Abstract

Introduction

Related work

3. Method

This study investigates racial, ethnic, and socioeconomic disparities in healthcare utilization among adult patients with disabilities, using electronic health record (EHR) data from OCHIN. The unit of analysis is the individual patient visit, with a dataset comprising 170,660 encounters from 2013 to 2021. Our analysis focuses specifically on patients aged 18–64. Variables such as insurance coverage (including Medicare and Medicaid), income level, and encounter type were examined in relation to the specific objectives of the study. We structure our methodology around three key research hypotheses, each of which requires tailored data cleaning, transformation, and statistical analysis, as detailed below.

3.1 Data Source: OCHIN EHR System

OCHIN operates a centralized, shared EHR system that aggregates real-time clinical, demographic, financial, and utilization data from community health centers across more than 30 U.S. states. Participating clinics contribute data through a secure infrastructure. To examine healthcare inequities among adults with disabilities, we used the following data categories, as shown in table 1:

Table 1: Selected variables used in the analysis.

Variable Type	Details
Encounter Data	• Ambulatory visits (AV) • Telehealth visits (TH) • Frequency of encounters per patient
Demographics	• Race and ethnicity • Federal Poverty Level (FPL)
Insurance Coverage	• Medicare • Medicaid

3.2 Software Used

We conducted all data processing, statistical analysis, and visualization using Python within Jupyter Notebook (Version 7.2.2). The following Python libraries were utilized, as shown in table 2:

Table 2: Python libraries used for the analysis.

Library	Purpose
pandas	Data cleaning, transformation, and manipulation
numpy	Numerical operations
matplotlib	Plot generation and visualizations
seaborn	Enhanced data visualizations
scipy.stats	Statistical tests (ANOVA, Shapiro-Wilk, Levene's test, Kruskal-Wallis H test)
statsmodels	Advanced statistical modeling and post hoc tests
scikit-posthocs	Post hoc testing (Dunn's test with Bonferroni correction)
openpyxl	Excel file handling

We executed all code within a reproducible research environment, allowing step-by-step documentation of the data analysis workflow.

3.3 **Variable Selection and Dataset Preparation**

We selected key variables based on their relevance to our research questions on healthcare access and disparities. Figure 1 illustrates a sample of the dataset along with the structure of the date-related variables. These variables included:

- PATIENT_NUM, BIRTH_DATE, and AGE_AT_ENCOUNTER_DATE: Used to identify individual patients, compute age, and apply the age filter (18–64 years).
- RACE_CD, HISPANIC_CD, and Race: Used to group patients into three primary racial/ethnic categories: White (Non-Hispanic), Black (Non-Hispanic), and Hispanic (any race).
- CURRENT_FPL_CD: Used to classify socioeconomic status via the Federal Poverty Level.
- STATE_CD: Used to filter out unknown or invalid geographic locations.

- START_DATE and derived visit_year: Used to determine the timing of visits and support longitudinal analyses.
- first_diagnosis_date: Used to identify the onset of chronic conditions or first healthcare encounter.
- ENC_TYPE and PAYOR_TYPE_RESEARCH: Used to classify the type of visit (AV or TH) and insurance (Medicare or Medicaid).

We cleaned and prepared the data separately for each research objective, as described in the following subsections.

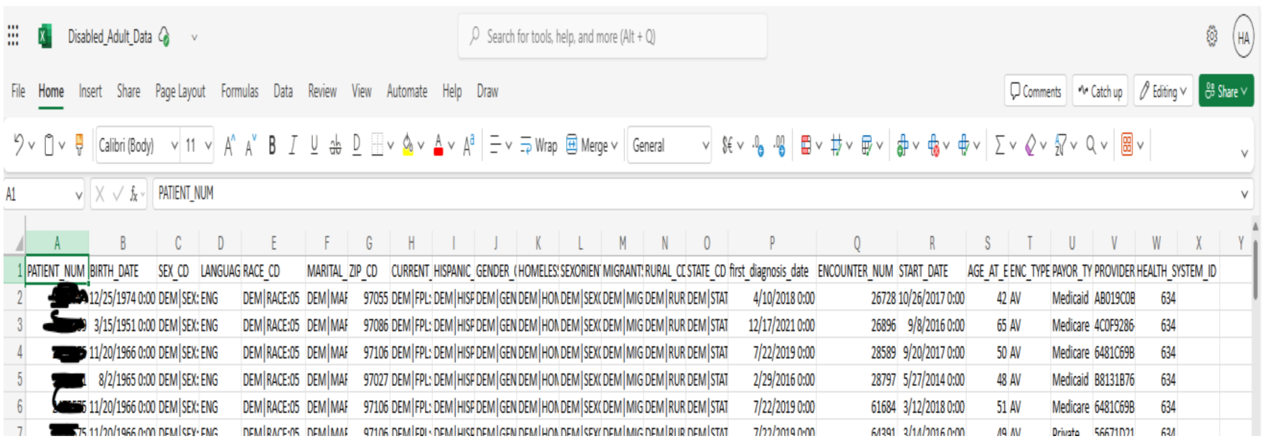

Figure 1 a sample of the dataset along with the structure of the date-related variables.

3.4 Objective 1: Annual Visit Count Disparities by Race/Ethnicity Among Medicaid and Medicare Patients (2013–2021)

To address this objective, we filtered the data to include only Medicaid or Medicare encounters for patients aged 18–64. We focused on three racial/ethnic groups: White Non-Hispanic, Black Non-Hispanic, and Hispanic (any race). We extracted the year from the START_DATE to create a visit_year variable. For patients with multiple visits in a single year, we retained only one visit per patient per year, ensuring each record reflected a unique patient-year. This avoided overrepresentation of frequent users and enabled longitudinal comparisons. Figure 2 illustrates the data cleaning and preparation workflow for Objective 1, using the dataset shown in Figure 1.

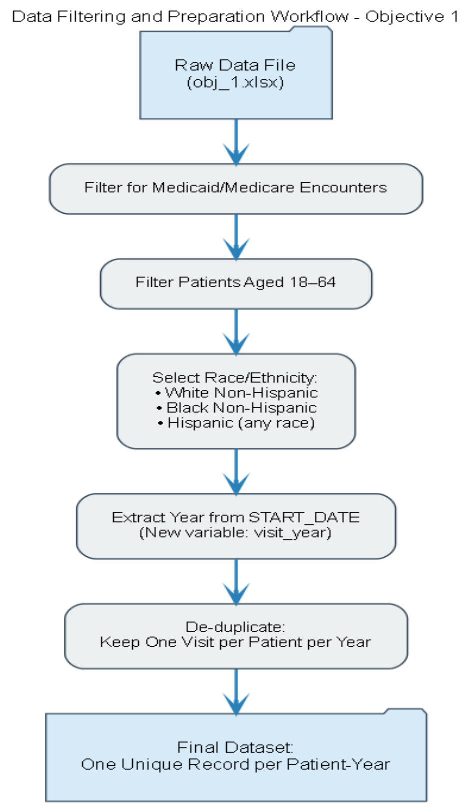

Figure 2 Data cleaning and preparation process for Objective 1

3.5 Objective 2: Examine Racial and Ethnic Disparities in Encounter Type (Ambulatory vs. Telehealth Visits)

We cleaned the data to test our hypothesis for Objective 2. Figure 3 illustrates the steps we follow to clean and prepare the data for analysis. We begin by ingesting the raw Excel file and loading it into our processing environment (Python, Jupyter Notebook Version 7.2.2). Next, we clean key columns by normalizing encounter types, converting date formats, and sorting the data by START_DATE, which we use to identify the visit date. We then categorize each encounter based on type (such as Both, Only AV, or Only TH) before removing redundant visits, keeping only the latest record for each patient. This step is essential to ensure data independence, which is required for the selected statistical test for this hypothesis. We include only data from 2020–2021, as data from earlier years was not significant for the analysis. Finally, we produce a clean, structured dataset ready for analysis. The figure summarizes each step we take in a clear, top-down format that reflects our systematic approach.

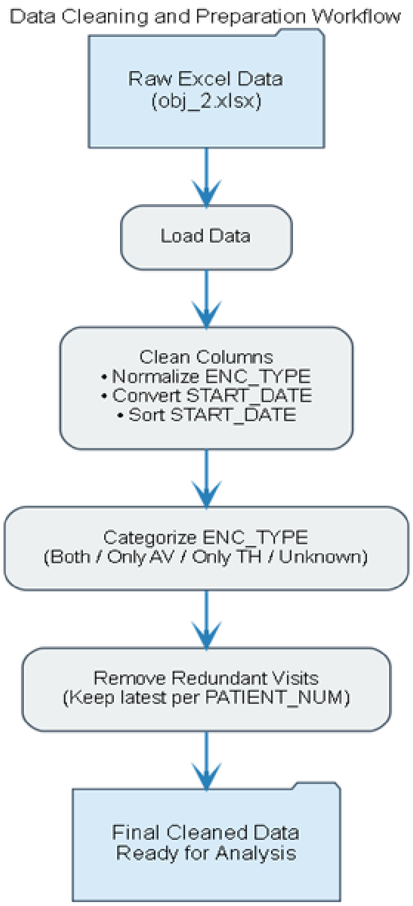

Figure 3 Data cleaning and preparation process for objective 2

3.6 Objective 3: Examine Racial and Ethnic Disparities in Healthcare Utilization Among Low-Income Patients

We began by cleaning and preparing the dataset to test our hypothesis for Objective 3. First, we computed the number of healthcare visits for each patient and created a new variable, Frequency_of_visit, to store these aggregated counts. To eliminate redundancy, we retained only one record per patient, reflecting their total number of visits. Next, we filtered the data to include only patients living at or below 100% of the Federal Poverty Level (DEM|FPL $\leq$ 100), ensuring the analysis targeted low-income individuals. Figure 4 summarizes the data preparation steps.

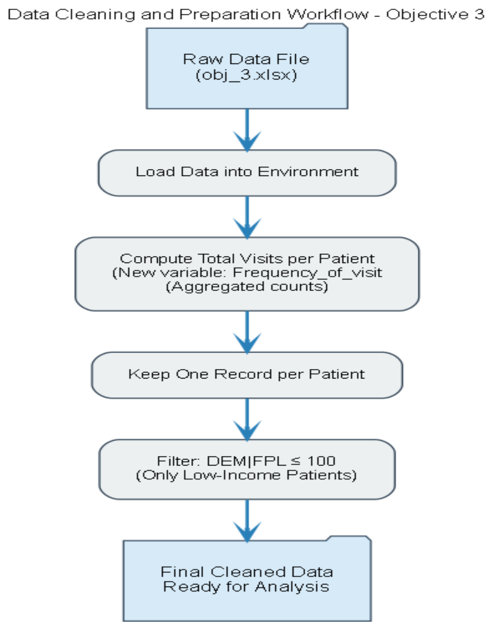

Figure 4 Data cleaning and preparation for objective 3

3.7 Ethical Considerations

This study used de-identified patient data, and no patient-identifiable information was accessed or disclosed. Data security and confidentiality were strictly maintained throughout all phases of data handling and analysis.

3.8 Limitations

Our analyses rely on EHR data, which may be subject to documentation inconsistencies and missing information. Race and ethnicity classifications depend on patient self-reporting or administrative records and may not capture the full diversity within groups. Additionally, our focus on patients aged 18–64 who are covered by Medicare, Medicaid, or classified as low-income limits the generalizability of our findings to other populations.

3.9 Data Governance

All data manipulations and analyses were conducted in secure computing environments with restricted user access.

3.10 Data Quality Assurance

We conducted initial data quality checks to ensure accuracy and completeness. Records with missing or inconsistent key variables were excluded from the analysis. Outliers in visit frequency were carefully examined and removed, particularly for Objective 3, due to inconsistencies in telehealth usage over the years. Telehealth visits were relatively rare before 2020 and became widespread during the COVID-19 pandemic. Therefore, data from 2019 were excluded because of the incomparable number of visits, which reflected the initial emergence of telehealth rather than normalized utilization patterns observed from 2020 onward.

4. Results and Analysis

In this section, we test hypotheses based on each of this study, we use a various statistically model to appropriate to test each hypothesis.

Objective 1: Annual Visit Count Disparities by Race/Ethnicity Among Medicaid and Medicare Patients (2013–2021)

To address the first objective, we analyzed racial and ethnic disparities in patient visit counts from 2013 to 2021, specifically among patients covered by Medicaid and Medicare. We compared annual visit counts across three racial/ethnic groups: Black (Non-Hispanic), Hispanic (any race), and White (Non-Hispanic). Figure 5 presents a box plot of total patient counts by race, while Figure 6 illustrates annual trends in patient visits across the study period. Each patient was counted only once per year in order to prevent duplication. To determine whether mean visit counts significantly differed among the racial/ethnic groups, we formulated and tested the following hypotheses using a one-way analysis of variance (ANOVA):

Null hypothesis (H_0): There is no significant difference in mean annual visit counts among the three racial/ethnic groups.

$$\mu_{Black} = \mu_{Hispanic} = \mu_{White}$$

Alternative hypothesis (H_1): At least one group differs significantly in mean annual visit counts.

$$\exists\ i, j \in \{Black, Hispanic, White\}\ such\ that\ u_i \neq u_j$$

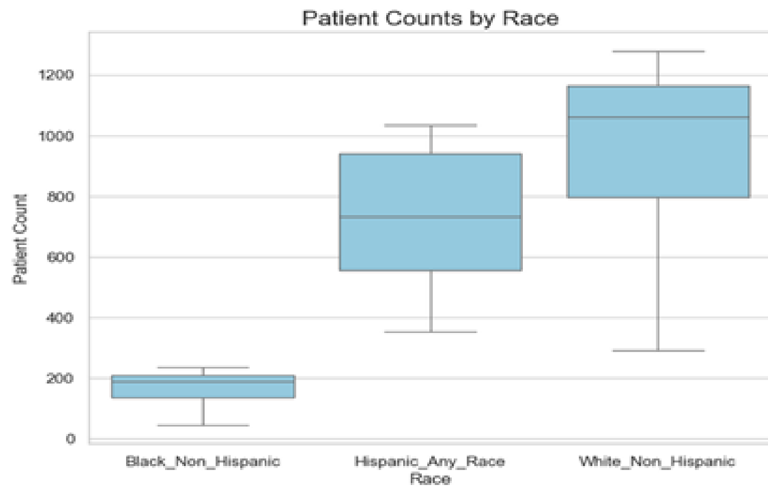

Figure 5 A box plot showing total patient counts by race

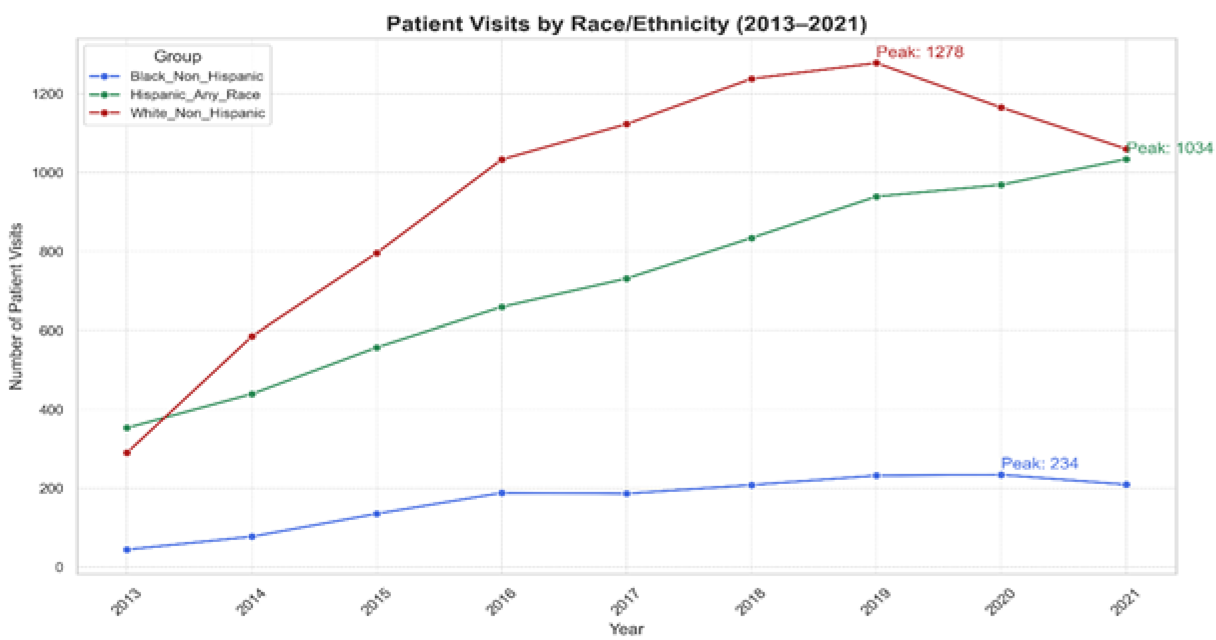

Figure 6 Annual patient visit trends by race during the 2013–2021 period

Before conducting the ANOVA, we evaluated whether the data met the assumptions necessary for parametric testing. We used the Shapiro-Wilk test to assess normality and found no significant deviations from normality for any group: Black (Non-Hispanic) ($p=0.0958$), Hispanic (any race) ($p=0.6934$), and White (Non-Hispanic) ($p=0.1289$). We also visually inspected the distributions using histograms and Q-Q plots, as shown in Figure 7, which supported the assumption of approximate normality.

We then tested for homogeneity of variances using Levene's test and found no significant differences in group variances ($p=0.0561$). Since all assumptions were met, we proceeded with the one-way ANOVA.

The ANOVA results shown in Table 3 indicate a statistically significant effect of race/ethnicity on patient visit counts, $F(2, 24) = 25.43$, $p < 0.001$, which exceeded the critical F-value ($F_{crit}=3.40$). Based on these results, we rejected the null hypothesis and concluded that at least one group's mean visit count differed significantly.

Table 3. One-Way ANOVA Summary: Patient Visit Counts by Race/Ethnicity

Summary Statistics

Group	Count	Sum	Average	Variance
Black (Non-Hispanic)	9	1,513	168.11	4,650.36
Hispanic (Any Race)	9	6,515	723.89	58,184.86
White (Non-Hispanic)	9	8,568	952.00	109,772.00

ANOVA Table

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2,926,216.22	2	1,463,108.11	25.43	1.18×10^{-6}	3.40
Within Groups	1,380,857.78	24	57,535.74	-	-	-
Total	4,307,074.00	26	-	-	-	-

We computed descriptive statistics to summarize the group means. Black (Non-Hispanic) patients had the lowest average visit count ($M = 168.11$), followed by Hispanic (any race) patients ($M = 723.89$), and White (Non-Hispanic) patients ($M = 952.00$). To determine which specific group differences were statistically significant, we conducted a Tukey Honestly Significant Difference (HSD) post hoc test.

We used the Tukey HSD test to examine pairwise differences in mean patient visit counts between racial and ethnic groups. The results showed that Black (Non-Hispanic) patients had significantly lower mean visit counts compared to both Hispanic (any race) patients ($p < 0.001$) and White

(Non-Hispanic) patients ($p < 0.001$). However, we did not find a statistically significant difference between Hispanic and White patients ($p = 0.130$). Table 4 summarizes the pairwise comparisons, and Figure 8 illustrates the differences in mean patient visit counts by race from 2013 to 2021.

Table 4. Tukey HSD Test Results for Pairwise Comparisons of Mean Patient Visits

(Family-Wise Error Rate = 0.05)

Group 1	Group 2	Mean Difference	*p*-adj	95% CI Lower	95% CI Upper	Significant
Black (Non-Hispanic)	Hispanic (Any Race)	555.78	0.0001	273.40	838.16	Yes
Black (Non-Hispanic)	White (Non-Hispanic)	783.89	<0.0001	501.51	1,066.27	Yes
Hispanic (Any Race)	White (Non-Hispanic)	228.11	0.130	–54.27	510.49	No

We also calculated the effect size using eta squared ($\eta2=0.6785$). This result indicates that approximately 67.85% of the variance in patient visit counts can be explained by racial/ethnic group membership, a large effect size.

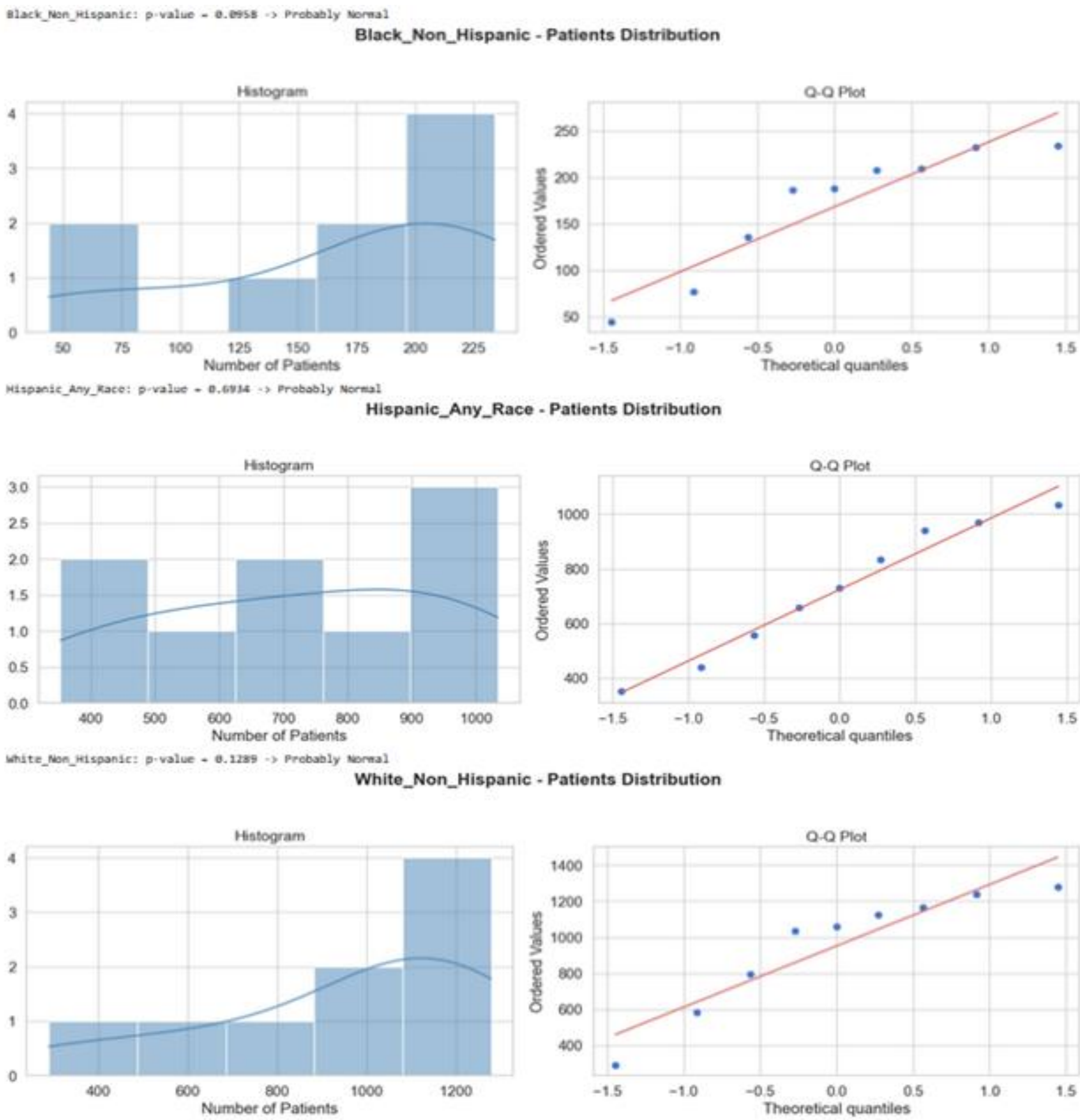

Figure 7 Assumption of approximate normality

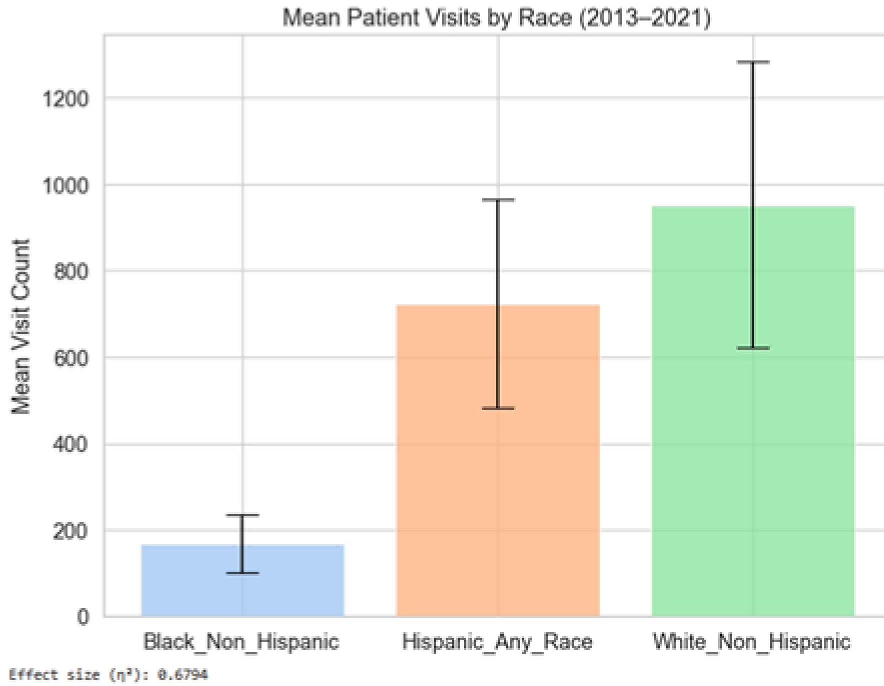

figure 8 Mean Patient Count Difference

Objective 1 Interpretation:

Our analysis reveals statistically significant racial and ethnic disparities in patient visit counts from 2013 to 2021. Specifically, Black (Non-Hispanic) patients had substantially fewer annual visits compared to both Hispanic (any race) and White (Non-Hispanic) patients. These differences were not only statistically significant ($p < 0.001$) but also practically meaningful, as reflected in the large effect size ($\eta^2 = 0.6785$), indicating that nearly 68% of the variance in visit counts can be attributed to racial/ethnic group membership.

The lack of a significant difference between Hispanic and White patients ($p = 0.130$) suggests that these two groups had relatively comparable levels of healthcare utilization within this dataset. In contrast, the consistently lower visit counts among Black patients point to a clear and persistent disparity. These findings may reflect systemic barriers to healthcare access and utilization that disproportionately affect Black communities. Taken together, our results call for targeted policy and intervention strategies to improve access, engagement, and equity in healthcare delivery, particularly for Black (Non-Hispanic) populations.

Objective 2: Examine Racial and Ethnic Disparities in Encounter Type (Ambulatory vs. Telehealth Visits)

To address the second objective of this study, we examined whether the distribution of patients who had both ambulatory visits (AV) and telehealth visits (TH) versus those who had only ambulatory visits differed by racial/ethnic group during the years 2020–2021. These distributions are visualized in Figure 9 and Figure 10. Table 5 presents descriptive statistics for each racial/ethnic group based on their use of ambulatory versus telehealth visits. To evaluate potential differences, we conducted a chi-square test of independence to test the following hypotheses:

Null Hypothesis (H_0): The distribution of patients who had both ambulatory visits (AV) and telehealth visits (TH) versus only ambulatory visits (AV) is the same across all racial/ethnic groups during 2020–2021.

$$H_0: P(Encounter\ Type | Race\ /\ Ethnicity) = P(Encounter\ Type)$$

Alternative Hypothesis (H_1): The distribution of patients across these encounter types differs by racial/ethnic group during 2020–2021.

$$H_1: P(Encounter\ Type | Race\ /\ Ethnicity) \neq P(Encounter\ Type)$$

In this context, “Encounter Type” refers to two categories: “Both AV & TH” and “Only AV.”

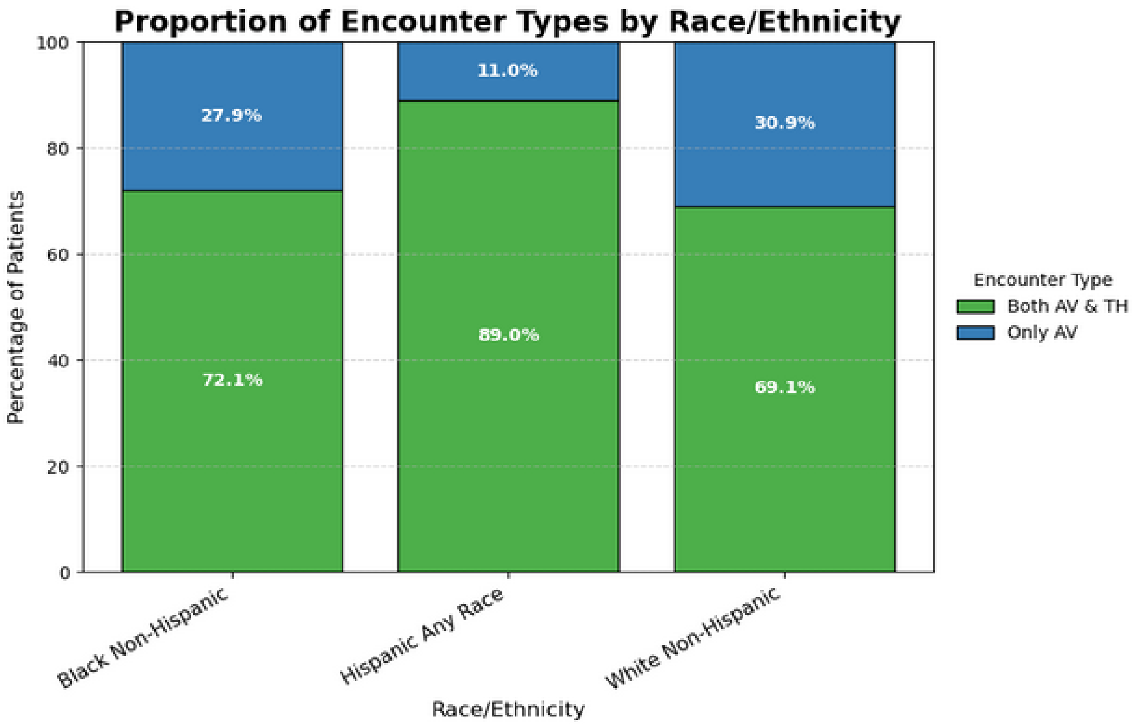

Figure 9 Observed Encounter Types by Race/Ethnicity

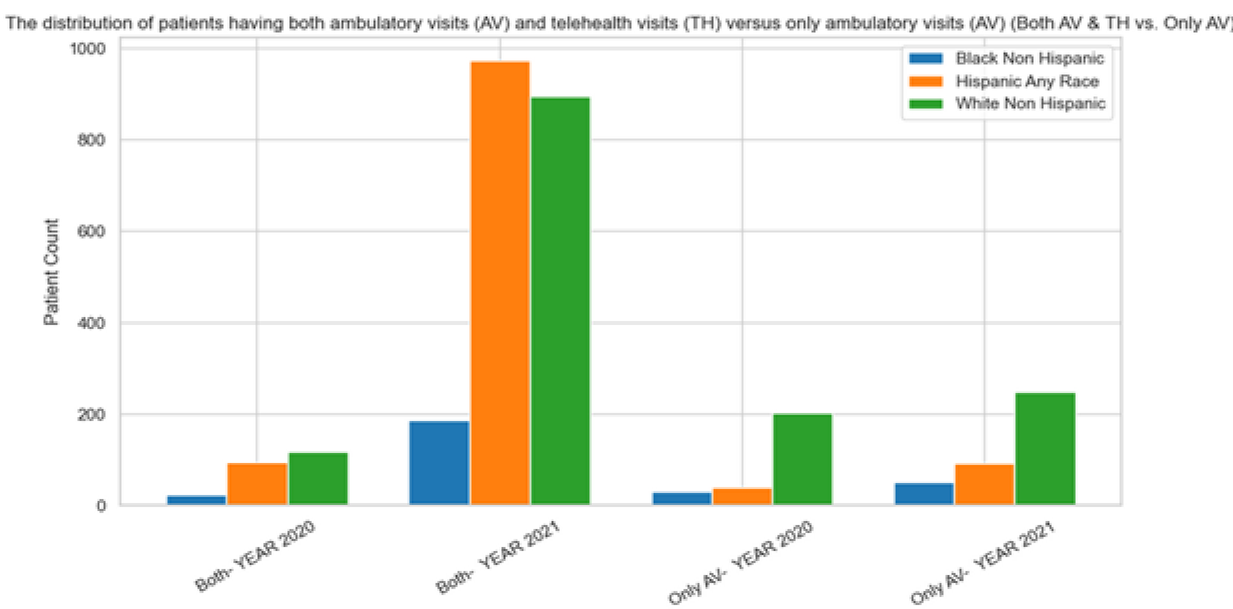

Figure 10 Distribution of patients having both ambulatory visits (AV) and telehealth visits (TH) versus only ambulatory visits (AV)

Table 5. Descriptive Statistics Table for Objective 2

STATISTIC	BLACK NON-HISPANIC	HISPANIC ANY RACE	WHITE NON-HISPANIC

COUNT	2.0	2.0	2.0
MEAN	147.0	600.0	732.5
STANDARD DEVIATION	91.92	661.85	395.27
MIN	82.0	132.0	453.0
25%	114.5	366.0	592.75
MEDIAN (50%)	147.0	600.0	732.5
75%	179.5	834.0	872.25
MAX	212.0	1068.0	1012.0
SUM	294.0	1200.0	1465.0
RANGE	130.0	936.0	559.0
COV (%)	62.53%	110.31%	53.97%

We conducted a chi-square test of independence by analyzing observed patient counts stratified by encounter type and race/ethnicity, as presented in Table 6. Before performing the test, we verified that all assumptions were satisfied. The data consisted of mutually exclusive and independent frequency counts, and all expected cell frequencies exceeded the minimum threshold of five, ranging from 66.27 to 1134.77. Table 7 presents the expected frequencies that confirm this assumption. Figure 11 provides a visual comparison between observed and expected patient counts.

We calculated a chi-square statistic of 155.29 with 2 degrees of freedom, resulting in a p-value of 1.9×10^{-34}. Given this extremely small p-value (less than 0.05), we reject the null hypothesis. These results indicate that the distribution of patients across encounter types significantly differs by racial/ethnic group during 2020-2021.

Table 6. Observed Patient Counts by Encounter Type and Race/Ethnicity (2020–2021)

Encounter Type	*Black Non-Hispanic*	*Hispanic Any Race*	*White Non-Hispanic*
Both AV & TH	212	1068	1012
Only AV	82	132	453

Table 7. Expected Patient Counts Under the Null Hypothesis

Encounter Type	*Black Non-Hispanic*	*Hispanic Any Race*	*White Non-Hispanic*
Both AV & TH	227.73	929.50	1134.77
Only AV	66.27	270.50	330.23

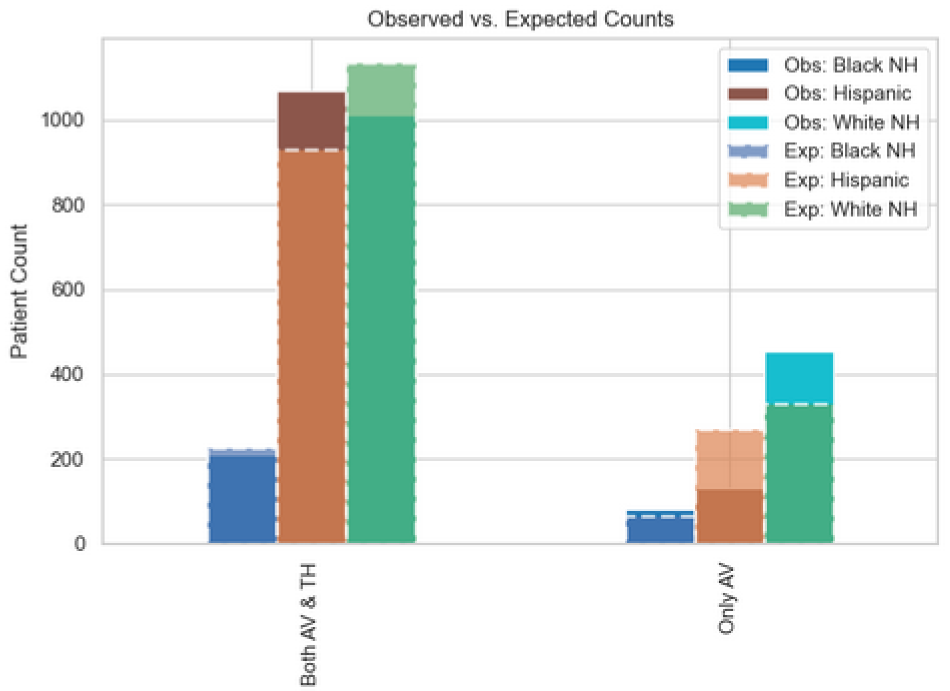

Figure 11 Comparison between observed and expected patient counts

Objective 2 Interpretation:

By rejecting the null hypothesis, we provide strong evidence that the distribution of encounter types differs significantly across racial/ethnic groups. This finding indicates that patients' utilization patterns of telehealth and ambulatory visits during 2020–2021 were not uniform among Black Non-Hispanic, Hispanic (Any Race), and White Non-Hispanic populations.

To further understand which specific racial/ethnic groups contributed most to the overall chi-square result, we examined the standardized residuals from each cell in the contingency table. Standardized residuals indicate how much the observed count in each cell differs from its expected count, measured in units of standard deviation. Values greater than +2 or less than –2 suggest meaningful, statistically significant deviations from what we would expect under the assumption of independence between encounter type and race/ethnicity.

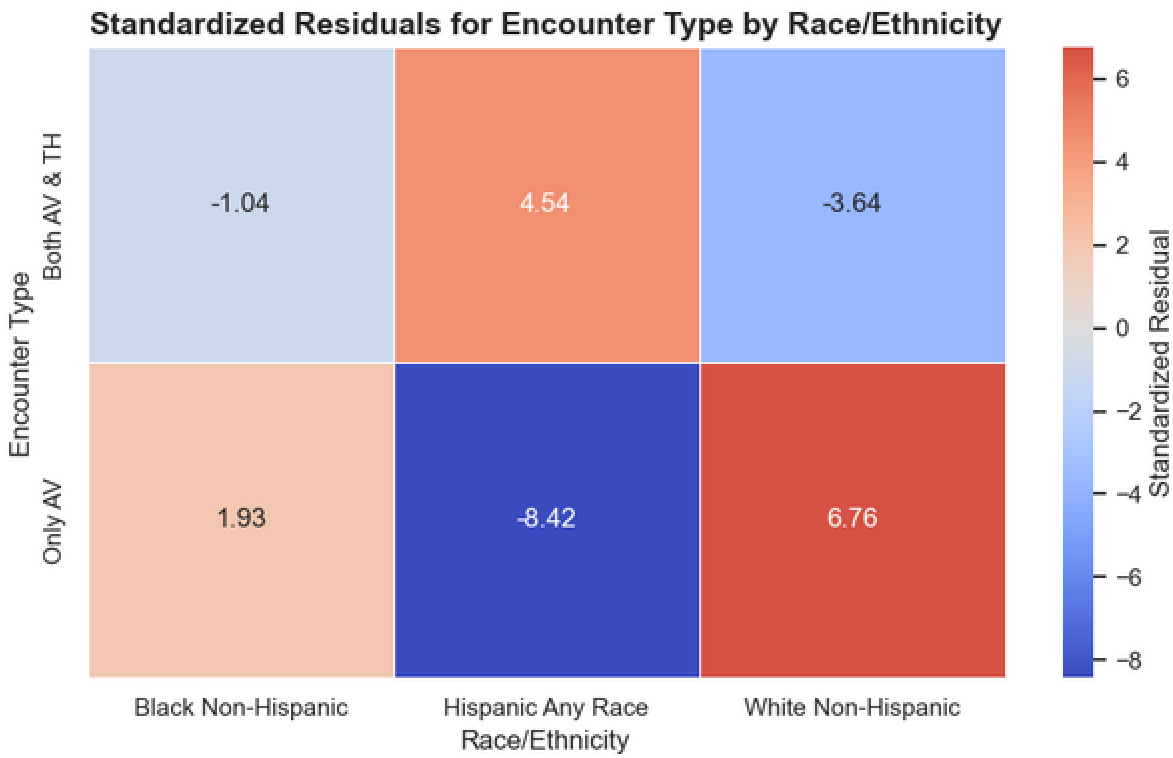

Figure 12 Standardized Residuals for Encounter Type by Race/Ethnicity

As shown in Figure 12, Hispanic patients exhibited the most notable pattern. We observed a standardized residual of +4.54 in the “Both AV & TH” category, meaning significantly more Hispanic patients received both telehealth and ambulatory care than expected. Conversely, the residual of –8.42 for the “Only AV” category among Hispanic patients indicates substantially fewer received ambulatory-only care than expected. This suggests that Hispanic patients may have had greater engagement with or access to telehealth services during the COVID-19 pandemic in 2020–2021.

Among White Non-Hispanic patients, the pattern reversed. We observed a residual of –3.64 for “Both AV & TH,” indicating that fewer White patients than expected used both visit types. Meanwhile, the residual of +6.76 for “Only AV” shows that more White patients relied exclusively on ambulatory care than the model anticipated. For Black Non-Hispanic patients, all residuals remained near ±2, suggesting no statistically significant deviation from the expected distribution. This indicates their visit pattern distribution aligned more closely with what we would expect under independence. These residuals support our earlier findings and help clarify which groups contributed most to the observed disparities. The combination of statistical testing and visual inspection of standardized residuals reveals meaningful racial/ethnic differences in how patients used ambulatory and telehealth services during 2020–2021.

Objective 3: Examine Racial and Ethnic Disparities in Healthcare Utilization Among Low-Income Patients

We examined whether healthcare utilization, measured by frequency of visits, differs significantly across racial and ethnic groups among low-income patients (only patients living at or below 100% of the Federal Poverty Level (DEM|FPL $\leq$ 100). Specifically, we compared visit frequencies among three groups: White Non-Hispanic, Black Non-Hispanic, and Hispanic of Any Race. Table 8 provides descriptive statistics by racial/ethnic group among low-income patients. To determine whether differences exist among these groups, we initially formulated and tested the following hypotheses using a one-way analysis of variance (ANOVA):

Table 8. Descriptive Statistics Table for Objective 3

STATISTIC	WHITE NON-HISPANIC	BLACK NON-HISPANIC	HISPANIC ANY RACE
COUNT	1465	287	848
MEAN	30.40	33.93	47.57
STD DEV	41.07	30.23	35.37
MIN	1	1	1
25%	6	11.5	18
MEDIAN	18	27	41
75%	38	46.5	69.25
MAX	505	216	226
SUM	44,539	9,737	40,343
RANGE	504	215	225
COV (%)	135.08	89.11	74.35

Hypotheses for ANOVA

Null Hypothesis (H_0): There is no significant difference in accessing healthcare, measured by the frequency of visits, among the racial/ethnic groups in low-income patients.

$$H_0 : \mu_{Black} = \mu_{Hispanic} = \mu_{White}$$

Alternative Hypothesis (H_1): There is a significant difference in accessing healthcare, measured by the frequency of visits, among the racial/ethnic groups in low-income patients.

$$H_1: \exists\, i, j \in \{Black, Hispanic, White\}\ such\ that\ u_i \neq u_j$$

Our dataset included approximately 2,600 patients distributed across the racial/ethnic groups. While our dependent variable (Frequency_of_visit) was continuous, it exhibited substantial positive skew.

Although we initially considered using a one-way ANOVA, we first tested the required assumptions. Shapiro-Wilk tests showed significant departures from normality across all three groups (all p-values < 0.001), violating the assumption of normal distribution. Figures 13 and 14 shows histograms and Q-Q plot confirming this through visible skew and heavy tails.

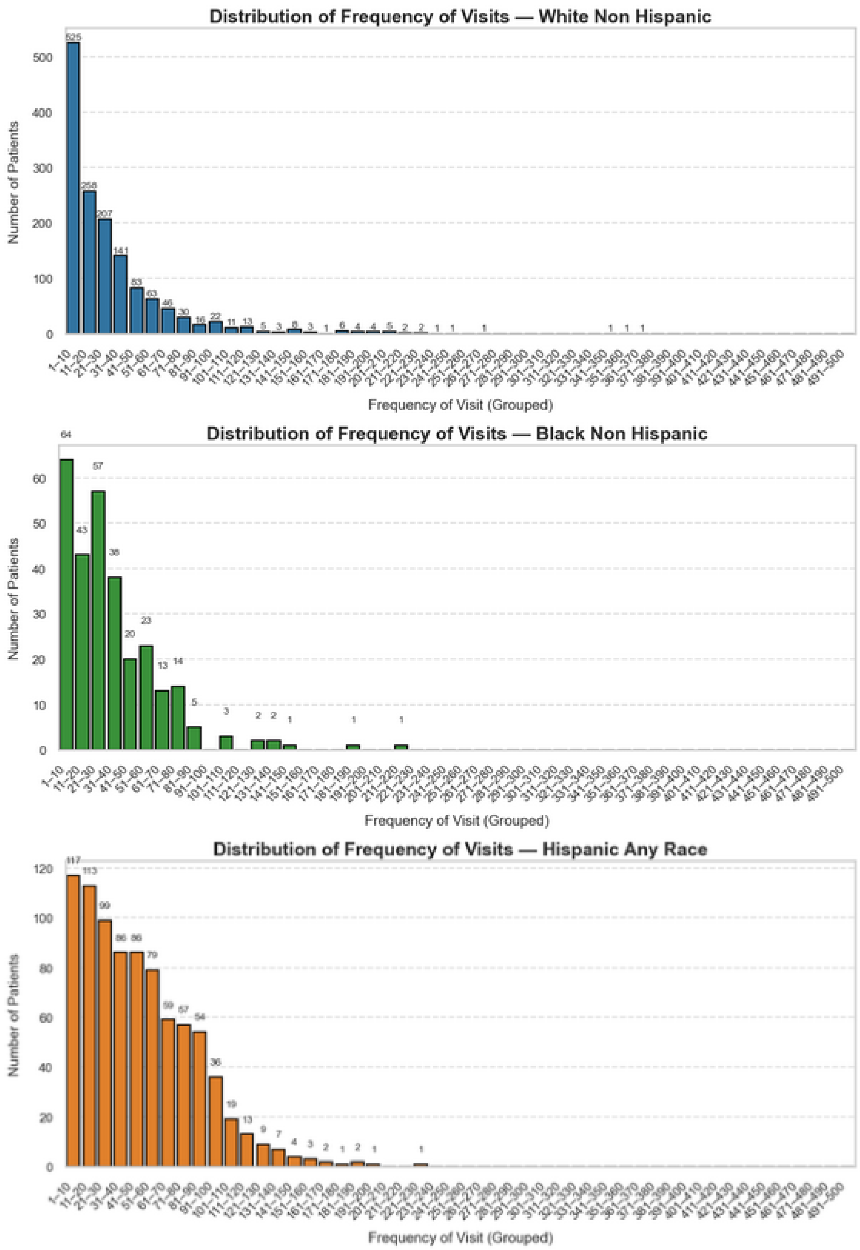

Figure 12 Histogram of Visit Frequency by Race/Ethnicity

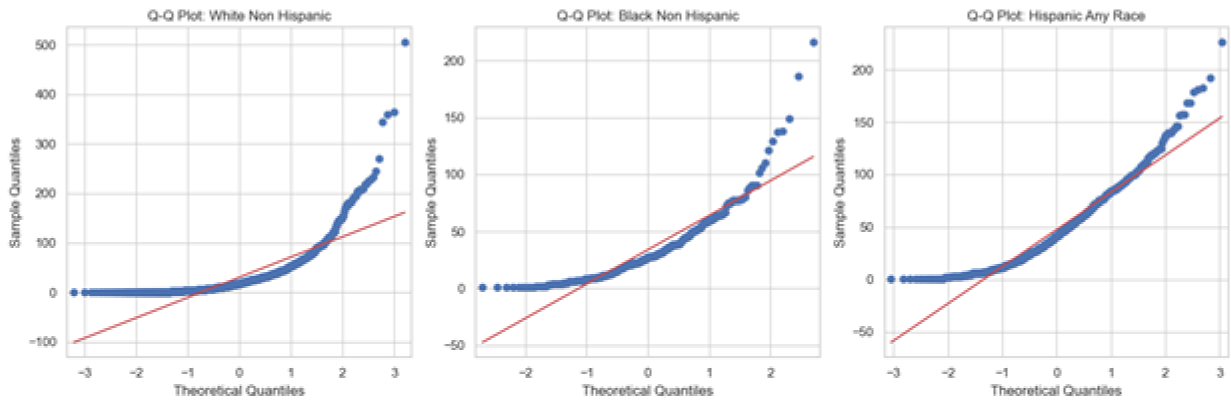

Figure 14 Q-Q Plots for Normality Testing by Group

To address this skewness and improve interpretability, we applied a log transformation: log(Frequency_of_visit + 1). The resulting log-scaled boxplot (Figure 15) showed improved symmetry while retaining between-group differences. Despite transformation, Levene's test indicated heterogeneity of variances ($p = 0.0003$), violating another core ANOVA assumption. Although sample sizes were above 30 per group, ordinarily allowing for parametric inference under the Central Limit Theorem, the data's extreme skew and variance inequality led us to use a more robust non-parametric approach.

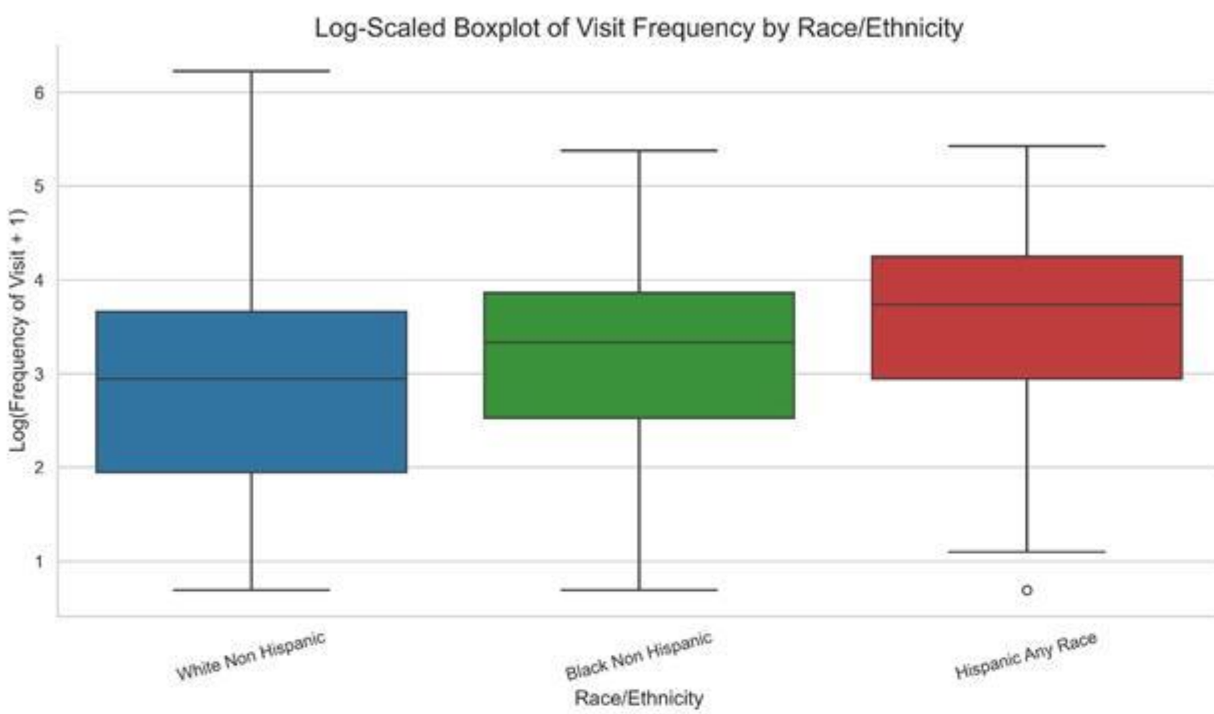

Figure 15 Log-Scaled Boxplot of Visit Frequency by Race/Ethnicity

Figure 16 (ECDF plots) and Figure 17 (violin plots) further illustrated non-normal, skewed distributions, justifying our decision to proceed with non-parametric testing.

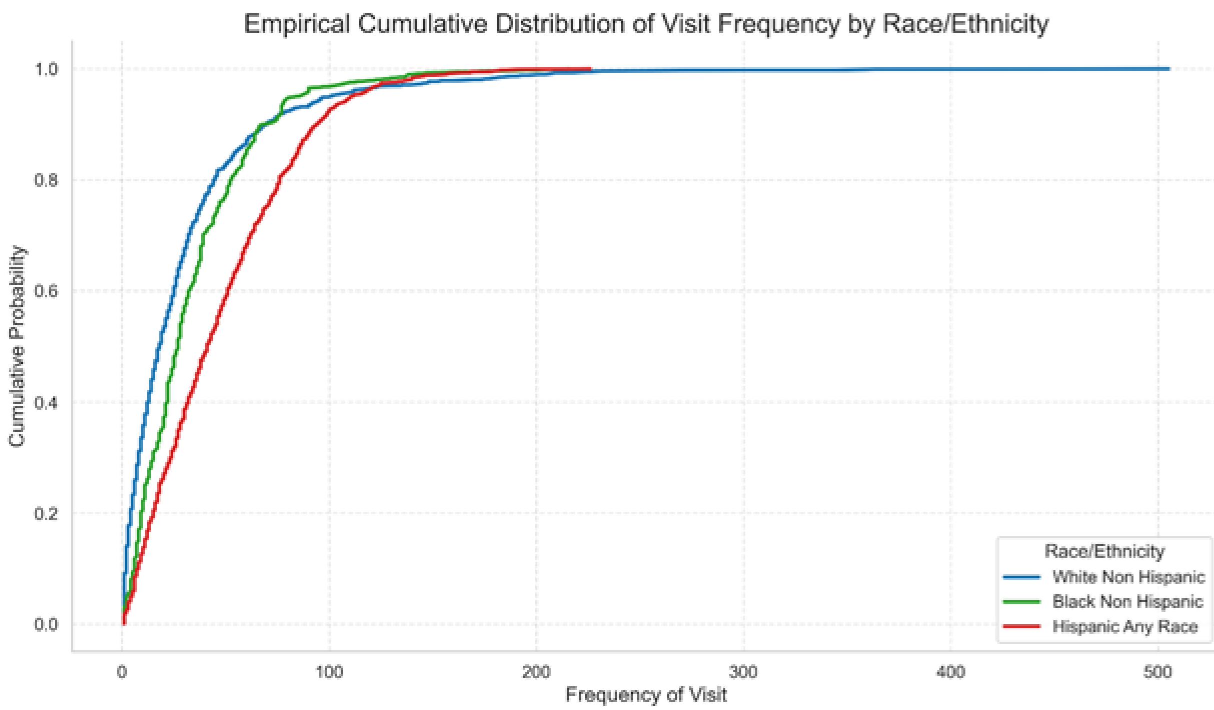

Figure 16 Empirical Cumulative Distribution Function (ECDF) Plot

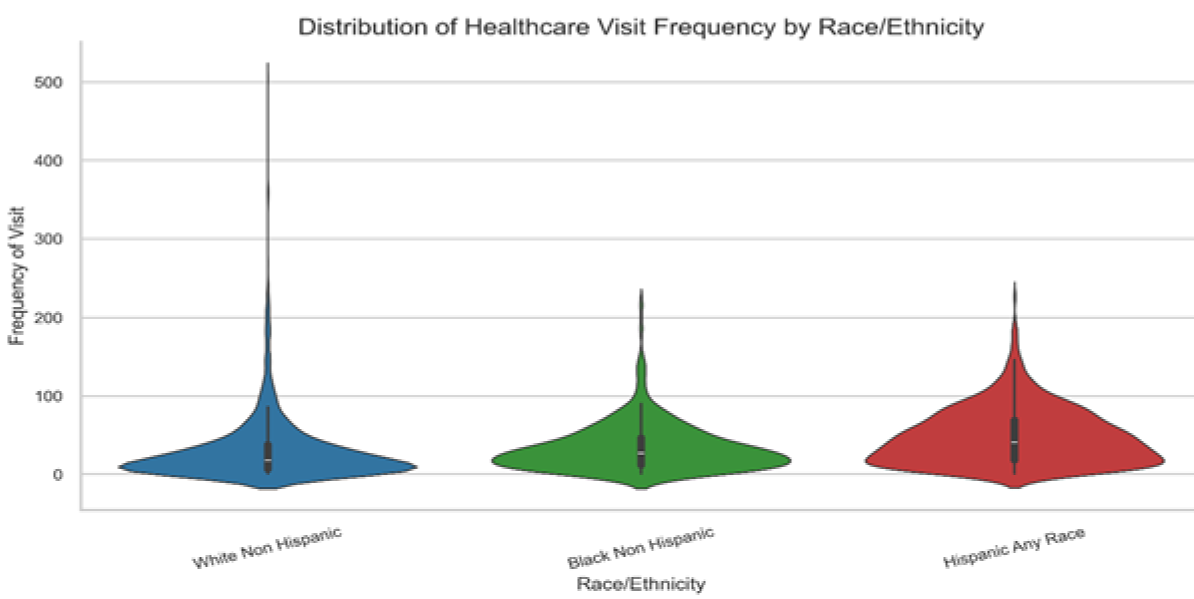

Figure 17 Violin Plot of Visit Frequency by Race/Ethnicity

Revised Hypotheses for Non-Parametric Testing (Kruskal-Wallis Test)

Given the violations of normality and homogeneity of variance assumptions required for parametric tests like ANOVA, we revised our hypotheses to be suitable for a non-parametric approach. The Kruskal-Wallis H test, which compares distributions across independent groups without assuming normality, is appropriate for our positively skewed data on visit frequency.

Null Hypothesis (H_0): There is no significant difference in the distribution of healthcare access (measured by frequency of visits) among the racial/ethnic groups (Black, Hispanic, and White) of low-income patients.

$$H_0 : F_{Black} = F_{Hispanic} = F_{White}$$

(All groups have the same distribution of healthcare access.)

Alternative Hypothesis (H_1): There is a significant difference in the distribution of healthcare access (measured by frequency of visits) among at least one of the racial/ethnic groups (Black, Hispanic, and White) of low-income patients.

$$H_1 : \exists\, i, j \in \{Black, Hispanic, White\} \text{ such that } F_i \neq F_j$$

(At least one group differs in distribution of healthcare access.)

We conducted a Kruskal-Wallis H test, a non-parametric alternative to one-way ANOVA, to test for differences in visit frequency between groups. The result was significant ($H = 238.16$, $p < 0.001$), indicating that at least one group differs in visit frequency distribution. These results support rejecting the null hypothesis that median visit frequencies are equal across groups. To identify specific group differences, we performed Dunn's test with Bonferroni correction. All pairwise comparisons were statistically significant ($p < 0.001$), as summarized in Table 9.

Table 9. Dunn's Post-Hoc Comparison of Visit Frequency by Race/Ethnicity

(Bonferroni-adjusted p-values)

Comparison	p-value	Significance
Black Non Hispanic vs Hispanic Any Race	1.48e-07	Yes
Black Non Hispanic vs White Non Hispanic	1.74e-05	Yes
Hispanic Any Race vs White Non Hispanic	4.04e-53	Yes

Objective 3 Interpretation:

Our analysis revealed that Hispanic patients utilized healthcare services more frequently on average than both White Non-Hispanic and Black Non-Hispanic patients. While White Non-Hispanic patients accounted for the highest total number of visits (44,539), they also made up the largest group in the dataset (n = 1,465). After adjusting for group size, we found that Hispanic patients (n = 848) averaged 47.6 visits per patient, followed by Black Non-Hispanic (33.9 visits, n = 287) and White Non-Hispanic (30.4 visits per patient). Figures 18 and 19 visually contrast total vs. average per-patient visit frequency. These figures emphasize the importance of considering per-capita metrics in understanding healthcare access.

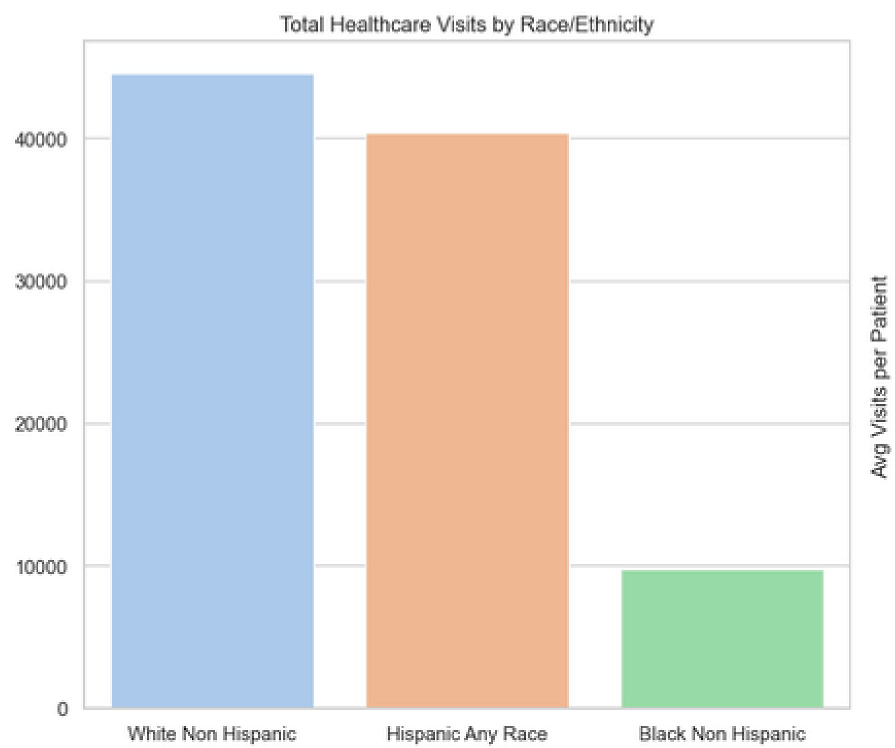

Figure 18 Total Sum Healthcare Visits by Race/Ethnicity

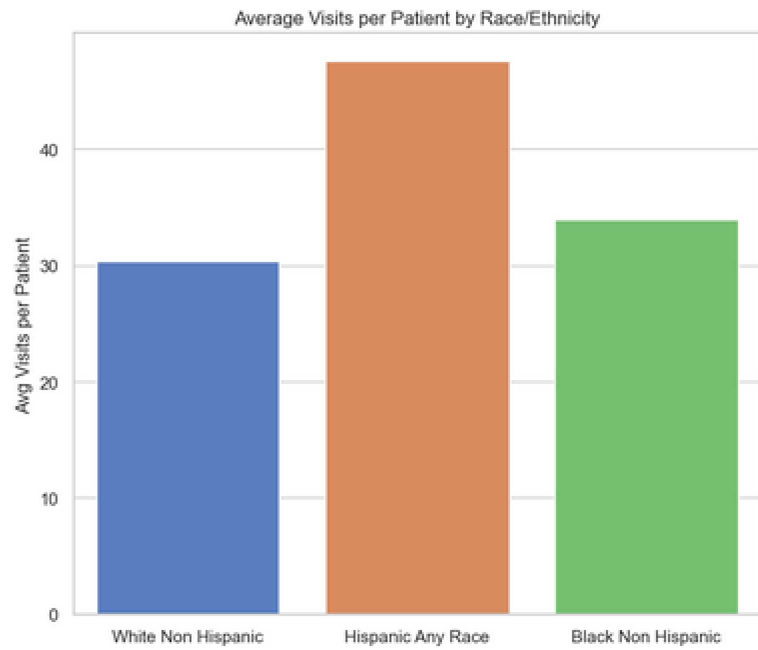

Figure 19 Average Visits per Patient by Race/Ethnicity

We conclude that significant differences in healthcare utilization exist among racial/ethnic groups within the low-income population. The analysis supports rejection of the null hypothesis. In particular, Hispanic patients show the highest individual engagement with healthcare, despite not having the highest total visit count. These findings underscore the need to further explore sociocultural and systemic factors influencing access and behavior within healthcare systems. Addressing these disparities is critical to advancing equitable health outcomes across all populations.