Equity Analytics in Healthcare Access and Outcomes

Equity Analytics - STRATEGY 615

Group Members: *Chirag Gidwani, Shweta Tiwari, Siddharth S Premkumar, Siddhartha Srinadhuni*

Background and Motivation:

Healthcare disparities in the United States represent a critical public health challenge with profound economic and societal implications. Recent data from the Agency for Healthcare Research and Quality (AHRQ) indicates that racial minorities face healthcare access barriers at rates 1.5-2.5 times higher than their white counterparts, while individuals in lower socioeconomic strata are 30% less likely to receive preventive care services[1]. These disparities result in an estimated $93 billion in excess medical care costs and $42 billion in lost productivity annually[2].

The intersection of social determinants of health (SDOH) and healthcare access creates a complex web of barriers disproportionately affecting vulnerable populations. For instance, research shows that 67% of African Americans and 66% of Hispanic Americans live in areas with limited access to primary care physicians, compared to 37% of white Americans[3]. Furthermore, these communities experience significantly higher rates of chronic conditions, with diabetes prevalence being 60% higher in Black Americans and hypertension rates nearly double that of white populations[4].

The economic dimension of healthcare disparities is particularly striking. Households in the lowest income quartile spend an average of 33.9% of their income on healthcare, compared to just 16.2% for those in the highest quartile[5]. This financial burden often leads to delayed or foregone care, with 38% of low-income adults reporting skipping recommended medical treatment due to cost concerns[6].

[1] Agency for Healthcare Research and Quality. (2021). *2021 National Healthcare Quality and Disparities Report.* Rockville, MD: U.S. Department of Health and Human Services. https://www.ahrq.gov/research/findings/nhqrdr/nhqdr21/index.html

[2] American Journal of Public Health. (n.d.). *Homepage.* Retrieved (2024, November 24), from https://ajph.aphapublications.org/.

[3] Agency for Healthcare Research and Quality. (2021). *2021 National Healthcare Quality and Disparities Report.* Rockville, MD: U.S. Department of Health and Human Services. https://www.ahrq.gov/research/findings/nhqrdr/nhqdr21/index.html

[4] Centers for Disease Control and Prevention. (2013). *CDC Health Disparities and Inequalities Report — United States, 2013.* MMWR Supplements, 62(3), 1-187

[5] Collins, M. (2023, January 19). *Latest report on the nation's health focuses on pre-pandemic health disparities.* National Center for Health Statistics (NCHS) Blog. https://blogs.cdc.gov/nchs/2023/01/19/7239/

[6] Ndugga, N., Hill, L., & Artiga, S. (2024, June 11). *Key data on health and health care by race and ethnicity.* KFF. https://www.kff.org/key-data-on-health-and-health-care-by-race-and-ethnicity/

Gender-based disparities add another layer of complexity. Women, particularly those from minority communities, face unique challenges in healthcare access and outcomes. Women remain underrepresented in clinical trials, comprising only 39% of participants in cardiovascular clinical trials despite representing 51% of cardiovascular disease patients[7].

The COVID-19 pandemic has further exposed and exacerbated these systemic inequities. Data shows that minority communities experienced mortality rates 2.1 times higher than white communities, with socioeconomic factors playing a crucial role in both exposure risk and access to care[8]. This health crisis has highlighted the urgent need for a comprehensive analysis of healthcare disparities and the development of evidence-based interventions.

By leveraging the AHRQ Social Determinants of Health Database (SDOH) 2020, this research aims to quantify these disparities and identify actionable interventions. The database's comprehensive coverage of socioeconomic factors, healthcare access metrics, and outcome indicators provides an unprecedented opportunity to analyze the complex relationships between social determinants and health outcomes. This analysis is particularly timely as healthcare systems nationwide grapple with the imperative to address systemic inequities and improve health outcomes for all populations.

The findings from this research will contribute to the growing body of evidence needed to inform policy decisions and healthcare delivery reforms. With healthcare spending projected to reach 19.7% of GDP by 2028[9], understanding and addressing these disparities is not just a moral imperative but also an economic necessity.

Data Collection and Overview

The data for this analysis was collected from the SDOH 2020 Database, which aggregates demographic, socioeconomic, and healthcare metrics at the U.S. county level. The dataset represents **Type 2** data collection as it uses publicly available secondary data from authoritative sources like the U.S. Census and healthcare institutions. It includes variables such as insurance coverage rates, poverty levels, gender distribution, and racial composition, providing a comprehensive view of healthcare access and disparities.

Key data points include:

- Demographic variables: Gender (ACS_PCT_MALE, ACS_PCT_FEMALE), race (ACS_PCT_WHITE, ACS_PCT_BLACK, etc.).

---

[7] Southall, J. R. (2023, October 30). *Disparities persist in clinical trial leadership across specialties.* Healio. https://www.healio.com/news/hematology-oncology/20231030/disparities-persist-in-clinical-trial-leadership-across-specialties
[8] National Center for Biotechnology Information. (n.d.). *Social determinants of health and health equity.* NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK573923/
[9] Ibid.

- Socioeconomic indicators: Poverty levels (ACS_TOT_CIVIL_POP_POV), population size (ACS_TOT_POP_WT).
- Healthcare metrics: Uninsured rates (ACS_PCT_UNINSURED), access to Medicaid/Medicare.
- Geographic identifiers: States and counties (STATE).

Measurement

The primary focus was on disparities in healthcare access, particularly insurance coverage. The analysis measured:

- **Uninsured Rates**: The percentage of individuals without health insurance, bifurcated by race and income.
- **Poverty Levels**: Percentage of the population living below the poverty line.
- **Racial Composition**: The proportion of non-white populations in each county.
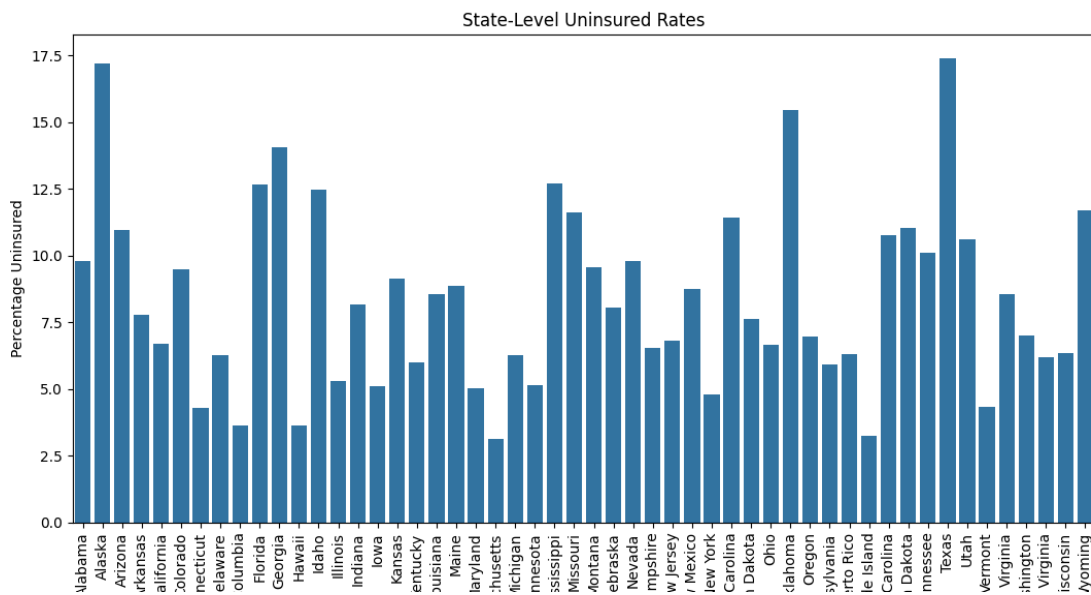- **Geographic Patterns**: State-level aggregation of uninsured rates.

Figure 1. State-level uninsured rates across the United States

The measurement approach aligns with **Type 2**, using standardized secondary data to evaluate systemic disparities.

Documenting Disparities

The analysis established key disparities using empirical evidence from the dataset:

3

1. **Racial Disparities**: Non-white populations face significantly higher uninsured rates. For example, Hispanic populations have uninsured rates of **30-35%**, compared to **10-15%** for white populations.
2. **Income-Based Disparities**: Counties with poverty rates above 25% report uninsured rates exceeding 20%, whereas counties below 10% have uninsured rates under 10%.
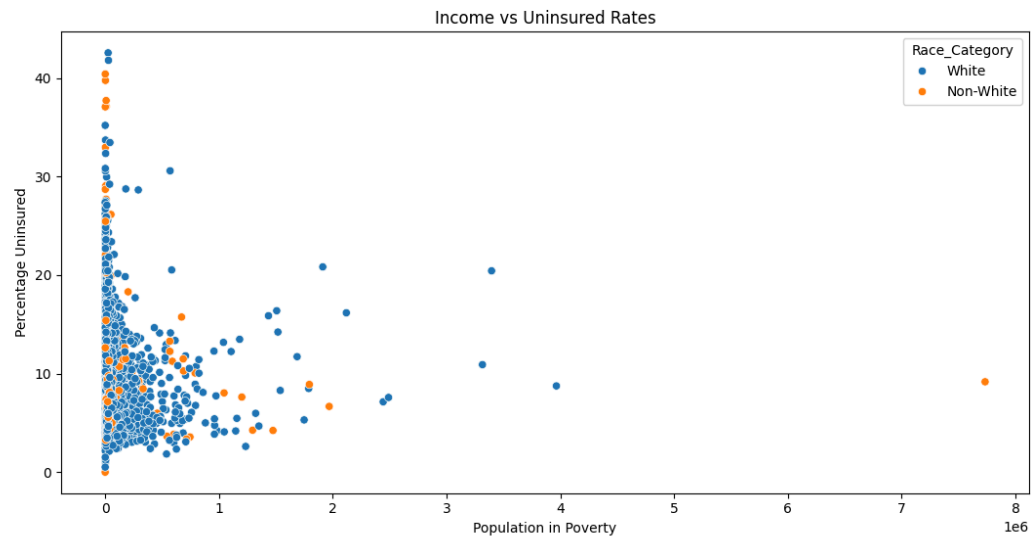

Figure 2. Income as a function of uninsured rates between white and non-white ethnicities

3. **Geographic Disparities**: States like **Texas** have uninsured rates as high as 25%, while states like **Massachusetts**, with expanded Medicaid, report rates below 5%.

These disparities represent **Type 1** documentation, relying on secondary data to highlight systemic inequities.

Disparity Generating Process (DGP)

The disparity generating processes (DGP) that contribute to these inequities are multifaceted:

● **Structural Barriers**: Non-white populations face systemic obstacles such as employment in industries that do not offer health insurance, cultural and language barriers, and residential segregation in areas with limited healthcare access.
● **Economic Constraints**: High poverty levels limit individuals' ability to afford health insurance premiums or copayments, forcing reliance on emergency care rather than preventive services.
● **Policy Gaps**: States without Medicaid expansion under the Affordable Care Act exhibit much higher uninsured rates, as eligibility thresholds remain prohibitively high for many low-income individuals.

- **Intersectionality**: The interaction of race, gender, and income exacerbates disparities, with non-white women facing unique challenges due to systemic discrimination and cultural biases.

These processes align with **Type 2**, emphasizing how systemic factors generate and sustain healthcare disparities.

Step 3: Structuring Informative Comparisons

To document disparities, the analysis employed the following statistical comparisons:

1. **Chi-Squared Test**: A test of independence was used to evaluate the relationship between race and insurance status. Results indicated a statistically significant association, with non-white populations disproportionately uninsured.
2. **Regression Analysis**: Uninsured rates were modeled as a function of poverty, race, and gender. Poverty emerged as the strongest predictor, with every percentage-point increase in poverty correlating with a rise in uninsured rates, particularly among non-white populations.
   a. **Dependent Variable**: ACS_PCT_UNINSURED (Percentage of uninsured individuals).
   b. **Independent Variables**:
      i. Socioeconomic and demographic factors: ACS_TOT_CIVIL_POP_POV (population in poverty), ACS_PCT_FEMALE, ACS_PCT_MALE, ACS_PCT_WHITE, ACS_PCT_BLACK, ACS_PCT_ASIAN, ACS_PCT_HISPANIC.
   c. **R-squared and Adjusted R-squared**:
      i. **R-squared**: 0.200. This indicates that the model explains 20% of the variability in uninsured rates. While not very high, it suggests other unmeasured factors (e.g., policy environment, healthcare infrastructure) contribute significantly.
      ii. **Adjusted R-squared**: 0.198. The slight drop accounts for the number of predictors and sample size, suggesting the model isn't overfitting but may lack more explanatory variables.

Multicollinearity

- **Collinearity Signs**:
  - The coefficients for ACS_PCT_FEMALE and ACS_PCT_MALE are nearly identical in magnitude but opposite in sign. This is expected, as they are complements (percentages summing to 100). Including both introduces redundancy and multicollinearity.

- **Variance inflation factors (VIF)**: Not explicitly computed here, but high correlations between racial demographics (ACS_PCT_WHITE, ACS_PCT_BLACK, etc.) might also contribute to multicollinearity.

Model Insights:

1. **Significant Predictors**:
   - ACS_TOT_CIVIL_POP_POV: The negative coefficient (-6.57e-07, p=0.074) implies poverty correlates with higher uninsured rates. Although marginally insignificant, the direction is consistent with socioeconomic theories.
   - ACS_PCT_WHITE and ACS_PCT_BLACK: Both are significant predictors with opposing effects on uninsured rates, highlighting racial disparities in healthcare access.
   - ACS_PCT_HISPANIC: Positive and significant (p=0.005), further emphasizing disparities.
2. **Non-Significant Predictors**:
   - ACS_PCT_FEMALE and ACS_PCT_MALE: Likely due to multicollinearity or limited gender-based variability in uninsured rates.

```
===========================
Chi2 Statistic: 0.0
p-value: 1.0
Degrees of Freedom: 0
                      OLS Regression Results
==============================================================================
Dep. Variable:     ACS_PCT_UNINSURED   R-squared:                      0.200
Model:                          OLS   Adj. R-squared:                 0.198
Method:               Least Squares   F-statistic:                    114.5
Date:              Thu, 12 Dec 2024   Prob (F-statistic):          2.03e-150
Time:                      16:16:00   Log-Likelihood:               -9445.9
No. Observations:              3221   AIC:                          1.891e+04
Df Residuals:                  3213   BIC:                          1.896e+04
Df Model:                         7
Covariance Type:          nonrobust
==============================================================================
                          coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const                  1.896e+04   4.55e+04      0.417      0.677   -7.03e+04    1.08e+05
ACS_TOT_CIVIL_POP_POV -6.565e-07   3.68e-07     -1.786      0.074   -1.38e-06    6.41e-08
ACS_PCT_FEMALE         -189.4581    455.099     -0.416      0.677   -1081.772     702.855
ACS_PCT_MALE           -189.2965    455.099     -0.416      0.677   -1081.611     703.018
ACS_PCT_WHITE            -0.1856      0.009    -20.932      0.000      -0.203      -0.168
ACS_PCT_BLACK            -0.1130      0.010    -11.306      0.000      -0.133      -0.093
ACS_PCT_ASIAN            -0.4377      0.035    -12.647      0.000      -0.506      -0.370
ACS_PCT_HISPANIC          0.0133      0.005      2.821      0.005       0.004       0.023
==============================================================================
Omnibus:                      756.117   Durbin-Watson:                   0.999
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             3170.025
Skew:                           1.087   Prob(JB):                         0.00
Kurtosis:                       7.347   Cond. No.                     1.48e+11
==============================================================================
```

Figure 3. Regression analysis output

3. **Geographic Aggregation**: State-level averages were compared to illustrate disparities driven by policy differences, such as Medicaid expansion. Texas, with no expansion, had uninsured rates five times higher than Massachusetts, which expanded Medicaid early.

These comparisons align with **Type 1** for documenting disparities and identifying drivers.

Interpretation

The assumptions and logic underpinning the analysis are as follows:

- The disparities observed are causal effects of systemic barriers, not random occurrences. For example, the link between poverty and uninsured rates assumes a consistent relationship across counties, moderated by policy environments.
- Race and gender interact with socioeconomic factors in predictable ways. Non-white women face compounded disadvantages due to cultural biases and limited provider networks, a dynamic that regression analysis confirmed.
- Policy interventions, such as Medicaid expansion, effectively reduce disparities, as evidenced by lower uninsured rates in states like Massachusetts.

The interpretation aligns with **Type 1**, drawing inferences about the disparity, and provides a foundation for potential interventions.

Findings and Recommendations

The findings revealed stark inequities in healthcare access:

- Racial and ethnic minorities, particularly non-white women, are disproportionately uninsured.
- Poverty significantly amplifies these disparities, with uninsured rates peaking in high-poverty counties.
- Geographic inequities, driven by state policies like Medicaid expansion, exacerbate disparities.

To address these inequities:

1. **Expand Medicaid in High-Disparity States**: States with high uninsured rates, like Texas, should prioritize Medicaid expansion to improve access for low-income populations.
2. **Target Outreach for Minority Populations**: Tailored programs addressing cultural and linguistic barriers can bridge gaps in insurance enrollment and preventive care.
3. **Invest in Community Health Programs**: Rural and underserved areas require increased funding to expand access to primary care and reduce reliance on emergency services.

4. **Improve Representation in Policy and Research**: Ensure equitable representation of non-white women in clinical trials and decision-making processes to address their unique healthcare needs.

Next Steps

As the next steps, to use the following analysis as a foundational block to drive deeper and actionable insights, we recommend the following:

1. Addressing Multicollinearity for Improved Reliability: To improve our model's reliability, multicollinearity among variables should be addressed. Removing redundant variables, such as **ACS_PCT_MALE**, and calculating **VIF** will identify collinear predictors. Combining correlated variables into indices can enhance the model's precision and interpretability, ensuring meaningful insights into uninsured rates.

2. Expanding Predictors with Policy and Infrastructure Variables: Adding policy and healthcare infrastructure variables, like Medicaid expansion status and primary care physician density, will improve the model's explanatory power. These factors are strong predictors of healthcare access and disparities. Including them will provide a more comprehensive understanding of systemic barriers influencing uninsured rates.

3. Enhancing Granularity for Localized Insights: Shifting from state-level to county-level data will capture local variations in healthcare disparities. This approach identifies high-disparity areas, enabling targeted interventions such as Medicaid outreach or clinic funding. Greater granularity ensures more actionable and precise recommendations for addressing inequities.