

---

**DEFINITION OF THE SAMPLING FRAME AND SAMPLE**

---

The sample was drawn from records of Army, Air Force, Marine Corps, and Navy personnel who were reported to have served in ODS/DS between August 1, 1990, and July 31, 1991. These records were subset to those personnel who served on the ground (as opposed to those located at sea in the Persian Gulf or who only flew over the area) in the Kuwaiti theater of operations (KTO). The survey sample was designed to meet the concurrent objectives of providing (1) an overall estimate of pesticide exposure across all services, (2) individual estimates by service, and (3) estimates for various situations or “pesticide scenarios.” It achieved this by drawing a diverse sample from across all the services and by oversampling various subgroups. In all, 3,264 records were sampled from 536,790 eligible records, evenly divided across the Army, the Air Force, and the Marine Corps combined with the Navy. Allowing for failure to locate some respondents, nonresponse by others, and errors in the original data (such as misclassifying some personnel as present in theater when they were not), the initial sample of 3,264 records was expected to yield 2,000 complete responses. The survey actually obtained 2,005 complete responses.

**THE SAMPLING FRAME**

The sampling frame was designed to be an enumeration of military personnel located on the ground in-theater between August 1, 1990, and July 31, 1991. Information from the Personnel and Monthly databases, acquired in October 1998 from USASCURR, augmented with unit location indicators derived from the Locations database by OSAGWI, was assembled to create the sampling frame. It consists of all Army and Marine Corps personnel located in Saudi Arabia, Kuwait, and Bahrain; all Air Force personnel located in Saudi Arabia, Kuwait, Bahrain, Qatar, the United Arab Emirates, and Oman (hereafter referred to as “in theater”); Navy personnel in units that could be identified as being ashore in theater; and all Coast Guard personnel. Thus, the sampling frame consists of the U.S. Armed Services Center for Unit Records Research

(USASCURR) Personnel database, merged with the Monthly database and derived location indicators, less the following records:

- All Marine Corps personnel having UICs with latitude and longitude coordinates consistently located at sea in the Persian Gulf and thus assumed to have been permanently stationed on a ship;
- All Air Force personnel with location codes outside Saudi Arabia, Kuwait, Bahrain, Qatar, the United Arab Emirates, and Oman; and
- All Navy personnel at sea.

The sampling frame included as many records as possible; that is, it defaulted in favor of including records rather than excluding them. For example, a sampling frame based solely on personnel location would by design exclude many records with missing data and thus would possibly introduce unknown biases. Note that the Air Force locations included in the sampling frame encompass the United Arab Emirates and Oman, as well as Saudi Arabia, Kuwait, Qatar, and Bahrain, all of which were designated as part of the KTO and geographically adjoin Saudi Arabia. They were included because the Air Force based a significant population in these two countries (approximately 10,000 personnel or roughly 13 percent of the force) under conditions that were likely to have been similar to conditions on bases in Saudi Arabia.

This left 536,790 records in the sampling frame, as shown in Table B.1, from which a stratified random sample of personnel was drawn.

**Table B.1**  
**Number of Personnel in the Sampling Frame by Service**

Army	349,622
Air Force	78,659
Marine Corps	95,441
Navy	12,220
Coast Guard	848
Total	536,790

### **The Data Files Used to Construct the Sampling Frame**

The U.S. Armed Services Center for Unit Records Research (USASCURR), located at Ft. Belvoir, Virginia, created and maintains the Department of Defense (DoD) Persian Gulf Registry of Personnel and Unit Movement databases in response to the National Defense Authorization Act for Fiscal Years 1992 and 1993, Public Law 102-190 (DoD to Establish Persian Gulf Registry, Section 734)

and Public Law 102-585 (Persian Gulf War Veterans' Health Status, Section 704, Expansion of Coverage of Persian Gulf Registry). USASCURR maintains or has archived the following databases:

- The Personnel Registry (“Personnel”) database contains information on more than 757,000 Army, Navy, Marine Corps, Air Force, and Coast Guard personnel who served in the Gulf between August 1, 1990, and December 31, 1991. This database was an extension of the Desert Shield/Desert Storm database originally compiled by the DMDC.
- The DMDC unit identification code (“Monthly”) database contains periodic UIC data for most of the personnel in the Personnel database. In particular, it contains UIC information for February 1991, the month the ground war started. This database was originally acquired from DMDC.
- The unit movement (“Locations”) database contains approximately 800,000 locations for units from all services that deployed to the KTO during ODS/DS. Gulf War unit locations were derived from a labor-intensive review of a wide variety and large number of ODS/DS unit records and through personal interviews with Gulf War Operations Officers (G3/S3) and Commanders. This review began in mid-1994 and the database was continuously updated as locations were obtained. USASCURR gathers unit location information from multiple sources including unit history data archives, operational logs, situation reports, after-action reports, and historical files. Though not complete, it was the most accurate record of unit locations for the war.

### **Data File Uses and Limitations**

The Personnel and Monthly databases provide detailed information about individuals, including demographic, professional, and unit assignment information. When merged, a combined record includes the UICs to which an individual was assigned over the year of interest. As might be expected, the Personnel database was imperfect. For example, OSAGWI has related anecdotes of listed personnel who did not participate in ODS/DS and, conversely, personnel who were in ODS/DS who were not listed in the database. Unfortunately, there was no measure of either type of inaccuracy. The inclusion of personnel who were not in ODS/DS was accounted for by drawing an initial sample of sufficient size to allow for respondents not meeting the survey entry requirements.

The Locations database links UICs to unit locations, which provides a way to infer the geographic location of personnel. Latitude and longitude coordinates from the Locations database were entered into geographic information system (GIS) software from which it was possible to select the UICs for units in particu-

lar geographic locations at specific times.<sup>1</sup> From this, for those personnel with matching UICs for the same time period, indicator variables were attached to the Personnel database to identify subpopulations of interest. In particular, as is discussed below, personnel with UICs that correspond to units located in urban areas on the day before the ground war started—February 23, 1991 (Julian date<sup>2</sup> 91054)—were identified.

There were some difficulties with this approach. First, approximately 100,000 non-Navy personnel records have UICs that were not in the Locations database; these personnel could not be placed in a specific geographic location. Second, the UIC data in the Personnel database was not completely accurate. OSAGWI personnel related that they had found that some UICs did not reflect the unit that personnel actually served with during ODS/DS. For example, the Air Force tended to assign personnel via temporary duty assignment (TDY) and the Personnel database records reflect either their originating unit or provisional unit, not necessarily their unit (or location) in the Gulf. Third, the location of a unit was only an indication of where the individual was likely to have been; even personnel correctly assigned to a unit may have had duties that placed them away from the unit's recorded location. Further, the Locations database itself has inaccuracies: It did not have data for all units, and for many units it did not have locations for all time periods of interest. Finally, almost 6,000 personnel records do not have any UICs. Because of these factors and more, it was not possible to design a sampling strategy strictly around personnel location. However, as is discussed below, location was used to select subpopulations that should have a high percentage of respondents in a particular type of location.

### Details of Merging the Data Files

The sampling frame was constructed first by merging the Personnel database and the Monthly database by service and then by selecting the members of the sampling frame.

**The Monthly Database.** The Monthly database, as received from USASCURR, consisted of 696,643 records, of which 113 were duplicate records by name and social security number (SSN). The duplicate records were eliminated by retaining the records that contained the largest number of UICs. This resolved to keeping 75 Army Guard records instead of Army Reserve records; keeping 25 Air Force Active records instead of Air Force Reserve records; keeping the record

---

<sup>1</sup>The GIS locations used in the sampling plan were derived by the Center for Health Promotion and Preventive Medicine (CHPPM), Edgewood Arsenal, Maryland.

<sup>2</sup>A Julian date consists of five digits; the first two digits indicate the year and the last three digits indicate the day of the year, sequentially numbered starting at one on January 1st. Thus, 91054 is the 54th day of 1991—February 23, 1991.

with the largest number of UICs for 11 other personnel who had duplicate records between various services and service components; and keeping the Army Guard records over Army Reserve for two personnel who had the same number of UICs in both records. Flags were added to these records so that if one of the 113 was chosen for the survey sample the alternative record could also be retrieved for interviewing purposes. Resolution of the duplicates resulted in a final database of 696,530 records.

**The Personnel Database.** The Personnel database was received in four sections from USASCURR, one for the Navy, one for the Marine Corps, one for the Army, and one for the Air Force and Coast Guard combined. These database sections were individually checked for duplicate records, merged with the Monthly database, subset to remove personnel not in the sampling frame, and then assembled into the final sampling frame of 536,790 personnel. For each service, this process is described separately and in detail below.

*Army.* The Army Personnel database was received in two sections, one for the Active forces and one for the Guard and Reserves. These were combined to total 357,879 records, of which 69 had duplicate names and SSNs (i.e., 69 personnel had two records each in the combined database). Removal of the duplicates resulted in 357,810 personnel records. These were then merged by SSN with the Army Monthly database of 351,297 records resulting in a combined database in which 79 records did not have UIC data and 38 records did not have personnel data. The majority of the duplicate Personnel records (65) corresponded to the duplicates identified in the Monthly database, and most of them (66) were between the Guard and the Reserves. In both the Monthly and Personnel databases the Guard records were in general much more complete and were kept; thus the choice of kept records in both databases was very consistent. The combined database was then further subset so that the sampling frame included only personnel with in and out dates (the dates they arrived in and then left the Persian Gulf region) overlapping with the survey period (August 1, 1990, to July 31, 1991). Personnel with missing in or out dates, which made it impossible to tell where they were during the survey period, were kept in the sampling frame. This resulted in a final set of 349,622 Army records.

*Marine Corps.* The Marine Corps Personnel database totaled 103,711 records, of which one person had duplicate entries. The duplication was resolved by choosing the record that matched the Monthly database by component, resulting in 103,710 personnel records. These were merged with the Monthly database for a combined set of 103,710 records, all of which had both personnel and UIC data. (The equivalent terminology in the Marine Corps for UIC is RUC, for Reporting Unit Code.) The data were then further subset so that the sampling frame included only personnel with in and out dates overlapping the survey period and who were not at sea in the Persian Gulf for the entire time. As

with the Army, personnel with missing in or out dates, which made it impossible to tell where they were during the survey period, were kept in the sampling frame. This resulted in a final set of 95,441 Marine Corps records for the sampling frame.

Marine Corps personnel were identified as being at sea as follows. First, the Office of the Special Assistant for Gulf War Illness (OSAGWI) and the Center for Health Promotion and Preventive Medicine (CHPPM) identified Marine Corps RUCs in the USASCURR Locations database with latitude and longitude coordinates that put them in the Persian Gulf during the entire time the units were in the Gulf region. Table B.2 lists the RUCs that were classified as Marine units at sea. Second, a Marine was classified at sea if all of their RUCs in the Monthly database between his in and out dates were contained in the list of at-sea RUCs. (Personnel with missing RUCs between their in and out dates were presumed to be ashore.)

*Air Force.* The Air Force Personnel database had a total 99,444 records, for which nine people had duplicate names and SSNs. These records were merged with the Monthly database of 82,676 records resulting in a combined set of

**Table B.2**  
**Marine Corps RUCs Identified as Being at Sea**

RUC	Unit Name
00044	MAG-40
00207	DET, MACG-28
00274	MWSS-274
00820	DET, MASS-1
00971	DET, MACG-38
00973	MACS-6
01296	MATCS-28
01331	VMA-331
12101	2D MARINES RLT
12110	1ST BN 2ND MAR
12110AC	C CO 1ST BN 2D MAR
12130	3D BN 2D MARINES
12130AW	WPNS CO 3D BN 2D MAR
20034	CMD ELE 4TH MEB
20034HS	HQ SVC CO 4TH MEB
20036HQ	HQ CO, 4TH MEB
20460	2D LAI BN
20460HQ	DET HQ SVC CO 2D LAI BN
21610	1ST ANGLICO 1ST SRI GROUP
21610DA	DET, 1ST ANGLICO, 13TH MEU
28390	MSSG-11, 11TH MEU
28390DM	DET, MEDICAL MSSG-11
28391DT	DET, 7TH MT BN

83,639 unique records, of which 973 records had personnel data but no UIC data and one person had UIC data but no personnel data. The data were then further subset so that the sampling frame included only personnel with location codes in the area of interest *and* whose in and out dates overlapped the survey period (or the in or out dates made it impossible to tell where they were during the survey period). Table B.3 provides the location codes for the areas included in the survey; personnel missing location codes were included by default. This resulted in a final set of 78,659 Air Force records.

*Navy.* The Navy Personnel database had 158,003 records, none of which were duplicates. When merged with the Monthly database of 158,000 records, three records were found for which no UIC data existed. Of these, only those personnel likely to have been ashore were selected for the sampling frame. This resulted in a final Navy population of 12,220 persons.

Ashore Navy personnel were identified by UIC as follows. First, UICs were identified as either being ashore or having a function that likely put the unit ashore by either:

**Table B.3**  
**Air Force Location Codes Defined to Be in the Survey Area**

Location Code	Location Name	Location Code	Location Name
AAVN	Abu Dhabi, UAE	MFFS	King Fahd IAP, SA
AAVS	Bateen, UAE	MMDL	Kuwait City, KU
ABFL	Ad Dammam PRT, SA	MMDN	Kuwait IAP, KU
ADKL	AL Kharj, SA	NZYR	Manama City, BA
ADKN	Buraymi West APT/Al Ain, UAE	PKVV	Masirah, Oman
ADL3	Al Dhafra, UAE	QEPQ	Thumrait, Oman
ADL5	Al Kharj APT, SA	QJGD	Minhad AB, UAE
ALBQ	Al Jouf OLD, SA	UGYJ	Riyadh New APT/KKIA, SA
ARMY	Use "ARMY UIC" For Locations	UGZX	Riyadh IAP, SA
ATXC	Bahrain City, BA	UGZY	Riyadh, SA
ATXK	Bahrain IAP, BA	UYNR	Sirsenk AB, Iraq
FFTJ	Dhahran IAP, SA	VGHV	Seeb IAP, Oman
FHLZ	Doha Intl, QA	VKWD	Shaikh ISA IAP, BA
FMAU	Dubai IAP, UAE	VLJD	Sharjah IAP, BA
KJAZ	KKMC, SA	WPPX	Tabuk/King Faisal, SA
LTWA	King Abdul Aziz AB, SA	WQLS	Taif, SA
LUTC	Jeddah, SA	XJAZ	Classified/unknown
LWEX	Jubail AFD, SA	XPQF	Classified/unknown
MEBG	Khamis Mushayt, SA	XQFT	Classified/unknown

- Identifying units in the USASCURR Locations database with latitude and longitude coordinates that put them ashore during the period of interest and which also were not obviously at-sea units (such as ships); or
- Identifying units by UIC name that had obvious ashore functions, such as Navy UICs attached to the Marine Corps, Navy construction battalions, and Seal units.

These units were then checked against the Monthly database to determine how many personnel were attached to those units during their time in the Gulf region, and units with abnormally low or zero personnel counts were removed.

A list of the Navy UICs that were classified as ashore units appears at the end of this appendix (Table B.8). Once the UICs were identified, Navy personnel were classified as ashore if one or more of the UICs that fell between their in and out dates were contained in the list of ashore UICs *and* either their service time in the Gulf overlapped with the survey period or missing in or out dates made it impossible to tell where they were during the survey period.

*Coast Guard.* Information on Coast Guard personnel was not available, so all Coast Guard records were included in the sampling frame.

**Assembling the Sampling Frame.** The sampling frame was assembled by merging the four services into one file. An indicator was attached to the records for 16,525 personnel (by SSN) who were identified by OSAGWI as potentially having lived or worked in an urban area. These personnel were identified in much the same way that the Marine at-sea population was created: First, OSAGWI and CHPPM identified units in the USASCURR database with latitude and longitude coordinates that put them in proximity (most within a five kilometer radius) of a primary city (as listed in the Living Conditions (by Geographic Location subsection) in the Kuwaiti Theater of Operations on February 23, 1991, the day before the start of the ground war. Then, personnel were classified as possibly having lived or worked in a built-up area if one or more of the UICs that fell between their in and out dates were contained in the list of built-up area UICs.

### **Stratification Variables**

The sample was chosen so that pesticide exposure estimates could be generalized to the whole in-theater ashore population of Army, Marine Corps, Air Force, and Navy personnel. Stratification was used to ensure that sufficient data were gathered on particular subpopulations to achieve a given precision in the exposure estimates. Thus, strata determination was largely an effort to define the important population characteristics that were underrepresented, us-

ing a simple random sampling scheme, and then specifying the precision desired for the estimates associated with each strata. Conversely, the decision not to stratify meant that sufficient data were gathered via random sampling of the rest of the population to analyze in the nonstratified dimensions.

Subpopulations of personnel useful for stratification were divided into two major categories: exposure-based and knowledge-based. In the former category, the population may be divided into groups that may have been exposed to different levels or types of pesticides; in the latter, personnel may be divided into groups that might have special knowledge of the use of pesticides. Exposure-based variables were principally related to branch (and component) of service, living conditions (primarily by branch of service and geographic location), geographic conditions, time of year, length of time in country, and some occupational specialties. Knowledge-based variables were primarily a function of occupational specialty and perhaps rank.

Each possible stratification variable is discussed in detail below, with a justification for why stratification was or was not necessary. In summary, the sample was drawn stratified by branch of service, food service military occupational specialty, senior enlisted personnel ranks (E-6 through E-9), and living conditions in “urban areas.” As mentioned above, the decision not to stratify meant that sufficient data were available to analyze those dimensions without oversampling.

**Branch and Component of Service.** All services were sampled: Army, Air Force, Marine Corps, Navy, and Coast Guard. As was shown in Table B.1, there were approximately 350,000, 79,000, and 95,000 Army, Air Force, and Marine Corps personnel on the ground in theater on 91054, respectively, which constituted 65 percent, 15 percent, and 18 percent of the personnel. Navy ashore constituted only a small part (roughly 2 percent) of the population and Coast Guard only 0.1 percent. For purposes of sampling, the Navy ashore and Coast Guard personnel were grouped with the Marine Corps. The Air Force and the Marine Corps/Navy were proportionally oversampled so that exposure estimates for the Army, the Air Force, and the Marine Corps had approximately the same precision.

Reservists and Guard members constituted approximately 24 percent of the Army, 14 percent of the Air Force, and 12 percent of the Marine Corps. However, since Reservists were integrated and operating as members of the Active force, their numbers were included in the total service counts above and service component was not considered as a stratification variable.

**Living Conditions by Branch of Service.** Since the Army, Air Force, and Marine Corps/Navy/Coast Guard were sampled separately for the reasons discussed above, this stratification requirement was already met. Note that anecdotal

evidence indicates that the living conditions for the Army and Marine Corps units were similar, so that sampling the two services separately to analyze their common living conditions may be uninformative. The Air Force, however, used distinctly different tents and was located largely at air bases, which warrants a separate analysis. In any case, stratification by service allows for separate evaluation of service-related living conditions.

**Living Conditions by Geographic Location.** It was hypothesized that there were four distinct types of living and working locations: (1) urban areas; (2) air bases; (3) permanent, relatively nonmobile “tent cities”; and (4) everything else (i.e., the rest of the troops in the desert). They were characterized as follows:

1. **Urban areas:** Cities and other urban areas that existed before the war and consisted of permanent structures such as apartment buildings and other converted buildings that U.S. military personnel lived in. The essential descriptors are: (a) urban and (b) existing permanent structures. Examples include personnel who lived or worked in buildings in Al Jubayl, Bahrain, Ad Dammam, Dhahran, Abu Dhabi, Hafar Al Batin, Khobar, and Riyadh.
2. **Air bases:** Air bases that primarily served the Air Force. They were located in the United Arab Emirates (UAE) and Oman, as well as Saudi Arabia and Bahrain. These air bases may have been similar to the tent cities discussed below, but they also may have had unknown differences, particularly for facilities located outside of Saudi Arabia. Examples include King Fahd (IAP, SA) airfield, Bateen in the UAE, and Seeb IAP in Oman.
3. **Relatively permanent, nonmobile “tent city” installations:** These were specifically erected by the U.S. military or U.S. and local contractors to house U.S. troops for the war effort. They generally consisted of tents and mobile trailers. These tent cities were distinguished from those erected by front-line troops in that they were relatively stationary for the entire war and tended to be built up and improved upon over the course of their occupation. The essential descriptors here are: (a) not previously existing and (b) relatively permanently located for the duration of the war (i.e., they were not intended to move, as would a front-line encampment). Examples include the logistics bases.
4. **Everything else:** These were intended to consist primarily of field accommodations for the troops in the desert. Although some may have looked similar to some of the tent cities, they would have been less permanent and may have been moved more often. In all likelihood, they were less densely populated than the tent cities.

Preliminary analysis (based on examination of Army personnel) showed that approximately 40 percent of in-theater personnel lived in relatively permanent,

nonmoving tent cities (category 3, above). For this reason, this category did not require oversampling. Similarly, most of the remaining personnel lived in field accommodations, category 4, and so this category also did not require oversampling. And, the majority of Air Force personnel lived or worked at air bases, so category 2 was evaluated as part of the stratification on Air Force as a service.

Of the four categories, only the number of personnel who lived/worked in buildings in urban areas was small enough to require oversampling. This was achieved by identifying units in the Locations database that were near urban areas and local cities on 91054, the day before the ground war started. The theory behind this approach was that on the day before the start of the ground war, units were highly likely to be in their functional locations. Thus, a unit located in the middle of a city on 91054 was assumed to have a long-term function there, so that the personnel assigned to that unit should be more likely to have either lived or worked in buildings in the city. Of course, proximity of a unit to an urban area does not guarantee that the unit's personnel lived or worked in buildings, but it was hoped that oversampling this population would provide a sufficiently large cohort of personnel who did.

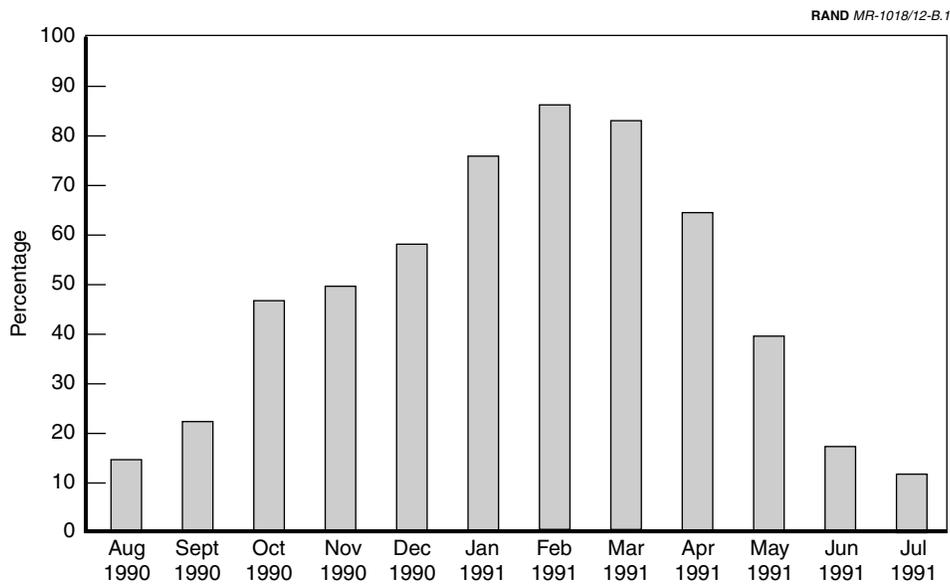
The following areas were identified as urban: Al Jubayl, Bahrain, Ad Dammam, Dhahran, Abu Dhabi, Dubai, Hafar Al Batin, Khobar, and Riyadh. Approximately 26,000 personnel were linked to units located in or near urban areas on 91054. Of this total, approximately 16,100 were Army personnel, 9,700 Air Force, and 400 Marine Corps. Each person was assigned an "urban" indicator, and the Army and Air Force were oversampled to gather a sufficiently large sample living or working in urban areas.

**Geographic Conditions.** It has been hypothesized that conditions differed between the inland desert and coastal areas, and between the desert and conditions around the Euphrates and Tigris Rivers in Iraq. Preliminary interviews with personnel who served in the War have not uncovered any indication of significant differences between the coastal and inland areas in terms of pests. In addition, troops were exposed to the river areas in Iraq for only a short time. Thus, neither hypothesis was used for stratification.

**Length of Time in Theater.** Eighty-seven percent of active duty personnel spent more than 60 days in country according to the data; more than 37 percent spent over 180 days in theater. The date data show evidence of rounding to the first of the month or the 31st of the month, but in spite of this, the distribution of length of times in theater was well distributed across the whole range of times. Almost 90 percent of reservists spent more than 60 days in country, but only slightly more than 16 percent spent over 180 days in theater. Thus it was not necessary to stratify on time in theater, as simple random sampling resulted in a robust selection of times with very few times in theater of less than 30 days.

**Time of Year.** It has been hypothesized that pesticide usage may have varied with the seasons, most notably because of a possible increase in pests during periods of warm weather. Figure B.1 shows the percentage of the sampling frame personnel that were in theater on a particular month. Forty-seven percent of active duty personnel were in theater in October 1990, although half of them arrived that month. The remainder arrived after October, with a peak of almost 17 percent in January 1991, and with much smaller percentages in the other months. Eighty percent of the reservists, on the other hand, arrived between November 1990 and February 1991. Under the assumption that warm weather in theater extends from May through October,<sup>3</sup> over half of the population were in theater either near the end of the 1990 summer or near the beginning of the 1991 summer.

Respondents were asked to report their locations in a randomly chosen month that they were in theater. Given that half of active duty respondents were in theater for between two and six months and half (not necessarily the same personnel) arrived in theater during or before October 1990, a random sample



**Figure B.1—Percentage of the Sampling Frame Personnel in Theater by Month**

<sup>3</sup>“Saudi weather was among the most inhospitable in the world, the temperatures in August and September sometimes reaching 140 degrees Fahrenheit. . . . Between November and March, temperatures moderate considerably” (Scales, p. 121).

asked about a random time was expected to yield a large cohort of respondents who experienced warm weather. Thus, it was unnecessary to stratify on this condition.

**Military Occupational Specialties.** Food service occupational specialties were hypothesized to have special knowledge of pesticide usage in mess halls (dining facilities). The Army's food service occupational specialties constituted approximately 2.7 percent of the force in theater, the Air Force's approximately 1.7 percent, and the Marine Corps's approximately 1.8 percent. Table B.4 lists the food service occupational specialty codes that were oversampled.

Supply occupational specialties were also identified as having special pesticide knowledge. However, survey time constraints did not permit inclusion of additional questions related to supply and food service specialties, so these were not oversampled. Military police was a third occupational category considered for oversampling because of possible exposure to delousing chemicals used with enemy prisoners of war. But delousing procedures were investigated by OSAGWI, so oversampling in this survey was not conducted.

**Rank.** It has been suggested that senior enlisted personnel were likely to have a broader knowledge of how pesticides were used and might be able to provide additional useful information. The majority of the personnel in the gulf were junior enlisted: 70.8 percent were E-5 and below, 17.7 percent were E-6 to E-9, and the officer and warrant officer corps constituted the remaining 11.2 percent (with 0.3 percent of the records missing rank information). Senior enlisted personnel, defined as E-6 to E-9, were therefore oversampled.

**Table B.4**

**Food Service Military Occupational Specialties to Be Oversampled**

Service	Occupational Specialty Code	Description
Army	91M	Hospital food service specialist
Army	94B	Food service specialist
Air Force	623**	Subsistence operations specialist
Marine Corps	3381	Food service specialist
Marine Corps	3311	Baker
Navy	MS**	Mess management specialist

\* indicates that any alphanumeric character was allowable in these positions.

### Summary of the Stratification Variables

In summary, the sample was stratified by branch of service, food service occupational specialty, senior enlisted personnel, and urban areas. Table B.5 tabulates the final sampling frame by service and by the three variables that were

**Table B.5**  
**Tabulations of the Sampling Frame by Service and**  
**Categories to Be Oversampled**

Service	Food Service	Senior Enlisted	Urban Area	Count	Total
Army	0	0	0	265,557	349,622
	0	0	1	12,404	
	0	1	0	58,733	
	0	1	1	3,240	
	1	0	0	6,828	
	1	0	1	334	
	1	1	0	2,388	
	1	1	1	138	
Marine Corps	0	0	0	82,847	95,441
	0	0	1	291	
	0	1	0	10,646	
	0	1	1	79	
	1	0	0	1,358	
	1	0	1	0	
	1	1	0	220	
Navy	0	0	0	9,604	12,220
	0	1	0	2,442	
	1	0	0	129	
	1	1	0	45	
Air Force	0	0	0	53,149	78,659
	0	0	1	7,640	
	0	1	0	14,418	
	0	1	1	2,019	
	1	0	0	1,092	
	1	0	1	110	
	1	1	0	203	
Coast Guard	1	1	1	28	848
				848	
Total	1 = 12,873	1 = 94,599	1 = 16,486	536,790	536,790

NOTE: Food service personnel occupational specialty codes were described in Table B.4, and senior enlisted personnel were defined as E-6 to E-9.

oversampled. A "1" in a column means that personnel meeting the description of the column title were included in the "count" total. For example, the last line for the Army shows that out of 349,622 Army personnel in the sampling frame, there were 138 senior enlisted food service personnel attached to units identified in urban areas.

As Table B.5 shows, there were insufficient personnel to completely cross the strata into all 24 possible combinations. (For example, there were no senior en-

listed food service Marines in urban areas.) So the oversampling categories were simplified as follows.

- Only Army and Air Force personnel in urban areas were oversampled. (Marine Corps/Navy had very few identifiable personnel in the selected urban areas, and they were restricted to only two Marine Corps units.)
- Food service occupational specialties were oversampled among the non-urban population.
- Senior enlisted personnel were oversampled among the nonurban, non-food service occupational specialty population.

This reduced the number of strata to 11, including three “all other” strata for the Army, Air Force, and Marine Corps/Navy/Coast Guard. As the next section will discuss, such a reduction was necessary to achieve an acceptable estimation precision within a reasonable sample size.

### **DEFINING THE SAMPLE: SIZE AND ADJUSTMENT**

The sample size calculations were based on a dichotomous question. The sample was sized so that the width of 95 percent confidence intervals for the percentage of personnel using a pesticide across all services was at most plus or minus 3, and for each individual service, it was at most plus or minus 4. The sample was also divided so that the confidence interval widths by service were approximately equal and the confidence interval widths by strata, particularly the rare strata, were minimized as much as possible within the constraints of the precision of the overall and service estimates.

Equal confidence intervals among the services (Army, Air Force, and Marine Corps/Navy/Coast Guard) were necessary under the assumption that it was desirable to report final results for the services with equal precision. Strata were limited, as described in the last section, to preserve precision in the overall sample estimates.

Assuming that 50 percent of personnel used a pesticide,<sup>4</sup> these requirements dictated a final sample of 2,000 people with 667 respondents for the Army and Air Force, and 666 respondents from the Marine Corps/Navy ashore/Coast Guard. Table B.6 shows how the total sample was divided for oversampling and

---

<sup>4</sup>Estimated confidence interval width was a function of the percentage, and 50 percent provides the “worst case” scenario; that is, it gives the widest confidence interval. Should the percentage vary from 50, then the smaller confidence intervals will result. Also, note that these calculations do not use a finite population correction, because the sampling fraction was kept to less than 5 percent of the sample frame population, both overall and within each strata (Cochran, 1997, p. 25).

**Table B.6**  
**The Oversampling Scheme**

Stratum	Number (%) in Sampling Frame	Number (%) in Sample	Width of 95% Confidence Interval Under SRS (%)	Width of 95% Confidence Interval with Oversampling (%)
<b>Army</b>	349,622(65.2)	66(33.3)	6	9
Urban area	16,116(4.6)	133(20.0)	26	17
Food service <sup>a</sup>	9,216(2.6)	67(10.0)	34	24
Senior enlisted <sup>b</sup>	58,733 (16.8)	167(25.0)	14	16
All else	265,557(76.0)	300(45.0)	6	12
<b>Air Force</b>	78,659(14.7)	667(33.3)	12	8
Urban area	9,797(12.5)	133(20.0)	33	17
Food service <sup>c</sup>	1,295(1.6)	67(10.0)	90	24
Senior enlisted <sup>d</sup>	14,418(18.3)	167(25.0)	27	16
All else	53,149(67.6)	300(45.0)	14	12
<b>Marine Corps/ Navy/Coast Guard</b>	108,509(20.1)	666(33.3)	10	8
Food service	1,752 (1.6)	67(10.0)	78	24
Senior enlisted <sup>e</sup>	13,167(12.2)	167(25.0)	28	16
All else	93,590(86.2)	432(65.0)	11	10
All services	536,790(100)	2,000(100)	4.4	6.4

<sup>a</sup>Personnel with food service occupations among nonurban.

<sup>b</sup>Senior enlisted personnel (E-6 to E-9) among nonurban, non-food service personnel.

<sup>c</sup>Personnel with food service AFSC among nonurban.

<sup>d</sup>Senior enlisted personnel (E-6 to E-9) among nonurban, non-food service AFSC personnel.

<sup>e</sup>Senior enlisted personnel (E-6 to E-9) among non-food service personnel.

gives the planned confidence interval widths by service and for the individual strata. The second column ("Number and % in Sampling Frame") provides the fraction of personnel by service in the sampling frame and a breakdown within each service by strata. Comparison of the percentages in this column with those in the third column ("Number and % in Sample") demonstrates the areas and degrees of oversampling; for example, the Air Force made up 14.7 percent of the sampling frame but was oversampled to constitute 33.3 percent of the sample. The fourth column gives the expected width of the confidence intervals by service and strata if simple random sampling (SRS) was employed; the last column shows the confidence interval widths using oversampling based on the sample sizes specified in the third column.<sup>5</sup> The table shows large gains in precision for the smaller strata which comes at the expense of: (a) some of the

<sup>5</sup>The oversampling confidence intervals were calculated using variances appropriately adjusted using the weights that would result from oversampling. See Cochran (1997, Chapter 5) for calculation details.

Army estimates and (b) increasing the aggregate confidence interval width across all the services to 3.2 percent from 2.2 percent under SRS.

### Sample Size Corrections

The calculations used to estimate the required sample sizes do not consider the various types of errors that must be accounted for when selecting the initial sample. That is, the initial sample must be inflated to allow for respondent nonresponse, the inability to locate some veterans, and general errors in the database (such as service members who were listed in the database but who were never in ODS/DS or the region of interest).

**Nonresponse.** An 85 percent response rate was assumed. Decoufle et al. (1991) reported a 92 percent response rate in a similar survey of Vietnam veterans. Of those respondents contacted who were in ODS/DS, the actual nonresponse rate—meaning that the potential respondent refused to participate in the survey—was only 3 percent.

**Unlocatability.** An 85 percent location rate was assumed. The actual unlocatability rate was 23 percent. Decoufle et al. (1991) reported a 93 percent location rate in their survey of Vietnam veterans, however, that survey used additional locating methods, such as IRS address records, which were not available to us.

**Database Errors.** A 15 percent overall database error rate was assumed to account for many possible types of errors, including personnel who did not participate in ODS/DS or whose location was misclassified so that they were not in the region of interest; and coding, administrative, and other types of records errors. The actual error rate was 7 percent.

As shown in Table B.7, incorporation of the nonresponse, unlocatability, and database error factors into the original sample size gives adjusted sample sizes of 1,088 for the Army, the Air Force, and Marine Corps/Navy/Coast Guard. This resulted in an initial combined sample of 3,264, which was drawn by strata according to the numbers listed in the last column of Table B.7.

**Table B.7**  
**The Initial Sample Size Defined to Achieve**  
**2,000 Complete Final Responses**

Stratum	Desired Num- ber in Final Sample	Initial Sample to Be Selected
<b>Army</b>		
Urban area	133	218
Food Service	67	109
Senior enlisted	167	272
All else	300	489
Total Army	667	1,088
<b>Air Force</b>		
Urban area	133	218
Food Service	67	109
Senior enlisted	167	272
All else	300	489
Total Air Force	667	1,088
<b>MarineCorps/ Navy/Coast Guard</b>		
Food Service	67	109
Senior enlisted	167	272
All else	432	707
Total Marines/Navy/C.G.	666	1,088
Total	2,000	3,264

Table B.8

## Navy Units That Were Classified as Ashore Units

UIC	Unit Name	UIC	Unit Name
N57100	NAV SPEC WARFARE GRU 1	N55103	MOBILE CONST BATT 3
N0031A	NAV SPEC WARFARE GRU 2	N55114	MOBILE CONST BATT 4
N55777	SEAL TEAM 1	N55115	MOBILE CONST BATT 5
N55778	SEAL TEAM 2	N55117	MOBILE CONST BATT 7
N44884	SEAL TEAM 3	N08864	MOBILE CONST BATT 24
N08943	SEAL TEAM 4	N55448	MOBILE CONST BATT 40
N08971	SEAL TEAM 5	N55488	MOBILE CONST BATT 74
N46985	SEAL TEAM 8	N55451	MOBILE CONST BATT 133
N55205	CG I MEF	N55163	CONSTRUCTION BATT UNIT 421
N55207	CG II MEF	N81123	NR CARGO HD BN 3
N55211	CG III MEF	N81124	NR CARGO HD BN 4
N67448	1ST MAR DIV FMF PAC	N82218	NR CARGO HD BN 13
N08321	2ND MAR DIV FMF LANT	N81464	RESERVE CARGO HAND FORCE S
N67360	3RD MAR DIV FMF PAC	N35010	T-AO 107 PASSUMPSIC MILDPT
N67339	CG FIRST MEB	N44291	T-AFS 9 SPICA MILDEPT
N55206	CG FIFTH MEB	N47842	ACADIA REPAIR CO
N55208	CG SIXTH MEB	N68684	FLT HOSP 500 BED CBTZ-4
N55356	CG SEVENTH MEB	N68685	FLT HOSP 500 BED CBTZ-5
N46616	CG FIRST MEB DET NMC HAWAI	N68686	FLT HOSP 500 BED CBTZ-6
N67446	1ST FSSG FMFPAC	N45399	FLT HOSP 500 BED CBTZ-15
N46621	1ST FSSG DET NH CAMP PENDL	N42221	SPECBOATU 12
N68408	2D FSSG FMF LANT	N42223	SPECBOATU 20
N46614	2D FSSG DET NH CAMP LEJEUN	N42224	SPECBOATU 24
N47438	2D MARDIV DET NAVHOSP LEJE	N44394	SPECBOATU 24 SEA DUTY
N46612	2ND MARDIV DET NH BETHESDA	N53210	ASSAULT CRAFT UNIT 2
N67436	3D FSSG FMFPAC	N42056	ASSAULT CRAFT UNIT 2 SHORE
N67683	4TH MARDIV 3RD ANGLICO	N47106	ASSAULT CRAFT UNIT 4 SHORE
N67803	4TH FSSG MEDLOGCO 4TH SUP	N46587	ASSAULT CRAFT UNIT 5 SHORE
N42320	CBAT SVC SUPP DET (CSSD) 1	N67408	1ST RADIO BN FMFPAC
N47114	CBAT SVC SUPP DET (CSSD) 1	N08973	SDVT 1
N41638	CBAT SVC SUPP DET (CSSD) 1	N45597	USCINCCENT SPACT RIYADH SA
N41629	CBAT SVC SUPP DET (CSSD) 2	N79109	USCINCCENT
N53212	BEACHMASTER UNIT 1	N81383	3RD RNCR
N44920	BEACHMASTER UNIT 1 DET A	N45454	ATTACHE OMAN
N44921	BEACHMASTER UNIT 1 DET B	N44349	NAVY IPO DET JEDDAH
N44922	BEACHMASTER UNIT 1 DET C	N79087	NAVY IPO DET JUBAIL
N44923	BEACHMASTER UNIT 1 DET D	N44350	NAVY IPO REP RIYADH
N44924	BEACHMASTER UNIT 1 DET E	N44691	NAVY IPO DET DHAHRAN
N44925	BEACHMASTER UNIT 1 DET F	N46026	NAVSEASYSKOMDET RSNF JUBAY
N41914	BEACHMASTER UNIT 1 SHORE D	N08991	VR 51
N53211	BEACHMASTER UNIT 2	N09014	VR 24
N42055	BEACHMASTER UNIT 2 SHORE C	N09031	HS 75
N66647	CONSTRUCTION BATT UNIT 408	N09043	VP 23
N66649	CONSTRUCTION BATT UNIT 405	N09179	VP 19
N66629	CONSTRUCTION BATT UNIT 407	N09244	VPU 2
N66676	CONSTRUCTION BATT UNIT 411	N09305	VP-91
N66923	CONSTRUCTION BATT UNIT 415	N09362	VP 48
N68571	CONSTRUCTION BATT UNIT 418	N09367	VP 11
N68680	CONSTRUCTION BATT UNIT 419	N09618	VP 1
N55101	MOBILE CONST BATT 1	N09619	VP 49

**Table B.8 (continued)**

UIC	Unit Name	UIC	Unit Name
N09623	VP 4	N09946	VQ 2
N09630	VP 5	N09962	VQ 4
N09632	VP 46	N30197	VC 6 DET DAM NECK
N09661	VP 8	N53855	VR 55
N09665	VP 45	N53869	VPU 1
N09674	VP 40	N53910	VR 57
N09804	VC 5	N53921	VR 59
N09806	VC 6	N53811	HCS 4
N09930	VQ 1	N53812	HELO LIGHT ATTACK SQ 5