Overview and Quality Assurance for the Oral Health Component of the National Health and Nutrition Examination Survey (NHANES), 2003-04

B. A. Dye, DDS, MPH; R. Nowjack-Raymer, MPH, PhD; L. K. Barker, MSPH; J. H. Nunn, PhD, MA, BDS; J. G. Steele, PhD, BDS; S. Tan, MS; B. G. Lewis, MPH; E. D. Beltran-Aguilar, DMD, MPH, MS, DrPH

Abstract

The 2003-04 National Health and Nutrition Examination Survey (NHANES) was a collaborative effort involving 28 federal funding partners with the National Center for Health Statistics. The collaborators for the 2003-04 NHANES oral health component included the National Institute of Dental and Craniofacial Research and the National Center for Chronic Disease Prevention and Health Promotion, Division of Oral Health. Oral health data are available on 8,272 persons aged 2 years or older. This report provides an overview of the 2003-04 oral health component including content descriptions and procedures for oral health assessments conducted for the first time in a national survey in the United States. These assessments include posterior functional contacts, tooth wear, and oral health-related quality of life. This report also provides evaluations of data quality in terms of examiner reliability statistics (percent agreements, kappas, and correlation coefficients) for various NHANES 2003-04 oral health examination components and analytical recommendations for producing 6-year estimates using the previous two NHANES data collection components (1999-2000 and 2001-02).

Key Words: NHANES, oral health, dental public health, epidemiology, data reliability, quality assurance

Introduction

Periodic national health examination surveys have been conducted in the United States by the Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS) for more than four decades. When a new data collection cycle began in 1999, the National Health and Nutrition Examination Survey (NHANES) was changed from a periodic survey to a continuous, annual survey. Current NHANES data are released in 2-year nationally representative periods to protectthe confidentiality of survey participants and increase statistical reliability. The NHANES 2003-04 followed the same protocol as the NHANES 1999-2000 and 2001-02, but included some minor modifications that are described in this report. Additionally, an expanded oral health section was added to the NHANES 2003-04 that produced oral health information on conditions that have never been assessed in prior national health surveys in the United States. This included assessments on functional occlusal contacts, tooth wear, and oral health-related quality of life.

The NHANES data set is a unique national and international resource, and its availability allows researchers to explore many aspects of health for a national sample. Furthermore, it allows comparisons with other national populations, with the potential to explore the impact of national and cultural characteristics on health. The aim of this methodological paper is to describe the various components of the oral health exam in 2003-04 and the steps taken to ensure data quality, including statistical analyses of reliability, to ensure that researchers are in a position to use the data properly for future research.

NHANES 2003-04 Overview. The civilian, noninstitutionalized population living in the 50 United States and the District of Columbia was the target population for the NHANES 2003-04. The survey used a stratified, multistage probability sampling design to select study participants. Oversampling was performed on selected subgroups, such as non-Hispanic Blacks, Mexican-Americans, persons aged 12-19 years and 60 years or older, to facilitate the calculation of more precise estimates for these groups. All study participants gave informed consent as

Send correspondence and reprint requests to Bruce Dye, DDS, MPH, CDC/NCHS/NHANES Program, 3311 Toledo Road, RM 4416, Hyattsville, MD 20782. Tel.: 301-458-4199; Fax: 301-458-4029; e-mail: bfd1@cdc.gov. B. A. Dye is with the Centers for Disease Control and Prevention/National Center for Health Statistics, Hyattsville, MD and the University of Maryland School of Dentistry, Baltimore, MD. S. Tan and B. G. Lewis are with the Centers for Disease Control and Prevention/National Center for Health Statistics, Hyattsville, MD. R. Nowjack-Raymer is with the National Institutes of Health/National Institute of Dental and Craniofacial Research, Bethesda, MD. L. K. Barker and E. D. Beltran-Aguilar are with the Centers for Disease Control and Prevention/National Center for Chronic Disease Prevention and Health Promotion/Division of Oral Health, Atlanta, GA. J. H. Nunn is with the University of Dublin Trinity College/Dental School and Hospital/Department of Public and Child Dental Health, Dublin, Ireland. J. G. Steele is with the University of Newcastle upon Tyne/Dental School/Department of Restorative Dentistry, Newcastle upon Tyne, UK.
Manuscript received: 4/3/07; accepted for publication: 9/2/07. **Ethics statement:** All study participants gave informed consent in accordance with the Institutional Review Board and study ethic guidelines at the Centers for Disease Control and Prevention. The authors do not have any financial or other competing interests to declare.

DOI: 10.1111/j.1752-7325.2007.00076.x

Table 1
Comparison of Sampling Design Characteristics for Previous and Current National Health and Nutrition Examination Surveys (NHANES) with Oral Health Content

Characteristic	NHANES 1999-2000	NHANES 2001-02	NHANES 2003-04
Data collection year	1999-2000	2001-02	2003-04
Age of the target population	From birth	From birth	From birth
Dental exam age	2 years or older	2 years or older	2 years or older
Number of survey exam locations	27	30	30
Eligible geographical areas for sample	50 states + DC	50 states + DC	50 states + DC
Groups targeted for oversampling	Persons 12-19 years; non-Hispanic Blacks; Mexican-Americans; those 60 years or older; and others*	Persons 12-19 years; non-Hispanic Blacks; Mexican-Americans; those 60 years or older; low-income Whites; and others	Persons 12-19 years; non-Hispanic Blacks; Mexican-Americans; those 60 years or older; low-income Whites; and others
Total interviewed	9,965	11,039	10,122
Total mobile examination center examined	9,282	10,477	9,643
Total with an oral examination record	8,082	9,010	8,272

* Low-income Whites were oversampled only for the NHANES 2000.
DC, District of Columbia.

approved by the NCHS Ethics Review Board. Sample design characteristics for the current NHANES are shown in Table 1. Additional information on the background and content of the NHANES can be found at http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm

NHANES 2003-04 Oral Health Component. The NHANES 2003-04 oral health component was a collaborative effort between the National Institute of Dental and Craniofacial Research (NIDCR), National Center for Chronic Disease Prevention and Health Promotion, Division of Oral Health, and NCHS. The oral health assessment included a group of questions asked during the home interview, followed by a second group of questions and dental examinations at the mobile examination center (MEC). Eligibility for a specific oral health assessment was age based and varied among study participants. Additional details pertaining to study design and participant selection for the NHANES 1999-2002 have been described elsewhere (1).

The NHANES 2003-04 expanded oral health assessments were conducted during the home interview and at the MEC. Although the target population for the oral health component during 2003-04 was persons 2 years of age or older, the expanded assessments include only adolescents and adults. The age eligibility requirements for the various oral health assessments are displayed in Table 2. The number of participants completing various components of the NHANES 2003-04 is shown in Table 3. Among adults aged 20 years or older who completed the home interview portion of the NHANES 2003-04, approximately 86 percent completed the oral health exam, whereas 90 percent of youths aged 2 to 11 years and nearly 94 percent of adolescents aged 12 to 19 years completed the oral health exam.

Home Interview Oral Health Assessment Healthy People 2010 is the third in a series of 10-year health promotion, disease prevention, and health goals for the nation. One of two overarching goals for *Healthy People 2010* is to increase the years and quality of life for individuals of all ages (2). Measures of oral health-related quality of life have been developed, tested, and applied in numerous countries (3-6). Oral health-related quality of life studies have provided health, health service, and health policy researchers with important information about the impact of oral disease and dental services on quality of life. The NHANES 2003-04 oral health quality of life assessment consisted of seven questions, which were derived from the 14-item Oral Health Impact Profile instrument (6), and was part of a series of oral health questions administered during the home interview. Additional questions new to the NHANES 2003-04 oral health section of the questionnaire inquired into perceived treatment needs and reasons for needing dental extractions, as well as reasons why dental care was needed in the past year but was not obtained.

MEC Oral Health Assessment. The NHANES 2003-04 oral health exam was performed by trained dentists with a dental license in at least one US jurisdiction. Other MEC personnel were trained as dental recorders. All NHANES data were collected by a dental examination team in each of the concurrently functioning MECs. The two primary dental examiners assigned to the MEC teams per-

Table 2
Oral Health Subcomponents Age Eligibility and Age Range of Participants for Interrater Evaluation, NHANES 2003-04

Assessment	Age in years	Analytical sample age range
Home interview		
Dental health perception, dental visits, and dental care utilization*	2+	
Dental mouth and chewing problems	18+	
Oral health quality of life*	16+	
Perceived dental treatment needs*	18+	
Mobile examination center exam		
Medical history screening	13+	
Denture use and wear questions*	25+	
Self-assessed dental condition question*	16+	
Tooth count	2+	5-90
Coronal caries	2+	5-90
Root caries	18+	18-90
Dental sealants	2-34	5-34
Incisor trauma	6-29	7-29
Dental fluorosis	6-49	7-49
Dental erosion and tooth wear*	13+	13-87
Functional occlusal contacts*	25+	25-87
Periodontal assessment	13+	13-87
Bleeding from probing	13+	13-87
Recommendation for care	2+	

* Data collected from these assessments are available only from the NHANES 2003-04 and are not in the NHANES 1999-2002.
NHANES, National Health and Nutrition Examination Surveys.

Table 3
Number of Sampled Persons Aged 2 Years or Older Completing the Interview and MEC and Oral Health Exams for the National Health and Nutrition Examination Survey (NHANES) 2003-04, by Selected Demographic Characteristics‡

Characteristic	Number of sample persons*			Number of participants interrater evaluation
	HIQ	MEC	OHX	
Age (years)				
2-5	926	865	790	0
6-11	1,034	992	973	16
12-19	2,303	2,248	2,155	66
20-39	1,742	1,656	1,501	63
40-59	1,398	1,336	1,221	46
60+	1,901	1,750	1,632	45
Sex				
Male	4,540	4,316	4,065	107
Female	4,764	4,531	4,207	129
Race/ethnicity†				
Non-Hispanic Blacks	2,463	2,356	2,220	64
Mexican-Americans	2,247	2,174	2,015	43
Non-Hispanic Whites	3,870	3,642	3,427	112
Others	724	675	610	17
Total	9,304	8,847	8,272	236

* Number of sampled persons completing a home interview questionnaire (HIQ), a mobile examination center (MEC) examination, and an oral health examination (OHX).
† RIDRETH1 race/ethnicity variable.
‡ Includes the number of persons participating in the OHX interrater evaluation during 2003-04.

formed nearly 95 percent of all the dental examinations undertaken in 2003-04. Trained interviewers administered the oral health questions during the home interview. For the 2003-04 data collection cycle, all primary and backup dental examiners were trained and calibrated by the author (B. A. Dye).

The sequence of assessments administered during the oral health exam is listed in order as shown in Table 3. The first assessment was a medical history inquiry conducted by the dentist to determine if a participant was to be excluded from the root caries and periodontal assessments. In 2003-04, 9.4 percent of the assessed adolescents and adults were excluded from some of the oral health exams because of medical history concerns. Among those excluded, approximately 40 percent was because of a positive response to the question asking about the need for antibiotic prophylaxis. Questions on denture use and wearing of dentures followed the medical screening. Additional questions new to the NHANES 2003-04 oral health exam included asking participants to rate the condition of their teeth using a five-point scale. If the response was "good," "fair," or "poor," a follow-up question was asked to assess for specific problems that the respondent may have had with their teeth.

The oral health exam continued with the same sequence used in the NHANES 1999-2002. These procedures included: a tooth count to identify the presence or absence of permanent and/or primary teeth, including the presence of surgical dental implants and retained dental root tips, for each of the 32 tooth positions; assessments for coronal caries, dental restorations, and replacement; and assessments for root caries, dental sealants, incisal trauma, and dental fluorosis. Following the dental fluorosis assessment, two new examination procedures were performed: an evaluation for tooth wear and an assessment of the quantity and quality of functional occlusal contacts. At the conclusion of the oral health examination, the

dental examiner assigned each survey participant to one of four care levels (immediate, urgent, earliest convenience, and routine) and identified generalized conditions (such as dental caries, oral hygiene instruction, periodontal needs, denture needs, etc.) that required attention. Additional information describing the examination methods for the oral health component is available at http://www.cdc.gov/nchs/about/major/nhanes/nhanes2003-2004/current_nhanes_03_04.htm

New exam assessments: tooth wear and functional contacts. Tooth wear can affect individuals across the life span. Many of the suggested etiologies include dietary factors, medications, health conditions, and socioeconomic status (7-14). Epidemiologic assessments of dental erosion have used the Smith and Knight Tooth Wear Index (TWI) (15) and its subsequent modifications (16). The NHANES 2003-04 protocol included a TWI derived from the 1998 United Kingdom Adult Dental Health Survey (17). The assessment required the visual examination of the facial, lingual, and incisal surfaces of all 12 anterior teeth and the occlusal surfaces of the first molars. Tooth wear scores (TWS) were based upon the extent of enamel loss and dentin exposure for each eligible tooth surface.

The functional contacts assessment was incorporated into the NHANES to fully describe the functional capacities of the dentition, with the intent to assess the impact of dental status on oral health-related quality of life. This exam was derived from a similar assessment used in the 1998 United Kingdom Adult Dental Health Survey (18) and counted the number of posterior functional contacts in such a way as to quantify an important aspect of the functional status of the dentition that simple counts of teeth and prostheses alone cannot provide. The occlusal contact assessment consisted of evaluating a total of 16 posterior zones: eight in the right premolar and molar region and eight in the left premolar and molar region, with the participant in centric occlusion. Coding for a functional contact was qualitatively based upon an assumed hierarchical order, that is, contact between upper and lower natural teeth, contact between a natural tooth and a prosthetic tooth, contact between upper and lower prosthetic teeth, and no contact present. A limited anterior assessment was performed requiring the evaluation of at least one anterior mandibular tooth in contact with an opposing anterior maxillary tooth to produce the "best" score for the anterior region. A measurement of maximal incisal opening was also made in whole millimeters using an endodontic ruler.

The protocol development and initial training of the tooth wear and functional contact assessments involved an international collaborative effort among researchers and oral epidemiologists from the University of Dublin Trinity College (Ireland) Dental School and Hospital (J. H. Nunn), the University of Newcastle upon Tyne (UK) Dental School (J. G. Steele), the NIDCR, and the NCHS. The NHANES reference examiner (B. A. Dye) received training and standardization at the Dublin Dental School and Hospital in July 2002. Professors Nunn and Steele led the training and reliability assessments in tooth wear and functional contacts, respectively. Both had served as content experts, trainers, and reference examiners in prior national oral health surveys in the UK. All training and calibration participants had been prescreened and selected based upon their case definitions to ensure a broad representation of conditions. Interrater reliability across all tooth sites for the tooth wear sessions yielded a percent agreement of 87.7 percent and a kappa of 0.80. For the functional contacts session, a 96.8 percent agreement and a kappa of 0.95 was achieved.

Statistical Issues. When using the NHANES data, complex estimation procedures are necessary to calculate valid population estimates because of the complex survey design and unequal probability for participant selection. The NHANES public-use data sets include individual sample weights for the interview and MEC examination for each of the 2-year data collection periods to account for the unequal probability of selection and nonresponse. If the analysis uses only information obtained from the interview sample, then the appropriate weight variable to use is WTINT2YR. If data are used from the MEC examination, alone or in combination with the home interview data, then WTMEC2YR would be the appropriate weight variable to use.

There are some important considerations for investigators when combining the 2-year NHANES data sets for analyses. The 1999-2000 sample weights were developed using population estimates provided by the Bureau of the Census before the Year 2000 Census counts became available, whereas after 2001, sample weights were created by incorporating information from the Year 2000 Census. Consequently, a 4-year sample weight, such as the variable WTMEC4YR for the MEC examination data, was created by the NCHS for analyses of the NHANES 1999-2002 data.

For analyses requiring the combination of other 2-year data sets, the NCHS has not provided additional special weights in the public-release data files. However, guidance has been provided to allow analysts to derive unique sample weights appropriate for a suitable multiyear analysis. For instance, when analyzing the 2001-04 data, an appropriate 4-year sample weight can be derived by taking half the value of the 2001-02 weight and half of the 2003-04 weight for each sampled person. If a 6-year weight variable is desired, two-thirds of the 1999-2002 special 4-year sample weight and one-third of the 2003-04 weight for each sampled person should be used. Because of the stratification and clustering of the NHANES samples, appropriate sample weights must be used to calculate proper estimates and variance statistics.

Masked variance units (MVUs) have been created from the primary

sampling unit (PSU) information for variance estimation. These MVUs are available for each 2-year period (1999-2000, 2001-02, and 2003-04), and their use is recommended for calculating sampling errors. Software such as SUDAAN (RTI, Research Triangle Park, NC, USA) or STATA (StataCorp, College Station, TX, USA) can be used to estimate sampling errors by the Taylor series method. The MVU (derived PSU) variable is SDMVPSU, and the stratum variable is SDMVSTRA. Moreover, investigators should be cognizant of the issues relevant to reporting estimates from small sample sizes. The NHANES data may be considered statistically unreliable when an estimate is based on fewer than 30 sample units in the denominator or if the relative standard error is greater than 30 percent. More detailed analytical guidelines for the NHANES are available at http://www.cdc.gov/nchs/about/major/nhanes/nhanes2003-2004/analytical_guidelines.htm

Modifications and Technical Notes. Because the stock of usable NIDR periodontal probes, which had been used since the NHANES III, was critically low and the probe was no longer commercially available, a similar probe (Hu-Friedy PCP2, Hu-Friedy Manufacturing Company, Chicago, IL, USA) was used for periodontal examinations in the NHANES 2003-04. The PCP2 is color-banded and graduated at 2, 4, 6, 8, 10, and 12 mm. Periodontal measurements were made at the distofacial, midfacial, and mesiofacial sites on all eligible permanent teeth following the 2001-02 periodontal protocols.

Like the NHANES 1999-2002 data files, the NHANES 2003-04 oral health data sets are organized into "chapters." For the 2003-04 data cycle, there are six chapters: Dentition Exam, Dentition-Addendum Exam, Periodontal Exam – Upper, Periodontal Exam – Lower, Miscellaneous, and Home Interview. The tooth wear and functional contacts data are located in the Dentition-Addendum Exam data file, and data pertaining to oral health-related quality of life are located in the Home Interview data file.

Additional information including valid questionnaire responses, code values, and criteria for assessment of dental conditions including tooth wear and functional contacts, as well as the 2003-04 oral health data sets are available at http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm. Oral health derived variables in major reports and SAS sample codes are available from the NIDCR/CDC Dental, Oral, and Craniofacial Data Resource Center at http://drc.hhs.gov/

Interrater Evaluation. Dental examiners undergo a comprehensive training and standardization period, including periodic field reviews and repeat examinations, to ensure quality oral health data. Details of the training procedures are available elsewhere (1). For the 2003-04 NHANES data collection cycle, second examinations conducted by the primary examiners were not performed. Consequently, the only reliability statistics available to assess data quality and examiner performance were derived from the repeat exams performed by the reference examiner on the study participants (interrater).

Interrater data were collected on a convenience sample of NHANES study participants. The age range of persons participating in the interrater evaluation is presented in Table 2, and Table 3 shows the number of participants by selected demographic characteristics. During 2003-04, the reference examiner evaluated the two primary dental examiners approximately four times each during regular MEC operations. Site visits were evenly spaced and all sample participants scheduled for MEC examinations were eligible for the interrater reliability evaluation during these site visits. The reference examiner completed a repeat dental exam immediately following a standard oral health exam conducted by the primary dental examiner. The reference examiner determined which study participants to examine, and both examiners were blinded to each other's observations. Given the constraints of appointing and examining study participants on the MEC, the primary dental examiners were aware of when interrater evaluations were to be performed. Although the calculated interrater agreements could have been affected by this knowledge, the interrater assessments were designed to minimize the potential for examiner bias.

The interrater reliability statistics produced for this report were percent agreement, kappa statistics, and interclass correlation coefficients (ICCs). SAS software (version 8.02, SAS Institute Inc., Cary, NC) was used to calculate the weighted and nonweighted kappa statistics for a number of oral health conditions, including *Healthy People 2010* derived oral health variables. These include caries experience and untreated decay in the primary and permanent dentitions, presence of dental sealants, tooth retention, and periodontal disease. Kappa statistics also were produced for incisal trauma and dental fluorosis using the same process as reported in the NHANES 1999-2002 quality assurance report (1). Evaluating inter-examiner reliability for the presence of at least one residual dental root also was performed by calculating the kappa scores. Kappa statistics are presented only if the number of observations was >1 for the concordant cells. Weighted kappa statistics were calculated for dental fluorosis to account for the difficulty in differentiating lower levels of fluorosis, particularly between questionable and very mild dental fluorosis. A weight of "1" was applied to observations that were in perfect agreement, a weight of "$\frac{2}{3}$" was applied to observations that were one category apart, and a weight of "$\frac{1}{3}$" was applied to observations that were two categories apart. Observations with values "cannot assess" were excluded from the calculations.

Additional kappa statistics for periodontal status were calculated for the prevalence of at least 4 mm or greater pocket depth and for a derived periodontitis variable that

Table 4
Summary of Inter-Rater Reliability Statistics (Primary Dental Examiner–Reference Examiner Comparisons) for Selected Oral Health Characteristics: National Health and Nutrition Examination Survey, 2003-04

	2003-04 interrater			
Characteristic	*n*•	% agree	Kappa	ASE
Untreated primary decay*				
Examiner A	15	86.7	0.67	0.21
Examiner B	23	100.0	1.00	0.00
Primary caries experience*				
Examiner A	15	93.3	0.86	0.13
Examiner B	23	95.7	0.88	0.12
Untreated permanent decay*				
Examiner A	96	87.5	0.65	0.09
Examiner B	139	94.2	0.73	0.09
Permanent caries experience*				
Examiner A	96	92.7	0.84	0.06
Examiner B	139	96.4	0.91	0.04
Dean's Fluorosis Index†‡				
Examiner A	61	41.0	0.56	0.07
Examiner B	117	64.1	0.69	0.05
Incisor trauma¶				
Examiner A	60	99.8	1.00	0.01
Examiner B	89	99.9	1.00	0.01
Sealant prevalence*				
Examiner A	56	98.2	0.96	0.04
Examiner B	92	98.9	0.98	0.02
Periodontal disease*				
Examiner A	70	94.3	0.72	0.13
Examiner B	99	90.9	0.64	0.11
Periodontitis§				
Examiner A	70	92.9	~	~
Examiner B	99	93.9	0.64	0.13
Residual dental root				
Examiner A	96	99.1	0.94	0.06
Examiner B	139	100	1.00	0.00
Tooth retention†				
Examiner A	96	100.0	1.00	0.00
Examiner B	139	97.8	0.94	0.03
Posterior functional contacts (five or less)				
Examiner A	57	100.0	1.00	0.00
Examiner B	77	100.0	1.00	0.00
Presence of contact in right molar zone				
Examiner A	57	100.0	1.00	0.00
Examiner B	77	98.7	0.95	0.05
Presence of contact in right premolar zone				
Examiner A	53	98.1	0.91	0.09
Examiner B	74	98.6	0.93	0.07
Presence of contact in left molar zone				
Examiner A	57	98.2	0.94	0.06
Examiner B	77	98.7	0.96	0.04
Presence of contact in left premolar zone				
Examiner A	53	96.2	0.78	0.15
Examiner B	74	98.6	0.93	0.07

has been used in prior NHANES data analyses (19,20). A participant was identified as having periodontal disease if they had at least one periodontal site with 3 mm or more loss of attachment and 4 mm or more pocket depth at the same site. Periodontal status interrater reliability also was assessed by comparing ICCs generated from subject-level means (mm) for loss of attachment and pocket depth using measurements obtained from the three periodontal sites: mesiofacial, midfacial, and distofacial.

For this report, five different posterior functional contact summary variables were used to report interrater reliability statistics. Kappa statistics for the presence of at least one contact in the premolar and molar zones for each of the right and left sides were calculated. A similar reliability statistic was calculated for the presence of five or fewer posterior functional contacts. ICCs were calculated for total posterior functional contacts, total posterior denture contacts, and the number of third molars present. Tooth wear was assessed by calculating kappa statistics for the presence of at least one surface with severe loss of enamel and dentin affecting the pulp (TWS = 3), with extensive loss of enamel and dentin (TWS = 2), and with moderate loss of enamel and dentin (TWS = 1). ICCs were also produced for total sites with a TWS = 3 and TWS = 2.

Additional reliability statistics for the oral health component are available from the authors. For oral health data quality comparisons between NHANES cycles since 1999, interested readers can refer to previously published information (1). The dentist identified as Examiner B for this 2003-04 report was previously identified as Examiner C for the 2001-02 data collection cycle.

Interrater Results. The interrater reliability statistics for the NHANES 2003-04 oral health examination are shown in Table 4. The statistics shown are for the two primary dental examiners employed during the 2003-04 data collection cycle. The kappa scores ranged from 0.65

Table 4
Continued

Characteristic	2003-04 interrater			
	n•	% agree	Kappa	ASE
Severe loss of enamel/ dentin (at least one tooth)				
Examiner A	83	98.8	0.85	0.15
Examiner B	117	98.3	~	~
Extensive loss of enamel/ dentin (at least one tooth)				
Examiner A	83	96.4	0.71	0.16
Examiner B	117	97.4	0.84	0.09
Moderate loss of enamel/ dentin (at least one tooth)				
Examiner A	83	94.0	0.83	0.07
Examiner B	117	85.5	0.71	0.06

* *Healthy people 2010* objective.
† Recommended derived variable – code available from the Dental, Oral, and Craniofacial Data Resource Center.
‡ Weighted kappa.
¶ Pooled kappa.
§ Periodontitis defined as having at least 3 mm of attachment loss and 4 mm of pocket depth at the same site.
• Number of paired observations.
ASE, asymptotic standard error (kappa). ~, at least one concordant cell did not have an observation >1, kappa was not calculated.

Table 5
Dental Examiner Interclass Correlation Coefficients (ICCs) for Selected Oral Health Characteristics: National Health and Nutrition Examination Survey, 2003-04

Characteristic	2003-04 interrater			
	n†	Reference‡ mean	Examiner mean	ICC
LOA mean*				
Examiner A	70	0.52	0.75	0.86
Examiner B	99	0.50	0.33	0.93
PD mean*				
Examiner A	70	0.71	0.67	0.61
Examiner B	99	0.83	0.74	0.86
Number of third molars present				
Examiner A	96	0.68	0.67	0.94
Examiner B	139	0.71	0.70	0.99
Total posterior functional contacts				
Examiner A	57	9.53	9.63	0.98
Examiner B	77	9.57	9.39	0.98
Total posterior denture contacts				
Examiner A	57	1.68	1.70	1.00
Examiner B	77	0.86	0.83	1.00
Total sites with severe or extensive tooth wear				
Examiner A	83	0.25	0.21	0.85
Examiner B	117	0.25	0.22	0.95

* Overall mean is for the combined mesio-, mid-, and distofacial periodontal sites for loss of attachment (LOA) and pocket depth (PD).
† Number of paired observations.
‡ Reference examiner's mean statistics.

to 1.00 for dental caries in the permanent and primary dentitions. For tooth retention, dental sealants, and dental fluorosis, the kappa scores varied from 0.94 to 1.00, 0.96 to 0.98, and 0.56 to 0.69, respectively. For the *Healthy People 2010* variable prevalence of periodontal disease in the United States (loss of attachment ≥4 mm at one or more sites), the kappa scores were 0.64 and 0.72 for each examiner. The kappa scores for periodontitis using the three case definitions ranged from 0.64 to 0.82 for Examiner B. Because of an insufficient number of periodontitis cases, we were unable to calculate the kappa statistics for Examiner A. For the various reliability assessments for functional posterior contacts and tooth wear, the kappa values ranged from 0.91 to 1.00 for contacts and from 0.71 to 0.85 for tooth wear.

Table 5 shows the interrater reliability statistics expressed as ICCs. For loss of attachment and pocket depth, the ICCs ranged from 0.86 to 0.93 and 0.61 to 0.86, respectively. An evaluation of mean differences in loss of attachment and pocket depth between the reference examiner and each primary examiner produced differences in measurement means across all periodontal sites ranging from 0.02 to 0.45 mm for loss of attachment and from 0.04 to 0.14 mm for pocket depth. For total posterior functional contacts and total posterior denture contacts, the ICC values ranged from 0.98 to 1.00. The ICCs for severe or extensive tooth wear were 0.85 and 0.95.

Discussion

This report presents a broad range of data-quality issues for the NHANES 2003-04 oral health component. Two primary dental examiners collected the majority of the 2003-04 oral health data, compared with five examiners during 1999-2002. Examiner concordance for the NHANES 2003-04 is consistent with the values reported for 1999-2002 in an earlier publication (1). For caries experience in the permanent dentition, kappa scores ranged from 0.85 to 0.96 during 1999-2002 and from 0.84 to

0.91 in 2003-04. Likewise, the kappa scores for periodontal disease varied from 0.58 to 0.77 during 1999-2002 and from 0.64 to 0.72 in 2003-04. The ICCs for mean pocket depth were similar as well. Coefficient values ranged from 0.55 to 0.87 during 1999-2002 and from 0.61 to 0.86 in 2003-04. However, the ICC values for mean loss of attachment were slightly higher in 2003-04 (0.86 to 0.93) compared with 1999-2002 (0.72 to 0.89).

This report also provides interrater reliability statistics for functional occlusal contacts and tooth wear, which are examination assessments new to the NHANES. Following published guidance for evaluating the strength of interexaminer agreement (21,22), the functional contacts interrater kappa scores for both the left and right premolar and molar areas indicate that excellent agreement was achieved for this assessment. Using the same guidance, excellent agreement was achieved for the more advanced levels of tooth wear.

We have presented a variety of reliability statistics for the NHANES 2003-04 oral health data. However, the interpretation of these reliability statistics should depend upon the context of the study design, data use, and clinical relevance. For the 2003-04 periodontal data, the ICC statistics ranged from 0.61 to 0.86 for pocket depth. Although it has been suggested that a threshold of 0.75 or greater would represent excellent reliability, examiner bias (the mean difference between reference examiner and survey examiners) was between 0.04 and 0.09 mm. This level of bias is comparable with previous bias estimates for mean pocket depth measures obtained during the NHANES III (absolute value ranged from 0.03 to 0.18 mm) (23). The susceptibility of the periodontal measurement data to measurement error is an important factor when evaluating quality control data and when comparing such data with other data sources. For example, obtaining pocket depth measurements requires consistent angulation and probe pressure by the examiners. Moreover, the periodontal probes are marked in 2-mm increments, and the examiners are trained to round down to the nearest millimeter.

Another difficult oral health assessment from the perspective of evaluating data reliability is enamel fluorosis. Fluorosis was assessed in all permanent teeth by assigning one of six categories (unaffected, questionable, very mild, mild, moderate, or severe) to each tooth using Dean's Fluorosis Index (24). For the NHANES 2003-04, the weighted kappa scores range from 0.56 to 0.69, indicating moderate to good agreement. For the NHANES 1999-2002, the weighted kappa scores ranged from 0.60 to 0.73 (1). One factor that may be affecting interrater reliability is the skewed distribution of the scores and the increased difficulty in discerning between the unaffected and less severe degree categories of fluorosis. Given the low prevalence of the more extreme scores and the high prevalence of scores in the nonaffected to very mild range, there is the likelihood of increased variability between examiners. Moreover, the interpretation of the kappa scores requires some caution because each examiner assessed different study participants.

The prevalence of tooth wear is a growing concern (7). Dental erosion affecting permanent teeth during childhood can have lifelong management consequences, including increasingly complex restorative treatment. Likewise, increasing tooth wear in an aging population with increasing tooth retention may have equally important consequences in a population that typically experiences less dental care compared with younger adults. The inclusion of the tooth wear assessment, as well as the functional contacts and quality of life assessments, represents an important addition to the oral health component of the NHANES. These assessments will provide unique opportunities for analysts to explore emerging concerns and hypotheses using a nationally representative data set.

Many of the NHANES 2003-04 oral health assessments are consistent with dental examination procedures used in the NHANES 1999-2000 and 2001-02. These data may be combined to improve the analytical sample size to test hypotheses. Moreover, the oral health data sets can be linked to a wide variety of other current NHANES data components by using the unique survey participant identifier to facilitate interdisciplinary research. The NHANES 2003-04 oral health data is the latest in a series of oral health examination data obtained by a dentist for a period of at least 4 years. During the NHANES 2005-06 and 2007-08, oral health data will be collected by a team of trained MEC health technologists using person-based assessments to obtain general measures of prevalence of specific oral health conditions. The protocol was derived from the Association of State and Territories Dental Directors' Basic Screening Survey (http://www.astdd.org/index.php?template=surveybss.html).

Acknowledgments

The 2003-04 NHANES oral health component was a funding and content collaborative effort between the NIH/NIDCR, the CDC/National Center for Health Promotion and Disease Prevention, Division of Oral Health, and the CDC/NCHS.

References

1. Dye BA, Barker LK, Horowitz AM, Tan S, Thornton-Evans G, Eke P, Li CH, Smith V, Lewis BG, Beltran-Aguilar ED et al. Overview and quality assurance for the National Health and Nutrition Examination Survey (NHANES) oral health component, 1999-2002. Community Dent Oral Epidemiol. 2007;35:140-51.
2. US Department of Health and Human Services. Healthy people 2010. 2nd ed. With understanding and improving health and objectives for improving health. 2 vols. Washington (DC): US Government Printing Office; 2000.
3. Cushing AM, Sheiham A, Maizels J. Developing socio-dental indicators – the social impact of dental disease. Community Dent. Health. 1986;3:3-17.
4. Locker D. Measuring oral health: a conceptual framework. Community Dent Health. 1988;5:3-18.
5. Gift HC, Atchison KA. Oral health, health, and health-related quality of life. Med Care. 1995;33:NS57-77.

6. Slade GD. Derivation and validation of a short-form oral health impact profile. Community Dent Oral Epidemiol. 1997; 25:284-90.
7. Lussi A. Erosive tooth wear – a multifactorial condition of growing concern and increasing knowledge. Monogr Oral Sci. 2006;20:1-8.
8. Bartlett DW. The role of erosion in tooth wear: aetiology, prevention and management. Int Dent J. 2005;55:277-84.
9. Bartlett D. Intrinsic causes of erosion. Monogr Oral Sci. 2006;20:119-39.
10. Addy M, Shellis RP. Interaction between attrition, abrasion and erosion in tooth wear. Monogr Oral Sci. 2006;20:17-31.
11. Dugmore CR, Rock WP. A multifactorial analysis of factors associated with dental erosion. Br Dent J. 2004;196:283-6.
12. Sheiham A. Dietary effects on dental diseases. Public Health Nutr. 2001;4:569-91.
13. Scheutzel P. Etiology of dental erosion – intrinsic factors. Eur J Oral Sci. 1996; 104:178-90.
14. Al-Dlaigan YH, Shaw L, Smith A. Dental erosion in a group of British 14-year-old school children. Part 1: prevalence and influence of differing socioeconomic backgrounds. Br Dent J. 2001;190:145-9.
15. Smith BN, Knight JK. An index for measuring the wear of teeth. Br Dent J. 1984;156:435-8.
16. Millward A, Shaw L, Smith AJ, Rippin JW, Harrington E. The distribution and severity of tooth wear and the relationship between erosion and dietary constituents in a group of children. Int J Paediatr Dent. 1994;4:151-7.
17. Kelly M, Steele J, Nuttall N, Bradnock G, Morris J, Nunn J, Pine C, Pitts N, Treasure E, White D. Adult Dental Health Survey: oral health in the United Kingdom. London: Office for National Statistics; 1998.
18. Steele JG, Sheiham A, Marcenes W, Walls AWG. National Diet and Nutrition Survey: people aged 65 years and over. Volume 2: report of the Oral Health Survey. London: The Stationery Office; 1998. p. 1-124.
19. Arbes S, Agustsdottir H, Slade G. Environmental tobacco smoke and periodontal disease in the U.S. Am J Public Health. 2001;91:253-7.
20. Dye BA, Choudhary K, Shea S, Papapanou PN. Serum antibodies to periodontal pathogens and markers of systemic inflammation. J Clin Periodontol. 2005;32:1189-99.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74.
22. Fleiss JL. Statistical methods for rates and proportions. 3rd ed. Hoboken (NJ): John Wiley and Sons, Inc.; 2003. p. 604.
23. Kingman A, Albandar JM. Methodological aspects of epidemiological studies of periodontal diseases. Periodontol 2000. 2002;29:11-30.
24. Dean HT. The investigation of physiological effects by the epidemiological method. In: Moulton FR, editor. Fluoride and dental health. Washington (DC): American Association for the Advancement of Science; 1942. p. 23-31.