



Article

Cohort Profile: ZOE 2.0—A Community-Based Genetic Epidemiologic Study of Early Childhood Oral Health

Kimon Divaris ^{1,2,*} , Gary D. Slade ¹, Andrea G. Ferreira Zandona ³, John S. Preisser ⁴, Jeannie Ginnis ¹, Miguel A. Simancas-Pallares ¹, Cary S. Agler ¹, Poojan Shrestha ^{1,2} , Deepti S. Karhade ¹, Apoena de Aguiar Ribeiro ⁵, Hunyong Cho ⁴, Yu Gu ⁴ , Beau D. Meyer ⁶ , Ashwini R. Joshi ⁷, M. Andrea Azcarate-Peril ⁸ , Patricia V. Basta ², Di Wu ^{4,9} and Kari E. North ^{2,10}

¹ Division of Pediatric and Public Health, Adams School of Dentistry, University of North Carolina-Chapel Hill, NC 27599-7450, USA; Gary_Slade@unc.edu (G.D.S.); Jeannie_Ginnis@unc.edu (J.G.); simancas@email.unc.edu (M.A.S.-P.); cagler@email.unc.edu (C.S.A.); poojansh@live.unc.edu (P.S.); deepti_karhade@unc.edu (D.S.K.)

² Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina-Chapel Hill, NC 27599-7400, USA; patricia_basta@unc.edu (P.V.B.); kari_north@unc.edu (K.E.N.)

³ Department of Comprehensive Dentistry, School of Dental Medicine, Tufts University, Boston, MA 02111, USA; Andrea.Zandona@tufts.edu

⁴ Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina-Chapel Hill, NC 27599-7400, USA; jpreisse@bios.unc.edu (J.S.P.); hunycho@live.unc.edu (H.C.); yugu@live.unc.edu (Y.G.); did@email.unc.edu (D.W.)

⁵ Division of Diagnostic Sciences, Adams School of Dentistry, University of North Carolina-Chapel Hill, NC 27599-7450, USA; apoena@email.unc.edu

⁶ Division of Pediatric Dentistry, College of Dentistry, The Ohio State University, Columbus, OH 43210, USA; meyer.781@osu.edu

⁷ Division of Surgery, School of Medicine, University of North Carolina-Chapel Hill, NC 27599-7050, USA; ashwini1@email.unc.edu

⁸ Center for Gastrointestinal Biology and Disease, Division of Gastroenterology and Hepatology, and UNC Microbiome Core, Department of Medicine, School of Medicine, University of North Carolina-Chapel Hill, NC 27599-7555, USA; andrea_azcarate-peril@med.unc.edu

⁹ Division of Oral and Craniofacial Health Sciences, Adams School of Dentistry, University of North Carolina-Chapel Hill, NC 27599-7450, USA

¹⁰ Carolina Center for Genome Sciences, University of North Carolina-Chapel Hill, NC 27514, USA

* Correspondence: Kimon_Divaris@unc.edu; Tel.: +1-919-537-3556

Received: 5 October 2020; Accepted: 29 October 2020; Published: 1 November 2020



Abstract: Early childhood caries (ECC) is an aggressive form of dental caries occurring in the first five years of life. Despite its prevalence and consequences, little progress has been made in its prevention and even less is known about individuals' susceptibility or genomic risk factors. The genome-wide association study (GWAS) of ECC ("ZOE 2.0") is a community-based, multi-ethnic, cross-sectional, genetic epidemiologic study seeking to address this knowledge gap. This paper describes the study's design, the cohort's demographic profile, data domains, and key oral health outcomes. Between 2016 and 2019, the study enrolled 8059 3–5-year-old children attending public preschools in North Carolina, United States. Participants resided in 86 of the state's 100 counties and racial/ethnic minorities predominated—for example, 48% ($n = 3872$) were African American, 22% white, and 20% ($n = 1611$) were Hispanic/Latino. Seventy-nine percent ($n = 6404$) of participants underwent clinical dental examinations yielding ECC outcome measures—ECC (defined at the established caries lesion threshold) prevalence was 54% and the mean number of decayed, missing, filled surfaces due to caries was eight. Nearly all (98%) examined children provided sufficient DNA from saliva for genotyping.

The cohort's community-based nature and rich data offer excellent opportunities for addressing important clinical, epidemiologic, and biological questions in early childhood.

Keywords: children; early childhood caries; community-based studies; oral health; genomics

1. Introduction

Dental caries is the most common chronic childhood disease worldwide. Early childhood caries (ECC) is an early-onset, aggressive form of the disease that affects an estimated 600 million children worldwide. Management of ECC is challenging—it frequently requires restorative care under sedation or general anesthesia, and there are populations wherein it remains largely untreated [1]. The disease imposes substantial human and economic costs on children, their families, and the public health care infrastructure [2]. This burden falls disproportionately on poor families, whose children are more likely to develop ECC compared to their non-poor counterparts, and who are frequently unable to access the requisite dental care [3,4].

ECC, similar to other forms of dental caries, is characterized by a biofilm-mediated, sugar-driven, multifactorial, dysbiotic shift at the interface of a tooth surface and its biofilm, which results in the progressive demineralization of dental hard tissues [5,6]. Despite a general decrease in caries prevalence in permanent teeth of older children and adults, the proportion of preschool children with ECC has persisted in recent decades [7]. This suggests that public health- and personal-preventive care, while effective in preventing caries in older children, has had little impact on ECC. An explanation for its persistence therefore must look beyond established behavioral, environmental, and societal risk factors. A better understanding of genetic factors underlying individual susceptibility and its interaction with oral health-related behaviors and the environment is needed [8].

Earlier observations emanating from twin, family, candidate gene, and recent genome-wide association studies (GWAS) support a considerable genetic component for dental caries [8–10]. Our group recently estimated ECC heritability to be 44% but this estimate was based on a very small sample size and was imprecise [11]. More recently, Haworth and colleagues [12] reported heritability estimates for primary dentition caries based on a consortium ($n = 7230$) meta-analysis to be 28% (95% confidence interval: 9–48%). Despite this evidence, efforts to identify genetic susceptibility factors associated with ECC have not paralleled recent genomics discoveries for dental caries among adults [13]. The first GWAS of ECC involved primarily European–American (white) children [14], whereas the recent consortium-based analysis [12] included small numbers of non-European children. Data collection and the resulting phenotypic heterogeneity was a limitation of the consortium GWAS, and both studies included predominantly children over the age of six, whereas 71 months is the uppermost age threshold for defining ECC [1]. Despite these issues, both studies provided evidence in support of loci implicated in childhood caries including *ALLC*, *MPPED2*, and *PKD2*.

The genome-wide association study of ECC (“ZOE 2.0”) is a community-based, multi-ethnic, genetic epidemiologic study of early childhood oral health seeking to add to the knowledge base of ECC genomics. The justification of ZOE 2.0 was supported by four key features: (1) a focus on 3–5-year-old children, an age range within which the ECC phenotype is defined and clinically-determined caries experience has been the least distorted by restorative treatment and shedding of primary teeth; (2) a large sample of >6000 children; (3) a community-based (non-clinical) sample; (4) a tested clinical examination protocol including tooth-surface level recording of lesions of different severity levels. Accordingly, the study's specific aims were to: (a) identify risk loci and gene sets/pathways for prevalent ECC and related traits (e.g., caries prevalence and severity) by conducting a multi-ethnic GWAS meta-analysis; (b) determine the prevalence and severity of dental caries, other oral disease outcomes, and putative risk factors for dental caries in a low-income population-based sample of European-, African-, and Latino/Hispanic–American preschool-age children.

2. Materials and Methods

ZOE 2.0 is an observational, cross-sectional, community-based epidemiologic investigation of early childhood oral health with focus on ECC. The study received approval (#14-1992) from the University of North Carolina—Chapel Hill Office of Human Research Ethics IRB on 18 September 2014. The study represents the evolution of previous oral health research initiatives in the public preschool child population in North Carolina, including the ZOE study [11,15] between 2012 and 2013, and the ZOE-pilot study between 2013 and 2014 [16]. Specifically, the ZOE 2.0 study population was children attending public preschools (Head Start) in North Carolina, United States, between August 2016 and February 2019. The federal Head Start program is authorized by the Improving Head Start for School Readiness Act of 2007 with the goal to promote school readiness of children from low-income families. Children were eligible to participate in ZOE 2.0 if they were between the ages of 3–5 years and currently enrolled in a North Carolina Head Start program. Other eligibility criteria were a legal guardian at least 18 years old who was able to understand the study's informational material and consent documents that were provided in English and Spanish language, and provided their written informed consent. At enrollment, children's legal guardians completed a one-page, self-administered, paper-and-pen questionnaire that included 15 demographic and oral health-related questions. To enable the production area-level visual summaries and analyses, participants' residential locations were geocoded using ArcGIS Pro software (Esri, Redlands, CA, USA).

The target analytical sample of ~6000 was set balancing the requirements of adequate power to detect genetic effects and feasibility—it was based upon a goal of >80% power to detect relative effect sizes (i.e., for the binary trait of established ECC case status) of 1.15 for single nucleotide polymorphisms (SNPs) with minor allele frequency (MAF) ≥ 0.45 , 1.20 for MAF ≥ 0.20 , and 1.30 for MAF ≥ 0.10 . We consider these scenarios and plausible detectable effect sizes realistic in light of the only genome-wide significant marker reported for the primary dentition caries to date, rs1594318 (MAF = 0.40, odds ratio = 0.85, $p = 4.1 \times 10^{-8}$) [12].

The study's design and sampling frame specified procedures for the initial selection of 24 out of the state's 52 Head Start programs, stratified into four geographical regions of NC. All centers within the invited programs and all children within each center were invited to participate. The selection of programs used probability proportionate to size (PPS) sampling within strata (i.e., 4 geographic regions) and assumed a 75% enrollment rate. Because the enrollment rate after the first year of the study was considerably lower than expected (42%), the study team invited 2–3 additional programs in each region, resulting in the enrollment of 10 additional programs (i.e., a total of 34). Of note, by the end of the study, the cumulative enrollment rate had increased to 62%.

After enrollment, 10 trained and calibrated clinical examiners (registered dental hygienists or dentists) conducted comprehensive examinations of participating children in their preschools. After toothbrushing among all children (and flossing as needed), children's teeth were examined using portable dental equipment, compressed air, and uniform artificial light and magnification conditions (i.e., custom-fit magnifying loupes with headlights; Orascoptic XV1™ Loupe + Light, Orascoptic, Middleton, WI, USA). Recording of clinical data was done using a Microsoft Access-based (Microsoft Corp., Redmond, WA, USA) custom-written data entry application. A detailed description of the study's clinical protocol and procedures has been reported by Ginnis et al. [17]. Briefly, tooth surface caries diagnoses were based on modified International Caries Detection and Assessment (ICDAS) criteria [18,19] and made at the levels of health (ICDAS code, 0, representing caries-free tooth surfaces), early stage (ICDAS codes, 1–2, representing initial caries lesions), and established stage (ICDAS codes, 3–6, representing established or severe caries lesions). These stages resemble earlier, historic classifications of caries lesions as “pre-cavitated” (i.e., analogous to early stage) or “cavitated” (i.e., analogous to established/severe) but do not correspond exactly to them—for example, ICDAS 4 lesions may or may not be cavitated, but they are nowadays considered established due to their extension in dentin [19]. Data were summed from all primary tooth surfaces to calculate the dmfs index, a count of primary tooth surfaces that were decayed, missing due to caries or filled/restored

(ranging between 0 and 88). Different caries lesion detection thresholds were used to calculate two values of dmfs: early-stage (ICDAS ≥ 1 ; classic ECC definition [1]) and established (ICDAS ≥ 3) [17]. Classic ECC is the internationally accepted case definition for ECC and includes the earliest detectable lesions compatible with enamel demineralization loss [1,19]. The established definition includes lesions with deeper demineralization and possibly tooth surface loss—as such, it has more clinical and public health relevance and is also the closest to what is measured using the National Health and Nutrition Examination Survey (NHANES) clinical examination protocol [7]. The collection of detailed surface-level data on experience allows for the creation of several other clinical traits for comparisons with previous and future studies, including or excluding restorations (e.g., fillings, crowns) and extracted teeth, as well as clinical patterns of ECC (e.g., pits and fissures, smooth surfaces, proximal surfaces, etc.). Additional clinical examination domains included skeletal and occlusal characteristics, developmental defects of the enamel, dental trauma, and restorative treatment needs. Children's height and weight were measured using a portable stadiometer (Seca[®] 213, Seca GmbH & Co. KG, Hamburg, Germany) and a portable digital scale (Doran[®] DS6150 Remote Indicator Scale, Doran, Batavia, IL, USA) [17]. Age- and sex-specific body mass index (BMI) Z-scores and corresponding BMI categories (i.e., underweight, normal weight, overweight, obese) were calculated using the zanthro program in Stata 16.1 (StataCorp LLC, College Station, TX, United States) and the 2000 Centers for Disease Control (CDC) Growth reference data.

Saliva samples were collected from children using the DNA Genotek Oragene DNA-575 kit (DNA Genotek, Ottawa, Ontario, Canada). This was done for children that presented for a clinical examination. DNA was extracted from these saliva samples, quantitated, and quality assessed using procedures detailed in the study's genomics analysis protocol [20]. The purified DNA material was carried forward to high-density genotyping at the Center for Inherited Disorders Research (CIDR), at Johns Hopkins University, using the Infinium[™] Global Diversity Array (Illumina, San Diego, CA, USA).

Two supragingival plaque samples were also collected for each participant; one pooled sample from all facial/buccal surfaces of the upper right quadrant and one pooled sample from the facial/buccal surfaces of the upper left quadrant. These samples were biobanked for future studies involving metagenomics, metatranscriptomics, and metabolomics [21,22]. To date, ~5% of them ($n = 300$, equally split between ECC cases and non-cases, selected from the first wave of participants and sequenced in 2 batches) were used in pilot, feasibility and discovery studies to facilitate the development of novel multi-omics analysis methods. Detailed, step-by-step descriptions of procedures for collection, transportation, storage, and processing of all biospecimens (i.e., saliva and two plaque samples for each participant) have been made available via three publicly available publications [17,20,21].

A sample of tap water from the child's home was obtained after the clinical examination and was used for fluoride ion concentration measurement by the North Carolina State Laboratory of Public Health using the United States Environmental Protection Agency (EPA) 300.0 method. An overview of data domains and measures collected in the study has been previously reported and is illustrated in Figure 1.

Quality assessment and assurance procedures were based upon periodic review of database reports, including but not limited to enrollment data, clinical data, biospecimen manifests, fluoride measurement reports, consents received and signed, physical and scanned questionnaires. Periodic study-wide quality control procedures were implemented both by the study team and an external independent monitor (Rho Inc., Durham, NC, USA), as required by the study sponsor. Targeted reviews of primary records were done by at least two study members in cases where: (a) free-text responses were encoded to study variables (e.g., children's medication list and measured height and weight); (b) participants' residential addresses could not be exactly geocoded due to spelling or data entry errors; (c) clinical examiners entered notes in "notes field" to be reviewed by the core study team.

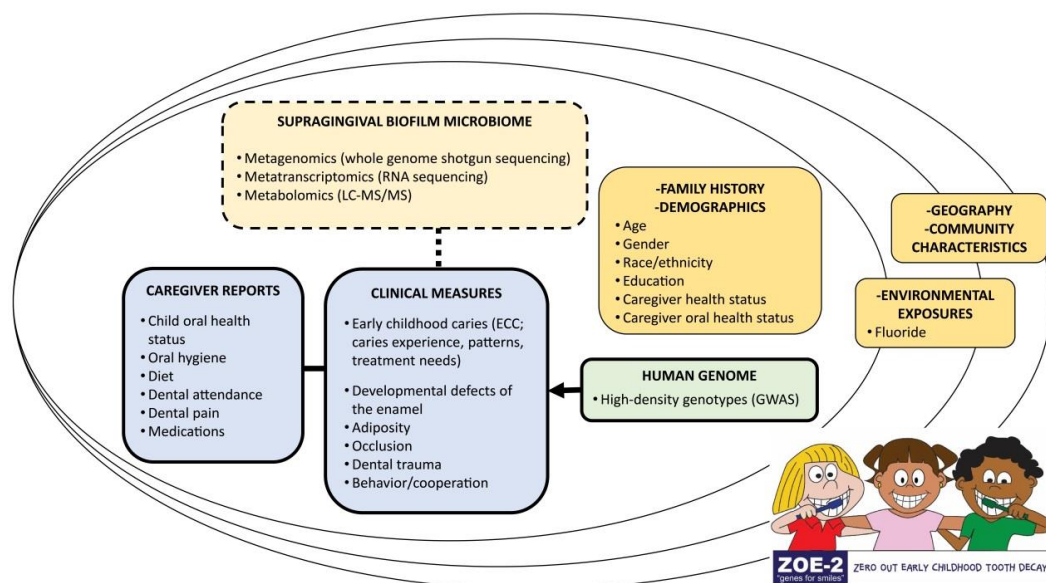


Figure 1. Overview of data domains and measures collected in the ZOE 2.0 study, North Carolina, United States. Adapted with permission from Divaris and Joshi 2020 [20]. To date, supragingival biofilm microbiome analyses have been carried out in a pilot subset of 300 ZOE 2.0 participants (~5% of those with clinical data).

For this cohort profile, descriptive statistics were estimated using Stata 16.1 (StataCorp LLC, College Station, TX, USA) and SAS 9.4 (SAS Institute Inc., Cary, NC, USA). Estimation accounted for the clustered nature of observations (i.e., children’s enrollment in 34 different Head Start programs) via the use of robust Taylor-linearized variances.

3. Results

3.1. Enrollment of Study Participants

The study’s target analytical sample for the conduct of the GWAS was 6000 examined and genotyped children. There were approximately 20,000 children enrolled in the North Carolina public preschool (Head Start) system, the target population. A flowchart describing participants’ enrollment, clinical examination, and the derivation of the GWAS sample is presented in Figure 2. Monthly and cumulative enrollment numbers over time, during the 30-month enrollment period are presented in Figure 3. The study team used information on enrollment and examination rates from pilot studies and from the first year of the study to initially design, and subsequently adjust and optimize the recruitment strategy to meet the target analytical sample size. Approximately 13,000 participants were invited via written material disseminated to Head Start centers and 62% of those ($n = 8059$) were enrolled. The overwhelming reason for exclusion after the assessment of eligibility was duplicate entries (i.e., more than one participation packet for the same participant), followed by incomplete documentation, such as missing or unsigned consent, and being outside the eligible age range.

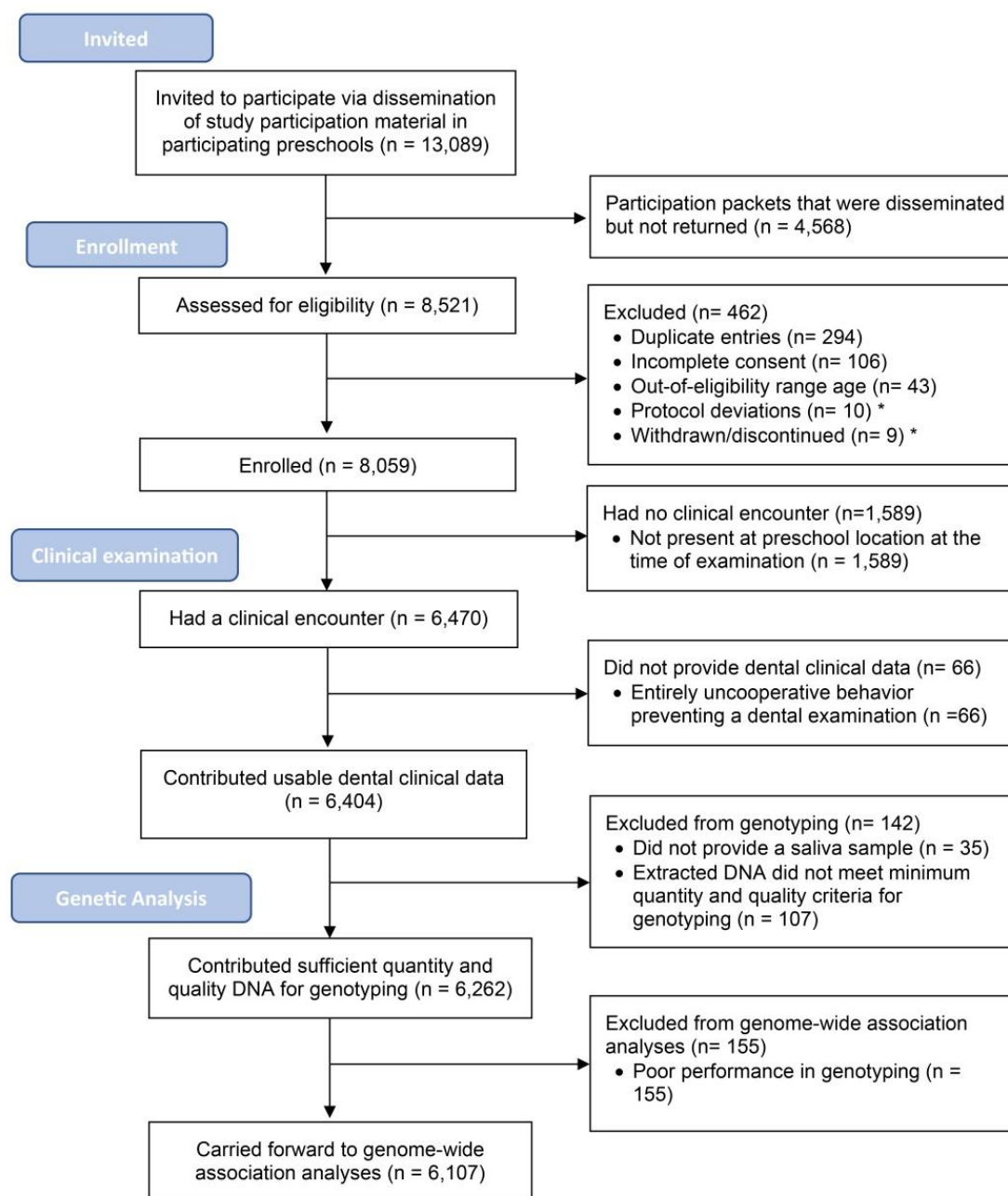


Figure 2. Flowchart of enrollment, clinical examinations, and derivation of the GWAS analytical samples in the ZOE 2.0 study, North Carolina, United States (the 19 participants that are identified with protocol deviations that were withdrawn, or discontinued were initially considered as part of enrollment, but were removed from the study enrollment and do not count towards the reported enrollment figure of 8059).

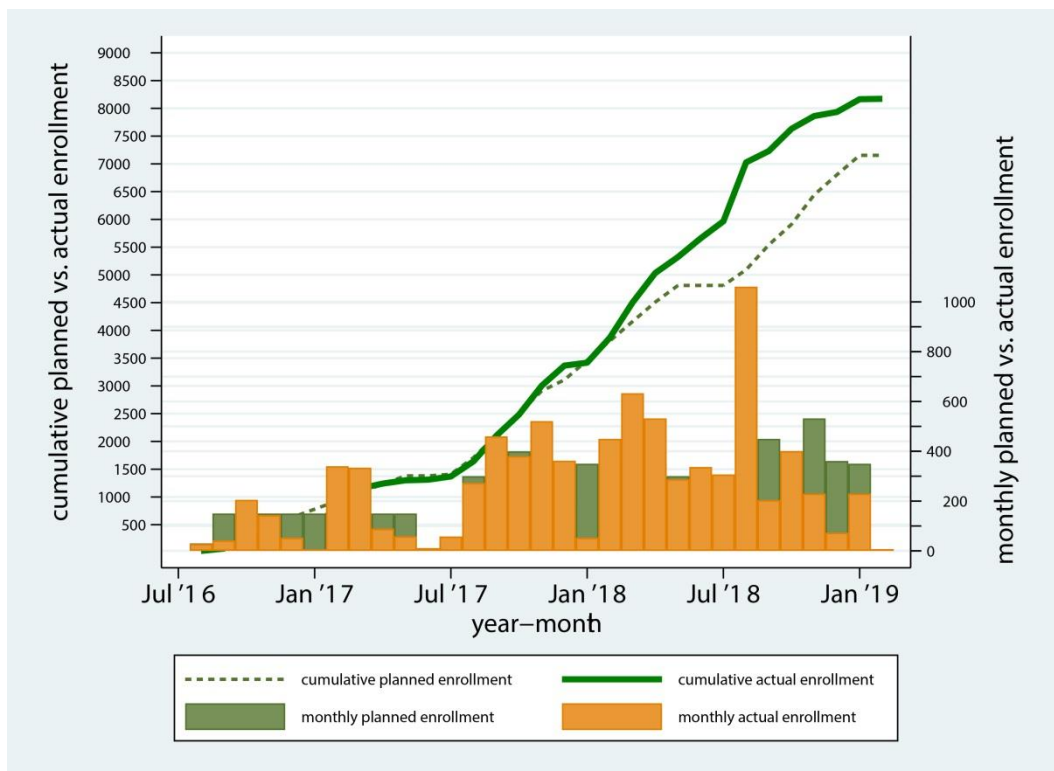


Figure 3. Monthly and cumulative enrollment during the 30-month enrollment period in the ZOE 2.0 study, North Carolina, United States.

As illustrated in Figure 4, enrolled participants’ residential locations were spread across North Carolina, and were in 86 (out of a total 100) counties of the state. Enrolled children attended 34 different Head Start programs (clusters), and 260 different Head Start centers nested within these programs. Based on the number and size of programs and centers, the Head Start program was determined as the appropriate level of clustering for analytical purposes.

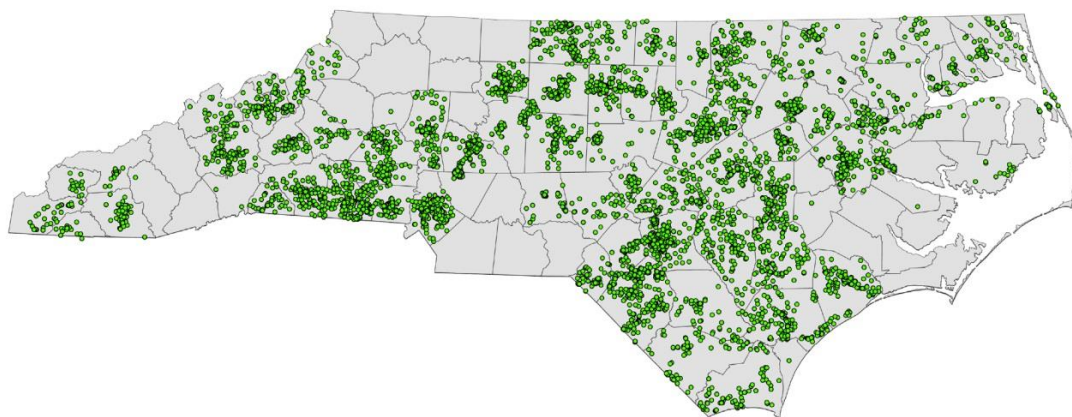


Figure 4. Geographical distribution of participants enrolled in the ZOE 2.0 study across the state of North Carolina, United States.

3.2. Clinical Examinations

Eighty percent of enrolled participants ($n = 6470$) were seen for clinical assessments and 99% of those encounters ($n = 6404$) yielded usable clinical information on ECC. Clinical encounters that did not lead to usable clinical data were due to children's uncooperative behavior, which is not unexpected in this age group. The monthly and cumulative numbers of clinical examinations are presented in Figure 5 (left panel) and generally paralleled enrollment. As illustrated on the right panel of Figure 5, most exams were conducted within 2 months after enrollment; median and mean days elapsed between enrollment and examination were 36 and 57, respectively.

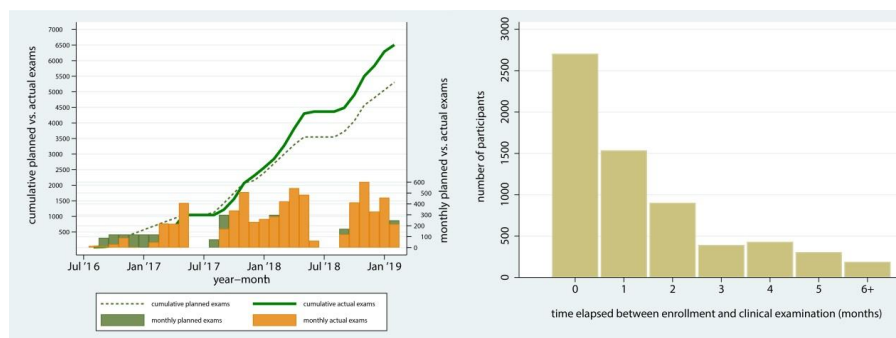


Figure 5. Monthly and cumulative clinical encounters during the 30-month study period and distribution of time elapsed (months) between enrollment and clinical examination in the ZOE 2.0 study, North Carolina, United States.

3.3. Data Completeness

We considered data completeness by data domain (e.g., questionnaire, clinical examination, biospecimens, water sample return, etc.) and for each individual questionnaire item. As presented in Table 1, data completeness was high for all data domains except for domestic water samples returned. Directly measured concentrations of participants' domestic water source fluoride content are available for 24% of those with ECC data. Fluoride measurement was done for virtually all water samples received—the most common reason for missing fluoride measurements of water samples received was over 30 days elapsed between the reported water sample collection and receipt of the sample in the lab. A smaller proportion of samples were not carried forward to measurement due to issues with labeling or shipping (e.g., leaked during transportation).

Nearly all examined participants (98–99%) provided saliva and supragingival plaque samples. Human genomic DNA of adequate quality and quantity was available for 98% of those who provided saliva samples. A similarly high percentage of samples with adequate quantity and quality DNA (98%; $n = 6107$) performed satisfactorily during genotyping and are eligible to be carried forward to a GWAS.

Similarly, data missingness was rare when considering responses to individual questionnaire items, with most items having 0 to 2% missing data. As presented in Table 2, the highest proportion of missing responses was noted for the question “What is your child's race” which was commonly left incomplete among Spanish language respondents and participants of Hispanic ethnicity. Of note, 11% of participants ($n = 923$) elected to complete study materials in the Spanish language. In terms of respondent-level missingness, 88% of respondents had no missing questionnaire responses; 9.4% had one, 1.4% had two, and 1.2% had three or more. Missingness was higher among those who completed the Spanish language questionnaire: 26% had at least one missing response, and 7% had at least two missing responses compared to 10% with at least one missing response and 2% with at least two missing responses among those who completed the English language questionnaire.

Table 1. Numbers and percentages of participants with non-missing data by domain.

	<i>n</i> (%)
Questionnaire data	8054/8059
(among those with verified, consented enrollment)	99.9%
Usable (i.e., non-missing) clinical dental data	6404/6470
(among those with attempted clinical exams)	99%
Usable (i.e., non-missing) anthropometric (i.e., height/weight) data	6442/6470
(among those with attempted clinical exams)	99%
Saliva collected	6369/6404
(among those that had exams and have non-missing clinical data)	99%
Supragingival plaque collected	6271/6404
(among those that had exams and have non-missing clinical data)	98%
Adequate quality and quantity of human DNA available	6262/6369
(among those with non-missing clinical data that provided saliva samples)	98%
Acceptable performance during genotyping	6107/6262
(among those with adequate quality and quantity of human DNA)	98%
Water sample provided for fluoride measurement	1613/6470
(among those with attempted clinical exams)	25%
Usable fluoride concentration information	1518/1613
(among those who provided a domestic water sample)	94%
Fluoride concentration data available	1518/6404
(among those with clinical data)	24%
Fluoride concentration and geocoding data available	1501/6404
(among those with clinical data)	23%

Table 2. Percentage of missing responses in individual questionnaire items.

Question	%
Who brushes your child's teeth at home?	0.4
How often are your child's teeth brushed?	0.4
Is toothpaste with fluoride used every time your child's teeth are brushed?	1.1
How many snacks and drinks containing sugar does your child usually have between meals?	0.5
Has your child ever gone to the dentist?	1.1
Does your child usually see a dentist for a check-up or because of dental problems?	1.3
Has your child ever been put to bed with a bottle containing something other than water?	2.1
How would you describe the condition of your child's mouth and teeth?	1.2
How would you describe the condition of your mouth and teeth?	2.6
Has your child ever had a toothache or other dental pain (not from teething)?	1.2
Has your child been given an over-the-counter medication during the last 30 days?	1.2
What is the highest grade-level of schooling you have completed?	2.6
Are you of Hispanic or Latino origin?	1.1
What is your child's race?	2.9
What is the primary source of your home drinking water for your child?	1.1

3.4. Protocol Deviations

A total of 10 protocol deviations occurred during the course of the study. None of them affected participant safety or the interpretability of the study data. Nine of these protocol deviations were due to children who were marginally outside the age range for eligibility and were inadvertently clinically examined during the first year of the study. The tenth protocol deviation pertained to the clinical examination of an incompletely consented participant. These participants were subsequently rendered ineligible and were reported as protocol deviations to the overseeing UNC-Chapel Hill IRB who determined that they did not increase participants' risk. These protocol deviations were also reported to the sponsor (National Institutes of Health). It was determined that these deviations did not meet the criteria for Unanticipated Problems Involving Risks to Subjects or Others (UPIRSO), or other noncompliance, and the study team's corrective action plan in response to these protocol deviations was deemed acceptable.

3.5. Enrolled and Examined Participants' Demographic Characteristics

Enrolled participants were mostly 4-year-olds (i.e., their mean and median age at enrollment was 52 months). African-Americans were the most represented racial group (48%), followed by whites (22%), and those of more than one race. One fifth ($n = 1611$) of respondents were of Hispanic ethnicity. Importantly, as presented in Table 3, there were no differences in the demographic profile of participants enrolled and those who were clinically examined and provided information for ECC.

Table 3. Demographic characteristics of all participants enrolled and all participants that provided early childhood caries (ECC) information via clinical examinations in the ZOE 2.0 study, North Carolina, 2016–2019.

	All Enrolled	w/ECC Information	p^*
	n (Column %)	n (Column %)	
Entire Sample	8059 (100)	6404 (100)	
Sex †			
male	4000 (50)	3189 (50)	0.6
female	4057 (50)	3215 (50)	
missing	2	0	
Age at enrollment (years)			0.2
3	2568 (32)	1992 (31)	
4	4234 (53)	3375 (53)	
5	1257 (16)	1037 (16)	
(months), mean (SD)	52 (7.5)	52 (7.4)	0.2
Race			
African-American	3872 (48)	3094 (48)	0.7
American-Indian or Alaskan Native	236 (3)	186 (3)	
Asian	42 (1)	32 (1)	
Native Hawaiian or other Pacific Islander	7 (0.1)	4 (0.1)	
White	1765 (22)	1385 (22)	
>1 race	1067 (13)	835 (13)	
other/missing	1070 (13)	868 (13)	
Hispanic ethnicity			0.8
yes	1611 (20)	1291 (20)	
no	6355 (80)	5042 (80)	
missing	93	71	

ECC, Early Childhood Caries; SD, Standard Deviation. * Bivariate tests accounting for the clustered nature of the data. †The parent questionnaire asked whether the participating child was a boy or girl; subsequently, biological sex was determined via genotyping.

3.6. Main Study Outcomes—ECC

ECC was common in this community-based sample of preschool-age children. As presented in Table 4, more than half of examined children (54%, $n = 3465$) were ECC cases at the established/severe caries lesion detection threshold (ICDAS ≥ 3).

Table 4. Estimates of early childhood caries (ECC) experience (case status, defined as one or more decayed, missing, or filled tooth surface (dmfs) due to dental caries) and burden (dmfs) overall and according to participants' demographic characteristics in the ZOE 2.0 study, North Carolina, 2016–2019. Two thresholds of caries lesion detection according to the International Caries Detection and Assessment System (ICDAS) are used—ICDAS ≥ 3 (established ECC) and ICDAS ≥ 1 (classic ECC).

	Established ECC, Defined at the ICDAS ≥ 3 threshold		Classic ECC, Defined at the ICDAS ≥ 1 threshold	
	dmfs > 0, <i>n</i> (row %)	dmfs, mean (se*)	dmfs>0, <i>n</i> (row %)	dmfs, mean (se*)
Entire Sample	3465 (54)	8 (0.6)	5882 (92)	15 (0.7)
Sex				
male	1748 (55)	8 (0.5)	2935 (92)	16 (0.6)
female	1717 (53)	8 (0.7)	2947 (92)	15 (0.7)
Age (years)				
3	673 (45)	5 (0.4)	1347 (90)	13 (0.6)
4	1836 (55)	8 (0.6)	3084 (92)	15 (0.7)
5	956 (61)	11 (0.8)	1451 (93)	17 (0.9)
Race				
African–American	1622 (52)	7 (0.5)	2845 (92)	14 (0.6)
American–Indian or Alaskan Native	127 (68)	15 (1.6)	180 (97)	22 (1.3)
Asian	24 (75)	16 (3.3)	32 (100)	23 (3.5)
Native Hawaiian or other Pacific Islander	4 (100)	21 (6.7)	4 (100)	30 (8.0)
White	725 (52)	8 (0.6)	1251 (90)	15 (0.9)
>1 race	426 (51)	7 (0.6)	741 (89)	13 (0.7)
other	534 (62)	11 (1.0)	825 (95)	18 (1.1)
Hispanic ethnicity				
yes	788 (61)	11 (1.0)	1209 (94)	18 (1.2)
no	2643 (52)	7 (0.5)	4604 (91)	14 (0.6)
missing	34 (48)	8 (2.4)	69 (97)	16 (2.4)

ECC, Early Childhood Caries; ICDAS, International Caries Detection and Assessment System; dmfs, decayed, missing, filled/restored tooth surfaces due to caries. *se—linearized standard errors accounting for the clustered nature of the data (i.e., attendance in 34 different public preschool programs).

The dmfs index was zero-inflated and right-skewed, with a sample mean of 8 and a standard deviation of 14. Because a child's dmfs index cannot reduce within this age range while the disease may progress, caries prevalence was substantially higher among older children (e.g., 45% among 3-year-olds vs. 61% among 5-year-olds). Children in the least represented racial groups and those of more than one race had higher ECC prevalence and burden compared to African-Americans and whites. Similarly, those of Hispanic ethnicity had higher ECC prevalence and burden compared to their non-Hispanic counterparts. These patterns were similar but less pronounced when considering the early-stage definition of ECC, defined at the ICDAS ≥ 1 caries lesion detection threshold. Ninety-two percent of children were ECC cases according to this definition, with a mean dmfs of 15. Differences in ECC prevalence between age groups were less pronounced using this definition.

In terms of BMI related outcomes, 10% of children ($n = 624$) were found to be underweight (i.e., <5th BMI percentile for age and sex), 68% had normal weight, 13% were overweight, and 9% ($n = 588$) were obese (i.e., ≥ 95 th BMI percentile for age and sex).

4. Discussion

The successful completion of the data collection for the ZOE 2.0 GWAS holds promise for new insights about genomic influences on ECC. We anticipate that this cross-sectional study's findings will offer novel molecular, clinical, and contextual insights for prevalent ECC.

Looking ahead, ample opportunities now exist for the validation of biomarkers (e.g., microbiome and genome) using prospective follow-up studies that are envisaged for this cohort [20]. The study enrolled a sizable sample of over 8000 preschool-age children for which guardian-provided measures of

child oral health were available and approximately 6500 children were clinically examined. Virtually all examined children provided saliva and supragingival plaque (i.e., biofilm) samples enabling a GWAS of ECC, microbiome, and other multi-omics studies.

We found that 54% of children examined in this low-income, community-based sample had caries experience at the established ECC threshold and children had, on average, eight caries-affected tooth surfaces per child. The observed ECC prevalence is greater than national estimates. According to NHANES 2011–2016 data [7], 23% of 2–5-year-old children in the United States had caries experience at the cavitation threshold, and prevalence was 34% among children living below the federal poverty level. Our finding of higher ECC prevalence among Hispanic (61%) vs. non-Hispanic (52%) participants mirrors the nationally-representative data, wherein Mexican–Americans had higher ECC prevalence compared to non-Hispanic whites and African–Americans (33% versus 18% and 28%, respectively). However, it is noteworthy that non-Hispanic white and African–American children had identical ECC prevalence in our study.

A key feature of ZOE 2.0 is its community-based design and the inclusion of substantial proportions of children from minority groups—this is important from a public health standpoint, because estimates of ECC experience can be reasonably generalized to this demographic group of public preschool-attending children in North Carolina. This feature is also important for the planned oral health genomics studies, where the under-representation of racial/ethnic minorities is an important problem [23]. We plan to account for population stratification (i.e., ancestry) by computing 10–20 ancestry principal components using Genome-wide Complex Trait Analysis (GCTA) software [24]. We will carry out genome-wide association analyses using Genetic Association Analysis Under Complex Survey Sampling (SUGEN) software [25], implementing established [26,27] and recently developed strategies for multi-ethnic GWAS [28,29].

The study also benefits from its collection of detailed tooth surface-level clinical data, enabling both “deep and wide” types of analyses [30]—these may include interrogations of the consensus ECC case definitions [1] and harmonization with consortia [12] and databases with similar phenotypes and genetic data. At the same time, the already-collected, detailed clinical data and the stored biospecimens can be used to generate multi-omics information (i.e., human genome, metagenome, transcriptome, and metabolome) and thus enable the creation and study of more refined, data-driven or biologically informed disease subtypes. We have previously demonstrated that biologically informed or biologically enriched dental traits (e.g., including information from the oral microbiome) contain more heritable (i.e., genetically explained) variance, and as such may be more promising GWAS traits [31]. While this hypothesis has not been systematically tested for dental caries and ECC, we posit that information on microbial taxonomy, activity, and metabolism that can be obtained from the study’s supragingival plaque samples will enhance our ability to understand biological pathways contributing to ECC and its postulated subtypes [6,32,33]. It must also be stressed that genetic risk factors do not operate in a vacuum and interactions with important behavioral and environmental factors (e.g., consumption of sugar-containing snacks and beverages and exposure to optimally fluoridated water) are likely to be informative for the study of ECC. To facilitate data sharing, clinical, genomics, and microbiome data generated in ZOE 2.0 are being deposited in the dbGaP repository [34], under the umbrella study name Trans-Omics for Precision Dentistry and Early Childhood Caries or TOPDECC (accession: phs002232.v1.p1).

The low rate of families’ rerunning home water samples was a challenge that was identified during the first year of the study. Two mitigating strategies exist to address this missing data problem. First, the parent questionnaire includes a question about the primary source of home water for the child (e.g., tap, well, bottled water, etc.); this information is important because most well water in NC and most bottled water is non-fluoridated, whereas most community water in NC is fluoridated. Second, a geographic information systems (GIS)-based imputation method is being developed for the purpose of inferring other participants’ fluoride levels based on their residential address and the observed data.

Childhood oral disease has important, multi-level consequences for children, their families, the communities they live in, and the health systems that serve them [35]. ECC is both a person-level clinical problem and an important public health issue characterized by pronounced, persistent disparities [36]. It is thus important that efforts and resources are invested in both precision health-driven programs [37,38] that aim to study and operationalize individual susceptibility, as well as precision public health initiatives [22,39], aiming to inform or develop community interventions and policy. ZOE 2.0 aims to foster the generation of novel molecular, clinical, behavioral, and contextual insights into early childhood oral health and ECC. It is envisaged that this new information will serve as the basis of advancing precision oral health and care, and precision dental public health. Future longitudinal studies, prospectively evaluating ZOE 2.0 cohort participants have the potential to add to the knowledge base of dental caries incidence and progression, validate genetic, microbial, or metabolomic biomarkers collected at baseline, and provide valuable information on prospectively-assessed oral health outcomes in childhood and adolescence. Taken together, this information can facilitate the refinement of existing disease taxonomies and the improvement of current caries risk assessment tools [6,40].

Importantly, the ZOE 2.0 study is a community-based epidemiologic study of early childhood oral health that enables the interrogation of both upstream influences (e.g., geography, social, and area-level determinants) and proximal determinants (e.g., oral health-related behaviors, biomarkers, etc.). In fact, the availability of detailed social and biological information on several thousands of participants with comprehensive dental examination data, presents opportunities for the examination of the crosstalk or links between these different levels [41]. Aside from addressing important scientific questions, we envision that information generated from the study can also inform ECC surveillance, public health planning, and resource allocation. It may also serve as the basis for the development of preventive or therapeutic intervention studies, and ultimately influence actions at the policy level. We envision that the publication of the study's clinical, genomics, and microbiome protocols, as well as the public sharing of its human genomics and biofilm multi-omics data will facilitate the conduct of similar studies in other areas of the world.

5. Conclusions

The ZOE 2.0 cohort is a sizable, community-based, multi-ethnic, genetic epidemiologic study of early childhood oral health. The study's publicly available, de-identified clinical, behavioral, human genomics, and biofilm multi-omics data create a rich information source that can help answer important questions about the multi-level determinants of early childhood oral health.

Author Contributions: Conceptualization, K.D., G.D.S and K.E.N.; methodology, K.D., G.D.S, A.G.F.Z., J.S.P., H.C., Y.G., D.W. and K.E.N.; data acquisition, K.D., J.G., M.S.P. and B.D.M.; methods development; data analysis, M.A.S.-P., C.S.A., P.S., D.S.K., H.C. and Y.G.; laboratory methods development and biospecimen processing, A.d.A.R., M.A.A.-P. and P.V.B.; data curation, M.A.S.-P. and A.R.J.; writing—original draft preparation, K.D.; visualizations, K.D. and P.S.; funding acquisition, K.D., G.D.S., A.G.F.Z., J.S.P., D.W. and K.E.N. All authors have read, reviewed, contributed, and agreed to the published version of the manuscript.

Funding: This research was funded by NIH grant numbers U01DE025046, X01HG010871, and R03DE028983.

Acknowledgments: The investigators thank all participants and their families for their contributions, the study's regional coordinators and staff, project managers and volunteer workers, and the North Carolina Head Start staff for their support and assistance during the implementation of the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pitts, N.B.; Baez, R.J.; Diaz-Guillory, C.; Donly, K.J.; Alberto Feldens, C.; McGrath, C.; Phantumvanit, P.; Seow, W.K.; Sharkov, N.; Songpaisan, Y.; et al. Early Childhood Caries: IAPD Bangkok Declaration. *Int. J. Paed. Dentistry* **2019**, *29*, 384–386. [[CrossRef](#)]
2. Casamassimo, P.S.; Thikkurissy, S.; Edelman, B.L.; Maiorini, E. Beyond the dmft: The Human and Economic Cost of Early Childhood Caries. *J. Am. Dent. Assoc.* **2009**, *140*, 650–657. [[CrossRef](#)] [[PubMed](#)]

3. Peres, M.A.; Macpherson, L.M.D.; Weyant, R.J.; Daly, B.; Venturelli, R.; Mathur, M.R.; Listl, S.; Celeste, R.K.; Guarnizo-Herreño, C.; Kearns, C.; et al. Oral Diseases: A Global Public Health Challenge. *Lancet* **2019**, *394*, 249–260. [[CrossRef](#)]
4. Lee, J.Y.; Divaris, K. The Ethical Imperative of Addressing Oral Health Disparities: A Unifying Framework. *J. Dent. Res.* **2014**, *93*, 224–230. [[CrossRef](#)]
5. Pitts, N.B.; Zero, D.T.; Marsh, P.D.; Ekstrand, K.; Weintraub, J.A.; Ramos-Gomez, F.; Tagami, J.; Twetman, S.; Tsakos, G.; Ismail, A. Dental Caries. *Nat. Rev. Dis. Primers.* **2017**, *25*, 1–16. [[CrossRef](#)]
6. Divaris, K. Predicting Dental Caries Outcomes in Children: A “Risky” Concept. *J. Dent. Res.* **2016**, *95*, 248–254. [[CrossRef](#)]
7. Lin, M.; Griffin, S.O.; Gooch, B.F.; Espinoza, L.; Wei, L.; Li, C.H.; Thornton-Evans, G.; Junger, M.L.; Robison, V.A.; Fleming, E.B.; et al. *Oral Health Surveillance Report: Trends in Dental Caries and Sealants, Tooth Retention, and Edentulism, United States: 1999–2004 to 2011–2016*; Centers for Disease Control and Prevention, US Dept of Health and Human Services: Atlanta, GA, USA, 2019.
8. Divaris, K. Precision Dentistry in Early Childhood: The Central Role of Genomics. *Dent. Clin. North. Am.* **2017**, *61*, 619–625. [[CrossRef](#)]
9. Divaris, K. The Era of the Genome and Dental Medicine. *J. Dent. Res.* **2019**, *98*, 949–955. [[CrossRef](#)]
10. Morelli, T.; Agler, C.S.; Divaris, K. Genomics of Periodontal Disease and Tooth Morbidity. *Periodontology 2000* **2020**, *82*, 143–156. [[CrossRef](#)] [[PubMed](#)]
11. Ballantine, J.; Carlson, J.C.; Zandona, A.; Agler, C.S.; Zeldin, L.P.; Rozier, G.; Roberts, M.W.; Basta, P.V.; Luo, J.; Antonio-Obese, M.E.; et al. Exploring the Genomic Basis of Early Childhood Caries: A Pilot Study. *Int. J. Paed. Dentistry.* **2018**, *28*, 217–225. [[CrossRef](#)] [[PubMed](#)]
12. Haworth, S.; Shungin, D.; van der Tas, J.T.; Vucic, S.; Medina-Gomez, C.; Yakimov, V.; Feenstra, B.; Shaffer, J.R.; Lee, M.K.; Standl, M.; et al. Consortium-Based Genome-Wide Meta-Analysis for Childhood Dental Caries Traits. *Hum. Mol. Genet.* **2018**, *27*, 3113–3127. [[CrossRef](#)] [[PubMed](#)]
13. Shungin, D.; Haworth, S.; Divaris, K.; Agler, C.S.; Kamatani, Y.; Keun Lee, M.; Grinde, K.; Hindy, G.; Alaraudanjoki, V.; Pesonen, P. Genome-Wide Analysis of Dental Caries and Periodontitis Combining Clinical and Self-Reported Data. *Nat. Commun.* **2019**, *10*, 1–13. [[CrossRef](#)] [[PubMed](#)]
14. Shaffer, J.R.; Wang, X.; Feingold, E.; Lee, M.; Begum, F.; Weeks, D.E.; Cuenco, K.T.; Barmada, M.M.; Wendell, S.K.; Crosslin, D.R. Genome-Wide Association Scan for Childhood Caries Implicates Novel Genes. *J. Dent. Res.* **2011**, *90*, 1457–1462. [[CrossRef](#)]
15. Born, C.D.; Divaris, K.; Zeldin, L.P.; Rozier, R.G. Influences on Preschool Children’s Oral Health-Related Quality of Life as Reported by English and Spanish-Speaking Parents and Caregivers. *J. Public. Health. Dent.* **2016**, *76*, 276–286. [[CrossRef](#)]
16. Born, C.D.; Jackson, T.H.; Koroluk, L.D.; Divaris, K. Traumatic Dental Injuries in Preschool-Age Children: Prevalence and Risk Factors. *Clin. Exp. Dent. Res.* **2019**, *5*, 151–159. [[CrossRef](#)]
17. Ginnis, J.; Ferreira Zandoná, A.G.; Slade, G.D.; Cantrell, J.; Antonio, M.E.; Pahl, B.T.; Meyer, B.D.; Shrestha, P.; Simancas-Pallares, M.A.; Joshi, A.R.; et al. Measurement of Early Childhood Oral Health for Research Purposes: Dental Caries Experience and Developmental Defects of the Enamel in the Primary Dentition. *Methods. Mol. Biol.* **2019**, *1922*, 511–523. [[CrossRef](#)]
18. Ismail, A.I.; Sohn, W.; Tellez, M.; Amaya, A.; Sen, A.; Hasson, H.; Pitts, N.B. The International Caries Detection and Assessment System (ICDAS): An Integrated System for Measuring Dental Caries. *Community Dent. Oral Epidemiol.* **2007**, *35*, 170–178. [[CrossRef](#)] [[PubMed](#)]
19. Young, D.A.; Nový, B.B.; Zeller, G.G.; Hale, R.; Hart, T.C.; Truelove, E.L. The American Dental Association Caries Classification System for Clinical Practice: A Report of the American Dental Association Council on Scientific Affairs. *J. Am. Dent. Assoc.* **2015**, *146*, 79–86. [[CrossRef](#)] [[PubMed](#)]
20. Agler, C.S.; Shungin, D.; Ferreira Zandoná, A.G.; Schmadeke, P.; Basta, P.V.; Luo, J.; Cantrell, J.; Pahl, T.D., Jr.; Meyer, B.D.; Shaffer, J.R.; et al. Protocols, Methods, and Tools for Genome-Wide Association Studies (GWAS) of Dental Traits. *Methods Mol. Biol.* **2019**, *1922*, 493–509. [[CrossRef](#)] [[PubMed](#)]

21. Divaris, K.; Shungin, D.; Rodríguez-Cortés, A.; Basta, P.V.; Roach, J.; Cho, H.; Wu, D.; Ferreira Zandoná, A.G.; Ginnis, J.; Ramamoorthy, S. The Supragingival Biofilm in Early Childhood Caries: Clinical and Laboratory Protocols and Bioinformatics Pipelines Supporting Metagenomics, Metatranscriptomics, and Metabolomics Studies of the Oral Microbiome. *Methods Mol. Biol.* **2019**, *1922*, 525–548. [CrossRef] [PubMed]
22. Divaris, K.; Joshi, A. The Building Blocks of Precision Oral Health in Early Childhood: The ZOE 2.0 Study. *J. Public Health Dent.* **2020**, *80*, S31–S36. [CrossRef]
23. Agler, C.S.; Divaris, K. Sources of Bias in Genomics Research of Oral and Dental Traits. *Community Dent. Health.* **2020**, *37*, 102–106. [CrossRef] [PubMed]
24. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A Tool for Genome-Wide Complex Trait Analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. [CrossRef] [PubMed]
25. Lin, D.Y.; Tao, R.; Kalsbeek, W.D.; Zeng, D.; Gonzalez, F., 2nd; Fernández-Rhodes, L.; Graff, M.; Koch, G.G.; North, K.E.; Heiss, G. Genetic Association Analysis under Complex Survey Sampling: The Hispanic Community Health Study/Study of Latinos. *Am. J. Hum. Genet.* **2014**, *95*, 675–688. [CrossRef] [PubMed]
26. Pasaniuc, B.; Zaitlen, N.; Lettre, G.; Chen, G.K.; Tandon, A.; Kao, W.H.; Ruczinski, I.; Fornage, M.; Siscovick, D.S.; Zhu, X. Enhanced Statistical Tests for GWAS in Admixed Populations: Assessment using African Americans from CARE and a Breast Cancer Consortium. *PLoS Genet.* **2011**, *7*, e1001371. [CrossRef] [PubMed]
27. Mägi, R.; Horikoshi, M.; Sofer, T.; Mahajan, A.; Kitajima, H.; Franceschini, N.; McCarthy, M.I.; COGENT-Kidney Consortium; T2D-GENES Consortium; Morris, A.P. Trans-Ethnic Meta-Regression of Genome-Wide Association Studies Accounting for Ancestry Increases Power for discovery and Improves Fine-Mapping Resolution. *Hum. Mol. Genet.* **2017**, *26*, 3639–3650. [CrossRef]
28. Baldassari, A.R.; Sitlani, C.M.; Highland, H.M.; Arking, D.E.; Buyske, S.; Darbar, D.; Gondalia, R.; Graff, M.; Guo, X.; Heckbert, S.R. Multi-Ethnic Genome-Wide Association Study of Decomposed Cardioelectric Phenotypes Illustrates Strategies to Identify and Characterize Evidence of Shared Genetic Effects for Complex Traits. *Circ. Genom. Precis. Med.* **2020**, *13*. [CrossRef]
29. Wojcik, G.L.; Graff, M.; Nishimura, K.K.; Tao, R.; Haessler, J.; Gignoux, C.R.; Highland, H.M.; Patel, Y.M.; Sorokin, E.P.; Avery, C.L. Genetic Analyses of Diverse Populations Improves Discovery for Complex Traits. *Nature* **2019**, *570*, 514–518. [CrossRef]
30. Divaris, K. Searching Deep and Wide: Advances in the Molecular Understanding of Dental Caries and Periodontal Disease. *Adv. Dent. Res.* **2019**, *30*, 40–44. [CrossRef]
31. Agler, C.S.; Moss, K.; Philips, K.H.; Marchesan, J.T.; Simancas-Pallares, M.; Beck, J.D.; Divaris, K. Biologically Defined or Biologically Informed Traits Are More Heritable Than Clinically Defined Ones: The Case of Oral and Dental Phenotypes. *Adv. Exp. Med. Biol.* **2019**, *1197*, 179–189. [CrossRef]
32. Zeng, Z.; Feingold, E.; Wang, X.; Weeks, D.E.; Lee, M.; Cuenco, D.T.; Broffitt, B.; Weyant, R.J.; Crout, R.; McNeil, D.W.; et al. Genome-Wide Association Study of Primary Dentition Pit-and-Fissure and Smooth Surface Caries. *Caries Res.* **2014**, *48*, 330–338. [CrossRef] [PubMed]
33. Psoter, W.J.; Pendry, D.G.; Morse, D.E.; Zhang, H.P.; Mayne, S.T. Caries Patterns in the Primary Dentition: Cluster Analysis of a Sample of 5169 Arizona Children 5–59 Months of age. *Int. J. Oral Sci.* **2009**, *1*, 189–195. [CrossRef] [PubMed]
34. National Center for Biotechnology Information, U.S. National Library of Medicine, Database of Genotypes and Phenotypes (dbGaP). Available online: <https://www.ncbi.nlm.nih.gov/gap/> (accessed on 31 October 2020).
35. Casamassimo, P.S.; Lee, J.Y.; Marazita, M.L.; Milgrom, P.; Chi, D.L.; Divaris, K. Improving Children’s Oral Health: An Interdisciplinary Research Framework. *J. Dent. Res.* **2014**, *93*, 938–942. [CrossRef]
36. Slade, G.D.; Sanders, A.E. Two Decades of Persisting Income-Disparities in Dental Caries Among U.S. Children and Adolescents. *J. Public Health Dent.* **2018**, *78*, 187–191. [CrossRef] [PubMed]
37. Collins, F.S.; Varmus, H. A New Initiative on Precision Medicine. *N. Engl. J. Med.* **2015**, *372*, 793–795. [CrossRef] [PubMed]
38. Divaris, K. Fundamentals of Precision Medicine. *Compend. Contin. Educ. Dent.* **2017**, *38*, 30–32. [PubMed]
39. Khoury, M.J.; Iademarco, M.F.; Riley, W.T. Precision Public Health for the Era of Precision Medicine. *Am. J. Prev. Med.* **2016**, *50*, 398–401. [CrossRef]

40. Halasa-Rappel, Y.A.; Ng, M.W.; Gaumer, G.; Banks, D.A. How Useful are Current Caries Risk Assessment Tools in Informing the Oral Health Care Decision-Making Process? *J. Am. Dent. Assoc.* **2019**, *150*, 91–102. [[CrossRef](#)]
41. Gomaa, N.; Glogauer, M.; Tenenbaum, H.; Siddiqi, A.; Quiñonez, C. Social-Biological Interactions in Oral Disease: A ‘Cells to Society’ View. *PLoS ONE* **2016**, *11*, e0146218. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).