Published in final edited form as:

Community Dent Oral Epidemiol. 2013 August ; 41(4): 364–373. doi:10.1111/cdoe.12016.

Demographic, socioeconomic, and behavioral factors affecting patterns of tooth decay in the permanent dentition: Principal components and factor analyses

John R. Shaffer^{A,B}, Deborah E. Polk^{B,C,D}, Eleanor Feingold^{A,B,E}, Xiaojing Wang^{B,F,G}, Karen T. Cuenca^{A,B,C,F,G}, Daniel E. Weeks^{A,E}, Rebecca S. DeSensi^{B,F,G}, Robert J. Weyant^{B,C}, Richard Crout^{B,H}, Daniel W. McNeil^{I,J}, and Mary L. Marazita^{A,B,F,G,K,L}

^ADepartment of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, 15261

^BCenter for Oral Health Research in Appalachia, University of Pittsburgh, Pittsburgh, PA, 15261 and West Virginia University, Morgantown, WV, 26506

^CDepartment of Dental Public Health, University of Pittsburgh, School of Dental Medicine, Pittsburgh, PA, 15261

^DDepartment of Behavioral and Community Health Sciences, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, 15261

^EDepartment of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, 15261

^FCenter for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, 15219

^GDepartment of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, 15261

^HDepartment of Periodontics, West Virginia University School of Dentistry, Morgantown, WV, 26506

^IDepartment of Psychology, Eberly College of Arts and Science, West Virginia University, Morgantown, WV, 26506

^JDental Practice and Rural Health, School of Dentistry, West Virginia University, Morgantown, WV, 26506

^KClinical and Translational Science Institute, School of Medicine, University of Pittsburgh, Pittsburgh, PA

^LDepartment of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA

Abstract

Corresponding author: John R. Shaffer, Ph.D., Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, 130 DeSoto St., Pittsburgh, PA 15261, jrs51@pitt.edu.
jrs51@pitt.edu; dpolk@pitt.edu; feingold@pitt.edu; xiw23@pitt.edu; ktc14@pitt.edu; weeks@pitt.edu; shue@pitt.edu; rjw1@pitt.edu; rcrouth@hsc.wvu.edu; daniel.mcneil@mail.wvu.edu; marazita@pitt.edu

Authors' contributions:

JRS conceived and designed this study. RJW, RC, DWM, and MLM conceived and designed the COHRA initiative. JRS analyzed the data. JRS, EF, WX, KT, DEW, RSD, and MLM managed, cleaned and quality checked the data. JRS, DEP, EF, WX, KT, DEW, RSD, SW, RJW, RC, DWM, and MLM interpreted the results. JRS wrote the manuscript with significant input from DEP. JRS, DEP, EF, XW, KT, DEW, RSD, SW, RJW, RC, DWM, and MLM read, revised and approved the manuscript.

Objectives—Dental caries of the permanent dentition is a multi-factorial disease resulting from the complex interplay of endogenous and environmental risk factors. The disease is not easily quantified due to the innumerable possible combinations of carious lesions across individual tooth surfaces of the permanent dentition. Global measures of decay, such as the DMFS index (which was developed for surveillance applications), may not be optimal for studying the epidemiology of dental caries because they ignore the distinct patterns of decay across the dentition. We hypothesize that specific risk factors may manifest their effects on specific tooth surfaces leading to patterns of decay that can be identified and studied. In this study we utilized two statistical methods of extracting patterns of decay from surface-level caries data in order to create novel phenotypes with which to study the risk factors affecting dental caries.

Methods—Intra-oral dental examinations were performed on 1,068 participants aged 18 to 75 years to assess dental caries. The 128 tooth surfaces of the permanent dentition were scored as carious or not and used as input for principal components analysis (PCA) and factor analysis (FA), two methods of identifying underlying patterns without *a priori* knowledge of the patterns. Demographic (age, sex, birth year, race/ethnicity, and educational attainment), anthropometric (height, body mass index, waist circumference), endogenous (saliva flow), and environmental (tooth brushing frequency, home water source, and home water fluoride) risk factors were tested for association with the caries patterns identified by PCA and FA, as well as DMFS, for comparison. The ten strongest patterns (i.e., those that explain the most variation in the data set) extracted by PCA and FA were considered.

Results—The three strongest patterns identified by PCA reflected (i) global extent of decay (i.e., comparable to DMFS index), (ii) pit and fissure surface caries, and (iii) smooth surface caries, respectively. The two strongest patterns identified by FA corresponded to (i) pit and fissure surface caries and (ii) maxillary incisor caries. Age and birth year were significantly associated with several patterns of decay, including global decay/DMFS index. Sex, race, educational attainment, and tooth brushing were each associated with specific patterns of decay, but not with global decay/DMFS index.

Conclusions—Taken together, these results support the notion that caries experience is separable into patterns attributable to distinct risk factors. This study demonstrates the utility of such novel caries patterns as new outcomes for exploring the complex, multifactorial nature of dental caries.

Keywords

dental caries; permanent dentition; pit and fissure surfaces; smooth surfaces; tooth surfaces; principal components analysis; factor analysis; tooth brushing

Introduction

Dental caries of the permanent dentition is a multi-factorial disease affecting a majority of adolescents and adults, worldwide. Many factors influence the risk of developing dental caries, including environmental agents such as bacterial flora and fluoride exposure; behavioral factors including diet and oral hygiene; endogenous features such as tooth position and morphology, enamel composition, saliva composition and flow rate; and demographic characteristics such as age, sex, race, ethnicity, socio-economic status, and access to oral health care (1). However, known predictors account for only a fraction of the variance in caries experience. The etiology of dental caries is further complicated by the fact that risk of decay is non-uniform across tooth surfaces of the complete permanent dentition (2, 3). In the primary dentition, specific patterns of tooth decay have been identified and attributed to specific risk factors. A notable example is the pattern of anterior maxillary tooth decay seen in young children (i.e., previously known as “baby bottle caries”) due, in part,

to feeding habits (4-8). Analogous patterns of decay may exist in the permanent dentition. We hypothesize that specific risk factors may manifest their effects on specific tooth surfaces leading to distinct patterns of decay that can be identified and studied. Permanent dentition caries experienced in an individual may therefore represent the cumulative result of multiple superimposed decay patterns, each due to specific risk factors over the life course (9).

Another challenge in understanding the etiology of dental caries is the complexity of the disease state: innumerable manifestations of caries across the dentition cannot be easily or succinctly quantified by a single metric. Common outcomes used in studying the epidemiology of dental caries, such as DMFT and DMFS indices (*i.e.*, totals of the number of decayed, missing, or filled, teeth/surfaces) do not distinguish between distinct decay patterns, and hence may be insufficient for investigating the risk factors that lead to specific patterns of decay. Therefore methods that separate caries experience into components attributable to distinct etiologies may be necessary to identify risk factors with modest effect sizes, and/or to fully understand the non-uniform effects across the permanent dentition for risk factors with moderate-to-large effect sizes.

Previous descriptive or epidemiological studies of caries patterns have been reported, usually based on *a priori* assumptions regarding the combinations of tooth surfaces that define the patterns of interest (4-6, 10-16). Groupings of tooth surfaces and/or *a priori* patterns differed across studies, so combined interpretation of results across studies is not possible. Moreover, there is little evidence that *a priori* caries patterns best represent underlying disease etiologies. In contrast, a few studies have used agnostic methods to successfully detect the patterns of tooth decay in children without relying on *a priori* surface classifications (3, 7, 8).

Agnostic methods of extracting caries patterns from surface level caries data may be useful for understanding the action of caries risk factors and may complement analyses of traditional caries outcomes such as DMFS index. As part of ongoing efforts by the Center for Oral Health Research in Appalachia (17), we have previously assessed the heritability of patterns of dental decay in the permanent dentition in this data set using two related methods: principal components analysis (PCA) and factor analysis (FA) (9). The results indicated that genes cumulatively accounted for 30% to 65% of the variation in some patterns of decay, whereas other patterns of decay were not due to genetic factors (9). These results are consistent with the view that specific patterns of decay have distinct etiologies. In the current study we test associations between these factors/components and a wide range of demographic, environmental, and behavioral factors and demonstrate that we are able to detect associations that are not evident when only global caries scores such as DMFS index are used.

Methods

Sample recruitment and data collection

The Center for Oral Health Research in Appalachia is a joint initiative between The University of Pittsburgh and West Virginia University to study the individual-, familial-, and community-based factors related to oral health outcomes in the Appalachian population. Participants were recruited by household, with eligible households including at least one biological parent-child pair. All members of eligible households were invited to participate. Altogether, 732 households were recruited, comprising 2,663 participants. Written informed consent was provided by adult participants and by parents/guardians on behalf of child participants. Data collection efforts were approved by the Center's research committee and

the Institutional Review Boards of respective Universities. Details of the data collection have previously been described (9, 17-21).

Data on dental caries experience were collected during intra-oral examinations by dentists or research dental hygienists (calibrated at least annually to a reference dentist). Each tooth surface was scored as sound, pre-cavitated caries, cavitated caries, restored, missing due to caries, or missing due reasons other than caries by visual inspection including the use of a dental explorer in accordance with the National Center for Health Statistics Dental Examiners Procedures Manual (Section 4.9.1.3) to maximize comparability with other national data sets. Third molars were excluded from caries assessment. Edentulous individuals excluded from the current analysis.

Data on a number of potential dental caries risk factors were collected. Self-reported sex (male, female), race (white, Asian, black, Pacific Islander, American Indian, bi- or multi-racial, or other), ethnicity (non-Hispanic, Hispanic), age (years), birth year (date), educational attainment (none, high school diploma, technical school/associate degree, some college but no degree, undergraduate degree, advanced degree), tooth brushing frequency (times/day), and home water source (city/public, well) were obtained via questionnaire and/or interview. Height (m), weight (kg), and waist circumference (cm; measured above the iliac crest) were obtained by physical examination. Body mass index (BMI; kg/m^2) was calculated from measured height and weight. Unstimulated saliva flow rate (ml/min) was measured by expectoration into a vial for three minutes. Home water source fluoride level (mg/L) was measured in water samples using a fluoride-specific electrode (Model #9609, Orion Research Inc., Boston, MA).

In the present study, which aimed to explore the role of risk factors on patterns of decay in the permanent dentition, we limited our study sample to 1,068 participants aged 18 to 75 years. Tooth surface-level data on 128 surfaces (i.e., buccal, distal, lingual, and mesial surfaces on all 28 permanent teeth, plus occlusal surfaces on pre-molars and molars, excluding third molars) were coded as 0 for sound or missing due to reasons other than decay, or coded as 1 for pre-cavitated, decayed, missing due to decay, or restored. Using this coding scheme, we generated a matrix of 1,068 participants by 128 surface-level caries affection statuses. This matrix was used as input for two statistical methods of identifying patterns within the data: principal components analysis (PCA) and factor analysis (FA) (9, 22). These methods have been used for over 100 years (with varying degrees of success) for numerous applications.

Pattern detection methods

The methods used to extract and test factor/components are described in full in Shaffer et al. (9), but we briefly summarize them here. PCA is a method of extracting uncorrelated variables (called principal components scores, PCs) using an orthogonal transformation. Each PC is calculated as a linear combination of the original variables weighted by their loadings. The loadings are generated such that the first PC (i.e., PC1) explains the largest possible amount of variability in the data in a single dimension, the second PC (i.e., PC2) explains the largest possible amount of *remaining* variability in the data in a single dimension orthogonal to PC1, the third PC (i.e., PC3) explains the largest possible amount of *remaining* variability in the data in a single dimension orthogonal to *both* PC1 and PC2, and so on. The output from this method is a set of PCs (equal in number to the original variables, in this case, 128), with each successive PC explaining less and less of the data variability. This method can be used to assess underlying patterns from a number of correlated phenotypic measurements, such as surface-level caries data in this case. The first several PCs, which explain the largest independent components of variability in the data, may represent important patterns in the data. The remaining PCs, which explain the smallest

independent components of variability in the data, may represent negligible patterns and/or noise in the data. The loadings provide a way of mapping PCs back to the original variables, and can assist in interpreting the patterns extracted by this method.

FA is similar to PCA in that it extracts latent variables called factor scores (FACs) from a number of measured correlated variables. The premise is that numerous measured variables may reflect the variation in a reduced number of unobserved variables. Unlike in PCA, the number of FACs is pre-determined by the investigator, and is less than the number of original variables because error terms are used to account for unexplained variation. Moreover, the FACs are not constrained to be orthogonal. As with PCs, the FACs are calculated as linear combinations of the original variables weighted by their loadings, thus the loadings provide means of interpreting the FACs in terms of the original variables. In many cases, FA and PCA perform similarly; however, for some data sets they perform differently, largely because they assume opposite perspectives in extracting patterns: PCA assumes that the patterns are based on the observed variables, whereas FA assumes that the observed variables are based on the underlying patterns. In this sense, FACs may better represent underlying caries “endophenotypes” (i.e., unmeasured phenotypes that manifest as decay patterns), whereas PCA may more appropriately be used for dimension reduction (i.e., summarizing a large number of variables with a few variables). For both PCA and FA, the loadings define a particular caries pattern, whereas the PC/FAC quantifies the severity of dental caries of that particular pattern.

Statistical analysis

The data matrix of 1,068 participants by 128 tooth surface-specific caries affection statuses was used as input for PCA and FA. The first 10 PCs and exactly 10 FACs (an arbitrary number chosen *a priori*) were considered as caries pattern outcomes in subsequent analysis. For comparison to the PCs and FACs, we also generated the commonly-used DMFS index, calculated as the number of pre-cavitated, decayed, missing due to decay, or filled/restored surfaces, summed across the complete permanent dentition.

In order to help distinguish caries patterns relevant to disease etiology from sporadic patterns due to noise, and in order to assess the stability of patterns identified by PCA and FA, we performed a sensitivity analysis (i.e., delete-d jackknifing) by repeating PCA and FA on ten random subsets of the data comprised of 80% of the participants from the full sample (i.e., delete-d jackknifing where $d = 214$, 20% of the total sample). We used Pearson correlations to compare PCs and FACs extracted from the random subsets to those from the full sample. Similarly, because our sample contains related individuals, we repeated our PCA and FA using the maximal set of unrelated individuals, which yielded very similar PCs and FACs indicating that the methods of pattern extraction are robust to the inclusion of relatives. Details regarding the sensitivity analysis have been previously published (9).

To assess the association between caries patterns and risk factors, Pearson correlations were calculated between PCs/FACs and caries risk factors. For statistical analysis, sex, race, ethnicity, home water source, and tooth brushing were treated as binary categorical variables, with race dichotomized as white vs. non-white, and tooth brushing dichotomized as twice or more daily vs. one or fewer times daily. Age, age², birth year, weight, BMI, waist circumference, saliva flow, and home water source fluoride were treated as continuous variables. Educational attainment was treated as an ordinal trait taking values of 0 to 5 for none, high school diploma, technical school/associate degree, some college, undergraduate degree, and advanced degree, respectively. Linear regression was used to model the effects of risk factors on PCs/FACs while adjusting for age and sex (which are widely-accepted caries risk factors for which data were available on all participants). Because the distribution of DMFS index was very non-Normal, we performed non-parametric (Spearman correlation)

tests of association between DMFS index and caries risk factors. Due to the fact that the sample contains biologically related participants, significance of associations was assessed with and without conditioning on the family structure within the sample. For all association tests, unadjusted p-values and family-adjusted p-values were nearly identical and in no cases meaningfully altered interpretation of the results. Bonferroni adjustment for 21 caries outcomes (i.e., DMFS plus 10 PCs and 10 FACs; $\alpha = 0.00238$) was used to conclude statistical significance.

Statistical analyses were primarily performed using the R software package (R Foundation for Statistical Computing, Vienna, AU). Principal components analysis was performed by singular value decomposition on the data matrix after mean-centering the original variables, as implemented in the *prcomp* function using default parameters. Factor analysis was performed using the Thomson's regression-based method (23) of calculating FACs for 10 factors as implemented in the *factanal* function using all other default parameters. Pearson and Spearman correlation coefficients and tests of association were calculated using the *cor.test* function, and linear regression was performed as implemented in the *lm* function. Statistical significance of association, while conditioning on the family structures within the data, was assessed using variance decomposition methods as implemented in SOLAR (24).

Results

Descriptive statistics of the study sample are shown in Table 1. The sample included 1068 adults (ages 18 to 75 years; mean age 34.7; 63.3% female), who were mostly self-reported white (89.9%) and mostly non-college graduates (84.3%) from rural communities in Northern Appalachia. This sample came from a population that, overall, has lower income, more limited access to oral health care, and poorer oral health compared to the general US population. Table 2 shows the surface-specific prevalences of caries organized by dental quadrant. As expected, surface-specific caries prevalences were highest for pit and fissure surfaces, including the occlusal surfaces of premolars and molars, lingual surfaces of maxillary molars, and buccal surfaces of mandibular molars. Surface-specific caries prevalences of approximal (i.e., mesial and distal) surfaces of the premolars and molars, with the exception of mandibular 1st premolars, were higher than most other smooth surfaces. Surface-specific caries prevalences were higher for maxillary anterior tooth surfaces (i.e., those of the incisors and canines) than for mandibular anterior tooth surfaces.

Surface-level data for the 128 tooth surfaces of the permanent dentition were analyzed using two methods of extracting the underlying patterns of caries: PCA and FA. For each of the two methods, the 10 patterns of decay explaining the most variation in the data (i.e., PC1 to PC10 and FAC1 to FAC10) were investigated. Cumulatively, PCs 1-10 and FACs 1-10 accounted for 53.2% and 44.7% of variation in the data, respectively. Details of the loadings for PCs and FACs, which facilitate interpretation of caries patterns in terms of the original surface-level variables, have previously been described in detail (9). Table 3 lists the major contributing teeth/surfaces for PCs 1-10 and FACs 1-10, providing general interpretations of these patterns of dental decay. For example, PC1 is attributable to near-uniform contributions of all maxillary tooth surfaces as well as mandibular premolar and molar surfaces, and therefore roughly represents the global extent of tooth decay. Indeed, PC1 is highly correlated with DMFS index (correlation coefficient, $r = 0.97$). We have previously shown that PC1, PC2 and PC3 closely recaptures variation attributable to global, pit and fissure, and smooth surface caries, respectively (results not shown; see (9)). Sensitivity analysis comparing PCs/FACs extracted from the full data set to those extracted from random sub-sets showed that most patterns were stable (Table 3).

Associations between caries patterns and potential risk factors are summarized in Table 4 (with and without adjustment for sex and age; statistical significance determined at Bonferroni adjustment for 21 outcomes). Age, age², and birth year were significantly associated with DMFS index, as well as several PCs/FACs, which is expected considering the continued action of cariogenic processes over the lifespan. Other demographic variables showed significant associations with specific patterns of decay, such as the effect of sex on PC3, the effects of race on PC4, PC8, FAC6, and FAC10, and the effects of educational attainment on PC2, PC4, FAC1, FAC2, and FAC6. These risk factors were not significantly associated with DMFS index (p-values > 0.05). Hispanic ethnicity was not significantly associated with any caries outcomes, probably due to lack of power given the small number of Hispanic individuals in this sample. Likewise, anthropometric variables (height, weight, and waist circumference), saliva flow, water fluoride, and public water source were not associated with any specific caries outcomes. Tooth brushing was significantly associated with PC3, FAC2, and FAC4, but not DMFS index.

Discussion

This study employed two methods of extracting patterns of tooth decay from surface-level caries data for the purpose of studying the environmental risk factors of dental caries. This approach afforded detection of significant associations that would otherwise have gone unrecognized if we had considered DMFS index alone. Specifically, we identified associations between separate decay patterns and sex, race, educational attainment, and tooth brushing frequency. These results are consistent with the premise that overall caries experience can be partitioned into patterns of decay with distinct (though not necessarily mutually exclusive) risk factors.

One of the major findings from this study is the increased utility of modeling patterns of decay over global caries outcomes (e.g., DMFS index) to identify risk factors. Surface-level caries data can be easily collected and analyzed using PCA and/or FA. Both of these statistical methods are well studied, widely used in many fields, and easily implemented in most data analysis software, and they produce results that are readily accessible to dental researchers. Indeed, PCA and/or FA have been previously used for epidemiological studies of tooth coloration, dentition morphology, dental education, gender identification, dental health behaviors, and more. Adding to this list, we therefore propose the use of agnostically-identified decay patterns to complement *a priori* caries outcomes (e.g., DMFS index) in epidemiological studies of dental caries. This approach may be particularly appropriate for identifying risk factors of modest effect size, such as genetic susceptibility loci (9, 21). Moreover, the distributions of PCs/FACs will typically not exhibit the extreme non-Normality (skewness and zero-inflation) of DMFS, and thus may be better-suited for common statistical analyses that assume data normality.

Particularly interesting findings of our study were the associations between educational attainment and PC2 and FAC6 ($r = 0.19$ and 0.17 , respectively), which are among the strongest unadjusted associations identified in this study aside from age, age², and birth year. The interpretation of PC2, for example, is that for a given level of global decay (*i.e.*, PC1), people with higher educational attainment experience comparatively more dental caries on molar surfaces and comparatively less dental caries on non-molar surfaces than individuals with lower educational attainment. These associations with educational attainment were significant even after adjusting for age and sex (p-values = 10^{-10} and 10^{-7} , respectively), and are consistent with other studies (25-31). It is unclear how educational attainment may affect patterns of tooth decay, however, we speculate that possible mechanisms may relate to socioeconomic status (SES) including access to oral health care or behavioral factors such as dietary choices and oral hygiene. A recent study on the effects of

SES factors on dental caries in Pennsylvanian adolescents showed that parental education was significantly correlated with many SES-related risk factors, including tooth brushing, sealants, and recency of dental visits, but that these factors did not entirely explain the difference in caries prevalence or severity across the SES spectrum (30).

Another noteworthy finding was the significant associations of tooth brushing frequency with PC3 and FAC4 (both caries patterns related to premolars), as well as FAC2 (roughly representing caries of the maxillary incisors). Interestingly, none of these caries patterns were heritable in our previous study (9), which further supports the involvement of environmental risk factors, such as tooth brushing, on these caries patterns. This finding also corroborates a number of recent studies demonstrating the preventative role of tooth brushing on dental caries (32, 33). Despite the widespread promotion of oral hygienic behaviors such as tooth brushing to the greater U.S. population (*e.g.*, through print and television advertisements, in public schools, and by dental health care providers), more than one third of the study population reported brushing their teeth fewer than two times daily. Efforts aimed at improving oral hygiene, including frequent tooth brushing with fluoridated tooth paste, may help in reducing oral health disparities in at-risk populations such as the rural Appalachian population.

As we have shown, agnostic methods of extracting patterns such as PCA and FA can be useful for studying dental caries, a notoriously complex phenotype, though these methods come at the price of reduced interpretability of the results. Making sense of patterns reflective of PCs/FACs can be very challenging. Interpretations of some patterns were clear, such as PC1, which was attributable to near-uniform contributions across most tooth surfaces, and thus described global extent of decay. Many patterns, however, were not so clear. Furthermore, distinguishing biologically-relevant patterns attributable to distinct risk factors from sporadic patterns due to chance alone can be difficult. We attempted to address this issue through sensitivity analysis which demonstrated that PC1 to PC9 and FAC1 to FAC6 were stable (and other PCs/FACs were moderately stable), which suggested that these represent meaningful patterns in the data rather than noise.

Our study profited from several strengths, including the large sample of participants with surface-level caries assessment making modeling caries patterns possible. Moreover, we demonstrated the utility of pattern extraction using two separate but related methods, both of which proved successful in identifying risk factors, especially compared with DMFS index. Lastly, our methods did not rely on *a priori* pattern definitions.

These strengths notwithstanding, important limitations of this study warrant consideration. First, dichotomizing surfaces as carious or not was an over-simplification, which did not take into account lesion severity and may have been blurred by aggressive preventative treatment inflating the observed number of restorations. Likewise, all surfaces of teeth missing due to decay were considered carious, even though lesions causing tooth extraction may have affected fewer surfaces. Similarly, restorations of approximal lesions may have affected adjacent non-carious occlusal surfaces, thus causing erroneous scoring of surfaces. These three sources of error, which may have led to inflated surface-specific prevalences, were offset by potential deficiencies in the visual method of caries detection, which may have under-represented disease. Additionally, our sample was recruited and data was collected at a number of separate clinic sites. This recruitment strategy, while necessary for collecting sufficient number of participants in communities of low population density such as rural Appalachia, may have introduced sample heterogeneity due to unknown differences among recruitment sites. However, we collected data on many of the important covariates that may differ among sites, such as water source fluoride and educational attainment. Other

unknown caries risk factors that differ among sites may exist, and is a limitation of this project.

In conclusion, this study showed the role of several demographic and behavioral factors, notably educational attainment, on patterns of dental caries. This study also served as a lesson that patterns of decay, which provide a more detailed characterization of caries experience than do global measures of decay, may be useful for identifying caries risk factors as a complementary approach to studying DFMS alone. These results are consistent with the notion that overall caries experience can be partitioned into specific patterns of dental decay due to distinct etiologies (including both biological (9) and environmental factors, and the interactions thereof). Although additional work to better characterize the subtleties of the dental caries phenotype is needed, this study shows the importance of considering decay patterns in epidemiological research studies of tooth decay in the permanent dentition.

Acknowledgments

This work would not be possible without the contributions of many individuals. Foremost, on behalf of the Center for Oral Health Research in Appalachia, we would like to extend our sincere thanks to the participants of the study for their contributions toward our mission to understand and improve the oral health of rural Appalachian communities. In addition, we would like to thank the following health care institutions for their involvement in data collection: The University of Pittsburgh, Bradford, Center for Rural Health Practice, Bradford, PA; McKean County Dental Center, Bradford, PA; Cornerstone Care Community Medical and Dental Center, Burgettstown, PA; UPMC Braddock Hospital, Braddock, PA; Camden-on-Gauley Medical Center, WV; Community Health Clinic of Nicholas County, WV; Richwood Area Community Hospital, WV; Summersville Memorial Hospital, WV; and the Webster County Memorial Hospital, WV. We would also like to thank the following organizations for their contributions: the GORGE Connection Rural Health Education Partnership Board, the Webster-Nicholas Rural Health Education Consortium Board, the West Virginia Rural Health Education Partnerships program, the Nicholas and Webster Boards of Education, and the UPMC Braddock Community Advisory Board.

Support for this study was provided by the National Institute of Dental and Craniofacial Research, including grants R01-DE014899 and R03-DE021425. Additional support was provided by the University of Pittsburgh School of Dental Medicine, the West Virginia University School of Dentistry and Eberly College of Arts and Sciences. The content presented herein is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Dental and Craniofacial Research, nor the National Institutes of Health. The funding sources had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Hunter PB. Risk factors in dental caries. *Int Dent J*. Dec; 1988 38(4):211–7. [PubMed: 3063664]
2. Singh KA, Spencer AJ. Relative effects of pre- and post-eruption water fluoride on caries experience by surface type of permanent first molars. *Community Dent Oral Epidemiol*. Dec; 2004 32(6):435–46. [PubMed: 15541159]
3. Batchelor PA, Sheiham A. Grouping of tooth surfaces by susceptibility to caries: a study in 5-16 year-old children. *BMC Oral Health*. Oct 28.2004 4(1):2. [PubMed: 15511295]
4. Douglass JM, Tinanoff N, Tang JM, Altman DS. Dental caries patterns and oral health behaviors in Arizona infants and toddlers. *Community Dent Oral Epidemiol*. Feb; 2001 29(1):14–22. [PubMed: 11153558]
5. O'Sullivan DM, Tinanoff N. Maxillary anterior caries associated with increased caries risk in other primary teeth. *J Dent Res*. Dec; 1993 72(12):1577–80. [PubMed: 8254125]
6. O'Sullivan DM, Tinanoff N. The association of early dental caries patterns with caries incidence in preschool children. *J Public Health Dent*. 1996 Spring;56(2):81–3. [PubMed: 8863291]
7. Psoter WJ, Pendrys DG, Morse DE, Zhang HP, Mayne ST. Caries patterns in the primary dentition: cluster analysis of a sample of 5169 Arizona children 5-59 months of age. *Int J Oral Sci*. Dec; 2009 1(4):189–95. [PubMed: 20690422]

8. Psoter WJ, Zhang H, Pendrys DG, Morse DE, Mayne ST. Classification of dental caries patterns in the primary dentition: a multidimensional scaling analysis. *Community Dent Oral Epidemiol.* Jun; 2003 31(3):231–8. [PubMed: 12752550]
9. Shaffer JR, Feingold E, Wang X, Tcuenko K, Weeks DE, Desensi RS, et al. Heritable patterns of tooth decay in the permanent dentition: principal components and factor analyses. *BMC Oral Health.* 2012; 12:7. [PubMed: 22405185]
10. Douglass JM, Yi W, Xue ZB, Tinanoff N. Dental caries in preschool Beijing and Connecticut children as described by a new caries analysis system. *Community Dent Oral Epidemiol.* Apr; 1994 22(2):94–9. [PubMed: 8205787]
11. Greenwell AL, Johnsen D, DiSantis TA, Gerstenmaier J, Limbert N. Longitudinal evaluation of caries patterns from the primary to the mixed dentition. *Pediatr Dent.* Sep-Oct;1990 12(5):278–82. [PubMed: 2095537]
12. Johnsen DC. Dental caries patterns in preschool children. *Dent Clin North Am.* Jan; 1984 28(1):3–20. [PubMed: 6584344]
13. Johnsen DC, Bhat M, Kim MT, Hagman FT, Allee LM, Creedon RL, et al. Caries levels and patterns in head start children in fluoridated and non-fluoridated, urban and non-urban sites in Ohio, USA. *Community Dent Oral Epidemiol.* Aug; 1986 14(4):206–10. [PubMed: 3461906]
14. Johnsen DC, Schechner TG, Gerstenmaier JH. Proportional changes in caries patterns from early to late primary dentition. *J Public Health Dent.* 1987 Winter;47(1):5–9. [PubMed: 3469410]
15. Johnsen DC, Schubot D, Bhat M, Jones PK. Caries pattern identification in primary dentition: a comparison of clinician assignment and clinical analysis groupings. *Pediatr Dent.* Mar-Apr;1993 15(2):113–5. [PubMed: 8327361]
16. Johnsen DC, Schultz DW, Schubot DB, Easley MW. Caries patterns in Head Start children in a fluoridated community. *J Public Health Dent.* 1984 Spring;44(2):61–6. [PubMed: 6588207]
17. Polk DE, Weyant RJ, Crout RJ, McNeil DW, Tarter RE, Thomas JG, et al. Study protocol of the Center for Oral Health Research in Appalachia (COHRA) etiology study. *BMC Oral Health.* 2008; 8:18. [PubMed: 18522740]
18. Wang X, Shaffer JR, Weyant RJ, Cuenco KT, Desensi RS, Crout R, et al. Genes and Their Effects on Dental Caries May Differ between Primary and Permanent Dentitions. *Caries Res.* 2010; 44(3): 277–84. [PubMed: 20516689]
19. Wendell S, Wang X, Brown M, Cooper ME, DeSensi RS, Weyant RJ, et al. Taste genes associated with dental caries. *J Dent Res.* Nov; 2010 89(11):1198–202. [PubMed: 20858777]
20. Shaffer JR, Wang X, Desensi RS, Wendell S, Weyant RJ, Cuenco KT, et al. Genetic Susceptibility to Dental Caries on Pit and Fissure and Smooth Surfaces. *Caries Res.* Jan 25; 2012 46(1):38–46. [PubMed: 22286298]
21. Shaffer JR, Wang X, Feingold E, Lee M, Begum F, Weeks DE, et al. Genome-wide association scan for childhood caries implicates novel genes. *J Dent Res.* 2011 in press.
22. Wang X, Kammerer CM, Anderson S, Lu J, Feingold E. A comparison of principal component analysis and factor analysis strategies for uncovering pleiotropic factors. *Genet Epidemiol.* May; 2009 33(4):325–31. [PubMed: 19048641]
23. Thompson, GH. *The Factorial Analysis of Human Ability.* London University Press; London: 1951.
24. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet.* May; 1998 62(5):1198–211. [PubMed: 9545414]
25. Acharya S, Pentapati KC, Singh S. Influence of socioeconomic status on the relationship between locus of control and oral health. *Oral Health Prev Dent.* 9(1):9–16. [PubMed: 21594202]
26. Al-Hosani E, Rugg-Gunn A. Combination of low parental educational attainment and high parental income related to high caries experience in pre-school children in Abu Dhabi. *Community Dent Oral Epidemiol.* Feb; 1998 26(1):31–6. [PubMed: 9511839]
27. Boyce WT, Den Besten PK, Stamperdahl J, Zhan L, Jiang Y, Adler NE, et al. Social inequalities in childhood dental caries: the convergent roles of stress, bacteria and disadvantage. *Soc Sci Med.* Nov; 2010 71(9):1644–52. [PubMed: 20870333]
28. Edelstein BL. Disparities in oral health and access to care: findings of national surveys. *Ambul Pediatr.* Mar-Apr;2002 2(2 Suppl):141–7. [PubMed: 11950385]

29. Ferrazzano GF, Scaravilli MS, Ingenito A. Dental and periodontal health status in Campanian children and relation between caries experience and socio-economic behavioural factors. *Eur J Paediatr Dent*. Dec; 2006 7(4):174–8. [PubMed: 17168626]
30. Heft MW, Gilbert GH, Shelton BJ, Duncan RP. Relationship of dental status, sociodemographic status, and oral symptoms to perceived need for dental care. *Community Dent Oral Epidemiol*. Oct; 2003 31(5):351–60. [PubMed: 14667006]
31. Polk DE, Weyant RJ, Manz MC. Socioeconomic factors in adolescents' oral health: are they mediated by oral hygiene behaviors or preventive interventions? *Community Dent Oral Epidemiol*. Feb; 2010 38(1):1–9. [PubMed: 19845711]
32. Javali SB, Pandit PV. Using zero inflated models to analyze dental caries with many zeroes. *Indian J Dent Res*. Oct-Dec;21(4):480–5. [PubMed: 21187609]
33. Tseveenjav B, Suominen AL, Hausen H, Vehkalahti MM. The role of sugar, xylitol, toothbrushing frequency, and use of fluoride toothpaste in maintenance of adults' dental health: findings from the Finnish National Health 2000 Survey. *Eur J Oral Sci*. Feb; 119(1):40–7. [PubMed: 21244510]

Table 1

Study population characteristics

variable	N ^a	% or mean (SD) [range]
female (%)	1068	63.2
race		
white (%)	964	89.9
black (%)	73	6.8
other (%)	35	3.3
Hispanic (%)	1052	1.1
age (years)	1068	34.7 (9.2)
birth year	1066	1971 [1929-1993]
height (m)	897	1.69 (0.10)
weight (kg)	874	85.5 (23.5)
body mass index (kg/m ²)	874	29.7 (7.7)
waist circumference (cm)	322	101.7 (17.9)
educational attainment		
none (%)	149	14.5
high school diploma (%)	454	44.0
technical school/associate degree (%)	136	13.2
some college (%)	128	12.4
undergraduate degree (%)	103	10.0
advanced degree (%)	61	5.9
saliva flow rate (ml/min)	1030	0.68 (0.48)
home water fluoride level (mg/L)	536	0.68 (0.42)
public water source (%)	999	79.3
tooth brushing twice or more per day (%)	983	58.2

^a sample size with available data

Table 2

Surface-specific dental caries prevalence

quadrant	surface	central incisor	lateral incisor	canine	1 st premolar	2 nd premolar	1 st molar	2 nd molar
right maxillary	buccal	0.16	0.13	0.11	0.10	0.09	0.15	0.18
	distal	0.16	0.11	0.07	0.18	0.20	0.21	0.18
	lingual	0.15	0.15	0.07	0.07	0.08	0.40	0.22
	mesial	0.17	0.15	0.07	0.12	0.19	0.27	0.16
	occlusal	-	-	-	0.26	0.30	0.63	0.60
left maxillary	buccal	0.15	0.13	0.11	0.09	0.10	0.15	0.20
	distal	0.16	0.13	0.08	0.19	0.21	0.20	0.18
	lingual	0.16	0.16	0.08	0.07	0.09	0.39	0.23
	mesial	0.18	0.16	0.09	0.12	0.20	0.25	0.18
	occlusal	-	-	-	0.28	0.31	0.63	0.59
right mandibular	buccal	0.03	0.04	0.08	0.09	0.11	0.41	0.29
	distal	0.03	0.03	0.02	0.08	0.19	0.26	0.16
	lingual	0.02	0.01	0.01	0.02	0.07	0.19	0.16
	mesial	0.02	0.02	0.03	0.05	0.11	0.25	0.22
	occlusal	-	-	-	0.12	0.27	0.60	0.64
left mandibular	buccal	0.03	0.03	0.07	0.09	0.10	0.39	0.28
	distal	0.02	0.03	0.02	0.08	0.18	0.26	0.15
	lingual	0.01	0.02	0.02	0.03	0.07	0.21	0.14
	mesial	0.03	0.02	0.03	0.04	0.10	0.25	0.21
	occlusal	-	-	-	0.14	0.26	0.59	0.61

Table 3

General interpretations of PCA and FA loadings.

pattern	major contributing teeth/surfaces	% variation	stability, r
PCA			
PC1	all maxillary teeth and mandibular premolars and molars	26.3	1.00
PC2	molars vs. ^a non-molars	6.7	1.00
PC3	premolars vs. non-premolars	3.9	0.99
PC4	mandibular teeth vs. maxillary teeth	3.3	0.96
PC5	2nd molars vs. mandibular 1st molars	2.6	0.84
PC6	mandibular premolars and 2nd molars vs. mandibular 1st molar and maxillary molars and 2nd premolar	2.4	0.74
PC7	maxillary premolars and mandibular molars vs. maxillary molars and mandibular premolars	2.3	0.85
PC8	complex contributions	2.1	0.88
PC9	complex contributions	1.9	0.71
PC10	right vs. left mandibular molars	1.7	0.62
FA			
FAC1	posterior teeth: premolars and molars	6.3	0.95
FAC2	maxillary anterior teeth: incisors and canines	6.3	0.95
FAC3	mandibular canines and premolars	6.1	0.91
FAC4	maxillary premolars	6.1	0.90
FAC5	mandibular incisors and canines	6.2	0.98
FAC6	non-occlusal premolar and molar surfaces, maxillary lateral incisors, and maxillary canines	6.0	0.75
FAC7	tooth 20 (left mandibular 2nd premolar)	5.9	0.72
FAC8	tooth 29 (right mandibular 2nd premolar)	6.0	0.59
FAC9	maxillary 2nd molars	5.8	0.83
FAC10	tooth 13 (left maxillary 2nd premolar)	6.0	0.51

% variation = percentage of total variability in the data explained by the PC/FAC. PCs 1-10 are additive and cumulatively account for 53.2% of data variability. FACs 1-10 are not additive and cumulatively account for 44.7% of data variability.

stability, r = the mean correlation coefficient comparing PCs/FACs extracted from the full data set to those extracted from 10 random subsets of 80% of the full data.

^a contributions of surfaces to PCs may differ in direction; here we separate surfaces with opposing contributions using “vs.”

Table 4

Associations (correlation coefficients) between caries outcomes and potential risk factors.

	sex ^a	white race	Hispanic	age ^a	age ^a ^a	birth year ^a	BMI	weight	waist	educational attainment	saliva flow	water fluoride	tooth brushing	public water
N	1068	1068	1052	1068	1068	1066	874	874	322	1031	1030	536	999	983
DMFS ^f	0.04	0.01	0.02	0.19 ^c	0.19 ^c	-0.15 ^c	-0.04	-0.04	0.04	0.01	-0.03	0.06	-0.02	-0.00
PCA														
PC1	0.03	0.05	0.05	0.29 ^c	0.3 ^{c,d}	0.26 ^{c,e}	0.00	0.00	0.09	0.00 ^d	0.01	0.07	0.02 ^d	0.04
PC2	0.07 ^b	0.04	0.01	0.00	0.02 ^e	0.02 ^e	0.01	0.00	0.02	0.19 ^{c,e}	0.01	0.01	0.07 ^{b,d}	0.06
PC3	0.10 ^c	0.03	0.01	0.14 ^c	0.15 ^{c,d}	0.15 ^{c,e}	0.05	0.01	0.04	0.01	0.02	0.03	0.13 ^{c,d}	0.05
PC4	0.03	0.09 ^{b,e}	0.01	0.04	0.03	0.06 ^e	0.01	0.02	0.08	0.10 ^{c,d}	0.01	0.09 ^{b,d}	0.04	0.04
PC5	0.02	0.02	0.00	0.12 ^c	0.10 ^{c,d}	0.13 ^{c,d}	0.01	0.03	0.06	0.08 ^b	0.03	0.03	0.06 ^b	0.00
PC6	0.01	0.08 ^{b,d}	0.01	0.03	0.03	0.01 ^e	0.03	0.03	0.03	0.02	0.00	0.05	0.01	0.05
PC7	0.03	0.01	0.06 ^{b,d}	0.01	0.00	0.01	0.02	0.04	0.04	0.02	0.03	0.03	0.03	0.01
PC8	0.04	0.11 ^{c,e}	0.05	0.04	0.04	0.05 ^e	0.04	0.01	0.01	0.06	0.02	0.09 ^{b,d}	0.04	0.07 ^{b,d}
PC9	0.00	0.05	0.02	0.08 ^b	0.08 ^b	0.08 ^b	0.03	0.01	0.07	0.02	0.03	0.00	0.02	0.01
PC10	0.01	0.06	0.03	0.00	0.01	0.00	0.02	0.04	0.01	0.04	0.07 ^{b,d}	0.07	0.01	0.01
FA														
FAC1	0.02	0.05	0.04	0.17 ^c	0.16 ^c	0.13 ^{c,e}	0.00	0.00	0.10	0.12 ^{c,d}	0.00	0.05	0.00	0.05
FAC2	0.08 ^b	0.05	0.03	0.09 ^b	0.09 ^c	0.09 ^b	0.02	0.00	0.03	0.09 ^{b,e}	0.00	0.01	0.11 ^{c,e}	0.07 ^{b,d}
FAC3	0.02	0.05	0.06	0.03	0.05 ^d	0.03	0.03	0.03	0.05	0.06 ^{b,d}	0.02	0.05	0.01	0.00
FAC4	0.07 ^b	0.02	0.07 ^{b,d}	0.22 ^c	0.22 ^c	0.22 ^c	0.03	0.01	0.05	0.08 ^b	0.04	0.00	0.11 ^{c,d}	0.03
FAC5	0.00	0.04	0.01	0.11 ^c	0.12 ^{c,d}	0.11 ^c	0.00	0.02	0.11	0.01	0.02	0.1 ^{b,d}	0.02	0.01
FAC6	0.03	0.11 ^{c,e}	0.06	0.01	0.03 ^d	0.02 ^e	0.01	0.01	0.05	0.17 ^{c,e}	0.01	0.08	0.07 ^{b,d}	0.06 ^d
FAC7	0.05	0.04	0.01	0.11 ^c	0.11 ^c	0.10 ^c	0.03	0.00	0.05	0.03	0.03	0.03	0.02	0.01
FAC8	0.00	0.00	0.00	0.11 ^c	0.12 ^{c,d}	0.12 ^c	0.00	0.00	0.00	0.01	0.05	0.02	0.01	0.03

	sex ^a	white race	Hispanic	age ^a	age ² ^a	birth year ^a	BMI	weight	waist	educational attainment	saliva flow	water fluoride	tooth brushing	public water
FAC9	0.04	0.02	0.01	0.04	0.03	0.05 ^d	0.02	0.05	0.06	0.04	0.02	0.06	0.03	0.00
FAC10	0.04	0.10 ^{c,d}	0.05	0.11 ^c	0.11 ^c	0.10 ^{c,d}	0.03	0.01	0.01	0.04 ^d	0.06 ^{b,d}	0.06	0.03	0.03

BMI = body mass index

Spearman (non-parametric) correlation coefficients are shown for DMFS due to its non-Normal distribution. Pearson correlation coefficients are shown for PCs and FACs. Note: The direction of associations for all PCs/FACs are listed as positive because the sign of PCs/FACs are arbitrary.

Note: Univariate associations (shown) were similar to multiple regression (i.e., sex- and age-adjusted) models (point estimates not shown).

^a age- and sex-adjusted regression models were not applicable for these risk factors

^b unadjusted suggestive significance; correlation p-value < 0.05

^c unadjusted Bonferroni significance; correlation p-value < 0.00238

^d age- and sex-adjusted suggestive significance; regression p-value < 0.05

^e age- and sex-adjusted Bonferroni significance; regression p-value < 0.00238