OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

# Socioeconomic Status and Patient Experiences at Dental Clinics: *An Analysis of Yelp Reviews*

Julie Haeri Kim
julierosekim0721@gmail.com

## ABSTRACT

Patients' experiences in dental clinics may be influenced by socioeconomic conditions that shape access to care, communication, and patient satisfaction. This study examined associations between community-level socioeconomic characteristics and Yelp-reported satisfaction using 35,132 reviews from 2,498 dental businesses across 13 U.S. states. Dental businesses were identified through dental-related Yelp categories, and Census-derived socioeconomic and demographic characteristics were linked to the geographic areas surrounding the clinics. Satisfaction was defined as a rating of 4 or 5 stars, while ratings of 1–3 stars were grouped as lower or mixed satisfaction.

Overall, 67.9% of reviews were classified as satisfied. Logistic regression analyses using clinic-clustered standard errors showed that area-level educational attainment and several demographic composition variables were associated with review satisfaction. Median household income was not statistically significant in the socioeconomic and demographic model ($p = 0.065$) but showed a small negative association after adjustment for review and business metadata ($p = 0.039$). Exploratory machine-learning models demonstrated that review text provided substantially greater predictive signal than socioeconomic and demographic variables alone. These findings identify associations within online review data rather than causal effects or individual-level socioeconomic differences. Because Yelp reviews represent voluntarily submitted experiences, favorable online ratings cannot establish equitable access to dental care.

**Keywords:** *socioeconomic status, patient experience, dental care, Yelp reviews, oral health disparities, healthcare equity*

## INTRODUCTION

Oral health is an important element of overall health and well-being, yet access to dental care stays uneven across the United States. Going to the dentist is more than just receiving treatment. How patients are treated, communication with providers, cost, and the overall experience all influence a patient's comfort returning for future appointments. For many people, however, external factors outside of the clinic, including where they live and their socioeconomic background, shape these experiences. People in lower-income communities regularly face obstacles such as financial costs, limited insurance, and fewer providers, which makes accessing consistent dental care difficult.

As more patients share their experiences online, platforms such as Yelp have become useful sources for examining how individuals perceive the care they receive. Yelp provides both quantitative information,

including 1- to 5-star ratings, and qualitative information through written reviews. Its publicly available dataset includes dental practices across multiple U.S. regions, allowing researchers to investigate associations between community characteristics and online patient satisfaction [5]. In this study, dental businesses were identified using dental-related Yelp categories rather than purposive selection based on clinic ratings. However, because the dataset is limited to businesses and reviewers represented on Yelp, it does not constitute a statistically representative sample of dental patients or practices nationwide.

Other studies have shown that socioeconomic status plays a major role in healthcare and access to dental care. For example, a systematic review found that children and adolescents from lower socioeconomic backgrounds were less likely to use dental services [4]. While these differences are well documented, less is known about whether they are reflected in patients' experiences with dental clinics. Understanding these differences can provide a clearer picture of oral health care and identify areas of patient-centered care. However, most of this research focuses on access to care, utilization, and clinical outcomes rather than on patients' publicly reported experiences with dental clinics.

This study investigates whether community-level socioeconomic characteristics are associated with patient satisfaction as reflected in Yelp reviews of dental practices across the United States. Specifically, it examines relationships between review satisfaction and the median household income, poverty rate, educational attainment, unemployment rate, and racial and ethnic composition of the geographic areas surrounding dental clinics. We hypothesized that socioeconomic characteristics would be associated with differences in the likelihood of receiving positive reviews. Logistic regression analyses were used to examine these associations, while exploratory machine-learning models evaluated the predictive contributions of review text, business metadata, and area-level demographic variables. Importantly, the socioeconomic measures describe clinic locations rather than the individual circumstances of reviewers, who may live outside the surrounding communities.

## LITERATURE REVIEW

Previous studies have shown that social and economic factors have a significant influence on oral health and on access to dental care. People with a low socioeconomic status experience greater difficulties in accessing preventive care, tend to delay treatment, and have worse dental health outcomes than those from higher socioeconomic groups. These inequalities are usually caused because of factors including inadequate insurance coverage, the high cost of treatment, problems with transportation, a smaller availability of dental services, and the difficulty of taking time off work to attend appointments. Consequently, individuals in areas of lower socioeconomic status may postpone seeing the dentist, which may affect oral health and their overall satisfaction with the dental services provided. In the end, these disparities impact how often people use dental services and their general experiences of receiving dental care.

The existence of web platforms for the purpose of reviewing services has led researchers to look at the possibility of using patient-written online reviews as a means of assessing reported experiences of dental care. In a recent study involving more than 4,700 reviews posted on Yelp and Google by patients at 65

dental schools in the United States, it was found that patients often referred to non-clinical aspects of the care they received, such as the pleasantness of the staff, the clarity of the financial information, the organisation of work at the clinic, and the communication [1]. Similar non-clinical concerns appear in research on negative reviews in healthcare and dentistry [6].

In a similar study, Larsen and his colleagues examined more than 11,000 online reviews of dentists in Norway in order to find out what factors influence patient satisfaction. They discovered that clear communication, reassurance, professional advice, and positive interactions with the dentist were strongly associated with higher levels of satisfaction, whereas negative treatment experiences had a lowering effect on it [2]. The study also showed that online reviews can give a useful insight into how patients perceive and assess dental care in real-life situations. Yet there are limitations to online reviews since the people who review on Yelp may not be representative of the general population, and those who have either very positive or very negative experiences are more likely to leave a review. Nevertheless, online reviews do provide useful information regarding patient experiences on a large scale.

Although earlier research has shown both the connection between socioeconomic status and oral health and the value of online reviews for recording patients' experiences, there have been fewer studies that have combined these two areas. Even though previous studies have pointed out socioeconomic discrepancies in access to dental care, there is not much known about whether these disparities also affect patients' experiences after they have received care. Specifically, it is not well understood whether the experiences mentioned in Yelp reviews vary according to the socioeconomic characteristics of the communities in which dental clinics are situated. The present study attempts to fill this gap via comparing the Yelp reviews of dental clinics in high-, middle-, and low-income areas of America in order to find out whether socioeconomic status is associated with differences in patient satisfaction.

## METHODS

This study used the Yelp Open Dataset to examine associations between community-level socioeconomic characteristics and satisfaction expressed in online dental reviews. Dental businesses were identified by restricting Yelp business records to U.S. states and filtering establishments according to dental-related terms in their business categories. Duplicate business records were removed using unique business identifiers. Reviews associated with these businesses were extracted and merged with business-level information, including clinic location, overall star rating, and review count.

The resulting dataset contained 35,132 Yelp reviews from 2,498 dental businesses across 13 U.S. states and 294 distinct city labels in the business records. Because the dataset was derived from Yelp rather than a probability-based sample of dental practices or patients, its geographic distribution was uneven and its findings cannot be generalized to all dental patients in the United States.

Community-level socioeconomic characteristics were incorporated using Census-derived data linked to the geographic areas surrounding dental businesses. These variables included median household income, poverty rate, educational attainment, unemployment rate, and racial and ethnic composition. Income bands were constructed using the median household income associated with each clinic's ZIP Code

Tabulation Area (ZCTA), rather than city-level income categories. These geographic measures describe the socioeconomic characteristics of clinic areas, not the individual incomes or demographic identities of reviewers.

To assess data completeness before modeling, missing values were examined across Census-derived socioeconomic and demographic variables (Figure 1). Approximately 0.61% of reviews had missing values in the displayed Census-derived fields. Reviews with missing predictors were excluded from the regression analyses through complete-case deletion, resulting in 34,919 observations in each model. To account for the clustering of multiple reviews within dental practices, cluster-robust standard errors were calculated using business identifiers.

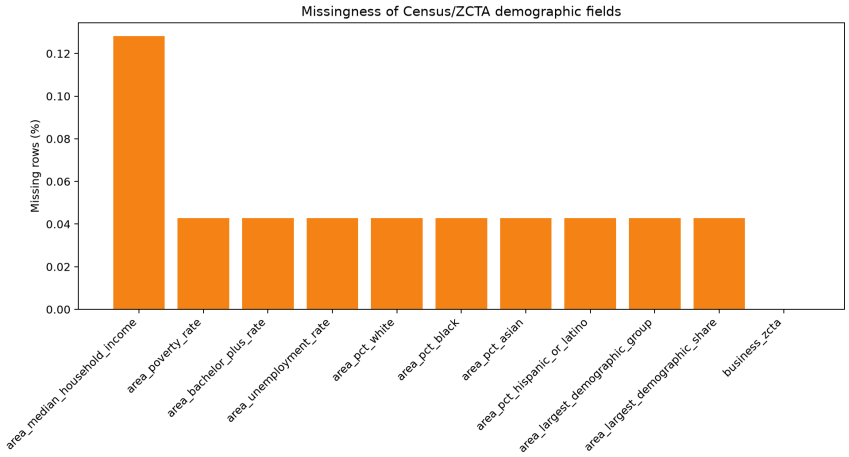



*Figure 1. Percentage of reviews with missing Census-derived socioeconomic and demographic information. Approximately 0.61% of reviews (213 of 35,132) lacked these variables, while 99.39% had complete data for the displayed Census-derived measures. Reviews with missing predictors were excluded from the complete-case regression analyses.*

The dental reviews gathered from different business locations being geographically clustered shows how the selected clinics are distributed over space. The data collected were analysed with the help of Python in order to detect patterns in patient experiences. Descriptive statistics were applied in order to compare the Yelp ratings and the number of reviews in high-, middle- and low-income areas. Furthermore, logistic regression analyses were carried out to see if socioeconomic characteristics were linked to patient satisfaction while taking into account the other variables in the dataset.

## RESULTS

The dataset consisted of 35,132 dental reviews from Yelp, the reviews having been grouped according to their original star ratings (from 1 to 5 stars) in order to display the general distribution of patient

sentiment in the dataset; most of the reviews were given five stars, whereas the lower ratings were seen less often.

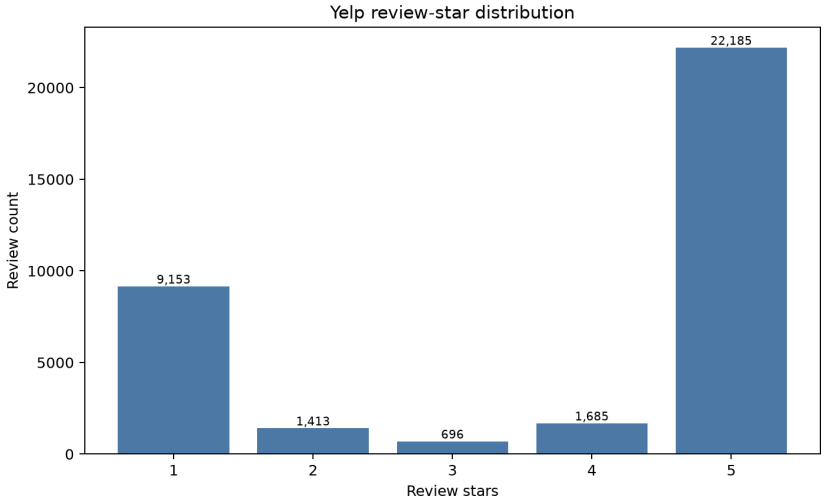



*Figure 2. Yelp review star-frequency distribution showing the breakdown of 1- to 5-star ratings across the 35,132 collected reviews.*

When binarized into target satisfaction classes, 23,870 reviews (67.9%) were classified as satisfied (4–5 stars), and 11,262 reviews (32.1%) were classified as lower or mixed satisfaction (1–3 stars).

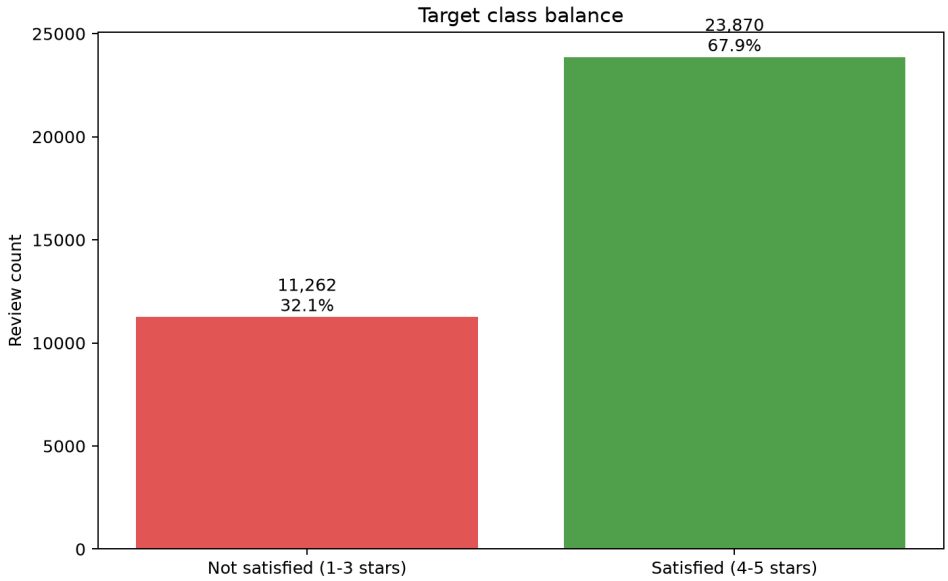



*Figure 3. Target class balance showing satisfied (4–5 stars) versus lower or mixed ratings (1–3 stars). The figure's "not satisfied" label includes neutral or mixed 3-star reviews.*

Furthermore, descriptive analysis across area income bands indicated variations in satisfaction percentages based on community-level wealth.

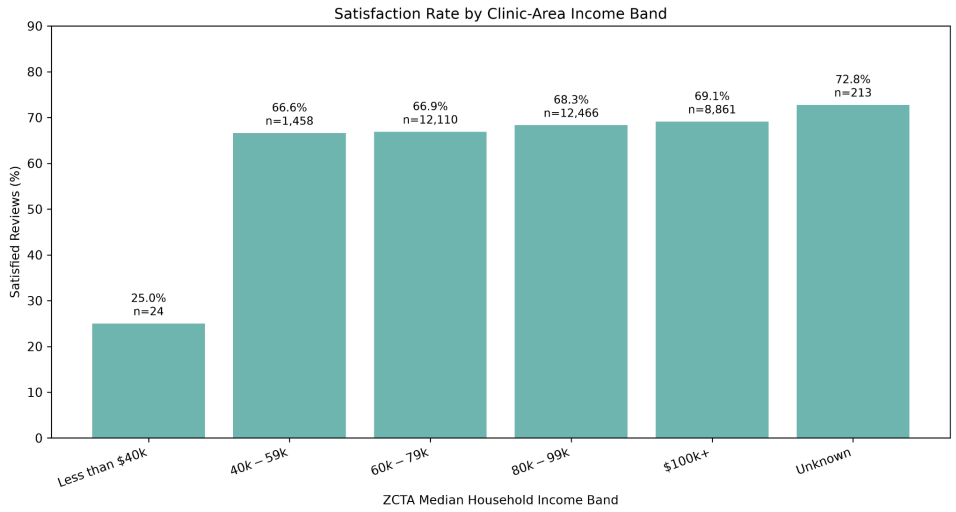



*Figure 4. Satisfaction rates and review counts by clinic-area ZCTA median household income band. Satisfaction was defined as a Yelp rating of 4–5 stars. The lowest-income category (below $40,000) contained only 24 reviews, so its 25.0% satisfaction rate should be interpreted cautiously.*

The logistic regression analysis examined associations between community-level socioeconomic characteristics and individual review satisfaction. Model 1 included area-level income, poverty, education, unemployment, and racial and ethnic composition. Model 2 additionally adjusted for business review count, review length, and review year. Both models used clinic-clustered standard errors to account for the grouping of reviews within dental practices. Median household income was not statistically significant in Model 1 ($\beta = -3.72 \times 10^{-6}$, $p = 0.065$), but was negatively associated with satisfaction in Model 2 ($\beta = -4.128 \times 10^{-6}$, $p = 0.039$). The proportion of residents with a bachelor's degree was positively associated with satisfaction in both models (both $p < 0.001$). Several racial and ethnic composition variables were negatively associated with satisfaction, although the association for the proportion of Asian residents was not statistically significant in Model 2 ($p = 0.080$). Poverty and unemployment rates were not statistically significant in either model. These findings describe associations rather than causal effects, and the community-level variables do not measure the socioeconomic circumstances of individual reviewers.

*Table 1. Logistic Regression Results Predicting Review Satisfaction*

| **Predictor** | **Model 1 β** | **Model 1 p-value** | **Model 2 β** | **Model 2 p-value** |
|---|---|---|---|---|
| Median household income | -3.72e-6 | 0.065 | -4.128e-6 | 0.039 |
| Poverty rate | 0.0121 | 0.274 | 0.0147 | 0.186 |
| Bachelor’s degree rate | 0.0147 | <0.001 | 0.0148 | <0.001 |

| Predictor | Model 1 β | Model 1 p-value | Model 2 β | Model 2 p-value |
|---|---|---|---|---|
| Unemployment rate | 0.0085 | 0.737 | 0.0045 | 0.862 |
| % White | -0.0433 | 0.009 | -0.0428 | 0.009 |
| % Black | -0.0475 | 0.005 | -0.0465 | 0.005 |
| % Asian | -0.0395 | 0.046 | -0.0348 | 0.080 |
| % Hispanic/Latino | -0.0429 | 0.018 | -0.0417 | 0.021 |
| Business review count | – | – | 7.904e-5 | 0.947 |
| Review length (words) | – | – | -0.0070 | < 0.001 |
| Review year | – | – | -0.0565 | < 0.001 |

Note. The dependent variable is review satisfaction, coded as 1 for reviews rated 4-5 stars and 0 for reviews rated 1-3 stars. Model 1 includes area-level socioeconomic and demographic variables. Model 2 additionally adjusts for business review count, review length, and review year. Both models use 34,919 complete observations and standard errors clustered by dental clinics. Coefficients are log-odds coefficients, not percentage changes. Statistical significance is assessed using two-sided tests.

Machine-learning models were evaluated using a clinic-grouped training and testing split to reduce the possibility of information leakage between reviews from the same dental practice. The training set contained 28,044 reviews from 1,998 clinics, while the held-out test set contained 7,088 reviews from 500 clinics, with no clinic overlap between the two sets. The highest-performing model combined review text, business metadata, and area-level demographic variables, achieving a balanced accuracy of 0.9542. A text-only TF-IDF logistic regression model performed nearly identically (0.9538), suggesting that review text contributed most of the predictive signal. In comparison, metadata-only and demographics-only models achieved balanced accuracies of 0.6632 and 0.5297, respectively. Adding demographic variables to business metadata did not improve predictive performance, resulting in a balanced accuracy of 0.6604. These exploratory results describe the models' ability to classify Yelp review ratings and do not establish causal relationships between socioeconomic conditions and patient satisfaction.

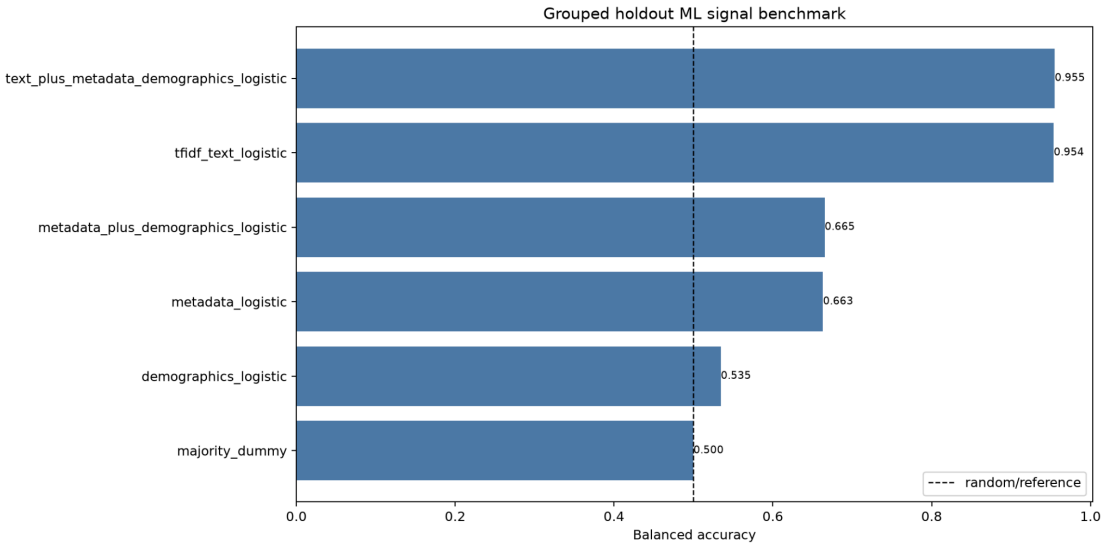



*Figure 5. Exploratory grouped holdout ML signal benchmark comparing balanced accuracies among different feature configurations and models.*

The confusion matrix for the best-performing model further illustrates classification performance among satisfied and lower- or mixed-satisfaction classes. Because the classification models predicted star-rating categories using review text and associated metadata, their performance reflects the predictability of Yelp ratings rather than an independent assessment of clinical care quality. Model comparisons were exploratory, and the best-performing configuration was identified using the reported holdout set.

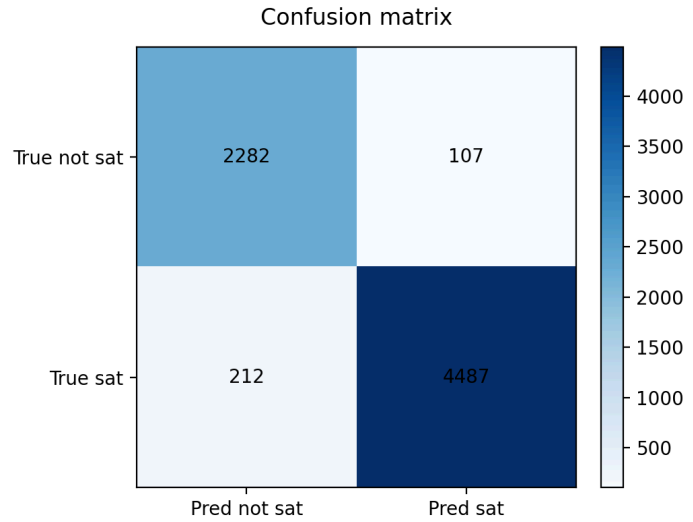



*Figure 6. Confusion matrix for the best-performing text-plus-metadata-and-demographics logistic regression model, evaluated on 7,088 held-out reviews from 500 clinics. The satisfied category includes reviews rated 4–5 stars, while the "Not sat" category includes reviews rated 1–3 stars, including neutral or mixed 3-star reviews.*

## DISCUSSION

The findings suggest that certain socioeconomic and demographic characteristics of the communities surrounding dental clinics were associated with Yelp-reported patient satisfaction. In Model 1, the proportion of residents with a bachelor's degree was positively associated with satisfaction, while several area-level racial and ethnic composition variables showed negative associations. Median household income, poverty rate, and unemployment rate were not statistically significant in this model. After adjusting for business review count, review length, and review year in Model 2, median household income showed a small negative association with satisfaction ($p = 0.039$), while educational attainment remained positively associated with satisfaction. These findings should be interpreted cautiously because the study is observational, socioeconomic variables are measured at the community rather than individual level, and the regression coefficients do not establish causal relationships.

The overall business star rating was excluded from both regression models. Because it is derived directly from individual patient ratings, its association with review satisfaction reflects shared information rather than an independent measure of a clinic’s external quality. The negative income coefficient in Model 2 contrasts with the generally higher unadjusted satisfaction rates observed in higher-income ZCTA categories in Figure 4, although the lowest-income category contains only 24 reviews. These unadjusted descriptive patterns and adjusted regression estimates represent different comparisons and should not be interpreted interchangeably.

Using online reviews from platforms like Yelp requires caution due to multiple potential sources of bias. Online platforms can contain bot reviews, and some clinics may actively solicit positive reviews through promotional incentives. Furthermore, as highlighted by Raihan et al. [3], low-income groups face unique challenges and barriers in accessing digital platforms, including disparities in internet availability, digital literacy, and technology access. These factors create plausible sources of bias in online review data. Because offline populations are missed and positive reviews can be solicited, the exact size and direction of any resulting bias cannot be determined directly from the online review data alone.

This study has several important limitations. First, because it relies on Yelp reviews, the sample may not represent the broader population of dental patients, particularly individuals who do not use online review platforms. Second, the selection of dental businesses through Yelp's available records and dental-related categories may introduce selection bias, and the dataset's geographic coverage is uneven. The high frequency of extreme star ratings may also reflect differences in who chooses to post reviews. Consequently, the overall satisfaction rate of 67.9% cannot be generalized to all dental clinics or patients in the United States. Third, community-level socioeconomic characteristics do not necessarily reflect the financial circumstances or demographic identities of individual reviewers. Future research could incorporate more representative sampling, individual-level socioeconomic information, and systematic analysis of review content to better understand the relationships between community characteristics and reported dental care experiences.

## CONCLUSION

This study examined associations between community-level socioeconomic characteristics and Yelp-reported satisfaction at dental clinics across selected U.S. cities. Educational attainment and several area-level demographic composition variables were associated with review satisfaction in the regression analyses, while median household income was statistically significant only in the model adjusted for review and business metadata. These findings suggest that community characteristics may be associated with online patient ratings, but they do not establish causation or demonstrate that socioeconomic factors are more or less influential than the characteristics of individual dental practices.

Overall, favorable online satisfaction ratings do not prove that dental care is equitable. Since Yelp reviews depend on people deciding to post online, the platform does not systematically capture vulnerable groups who could not afford treatment, did not have dental insurance, or were unable to access care at all. It is important to take into account the wider socioeconomic circumstances as well as the limitations of crowd-sourced data if we are to realise that high online satisfaction does not mean that access to care is equitable, and this underlines the need for specific policies intended at achieving real healthcare equity.

## REFERENCES

[1] Gupta R, Alraqiq HM. Patient experiences at dental school clinics: A comparative analysis of online reviews. J Dent Educ. 2022;86(1):77–87. doi:10.1002/jdd.12782.

[2] Larsen M, Holde GE, Johnsen JAK. Investigating Patient Satisfaction Through Web-Based Reviews of Norwegian Dentists: Quantitative Study Using the Meaning Extraction Method. J Particip Med. 2024;16. doi:10.2196/49262.

[3] Raihan MMH, Subroto S, Chowdhury N, Koch K, Ruttan E, Turin TC. Dimensions and barriers for digital (in)equity and digital divide: a systematic integrative review. Digital Transformation and Society. 2025;4(2):111–127. doi:10.1108/DTS-04-2024-0054.

[4] Senavirathna N, Kabir E, Abebe AM, Khanam R. Socioeconomic status and dental service utilization among children and adolescents: systematic reviews and meta analysis. BMC Oral Health. 2025;25:1370. doi:10.1186/s12903-025-06742-4.

[5] Tran NN, Lee J. Online Reviews as Health Data: Examining the Association Between Availability of Health Care Services and Patient Star Ratings Exemplified by the Yelp Academic Dataset. JMIR Public Health Surveill. 2017;3(3). doi:10.2196/publichealth.7001.

[6] Tsoukatos K, Thomaidis S, Tzaferi EI, Antoniadou M. Online Negative Reviews and Complaints in Healthcare and Dentistry: A Narrative Review of Patient Experiences. Hygiene. 2026;6(3):49. doi:10.3390/hygiene6030049.