5. STATISTICAL ISSUES

This section discusses a number of topics related to statistical issues and data analysis. These topics include methods used to resolve problems that arose when confirming CLIA ID numbers; weighting and nonresponse adjustments; and survey results.

5.1 Phase I

5.1.1 CLIA ID Number Problem Resolution

Strictly speaking, the NICLTS sample frame consisted of CLIA ID numbers listed on HCFA's July 1996 OSCAR database as certified for performing moderate, high, waived, or Provider Performed Microscopy (PPM) testing. Every attempt was made to confirm each laboratory's CLIA ID number both at the point of enrollment and during the field tabulation. After cataloging the types of CLIA ID number problems discovered during the startup period, NICLTS clinical and statistical managers prepared a memo that explained how the field tabulators should handle them. Possibilities included the following:

1. **The CLIA ID number of the laboratory did not match the sampled CLIA ID number or could not be found.** In this case, the tabulator asked questions to establish a possible explanation for this situation. If the number did not match, the tabulator recorded the laboratory's CLIA ID number and collected data under some circumstances. The memo specified the circumstances when tabulation proceeded and when it did not.

2. **More than one CLIA ID number was found at the location.** The tabulator asked if separate data were available for testing performed under the sampled CLIA ID number, and, if so, tabulated them. Otherwise, all the data were tabulated and the statistician was alerted to this fact.

3. **More than one location was included under the same CLIA ID number** (other than multiple locations identified during enrollment). The tabulator asked for the names and addresses of all the locations. All locations were tabulated if the number of additional locations was three or less. If four or more locations were identified, the case was pulled for further consideration.

A number of cases fell into each of the three categories and most were dealt with easily by following the protocol in the memo. The last possibility proved to be the most complicated and the most interesting. One CLIA

ID number was associated with 80 separate locations at some distance from each other and would have required days of traveling and tabulating to complete. The laboratory's central office determined by phone which analytes were performed, the biological specimens used, and the annual volumes for the year of interest but could not determine the names of the test system manufacturers. Test systems differed from location to location. The problem was how to tabulate efficiently and accurately. The solution was to collect summary data on analytes on site at the central office and then to collect detailed data on analytes, test systems, and volumes by telephone from a representative sample of locations.

Several other CLIA ID numbers were also associated with multiple locations identified during enrollment or after the field tabulator established contact with the laboratory. The central office generally confirmed that test menus were homogeneous. This allowed Westat to select a sample of laboratories for visiting and tabulating and to visit all nonhomogeneous locations. If all locations under a CLIA ID number could be tabulated within 2 days, they were all tabulated on site.

5.1.2 Weighting and Nonresponse Adjustments

This section describes the steps taken for weighting and related postprocessing of the NICLTS data. The discussion covers each step of the calculation of the sampling weights, including construction of the base weights, nonresponse adjustments, raking, and adjustments for the various types of subsampling used within laboratories. These steps apply to both Phase I and Phase II of the study. Phase II adjustments were, however, somewhat simpler since there was no daily log or nursing station subsampling.

To clarify the discussion, some definitions are in order. For purposes of the NICLTS, a **laboratory** was defined as the unit corresponding to a specific CLIA ID number. Thus, laboratories (i.e., CLIA ID numbers) were the sampled units and the units of analysis. To facilitate data tabulation, further operational subdivisions were created. A laboratory could have multiple geographic **locations**--for example, when a health clinic had several clinics distributed across the state. Each location was sent to the field for a separate tabulation (although multiple locations of a single laboratory were handled operationally by the same tabulator). A location, in turn, could have multiple **sites** within it--for example, a nursing home with several nursing stations.

Review of Tabulation Problem Sheets

The Tabulation Problem Sheets from the field were reviewed to see if any of the situations described affected the weighting--in particular, cases involving multiple locations, multiple or mismatching CLIA ID numbers, and the use of billing or partial year data.

Assign Response Codes to Laboratories

Westat reviewed the specific operational result codes for every laboratory location to assign an overall response code for the laboratory. The result codes for locations were reduced to four response codes for weighting purposes: respondent, eligible nonrespondent, ineligible, and nonrespondent with unknown eligibility. For laboratories with multiple locations, it was necessary to assign a single response code to characterize the final response outcome at the level of the sampled CLIA ED number.

Table 5-1 shows how overall response codes were assigned to CLIA ID numbers whose laboratory locations had different combinations of operational result codes. The response codes were respondent (R), eligible nonrespondent (NR), nonrespondent with unknown eligibility (NU; this group included laboratories which could not be located) and ineligible (I; this group included laboratories that were out of business). No combinations other than those shown in the table occurred.

Table 5-1. Overall response codes and CLIA ID numbers

Laboratories with location operational result codes that were:	Had overall response codes of:
All R	R
All NR	NR
All NU	NU
All I	I
Any combination of R and NR	R
Any combination of R and NU	R
Any combination of R and I	R
Any combination of NR and I	NR
Any combination of NU and I	NU

Calculate Laboratory Base Weights

In general, the laboratory (i.e., CLIA ID number) base weight was simply the inverse of the probability of selection. The laboratory base weight formula accounts for any different probabilities of selection and was calculated as:

$$\text{Base weight} = \frac{1}{p_1 x p_2 + (1 - p_1 x p_2) x p_3}$$

where

p_1 = probability of selecting a laboratory for the initial sample;

p_2 = conditional probability of selecting a laboratory for the primary sample; and

p_3 = conditional probability of selecting a laboratory from the reserve sample.

Calculate Final Laboratory Weights

The final laboratory weight was the product of the base weight, nonresponse adjustments, and raking adjustments. Westat did three nonresponse adjustments: one to account for nonrespondents of unknown eligibility at the enrollment stage, one to account for eligible nonrespondents at the enrollment stage, and one to account for all nonresponse at the field data collection stage.

Nonresponse adjustments are made by computing an adjustment factor:

$$\text{Adjustment factor} = \frac{\sum_{\text{Full sample}} \text{Base weights}}{\sum_{\text{Responding laboratories}} \text{Base weights}}$$

This adjustment factor may be calculated over the entire sample or within various subgroups. For the NICLTS, nonresponse adjustment factors were computed separately for each phase, using region and laboratory type as subgroups. Some laboratory types were combined where the sample sizes were small.

Once the adjustment factors were computed, the nonresponse-adjusted weights were computed as

Adjusted weight = (Adjustment factor) x (Base weight).

Raking is a procedure where survey estimates are adjusted to match certain known values. In the NICLTS, after performing the nonresponse adjustments, Westat raked the resulting adjusted weights to counts of laboratories by the six levels of laboratory group (see Table 2-3) and ten levels of region (see Table 2.2). These counts were obtained from the original sampling frame (i.e., the July 1996 OSCAR database). The raking adjustments and final laboratory weights were calculated using Westat's WESWGT computer software.

Creating Total Volumes of Distinct Clusters within each Laboratory

In creating total volumes of distinct clusters within each laboratory, Westat adjusted for the following factors:[10]

- Subsampling of daily logs and nursing stations;
- Nonresponse among sampled logs and nursing stations; and
- Subsampling of locations.

Each of these adjustments is discussed separately.

Accounting for Daily Log Subsampling and Nonresponse

Dally logs were subsampled separately by site. For the sites that subsampled daily logs, Westat applied a factor to the volume amounts so that they represented the volume at the site as if subsampling and nonresponse had not occurred. Westat calculated the daily log adjustment factor as:

[10] Subsampling of daily logs, nursing stations, and locations occurred only during Phase I. Conceptually, the corresponding adjustment factors during Phase II are all "1."

$$\text{Adjustment factor} = \begin{cases} \dfrac{\text{days site was open}}{\text{days with data}} & \text{for sites that subsampled daily logs} \\ 1 & \text{for sites that did not subsample daily logs} \end{cases}$$

The adjusted volume amount was calculated by first subtracting the QC amount[11] multiplying by the daily log factor as follows:

$$\text{Volume adjustment} = (\text{Volume} - \text{QC}) \times (\text{Adjustment factor})$$

In a few cases, volumes were not collected for some of the subsampled days. This occurred in some laboratories because the selected days were holidays during which the laboratories were closed. These laboratories were treated as if the holidays were never selected. For instance, if a laboratory was open for 305 days during the year and one of the 20 subsampled days occurred on a holiday, then the adjustment factor was 305/19 instead of 305/20.

Data also could not be collected for some days in three laboratories. In one, data were not available for some analytes at various points during the year. This affected 40 of 300 potential data points. At the second laboratory, no daily log records were available from January 1, 1996 through May 24, 1996, and the contact person who might have provided estimates was out of the office during the tabulation visit. In the third laboratory, daily log records were available only from August 1, 1996 through December 31, 1996 for some analytes, and no one could provide estimates for the rest of the year.

For tests performed by nurses, daily log data were available for the full year. The total number of daily log data points not collected from all laboratories was 399, or 0.8 percent of the total of 47,699 data points selected for daily log tabulation. Table 5-2 summarizes the characteristics of the daily log nonresponses.

[11] QC is the nonpatient care volume included in the Volume value. See Section 3.1.1-3.

Table 5-2. Characteristics of daily log nonresponse

Location	Number of days with some data missing	Number of Distinct clusters	Number of clusters with some data unavailable	Expected number of data points (20 x number of distinct triples)	Number of data points unavailable	Percent of data points unavailable
23362-01	20	15	12	300	40	13.3
25599-01	12	35	28	700	335	47.9
34939-01	8	7	3	140	24	17.1

Accounting for Nursing Station Subsampling and Nonresponse

At some locations, homogeneous nursing stations were subsampled. To adjust the volumes collected to represent the volumes at all homogeneous nursing stations, Westat calculated the nursing station factor as:

$$\text{NURSEFC} = \begin{cases} \dfrac{\text{total nursing stations}}{\text{nursing stations subsampled}} & \text{for laboratories where nursing stations were subs} \\ 1 & \text{for laboratories where nursing stations were not sul} \end{cases}$$

The adjusted volume amounts were calculated by multiplying the amounts after adjusting for daily logs by the nursing station factor. Because all nursing stations at participating locations responded, no nonresponse adjustments were needed.

Accounting for the Subsampling of Locations

In some laboratories with many geographically dispersed locations, the locations were subsampled. Westat adjusted the volumes for these locations to represent the volumes as if locations had not been subsampled

by calculating a location subsampling factor; new volumes were calculated as:

$$\text{LOCFC} = \begin{cases} \dfrac{\text{locations at laboratory}}{\text{locations subsamples}} & \text{for laboratories where locations were subsampled} \\ 1 & \text{for laboratories where locations were not subsampled} \end{cases}$$

$$\text{Adjusted volume} = \text{LOCFC} \times \text{Volume}$$

5.1.3 Phase I Summary Case Result Reports

This section presents summary statistics and a discussion of the operational results for each laboratory sampled for Phase I. Summary results are presented separately for each Phase I operational stage: Telephone Enrollment, Field Tabulation, Telephone Verification, Validation Telephone Enrollment, and Validation Field Tabulation.

5.1.3.1 Phase I Telephone Enrollment Results[12]

The final outcomes of the Phase I Telephone Enrollment process for all 930 laboratories (i.e., CLIA ID numbers) released as active sample appear in Table 5-3. The categories appearing in this table are further defined as follows:

Enrolled - The laboratory agreed to participate in the study.

Ineligible - The facility was confirmed as not having performed any Complexity Model tests in calendar year 1996 under the sampled CLIA ID number.

Refusal - The laboratory refused to respond to the enrollment interview or, during the interview, explicitly refused to participate in the on-site tabulation.

Out of Business - Contact with someone associated with the laboratory or with a third-party source provided definitive information that the lab was no longer in business (ceased operations, physician

[12] The numbers of laboratories assigned to the various result categories in the tables appearing throughout this section may represent a slightly different distribution of result categories from that used for statistical weighting purposes. These tables present descriptive statistics of the outcomes of operational processes, specific to NICLTS. Weighting uses standard rules for assigning cases to weighting classes.

retired, bought out by/merged with facility having its own CLIA ID number)

Nonlocatable - Unable to locate the laboratory through any available telephone numbers or telephone tracing efforts (e.g., Directory Assistance calls, lookups in telephone databases, AHA yearbook; forwarding address information provided by the U.S. Postal Service on advance notification letters that were returned undelivered because the laboratory had relocated and the forwarding order had expired.).

Other - Enrollment process concluded for any other reason, such as inability to identify or achieve contact with a respondent after repeated attempts, respondent not available during telephone enrollment period, laboratory not available for visit during the time period for field operations, and miscellaneous individual situations not otherwise categorizable.

Table 5-3 presents the absolute number of cases falling into these categories, as well as the percentage of the total sample represented by each category. Overall, 773 (83.1%) of the sample was enrolled for the Field Tabulation and 91 (9.8%) refused to participate at the Telephone Enrollment stage.

Table 5-3. Phase I telephone enrollment results

Result	Number of laboratories	Percent
Enrolled	773	83.1
Ineligible	13	1.4
Refusal	91	9.8
Out of Business	12	1.3
Nonlocatable	16	1.7
Other	25	2.7
Total	930	100.0

Response Rate

The response rate excludes from the calculations cases that were confirmed as not belonging to the population the sample is meant to represent. Put another way, it excludes cases that were on the sampling frame but, if perfect information had been available, would not have appeared on the frame. For NICLTS, the population covered is those laboratories that performed at least one test during calendar year 1996. Thus, laboratories identified during the NICLTS operations as being ineligible (no 1996 testing) were excluded from the response rate calculation.

Less clear-cut is the situation of laboratories that were confirmed as out of business. This group can be divided into two logical groups: those out of business prior to calendar year 1996 and those that went out of business any time on or after January 1, 1996. It is easy to see that the first group is ineligible for the study, since they could not have performed any tests in 1996. In contrast, it is possible, and even likely, that members of the latter group were part of the population covered by NICLTS, i.e., did perform eligible tests at some time during calendar year 1996. Practically speaking, however, it was generally impossible to determine during the Phase I data collection period (1997) to which group an out-of-business laboratory belonged. Moreover, even if an out-of-business laboratory's 1996 eligibility could be established after the fact, it was almost certainly impossible to collect useful, detailed data after the laboratory ceased operations.

Given these considerations, the response rate calculated for the NICLTS Phase I Telephone Enrollment excluded both ineligible and out-of-business laboratories from the denominator of the response rate calculation. This produced a response rate of 85.4 percent:

(773 enrolled) / (930 sampled - (13 ineligible + 12 out of business))

The complement of the response rate was the nonresponse rate, i.e., the percentage of the eligible sample population that did not provide a useful response that could be incorporated into the final estimates. For the Phase I Telephone Enrollment, the nonresponse rate was 14.6 percent.

Because it excludes entities that did not properly belong to the population of interest (ineligibles), the response rate is generally considered an indicator of potential bias that could be associated with nonresponse. In simplified terms, the more the 14.6 percent who did not respond differed from the 85.4 percent who did respond, the greater the chance that the final estimates could be biased because the 14.6 percent are not represented in the final sample. It should be emphasized that the issue is how different the estimates would be if the nonresponders had been 'included. As the response rate increases, there is a concomitant reduction in the size of the effect that the decreasing number of nonresponders could have on the estimates derived from the responders. Hence, all other things being equal, higher response rates are associated with more accurate estimates.

5.1.3.2 Phase I Field Tabulation Results

The final outcomes of the Phase I Field Tabulation process for the 827 enrolled laboratory locations appear in Table 5-4. The categories are nearly identical to those described in Section 5.1.3.1 for Telephone Enrollment, with two differences. First, Tabulated replaces Enrolled as the desired successful outcome of this stage of the process. A laboratory location was classified as "Tabulated" if and only if data were successfully tabulated at the laboratory. Second, the category Nonlocatable does not appear, since all laboratories that passed to this stage were, by definition, located at the Telephone Enrollment stage.[13]

Table 5-4. Phase I Field Tabulation results

Result	Number of laboratories	Percent
Tabulated*	757	91.5
Ineligible	29	3.5
Refusal	36	4.4
Out of Business	2	0.2
Other	3	0.4
Total	827	100.0

* Includes three satellite locations visited but tabulated on the single record of the central office location.

To knowledgeably interpret the Phase I Field Tabulation results, it is essential to remember that the cases represented in Table 5-4 are laboratory locations, not CLIA ID numbers. In situations where a CLIA ID number covered more than one geographical laboratory location, separate cases were created for field operational purposes. After the Telephone Enrollment stage, tracking and accounting for cases was always done at the location level. Examination of Tables 5-3 and 5-4 shows that the 773 CLIA ID numbers enrolled at the Phase I Telephone Enrollment stage became the 827 total laboratory locations sent to the field in the Phase I Field Tabulation stage. All results presented in Table 5-4 apply to the individual locations.

[13] In that context, another point must be mentioned. While the same logic should apply in regard to the determination of ineligible and out-of-business laboratories during telephone enrollment, in reality the on-site tabulator could investigate these issues in greater depth. This could result in recategorizing of laboratories deemed eligible and in business in the Telephone Enrollment phase as actually ineligible or out of business after the Field Tabulation stage was implemented. Ineligibles and, more rarely, out-of-business laboratories were identified in the Tabulation stages of Phase I and Phase II.

Table 5-4 presents both the absolute number of Phase I Field Tabulation cases falling into the result categories and the percentage of the total sample represented by each of these categories. Overall, 757 (91.5%) of the fielded laboratory locations were tabulated in the Phase I Field Tabulation, and 36 (4.4%) refused to allow the on-site tabulation after initially agreeing to be enrolled in the study during the Telephone Enrollment stage. An additional 29 (3.5%) locations were recategorized as ineligible during the Phase I Field Tabulation and 2 (0.27%) were determined to be out of business.[14]

Response Rate

For the Field Tabulation stage, as for the Telephone Enrollment stage, laboratories identified during NICLTS operations as ineligible or out of business were excluded from the response rate calculation. This produced a response rate of 95.1 percent:

(757 tabulated) / (827 fielded - (29 ineligible + 2 out of business))

5.1.3.3 Phase I Telephone Verification Results

Table 5-5 presents the results of the Phase I Telephone Verification quality assurance process (described more fully in Section 6.1.1). It presents the verification outcomes for the 757 laboratory locations that were tabulated during the Field Tabulation stage, showing that the verification process was completed for all but four tabulated laboratories. These four incomplete verifications were the result of inability to re-establish telephone contact with the laboratory's field contact person or a suitable substitute.

[14] The 29 ineligible locations identified in the Field Tabulation stage were more than twice the number identified as ineligible at the Telephone Enrollment stage. This phenomenon was due less to the superior opportunity to confirm eligibility on site than to the fact that the Telephone Eligibility Screener was implemented several weeks into the Phase I Telephone Enrollment process. This occurred when it became apparent that there were more ineligible laboratories than anticipated and that it was undesirable to expend field resources unnecessarily on ineligible laboratories. Before the NICLTS protocol was revised to implement a Telephone Eligibility Screener several ineligibles had already passed to the Field Tabulation stage; consequently, these were categorized as ineligibles at that later stage. Many of the 29 ineligibles at the Field Tabulation stage are explained by this phenomenon. Fortunately, the NICLTS home office field support staff administered the screener to most of these cases before the field tabulator made an unnecessary trip to the laboratory. In such cases, the tabulator was instructed by telephone to record the final result of the case as Ineligible.

Table 5-5. Phase I Telephone Verification results

Result	Number of locations	Percent
Verified (unconditionally)	744	98.3
Verified (after specific review)	9	1.2
Unable to reach respondent	4	0.5
Total	930	100.0

The verification process confirmed that the tabulators adhered to the protocol in the tabulation of the other 753 laboratories. Of these 753, all but nine received perfect scores on the verification questionnaire (i.e., every protocol-related question was answered with the desired response). The findings of the verification process are further discussed in Section 6.1.1.3.

5.1.3.4 Phase I Validation Telephone Enrollment

Table 5-6 shows the final outcomes of the Phase I Validation Telephone Enrollment process for the 51 tabulated laboratory locations released as the validation sample. The categories appearing in this table are the same as those defined in Section 5.1.3.1. Phase I Validation is more fully described in Section 6.1.2.

Table 5-6. Phase I Validation Telephone Enrollment results

Result	Number of laboratories	Percent
Enrolled	31	60.8
Ineligible	0	0.0
Refusal	11	21.6
Out of Business	0	0.0
Nonlocatable	0	0.0
Other	9	17.6
Total	51	100.0

* Only one location was retabulated at each validation laboratory

Thirty-one (60.8%) of the validation laboratories sample were enrolled for the Phase I Validation study and 11 (21.6%) refused to participate. Since the laboratories had recently gone through a process they were now being asked to repeat, this rate of refusal was actually lower than expected; a refusal rate as high as 50 percent could have been expected.

Much of the credit for the enrollment success rate of 60 percent belongs to the laboratory staff for their willingness to accommodate the needs of the study, their understanding of its importance, and their appreciation of the purpose of quality assurance. As in the Telephone Enrollment for the Phase I main study, CDC's sponsorship was also a contributory factor, as were the technical and interpersonal skills of the enrollment specialists who handled this specialized enrollment. Additionally, in the case of the Phase I Validation study, testimonial evidence pointed clearly to the professionalism and courtesy of the field tabulators during the original site visits as the principal factor in the willingness of the laboratories to bear the burden of a second tabulation.

There is no quantitative indicator of the specific effect of any one of these factors, nor of their relative effect, for any of the enrollment activity that occurred at this stage, or at any of several other stages of both phases. However, for all such activity, anecdotal evidence provided by the telephone enrollment staff is that CDC's sponsorship opened doors that might otherwise have been closed to the tabulation. While no count is available, the enrollment staff frequently found that it made a difference to fax another copy of the original CDC advance letter whenever they encountered any resistance or uncertainty. Unfortunately, it is not possible to determine how much of this effect is attributable to CDC's name on the letterhead, as compared to that of any other reputable sponsor.

As expected, the enrollment staff also encountered some expressions of dissatisfaction with the CLIA regulatory process itself. However, they again found that emphasizing CDC's sponsorship of NICLTS usually mitigated such negative attitudes. Finally, as the enrollment staff developed their learning curve over the course of the project, they found that, when they encountered certain specific circumstances or specific barriers in the enrollment process, a well-timed reference to CDC's sponsorship often was instrumental in breaking through resistance.

The nine cases (17.6%) of the Phase I Validation Survey Enrollment sample with other results consisted mainly of laboratories whose staff were not available to respond to the enrollment request during the short period allotted for this process. Or, if they were available for the enrollment request, they informed the enrollment specialist that the appropriate laboratory staff to coordinate the site visit would not be available during the relatively brief period allotted for the Phase I Validation Field Tabulation activity. Since the Phase I Validation study took place during the July 4th holiday and the start of summer vacations, this outcome is not surprising.

Response Rate

The response rate for the NICLTS Phase I Validation Survey Enrollment was 60.8 percent, as shown by the following response rate calculation:

(31 enrolled) / (51 sampled - (0 ineligible + 0 out of business))

5.1.3.5 Phase I Validation Field Tabulation Results

The final outcomes of the Phase I Validation Field Tabulation process for the 31 enrolled laboratory locations appear in Table 5-7. For the sake of parallelism, this table presents the same categories as in the comparable table for the main Phase I Field Tabulation, Table 5-4. Thirty of the 31 fielded cases were tabulated; the other laboratory was visited but could not be successfully tabulated because of a Tabulation Device problem. Using the same formula as in the preceding sections, the Validation Field Tabulation response rate was 96.8 percent.

Table 5-7. Phase I Validation Field Tabulation results

Result	Number of locations	Percent
Tabulated	30	96.8
Ineligible	0	0.0
Refusal	0	0.0
Out of Business	0	0.0
Other	1	32
Total	31	100.0

5.1.3.6 Summary of Phase I Response Rates

Table 5-8 is a convenient overview of the response rates for each of the Phase I operational stages.

Table 5-8. Summary of Phase I response rates

Operational stage	Total sample (n)	Response rate (percent)
Telephone Enrollment	930	85.4
Field Tabulation	827	95.1
Validation Telephone Enrollment	51	60.8
Validation Field Tabulation	31	96.8

5.2 Phase II

5.2.1 CLIA ID Number Problem Resolution

To resolve CLIA ID number problems in Phase II, the telephone tabulators used decision rules similar to those in Phase 1. Because tabulation was a very short process in every facility during Phase II, the issue of whether the laboratory would require more or less than 2 days of tabulation time was eliminated. A supervisor was in the telephone center at all times and, if necessary, could help the tabulator with any decisions about how to proceed. Since tabulators were not physically touring a facility, the decision was made to collect information about all testing performed and to record all CLIA ID numbers found at the primary location.

5.2.2 Phase II Summary Case Result Reports[15]

This section presents summary statistics and discussions concerning the operational results (case outcomes) for each laboratory sampled for Phase II. The summary results are presented separately for each Phase II operational stage: Telephone Enrollment, Telephone Tabulation, Validation Telephone Enrollment, and Validation Field Tabulation. The organization and content of this section are similar to that of Section 5.1.3, Phase I Summary Case Result Reports. Issues and topics covered in the earlier section are covered here by reference to that section. Where relevant, the discussion addresses meaningful differences between the results of the two phases.

[15] The numbers of laboratories assigned to the various result categories in the tables appearing throughout this section may represent a slightly different distribution of result categories from that used for statistical weighting purposes. These tables present descriptive statistics of the outcomes of operational processes, specific to NICLTS. Weigh ting uses standard rules for assigning cases to weighting classes.

5.2.2.1 Phase II Telephone Enrollment Results

The final outcomes of the Phase II Telephone Enrollment process for all 1,859 laboratories (i.e., CLIA ID numbers) released as active sample appear in Table 5-9. The categories in this table were defined in Section 5.1.3.1.

Table 5-9. Phase II Telephone Enrollment results

Result	Number of laboratories	Percent
Enrolled	1,473	79.2
Ineligible	78	4.2
Refusal	77	4.1
Out of Business	73	3.9
Nonlocatable	108	5.8
Other	50	2.7
Total	1,859	100.0

Table 5-9 presents both the absolute number of cases falling into these categories and the percentage of the total sample represented by each category. Overall, 1,473 (79.2%) of the sampled laboratories were enrolled for the Telephone Tabulation, and 77 (4.1%) refused to participate at the Telephone Enrollment stage. This percentage of enrolled laboratories is very close to the Phase I percentage (83.1%). The Phase II enrollment refusal percentage of 4.1 percent is less than half of the Phase I percentage (9.8%). This difference in the rate of refusals is further discussed in Section 5.2.2.2.

While still fairly small, the percentages of laboratories that were ineligible and out of business in the Phase II Telephone Enrollment (4.2% and 3.9%, respectively) were both three times as high as in the Phase I Telephone Enrollment. This trend is attributable to three factors. First, the Phase II sample contained a higher concentration of physician offices and other small laboratories. In terms of ineligibility, they were more likely to have acquired a CLIA ID number when they did not need one because they misunderstood when CLIA certification was required. Alternatively, they may have needed one because they had been performing one or two tests requiring CLIA certification but subsequently stopped performing them. They were more likely to go out of business because of retirement, sale, consolidation, or absorption than the larger laboratories in the Phase I sample. These explanations of the higher prevalence of ineligible and out-of-business laboratories are all

functions of the sample population's own characteristics.

The second factor, which applies only to explaining the higher percentage of ineligibles, is a function of the NICLTS protocol. The Telephone Eligibility Questionnaire was added to the Phase I Telephone Enrollment protocol several weeks after enrollment was underway. As discussed in Section 5.1.3.2, this caused a number of ineligible cases to be enrolled that would otherwise have been detected by the screener and immediately categorized as ineligible at the Telephone Enrollment stage. Thus, the percentage of cases found to be ineligible during the Phase I Telephone Enrollment was artificially low and, as a corollary, the percentage of ineligibles found at the Phase I Field Tabulation stage was artificially high. This view is supported by Table 5-10, which shows the relative ineligibility percentages between the Telephone Enrollment and Field Tabulation results of Phase I and Phase II, respectively. Whereas the Phase I Field Tabulation had over twice the number and percentage of ineligibles as Phase I Telephone Enrollment, Phase II Field Tabulation had a smaller number and only slightly larger percentage of ineligibles than the Phase II Telephone Enrollment.

Table 5-10. Ineligibility, Telephone Enrollment and Field Tabulation stages, Phase I and Phase II

	Telephone Enrollment Ineligibles		Field Tabulation Ineligibles	
	Number	Percent	Number	Percent
Phase I	13	1.4	29	3.5
Phase II	78	4.2	71	4.6

The third factor, which applied only to the out-of-business prevalence, was a function of timing. Phase II operations began 10 months after Phase I began, and ended 8 months after Phase I ended. Since both the Phase I and Phase II samples were drawn from the same version of the OSCAR file, there was a longer period during which the Phase II laboratories might have ceased operations.

Thus, as further discussed in Section 5.2.2.2, the overall percentage of ineligibles was higher in Phase II than in Phase I, for the various reasons already discussed. However, the Telephone Eligibility Screener improved the detection of ineligibles at the Phase II Telephone Enrollment Stage by detecting a higher proportion of them at the preferable, earlier stage. Notwithstanding this implementation of the Telephone Eligibility Screener, however, nearly half of all ineligibles identified during Phase II were still not detected until the Field Tabulation stage. This seeming anomaly is discussed in Section 5.2.2.2.

Response Rate

The response rate for the Phase II Telephone Enrollment was 86.2 percent:

(1,473 enrolled) / (1,859 sampled - (78 ineligible + 73 out of business))

This was very close to the Phase I Telephone Enrollment response rate reported in Section 5.1.3.1 (85.4%). The complement, the nonresponse rate, was 13.8 percent.

5.2.2.2 Phase II Telephone Tabulation Results

The final outcomes of the Phase II Telephone Tabulation process for the 1,544 enrolled laboratory locations appear in Table 5-11. The categories appearing in this table are the same as those described in Section 5.1.3.2, Phase I Field Tabulation Results. The cases represented in Table 5-11 are laboratory locations, not CLIA ID numbers. Examination of Tables 5-9 and 5-11 shows that the 1,473 CLIA ID numbers enrolled at the Phase II Telephone Enrollment stage became the 1,544 total laboratory locations sent to the Phase II Telephone Tabulation stage. All results presented in Table 5-11 apply to the individual locations.

Table 5-11. Phase II Field Tabulation results

Result	Number of locations	Percent
Tabulated*	1,322	85.6
Ineligible	71	4.6
Refusal	138	8.9
Out of Business	9	0.6
Other	4	0.3
Total	1,544	100.0

* includes 24 locations at multiple-location laboratories collected without separate tabulation: consolidated by respondents on single data form for each laboratory CLIA ID number overall.

Overall, 1,322 (85.6%) of the laboratory locations sent to the Phase II Telephone Tabulation process were tabulated, and 138 (8.9%) refused to respond to the telephone tabulation interview (i.e., after initially agreeing to

be enrolled in the study during the Telephone Enrollment stage). An additional 71 (4.6%) of locations were categorized as ineligible during the Phase II Telephone Tabulation and 9 (0.6%) were determined to be out of business.

The preceding section discussed various reasons why the Phase II Telephone Enrollment had higher percentages of ineligible and out-of-business laboratories than Phase I. Except for issues related to the full implementation of the Telephone Eligibility Screener in Phase II, these reasons apply to all ineligible and out-of-business laboratories discovered in Phase II, regardless of whether they were detected at the Telephone Enrollment stage or the Telephone Tabulation stage. The net effect was that the percentage of ineligible and out-of-business laboratories was higher for Phase II than for Phase I, at each stage and cumulatively. The Telephone Eligibility Screener did show some effect in Phase II, by increasing the proportion of ineligible laboratories that were detected at the Telephone Enrollment stage. In Phase I, about one-third of all ineligibles were detected during Telephone Enrollment. In Phase II, more than half were detected during enrollment (Table 5-10). In contrast, all but two out-of-business laboratories were detected in the Phase I enrollment but, in Phase II, nine enrolled laboratories were ultimately determined to have been out of business at the time of enrollment, after the telephone tabulator discussed the situation in greater detail during the Telephone Tabulation stage. The question that remains is why a fairly high number of ineligible laboratories were initially enrolled in Phase II, and why any out-of-business laboratories were enrolled in either phase.

The second question is easier to answer. As previously discussed, the out-of-business status was not limited to laboratories that simply closed their doors. It included a continuum of situations, such as laboratories that were bought by other laboratories and consolidated their operations at the acquirer's facility. While the CLIA regulations are clear on the matter, a degree of confusion often remained in the minds of the laboratory managers about the status of the CLIA ID number of the merged laboratory. Thus, it was feasible for the enrollment specialists to trace the CLIA ID number to a fully operational laboratory, where staff were quite able to make a connection with the information about the sampled laboratory, by name and CLIA ID number. Since the enrollment protocol employed a chain-of-evidence approach to locating the sampled laboratory when there was instability in the name, address, and so forth, it was a natural outcome for such laboratories to pass successfully through the enrollment process. That process attempted to clarify the situation as much as possible, but the protocol always anticipated that the most confusing and complex situations would require the in-depth investigation that the tabulators or their supervisors would be in a better position to pursue knowledgeably.

Most of the foregoing explanation for the initial enrollment of a few out-of-business laboratories applied as well to the enrollment of ineligible laboratories. Here the issues of ambiguity and confusion about the status of the sampled CLIA ID number came into play with full force. The issues were diverse: multiple simultaneous CLIA ID numbers; accretion of CLIA ID numbers over time; and CLIA ID numbers associated with different providers, locations, or acquired facilities. The enrollment protocol also explicitly sought to enroll a laboratory even when the CLIA ID number confirmation was in doubt, again with the goal of allowing the tabulator to make a final determination after further investigation.

When all of these factors are considered, it becomes clear why a number of ineligibles were encountered at the tabulation stage. The Telephone Eligibility Screener was designed primarily to address the issue of the types of tests being performed, rather than the CLIA ID number issue. It ensured the inclusion of facilities that did not consider themselves as performing any clinical tests, but actually did perform testing. It filtered out those that clearly did not perform any tests in calendar year 1996. It did not exclude a laboratory that could not confirm its CLIA ID number. Absent an enrollment protocol that stringently required positive confirmation of the sampled CLIA ID number and vigorously investigated multiple location/multiple CLIA ID number/multiple provider situations before classifying a laboratory as eligible to proceed to the tabulation stage, a percentage of ineligible cases inevitably passed through enrollment before being detected at the tabulation stage.

Response Rate

For the Telephone Tabulation stage, as for the Telephone Enrollment stage, laboratories identified during NICLTS operations as being ineligible or out of business were excluded from the response rate calculation. This produced a response rate of 90.3 percent:

(1,322 tabulated) / (1,544 sent to telephone tabulation - (9 out of business + 71 ineligible))

While this was considered a successful result, it was somewhat lower than the analogous Phase I Field Tabulation response rate (94.9%). Most of the difference was due to the higher percentage of laboratories that refused to respond at this stage of Phase II (8.9%, versus 4.4% for Phase I). There are several likely explanations for this. In Phase I, the field tabulator's presence may have had a "foot-in-the-door" effect, making it more awkward for the laboratory to decline participation. If laboratory contacts were reticent to turn down someone in

person, in Phase II they may have been less averse to refusing someone on the telephone.

The fact that the Phase I field tabulator did the work of reviewing the records and assembling the test data was presented as an incentive for the Phase I laboratories to participate. In Phase II, the laboratory contact bore the burden of assembling the test data on the mailed Test Inventory Form. This may have had a three-part effect on the comparative rates of refusal between Phase I and Phase II tabulations. First was the simple difference in the burden (work performed by tabulator versus work performed by laboratory). Second, the laboratory had agreed to enrollment after hearing a brief description of the process. When the Phase II laboratory then received the actual tabulation materials in the mail, they may have decided that the process was more involved or time-consuming than they had anticipated. Third, in POL's it was more common than in other types of laboratories that a physician was identified as the only feasible respondent. Given that a physician may have less time or inclination to review and fill out the NICLTS material, and given the high concentration of POL's in the Phase II sample, it is not surprising that more refusals would occur in Phase II after a laboratory initially agreed to be enrolled in the study.

5.2.2.3 Phase II Validation Telephone Enrollment

The final outcomes of the Phase II Validation Telephone Enrollment process for the 204 tabulated laboratory locations released as the validation sample appear in Table 5-12. (Phase II Validation is more fully described in Section 6.2.2.2.) The categories appearing in this table are the same as those defined in Section 5.1.3.1.

Table 5-12. Phase II Validation Telephone Enrollment results

Result	Number of locations	Percent
Enrolled	127	62.3
Ineligible	0	0.0
Refusal	72	35.3
Out of Business	0	0.0
Nonlocatable	0	0.0
Other	5	2.5
Total	204	100.0

Overall, 127 (62.3%) of the Phase II validation sample was enrolled, while 72 (35.3%) refused to participate. As in the Phase I Validation Enrollment, this rate of refusal was still lower than expected.

The 35.3 percent who refused to participate in the Phase II Validation Field Tabulation did, however, represent a substantially higher refusal rate than in the Phase I Validation Field Tabulation (21.6%). This higher rate was likely due to three factors. First, because of a variety of design and operations considerations, more time had elapsed between the original tabulation and the validation request for many of the Phase I sample laboratories as compared to the Phase II sample laboratories. Thus, the Phase II request may have seemed more abrupt or pressing than the Phase I request.

Second, the Phase I laboratories were asked to do nothing more than they had already done in the main study. In contrast, the Phase II laboratories had participated in the primary data collection by telephone, and the request was now for the more intrusive activity of allowing a tabulator to enter their laboratory. Furthermore, most Phase I participants had already surrendered any initial concerns about burden or intrusiveness after experiencing the well-planned protocol and professionalism of the field staff. In contrast, the Phase II sample were asked to agree to an unknown experience, with the potential perception of extra burden or intrusion.

A third possible factor was the higher percentage of physician office laboratories in the Phase II sample. In perception and possibly in reality, physician office operations may have been more affected by the on-site tabulation process or concerned about patient confidentiality issues. In POL's, the physician was often the only feasible contact person, with less time or inclination to accommodate the field tabulator. In some combination, all of these factors were likely to have contributed to the higher percentage of refusals to the Phase II Validation Enrollment.

As in Phase I, the small percentage of the Phase II Validation Telephone Enrollment sample falling into the "Other" result category consisted mainly of laboratories whose staff were not available to respond to the enrollment request or during the period allotted for the Phase II Validation Field Tabulation activity.

Response Rate

The response rate for the Phase II Validation Telephone Enrollment was 62.3 percent:

(127 enrolled) / (204 sampled - (0 ineligible + 0 out of business))

Achieving this high a response rate was attributable to the factors suggested in Section 5.1.3.4 for Phase I: laboratory good will, CDC sponsorship, the positive effect created by the professionalism and courtesy of the telephone tabulators during the original tabulation, and the skill of the telephone tabulators who were given the special assignment of performing as the Phase II validation enrollment.

5.2.2.4 Phase II Validation Field Tabulation Results

The final outcomes of the Phase II Validation Field Tabulation process for the 127 enrolled laboratory locations appear in Table 5-13. For the sake of parallelism, this table presents the same categories as those appearing in the comparable table for the main Phase I Validation Field Tabulation, Table 5-7.

Table 5-13. Phase II Validation Field Tabulation results

Result	Number of locations	Percent
Tabulated	110	86.6
Ineligible	0	0.0
Refusal	4	3.1
Out of Business	0	0.0
Other	13	10.2
Total	127	100.0

Of the 127 laboratories enrolled, 110 (86.6%) were tabulated. There were four refusals (3. 1 %). The 13 laboratories falling into the "Other" result category consisted mainly of ones where it was not feasible to complete a tabulation, because of tabulator time and distance considerations at the end of the Validation Field Tabulation period or because the laboratory staff proved to be unavailable during the field period. Using the same formula as

in preceding sections, the Phase II Validation Field Tabulation response rate was 86.6 percent.

5.2.2.5 Summary of Phase II Response Rates

Table 5-14 is a convenient overview of the response rates for each of the operational stages.

Table 5-14. Summary of Phase II response rates

Operational stage	Total sample (n)	Response rate (percent)
Telephone Enrollment	1,859	86.2
Field Tabulation	1,544	90.3
Validation Telephone Enrollment	204	62.3
Validation Field Tabulation	127	86.6

5.3 Overall Survey Results

Table 5-15 shows the number of distinct clusters by region and group. These data are the distinct clusters over the sampled laboratories in the categories indicated. For example, POL's in the Northeast region tested 148 distinct clusters overall, while POL's as a whole tested a total of 1,604 distinct clusters. The full sample of laboratories that participated in the NICLTS tested 8,164 distinct clusters. Table 5-16 shows the average number of distinct clusters per laboratory. For example, POL's nationally tested an average of nine distinct clusters, though this is somewhat higher for POL's in the Northwest region (13.7 distinct clusters). The average among all laboratories in the NICLTS sample was 15.2 distinct clusters per laboratory.

The estimated mean number of distinct clusters per laboratory is given in Table 5-16. Nationally, the estimated mean number of distinct clusters tested per laboratory is 15.2 (95% CI 13.9 to 16.4).

Table 5-15: Number of distinct clusters by region and group

	Laboratory Group						
Region	POL	Other Ambulatory	Hospice, Nursing home	Hospital	Independent, Blood bank	Specialty	Total
1. Northeast	148	105	29	818	127	49	1,075
2. New York, New Jersey	348	79	79	976	231	183	1,582
3. Mid-Atlantic	433	197	30	1,361	171	224	1,968
4. Southeast	650	209	31	1,168	484	333	2,124
5. Midwest (North)	602	114	46	2,118	508	416	3,130
6. South (Central)	264	231	30	2,052	180	846	2,883
7. Midwest (Central)	204	174	25	970	90	387	1,474
8. Mountain	142	117	13	867	49	393	1,323
9. West	327	115	19	834	410	471	1,708
10. Northwest	209	52	14	223	208	217	733
Total	1,604	822	151	5,698	1,679	2,322	8,164

NOTE: This table shows the number of distinct clusters in the sample in each cell. For example, data collected from sampled POL's in the Northeast region had a total of 148 distinct clusters.

Table 5-16: Estimated mean number of distinct clusters per laboratory by region and group

	Laboratory Group						
Region	POL	Other Ambulatory	Hospice, Nursing home	Hospital	Independent, Blood bank	Specialty	Total
1. Northeast	7.0	7.9	2.2	148.2	29.5	4.3	13.3
2. New York, New Jersey	7.6	7.9	6.6	131.4	55.5	9.9	14.2
3. Mid-Atlantic	10.4	13.4	2.2	113.4	27.7	7.9	16.7
4. Southeast	10.4	7.6	1.6	64.7	45.7	7.3	13.3
5. Midwest (North)	9.0	5.0	1.7	126.9	50.0	9.0	15.6
6. South (Central)	7.7	5.9	1.4	104.9	24.2	19.5	16.6
7. Midwest (Central)	9.6	12.7	1.3	104.0	24.2	17.9	17.8
8. Mountain	10.0	11.4	1.6	139.6	14.1	33.1	22.5
9. West	7.6	5.6	1.4	56.5	40.0	15.6	12.3
10. Northwest	13.7	7.5	3.0	64.9	73.7	19.2	18.0
Total	9.0	7.5	1.9	100.6	39.0	12.4	15.2

Based on NICLTS data, the estimated total national volume of testing performed in calendar year 1996 is 7,250,519,342 (7.25 billion) tests; the 95 percent confidence interval around this point estimate is 5.12 to 9.38 billion tests. The average volume per laboratory for all laboratories is 51,114 tests; the 95 percent confidence interval ranges from 36,119 to 66,109 annual tests per laboratory.

Table 5-17 gives estimates of volume by DHHS region and laboratory group. The greatest volume of testing is performed in Region 4 (Southeast); across the six major laboratory groups, the greatest volume of testing is performed by hospitals. Table 5-18 gives the estimated mean volume per laboratory. The Northeast (Region 1) has the highest average volume per laboratory. As might be expected, hospitals show the highest average volume per laboratory. Table 5-19 gives estimates for total volume and mean volume per laboratory by more detailed categories of laboratory type.

Table 5-17. Estimated volume of tests by region and group (000,000)

	Laboratory Group						
Region	POL	Other Ambulatory	Hospice, Nursing home	Hospital	Independent, Blood bank	Specialty	Total
1. Northeast	24	24	6	441	91	10	596
2. New York, New Jersey	68	5	8	551	38	44	715
3. Mid-Atlantic	71	14	6	543	30	41	705
4. Southeast	167	21	12	302	351	912	1,765
5. Midwest (North)	115	6	19	672	126	53	991
6. South (Central)	44	21	8	507	26	295	901
7. Midwest (Central)	36	12	3	154	36	39	279
8. Mountain	25	9	3	123	2	37	199
9. West	113	12	6	212	347	175	865
10. Northwest	70	1	2	13	133	16	236
Total	732	126	73	3,518	1,181	1,621	7,251

Table 5-18. Estimated mean volume per laboratory by region and group

	Laboratory Group						
Region	POL	Other Ambulatory	Hospice, Nursing home	Hospital	Independent, Blood bank	Specialty	Total
1. Northeast	5,456	29,255	4,927	1,258,852	280,753	11,413	74,880
2. New York, New Jersey	7,623	7,828	10,922	1,070,788	123,187	38,786	58,580
3. Mid-Atlantic	8,492	13,618	5,168	616,823	66,559	21,298	51,080
4. Southeast	10,105	8,767	5,639	186,939	404,811	229,760	64,123
5. Midwest (North)	7,753	2,845	5,843	435,960	176,862	12,310	36,818
6. South (Central)	5,097	6,480	4,336	352,093	39,329	101,682	47,924
7. Midwest (Central)	9,639	12,541	2,050	230,110	139,630	26,803	32,851
8. Mountain	11,388	14,523	3,976	336,312	9,879	55,598	42,227
9. West	11,214	6,635	5,367	223,065	404,990	90,048	51,392
10. Northwest	26,111	2,561	4,857	66,691	643,829	22,305	50,960
Total	9,118	8,815	5,173	412,762	243,192	81,612	51,114

Table 5-19. Estimated distinct clusters and volume by laboratory type

Laboratory type	Distinct clusters	Mean distinct clusters per laboratory	Total volume (000,000)	Mean Volume per laboratory
1. Ambulatory surgery center	452	20.6	243.54	188,618
2. Community clinic	675	14.2	66.04	12,850
3. Comprehensive outpatient rehab.	44	44.0	11.50	207,853
4. Ancillary test site	394	14.0	71.31	35,347
5. End stage renal disease dialysis	29	2.1	7.00	4,154
6. Health fair	5	2.0	0.08	501
7. Health Maintenance Organization (HMO)	215	20.2	44.45	51,714
8. Home health agency	52	1.5	13.24	1,862
9. Hospice	6	1.3	0.01	20
10. Hospital	5,698	100.6	3,517.66	412,762
11. Independent	1,656	41.6	1,159.37	258,384
12. Industrial	89	4.7	8.31	7,173
13. Insurance	2	2.0	0.02	402
14. Interm. care facil. mentally retarded	52	6.0	1.97	3,210
15. Mobile unit	21	1.7	0.23	341
16. Pharmacy	9	3.3	0.02	158
17. School/student health service	65	4.5	1.91	1,675
18. Skilled nursing/nursing facility	151	2.0	72.78	5,293
19. Physician office	1,604	9.0	732.02	9,118
20. Other practitioner	496	19.7	57.04	30,106
22. Blood banks	33	6.9	21.64	58,608
23. Other	1,490	13.6	1,220.38	120,218

NOTE: "Distinct clusters" gives the number of distinct clusters in the sample for each laboratory type. "Mean distinct clusters" gives the mean number of distinct clusters per laboratory "Total volume" represents the estimated national volume for each laboratory type. "Mean volume" represents the estimated mean per laboratory.