Neuropsychological and Renal Effects of Dental Amalgam in Children
A Randomized Clinical Trial

David C. Bellinger, PhD, MSc

Felicia Trachtenberg, PhD

Lars Barregard, MD, PhD

Mary Tavares, DMD, MPH

Elsa Cernichiari, MS

David Daniel, PhD

Sonja McKinlay, PhD

ALTHOUGH IT IS ESTIMATED that more than 70 million dental amalgam restorations are placed annually in the United States,[1] the health risks posed by the potential chronic release of metallic mercury vapor from amalgams (40%-50% mercury by weight) remain unclear. Occupational exposures resulting in urinary mercury levels greater than 50 µg/L have been associated with various neurological, renal, and immunological impairments.[2] Potential effects of lower occupational levels of mercury have also been evaluated, but results are inconsistent. Studies of dentists have found urinary mercury levels as low as 4 to 10 µg/L to be inversely associated with scores on tests of neurobehavioral function, including memory, attention, motor coordination and steadiness, and mood,[3-5] but others failed to confirm a statistically significant association between urinary mercury and neurobehavioral function among dentists.[6]

For the most part, studies in the general adult population, which presume that exposure to metallic mercury is pri-

Context No randomized trials have been published that address the concern that inhalation of mercury vapor released by amalgam dental restorations causes adverse health effects.

Objective To compare the neuropsychological and renal function of children whose dental caries were restored using amalgam or mercury-free materials.

Design and Setting The New England Children's Amalgam Trial was a 2-group randomized safety trial involving 5 community health dental clinics in Boston, Mass, and 1 in Farmington, Me, between September 1997 and March 2005.

Participants and Intervention A total of 534 children aged 6 to 10 years at baseline with no prior amalgam restorations and 2 or more posterior teeth with caries were randomly assigned to receive dental restoration of baseline and incident caries during a 5-year follow-up period using either amalgam (n=267) or resin composite (n =267) materials.

Main Outcome Measures The primary neuropsychological outcome was 5-year change in full-scale IQ scores. Secondary outcomes included tests of memory and visuomotor ability. Renal glomerular function was measured by creatinine-adjusted albumin in urine.

Results Children had a mean of 15 tooth surfaces (median, 14) restored during the 5-year period (range, 0-55). Assignment to the amalgam group was associated with a significantly higher mean urinary mercury level (0.9 vs 0.6 µg/g of creatinine at year 5, $P<.001$). After adjusting for randomization stratum and other covariates, no statistically significant differences were found between children in the amalgam and composite groups in 5-year change in full-scale IQ score (3.1 vs 2.1, $P=.21$). The difference in treatment group change scores was 1.0 (95% confidence interval, −0.6 to 2.5) full-scale IQ score point. No statistically significant differences were found for 4-year change in the general memory index (8.1 vs 7.2, $P=.34$), 4-year change in visuomotor composite (3.8 vs 3.7, $P=.93$), or year 5 urinary albumin (median, 7.5 vs 7.4 mg/g of creatinine, $P=.61$).

Conclusions In this study, there were no statistically significant differences in adverse neuropsychological or renal effects observed over the 5-year period in children whose caries were restored using dental amalgam or composite materials. Although it is possible that very small IQ effects cannot be ruled out, these findings suggest that the health effects of amalgam restorations in children need not be the basis of treatment decisions when choosing restorative dental materials.

Trial Registration clinicaltrials.gov Identifier: NCT00065988

Author Affiliations: Department of Neurology, Children's Hospital Boston, and Harvard Medical School and Department of Environmental Health, Harvard School of Public Health, Boston, Mass (Dr Bellinger); New England Research Institutes, Watertown, Mass (Drs Trachtenberg and McKinlay); Department of Occupational and Environmental Medicine, Sahlgrenska University Hospital and Academy, Goteborg University, Goteborg, Sweden (Dr Barregard); The Forsyth Institute, Boston, Mass (Dr Tavares); Department of Environmental Medicine, University of Rochester School of Medicine, Rochester, NY (Ms Cernichiari); and the Department of Psychology, University of Maine at Farmington (Dr Daniel).
Corresponding Author: Sonja M. McKinlay, PhD, New England Research Institute Inc, 9 Galen St, Watertown, MA 02472 (smckinlay@neriscience.com).

marily a result of dental amalgams, have not found significant associations between neuropsychological function and various amalgam exposure indexes, including urinary mercury level (when measured, generally <5 µg/L), number of amalgam restorations, total number of amalgam surfaces, and number of occlusal amalgam surfaces.[7-12] Some studies suggest that dental amalgams are associated with neurodegenerative disorders such as Alzheimer disease[13] and multiple sclerosis.[14] In other studies, interventions such as the administra-

tion of chelating agents or the removal of dental amalgam have failed to demonstrate health benefits.[15,16] Caution is warranted in drawing inferences from the available data, however, insofar as none of the studies evaluating the health effects of dental amalgam were randomized clinical trials.

A larger concern is that few data are available on the possible effects of amalgam on children, who might be more vulnerable to mercury toxicities because of their developmental immatu-

rity during the period in which the risk of caries is greatest, and thus the placement of amalgam is most frequent. Amalgam fillings in a child's mouth are associated with greater exposure to mercury, as determined by significantly higher urinary mercury levels.[17-20] Whether the exposure levels that result from the placement of amalgam are sufficiently high to adversely affect children's health remains uncertain.[21,22]

We report herein the results of The New England Children's Amalgam Trial (NECAT), a randomized clinical trial comparing the health of children whose caries were restored using either dental amalgam or mercury-free composite materials.

METHODS

Study Design and Participants

A detailed discussion of the design of the NECAT has been previously published.[23] The study was approved by the institutional review boards of all participating sites, which included the independent research organization New England Research Institutes, the nonprofit independent Forsyth Institute, and hospital-affiliated dental clinics (affiliated with Franklin Memorial Hospital in Maine, and the Cambridge Health Alliance, Boston University Medical Center, or Children's Hospital Boston in Massachusetts).

Children were eligible if they were 6 to 10 years of age at last birthday; fluent in English; had no known prior or existing amalgam restorations; had 2 or more posterior teeth with dental caries such that restoration would include the occlusal surfaces; and, by parent report, had no physician-diagnosed psychological, behavioral, neurological, immunosuppressive, or renal disease. Race or ethnicity was self-reported by the parents of the children from a list including non-Hispanic white, non-Hispanic black, Hispanic, Asian or Pacific Islander, Native American (including Alaskan), biracial or multiracial (specify), or other (specify).

A total of 5116 children were screened for eligibility (**FIGURE 1**). Up to 3 baseline visits were required to con-

Figure 1. Profile of Recruitment, Randomization, and Follow-up in the New England Children's Amalgam Trial
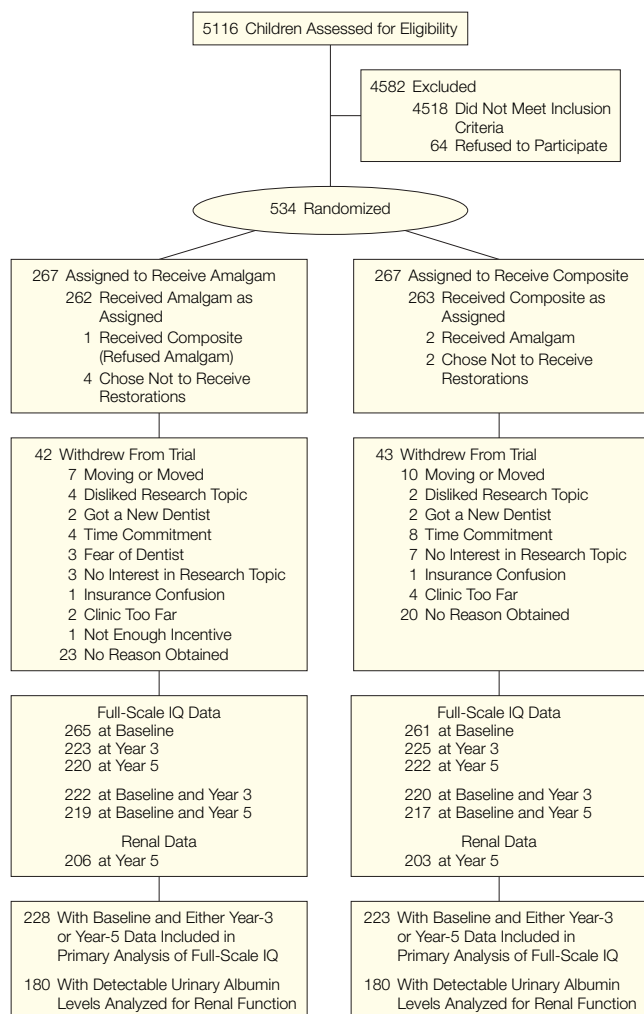
The recruitment period ran from September 1997 through September 1999 with follow-up ending March 2005. Some participants expressed more than one reason for withdrawal all of which are recorded in the Figure. Any missing IQ data were not obtained due to a missed visit. Missing renal data are due to either a missed visit or insufficient volume of urine sample.

firm study eligibility and to collect baseline data. The baseline visits included a dental examination by a NECAT dentist, x-rays, standard preventive dental care (eg, cleaning, application of sealants), phlebotomy, urine sample, anthropometric measurements (height, weight, body fat), health interviews, and neuropsychological testing of the child and his or her guardian. Eligibility was confirmed for 598 children, and written parental consent and written child assent was obtained for 534.

After completion of baseline data collection visits, these 534 children were randomly assigned to 1 of 2 study treatment groups. Randomization was stratified by geographic location (Boston vs Maine) and number of teeth with caries (2-4 vs $\geq$5) using randomly permuted blocks within each of the 4 strata (Boston, 2-4 caries, n=103; Boston, $\geq$5 caries, n=188; Maine, 2-4 caries, n=141; Maine, $\geq$5 caries, n=102). Assignment was made via telephone using software and encrypted files by New England Research Institutes personnel who were not involved in data collection.

Interventions and Follow-up

For children assigned to the amalgam group, a dispersed phase amalgam was used to restore all posterior teeth with caries at baseline and to restore incident caries during the 5-year trial period. For children assigned to the composite group, resin composite material (white filling) was used for all restorations. Following standard clinical practice, however, for both groups, composite material was used to restore caries in the front teeth. In both groups, stainless steel crowns were used to restore primary teeth with extensive lesions that could not be restored using either assigned restorative material. The choices of dental materials and techniques were standardized across sites and dentists. Participants and dentists could not be blinded to treatment assignment. All individuals who collected outcome data or analyzed specimens were blinded to children's treatment assignments.

Children in both groups had semi-annual dental examinations, as well as additional visits required to meet any treatment needs identified at these examinations. At every examination and treatment visit, pertinent dental data, including the status of each tooth surface and reasons for the placement of restorations, were documented. At the annual visits, anthropometric measurements were made and a urine sample collected. Initially, attempts were made to collect timed overnight urine samples but, mid trial, a switch was made to spot samples. Hair samples were collected biennially. In addition, children assigned to the amalgam group participated in 2 additional visits for safety monitoring, at 2 months and 6 months after restoration. These visits included anthropometric measurements, urine collection, and, at 6 months only, a blood draw.

Neuropsychological Outcome Measures

Because the potential neuropsychological effects of long-term exposure to low-doses of elemental mercury in children are not known, full-scale IQ on the Wechsler Intelligence Scale for Children-Third Edition (WISC-III), an apical score that integrates a child's performance over a diversity of cognitive domains, was selected as the primary outcome measure. The WISC-III was administered 3 times: at baseline prior to caries restoration, and at 3 and 5 years after baseline.

The primary end point is the difference between the baseline and full-scale IQ scores 5 years after baseline. The Wechsler Individual Achievement Test and the Behavior Assessment System for Children were also administered at the same visits as the IQ test. To provide insight into the mechanism(s) of any treatment group differences in full-scale IQ, a battery of additional neuropsychological tests was administered at baseline, and at years 1, 2, and 4. This battery consisted of tests that prior studies with adults suggested might be sensitive to inorganic mercury exposure: auditory memory, visual-motor integration, at-tention, and emotional state. The tests in the battery were the Wide Range Assessment of Visual Motor Ability, the Wide Range Assessment of Memory and Learning, the Stroop Color-Word Interference Test, the Wisconsin Card Sorting Test, the Trail-Making Test, a verbal cancellation task, tests of verbal fluency, finger tapping, and reaction time. The 2 scores selected as secondary end points were the changes between baseline and year-4 evaluations in the visual motor composite of the visuomotor ability assessment and the general memory index of the memory and learning assessment.

Quality control of the neuropsychological assessments was ensured by having all examiners trained and certified (by D.C.B.) before conducting assessments of trial participants. During the course of the trial, a total of 14 testers were used at the Boston site and 5 at the Maine site. Each tester was observed in-person annually. In addition, each completed testing protocol was rescored by a second tester and errors were corrected.

Analytical Methods

Total mercury was measured in urine and hair. The method is based on the rapid conversion of mercury compounds into atomic mercury suitable for aspiration through the cell of a flameless atomic absorption monitor (Laboratory Data Control Model 1235, Interstate Industrial Park, Riviera Beach, Fla).[24,25] Biological samples are digested in 45% (weight to volume) sodium hydroxide solution in the presence of 1% cysteine. In the presence of stannus chloride at high pH, cadmium chloride breaks the carbon bond, with a subsequent reduction of mercuric mercury (Hg^{2+}) to elemental mercury (Hg^0). The detection limit, initially 1.5 ng/mL, was reduced to 0.45 ng/mL after February 1, 2000, as a result of increasing the volume of urine analyzed from each child. Nondetectable concentrations ($<$0.45 ng/mL) were imputed as $0.45/\sqrt{2}$.[26]

Blood lead levels were measured by the Strong Hospital Clinical Laboratory by

an electrothermal process using an atomic absorption spectrometer with Zeeman background correction. Blood samples, blood-based quality control materials, and aqueous standards were diluted 1:9 with a matrix modifier solution containing nitric acid, Triton X-100 (Dow Chemical Co, Midland, Mich) and ammonium dihydrogen phosphate.

Urinary albumin was determined at the Sahlgrenska University Hospital, Goteborg, Sweden, by an automated nephelometric immunochemical method using reagents and calibrator from Beckman Coulter (Fullerton, Calif). The detection limit was 2.4 mg/L. The excretion of albumin was expressed in milligrams per gram of creatinine.

Sample Size Determination

Prior studies did not provide information about the likely magnitude of an effect of dental amalgam on children's IQ scores. We based sample size calculation on the literature showing that, in children, a 10- to 15-µg/dL (0.483 µmol/L) increase in blood lead level is associated with a 3-point decline in IQ.[27,28] This is widely considered to be an effect of public health importance. The NECAT trial was therefore designed to achieve 80% power to detect a 3-point difference between treatment groups of 186 each in 5-year change in full-scale IQ score, adjusted for baseline IQ score and randomization stratum.[23] Assuming a retention rate of 75% over the 5-year follow-up period with a 2-sided α level of .045 (to account for spending .005 of an overall α of .05 on interim looks), the recruitment goal was 250 children per treatment group, for a total sample size of 500 children.

Statistical Analysis

Wilcoxon rank-sum tests were used to compare treatment groups with respect to exposure to dental materials (number of restored surfaces) and urinary mercury levels. Incidence of adverse health events was compared between treatment groups using the Fisher exact test.

In intention-to-treat analyses, analysis of covariance was used to model 5-year change in IQ and 4-year change in general memory index and visuomotor composite scores as a function of assigned treatment group, adjusting for baseline score and randomization stratum. In secondary analyses, adjustments were made for baseline covariates, including age, sex, socioeconomic status, hair mercury concentration, and blood lead level. Socioeconomic status was calculated using the method developed by Green.[29] Hair mercury was included to control for dietary sources of mercury. In addition, a repeated-measures model with both 3-year and 5-year change in IQ was fit with and without the interaction between treatment group and year. Further sensitivity analyses included an adjustment for the time between baseline and follow-up, an adjustment for potential interexaminer differences, an as-treated analysis, and a dose-response model (using amalgam exposure measured in surface-years of amalgam fillings).

In the primary analysis of 5-year change in IQ, missing data were handled by the method of last observation carried forward. In sensitivity analyses, we used multiple imputation[30] of missing outcome data assuming data were missing at random and multiple imputation assuming that children with missing data in the amalgam group scored 3 points below[27,28] what would otherwise be expected. The algorithm for multiple imputation used 5 imputations and included variables found to be associated with dropout, based on a multivariate logistic regression model.

In analyses of albumin, analysis of variance was used to model year-5 creatinine-corrected albumin as a function of assigned treatment group, adjusting for randomization stratum. In secondary analyses, adjustments were made for baseline covariates but also included urine collection type (overnight or spot), urinary creatinine concentration, lean body mass, and sample storage time. We controlled for collection type (ie, time) and creatinine concentration to take into account urinary flow rate. Storage time is included because of data suggesting that albumin measurements can be affected by the duration of storage.[31] A log-transformation was used because albumin was log-normally distributed. In addition, a repeated-measures model with both 3-year and 5-year albumin was fit with and without the interaction between treatment group and year.

The data and safety monitoring board reviewed interim analyses comparing the mean scores for full-scale IQ, the general memory index, and the visuomotor composite of children in the 2 treatment groups at 3 years after baseline. The data and safety monitoring board also monitored individual trajectories of test scores, extreme outcome values, and adverse health effects. Parents were notified of any adverse health effects or of outcome values outside established normal ranges. All statistical tests were 2-sided, performed at an α level of .05, and conducted using SAS version 9.1 software (SAS Institute Inc, Cary, NC).

RESULTS

Baseline Characteristics

Children in the 2 treatment groups were similar in terms of most baseline characteristics, including age, race, household income, education of primary caregiver, full-scale IQ, hair and urinary mercury concentrations, blood lead level, and number of decayed tooth surfaces (TABLE 1). The numbers of girls and boys were comparable in the amalgam group, but girls outnumbered boys in the composite group. Participants were primarily non-Hispanic white (62%), with non-Hispanic blacks comprising 19% of the sample.

The mean number of total caries recorded at baseline was 9.5 decayed tooth surfaces, with 1.7 of the surfaces being in permanent teeth. Slightly more than half of the children (54%) had 5 or more teeth with caries that required restoration with the rest having 2 to 4 carious teeth. Children from the Boston site tended to have more caries than children from Maine (10.3 vs 8.6 carious surfaces, respectively). At baseline, 93% of children had urinary mercury levels below the limit of detection.

Completion of Follow-up Assessments

The percentages in each treatment group of children who completed the yearly visits were comparable. Annual neuropsychological outcome data were available for at least 75% of enrolled children except in year 2, in which a data collection hiatus occurred due to funding uncertainty. Renal outcome data were available for the majority of the children, with the primary reason for its unavailability being a urine sample that was insufficient in volume.

Table 1 also shows the baseline characteristics of children who later withdrew from the trial. Children who withdrew tended to have lower baseline IQ; be from Boston; be of minority race, es-pecially Hispanic; and have lower parental income and educational achievement. However, characteristics of children who completed the study remained comparable by treatment group.

Exposure to Dental Materials

Few children were not treated according to random assignment: 1 child in the amalgam group whose parent refused to allow amalgam fillings, 2 children in the composite group who received amalgam from out-of-study dentists, and 6 children (4 in the amalgam and 2 in the composite group) who chose not to receive needed restorations but continued with follow-up measurements.

Neither the mean number of restored surfaces in place at the end of the study nor the mean cumulative number of surfaces restored over the course of the study differed significantly between treatment groups ($P = .16$ and $P = .10$, respectively; TABLE 2). The numbers of restored surfaces were greatest shortly after entry into the study due to unmet dental needs. However, most baseline fillings were placed in primary teeth, which were then lost over the course of the trial. The children did have recurrent treatment needs, averaging approximately 1 additional filled surface per year.

Mercury Exposure

Children assigned to the amalgam group had a significantly higher mean (SD) urinary mercury level 5 years af-

Table 1. Baseline Characteristics of All Participants and Those Who Withdrew During Follow-up, by Treatment Group

	All Participants (N = 534)*		Withdrawals (n = 85)	
	Amalgam Group (n = 267)	Composite Group (n = 267)	Amalgam Group (n = 42)	Composite Group (n = 43)
Study site, No. (%)				
Boston	144 (53.9)	147 (55.1)	28 (66.7)	28 (65.1)
Maine	123 (46.1)	120 (44.9)	14 (33.3)	15 (34.9)
Carious surfaces, mean (SD) [range]	9.8 (6.9) [2-39]	9.3 (6.2) [2-36]	11.2 (7.2) [2-35]	9.5 (5.6) [2-22]
Age, mean (SD), y	7.9 (1.3)	7.9 (1.4)	8.1 (1.5)	8.0 (1.3)
Sex, No. (%)				
Female	131 (49.1)	156 (58.4)	20 (47.6)	30 (69.8)
Male	136 (50.9)	111 (41.6)	22 (52.4)	13 (30.2)
Race or ethnicity No. (%)†				
Non-Hispanic white	165 (64.0)	158 (60.3)	15 (42.9)	22 (53.7)
Non-Hispanic black	49 (19.0)	49 (18.7)	7 (20.0)	9 (22.0)
Hispanic	15 (5.8)	23 (8.8)	5 (14.3)	5 (12.2)
Other	29 (11.2)	32 (12.2)	8 (22.9)	5 (12.2)
Household income, $, No. (%)				
≤20 000	74 (29.2)	86 (33.1)	12 (35.3)	19 (47.5)
20 001-40 000	113 (44.7)	109 (41.9)	14 (41.2)	13 (32.2)
>40 000	66 (26.1)	65 (25.0)	8 (23.5)	8 (20.0)
Education of primary caretaker, No. (%)				
<High school	34 (13.2)	38 (14.6)	7 (20.0)	9 (22.5)
High school graduate	197 (76.4)	194 (74.3)	22 (62.9)	29 (72.5)
College graduate	18 (7.9)	17 (6.5)	3 (8.6)	2 (5.0)
Postcollege degree	9 (3.5)	12 (4.6)	3 (8.6)	0
WISC-III Full-Scale IQ score, mean (SD) [range]	95.1 (14.5) [65-141]	96.1 (12.1) [62-123]	91.3 (15.2) [69-122]	93.6 (12.5) [62-123]
Detectable urinary mercury concentration, No. (%)‡	21 (8.4)	11 (4.5)	3 (7.5)	3 (8.1)
Hair mercury concentration, mean (SD) [range], µg/g of hair	0.4 (0.5) [0.1-4.4]	0.4 (0.5) [0.1-4.5]	0.7 (0.8) [0.1-4.4]	0.4 (0.3) [0.1-1.2]
Blood lead concentration, mean (SD) [range], µg/dL	2.4 (1.9) [1-13]	2.3 (1.5) [1-11]	2.5 (2.0) [1-10]	2.7 (2.0) [1-10]

Abbreviation: WISC-III, Wechsler Intelligence Scale for Children, Third Edition.
SI conversion: To convert lead from mg/dL to µmol/L, multiply by 0.0483.
*The number for all trial participants includes those who later withdrew (85 of 534). For race and lead, data were available for 520 participants; for income, 513; for education and hair mercury, 519; for WISC-III, 526; for urinary mercury, 498.
†Race was self-reported by the parents of the children. The other category included individuals who identified themselves as Asian, Pacific Islander, Native American, biracial, or other, which they were asked to specify.
‡Defined as urinary mercury concentration 1.5 ng/mL or higher.

Table 2. Dental Treatment and Amalgam Exposure at End of the 5-Year Trial, by Treatment Group

	Mean (SD)		Median (Range)	
	Amalgam Group	Composite Group	Amalgam Group	Composite Group
No. of restored surfaces in mouth at year 5*	5.3 (5.2)	6.1 (6.0)	4 (0-36)	5 (0-36)
No. of restored amalgam surfaces in mouth at year 5†	4.0 (4.0)	0.05 (0.6)	3 (0-21)	0 (0-9)
Cumulative No. of surfaces restored over 5 years‡§	14.6 (9.6)	15.8 (9.9)	13 (0-55)	15 (0-51)
Cumulative No. of surfaces restored with amalgam over 5 years†§	11.5 (7.1)	0.04 (0.6)	10 (0-35)	0 (0-9)

*P=.16 for difference between amalgam and composite groups.
†Two children in the composite group received amalgam fillings from an out-of-study dentist.
‡P=.10 for difference between amalgam and composite groups.
§Cumulative numbers do not include children who withdrew from the study. Six children chose not to receive any restorations but completed follow-up measurements nevertheless.

Figure 2. Urinary Mercury Excretion by Year and Treatment Group
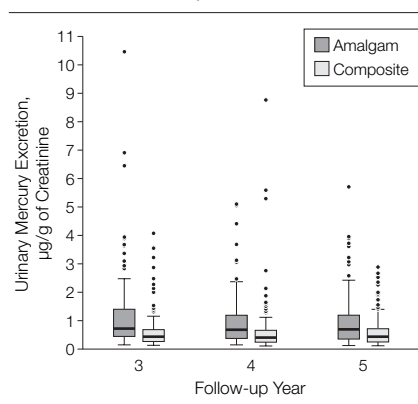
Boxes indicate upper and lower quartiles, and error bars indicate 2.5% and 97.5% values with points for outliers. P<.001 for the difference between amalgam and composite groups at year 5.

ter baseline than did children assigned to the composite group (0.9 [0.8]; range, 0.1-5.7 µg/g of creatinine vs 0.6 [0.5]; range, 0.1-2.9 µg/g of creatinine, P<.001; FIGURE 2), although the overlap in the distributions was considerable. Urinary mercury was detected in 63% of the amalgam group and in 45% of the composite group. Hair mercury was similar in the treatment groups (amalgam group, 0.4 [0.4]; range, 0.1-2.3 µg/g vs composite group, 0.5 [0.7]; range, 0.04-6.5 µg/g).

Neuropsychological Function

Full-scale IQ, general memory index, and visuomotor composite scores in-creased between the baseline and 5-year assessments in both treatment groups. None of the differences between the change scores in the 2 treatment groups, adjusting for baseline score and randomization stratum, were statistically significant, although for all 3 tests, the differences favored the amalgam group (TABLE 3). Adjusting for additional covariates did not change the results appreciably.

In the repeated-measures model using both 3-year and 5-year change in IQ, treatment group was not significant. However, year was significant (an increase of 0.43 per year, P=.005). In the model that included the interaction between treatment group and year, the interaction was not significant (P=.54), indicating that the secular trend in IQ score was independent of treatment group.

Additional sensitivity analyses confirmed the results of the primary analysis. For example, neither treatment group nor time was significant in the analysis controlling for time between baseline and 5-year follow-up with or without a treatment group × time interaction term. Likewise, the analysis adjusted for potential interexaminer differences revealed no statistically significant associations with treatment group, as well as no statistically significant interexaminer differences. The as-treated analysis again showed no significant effect of treatment group (P=.24). Further-more, a dose-response model, using surface-years of amalgam instead of treatment group as the exposure index, found no significant effect of amalgam exposure (P=.40); in fact, the mean (SE) 5-year improvement in IQ score increased very slightly with greater exposure to amalgam (ie, children with more amalgam fillings scored slightly higher, 0.016 [0.019] points increase per surface-year of amalgam exposure).

Our results were also confirmed in analyses using multiple imputation for missing outcome data. Missing outcome data were imputed from prior scores, marital status of the primary caregiver, and immigration status of the primary caregiver. The latter 2 characteristics were independently associated with dropout status. Assuming data were missing at random yielded a mean (SE) IQ improvement of 3.2 (0.6) points in the amalgam group and 2.0 (0.5) points in the composite group (P=.09). The assumption of a worst-case scenario, in which children with missing data in the amalgam group scored 3 points lower than would be otherwise expected, yielded an average IQ improvement of 2.7 (0.6) points in the amalgam group and 2.0 (0.5) points in the composite group (P=.34). Even under this worst-case scenario, the treatment-group difference continued to favor the amalgam group.

Renal Function

At year 3, albumin was detected in 87% of the samples provided by children in the amalgam group and 88% in the composite group. At year 5, these percentages were 87% for the amalgam group and 90% for the composite group. Albumin levels at year 5 did not differ significantly between treatment groups. Among the 180 participants in the amalgam group, the unadjusted mean (SE) albumin level at year 5 was 32.8 (6.9) mg/g of creatinine (median, 7.5) and among the 183 in the composite group, it was 23.7 (5.0) mg/g of creatinine (median, 7.4) with no significant difference between treatment groups in the log-transformed analy-

Table 3. Neuropsychological Outcomes: Baseline and Follow-up Scores, Change Scores, and Group Differences, by Randomization Assignment*

Neuropsychological Outcome	Amalgam Group				Composite Group				Treatment Group Difference in Change Score†	
		Mean				Mean				
	No.	Baseline	Year 4/5‡	Change (SE)	No.	Baseline	Year 4/5‡	Change (SE)	Difference (95% CI)	P Value
WISC-III full-scale IQ	228	95.1	98.9	3.1 (0.6)	223	96.1	98.3	2.1 (0.6)	1.0 (–0.6 to 2.5)	.21
General memory index	212	91.6	100.0	8.1 (0.7)	203	92.3	99.0	7.2 (0.7)	0.9 (–0.9 to 2.7)	.34
Visuomotor composite	211	100.1	104.7	3.8 (0.8)	203	100.4	104.5	3.7 (0.8)	0.1 (–2.0 to 2.2)	.93

Abbreviations: CI, confidence interval; WISC-III, Wechsler Intelligence Scale for Children, Third Edition.
*From analysis of covariance, adjusted for randomization stratum and baseline neuropsychological test scores.
†A positive difference (>0) indicates that the amalgam group scores improved more than the composite group scores.
‡WISC-III was administered at year 5. General Memory Index and Visual Motor Composite were administered at year 4.

sis of covariance (amalgam group, 0.1 mg/g higher than the composite group; 95% confidence interval, –0.2 to 0.3; P=.61). Adjustment for the additional covariates did not affect the results appreciably. However, mean (median) albumin was higher for girls than for boys (36.9 [7.3] vs 18.3 [3.4]; P=.02).

In the repeated-measures model using both year-3 and year-5 albumin levels, treatment group was not statistically significant nor was year or the interaction between treatment group and year. Albumin levels did not change over time in either treatment group (data not shown).

Adverse Events

No child had a urinary mercury level greater than 20 µg/g of creatinine at any time in the trial, and no child's neuropsychological test scores consistently decreased over time. There were 77 children with microalbuminuria (albumin >30 mg/g of creatinine) during the trial with no significant difference between treatment groups. Adverse health events were recorded similarly in both treatment groups (TABLE 4).

COMMENT

This randomized trial was powered to address the hypothesis that children exposed to low levels of elemental mercury from dental amalgam would, on average, have a 5-year change in full-scale IQ score that is 3 points lower than those exposed to composite restoration material. There was no support for this hypothesis. Despite the increase in elemental mercury exposure in the amal-

Table 4. Adverse Health Conditions Reported During 5-Year Follow-up*

Condition	No. (%)†		P Value
	Amalgam Group	Composite Group	
Allergy	45 (16.9)	47 (17.6)	.91
Anemia	4 (1.5)	3 (1.1)	>.99
Asthma	19 (7.1)	17 (6.4)	.86
Cancer	1 (0.4)	0 (0.0)	>.99
Cardiovascular disorders	12 (4.)	16 (6.0)	.56
Central nervous system disorders	2 (0.8)	3 (1.1)	>.99
Diabetes	0	0	>.99
Gastrointestinal disorders	6 (2.3)	9 (3.4)	.60
Kidney disorders other than diabetes	2 (0.8)	2 (0.8)	>.99
Migraine	16 (6.0)	14 (5.2)	.85
Neurological illness	4 (1.5)	1 (0.4)	.37
Psychological disorders	24 (9.0)	18 (6.7)	.42
Respiratory disorders	13 (4.9)	7 (2.6)	.25
Sensory disorders	36 (13.5)	28 (10.5)	.35
Sickle cell disease	1 (0.4)	1 (0.4)	>.99
Skin disorders	23 (8.6)	23 (8.6)	>.99
Weakness, fatigue, edema, or joint pains	6 (2.3)	6 (2.3)	>.99
No adverse health conditions reported	132 (49.4)	122 (45.7)	.44

*Adverse health conditions were self-reported by primary caregiver of participants at 6 annual visits.
†Percentages are calculated from all randomized participants in each group (n = 267) and sum to more than 100% in each group because participants may be counted for more than 1 condition.

gam treatment group compared with the composite treatment group, the average full-scale IQ score 5-year differences adjusted for baseline values were statistically equivalent. The increased mercury exposure in the amalgam group was still well within established background population levels[32] and comparable with average levels reported for US adults.[33,34] Moreover, for 3 of the 4 end points, the small differences observed favored the amalgam group.

Eligibility criteria for the trial required at least 2 posterior teeth with caries and no prior amalgam restorations,

resulting in high use of mercury amalgam in the children assigned to that group relative to children in the general US population. It is notable that, despite this relatively high exposure, urinary mercury levels were low. In light of these considerations, our findings indicate that for US children exposure to elemental mercury secondary to the restoration of dental caries with mercury amalgam is unlikely to cause a reduction in IQ of at least 3 points. This conclusion is strengthened by the consistent lack of differences found on the other neuropsychological end points that

measured visuomotor and general memory functions. Similarly, a statistically significant increase was not found in albumin excretion, a marker of renal glomerular integrity.

This trial has several strengths. First, the recruitment of children in whom no dental amalgam restorations had ever been placed and children's random assignment to treatment group ensured not only equivalence of treatment groups at baseline but also ensured no impact on results by possible variable prior amalgam exposure. Second, the sampling frames used resulted in the recruitment of children with many dental caries, providing a setting of high restoration rates in which the study hypotheses could be adequately tested. Third, the primary and secondary neuropsychological end points were well-standardized tests of domains of unquestioned importance for a child's well-being: intelligence, memory, learning, and visuomotor skills. Fourth, the primary neuropsychological end point, full-scale IQ, was measured 3 times, at baseline and years 3 and 5 after initial treatment. Secondary neuropsychological end points were measured on 4 occasions over the follow-up interval. This density of assessments provides a greater weight of evidence and a lesser role for chance variability to influence the inferences drawn than would a sparser schedule. Fifth, attrition was relatively low because 5-year neuropsychological data were obtained for 83% of the children enrolled and 5-year renal outcome data for 77% of the children enrolled.

Our study was designed to answer a specific question on the safety of amalgam restorations as the standard of care for US children. Because the children's first exposure to mercury from amalgam occurred between the ages of 6 and 10, our findings might not apply to children who receive amalgam restorations before age 6 years, when sensitivity to mercury toxicity might be greater. A follow-up period longer than 5 years might be needed to appreciate subtle toxic effects associated with exposure to dental amalgam. Also, few data were available to guide our selection of health end points for the trial. It is possible that we would have detected toxic effects had we measured different end points. This trial was not designed to detect rare adverse effects but an average response. Although the study was powered to detect at least a 3-point reduction in IQ scores, the sample size was insufficient to detect smaller between-group differences in the IQ change scores. The 95% confidence interval surrounding the treatment group difference suggests that the difference in IQ change scores may be as much as 0.6 points lower or 2.5 points higher for children who received mercury amalgam. Thus, the possibility of very small adverse effects of amalgam on IQ score cannot be completely ruled out.

Moreover, a small fraction of children and adults have a considerably higher mercury uptake from dental amalgam than average[35] and although it is possible that certain especially sensitive children could be affected by low-dose mercury exposure from amalgam,[5] the factors that might produce enhanced sensitivities are unknown. Finally, the choice of composite for comparison was based on widespread use and availability. The safety of the composite used is itself not established nor could it be assessed in this trial.

Clinically, implications from the results of this trial are clear. Under the conditions of use represented in this trial, there is no reason to discontinue use of mercury amalgam as the standard of care for caries in posterior teeth. This is a particularly important consideration for areas both in the United States and in other countries, where the replacement of mercury amalgam with a composite restoration material may not be feasible with respect to factors such as cost, storage, and expertise in handling, and thus could adversely affect the dental as well as general health of the population being served.

REFERENCES

1. American Dental Association. *1999 Survey of Dental Services Rendered.* Chicago, Ill: ADA Survey Center; 2002.
2. World Health Organization (WHO). *Concise International Chemical Assessment Document 50: Elemental Mercury and Inorganic Mercury Compounds: Human Health Aspects.* Geneva, Switzerland: WHO; 2003.
3. Echeverria D, Aposhian HV, Woods JS, et al. Neurobehavioral effects from exposure to dental amalgam Hg(0): new distinctions between recent exposure and Hg body burden. *FASEB J.* 1998;12:971-980.
4. Aydin N, Karaoglanoglu S, Yigit A, Keles MS, Kirpinar I, Seven N. Neuropsychological effects of low mercury exposure in dental staff in Erzurum, Turkey. *Int Dent J.* 2003;53:85-91.
5. Echeverria D, Woods JS, Heyer NJ, et al. Chronic low-level mercury exposure, BDNF polymorphism, and associations with cognitive and motor function. *Neurotoxicol Teratol.* 2005;27:781-796.
6. Ritchie KA, Gilmour WH, Macdonald EB, et al. Health and neuropsychological functioning of dentists exposed to mercury. *Occup Environ Med.* 2002;59:287-293.
7. Kingman A, Albers JW, Arezzo JC, Garabrant DH, Michalek JE. Amalgam exposure and neurological function. *Neurotoxicology.* 2005;26:241-255.
8. Factor-Litvak P, Hasselgren G, Jacobs D, et al. Mercury derived from dental amalgams and neuropsychologic function. *Environ Health Perspect.* 2003;111:719-723.
9. Dalen K, Lygre GB, Klove H, Gjerdet NR, Askevold E. Memory functions in persons with dental amalgam. *J Dent.* 2003;31:487-492.

10. Nitschke I, Muller F, Smith J, Hopfenmuller W. Amalgam fillings and cognitive abilities in a representative sample of the elderly population. *Gerodontology*. 2000;17:39-44.

11. Bjorkman L, Pedersen NL, Lichtenstein P. Physical and mental health related to dental amalgam fillings in Swedish twins. *Community Dent Oral Epidemiol*. 1996;24:260-267.

12. Saxe SR, Snowdon DA, Wekstein MW, et al. Dental amalgam and cognitive function in older women: findings from the Nun Study. *J Am Dent Assoc*. 1995;126:1495-1501.

13. Mutter J, Naumann J, Sadaghiani C, Schneider R, Walach H. Alzheimer disease: mercury as pathogenetic factor and apolipoprotein E as a moderator. *Neuroendocrinol Lett*. 2004;25:331-339.

14. Bates MN, Fawcett J, Garrett N, Cutress T, Kjellstrom T. Health effects of dental amalgam exposure: a retrospective cohort study. *Int J Epidemiol*. 2004;33:894-902.

15. Grandjean P, Guldager B, Larsen IB, Jorgensen PJ, Holmstrup P. Placebo response in environmental disease: chelation therapy of patients with symptoms attributed to amalgam fillings. *J Occup Environ Med*. 1997;39:707-714.

16. Nerdrum P, Malt UF, Hoglend P, et al. A 7-year prospective quasi-experimental study of the effects of removing dental amalgam in 76 self-referred patients compared with 46 controls. *J Psychosom Res*. 2004;57:103-111.

17. Trepka MJ, Heinrich J, Krause C, et al. Factors affecting internal mercury burdens among eastern German children. *Arch Environ Health*. 1997;52:134-138.

18. Evens CC, Martin MD, Woods JS, et al. Examination of dietary methylmercury exposure in the Casa Pia study of the health effects of dental amalgams in children. *J Toxicol Environ Health A*. 2001;64:521-530.

19. Gabrio T, Benedikt G, Broser S, et al. 10 years of observation by public health offices in Baden-Wurttemberg–assessment of human biomonitoring for mercury due to dental amalgam fillings and other sources [in German]. *Gesundheitswesen*. 2003;65:327-335.

20. Levy M, Schwartz S, Dijak M, Weber J-P, Tardif R, Rouah F. Childhood urine mercury excretion: dental amalgam and fish consumption as exposure factors. *Environ Res*. 2004;94:283-290.

21. Altmann L, Sveinsson K, Kramer U, et al. Visual functions in 6-year-old children in relation to lead and mercury levels. *Neurotoxicol Teratol*. 1998;20:9-17.

22. Walkowiak J, Altmann L, Kramer U, et al. Cognitive and sensorimotor functions in 6-year-old children in relation to lead and mercury levels: adjustment for intelligence and contrast sensitivity in computerized testing. *Neurotoxicol Teratol*. 1998;20:511-521.

23. Children's Amalgam Trial Study Group. The Children's Amalgam Trial: design and methods. *Control Clin Trials*. 2003;24:795-814.

24. Magos L, Clarkson TW. Atomic absorption determination of total, inorganic, and organic mercury in blood. *J Assoc Off Anal Chem*. 1972;55:966-971.

25. Barber TE, Wallis G. Correction of urinary mercury concentration by specific gravity, osmolality, and creatinine. *J Occup Med*. 1986;28:354-359.

26. Hornung RW, Reed LD. Estimation of average concentration in the presence of nondetectable values. *Appl Occup Environ Hyg*. 1990;5:46-51.

27. Schwartz J. Low-level lead exposure and children's IQ: a meta-analysis and search for a threshold. *Environ Res*. 1994;65:42-55.

28. Pocock SJ, Smith M, Baghurst P. Environmental lead and children's intelligence: a systematic review of the epidemiological evidence. *BMJ*. 1994;309:1189-1197.

29. Green LW. Manual for scoring socioeconomic status for research on health behavior. *Public Health Rep*. 1970;85:815-827.

30. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*. 1999;8:3-15.

31. Schultz CJ, Dalton RN, Turner C, Neil HA, Dunger DB; The Oxford Regional Prospective Study Group. Freezing method affects the concentration and variability of urine proteins and the interpretation of data on microalbuminuria. *Diabet Med*. 2000;17:7-14.

32. World Health Organization (WHO). *Environmental Health Criteria 118: Inorganic Mercury.* Geneva, Switzerland: WHO; 1991.

33. Kingman A, Albertini T, Brown LJ. Mercury concentrations in urine and whole blood associated with amalgam exposure in a US military population. *J Dent Res*. 1998;77:461-471.

34. Dye BA, Schober SE, Dillon CF, et al. Urinary mercury concentrations associated with dental restorations in adult women aged 16-49 years: United States, 1999-2000. *Occup Environ Med*. 2005;62:368-375.

35. Barregard L. Mercury from dental amalgam: looking beyond the average. *Occup Environ Med*. 2005;62:352-353.