TITLE: Comparing Providers for the Treatment of Early Childhood Caries with a Value-based Rating Methodology: A Cross-sectional study

AUTHORS:
David Okuji, DDS, MBA, MS1, Betty Ben Dor, BS2, Ashiana Jivraj, MBA, DMD3, Yinxiang Wu, MA4

1Senior Associate Director, NYU Langone Dental Medicine, Hansjörg Wyss Department of Plastic Surgery, Division of Dental Medicine, NYU Grossman School of Medicine, 5800 Third Avenue, 3rd Floor, Brooklyn, NY 11220

2Pre-doctoral student, Harvard School of Dental Medicine, 188 Longwood Avenue, Boston, MA 02115

3Graduate, Harvard Business School, Dillon House, Soldiers Field, Boston, MA 02163 and Harvard School of Dental Medicine, 188 Longwood Avenue, Boston, MA 02115

4Former Assistant Research Scientist, NYU School of Medicine-Division of Biostatistics, Department of Population Health, 180 Madison Avenue, Fifth Floor, New York, NY 10016

Author Contributions
David Okuji, DDS, MBA, MS: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Roles/Writing - original draft; Writing - review and editing.

Betty Ben Dor, BS: Roles/Writing - original draft; Writing - review and editing.

Ashiana Jivraj, MBA, DMD: Formal analysis; Roles/Writing - original draft; Writing - review and editing.

Yinxiang Wu, MA: Data curation; Formal analysis; Software; Validation; Visualization; Roles/Writing - original draft; Writing - review and editing.

Declaration of Interest
The authors state no conflicts of interest.

Role of the Funding Source
This research was funded by the Hansjörg Wyss Department of Plastic Surgery, NYU Langone Hospitals, New York, NY and had no involvement with the study design; collection, analysis, and interpretation of data; writing of the report; and the decision to submit the article for publication.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
SHORT TITLE: Rating providers for treatment of ECC by value-based method

KEYWORDS: value-based, rating-survey, early-childhood-caries

CORRESPONDING AUTHOR:

David Okuji, DDS, MBA, MS
NYU Langone Hospitals
5800 Third Avenue, 3rd Floor
Brooklyn, NY 11220
Phone 617-903-7117
Email: david.okuji@nyulangone.org

WORD COUNT: Text: 4,252, Abstract: 279

Highlights

- Non-surgical management of early childhood caries yields the most value.
- Surgical intervention with or without general anesthesia is less valued by patients.
- Ranking treatment modalities by value is a useful tool for patients.
- Using a parental-reported symptom and service quality questionnaire assesses quality.
- Provider value can be determined by dividing quality of care by the cost of delivery.
ABSTRACT

Purpose: An adapted Porter-Teisberg model of value-based health outcomes and cost was applied to comparatively rate providers who treat children with early childhood caries, by treatment modalities, facilities, and geographic location.

Methods: Quality was measured by a Parental-Reported Symptom and Service Quality questionnaire, adapted from the Early Child Oral Health Impact Scale, Child-Oral Impacts on Daily Performance, and Dental Consumer Assessments of Healthcare Providers and Systems tools. Cost was measured by the labor cost for all providers involved in treatment and adjusted by dividing by total relative value units for the types and total number of procedures for all treatment visits. Value was calculated as the quotient of the weighted quality score and the adjusted cost, multiplied by a coefficient. To compare the value score by states and by facilities, Tukey’s multiple comparison test was used and a sensitivity analysis using the non-parametric Dunn’s Kruskal-Wallis multiple comparison test was conducted.

Results: For surgical treatment with general anesthesia, Tennessee ranked the highest, followed by Arizona, Hawaii, and New York. These four states had significantly higher value metrics than Florida, Massachusetts, and Maryland. Maryland had a significantly higher value score for surgical treatment without general anesthesia. While patients were seen for non-surgical treatment only in Maryland and New York, Maryland also had a significantly higher value for this treatment modality.

Conclusion: A value-based rating system of providers, by geographic region and treatment facility, can assist patients to better compare and select providers; third-party payers, such as insurance companies and government programs, to better compare outcomes and costs for provider systems and clinic facilities; and providers to benchmark their comparative ratings on an internal and external basis for continuous quality improvement.
INTRODUCTION

Recent years have seen an increased legislative and political effort to establish a value-based health care system in an attempt to deliver better care to patients while decreasing overall costs.1 A predominant model of evaluating and rating dental providers in the United States (U.S.) is through the use of online customer satisfaction and review digital media platforms, such as Yelp.com and Healthgrades.com.2 These public platforms allow patients to freely comment on their experiences with particular offices and providers in a manner that is easily accessible to potential and current patients, allowing them to use this information for the selection of a dentist.

Abbreviations in sequential order of first presentation:
United States (U.S.)
Centers for Medicare and Medicaid Services (CMS)
International Consortium for Health Outcomes Measurement (ICHOM)
American Dental Association (ADA)
Dental Quality Alliance (DQA)
early childhood caries (ECC)
federally qualified health center (FQHC)
minimally invasive non-surgical care (NSC)
surgical restorative care without adjunctive sedation or general anesthesia (SC)
surgical restorative care with general anesthesia (SCGA)
parental-reported symptom and service quality questionnaire (PRSSQ)
silver diamine fluoride (SDF)
health care providers (HCP)
general anesthesia (GA)
quality (Q)
Early Child Oral Health Impact Scale (ECOHIIS)
Child-Oral Impacts on Daily Performance (C-OIDP)
Dental Consumer Assessments of Healthcare Providers and Systems (D-CAHPS)
cost (C)
Relative Value Units (RVU)
value (V)
American Society of Anesthesiologists (ASA)
University of Pennsylvania Leonard Davis Institute of Health Economics (LDI)
Affordable Care Act (ACA)
Fédération Dentaire Internationale (FDI)
Adult Oral Health Standard Set (AOHSS)
Customers who have positive experiences tend to submit high ratings and those who have negative experiences do not typically submit ratings, which can artificially inflate the overall ratings.\(^3\)

One of the most comprehensive value-based care online platforms is the Medicare Care Compare tool, developed by the Centers for Medicare and Medicaid Services (CMS) and based on outcomes data.\(^4\) However, one large study showed Yelp.com was the most frequently used online surgeon rating site, and was nearly three times as likely to be used as the Medicare Physician Compare website.\(^2\) The Medicare Care Compare tool provides the public with a comprehensive website that patients can use to find in-network medical providers.\(^5\) Patients are provided information on the providers’ certifications and locations and have access to federal government provider-rating data. Few dentists are listed as providers, except participating oral and maxillofacial surgeons. Quality scores are listed as star ratings, are not typically available at the individual provider level, and are available at the “group affiliation” level if the individual provider and group opt to participate in the quality initiatives. The quality star ratings “are only publicly reported if the measured data meet the established public reporting standards and resonate with users.\(^6\)” The quality star rating methodology is based upon establishing statistically derived benchmarks based upon mean performance across top-ranked performers.\(^6\) Participating providers who meet or exceed the benchmark are assigned 5 stars. An equal range method is utilized to assign 4-, 3-, 2-, and 1-star ratings.\(^6\) The benchmarks for providers include innovative model participation, electronic health record technology performance information, improvement activities, quality performance information, and/or patient survey scores.\(^7\)

Neither customer satisfaction and review platforms nor the statistically derived CMS quality star ratings include the cost of delivering care as part of their rating methods. Hence, this
opens an opportunity for an improved, standardized, evidence-based, and easily utilized approach for rating providers, including dentists, for the delivery of value-based health care.

Multiple organizations are studying evidence-based structure and processes toward value-based outcome models. The International Consortium for Health Outcomes Measurement (ICHOM) has based their approach on the 2006 book, *Redefining Health Care*, by Harvard Business School Professors Michael E. Porter and Elizabeth O. Teisberg, who respectively lead the Harvard Business School, Institute for Strategy & Competitiveness, Health Care Team and the Value Institute for Health and Care out of the University of Texas at Austin Dell Medical School and McCombs School of Business. Although it is quality-focused, not value-focused, the American Dental Association (ADA) has developed a Dental Quality Alliance (DQA) with the expressed mission of expanding a set of performance measures for the delivery of oral health care.

Value-based studies which focus on early childhood caries (ECC) allow investigators to conduct research for a disease condition confined to a relatively short time-period, with known risk factors, and standard treatment modalities with wide variations in cost of care. ECC refers to caries found in the primary dentition of children under the age of six years old and globally presents as one of the most common chronic childhood diseases, primarily affecting those in low income and minority groups. It can lead to pain, infection, and decreased body mass index, along with negative effects on childhood learning and quality of life. ECC care presents a tremendous burden not only on families and communities, but on health care systems as well. As of 2013, pediatric oral health care generated $26.5 billion in expenditures. Given the negative clinical and economic impacts of ECC on the pediatric population, it makes financial sense to support patients and third-party payers with value-based tools to better select providers who improve oral health outcomes and reduce costs.

The Porter and Teisberg model for value-based health care focuses on the delivery of cycles of care for medical conditions. The model defines health care value as the quotient of
quality and cost. The aim with this study is to use an adapted formulation of the Porter-Teisberg value-based model to rate provider systems for their management of ECC, with the specific objective to compare value scores and letter grade ratings of providers at the geographic location and clinic facility levels. The null hypothesis was that providers would demonstrate equal value scores and letter grade ratings, as determined by a value-based statistical model used in this study.

**METHODS**

**Study Design and Setting**

This study utilized an analytical, observational, cross-sectional design to compare a value score and letter grades for providers at the levels of treatment facilities and geographic regions identified by specific U.S. states. The study was approved by the NYU Grossman School of Medicine Institutional Review Board under protocol number 16-00331.

A convenience sampling methodology was utilized. Children under six years old, who were patients of federally qualified health centers (FQHCs) with a diagnosis of ECC, were recruited and stratified by the treatment modality selected by their legal caregiver via the informed consent process. The three treatment modalities included minimally invasive non-surgical care (NSC), surgical restorative care without adjunctive sedation or general anesthesia (SC), and surgical restorative care with general anesthesia (SCGA).

Value of care was calculated as the parental-reported quality divided by the total cost of care. Quality was measured by a Parental-Reported Symptom and Service quality Questionnaire (PRSSQ), which was completed by the caregiver approximately one-month after the last treatment visit (Figure 1). Total cost of care was calculated by multiplying the total number of hours to complete the treatment by the total hourly labor costs of health care providers (HCP) who were directly involved with the treatment. Value was then calculated by dividing the quality PRSSQ score by the total cost of care. The comparative value for each
treatment facility and each region by state were statistically clustered and assigned a letter score, with “A” as the highest and “E” as the lowest ratings.

**Study Size and Participants**

Patient care was carried out at 16 clinic facilities located at FQHCs, hospitals, and surgery centers affiliated with the NYU Langone Dental-Advanced Education in Pediatric Dentistry program in seven states: Arizona, Florida, Hawaii, Maryland, Massachusetts, New York, and Tennessee.

Four-hundred, eighty-seven (487) subjects were recruited from 12/01/2015 to 10/03/2016. The patient populations served at these health centers were from predominantly underserved communities and made up an ethnically diverse group. Parents or legal caregivers of pediatric patients with a diagnosis of ECC were recruited for the study and underwent the informed consent process if they assented to participation.

The post-hoc power analysis was calculated at 99.4%, based upon the study’s sample size, treatment size effects, and alpha at 0.05.

Inclusion criteria for the study were 1) registration as a patient at one of the FQHCs, hospitals, or surgery centers affiliated with the NYU Langone Dental-Advanced Education in Pediatric Dentistry program, 2) age six or younger at the time of enrollment into the study, 3) a diagnosis of ECC, 4) a full complement of 20 primary teeth, 5) American Society of Anesthesiology status of 1, 2, or 3, and 6) consent for treatment by the parent/caregiver in either Spanish or English.

**Variables**

The independent variables for treatment modality were defined as follows:

**Non-surgical care (NSC):** Non-surgical care was defined as treatment with topical fluoride varnish and silver diamine fluoride (SDF) applications, interim therapeutic restorations, motivational interviewing, frequent follow-up treatment visits, and active surveillance visits.
Surgical care without sedation/general anesthesia (SC): Surgical care without sedation or GA was defined as full- or partial-mouth treatment with local anesthesia in one or more treatment visits with or without the use of nitrous oxide analgesia.

Surgical care with general anesthesia (SCGA): Surgical care with general anesthesia (GA) was defined as full-mouth dental rehabilitation in one visit under GA.

The dependent variables for quality and cost were defined as follows:

Quality (Q): Quality of care was assessed using a 26-question PRSSQ. This post-treatment survey was created for deployment to the parents or legal guardians of the patients one month after completion of treatment. The PRSSQ consists of questions adapted and compiled from three validated satisfaction surveys. Thirteen questions came from the Early Child Oral Health Impact Scale (ECOHIS), which looks at the effects of pediatric oral health on children and caregivers. Another six questions came from the Child-Oral Impacts on Daily Performance (C-OIDP), which assesses the impact of oral health on eating, speaking, cleaning teeth, smiling, emotional stability, relaxing, doing schoolwork, and social contact. Nine questions were sourced from the Dental Consumer Assessments of Healthcare Providers and Systems (D-CAHPS), which is used to measure satisfaction with dental care among consumers. The score for quality was calculated by using the equally weighted average score of the three components that made up the PRSSQ. Thus, as a metric, the PRSSQ is a compilation of existing, validated assessments for parental satisfaction, parent clinical observations, and impact on families.

Cost (C): As a proxy for total direct cost, as prescribed by the Porter-Teisberg model, total direct labor costs associated with treatment were utilized since direct labor cost contributes to nearly 60% of operational expenses for a pediatric dental practice. Accountants define “total direct cost” as the cost of materials and supplies, in addition to labor. This study defined “total direct labor cost” as the total labor cost for HCPs required to complete the treatment of each child. The HCPs involved in delivering care included dental residents, attending dentists, anesthesiologists, nurse-anesthetists, nurses, surgery technicians, dental hygienists, and/or
dental assistants. Direct labor cost was measured, and adjusted for the effects of the types and number of total procedures, by calculating the total duration of treatment time for the total number of visits, multiplying by the labor rate for the aggregate of each clinical staff member involved with the child's treatment, and dividing by the total Relative Value Units (RVU) for the types and total number of procedures performed for all treatment visits.\textsuperscript{20} In this way, the RVU-adjusted cost provided a basis for equal comparison among the three treatment modalities.

**Value (V):** Value was calculated as quality (PRSSQ score), divided by RVU-adjusted cost (C), times a coefficient of 100, with the multiplier product resulting in a value numeric which is greater than zero.

**Letter grade ratings:** Based upon the statistical model, letter grade ratings were assigned to provider groups which were clustered to have statistically similar mean value scores. The highest assigned letter grade was (A) and the lowest (E).

The dependent variables for provider levels were defined as follows:

**Geographic location:** The location was defined as the U.S. state in which the treatment facilities are located and identified as Arizona (AZ), Hawaii (HI), Florida (FL), Maryland (MD), Massachusetts (MA), New York (NY), and Tennessee (TN).

**Treatment facility:** The 16 treatment facilities were classified by type of clinic facility, geographic location, and identified as 1) Out-patient dental clinics: AZ-1, FL-1, FL-2, MA-1, MD-1, NY-1, TN-1, 2) Out-patient dental clinics with mobile general anesthesiology services: AZ-1, NY-1, and TN-1, and 3) Hospital or ambulatory surgery centers: HI-1, HI-2, FL-3, FL-4, MA-2, MD-2, MD-3, NY-2, TN-2.

**Individual dentist-provider:** Although data was available at the individual dentist-provider level, each individual dentist-provider did not treat a sufficiently large number of patients to yield meaningful and comparative statistical analyses.
Data Sources, Measurements, and Data Collection

Dental record:

Patients first received an oral examination, with a diagnosis of ECC as defined by the American Academy of Pediatric Dentistry, followed by development of treatment plan options to be provided to the parent/caregiver. The children were then classified as an American Society of Anesthesiologists Physical (ASA) Status Classification 1, 2, or 3 and were categorized to receive either non-surgical care (NSC), surgical treatment without general anesthesia (SC), or surgical treatment with general anesthesia (SCGA).

Patients were treated by the assigned individual dentist-provider on any given day. Given the large number of dentist-providers rendering the treatments across all patients, the assumption was there should be very little variation in the treatment outcomes on average. Value was not able to be compared at the individual dentist-provider level because there was, on average, a very low number of treatments rendered per each dentist-provider. Individual providers were therefore grouped by state and clinic facilities. The maximum number of treatment visits for each treatment modality listed was one for SCGA and up to five visits for SC and NSC.

Third party sources:

Direct labor cost of care was calculated by multiplying the duration of treatment time in hours by the aggregate hourly wage rate for HCPs involved in the treatment. The duration of treatment time was measured from the time of ingress to time of egress from the treatment room. The average hourly wage rate for HCPs was calculate from U.S. Department of Labor statistics.

PSSRQ survey:

The quality of the provider was assessed using parental responses to the PRSSQ. Average scores from the component surveys of the PRSSQ were summed and given weight to adjust for their differing sizes of contributions to the aggregated PRSSQ metric.
One month after completion of patient treatment, the PRSSQ was administered to parents or legal caregivers of the patients via telephone or in-person.

Bias

The primary sources of bias were selection and measurement error. Selection bias was addressed by recruiting a large sample of subjects across a large and diverse population in seven states. Measurement errors were addressed by orienting and training the data collectors on how to measure the direct labor costs and how to objectively collect PRSSQ survey data for each subject.

Statistical Methods and Quantitative Variables

Means and standard deviations were calculated for continuous variables. Counts and percentages were calculated for categorical variables. These descriptive statistics were presented both overall and by treatment modalities. Comparisons of distributions by treatment modalities were conducted by the Chi-square test (or the Fisher’s exact test, as appropriate) for categorical variables and by the one-way ANOVA test (or the Kruskal-Wallis test, as appropriate) for continuous variables.

To compare the value score by states and by facilities in which patients received their treatment, Tukey’s multiple comparison test was used. Preliminary data analysis demonstrated there was significant difference in the value score among three treatment modalities. Hence, comparisons of the value score by states and by facilities were conducted within each category of treatment modality. Natural logarithm transformation was applied to the value score to approximate a normal distribution. States or facilities with less than ten observations were excluded from the multiple comparison. Letter grade ratings were assigned to provider groups which were clustered to have statistically similar mean value scores.

To assess the robustness of results based on the Tukey’s multiple comparison test to potential heterogeneity of variances of the ln-transformed value score across states and
facilities, a sensitivity analysis using the non-parametric Dunn’s Kruskal-Wallis multiple comparison test was conducted.

All analyses were conducted in R version 3.5.0 for Windows,\textsuperscript{24} with ‘lsmeans’ package for Tukey’s multiple comparison test,\textsuperscript{25} and ‘FSA’ package for Dunn’s Kruskal-Wallis multiple comparison test.\textsuperscript{26}

RESULTS

Participants and Descriptive Data:

The analytic sample population included 487 patients, with an average age of 3.99 years (standard deviation (sd)=1.22) and characteristics comprised predominantly of female (50.1 percent), Hispanic ethnicity (56.2 percent), Other race (45.2 percent), and Medicaid enrollees (94.9 percent). The distribution by treatment modality group was SCGA (53 percent), SC (32 percent), and NSC (15 percent) (Table 1). Significant differences in the distributions of nearly all demographic variables were observed among three treatment modalities. Significant differences were also observed in the quality, cost, and value score. Specifically, NSC had on average the highest quality score, lowest cost score, and hence the highest value score among the three treatment modalities. The rating hierarchy of the value score was NSC > SC > SCGA. Comparisons of the value score by states and facilities were therefore carried out within each treatment modality.

Outcomes Data and Main Results:

Figure 2 presents the results from Tukey’s multiple comparison test for the comparisons of the ln-transformed value score by states, with all results significant at p<0.001. For SCGA, Tennessee had the highest average value score with letter grade (A), followed by the respective states and letter grades of Arizona (A-), New York (B), and Hawaii (B-). These four states also have on average significantly higher value scores than the other three states with letter grades of (C), namely, Florida, Maryland, and Massachusetts, with no significant differences observed between these three states. For SC, Maryland had significantly higher value score and letter
grade (A) than the other states, followed by Arizona (B), New York (B-), Tennessee (C). The value score for these four states outnumbered that of Florida and Massachusetts, both with letter grade (D). For NSC, patients were only seen in Maryland and New York. The respective (letter grade) and [ln-value mean score (standard deviation)] were significantly higher for Maryland (A) [4.81 (0.60)] than that of New York (B) [4.30 (0.44)].

Figure 3 shows the results from Tukey’s multiple comparison test by facilities, with all results significant at p<0.001. For SCGA, facilities with letter grade (A) included TN-2 and AZ-1, followed by letter grade (B) at HI-1, FL-4, NY-2, MA-2, and MD-2. For SC, the letter grade groups included (A) at MD1; (B) at AZ-1; (B-) at NY-1; (C-) at TN-1; (D-) at FL-2; and (E) at MA-1. For NSC, patients were only seen at MD-1 and NY-1. The respective (letter grade) and [ln-value mean score (standard deviation)] were significantly higher for MD-1 (A) [4.81 (0.60)] than for NY-1 (B) [4.30 (0.44)].

To assess the robustness of results based on the Tukey’s multiple comparison test to potential heterogeneity of variances of the ln-transformed value score across states and facilities, a sensitivity analysis using the non-parametric Dunn’s Kruskal-Wallis multiple comparison test was conducted. Table 2 presents the results of the sensitivity analysis by state and Table 3 shows the results by facility. The sensitivity analyses based on the Dunn’s Kruskal-Wallis multiple comparison test yielded similar results and confirmed the findings shown in Figures 2 and 3.

DISCUSSION

Findings and Hypothesis:

The key results showed statistical significance for rating providers by geographic region and specific clinic facilities. For analyzed U.S. states, Arizona and Tennessee received the highest mean ln-value scores with the letter grade (A) rating for SCGA; Maryland received the highest mean ln-value score with letter grade (A) rating for SC and NSC. For analyzed clinic facilities,
two dental clinics with mobile dentist-anesthesiology services, AZ-1 and NY-1, and a surgery center, TN-2, received the highest mean ln-value scores with letter grade (A) rating for SCGA; Maryland received the highest mean ln-value score with letter grade (A) for SC and NSC.

This study’s findings rejected the null hypothesis and demonstrated that the statistical model identified distinct geographic and clinic facility provider groups, which were rated by mean ln-value scores, and associated letter grade ratings. This evidence-based model demonstrated a hierarchal, value-based, health care provider rating system.

Comparison to Literature:

Except for another article by this study’s principal investigator, a search of the literature did not yield studies which quantified oral health care value by quality and cost. Instead, the literature exclusively focused on either the rationale for value-based health care, quality (i.e., patient-reported and/or clinical outcomes), or cost.

Multiple studies support the findings presented here regarding the rationale of value-based rating systems. In 2008 the Institute of Medicine (renamed as the National Academy of Medicine in 2015) and the Institute for Healthcare Improvement developed a framework of the “Triple Aim” goals for the future of high-value U.S. health care delivery, which consisted of “improving the individual experience of care, improving the health of populations, and reducing the per capita costs of care for populations.” A 2021 white paper from the University of Pennsylvania Leonard Davis Institute (LDI) of Health Economics looked at how far the U.S. health care system has progressed toward establishing a value-based institution since the passage of the Affordable Care Act (ACA) in 2010 and concluded that reimbursement systems should “push providers away from fee-for-service payment” and well-designed alternative payment models have resulted in improved value with cost savings.

Quality-focused publications, which included patient-reported and clinician-outcomes studies, were found in the dental literature. One research team found that online reviews in platforms such as Google, Yelp, and Healthgrades form a large and rich dataset reflective of the
patient experience, which can have the effect of creating a de facto patient-reported quality metric for providers. However, these websites are more susceptible to selection bias depending on the demographic of the patients that choose to leave publicly visible reviews.

A collaboration between the Fédération Dentaire Internationale (FDI) World Dental Federation and International Consortium for Health Outcomes Measurement (ICHOM) developed an Adult Oral Health Standard Set (AOHSS) of patient- and clinician-reported outcome measures, specifically focusing on caries and periodontal disease. One author outlined the challenges for developing value-based oral health care reporting system, which included the need for creating multi-disciplinary care, creating quantitative value measurement, identifying cost attribution for care, and establishing bundled care reimbursement models.

Cost-focused medical research on peri-operative care, musculo-skeletal and diabetes care, cystectomy, and extracorporeal life support exclusively emphasized the identification, measurement, and implementation of cost efficiencies. One author reported that, although “medicine has built a body of knowledge from cost-effectiveness, cost-benefit and cost-utility studies, dentistry has not addressed the same questions to any significant degree” and, thus, without a firm estimation of cost, an accurate estimate of value cannot be determined.

Study Strengths, Limitations, and Generalizability:

Strengths of this study include geographic diversity, given the range of states that were included. The patient population also had extensive racial and ethnic diversity. Furthermore, the post-hoc power analysis yielded a power level of 99.4 percent.

One of the limitations of this study is that the PRSSQ quality score in this study was a subjective value. Another limitation is that the sample population was recruited at FQHCs, which serve low SES populations where the cost and quality variables may differ compared to non-FQHC populations.

The findings may not be generalizable due to the high percentage of Medicaid enrollees in the study. Additionally, direct labor costs were used as a proxy for total cost, which could
skew the cost metric. The study was also limited by the number of observations of the individual providers. Furthermore, there was no randomization process for assigning patients to providers and some of the states did not see patients for each of the three procedure types, limiting the analysis of provider value by facility for a given treatment.

**Future Studies:**

Future studies might apply this model of value-based health care to incorporate a sufficiently large data set at the individual provider level for greater granularity for the provider rating system. Additionally, a randomized, controlled study design which includes both FQHC and non-FQHC populations would strengthen the validation of the PRSSQ survey and of the value model.

**Significance and Implications:**

This study allowed for the use of an evidence- and value-based approach to assess provider value in a standardized format. With some potential modifications to allow for this system to be used at a single provider level, these findings will allow for the instrumentation of an evidence-based modality for ranking providers by value, which can inform patients in their decision of where to seek care, motivate providers to increase value, and provide data so payers might identify high-value providers.

**CONCLUSION:**

Based on the results of this study, the following conclusions can be made:

1. An objective, value- and evidence-based rating system of providers, by geographic region and treatment facility, can assist:
   a. Patients to better compare and select providers.
   b. Providers to benchmark their comparative ratings on an internal and external basis for continuous quality improvement.
c. Third-party payers, such as insurance companies and government programs, to better compare outcomes and costs for provider systems and clinic facilities.

ACKNOWLEDGMENTS

The authors extend appreciation and acknowledgments to Jay Balzer, DDS, MPH, a dental public health specialist and reviewer, and the former NYU Langone Hospitals-Advanced Education in Pediatric Dentistry residents who supported the study as research-researchers: Jared Cattron, DMD, Danielle Fernandez, DMD, Shanti Gopalan, DDS, Mehedia Haque, DMD, MPH, Mariam Javaid, DMD, Stephany Liu, DDS, Andrea Lochan, DMD, Jeremy Morris, DDS, Olanrewaju Oye-Somefun, DDS, Andrea Salazar, DDS, Erin Saucier, DMD, Silpa Velivela, DMD, and Pavel Zeylikman, DMD.

REFERENCES:

   https://catalyst.nejm.org/doi/full/10.1056/CAT.17.0558

doi:10.1001/jamasurg.2016.4993

   https://hbr.org/2019/11/designing-better-online-review-systems


   https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/care-
compare-dac-initiative


    Statistical Software, 69(1), 1-33. doi:10.18637/jss.v069.i01
27. Okuji D, Jivraj A, Wu Y. From Volume to Value: Comparative value, quality, and cost
    between three treatment modalities for early childhood caries. Published online
    Aff (Millwood). 2008;27(3):759-769. doi:10.1377/hlthaff.27.3.759
29. Werner R, Emanuel E, Pham H, Navathe A. The Future of Value-Based Payment: A
    Road Map to 2030. Issue Briefs. Published online February 18, 2021. Accessed April
    16, 2023. https://repository.upenn.edu/ldi_issuebriefs/149
    of Dental Care Using Patient Online Reviews in the United States: Mixed Methods
    doi:10.2196/18652
31. Riordain RN, Glick M, Mashhadani SSAA, et al. Developing a Standard Set of Patient-
    centered Outcomes for Adult Oral Health - An International, Cross-disciplinary


Table 1. Summary of sample characteristics and comparisons by treatment modalities

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Surgical with GA</th>
<th>Surgical without GA</th>
<th>Non-Surgical</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>487</td>
<td>260</td>
<td>156</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>age (mean (sd))</td>
<td>3.99 (1.22)</td>
<td>3.94 (1.18)</td>
<td>4.55 (1.08)</td>
<td>2.94 (0.90)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gender = male (%)</td>
<td>243 (49.9)</td>
<td>137 (52.7)</td>
<td>68 (43.6)</td>
<td>38 (53.5)</td>
<td>0.16</td>
</tr>
<tr>
<td>Ethnicity = Hispanic (%)</td>
<td>272 (56.2)</td>
<td>119 (45.8)</td>
<td>104 (68.0)</td>
<td>49 (69.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>White</td>
<td>135 (27.7)</td>
<td>74 (28.5)</td>
<td>54 (34.6)</td>
<td>7 (9.9)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>80 (16.4)</td>
<td>47 (18.1)</td>
<td>26 (16.7)</td>
<td>7 (9.9)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>21 (4.3)</td>
<td>10 (3.8)</td>
<td>5 (3.2)</td>
<td>6 (8.5)</td>
<td></td>
</tr>
<tr>
<td>Native Hawaiian/Other Pacific Islander</td>
<td>31 (6.4)</td>
<td>31 (11.9)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>American Indian/Alaskan Native</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>220 (45.2)</td>
<td>98 (37.7)</td>
<td>71 (45.5)</td>
<td>51 (71.8)</td>
<td></td>
</tr>
<tr>
<td>Payer source (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Federal Medicaid</td>
<td>462 (94.9)</td>
<td>254 (97.7)</td>
<td>137 (87.8)</td>
<td>71 (100.0)</td>
<td></td>
</tr>
<tr>
<td>CHIP</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Employer paid commercial insurance</td>
<td>1 (0.2)</td>
<td>1 (0.4)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Individually paid commercial insurance</td>
<td>1 (0.2)</td>
<td>0 (0.0)</td>
<td>1 (0.6)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Self-pay</td>
<td>22 (4.5)</td>
<td>4 (1.5)</td>
<td>18 (11.5)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1 (0.2)</td>
<td>1 (0.4)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>ASA¹ = 2 (%)</td>
<td>61 (12.5)</td>
<td>42 (16.2)</td>
<td>17 (10.9)</td>
<td>2 (2.8)</td>
<td>0.008</td>
</tr>
<tr>
<td>Sum ECOHIS² (mean (sd))</td>
<td>57.45 (7.61)</td>
<td>55.59 (7.10)</td>
<td>58.19 (8.57)</td>
<td>62.66 (3.25)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Average ECOHIS² (mean (sd))</td>
<td>4.42 (0.58)</td>
<td>4.28 (0.55)</td>
<td>4.48 (0.66)</td>
<td>4.82 (0.25)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sum D-CAHPS³ (mean (sd))</td>
<td>40.30 (4.52)</td>
<td>40.07 (3.92)</td>
<td>39.13 (5.24)</td>
<td>43.68 (3.09)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Average D-CAHPS³ (mean (sd))</td>
<td>4.48 (0.50)</td>
<td>4.45 (0.44)</td>
<td>4.35 (0.58)</td>
<td>4.85 (0.34)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sum C-OIDP⁴ (mean (sd))</td>
<td>19.30 (2.31)</td>
<td>19.63 (1.62)</td>
<td>18.53 (3.32)</td>
<td>19.79 (1.01)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Average C-OIDP⁴ (mean (sd))</td>
<td>4.83 (0.58)</td>
<td>4.91 (0.40)</td>
<td>4.63 (0.83)</td>
<td>4.95 (0.25)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total PRSSQ⁵ (mean (sd))</td>
<td>117.05 (10.48)</td>
<td>115.28 (8.22)</td>
<td>115.85 (13.21)</td>
<td>126.13 (5.20)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>** Average PRSSQ⁵ (mean (sd))</td>
<td>4.57 (0.37)</td>
<td>4.55 (0.28)</td>
<td>4.49 (0.49)</td>
<td>4.87 (0.20)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cost care (mean (sd))</td>
<td>381.91 (343.57)</td>
<td>635.61 (275.75)</td>
<td>110.80 (99.58)</td>
<td>48.59 (26.33)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RVUs (mean (sd))</td>
<td>29.77 (21.06)</td>
<td>46.15 (13.02)</td>
<td>12.49 (11.36)</td>
<td>7.76 (3.59)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>† Cost per RVU⁶ (mean (sd))</td>
<td>12.66 (9.60)</td>
<td>14.23 (6.22)</td>
<td>12.71 (14.09)</td>
<td>6.82 (3.68)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‡ Value score (mean (sd))</td>
<td>58.19 (56.92)</td>
<td>40.82 (30.08)</td>
<td>70.45 (70.61)</td>
<td>94.83 (72.76)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Ln-value score (mean (sd))</td>
<td>3.79 (0.69)</td>
<td>3.57 (0.49)</td>
<td>3.90 (0.84)</td>
<td>4.40 (0.52)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Notes:
Means (standard deviations) are presented for continuous variables; frequencies (%) are presented for categorical variables.
* p < 0.001 by the Kruskal-Wallis test.
** Quality = Average PRSSQ, with higher score indicating higher quality.
GA=General Anesthesia.
† Cost = Cost per RVUs, with lower score indicating lower cost
‡ Value = Value score = 100 times Quality (** divided by Cost (†), with higher score indicating higher value
α ECOHIS: Early Childhood Oral Health Impact Score
β D-CAHPS: Dental Consumer Assessments of Healthcare Providers and Systems
χ C-OIDP: Child-Oral Impacts on Daily Performance
δ PRSSQ: parental-reported symptom and service quality questionnaire
ε RVU: Relative Value Unit
θ CHIP: Children’s Health Insurance Program
λ ASA: American Society of Anesthesiologists Physical Status Classification System
Table 2. Comparisons of value scores (median [IQR]) by states. Sensitivity analysis based on the Dunn’s Kruskal-Wallis multiple comparison test.

<table>
<thead>
<tr>
<th>State</th>
<th>N</th>
<th>Value (median [IQR])</th>
<th>Letter Grade Rating*</th>
<th>N</th>
<th>Value (median [IQR])</th>
<th>Letter Grade Rating*</th>
<th>N</th>
<th>Value (median [IQR])</th>
<th>Letter Grade Rating*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td>16</td>
<td>48.5 [37.6, 54.6]</td>
<td>A</td>
<td>34</td>
<td>66.86 [50.45, 113.34]</td>
<td>A(-)**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hawaii</td>
<td>35</td>
<td>33.6 [28.9, 40.0]</td>
<td>B</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>29</td>
<td>27.2 [23.1, 31.6]</td>
<td>C</td>
<td>33</td>
<td>24.20 [20.77, 28.49]</td>
<td>E</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Maryland</td>
<td>29</td>
<td>27.0 [22.7, 32.7]</td>
<td>C</td>
<td>23</td>
<td>140.68 [78.97, 187.01]</td>
<td>A</td>
<td>14</td>
<td>108.77 [86.44, 152.66]</td>
<td>A</td>
</tr>
<tr>
<td>New York</td>
<td>85</td>
<td>32.8 [24.0, 51.4]</td>
<td>B</td>
<td>23</td>
<td>53.92 [42.66, 86.95]</td>
<td>B(-)**</td>
<td>57</td>
<td>73.11 [56.68 101.30]</td>
<td>B</td>
</tr>
<tr>
<td>Tennessee</td>
<td>39</td>
<td>54.5 [47.9, 62.3]</td>
<td>A</td>
<td>18</td>
<td>45.50 [26.50, 73.96]</td>
<td>C(-)**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

p-value*   | <0.001 | <0.001 | 0.004

IQR: inter-quartile range.
* overall test difference.
GA=General Anesthesia.
α: A(-) is not significantly different than A or B.
β: B(-) is not significantly different than B or C.
χ: C(-) is not significantly different than C or D.
δ: D(-) is not significantly different than D or E.
Table 3. Comparisons of value scores (median [IQR]) by facilities. Sensitivity analysis based on the Dunn’s Kruskal-Wallis multiple comparison test.

<table>
<thead>
<tr>
<th>State-Facility Number</th>
<th>Surgical with GA</th>
<th>Surgical without GA</th>
<th>Non-surgical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Value (median [IQR])</td>
<td>Letter Grade Rating*</td>
</tr>
<tr>
<td>AZ-1</td>
<td>16</td>
<td>48.53 [37.55, 54.55]</td>
<td>A(-)α</td>
</tr>
<tr>
<td>HI-1</td>
<td>34</td>
<td>32.90 [28.85, 38.89]</td>
<td>C</td>
</tr>
<tr>
<td>HI-2*</td>
<td>1</td>
<td>174.3 [174.3, 174.3]</td>
<td>-</td>
</tr>
<tr>
<td>FL-1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FL-2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FL-3</td>
<td>5</td>
<td>21.11 [15.58, 22.15]</td>
<td>D</td>
</tr>
<tr>
<td>FL-4</td>
<td>24</td>
<td>27.37 [24.27, 31.95]</td>
<td>C(-)χ</td>
</tr>
<tr>
<td>MA-1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MA-2</td>
<td>27</td>
<td>27.19 [23.05, 31.55]</td>
<td>C(-)χ</td>
</tr>
<tr>
<td>MD-1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MD-2</td>
<td>23</td>
<td>26.39 [21.96, 31.27]</td>
<td>D</td>
</tr>
<tr>
<td>MD-3</td>
<td>6</td>
<td>34.56 [28.20, 37.95]</td>
<td>C(+/-)ε</td>
</tr>
<tr>
<td>NY-1</td>
<td>41</td>
<td>40.44 [31.58, 73.21]</td>
<td>A(-)α</td>
</tr>
<tr>
<td>NY-2</td>
<td>44</td>
<td>26.79 [20.51, 33.42]</td>
<td>C(-)χ</td>
</tr>
<tr>
<td>TN-1*</td>
<td>1</td>
<td>320.3 [320.3, 320.3]</td>
<td>-</td>
</tr>
<tr>
<td>TN-2</td>
<td>38</td>
<td>54.44 [47.49, 61.45]</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>0.004</td>
</tr>
</tbody>
</table>

* facilities with n <5 did not enter statistical testing
**overall test of difference.
GA=General Anesthesia.
α: A(-) is not significantly different than A or B.
β: B(-) is not significantly different than B or C.
ε: C(+/-) is not significantly different than B or D.
χ: C(-) is not significantly different than C or D.
δ: D(-) is not significantly different than D or E.
Early Childhood Oral Health Impact Scale (Adapted from ECOHIS)

*Problems with the teeth, mouth or jaws and their treatment can affect the well-being and everyday lives of children and their families. For each of the following questions please circle the number next to the response that best describes your child’s experiences or your own. Consider the child’s entire life from the last dental treatment visit until now when answering each question.*

**Response options:** 0=Don’t know; 1=Very often; 2=Often; 3=Occasionally; 4=Hardly ever; 5=Never

1. How often has your child had pain in the teeth, mouth or jaws? (Child symptom domain)
2. How often has your child had difficulty drinking hot or cold beverages? (Child function domain)
3. How often has your child had difficulty eating some foods? (Child function domain)
4. How often has your child had difficulty pronouncing any words? (Child function domain)
5. How often has your child missed preschool, daycare or school? (Child function domain)
6. How often has your child had trouble sleeping because of dental problems or dental treatments? (Child psychological domain)
7. How often has your child been irritable or frustrated because of dental problems or dental treatments? (Child psychological domain)
8. How often has your child avoided smiling or laughing when around other children because of dental problems or dental treatments? (Child self-image/social interaction domain)
9. How often has your child avoided talking with other children because of dental problems or dental treatments? (Child self-image/social interaction domain)
10. How often have you or another family member been upset because of your child’s dental problems or dental treatments? (Parent distress domain)
11. How often have you or another family member felt guilty because of your child’s dental problems or dental treatments? (Parent distress domain)
12. How often have you or another family member taken time off from work because of your child’s dental problems or dental treatments? (Family function domain)
13. How often has your child had dental problems or dental treatments that had a financial impact on your family? (Family function domain)

Patient Satisfaction Questionnaire (PSQ - Adapted from RAND Health Insurance Experiment)

The following questions include some things people say about dental care. Please read each one carefully, keeping in mind the dental care you are receiving now. (If you have not received care recently, think about what you would expect if you needed care today.) We are interested in your feelings about the dental care you have received.

**Response options:** 0=Don’t know; 1=Strongly disagree; 2=Disagree; 3=Uncertain; 4=Agree; 5=Strongly agree

**How strongly do you AGREE or DISAGREE with each of the following statements?**

1. The dental care my child is receiving is just about perfect.
2. My child’s dentist treats me in a very friendly and courteous manner.
3. Those who provide my child’s dental care sometimes hurry too much when they treat my child.
4. My child’s dentist’s office sometimes ignores what I tell them.
5. My child’s dentist usually spend plenty of time with my child.
6. I find it hard to get an appointment for my child’s dental care right away.
7. I am dissatisfied with some things about the dental care my child receives.
8. I would recommend my child’s dentist to relatives and friends.
9. I would recommend my child’s dentist’s clinic to relatives and friends.

Parental-Reported Oral Health (Adapted from Oral Impacts Daily Performance survey-OIDP)

Please answer the next set of questions which describes your child’s dental condition since the last dental treatment visit, including today.

**Response options:** 0=Yes; 1=I don’t know; 5=No

1. My child is experiencing pain of the teeth, mouth or face.
2. My child is experiencing swelling of the teeth, mouth or face.
3. My child is experiencing broken or lost fillings or crowns of the teeth.
4. My child is experiencing new cavities of the teeth since his/her last dental treatment visit.

Figure 1. Post-treatment Parental-Reported Symptom and Service quality Questionnaire (PRSSQ), 26-questions, English version.
Figure 2. Boxplots of ln-transformed value score by state. Tukey's multiple comparison test was used.

*Treatment modality group

**(Letter grade) [mean ln-value score (standard deviation)]

GA=General Anesthesia.
A(-) is not significantly different than A or B.
B(-) is not significantly different than B or C.
Figure 3. Boxplots of ln-transformed value score by facilities. Tukey’s multiple comparison test was used.

*Treatment modality group
**(Letter grade) [mean ln-value score (standard deviation)]
GA=General Anesthesia.
B(-) is not significantly different than B or C.
C(-) is not significantly different than C or D.
D(-) is not significantly different than D or E.