Original Paper

Caries Research

Caries Res 2014;48:91–95
DOI: 10.1159/000353140

Received: January 22, 2013
Accepted after revision: May 20, 2013
Published online: November 26, 2013

Estimation of Caries Experience by Multiple Imputation and Direct Standardization

A.A. Schuller[a] S. van Buuren[a, b]

[a]Netherlands Organization for Applied Scientific Research TNO, Leiden, and [b]University of Utrecht, Utrecht, The Netherlands

Key Words
Bias · Biostatistics · Dental public health · DMF index · Epidemiology · Oral health

Abstract
Valid estimates of caries experience are needed to monitor oral population health. Obtaining such estimates in practice is often complicated by nonresponse and missing data. The goal of this study was to estimate caries experiences in a population of children aged 5 and 11 years, in the presence of nonresponse and missing data. Four estimation methods are compared. Each method makes implicit assumptions about the processes that caused the nonresponse and the missing data. Three of the four methods are based on unrealistic assumptions about the missing data and underestimate caries experience. Under the missing at random assumption, multiple imputation in combination with direct standardization corrects for the deficiencies of current methodology. In the presence of missing data and nonresponse, we recommend a combination of multiple imputation and direct standardization to obtain correct estimates of caries experience.

A common task in dental epidemiology is to estimate the prevalence of a disease or a health status in a population. According to theory one may take a representative sample of individuals from the population of interest, examine their health status, and calculate the percentage of health problems found in the sample as an estimate of the prevalence in the population. In reality, applying this procedure is less straightforward. For example, the prevalence of the health status may be very low, the sample may not be representative for the population, or there could be nonresponse among the members included in the sample. This paper presents a method to deal with the latter two problems.

In a representative sample, each member of the population has an equal chance to be included into the sample. In reality, it is often useful to zoom in on certain subgroups, leading to different inclusion probabilities. For example, one could deliberately oversample ethnic minorities so as to obtain a more precise estimate for these subgroups. One may reweight the subgroup estimates by design-based weights to obtain the population estimate. The relevant statistical methodology is classic [Cochran, 1977] and is known in epidemiology as direct standardization.

KARGER

E-Mail karger@karger.com
www.karger.com/cre

0008–6568/14/0482–0091$39.50/0

Dr. Annemarie A. Schuller
Netherlands Organization for Applied Scientific Research TNO
PO Box 2215
NL–2301 CE Leiden (The Netherlands)
E-Mail Annemarie.Schuller@tno.nl

The problem becomes murkier when differences between the sample and the population are caused by factors that are not under the control of the investigator. Some persons may deny participating in the study, they may skip items in questionnaires, or they may not want to take certain examinations. This problem is known as the missing data problem and leads to incomplete data in the sample. When confronted with incomplete data, the analyst can choose a variety of strategies: ad hoc methods (e.g. analysis of the complete cases only, available case methods, substitution), likelihood-based approaches, weighting or imputation-based models; see Little and Rubin [2002] for an overview of the relative merits of these approaches.

Persons with increased caries experience or oral health risk often have a lower probability to be examined. For example, Armfield et al. [2009] found that extreme dental fear and participation in an epidemiological survey were related (OR = 0.66, 95% CI 0.56–0.77). Armfield argued that in his application the group with extreme fear was relatively small, so the impact of ignoring this group on the total was small. This argument may not apply, however, to other settings. In general, ignoring selectivity of the sample is likely to underestimate caries experience.

Relatively little is known about the effect of oral health on participation, but we may expect that it exists [Armfield et al., 2007; Armfield, 2013]. Oral health has a social gradient: participants with higher socioeconomic status (SES) have less caries experience than participants with a lower SES. This difference in caries experience already exists in the youngest age groups [Schuller et al., 2013]. The response rates for epidemiological oral health studies are often found to be higher in high SES groups than in low SES groups. When combined, these two factors can underestimate caries experience or oral disease.

The purpose of this paper is to estimate caries experiences in a population of children aged 5 and 11 years, in the presence of nonresponse and missing data. The text outlines the consequences of missing data on three commonly used methods, and presents a way to deal with these by means of multiple imputation and direct standardization. The popularity of multiple imputation in dental epidemiology is growing. Most of the work attempts to estimate prevalence [Tellez et al., 2006; Ismail et al., 2008; Liao et al., 2010; Mejia et al., 2011; Newton et al., 2011], but there is also methodological work advocating a wider use of multiple imputation in dental research [Plutzer et al., 2010; Pahel et al., 2011].

Materials and Methods

Sample and Population

The data used in this paper originate from the study 'Oral Health in Children' (aged 5 and 11 years) and adolescents (aged 17, 21 and 23 years) in the Netherlands [Schuller et al., 2013]. The participants were living in four medium-sized cities in the Netherlands. The sample has been shown to be representative regarding demographic variables. Random samples were drawn from the municipal population records of each city. The project was approved under the Personal Data Protection Act (No. m1383077). In this paper, data from children aged 5 and 11 years were used.

Data Collection

Parents of children aged 5 and 11 years (n = 3,090) received a letter about the purpose of the study. Informed consent for participating in the clinical examination was signed and returned by persons with parental authorization. Persons who did not respond were contacted face-to-face by trained interviewers who emphasized the importance of the study. In case of noncontact, the interviewer returned up to a maximum of three contact attempts. Individuals who refused participation were asked to fill out a nonresponse questionnaire, with questions about inter alia gender and SES. The power calculation indicated that 450 children per age group had to be included in the clinical examination. Recruitment of new participants stopped when this number was reached.

The data collection consisted of a questionnaire and a clinical oral examination. The questionnaire was sent to all eligible parents. It measured background variables (ethnicity and educational level) and their children's dental attendance, oral self-care and dental anxiety. The clinical assessment comprised visual inspection of the teeth with a registration of caries lesions and any subsequent treatment (restoration or extraction). The study protocol is written in Dutch and available on request. Clinical examinations were performed by four calibrated dentists in a mobile oral health facility. To evaluate interexaminer agreement, 11% of the participants (all ages) were reexamined by a second examiner. The Pearson correlation between raters for the dmfs+DMFS was 0.91, the intracluster coefficient was 0.91, and the average dmfs+DMFS were 4.0 (SD 5.9) and 4.1 (SD 5.7), respectively. These results indicate a satisfactory interexaminer agreement.

Population references on SES were obtained from the Permanent Study of Living Conditions (POLS) survey of Statistics Netherlands. The POLS survey is a continuous cross-sectional survey on many aspects of the Dutch population held among a representative sample of private households in the Netherlands. The data are available to researchers via the Data Archiving and Networked Services (DANS).

Definitions

The DMF score was used to describe caries experience [Klein et al., 1938]. The DMF score is the sum of decayed (D), missing (M) and restored (F, filled) surfaces (S) or teeth (T). Uppercase letters refer to permanent teeth and lowercase letters to deciduous teeth. SES was operationalized according to the classification of Statistics Netherlands as the level of highest completed education of the mother. Level of education was stratified into low, medium and high based upon the intellectual challenges posed by the edu-

cational system in the Netherlands. Generally, people with a low education had 10 years of education or less, people with a medium level of education had 10–14 years of education and those with a high level of education had at least 15 years of education. Ethnicity of the child was defined as mother being born in the Netherlands versus being born abroad. Each address in the Netherlands has a postal code. The average educational level per postal code was calculated using the data from responders who filled out the questionnaire.

Statistical Analysis

The DMF score was available for all children who participated in the clinical examination, regardless of whether their questionnaire data were available. We considered four methods to estimate the average DMF score in the population.

Method A. Calculate the average DMF score in the subsample with known DMF and known SES. The implicit assumptions are 2-fold: the subsample of children with known DMF is representative for all children, and the distributions of DMF are identical in the subsamples of known and unknown SES.

Method B. Calculate the average DMF score in the subsample with known DMF, irrespective of SES. The implicit assumption is the subsample of children with known DMF is representative for all children.

Method C. Calculate the average DMF score per SES subgroup in participants with known SES, and reweight the SES subgroup averages to match the distribution of SES in the population. This is the conventional direct standardization method to correct for systematic differences in SES between sample and population. The implicit assumption is that within each SES subgroup, the children with known DMF are representative for all children within that subgroup. In addition, the distribution of DMF for those with known and unknown SES is assumed to be identical within each SES subgroup.

Method D. Replace missing SES by multiple imputation and apply method C to all participants with known DMF scores. The implicit assumption is that within each SES subgroup, the children with known DMF are representative for all children in that subgroup.

Ideally, the method should provide estimates that are unbiased, and should use the data from all children that were clinically examined. Although the assumptions required by methods B and C are less stringent than those of A, as we will see, methods B and C can still provide unrealistic estimates. Method D is an attempt to combine the advantages of B or C.

Multiple imputation [Rubin, 1987; van Buuren, 2012] finds plausible replacements for the missing SES data. Since the true SES value for a given person is unknown, we cannot just take the imputed (filled in) data and act as if they were the real data. In multiple imputation, $m > 1$ imputed data sets are constructed. For each missing value, m replacements are drawn from the predictive distribution. The predictive distribution describes the variation in a person's unknown SES, conditional on what we know of that person. For example, the fact that the person lives in an area with many highly educated people increases his or her chance of a higher SES. The spread of the predictive distribution reflects the uncertainty about what to impute. Each of the completed data sets is analyzed by the usual method for complete data, and the results are combined across the m analyses by some simple rules. For moderate missing data problems the number of imputations is classically taken as $m = 5$.

An imputation model was specified by fully conditional specification, also known as chained equations or MICE [van Buuren et al., 2006]. A regression model per incomplete variable describes how its distribution depends on the other variables in the data. Since the variables mutually depend on each other, imputations are drawn in an iterative fashion. This method is now available in various statistical software packages (SAS V9.3, SPSS 17.0, Stata 11 and R 2.06). In this paper, SPSS 20.0 was used to set up the imputation model with variables 'educational level mother' (nominal, 3 categories), 'ethnicity mother' (nominal, 2 categories), 'average educational level per postal code' (continuous), and 'dmf' (5-year-olds, continuous) or 'DMF' (11-year-olds, continuous).

Results

Table 1 shows the response rates per age group and the availability of data. This paper is restricted to individuals with a clinical examination (5-year-olds: n = 486; 11-year-olds: n = 658), or with a filled nonresponse questionnaire (n = 114 and n = 92, respectively). The distribution of SES in participants aged 5 years was 8% (low), 23% (middle), 29% (high) and 39% (unknown), and in nonparticipants aged 5 years the distribution was 16, 33, 40 and 11%. For participants aged 11 years we found 11, 24, 33 and 32%, and for nonparticipants the distribution was 27, 42, 23 and 8%.

Table 2 shows that the average dmf in children aged 5 years with unknown SES was located between the subgroups of low and middle SES. For children aged 11 years we found that the average DMF of those with unknown SES was even higher than in the low SES group. This suggests that methods that ignore the children with an unknown SES, i.e. methods A and C, are likely to underestimate caries experience.

Table 3 presents average dmfs and dmft in children aged 5 years, and DMFS and DMFT in children aged 11 years, as well as the percentages of children without caries experience according to methods A, B, C and D. Methods A and B just take the sample averages and do not reweight the sample to the population, resulting in DMF estimates that are too low. Method C shows, as expected, more realistic results since the subgroup of low SES weighs heavier than the subgroup of high SES. However, method C is still biased as it ignores the DMF of the subgroup with unknown SES, which is high (table 2). Method D reweights the sample using all cases with clinical examinations and imputed SES, and is unbiased under the stated assumptions.

Table 1. Number of eligible individuals and response rates according to age

Total sample	5-year-olds[1]		11-year-olds[2]	
	n	%	n	%
Participants	700	46	831	53
Data available				
Q+C	302	43	453	55
C	184	26	205	25
Q	214	31	173	21
Nonparticipants	827	54	732	47
Data available				
NR	114	14	92	13

[1] n = 1,527. [2] n = 1,563.

Q = Questionnaire; C = clinical examination; NR = nonresponse questionnaire.

Table 2. Average dmfs, dmft, DMFS, DMFT, percentage caries-free according to age and SES, observed and imputed data

SES	Observed data				Imputed data			
	n	dmfs	dmft	caries-free	n	dmfs	dmft	caries-free
5-year-olds								
Low	38	3.26	2.37	42%	72	4.13	2.72	43%
Middle	114	1.59	1.16	64%	196	1.98	1.44	61%
High	143	0.73	0.60	72%	218	0.85	0.70	71%
Unknown	191	2.43	1.69	59%				
Total	486				486			
11-year-olds								
Low	72	0.76	0.60	76%	121	1.18	0.92	66%
Middle	159	0.37	0.32	81%	236	0.55	0.46	76%
High	217	0.35	0.31	81%	302	0.54	0.45	76%
Unknown	210	1.18	0.93	61%				
Total	658				658			

Table 3. Averages of dmfs, dmft, DMFS, DMFT and percentage caries-free, according to age and method

Method	5-year-olds				11-year-olds			
	n	dmfs	dmft	caries-free	n	DMFS	DMFT	caries-free
A	295	1.39	1.04	65	448	0.42	0.36	80%
B	486	1.80	1.30	63	658	0.66	0.54	74%
C	295	1.77	1.31	61	448	0.47	0.39	80%
D	486	2.21	1.55	59	658	0.72	0.58	73%

Table 4. Weight factors, sample versus DANS dataset

SES	5-year-olds		11-year-olds	
	observed[1]	after imputation[2]	observed[1]	after imputation[2]
Low	2.04	1.77	1.63	1.43
Middle	1.13	1.09	1.23	1.22
High	0.62	0.67	0.62	0.65

[1] Method C. [2] Method D.

Table 4 lists the weights used to calculate the population estimates from the SES subgroup averages in methods C and D. Observe that the weights for method D are closer to 1 since method D needs less correction to obtain an unbiased estimate.

Discussion

The estimation of caries experience in a population is often complicated by incompleteness in the sample data. Of course, the best way to deal with missing data is not to have any. Locker [2000] describes various strategies to optimize the data collection: sending a letter to establish the legitimacy of the survey prior to contact, giving the responders some (financial) incentives and working with trained interviewers. We implemented these suggestions in our project. However, despite the best efforts, in practice it is impossible to obtain complete data on every sampled subject.

The nonresponders in our survey had generally a lower SES. Since SES and oral health are related, naïve estimates that use only the complete data will underestimate caries experience. This paper proposed a solution that relies on multiple imputation and direct standardization. Standardization was based on SES, which was unknown for about 30–40% of the children with clinical examinations. Ignoring this group underestimates DMF. Multiple imputation redistributes the children with unknown SES over the groups with known SES, thus enabling the application of direct standardization on all clinical data. Note that direct standardization and imputation address distinct aspects of the problem of missing data, so the steps complement each other.

The introduction outlined that participation, caries and SES are likely to be related. Although method D improves upon the other methods, it is not free of assumptions. We emphasize that method D produces unbiased estimates only if the participation rate does not depend on caries within each SES subgroup, i.e. under the assumption of missing at random [Little and Rubin, 2002]. If this assumption is suspect, we may try including additional covariates that can explain major differences in participation rate.

The analysis in this paper was restricted to the children who had had a clinical examination. We can also incorporate children with just the SES score and no clinical score (table 1, group Q), and impute their clinical examination scores. There is, however, little benefit in doing so as long as the imputation model adequately explains differences in the probabilities of nonresponse [Little and Rubin, 2002]. In addition, we felt that it would be harder to justify such estimates for practitioners.

Multiple imputation represents the state-of-the-art for dealing with missing data, but it is by no means the only approach. One could cast the estimation problem as a problem of likelihood optimization, or reweight the data to correct for the missing data [Little and Rubin, 2002]. To our knowledge, such ideas have yet to be worked out for cases like this, and no software exists that may assist us. In contrast, applying multiple imputation is straightforward. Major statistical packages like SAS, SPSS, Stata and R have implemented multiple imputation under fully conditional specification. In addition, multiple imputation is highly modular and allows for extensive checking for each modeling step.

To conclude, in the presence of missing data and nonresponse and given that the data are missing at random, our recommendation is to combine multiple imputation and direct standardization to obtain unbiased prevalence estimates of caries experience. Our SPSS syntax is freely available on request. We encourage readers to try out the method by adapting and tweaking the SPSS syntax to their own needs.

Acknowledgment

The study 'Oral Health in Children' was financed by the Health Care Insurance Board (CVZ), Diemen, The Netherlands.

Disclosure Statement

This research is free of conflict of interest.

References

Armfield JM: What goes around comes around: revisiting the hypothesized vicious cycle of dental fear and avoidance. Community Dent Oral Epidemiol 2013;41:279–287.

Armfield JM, Slade GD, Spencer AJ: Are people with dental fear under-represented in oral epidemiological surveys? Soc Psychiatry Psychiatr Epidemiol 2009;44:495–500.

Armfield JM, Stewart JF, Spencer AJ: The vicious cycle of dental fear: exploring the interplay between oral health, service utilization and dental fear. BMC Oral Health 2007;7:1.

Cochran WG: Sampling Techniques. New York, Wiley, 1977.

Ismail AI, Sohn W, Tellez M, Willem JM, Betz J, Lepkowski J: Risk indicators for dental caries using the International Caries Detection and Assessment System (ICDAS). Community Dent Oral Epidemiol 2008;36:55–68.

Klein H, Palmer CE, Knutson JW: Studies on dental caries. I. Dental status and dental needs of elementary school children. Public Health Rep 1938;53:751–765.

Liao CC, Ganz ML, Jiang H, Chelmow T: The impact of the public insurance expansions on children's use of preventive dental care. Matern Child Health J 2010;14:58–66.

Little RJ, Rubin DB: Statistical Analysis with Missing Data. New York, Wiley, 2002.

Locker D: Response and nonresponse bias in oral health surveys. J Public Health Dent 2000;60:72–81.

Mejia GC, Weintraub JA, Cheng NF, Grossman W, Han PZ, Phipps KR, Gansky SA: Language and literacy relate to lack of children's dental sealant use. Community Dent Oral Epidemiol 2011;39:318–324.

Newton KM, Chaudhari M, Barlow WE, Inge RE, Theis MK, Spangler LA, Hujoel PP, Reid RJ: A population-based study of periodontal care among those with and without diabetes. J Periodontol 2011;82:1650–1656.

Pahel BT, Preisser JS, Stearns SC, Rozier RG: Multiple imputation of dental caries data using a zero-inflated Poisson regression model. J Public Health Dent 2011;71:71–78.

Plutzer K, Mejia GC, Spencer AJ, Keirse MJ: Dealing with missing outcomes: lessons from a randomized trial of a prenatal intervention to prevent early childhood caries. Open Dent J 2010;4:55–60.

Rubin DB: Multiple Imputation for Nonresponse in Surveys. New York, Wiley, 1987.

Schuller AA, Kempen van CPF, Poorterman JHG, Verrips GHW: Kies Voor Tanden: Een Onderzoek Naar Mondgezondheid En Preventief Tandheelkundig Gedrag Van Jeugdigen. Leiden, Netherlands Organization for Applied Scientific Research TNO, 2013.

Tellez M, Sohn W, Burt BA, Ismail AI: Assessment of the relationship between neighborhood characteristics and dental caries severity among low-income African-Americans: a multilevel approach. J Public Health Dent 2006;66:30–36.

van Buuren S: Flexible Imputation of Missing Data. Boca Raton, Chapman & Hall/CRC Press, 2012.

van Buuren S, Brand JPL, Groothuis-Oudshoorn CGM, Rubin DB: Fully conditional specification in multivariate imputation. J Stat Comput Simul 2006;76:1049–1064.