

NONSAMPLING ERROR CONSIDERATIONS IN THE DESIGN OF THE NATIONAL  
HOUSEHOLD PESTICIDE USAGE SURVEY

Douglas J. Drummond and Richard D. Waddell, Research Triangle Institute

### 1. BACKGROUND

The Environmental Protection Agency (EPA) is the principal regulator of pesticides in the United States. EPA's pesticide regulatory program focuses on the registration of new pesticides, development of registration standards, and a special pesticide review program called rebuttable presumption against registration (RPAR). Under this latter program, EPA weighs the risks and benefits of pesticides suspected of posing danger to the public or environment and decides whether or not to take regulatory action. EPA also conducts pesticide research, monitoring, enforcement, and training programs, many of which involve cooperation with states and other federal agencies.

To administer all of these pesticide programs, EPA needs information about where pesticides are used and in what quantities. Currently this information is derived from a variety of sources including private reporting services and United States Department of Agriculture (USDA) surveys. Data obtained from such sources, however, are generally not intended to be comprehensive and are frequently out-of-date or otherwise not reliable enough to meet the needs of most users of such information. Data on pesticides used in the home and urban environments are particularly prone to those criticisms.

In response to this need for more adequate pesticide use information, EPA has initiated a program to provide quantitative pesticide usage data for urban/nonfarm sites in the United States by conducting a series of surveys to be repeated on a three- to five- year cycle. Discussions in this paper will address the first of these efforts, namely, the National Household Pesticide Usage Survey (NHPUS).

Specifically, the intent of this paper is threefold: (1) to make the statistical community aware of the body of literature associated with the NHPUS design efforts; (2) to discuss nonsampling error considerations associated with measuring total product usage at a housing unit during a twelve month reference period; and (3) to overview major design decisions made to differentially control for sampling versus nonsampling errors in assessing total product usage. No attempt is made to describe in any detail the NHPUS sample design, instrumentation or field procedures.

### 2. OVERVIEW OF NHPUS DESIGN

The target population for the NHPUS is the noninstitutionalized, nonfarm household population of the 48 coterminous states and the District of Columbia for a specified 12-month reference period. Data are to be gathered about household usage of pesticides by the target population during this reference period. These data will include site/pest and quantity/frequency information about the pesticides used (purchased) by the households, safety precautions taken, methods of application, container and unused chemical disposal practices, efficacy, label comprehension, storage location(s) for pesticide containers,

storage location(s) for pesticide containers, sources of recommendation for products used/purchased, and information concerning the use of the services of a commercial pest control operator. Such data are needed by EPA in conducting human/environmental exposure and risk analyses and for risk/benefit analyses of products or groups of products used in the household.

Collecting data of this nature and specificity has been demonstrated to require that household members be educated as to what constitutes a pesticide product and that repeated measurements be taken in that pesticide usage is not so salient an event that a household can reliably recall their usage for an entire twelve month period. Moreover, the realization of accurate usage data mandates the imposition of a recurring physical measurement process (i.e., weighing containers over time). Finally, a mechanism is needed to account for usage associated with products no longer on the premises at time of interview.

In response to these needs, the NHPUS will be supported by a longitudinal housing unit design requiring the collection of pesticide usage data via personal interviews on three separate occasions. In addition, a baseline interview will be conducted at each sample housing unit prior to the start of the study reference period in order to weigh and tag all containers of pesticides on the premises and to initiate completion of a Purchase/Disposal Ledger. All pesticide containers found on the premises during any subsequent round will be weighed and tagged (if necessary) and purchases/disposals during each reporting period abstracted from the Purchase/Disposal Ledger. Reporting periods under a recurring physical measurement process will vary by housing unit and household (i.e., are determined by interview dates and/or period of residency in the housing unit). Moreover, the non-uniform nature of pesticide usage over time by a household generally precludes any analytic effort to reliably associate such data with prespecified intervals of time. Finally, to address the perceived decision-making needs of the Agency, the NHPUS will be supported by an initial self-weighting sample of 6,480 housing units distributed over 1,080 penultimate sample units consisting of compact area segments within a large, highly stratified first-stage sample of 180 counties. It is anticipated that the survey procedures being proposed (including cash incentives) should realize a response rate in excess of 80 percent across all rounds of data collection.

Implementation of the NHPUS design will entail field staff involvement over a 24-month period -- Table 1 provides the details. Household interviews would be scheduled so as to realize a maximum five month recall period in each of the three enumeration rounds, while at the same time minimizing downtime on the part of the permanent field staff. Moreover, such a

schedule addresses the need to carry out baseline interviews prior to the start of the reference period, and the final enumeration round after the end of same.

Completed data collection instruments will be edited and processed by round, and certain pertinent data elements (e.g., control cards, product summaries) will be generated for field use during each of the rounds following the first interview. Moreover, activities underlying the implementation of the NHPUS design will be monitored on a continual basis with particular attention paid to response rates, data quality and the expected precision of study estimates.

It must be recognized at the outset that little is currently known about household pesticide usage patterns and that the proposed field procedures and instrumentation have been exposed to very limited field testing. As such, field procedures may have to be modified, additional training furnished and/or adjustments made in the sample size based on initial implementation efforts.

### 3. SOURCES OF NONSAMPLING ERROR IN MEASURING PESTICIDE USAGE

The NHPUS instrumentation incorporates five mechanisms geared to the collection of quantitative information on pesticide usage:

- a. direct measurement (via weighing containers over time);
- b. respondent estimates of product usage (by indoor/outdoor);
- c. amount of product purchased;
- d. container level estimates by respondents of date purchased, date disposed of, and percent of contents remaining at time of disposal; and
- e. respondent estimates of the frequency of application by site as well as pests treated by the aggregate of such applications at each site.

For the most part, the multiple approaches taken to assessing pesticide usage reflect the reality that no one affordable approach can successfully work for all types of containers, nor can a single method work in all contexts that the data are being sought.

#### 3.1 Eligibility Requirements

The highest priority item in the NHPUS is the estimation of the total amount of active ingredient used during a 12-month period. Providing such an estimate must proceed in two phases: (1) obtaining the amount of formulated product used; and (2) converting such usage to equivalent active ingredient units by using the product EPA Registration Number to access the appropriate EPA database. Clearly, the absence of an EPA Reg. No. will preclude any such analysis. Accordingly, for the purposes of the NHPUS, a pesticide is defined as any product present in its original container and which has a discernible EPA Reg. No.<sup>1</sup> Moreover, the amount of formulated product must reflect pesticide usage by households during the reference period while they were residents of sample housing units.

#### 3.2 Estimating Amount of Formulated Product Used (Ideal Case)

Consider the following scenario applied to each household residing at a sample housing unit

- during any portion of the reference period.
- on the first day of residence at a sample housing unit, all containers of pesticides stored on the premises are weighed.
  - any containers of pesticide purchased by the household<sup>2</sup> while in residence at the housing unit during the reference period are weighed prior to use.
  - before disposing of any pesticide container, weight of same is recorded
  - all pesticide containers stored on the premises of the sample housing unit are again weighed on the last day of the reference period or last day of residence for the household in that sample housing unit, whichever occurs first.
  - differencing container weights (i.e. initial minus final) yields an estimate of formulated product usage from that container.
  - total formulated product usage for that household and housing unit is then obtained by summing over all containers of pesticide having the same EPA Reg. No.

Even this "ideal" scenario is subject to nonsampling error. For example: (1) potentially includes off-site usage; (2) weight difference may not be attributable to pesticide usage (e.g., anhydrous substances; propellant discharges; scale errors); (3) excludes pesticides used on premises but not stored there (e.g., pesticides used on premises by maid service, gardner or commercial pest control operators); and (4) unable to account for usage of pesticide products not stored in their original container. Moreover, direct weighing even in this ideal setting is not capable of estimating usage of a given chemical (i.e., active ingredient) specific to a site or pest at a particular sample housing unit.

#### 3.3 Estimating Amount of Formulated Product Used Under NHPUS Design

The proposed NHPUS instrumentation and field procedures address three naiveties inherent to the ideal measurement scenarios described in the previous subsection: (1) not all products are amenable to weighing; (2) weighing can only occur during the four visits made to each sample housing unit; and (3) logistics of identifying and tracking all pesticide containers at a housing unit are non-trivial. Each will be addressed in turn followed by a statement summarizing the joint impact of efforts taken to address these concerns.

##### 3.3.1 Estimating Pesticide Usage from Containers Unamenable to Weighing

Not all containers of pesticides are amenable to weighing. Specifically, some containers are too bulky and/or exceed the limits of the scale. Still others are just too messy to be subjected to weighing due to potential risk to the enumerator. Finally, some containers are designed to continuously emit pesticide (e.g., rat baits, no-pest strips, pet flea collars) which are better accounted for in whole units used (as opposed to weight). Unfortunately, the only recourse to estimating usage via direct

weighing in such cases is that of respondent recall of usage which is subject to excessive nonsampling error. In an effort to control same, the NHPUS design proposes to: (1) employ scales which can be calibrated on-site and which have adequate range/precision (e.g., 0-22 lbs. in quarter ounce increments); (2) educate interviewers as to label information that can alert them to particularly toxic products; (3) require that interviewers wear gloves while handling pesticides; and (4) account for continuous use products on a unit basis.

### 3.3.2 Coverage Problems in Using Direct Weighing

Direct weighing intends to measure usage from a container between interviews as opposed to usage from that container for the total intended reporting period for that household. Three scenarios will be described which identify deficiencies in the sole use of direct weighing as a method of studying product usage.

#### 3.3.2.1 Inclusion of Ineligible Periods of Usage

Under direct weighing, usage during ineligible periods occurs in three circumstances: (1) for households in residence at a sample housing unit for the baseline and first round enumeration (i.e., direct weighing incorrectly includes usage between the date of the baseline interview and the start of the reference period); (2) for households in residence at a sample housing unit for the second and third enumeration rounds (i.e., direct weighing incorrectly includes usage between the end of the reference period and the date of the third enumeration); and (3) for households which move from a sample housing unit and which are not enumerated until they are in their new (nonsample) housing unit (i.e., direct weighing incorrectly includes usage between the date of moving and the date of the enumeration).<sup>3</sup> These periods of ineligibility are structurally imposed by the direct weighing schedule and can only be ameliorated not eliminated. This notwithstanding, the following control mechanisms are proposed.

- times between the baseline interview and the start of the reference period and between the end of the reference period and the third round interview at a housing unit be kept as short as possible (subject to resource availability, etc.), and to correspond with periods of slack pesticide usage (e.g., Oct.-March).
- every effort made to secure an interview with out-movers immediately prior to their departure.
- respondents asked to guesstimate their usage (by product) during these periods of ineligibility.
- reference period chosen to be easily identified with (e.g., calendar year) to reduce telescoping in respondent recall.
- households will be sensitized to the role/importance of the start/end dates for the reference period.
- consideration will be given to employing randomization in scheduling the Baseline Interview so as to better control for differential biases

inherent in adjusting for usage outside the reference period based on respondent recall.

#### 3.3.2.2 Exclusion of Eligible Periods of Usage

Whenever the first interview with a household occurs during the reference period, subsequent direct weighing of containers will be unable to reflect usage from that container from the first day of residence at the sample housing unit (or start of reference period, whichever occurs latest) to the date of first interview. The guesstimate of such usage, by product, must be provided by respondent recall, which at best is poor. Moreover, such guesstimation has been found to be impossible whenever no container of the product remains on the premises to be inventoried during the first interview (e.g., product used in reporting period and all containers of same discarded prior to first interview). Little can be done to address this source of undercoverage. Nevertheless, NHPUS field procedures do attempt to determine from the out-moving household some estimate of arrival date for the in-moving household. When practical, the first interview with the latter household will occur as soon after moving in as is possible and prudent. Moreover, no attempt will be made to estimate usage for products not on the premises at the time of first interview.

#### 3.3.2.3 Lack of Initial/Final Container Weight

Containers (for the most part) cannot be weighed immediately prior to first use nor immediately prior to disposal. Moreover, storing discarded pesticide containers (with remaining contents) is not viable for most products due to obvious risks associated with same. The NHPUS design assumes that an external estimate of initial weight can be secured (e.g., from manufacturer and/or by purchasing product and weighing that container).<sup>4</sup> For disposal weight the NHPUS instrumentation (specifically, the Purchase/Disposal Ledger) solicits information on the date of disposal and the categorized percent of contents remaining at time of disposal. Information would then be needed as to the empty container weight in order to arrive at an estimate of usage. Clearly, pesticide products purchased and disposed of between interviews pose the greatest problem in these regards. Moreover, household moves tend to exacerbate these recognized measurement deficiencies. Specifically, out-movers have a tendency to discard certain pesticides rather than transport same to their new residence whereas in-movers are known to purchase and use increased quantities of some products (e.g., disinfectants). Unfortunately, households are least likely to incur the burden and imposition of maintaining their Purchase/Disposal Ledger during these critical periods. Nevertheless, NUPAS field procedures intend to emphasize the importance of completing the Purchase/Disposal Ledger during the Baseline (initial) Interview with a household and to employ postcard reminders of same between rounds. Moreover, a cash incentive will be provided to all out-movers who mail in their Purchase/Disposal Ledger so as to avoid losing this important source of information.

### 3.3.3 Difficulties in Identifying and Tracking Pesticide Containers

Estimating product usage by weighing containers of same over time is predicated on being able to identify and track such containers for an extended period of time. Unfortunately, there is substantial variability between households in their perceptions as to what constitutes a pesticide product. Accordingly, NUPAS instrumentation attempts to screen households for the presence of pesticide products based on correlated housing unit characteristics (e.g., pool, garden, fruit trees) and by having respondents report on their use of products to combat a series of pest scenarios. All products identified during screening must later be accounted for during the formal inventorying of pesticide containers. Finally, households are encouraged to allow the enumerator to visit all storage locations for the purposes of inventorying all pesticide containers on the premises.<sup>5</sup> As part of the inventory, all containers are assigned a unique computer-generated id number and the EPA Reg. No, name and initial weight are recorded on the appropriate survey instrument. In any subsequent interview(s) with the same household, the initial screening will be supplemented by a Pesticide Summary that profiles all (active) pesticide containers inventoried as of the last interview. All containers listed either on the summary or on the Purchase/Disposal Ledger will have to be accounted for (i.e., processed on current inventory, disposed of as per the Purchase/Disposal Ledger, or of unknown disposition after probing). The Pesticide Summary will also serve as a flagging device for containers processed during previous rounds but have missing or inconsistent data elements for key analysis characteristics (i.e., name, EPA Reg. No., tag number, and previous weight). It should be recognized that pesticides are stored in a host of locations in and around the home (e.g., garage, basement, under kitchen sink, laundry room, etc.) and that no one single respondent may be knowledgeable concerning the entire spectrum of products.

### 3.3.4 Nonsampling Errors Induced by NHPUS Field Procedures and Instrumentation

The current NHPUS field procedures and instrumentation are burdensome, repetitive and themselves a source of further nonsampling error. Specifically, our concerns rest with response rate, enumerator errors, and errors due to respondent recall. In the former regard, pesticide usage may often be regarded as unimportant by the respondent, thus jeopardizing their willingness to incur large amounts of burden, inconvenience or sacrificing of their privacy before refusing to participate further in the survey. With respect to enumerator error, two sources of same are particularly prevalent: failure to identify the most knowledgeable respondent(s) at a household; and employment of incorrect field procedures brought on by the extreme complexity in same (e.g., reporting periods vary by household and housing unit for a given round). It is currently proposed to train enumerators at two points in time (i.e., Baseline and before first enumeration round). Moreover, consideration will be

given to simplifying the field procedures/instrumentation to further control for enumerator errors. Finally, an attempt is made to control for errors associated with respondent recall (e.g., total product usage during period of eligibility/ineligibility) by inventorying all containers of the product before posing the question, by bounding length of recall period (i.e., maximum of five months) and by sensitizing respondents as to the need for usage estimates for periods of ineligibility. Consideration may yet be given to reducing the number/nature of products for which pesticide usage data is sought. Clearly, reductions of this latter sort could have a dramatic effect on burden and saliency to the respondent.

### 3.4 Sources of Nonsampling Error in Collecting Usage Data Specific to a Site/Pest

Weighing containers over time fails to address the Agency need for quantitative usage data specific to a chemical/site/pest (i.e., limited to aggregate usage data for all sites/pests at a particular housing unit). To address this deficiency, respondents in the NHPUS are asked for their subjective estimate of the total product usage (by indoor/outdoor) and for the number of times they applied the product during the reporting period at a series of mutually exclusive site categories (separately for indoor/outdoor). Estimates of site-specific usage data would thus proceed in three stages: estimate total product usage in reporting period (by direct weighing); use subjective estimate of proportion of product used indoors (outdoors) to estimate indoor (outdoor) usage; and assume an equal amount of product usage at each site/application (which could vary by indoor/outdoor) to estimate site-specific usage. Information is then collected on the identity of all pests treated at a site by the collection of applications at same. Clearly, more than one target pest can be associated with a single application.

The nonsampling errors associated with estimating usage specific to a site/pest are primarily due to the recall of non-salient events (which in turn may be aggravated by a failure to solicit the required information from the most knowledgeable respondent for same). This notwithstanding, reporting errors are also due to the inability of a respondent to correctly identify the pest(s) being treated at a site. The following features of the NHPUS field procedures and instrumentation address these recognized sources of nonsampling error: (1) respondent recall of total product usage is only used to determine relative use (by indoor/outdoor); (2) length of recall period limited to 5 months; (3) all containers of a product are inventoried before respondents are asked to recall usage data (i.e., concrete reference set); (4) bounded reference periods used (e.g., since last interview, etc); (5) interviewers trained to recognize pests indigenous to their survey area and to be knowledgeable about the typical products used to combat same; and (6) site and pest categories used for reporting are made as small in number as possible and were determined with training requirements for enumerators in mind.

### 3.5 Other Factors Affecting Control of Non-sampling Errors Associated with Assessing Usage

The primary objective of the NHPUS is to provide quantitative pesticide usage data for households during a twelve month reference period. Nevertheless, information will be collected and reported on characteristics related to such usage, including storage, efficacy, safety precautions, application equipment, and disposal practices. For the most part, these latter estimates are of secondary importance to the Agency despite having a direct bearing on risk/benefit analyses. Accordingly, design efforts in support of NHPUS attempted to express these aforementioned study objectives and priorities in terms of the relative total error of desired parameter estimates. The priority placed on estimating usage parameters and the obvious dominance of the nonsampling error component of total error for same gave rise to two design features: (1) minimum sample sizes<sup>6</sup>; and (2) suppression of varying degrees of specificity in the collection of characteristics associated with pesticide usage. For example, NHPUS instrumentation collects information on storage location for each container inventoried (by round) whereas information on the application equipment utilized is only collected for the aggregate of all products used in a given round. Unfortunately, the collection of any auxiliary characteristics further increases the respondent burden and hence may negatively impact on response rate. Consideration may yet be given to only collecting non-usage characteristics on a subset of sample housing units.

#### 4. CONCLUDING REMARKS

EPA has a need for usage data specific to a chemical/pest/site. Indeed, the decision-making needs of the Agency would ideally be supported by a survey in which the measurement process would entail the tracking of containers of pesticide over time and would collect information relative to amount used, where/why/how/by whom used, safety precautions taken, storage location of container prior to use, and efficacy for each application. Clearly, however, such an information network is not feasible in a real world survey context under limited funding and voluntary study participation. Proposed NHPUS instrumentation and procedures do, however, attempt to collect complete data on total pesticide usage by product as well as partial information concerning activities related to such usage. For the most part, the NHPUS derives such usage data from the direct weighing of pesticide containers over time. The intent of this paper has been to overview sources of nonsampling error associated with such a methodology and to describe the steps taken in the NHPUS instrumentation and field procedures to control for same. Discussions are intended to highlight the immense complexities associated with the proposed methodology and to argue that a compromise has (and must) be struck between sampling and nonsampling errors within the limits of available resources, priorities and constraints. Unfortunately, funding cut-backs in the Agency have delayed implementation of the

current NHPUS design. As such, no quantitative assessment is currently available as to the success, indeed viability, of the measures proposed. It is, however, anticipated that the NHPUS design will eventually be implemented perhaps on a smaller scale and that usage estimates will be updated every five years.

#### FOOTNOTES

<sup>1</sup>In cases where the absence of an EPA Reg. No. cannot be verified (e.g., part of the label missing or obliterated), the container will be inventoried providing that the complete name can be recorded. The product name will then be matched with the appropriate EPA database in order to determine its EPA Reg. No. (if any).

<sup>2</sup>The population at risk at a sample housing unit in a given reporting period consists of all non-institutionalized individuals who have been a resident of that housing unit at any time during the reporting period.

<sup>3</sup>Households moving more than 50 miles from a sample PSU will not be followed up by personal interview. Instead, a cash incentive is currently being proposed for return of the Purchase/Disposal Ledger.

<sup>4</sup>Variability between containers of initial weight for a product is assumed negligible.

<sup>5</sup>Respondents may opt to bring all pesticide containers on their premises to a central site for inventorying, etc.

#### BIBLIOGRAPHY

- [1] "NHPUS Pilot Study Final Report" RTI Report No. 1864/18-01F and 1864/18-02F dated January, 1981.
- [2] "Analytic Capabilities of the Proposed NHPUS" RTI Report No. 1864/20-02W dated 6 May, 1981.
- [3] "NHPUS: California Pretest Report" RTI Report No. 1864/20-04F dated June, 1981.
- [4] "NHPUS Interviewer's Field Manual (Baseline Round)" RTI Report No. 1864/20-08F dated September, 1981.
- [5] "NHPUS Interviewer's Field Manual (Rounds 1-3)" RTI Report No. 1864/20-09F dated September, 1981.
- [6] "OMB Submission Package" RTI Report No. 1864/20-10F dated September, 1981.
- [7] "Documentation of NHPUS Sample Members: Stage I and II" RTI Report No. 1864/20-11F dated September, 1981.
- [8] "Implementation Concerns Regarding the Current NHPUS Materials" RTI Report No. 1864/20-12F dated 16 October, 1981.

[9] "Nonsampling Error Considerations in the Design of the NHPUS (Part I: Main Text; Part II: Instrumentation)" Paper distributed in support of presentation of contributed paper at 1982 ASA meetings.

Table 1. Overview of Proposed Schedule for Field Staff Activities<sup>1</sup>

Activity	Schedule <sup>2</sup>		
	Starting Month	Duration [Months]	Completion Month
Hire and train counting and listing staff.	1	1	1
Counting and listing.	2	2	3
Select housing unit sample, prepare Baseline instruments, hire additional interviewers, and train field staff for Baseline Interviews.	4	1.5	4.5
Administer Baseline Interview.	4.5	2.5	7
Prepare for first round enumeration and train interviewers on remaining instruments/procedures.	8	1	8
First enumeration round.	9	4	12
Second enumeration round.	14	5	18
Third enumeration round.	20	4	23
Final Field Staff debriefing.	24	1	24

<sup>1</sup>Reference period under this schedule corresponds to months 8 through 19. (inclusive)

<sup>2</sup>Activities are assumed to begin on the first day of the starting month and end on the last day of the completion month (unless otherwise specified).