First National Survey of Lead and Allergens in Housing: Survey Design and Methods for the Allergen and Endotoxin Components.

Patrick J. Vojta

Warren Friedman

David A. Marker

Please see article for additional authors.
We conducted the first National Survey of Lead and Allergens in Housing (NSLAH I) from July 1998 to August 1999 under the sponsorship of the U.S. Department of Housing and Urban Development (HUD) and the National Institute of Environmental Health Sciences (NIEHS). The purpose of the survey was to assess children’s potential household exposure to lead, allergens, and bacterial endotoxins. We surveyed a sample of 831 homes, representing 96 million permanently occupied, noninstitutional housing units that permit resident children. We administered questionnaires to household members, made home observations, and took environmental samples. This article provides general background information on the survey, an overview of the survey design, and a description of the data collection and laboratory methods pertaining to the allergen and endotoxin components. We collected dust samples from a bed, the bedroom floor, a sofa or chair, the living room floor, the kitchen floor, and a basement floor and analyzed them for cockroach allergen Bla g 1, the dust mite allergens Der f 1 and Der p 1, the cat allergen Fel d 1, the dog allergen Can f 1, the rodent allergens Rat n 1 and mouse urinary protein, allergens of the fungus Alternaria alternata, and endotoxin. This article provides the essential context for subsequent reports that will describe the prevalence of allergens and endotoxin in U.S. households, their distribution by various housing characteristics, and their associations with allergic diseases such as asthma and rhinitis. Key words: asthma, data collection, house dust mite, indoor allergens. Environ Health Perspect 110:527–532 (2002). [Online 3 April 2002] http://ehpnet1.niehs.nih.gov/docs/2002/110p527-532vojtajabstract.html

The purpose of the allergen and endotoxin components of NSLAH I was to quantify the concentrations and loads of the major indoor allergens and endotoxins in floor, furniture, and bedding dust in the nation’s housing stock. The specific objectives were to a) estimate the number and percentage of homes with allergen and endotoxin levels above established thresholds (if known) for sensitization and disease; b) describe the distribution of allergen and endotoxin levels by housing and household characteristics, such as geographic region, race, ethnicity, and socioeconomic status; and c) provide baseline data for developing the nation’s Healthy People 2010 objective for reducing indoor allergen levels and for establishing a reference point for future allergen and endotoxin surveys.

The indoor allergens that were the focus of the laboratory analyses are the cockroach allergen Bla g 1, the dust mite allergens Der f 1 and Der p 1, the cat allergen Fel d 1, the dog allergen Can f 1, the rodent allergens Rat n 1 and mouse urinary proteins (MUPs), and the allergens of the fungus Alternaria alternata.

Methods

Survey design. The target population for the survey was the national housing stock of approximately 96 million permanently occupied, noninstitutional housing units that permitted resident children, including multi-family buildings and manufactured housing.
The first stage of the design was the selection of 75 primary sampling units (PSUs). The PSUs consisted of metropolitan statistical areas (MSAs), counties, or groups of counties. We assigned every area in the 50 states and the District of Columbia to a PSU. In some cases, we grouped adjacent counties to form PSUs above the minimum PSU population of 15,000. We divided several large MSAs into multiple PSUs and combined several smaller MSAs into one PSU. We identified and grouped a frame of 1,404 PSUs into 62 strata and assigned and chose with certainty the 24 most populous PSUs to one stratum each (i.e., their probability of selection was 100%). From each of the remaining 38 noncertainty strata, we randomly selected two PSUs, with each PSU’s probability of selection proportional to its 1990 U.S. census population. This produced a sample of 100 PSUs. Because of budgetary constraints, we randomly drew a subsample of 75 PSUs from these 100 PSUs, in which we included all 24 certainty PSUs and at least one of the PSUs from each of the noncertainty strata. Figure 1 shows the locations of the 75 PSUs. [For a detailed list of the 75 PSUs, see National Institute of Environmental Health Sciences Web site (16).]

The second stage of the design was the sampling of segments within the 75 PSUs. A segment consisted of one or more contiguous blocks, depending on the number of housing units in the block. In most PSUs, we selected 10 segments, with the probability of selection proportional to the number of housing units in each segment, as reported in the 1990 U.S. census. However, we modified the number of segments in a few PSUs to reduce variation in selection probabilities. In all, we selected 754 segments.

The third stage was the sampling of eligible housing units within the selected segments. Within each of the selected segments, we developed a list of all housing units by listing, which involved having trained personnel visit each segment to verify and update information on segment boundaries, roads, and numbers of housing units and to list all housing unit addresses. If a segment contained too many homes to list, we divided the segment into chunks and listed housing units within one randomly selected chunk. With a target survey sample size of approximately 800–1,000 housing units and an expected participation rate of about 50%, we decided to sample two or three housing units per segment. In all, we listed 39,071 housing units and selected 1,984 of these housing units for recruitment into the survey; however, we had not established the eligibility of the housing units at this point.

The fourth design stage was the selection of rooms within the recruited housing units. We inventoried the number of rooms in a housing unit and grouped them into four strata according to room type—kitchen, common living area (living room, den, or family room), bedroom, and other rooms (such as guest bedrooms, dining room, office, bathroom). We developed two sets of four strata depending on whether or not the housing unit contained one or more bedrooms for children. If the house contained children’s bedrooms, they became one of the four strata and we listed the adult bedrooms as other rooms. Otherwise, we listed the adult bedrooms as one of the four strata. We selected an equal-probability sample of one room within each room stratum; however, we selected two rooms if one of the strata had seven or more rooms.

Response rate. Of the 1,984 housing units selected for recruitment into the study, we surveyed 831. We counted a housing unit as surveyed (completed) if we had collected at least one environmental sample in at least one of the target rooms. In the calculation of the response rate, we considered the fact that not all selected housing units were eligible for the survey. Of the 1,153 housing units that did not participate in the survey, 149 were known to be eligible, 229 were known to be ineligible, and 775 were of unknown eligibility. Unknown eligibility usually resulted from failure of the field team to make contact with a householder or a refusal on the part of the householder to complete enough of the screening to establish eligibility. Assuming that the eligibility rate among homes with unknown eligibility was the same as among the homes with known eligibility (81%), we calculated the overall response rate for the survey as follows:

\[
\text{Response rate} = \frac{831}{1984 - 229 - 775(0.19)} \times 100\% = 52\%.
\]

Statistical weighting. Each of the 831 surveyed housing units received a statistical weight that permitted them to be expanded to represent the U.S. population of all permanently occupied, noninstitutional housing units that allow resident children. The weight for a particular housing unit is the number of housing units it represents. The final housing unit weight is the product of the base weight and three weighting adjustment factors—the nonresponse adjustment factor, the trimming factor, and the poststratification factor. The base weight is the inverse of the housing unit’s overall probability of selection, which is the product of its probability of selection at each of the first three stages. The nonresponse adjustment factor inflates the base weights of the surveyed housing units so that they also represent the eligible housing units that did not participate in the survey (17). The trimming factor reduces excessively large weights. This reduction generally improves the overall accuracy of survey estimates (18,19). The poststratification factor aligns the housing unit counts with the counts of another survey, which in this case was the 1997 American Housing Survey (20). Poststratification improves the accuracy of the sample estimates (17,21,22). We poststratified the weights by census region, age of housing construction,
and whether or not a child younger than 18 years lived in the home.

For analyses that may involve the room as the unit of analysis, we calculated room weights. Poststratified housing unit weights were the basis for room weights. We divided these housing unit weights by the room probabilities of selection to produce room base weights. We then made a nonresponse adjustment to account for noncompleted rooms. We considered a room completed only if we had collected some samples and data from the room.

Respondent contact, screening, and recruitment. For each of the 75 PSUs, we assigned a field interviewer to contact, screen, and recruit housing units. Before the interviewer’s home visit, we sent each of the 1,984 sampled households an initial contact letter signed by the directors of the HUD Office of Healthy Homes and Lead Hazard Control and of NIEHS that explained the purpose of the study and informed the householder that an interviewer would be visiting in the near future. To maintain the probability sample, field interviewers attempted to recruit all assigned housing units. The field interviewer made no less than four attempts to contact an adult householder, spaced over various days of the week and times of the day.

After making contact with an adult resident of a home, the interviewer administered a short recruiting questionnaire to determine whether or not the home met the inclusion criteria, such as the allowance of child residents. If the housing unit was eligible, the interviewer attempted to recruit the home for the survey. We offered the householder a financial incentive to participate in the survey. Initially, we offered householders $50; however, to improve response rates, we increased this amount incrementally up to a maximum of $200. For recruited homes, the interviewer listed all rooms and the major furniture units. The field interviewer made no less than four attempts to contact an adult householder, spaced over various days of the week and times of the day.

Field data collection. Field personnel. We assembled two-person teams consisting of an interviewer and a technician for the field data collection, which took place from July 1998 to August 1999. The interviewer was responsible for introducing the team to the home occupants, obtaining informed consent, explaining the tasks involved, administering the resident questionnaire, and collecting the allergen dust samples. The technician was responsible for collecting samples for the lead component of the study. On average, the data collection took approximately 2–3 hours to complete in each home. A field director, assisted by two field supervisors, scheduled, coordinated, and supported all field team activities.

Resident questionnaire. Once we obtained informed consent from the adult contact, the interviewer administered the resident questionnaire, which collected the following information:

1. Housing-unit information: housing unit age, how long the respondent has lived in the home, number of stories, type of heating and air conditioning, type of flooring, presence of dehumidification system, cleaning schedules, presence of pets, presence of cockroaches and rodents, insecticide application, and current allergen avoidance practices
2. Household information: household size; household income; smoking patterns; the name, age, sex, race, and education of each resident; indoor activity patterns; and the presence of asthma and allergy diagnoses and symptoms for each resident
3. Lead-related occupations or hobbies
4. Allergen-related occupation or hobbies (e.g., veterinarian, exterminator, farm worker)

The resident questionnaire can be found in its entirety at the NIEHS Web site (16).

Selection of surfaces for sampling. For each housing unit, we collected five or six dust samples, one from each of the following locations in the stratified and randomly selected rooms: floor in the kitchen, floor and sofa (or chair) in a common living area, floor and a bed in a bedroom, and the basement floor (if present and if none of the other sampled rooms was in the basement).

We selected the sofa most often used by a child; if no child resided in the home, we selected the most used sofa. We sampled only upholstered sofas; if none was present, we selected a chair. If one or more children younger than 18 years resided in the home, we selected a bedroom and bed used by a child. Otherwise, we randomly selected the bedroom and bed from all bedrooms and beds in the home. Because not all homes had all of the target rooms and because field technicians were sometimes denied access to a particular room, we did not collect vacuumed dust samples from all of the sample sites in all of the housing units surveyed.

Dust collection procedures. We collected dust samples using the Eureka Mighty-Mite 7.0-ampere vacuum cleaner (Eureka Company, Bloomington, IL, USA). We placed a 19 mm × 90 mm cellulose extraction thimble (Whatman International, Ltd., Maidstone, UK) into the distal end of the vacuum’s extension tube and sealed it with a rubber O-ring gasket. We placed a clean crevice tool over the distal end of the extension tube.

We collected dust from each location as follows:

- Kitchen floor: We vacuumed dust along the perimeter of the floors, including spaces between floor appliances and walls or cabinets, from beneath tables and chairs, and then from the open floor area, until 5 min had elapsed.
- Sofa (or chair): We vacuumed seat cushions (including both sides of reversible cushions), arms, seat backs, and throw pillows for a total of 5 min. We did not vacuum areas under cushions or in deep crevices.
- Common living area floor: We vacuumed an area 91 cm × 183 cm (36 inches × 72 inches), or of comparable area, directly adjacent to the sampled sofa for 5 min. We collected floor samples before sampling the sofa.
- Bedroom bed: We vacuumed all bedding layers for a total of 2.5 min, the primary sleeping pillow for 30 sec, and the mattress surface for 2 min. We did not remove fully encasing mattress covers, if present.
- Bedroom floor: We vacuumed an area 91 cm × 183 cm (36 inches × 72 inches) adjacent to the sampled bed and with one-eighth of the area located under the bed where possible for 5 min. We collected the floor sample prior to sampling the bed to avoid floor sample contamination.
- Basement floor: If the basement room was a kitchen, bedroom, or common living area, we followed procedures described for those rooms; otherwise, we vacuumed a 91 cm × 183 cm area in the center of the largest open area for 5 min.

We recorded sample location data on dust sampling logs along with information such as room dimensions, floor surface type and condition, dimensions of areas vacuumed, the type of material for the sofa and bedding, the temperature and humidity in each room, the presence of air conditioning devices, and evidence of smoking, food debris, moisture, mildew, cockroaches, and rodents. We sealed dust samples in reclosable plastic bags and shipped them to the field office via an overnight carrier, but we did not ship them over a weekend. We placed dust samples into a freezer at −20°C immediately upon receipt at the field office.

Laboratory protocols. We shipped dust samples on dry ice via overnight delivery to the central laboratory at the Harvard School of Public Health. At the laboratory, we sieved each dust sample through a 425-µm pore-size grating and determined the weight of the recovered dust. We placed approximately
100 mg (exact weight measured and recorded) of dust in each of multiple screw-cap microcentrifuge tubes labeled with the sample identification number, date, and amount of dust. We stored these aliquots of fine dust at –20°C.

In preparation for allergen analyses, we extracted dust aliquots at room temperature for 2 hr on a wrist-action shaker in borate-buffered saline (pH 8.5). We extracted 2 mL per 100 mg of dust. We then clarified extracts by centrifugation at 1,300 g and pipetted the supernatant into screw-top storage tubes, which we froze until needed. We measured the cockroach antigen Bla g 1 by a two-site, monoclonal antibody enzyme-linked immunosorbent assay (ELISA) as described by Pollart et al. (23); the dust mite allergens Der f 1 and Der p 1 and the cat allergen Fel d 1 using antigen-capture ELISA assays employing monoclonal capture and detector antibodies (24-25); and the dog allergen Can f 1 using a monoclonal capture and polyclonal detector antibody protocol similar to those described by Schou et al. (26) and Ingram et al. (27); and the rat allergens according to previously published methods (28). We assayed MUP using purified antigen and a polyclonal rabbit anti-MUP antibody (Greer Laboratories, Inc., Lenoir, NC, USA) and goat anti-rabbit antibody (Sigma Chemical Co., St. Louis, MO, USA) in a competitive inhibition ELISA as described by Miller et al. (29). We performed A. alternata allergen analyses with a similar competitive inhibition ELISA using a polyclonal anti-Alternaria antibody (Greer Laboratories). We performed all ELISAs in Immulon 2 microtiter plates (Thermo Labsystems, Helsinki, Finland) and read them kinetically over 5 min on an OPTIMax microplate reader (Molecular Devices Corp., Sunnyvale, CA, USA). Using SOFTMax Pro software (version 2.4.1; Molecular Devices Corp.), we plotted standard curves of the maximum kinetic rate (milli-optical density units per minute) of duplicate allergen standards against the concentrations of the allergen standards using four-parameter curve fitting. We used maximum kinetic rates of test samples to determine their allergen concentrations from the standard curve.

For endotoxin assays, we determined the concentration of endotoxins in the sieved dust using the Limulus amebocyte lysate assay (30). We extracted precision weighed samples in sterile, pyrogen-free water containing 0.05% Tween-20 for 1 hr at 22°C with continuous shaking, and then centrifuged extracts and transferred supernatants to pyrogen-free cryotubes. We prepared 2-fold serial dilutions of endotoxin standards and sample extracts using sterile, pyrogen-free water with Tween-20 in borosilicate glass tubes that had been heated for 4 hours at 200°C to remove endotoxin activity. We performed a 12-point calibration curve with four-point endotoxin determination. The standard curve ranged from 0.05 to 100 EU/mL of standard endotoxin from Escherichia coli 0111:B4. We pipetted aliquots (100 µL) of the serial dilutions of endotoxin standards and extracts into pyrogen-free polystyrene microplates and assayed them via the addition of the Limulus amebocyte lysate reagent and substrate. We measured the absorbance in each well at 405 nm every 30 sec for 90 min. We based endotoxin determinations upon the maximum slope of the absorbance versus time plot for each well. Four assay reagent blank wells served as reference and control for the pyrogen-free status of the reagent water, centrifuge tubes, pipette tips, and microplates. We calculated the endotoxin value for an individual sample from the arithmetic mean of those dilutions that fell in the middle two-thirds of the standard curve.

Quality assurance. Before we implemented the survey, all required authorities reviewed and approved the protocol. First, we submitted the proposed information collection to the U.S. Office of Management and Budget (OMB) for review and published a notice of the proposed data collection in the Federal Register on 3 October 1997 (31). Next, in accordance with OMB’s requirements, we submitted an Information Collection Request to OMB, which it approved on 30 April 1998. Finally, the Westat Institutional Review Board approved the survey plan on 20 May 1998, and the NIEHS Institutional Review Board approved it on 16 June 1998. HUD and NIEHS employed Westat to assist with the design of the survey, the development of the data collection protocols, the data collection, and the development of the survey database.

We developed written protocols and field tested them for all aspects of data collection: listing of housing units, interviewing home residents, measuring and recording physical data, collecting environmental samples, and handling of equipment and samples. We recruited a number of homes in the Washington, DC, metropolitan area to field test screening and recruiting procedures, the completeness of the introductory letter, and the effectiveness of recruitment strategies, including incentives to participate. Senior staff and subject matter experts conducted field data collection activities at five homes in the Washington, DC, area to test procedures and forms. Following the pilot studies, we modified recruitment efforts and the written field protocols, forms, and questionnaires as necessary. Data coding and data entry staff reviewed the protocols and forms to ensure efficient and accurate transfer of the data into a database.

During the data collection visit, the field team, using a checklist, performed a manual edit of all data and samples collected. The edit entailed an item-by-item proofreading of all forms to ensure that all required information had been collected and properly recorded, all required samples were collected, and all information was legible and consistent. We labeled all samples with preprinted, bar-coded labels using a standard numbering scheme. We returned data collection materials for each housing unit to the field office by overnight carrier as soon as the detailed review was completed. Once the data collection materials arrived at the field office, the field supervisor checked all field data and samples for completeness and consistency. We reconciled any errors as soon as possible with the field team prior to submission of the samples to the laboratory.

During the survey, the field director and supervisors contacted a random sample of completed households by telephone to verify the field team’s activities. Representatives of HUD, NIEHS, and Westat conducted random field audits of the field teams to verify the accuracy and completeness of data collected. We conducted audits by accompanying and observing the field team during data collection activities. We recorded observations and recommendations on a field team audit form and reviewed the findings with the team immediately following the data collection visit. We attempted to audit each team member at least once. If we noted problems, we conducted a second audit.

Sources of data error. Two broad types of error can occur in any survey such as the NSLH I: sampling error and nonsampling error. When present, these sources of error produce a parameter estimate, such as a mean or an odds ratio, that differs from the true population parameter. Sampling error arises from surveying a sample of the population rather than the complete population. Sampling error is measured by the standard error, which is a function of the sample size and the sample design. To produce valid standard errors for this survey, design features such as stratification, clustering, and unequal probabilities of selection must be accounted for using specialized statistical methods designed specifically for that purpose, which can be implemented in practice with special statistical software. Otherwise, the standard errors will be misleading and will result in invalid confidence intervals and p-values.

Nonsampling error occurs because of important differences between responders and nonresponders in a survey and because of inaccuracies in obtaining study information.
In this survey, potential sources of nonsampling error include response rates that differ by demographic characteristics, types of housing units, and geographic region; differences between the sampling frame and the target population of all permanently occupied, noninstitutional housing units that allow resident children; and measurement error in dust sample collection and laboratory analyses.

We conducted two analyses to estimate the potential impact of survey response rates on the estimated prevalence of allergens in the survey. First, we examined differences in response rates by 1990 U.S. census block group characteristics. Response rates differed significantly by the following block group characteristics: percentage Hispanic population, percentage black population, percentage below the poverty level, housing age, and percentage of housing units owned. We found the highest response rates among block groups with 30% or more Hispanic population, 30% or more in poverty, and in newer housing (1978 or later). We believe that the monetary incentive and the fact that we matched field technicians, whenever possible, to households on the basis of race or ethnicity helped to achieve a higher response rate among Hispanic and low-income groups. Because we calculated nonresponse adjustment factors within cells defined by the block group characteristics percentage of low-income population, percentage of Hispanic or black population, and percentage of pre-1940 and pre-1960 housing, use of the final adjusted statistical weights should greatly reduce this potential bias.

Second, we compared the weighted characteristics of the survey sample with characteristics of two other national surveys: the 1995 and 1997 American Housing Survey (AHS) (20) and the 1998 and 1999 Current Population Surveys (22) (Table 1). The 831 homes in the NSLAH I represent approximately 96 million permanently occupied, noninstitutional housing units that allow resident children. As shown in Table 1, the 95% confidence intervals (CI) for the NSLAH I contained the American Housing Survey or Current Population Survey estimate for most of the variables. For the remaining variables, the disparities were marginal.

**Conclusion**

The NSLAH I was the first survey to collect allergen and endotoxin samples from a nationally representative sample of the U.S. housing stock. In this article we provide the essential context for subsequent reports that will describe the prevalence of allergens and endotoxins in the U.S. housing stock, their distribution by various housing characteristics, and their associations with asthma and other allergic diseases. We also provide the framework for future population-based surveys of indoor allergens and endotoxins.

**REFERENCES AND NOTES**


