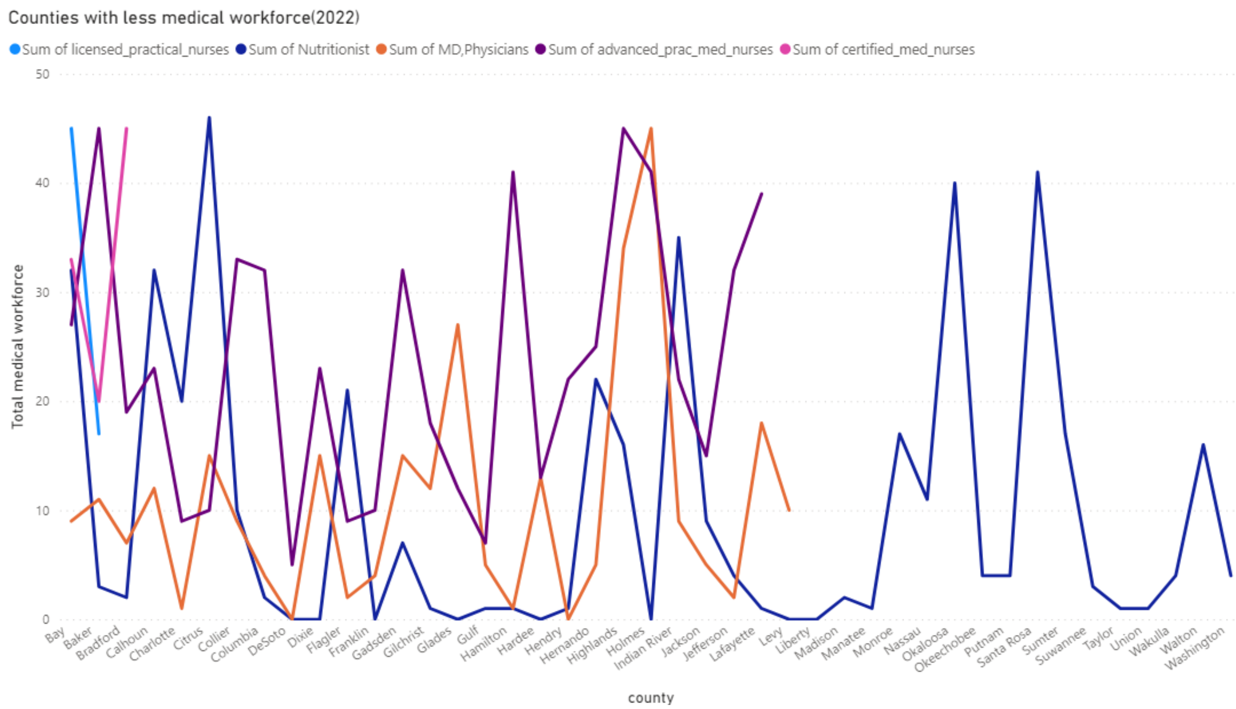




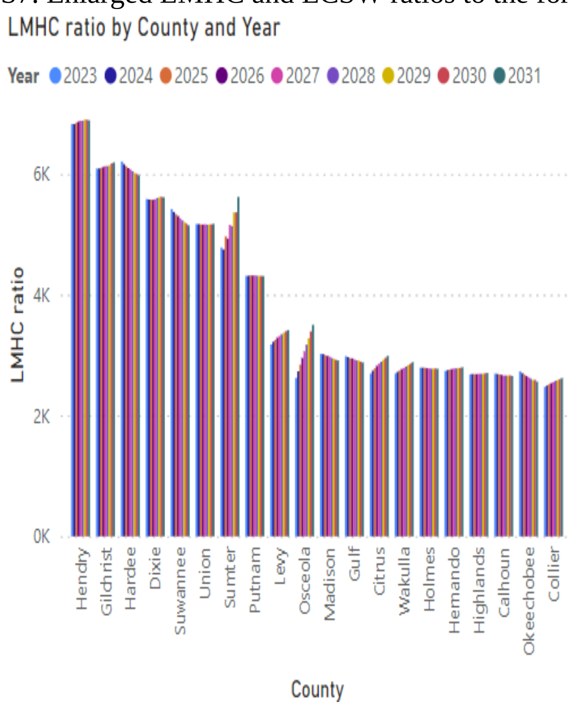


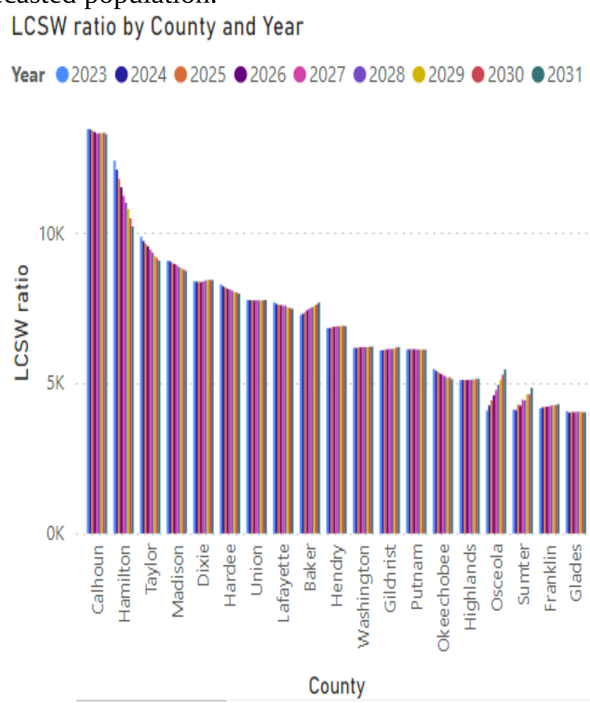
S6. Enlarged COPD Medical Workforce Charts



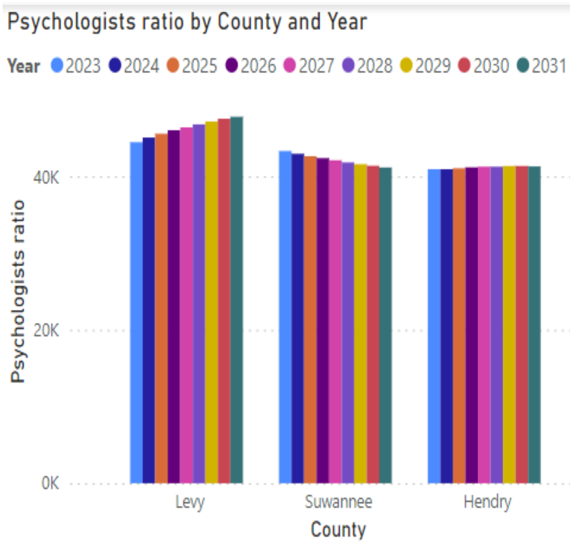


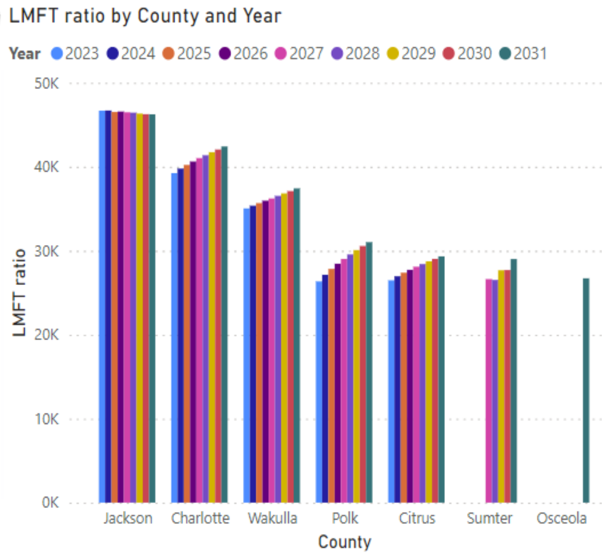
S7. Enlarged LMHC and LCSW ratios to the forecasted population.





S8. Enlarged LMFT to population ratio and psychologists to population ratio.


Psychologists ratio by County and Year
Year 2023 2024 2025 2026 2027 2028 2029 2030 2031
Psychologists ratio
40K
20K
0K
Levy
Suwannee
Hendry
County


LMFT ratio by County and Year
Year 2023 2024 2025 2026 2027 2028 2029 2030 2031
LMFT ratio
50K
40K
30K
20K
10K
0K
Jackson
Charlotte
Wakulla
Polk
Citrus
Sumter
Osceola
County