

# Predicting Dental Cleaning Uptake

## Introduction

The World Health Organization estimates that around 3.7 billion people worldwide are affected by oral disease<sup>1</sup>. While largely preventable, oral diseases can pose a major health concern if left untreated<sup>2</sup>. Regular dental checkups and cleanings serve as significant measures to identify and prevent oral disease, which is significantly cheaper and less invasive than treatment of advanced decay (fillings, root canals, extractions, etc.)<sup>3</sup>. According to the American Dental Association Health Policy Institute, an estimated 39% of adults between the ages of 19-64 visited a dentist in 2021<sup>4</sup>.

To better understand preventative oral health care behavior, I seek to identify predictors for dental cleaning utilization. Understanding what predicts dental cleaning utilization can help inform policymakers or other stakeholders on how to encourage uptake, leading to reduced costs for individuals and insurance companies as well as improved health outcomes.

## Data & Methods

### Data Sources

I use recently published data from the Medical Expenditure Panel Survey (MEPS), a nationally representative survey of healthcare utilization and expenditures conducted by the Agency for Healthcare Research & Quality. I constructed a two-year panel dataset by linking four MEPS files: person-level consolidated files that I rely on for socioeconomic and demographic data for 2022 and 2023, and corresponding dental visit files contained detailed information on dental care utilization.

### Sample Construction

Since MEPS follows respondents for two years, we restrict the sample to individuals observed in both survey years. The analysis is limited to adults aged 26 and older to reflect independent dental care utilization. After excluding observations with missing data on key

---

<sup>1</sup> World Health Organization, *Oral Health* (Fact sheet, March 31, 2023), <https://www.who.int/news-room/fact-sheets/detail/oral-health>.

<sup>2</sup> World Health Organization, *Oral Health*

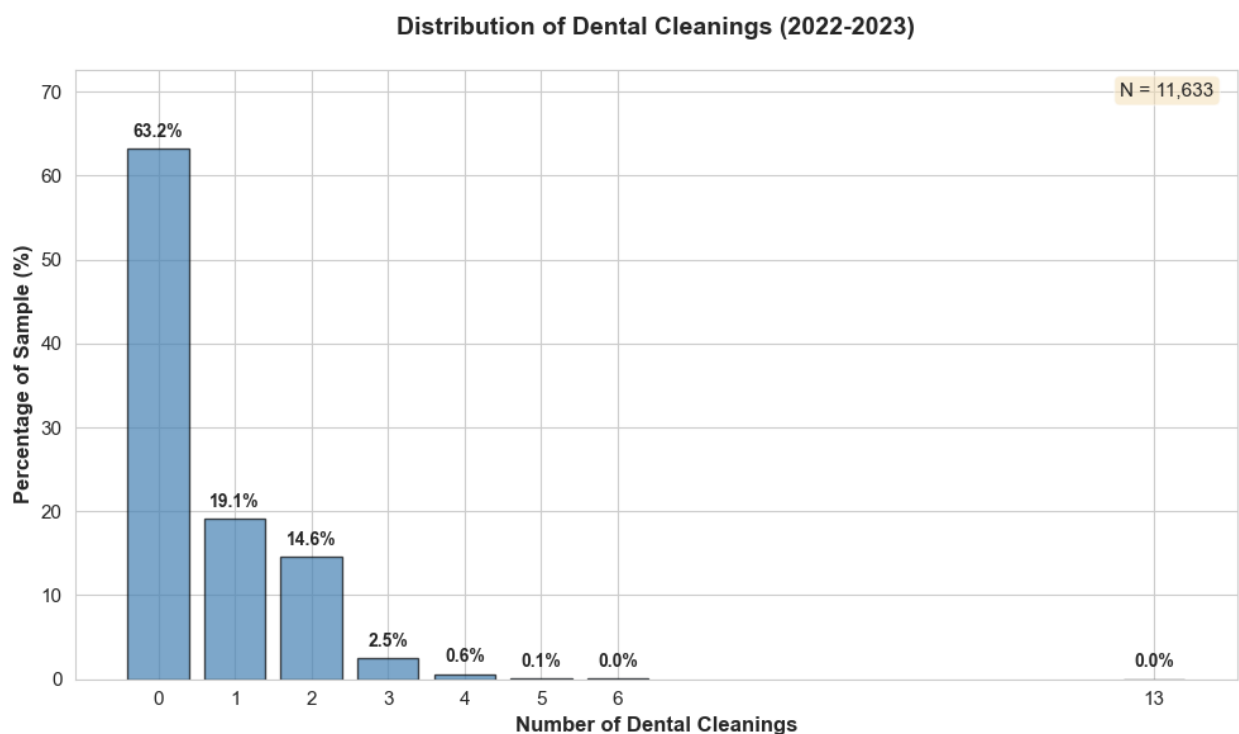
<sup>3</sup> Anuradha Chougule, Supriya Ayachit, and Minal Jamdhade, "AI in Healthcare: Opportunities, Challenges, and the Road Ahead," *Biyani International Journal of Nursing and Reproduction* 4, no. 1 (May 2024): 1–5, <https://bijnr.in/wp-content/uploads/2024/05/BIJNR202406a.pdf>.

<sup>4</sup> American Dental Association, Health Policy Institute, *National Trends in Dental Use, Benefits, and Barriers, 1998–2018* (Chicago, IL: American Dental Association, 2020), [https://www.ada.org/-/media/project/ada-organization/ada/ada-org/files/resources/research/hpi/national\\_trends\\_dental\\_use\\_benefits\\_barriers.pdf](https://www.ada.org/-/media/project/ada-organization/ada/ada-org/files/resources/research/hpi/national_trends_dental_use_benefits_barriers.pdf).

variables, the final analytic sample consists of 11,633 person-year observations from 6,038 individuals.

### Outcome Variable

I conduct a two-stage hurdle model analyze dental cleaning utilization. In the first stage, the dependent variable is a binary indicator equal to 1 if the individual received one or more dental cleanings during the survey year and 0 otherwise. This measure captures utilization of preventative dental care, which is usually recommended annually or biannually by dental professionals. In the sample, 63.2% of person-year observations recorded zero dental cleanings, highlighting substantial gaps in preventative care utilization. In the second stage, the outcome variable represents the number of dental cleanings among those who received at least one cleaning.



### Explanatory Variables

We included 10 explanatory variables in our analysis. To capture demographic and socioeconomic effects, we included: sex, race/ethnicity, veteran status, poverty level as a percentage of federal poverty line (scaled by 1/10,000), highest educational attainment, marital status, and employment status. We proxy for healthcare access by including: dental insurance coverage and whether or not an individual has a usual source of care.

## Statistical Analysis

I employ a two-part hurdle model to analyze dental cleaning utilization, recognizing that the decision to seek any preventative care may differ from the frequency of care among users.

In the first stage, I used logistic regression to model the probability of receiving any dental cleaning, accounting for the panel structure through clustered standard errors by individual ID. This allows for addressing expected correlations in outcomes for the same individual across years.

The model development proceeded in three stages. We began with a baseline model of predictors that we evaluated with a link test. After failing the link test, I added polynomial terms for continuous variables (age and poverty level) to capture potential non-linear relationships. The second model also failed the link test, so I added interaction terms of dental insurance and poverty level and dental insurance and usual source of care. These interaction terms were added to capture additional effects for individuals with dental insurance that may depend on one's income or additional healthcare access. Each model was also assessed by split sample cross validation. Though the final model also failed the link test, it had the strongest AUC measure (.7672 vs .7668 for the second model).

To assess the calibration of the model, I use the Hosmer–Lemeshow test by sorting our model probabilities in ten separate bins to compare our predicted probabilities with the actual outcomes in our data. This gives an extra check of our model quality beyond relying on just the AUC from our cross-validation.

Decile	N	Mean Predicted	Actual Rate	Difference
1	1,508	0.054	0.059	+0.005
2	1,816	0.112	0.101	-0.010
3	1,595	0.176	0.193	+0.016
4	1,585	0.246	0.230	-0.016
5	1,549	0.321	0.277	-0.044
6	1,467	0.393	0.425	+0.031
7	1,242	0.468	0.478	+0.010
8	730	0.544	0.555	+0.011
9	141	0.628	0.636	+0.008

Hosmer-Lemeshow  $\chi^2 = 11.839$ ,  $df = 8$ ,  $p = 0.159$   
Mean Calibration Error = 0.016 | Max Calibration Error = 0.044 | Brier Score = 0.185  
Interpretation: Good calibration ( $p > 0.05$ ). Model predictions closely match observed outcomes.

The second stage of our model employs a zero-truncated negative binomial regression that is conditional on receiving at least one dental cleaning. I used the same predictor set as Part 1 and computed clustered standard errors by individual ID. The second stage model was also assessed with split sample cross-validation which showed a mean absolute error of 0.62 cleanings and root mean squared error of 0.75. Finally, I assess the calibration of the model by comparing mean predicted versus mean actual counts across deciles of predicted probabilities.

This two-part approach gives an extra layer of analysis by distinguishing between barriers to initial access (Part 1) and factors influencing care intensity among those already engaged (Part 2).

Decile	N	Mean Predicted	Mean Actual	Difference
1	429	1.482	1.389	+0.093
2	428	1.514	1.430	+0.084
3	429	1.532	1.471	+0.061
4	428	1.548	1.547	+0.001
5	429	1.564	1.618	-0.054
6	428	1.586	1.617	-0.030
7	428	1.615	1.685	-0.070
8	429	1.650	1.709	-0.058
9	428	1.699	1.792	-0.093
10	429	1.791	1.667	+0.124

## Results and Policy Implications

### *First Stage: Any Cleanings*

Variable	Coefficient	Odds Ratio	95% CI
Female	0.207***	1.23	[1.11, 1.37]
No Degree	-0.927***	0.40	[0.31, 0.51]
Bachelor's Degree	0.664***	1.94	[1.70, 2.22]
Veteran	-0.158	0.85	[0.70, 1.04]
Employed	-0.186**	0.83	[0.73, 0.95]
No Usual Source of Care	-0.714***	0.49	[0.41, 0.58]
Has Dental Insurance	0.688***	1.99	[1.67, 2.37]
Income Level (scaled)	2.256***	9.54	[6.87, 13.25]
Income Level <sup>2</sup> (scaled)	-0.593***	0.55	[0.48, 0.64]
Dental Insurance × Income	-0.397**	0.67	[0.52, 0.87]
Dental Insurance × No USC	0.353**	1.42	[1.13, 1.79]
Not Married	-0.194***	0.82	[0.74, 0.92]
Non-White Race/Ethnicity	-0.506***	0.60	[0.54, 0.67]
Advanced Degree	0.497***	1.64	[1.45, 1.86]

The first stage model identifies several insights about dental cleaning utilization. The analysis reveals a non-linear relationship between income and dental cleanings. The significant quadratic income term demonstrates diminishing marginal returns, where each additional percentage point increase relative to the federal poverty line provides smaller benefits at high income levels. The statistically significant negative coefficient on the interaction term between income and dental insurance also indicates that insurance coverage is most beneficial for low-income populations. This finding suggests that expanding dental insurance would disproportionately help those who need it the most, indicating that targeted insurance programs for low-income individuals could achieve substantial impact at lower costs relative to universal dental insurance.

Lacking a usual source of care emerges as one of the strongest predictors of less dental cleaning uptake (OR = 0.49,  $p < .001$ ). The significant positive interaction between dental insurance and lack of usual source of care (OR = 1.43,  $p < .01$ ) indicates that insurance coverage may partially compensate for fragmented healthcare access. Additionally,

despite controlling for income, education, insurance, and healthcare access, substantial racial disparities persist. Non-white individual show 40% lower odds of receiving dental cleanings (OR = 0.61,  $p < 0.001$ ).

### ***Second Stage: Frequency of Cleanings Among Users***

Variable	Coefficient	IRR	95% CI
Age	-0.015	0.99	[0.97, 1.00]
Female	0.011	1.01	[0.93, 1.10]
No Degree	-0.032	0.97	[0.76, 1.24]
Bachelor's Degree	0.042	1.04	[0.94, 1.15]
Veteran	-0.022	0.98	[0.86, 1.11]
Employed	-0.027	0.97	[0.87, 1.09]
No Usual Source of Care	-0.031	0.97	[0.83, 1.13]
Has Dental Insurance	-0.008	0.99	[0.87, 1.14]
Agex Sq	0.000*	1.00	[1.00, 1.00]
Income Level (scaled)	0.195	1.22	[0.97, 1.52]
Income Level <sup>2</sup> (scaled)	-0.060	0.94	[0.85, 1.04]
Dental Insurance x Income	0.010	1.01	[0.85, 1.20]
Dental Insurance x No USC	0.010	1.01	[0.83, 1.23]
Not Married	0.066	1.07	[0.98, 1.16]
Non-White Race/Ethnicity	-0.050	0.95	[0.87, 1.04]
Advanced Degree	0.038	1.04	[0.95, 1.14]

The results revealed limited variation in the number of cleanings among users. After accounting for clustering, few covariates achieved statistical significance. Age exhibited a significant quadratic relationship with the number of cleanings (age:  $\beta = -0.015$ ,  $p = 0.075$ ; age<sup>2</sup>:  $\beta = 0.00019$ ,  $p = 0.013$ ). The turning point occurs at approximately age 39, after which the negative association between age and cleaning frequency begins to attenuate.

Notably, dental insurance coverage was not significantly associated with the number of cleanings among those who had at least one cleaning ( $\beta = -0.008$ ,  $p = 0.906$ ). Similarly, the interaction terms between dental insurance and income ( $p = 0.912$ ) and between dental insurance and lacking a usual source of care ( $p = 0.922$ ) were not statistically significant. Other demographic and socioeconomic factors, including gender, education, marital

status, race/ethnicity, employment status, and veteran status, also showed no significant associations with cleaning frequency.

To get more interpretable results, I implement counterfactual policy simulations using our first stage model with statistically significant predictors.

### ***Universal Dental Insurance***

To simulate the “best case scenario”, I simulate a potential outcome of expanding dental insurance to everyone, holding all other factors constant. At baseline, dental cleaning utilization was measured at 36.8%; under the universal dental insurance scenario, we see a 6.2 percentage point increase to 43.1%. Nominally, this brings an estimated 363 additional people from our sample into regular dental care annually.

Universal dental insurance is a costly solution. So, to target more realistic policy interventions, we simulate expanding dental insurance coverage to lower income individuals, which still reveals promising efficiency. We identify lower income individuals as those who are at or below the median poverty level as a percentage of the Federal Poverty Level. This policy would increase utilization from 36.8% to 40.5%, achieving 58% of universal dental insurance’s impact (211 new users vs 363) while only affecting 37% of the sample population. This policy reveals itself as being highly cost effective due to the interaction effect showing greater insurance benefits for low-income populations.

### ***Universal Usual Source of Care***

Assigning a universal usual source of care shows more moderate effectiveness. While less impactful than insurance interventions, we still see a 2.3 percentage point increase in dental cleaning uptake from the baseline. This policy may be more politically feasible as it could be promoted through awareness campaigns surrounding the importance of having a usual source of care, as 10% of our sample that have dental insurance don’t have a usual source of care.

### ***Income Support Programs***

To simulate the effectiveness of income support programs, we increase income for the bottom quartile of earners by 50%. The impact is marginal however, a meager 0.3 percentage point increase in dental cleanings. The diminishing returns to income captured by our quadratic term may explain why direct income transfers are less effective than target insurance coverage.

## Conclusion

This analysis of dental cleaning utilization among U.S. adults reveals significant disparities in preventative oral healthcare utilization and key policy levers for improvement. Using the nationally representative MEPS dataset spanning from 2022-2023, we find that nearly two-thirds of adults receive no dental cleanings in a given year, highlighting behaviors that could lead to more costly and painful interventions later.

Our two-part hurdle model reveals an important distinction between access to care and intensity of care among users. The first stage demonstrates that dental insurance coverage, usual source of care, income, and race are the primary drivers of whether individuals receive any dental care. Importantly, the significant interaction between dental insurance and income reveals that insurance benefits are not distributed equally and that coverage provides the greatest benefit to low-income populations. This finding has direct policy implications: targeted dental insurance programs for low-income individuals can achieve substantial impact while being more cost-effective than universal coverage.

In contrast, the second stage analysis reveals that among individuals who already access dental care, utilization patterns are relatively similar across socioeconomic groups. Dental insurance, income, education, race/ethnicity, and other demographic factors showed no significant associations with the number of cleanings received. Only age exhibited a significant (though modest) quadratic relationship. This suggests that disparities in dental care are concentrated at the point of initial access rather than in the intensity of care received.

This finding has important policy implications: interventions should focus on reducing barriers to initial access rather than increasing utilization among existing users. The policy simulations from our first stage model underscore this potential. Universal dental insurance would increase cleaning utilization by 6.2 percentage points, but targeted insurance for low-income uninsured individuals achieves 58% of this impact while affecting only 37% of the population. This efficiency gain stems from the interaction effect showing that insurance helps low-income individuals more than their higher-income counterparts.

Beyond insurance, our analysis reveals that lacking a usual source of care is one of the strongest predictors of foregone dental care. The 10% of individuals who have dental insurance but lack a usual source of care represent a particularly important target for intervention through care coordination and provider network expansion.

These findings suggest that effective policy interventions should prioritize: (1) expanding dental insurance coverage and (2) improving care coordination and usual source of care

access. While income support programs show minimal impact on dental utilization due to diminishing returns, direct interventions addressing insurance coverage and healthcare access barriers offer more promising pathways to improving preventative oral health behavior.

Future research should explore other factors related to dental care uptake not available in our chosen dataset. A follow-up analysis of the downstream effects of dental care uptake on healthcare expenditures is also relevant and feasible with the available data in MEPS.

## References

American Dental Association, Health Policy Institute. *National Trends in Dental Use, Benefits, and Barriers, 1998–2018*. Chicago, IL: American Dental Association, 2020. [https://www.ada.org/-/media/project/ada-organization/ada/ada-org/files/resources/research/hpi/national\\_trends\\_dental\\_use\\_benefits\\_barriers.pdf](https://www.ada.org/-/media/project/ada-organization/ada/ada-org/files/resources/research/hpi/national_trends_dental_use_benefits_barriers.pdf).

Chougule, Anuradha, Supriya Ayachit, and Minal Jamdhade. “AI in Healthcare: Opportunities, Challenges, and the Road Ahead.” *Biyani International Journal of Nursing and Reproduction* 4, no. 1 (May 2024): 1–5. <https://bijnr.in/wp-content/uploads/2024/05/BIJNR202406a.pdf>.

World Health Organization. *Oral Health*. Fact sheet. March 31, 2023. <https://www.who.int/news-room/fact-sheets/detail/oral-health>.