Chapter 4
Sampling Design and Participant Recruitment

4.1 Sampling Design

A population-based, stratified random sampling design (Figure 4.1.1) was developed to collect the data needed to meet the objectives of the study. In each state, four urban and two rural counties, representing three distinct geographical areas in the state, were randomly selected. Within these counties, there were two sampling frames (components), which were designed to allow testing of the study hypotheses, and in particular, to test whether the children's exposures are significantly different at day care versus at home. The first sampling component, the telephone component, was composed of households that were selected randomly through list-assisted telephone sampling. The telephone component enrolled households with preschool children who did not attend day care. The second sampling component, the day care component, was composed of child day care centers that were randomly selected and enrolled households with preschool children who did attend day care. Within these components, the households and child day care centers were stratified by income.

In both North Carolina (NC) and Ohio (OH), six counties were selected using stratified random sampling. Because of stratification, the samples represented different regions, urban and rural areas, and low-income and middle/high-income areas of each state. The sample selection process targeted counties with larger population and in particular, larger population in the low-income groups, by selecting counties using probabilities proportional to size (PPS) within each stratum. The county population in the low-income segment was used as a measure of size. This approach ensured greater representation of low-income families than would have occurred otherwise. The locations of these counties in the two states are shown in Figure 4.1.2. The selected counties were in three distinct geographical areas in each state. In NC, these geographical areas were the coastal plain, the Piedmont, and the mountains. In OH, the areas were the northern, central, and southern regions.

Within each of the two states, the samples were further stratified according to degree of urban character (urbanicity) and family income. The urbanicity stratification was imposed at the first stage of selection by classifying counties as predominantly urban or rural. A county was considered urban if it was within or contained wholly or in part a Metropolitan Statistical Area (MSA) as defined by the Office of Management and Budget (OMB Bulletin No. 99-04). Income stratification was performed at subsequent stages of selection for the day care component and the telephone component. This stratification was used to distinguish between low-income and middle/high-income households and day care centers. Day care centers were classified as low-income if they received Federal assistance to serve low-income clients under the Head Start

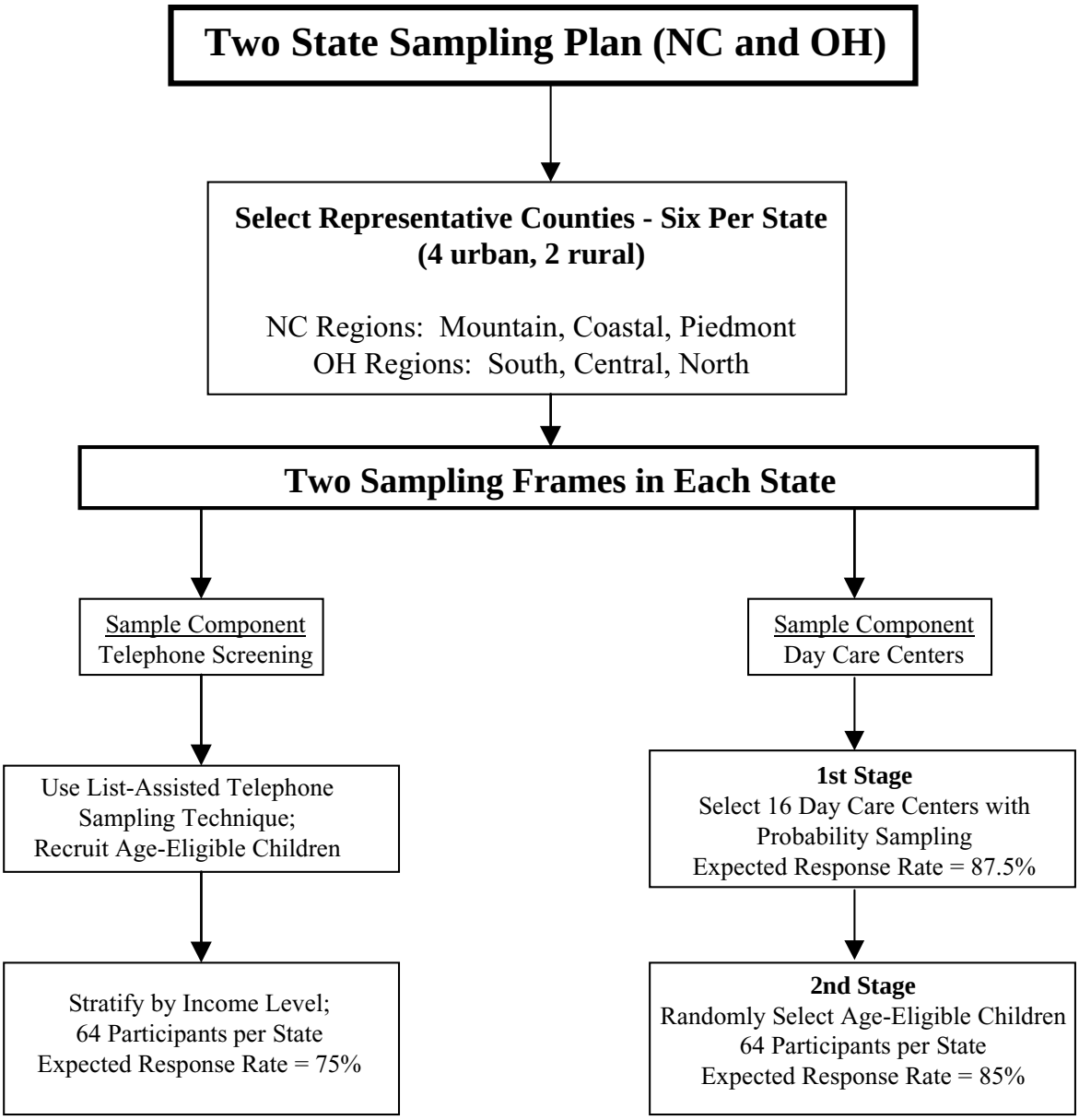
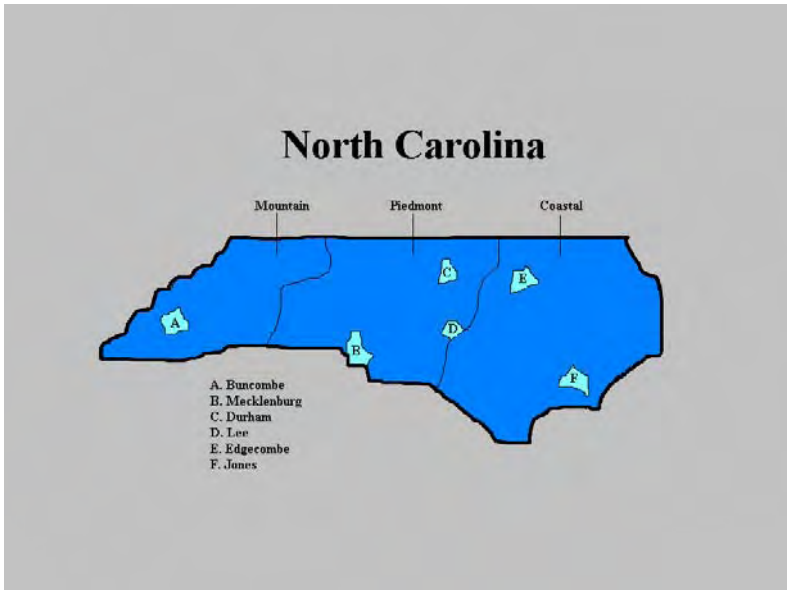
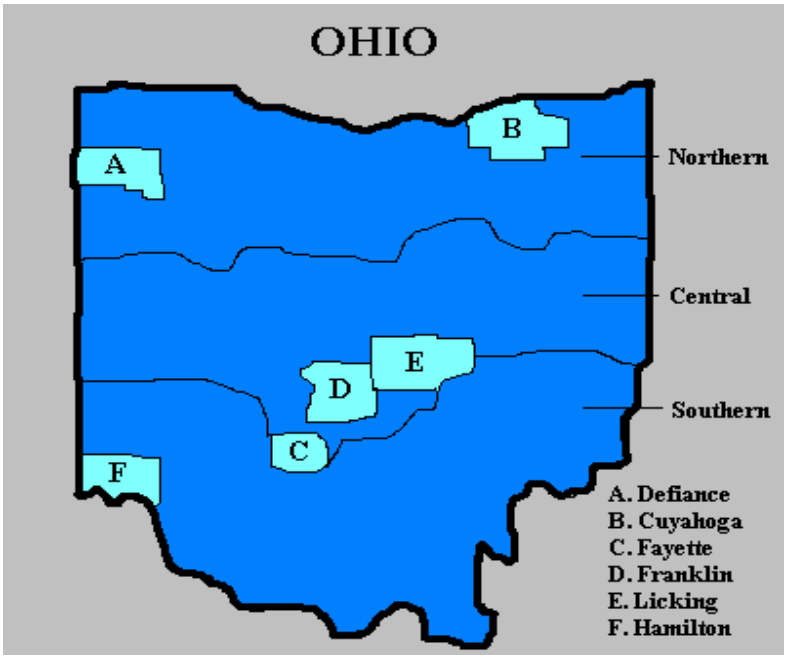


Figure 4.1.1 CTEPP Overall Sampling Design



A



B

Figure 4.1.2 Six Counties in North Carolina (A) and Ohio (B) Selected by Stratified Random Sampling

program. Low-income families were classified according to the federal guidelines for assistance eligibility under the Women, Infants, and Children program (WIC, 2000). A household was classified as low-income if its household income was below 185% of the federal poverty guidelines (Federal Register, 2000). In 2000, the WIC eligibility level for a family of two was $20,813 and for a family of four was $31,534.

In the day care component, all eligible child day care centers in the six selected counties were identified. A child day care center was considered eligible if it was a commercial or not-for-profit service provider, which provided child care services to seven or more preschool children at a location other than the service provider's personal residence. During the second-stage sampling frame, these centers were divided into the two income strata. From these strata, a random sample of targeted centers and a random sample of eligible children within each participating center were selected. In the telephone component, a random sample of telephone numbers was selected, using list-assisted telephone sampling techniques in the six counties in each state. The anticipated sample size was 128 children in each state, with half (64) from the day care center sample (children who attended day care) and the other half (64) from the telephone sample (children who did not attend day care). This dual frame approach provided maximum coverage for the target population.

4.2 Recruitment

4.2.1 Recruitment of the Day Care Center Component

Recruitment of the day care center component was conducted in two stages, as diagramed in Figure 4.2.1. In the first stage, master lists of all day care centers in NC and of all those in OH were compiled. For the six target counties in each state, a complete list of day care centers in each county was prepared and sorted by urbanicity and income. From these lists, approximately 16 centers were targeted for selection; of these at least four were Head Start centers, which served primarily low-income clients. The centers were contacted through telephone calls and mailings. In the second stage of the day care center component, eligible children who attended the day care centers were selected randomly from up to two classrooms in each participating center. Classroom information was requested from each of the centers. Parents or primary caregivers were contacted through the centers, as discussed below, to obtain informed consent for study participation.

Because every eligible child day care center must be licensed to operate in its state, the state licensing agencies were the main sources of comprehensive lists of centers in both NC and OH. Additionally, to ensure the completeness of the master lists of child day care centers, the lists obtained from the state agencies were supplemented with information on centers from other sources. The most updated CD-ROM national telephone database (Pro-CD, 1999-2000, infoUSA Inc.) was searched, and a list of eligible day care centers in the target counties was prepared. In addition to the CD-ROM national telephone database, an Internet search was done. Centers that appeared on the CD-ROM national telephone database and/or the Internet were cross-checked against the lists provided by state licensing agencies. Centers that appeared on the

CD-ROM national telephone database and/or Internet, but did not appear on the list from state licensing agencies, were called to determine the eligibility status of the center. Additional eligible centers were then added to the master list.

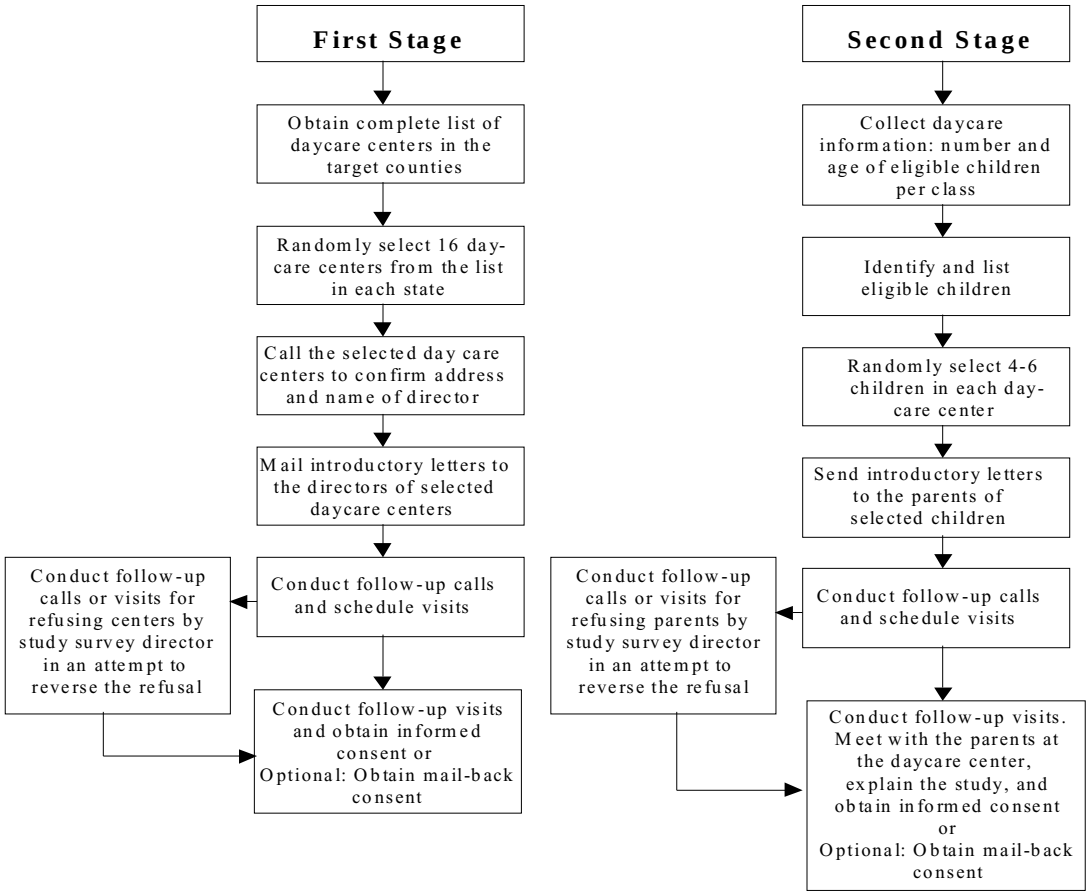


Figure 4.2.1 Procedures for Recruiting Day Care Center Component

This sampling component was then stratified by county and by whether or not the center received Federal assistance to serve low-income clients (Head Start centers). Within each stratum, day care centers were selected, with probability proportional to the number of children enrolled in the center. A total of 16 centers, including at least four Head Start centers, were targeted for recruitment in each state. Further details on the day care center sample recruitment can be found in the recruitment reports from NC and OH (Appendix B).

Screening calls were conducted by the recruitment team, to confirm the addresses of the selected centers and the names of the center directors. After confirmation, the recruitment team sent an introductory letter, a study brochure, and a gift certificate (as incentive for the center to participate) to each day care center director by overnight express mail. Approximately three days after the letters were mailed, the recruitment team made follow-up calls to each director. To encourage participation of each center, the team made follow-up visits to the center director, and the Battelle field team leader contacted the center as needed. The first stage recruitment activities were completed by obtaining informed consent forms from each day care center.

The second sampling stage of the day care component involved selecting a random sample of eligible children from up to two classrooms in the selected centers. Children in the child day care center component were eligible if they were between the ages of 18 months and 5 years, toilet-trained or able to provide at least one urine sample, and not being breast-fed. In addition, they had to attend a state-licensed child day care center, serving seven or more children, on three consecutive days, for at least 25 h per week.

The second stage recruitment activities began with the determination of the number of age-eligible children in each classroom. Classroom Information sheets were sent to and completed by the day care director. These sheets requested the following information for each classroom: name of the classroom, total number of children in the classroom, and the initials and ages of eligible children. Two classrooms and five children in each classroom were selected randomly. Following the selection of the children, the recruitment team asked the day care director to distribute the recruitment package, which contained an introductory letter, a study brochure, and a gift certificate (as incentive for the household to participate), to the parents of the selected children. Parents were encouraged to call the project toll-free number to ask about the study. In consultation with the day care center director, the recruitment team also set up an appropriate time, typically two or three days after the letters were sent, to meet with the parents at the day care center.

During the meeting with the parents, the recruitment team established rapport with the parents and the child, and gave a small gift to the child, such as a book or small toy. The recruitment team emphasized the positive experiences that we and the participants had in our previous pilot studies. An informed consent form was obtained from the parents, and they were asked to complete the Recruitment Survey (Form #1; Table 5.2.2). The recruitment team then scheduled an initial sampling date with each family.

4.2.2. Recruitment of the Telephone Sample Component

The procedures for recruiting households by telephone sampling are diagramed in Figure 4.2.2. A telephone sample list, which included addresses, was ordered from a commercial survey sampling firm (Marketing Systems Group [MSG], Genesys Sampling System, http://www.genesys-sampling.com). The sample design used for the telephone component was: (1) to identify efficiently, through telephone contact, households having one or more children in the eligible age range, that met the sampling targets in the household low-income or middle/high-income domains, and (2) to provide coverage of households with unlisted telephone numbers.

The survey sampling firm used Census data, marketing research data, and other sources to classify directory-listed households as having either one or more children in the age range of 18 months to 5 years, or having no children in that age group. The same data were used to assign the directory-listed households to an income range. All directory-listed households in each of the six counties were assigned to one of the following four strata:

1. Directory-listed households with income above $25,000 and having one or more children in the target age range
2. Directory-listed households with income below or equal to $25,000 and having one or more children in the target age range
3. Directory-listed households with income above $25,000 and having no children in the target age range
4. Directory-listed households with income below or equal to $25,000 and having no children in the target age range

In some counties, as many as 30% of households could have unlisted telephone numbers. To ensure inclusion of those households that did not appear in the directories, a Random Digit Dialing (RDD) approach was used. To implement the RDD approach, the survey sampling firm first identified all telephone exchanges in the selected county. Telephone exchanges having very low percentages of directory-listed households, primarily nonresidential or business areas, in the selected county were deleted. From the remaining exchanges, a systematic random sample of all numbers was drawn. Some of these telephone numbers were residential, and some were business or nonworking numbers. To prevent a directory-listed telephone number from being sampled in both the RDD frame and the directory-listed frame, the survey sampling firm selected the RDD sample of telephone numbers first. The sampled telephone numbers were compared to the database of directory-listed telephone numbers. Those telephone numbers that were directory-listed were removed from the directory-listed frame, prior to the stratification described above. The list-assisted samples, corresponding to the four strata above, and the RDD samples were combined in replicate files. This telephone sample selection did not include households without home telephones; however, they were represented in the day care sample component.

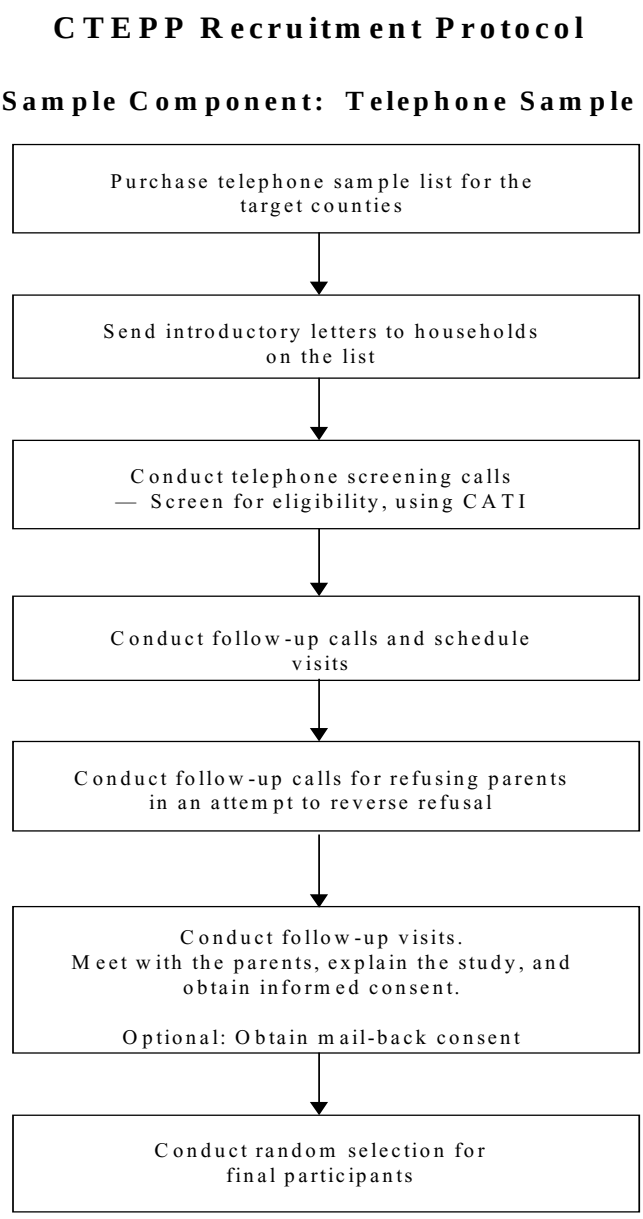


Figure 4.2.2 Procedures for Recruiting Telephone Sample Component

Introductory letters and a study brochure were sent to households in the telephone list that had valid addresses. A Computer Assisted Telephone Interview (CATI) system was developed to facilitate the screening process. All numbers in the files were called and screened for eligible subjects. Children were eligible for this telephone sample component if they were between the ages of 18 months and 5 years, toilet-trained or able to provide at least one urine sample, not being breast-fed, and not attending a day care center. The final participants were randomly selected from the eligible subjects. Staff visited those households that tentatively agreed to participate in the study. At these visits, the staff explained the study further and obtained informed consent.

4.3 Recruitment Results

4.3.1 North Carolina

Recruitment of subjects for the NC field study was conducted in two phases. Recruitment of Phase I participants began in four NC counties (Durham, Buncombe, Lee, and Mecklenburg) in early February 2000, but was suspended on February 29 for four months due to the OMB 2000 Census requirement. The OMB prohibited other federally-sponsored surveys from occurring during the period from March to June 2000 while the 2000 U.S. Census was conducted.). Recruitment of subjects in these counties resumed in July 2000 and continued through December 2000. Phase I field sampling activities were completed with 48 households in December 2000. Recruitment of Phase II subjects was conducted for the two eastern NC counties affected by severe flooding from Hurricane Floyd (Edgecombe and Jones) from February 26 through March 30, 2001. Twelve additional subjects and their adult caregivers from the day care center sample component were enrolled in Phase II. In Jones County, although one day care center agreed to participate in the study, no parents were willing to participate, because they were still dealing with the flooding problems from the hurricane.

A conservative approach was used to calculate the final response rate. During the recruitment period, some people refused to be screened and some could not be reached. As a result, their eligibility status was unknown. A calculated eligibility rate was used to estimate the number of eligible subjects in this group of status-unknown subjects. This eligibility rate, which was determined from the known responses, was calculated as the total number of eligibles divided by the sum of the total number of eligibles and ineligibles. To calculate the final response rate, the number of eligible subjects who agreed to participate was divided by the estimated total number of eligible subjects – the total of those eligibles who responded plus the estimated eligibles. This approach tends to underestimate the final response rate, because it does not include the number of status-unknown subjects who might be eligible and agree to participate in the study but could not be reached.

Table 4.3.1 summarizes the response rates for the NC study. Overall, 98% of the recruitment target for day care participants in NC was achieved through enrollment of a total of 63 of 64 target households. Overall, 105% of the targeted number (67 of 64 targeted) of telephone sample households in NC were enrolled in the CTEPP study. All recruitment activities for NC were completed by March 30, 2001.

Table 4.3.2 provides the overall recruitment results for NC for the children who were recruited at home or at day care. The final recruitment results for the NC field study led to the enrollment of 130 children, ranging in age from 20 to 64 months, and their primary adult caregivers.

Table 4.3.1 Summary of CTEPP North Carolina Response Rates

Sampling Frame	Summary
Child Day Care Component: Child Day Care Centers	
(A) Eligible and Recruited Child Day Care Centers	13
(B) Eligible Child Day Care Centers	17
(C) Ineligible Child Day Care Centers	5
(D) Unknown Eligibility	10
(E) Calculated Response Rate[a]	53%
Child Day Care Component: Day Care Parents	
(A) Eligible and Recruited Day Care Parents	69
(B) Eligible Day Care Parents	85
(C) Ineligible Day Care Parents	26
(D) Unknown Eligibility	71
(E) Calculated Response Rate[a]	50%
Telephone Screening Component	
(A) Eligible and Recruited Stay-at-Home Parents	272
(B) Eligible Stay-at-Home Parents	333
(C) Ineligible Stay-at-Home Parents	6547
(D) Unknown Eligibility	2807
(E) Calculated Response Rate [a]	58%

[a] Calculated Response Rate, E = (A)/(B + (B/(B + C)) x D)

Table 4.3.2 Summary of CTEPP North Carolina Participant Characteristics

Final NC Results		Telephone Sample				Day Care Sample				
		Unknown	**Low-income**	**Mid-income**	**Subtotal**	**Unknown**	**Low-income**	**Mid-income**	**Subtotal**	**Total**
Urban	Buncombe		6	1	**7**		6	4	**10**	**17**
	Durham		5	21	**26**		5	12	**17**	**43**
	Mecklenburg	3	2	15	**20**	1	11	3	**15**	**35**
	Edgecombe		1	1	**2**	1	11	0	**12**	**14**
	Total Urban	**3**	**14**	**38**	**55**	**2**	**33**	**19**	**54**	**109**
Rural	Lee		4	3	**7**	1	5	3	**9**	**16**
	Jones	1	3	1	**5**		0	0	**0**	**5**
	Total Rural	**1**	**7**	**4**	**12**	**1**	**5**	**3**	**9**	**21**
	Total NC	**4**	**21**	**42**	**67**	**3**	**38**	**22**	**63**	**130**
	% of Total	**6%**	**31%**	**63%**	**100%**	**5%**	**60%**	**35%**	**100%**	

Thirteen NC day care centers (eight regular day care and five Head Start) participated in the study. Sixty-three day care children, day care teachers, and their caregivers successfully completed the field activities of the study. Sixty-six stay-at-home children and their caregivers, successfully completed the field activities of the study. One stay-at-home participant did not complete the study. The distribution of low-income and middle/high-income of the NC families in the telephone sample component was very close to the original sampling design. However, in the day care sample, low-income families were over-enrolled, with 60% of the day care sample classified as low-income. This over-enrollment of low-income families in the day care sample occurred because many of the children in the regular day care centers, those not catering specifically to low-income families through the Federally funded Head Start program, came from families that were classified as low-income. Further information on the NC field study can be found in the NC Recruitment Report (Appendix B) and in our published paper on the CTEPP sampling design and field methodology (11).

4.3.2 Ohio

Recruitment of subjects for the OH field study began in January 2001 and was completed in November 2001. Fifty-eight households were successfully recruited. Table 4.3.3 summarizes the response rates for the OH study. For the day care sample component, 91% of the recruitment target for day care participants in OH was achieved through enrollment of a total of 58 of 64 target households. For the telephone sample component, a total of 165 potentially eligible households were identified. Overall, 108% of the target stay-at-home participants were recruited through enrollment of a total of 69 of 64 target households. All recruitment for OH was completed in November 2001.

Table 4.3.4 provides the overall recruitment results for OH, for both the stay-at-home and day care children. The final recruitment results for the OH field study led to the enrollment of 127 children, ranging in age from 20 to 65 months, and their primary adult caregivers.

Sixteen OH day care centers (12 regular day care and 4 Head Start) participated in the study. Fifty-eight day care children and their caregivers, participated successfully in the field activities of the study, with simultaneous sampling both at the centers and at the children's homes. Sixty-nine households in which the children did not attend day care participated successfully in the field activities of the study, with sampling for the children and their primary caregivers at the children's homes. The distribution of low-income and middle/high-income families in the OH telephone sample component is very close to the original sampling design, with 26% of the stay-at-home participants classified as low-income. However, as in NC, the low-income families were over-enrolled in the day care sample component, with 50% of the day care participants classified as low-income. Further information on the OH field study can be found in the OH Recruitment Report (Appendix B) and in our published paper on the CTEPP sampling design and field methodology (11).

Table 4.3.3 Summary of CTEPP Ohio Participant Response Rates

Sampling Frame	Summary
Child Day Care Component: Child Day Care Centers	
(A) Eligible and Recruited Child Day Care Centers	16
(B) Eligible Child Day Care Centers	24
(C) Ineligible Child Day Care Centers	4
(D) Unknown Eligibility	5
(E) Calculated Response Rate[a]	57%
Child Day Care Component: Day Care Parents	
(A) Eligible and Recruited Day Care Parents	71
(B) Eligible Day Care Parents	100
(C) Ineligible Day Care Parents	8
(D) Unknown Eligibility	141
(E) Calculated Response Rate[a]	31%
Telephone Screening Component	
(A) Eligible and Recruited Stay-at-Home Parents	165
(B) Eligible Stay-at-Home Parents	191
(C) Ineligible Stay-at-Home Parents	4598
(D) Unknown Eligibility	2449
(E) Calculated Response Rate [a]	57%

[a] Calculated Response Rate, E = (A)/(B + (B/(B + C)) x D)

Table 4.3.4 Summary of CTEPP Ohio Participant Characteristics

Final OH Results		Telephone Sample				Daycare Sample				
		Unknown	Low-income	Mid-income	Subtotal	Unknown	Low-income	Mid-income	Subtotal	Total
Urban	Cuyahoga	1	4	11	**16**		10	10	**20**	**36**
	Licking			7	**7**		4		**4**	**11**
	Franklin		7	13	**20**	2	6	8	**16**	**36**
	Hamilton		2	15	**17**	1	9	0	**10**	**27**
	Total Urban	**1**	**13**	**46**	**60**	**3**	**29**	**18**	**50**	**110**
Rural	Defiance		2	4	**6**	2		2	**4**	**10**
	Fayette		3		**3**			4	**4**	**7**
	Total Rural	**0**	**5**	**4**	**9**	**2**	**0**	**6**	**8**	**17**
	Total OH	**1**	**18**	**50**	**69**	**5**	**29**	**24**	**58**	**127**
	% of Total	**1%**	**26%**	**72%**	**100%**	**9%**	**50%**	**41%**	**100%**	

In addition to the field sampling and data collection described above for both NC and OH, 26 children in OH were videotaped for about two hours in their homes, in order to supplement the information collected within activity diaries and other observations. Videotaping started in OH in April 2001 and ended in October 2001. Sixty-nine percent of these 26 OH children were stay-at-home children; 88% percent of them lived in urban counties; and 38% percent of them were from low-income families. Fifty percent of the participants were female, and the children's ages ranged from two to five years.

4.4 Evaluation

Recruitment strategies included minimizing the burden on participants, ensuring confidentiality, providing incentives for participation, and using carefully selected and trained field staff. Throughout the study, the staff were encouraged to be sensitive to participants' concerns and to persevere in recruitment.

The most frequent concern related to participant burden was the lack of center staff or parent time. Day care teachers in particular were concerned about collection and storage of urine samples. Several ways of reducing participant burden were used. These included providing individual training to participants prior to the field sampling, providing assistance for urine collection at the centers, offering flexible sampling schedules, and providing a project toll-free telephone number to call for assistance. Additionally, actual contact time between staff and participants during sampling was kept as short as possible.

A major concern of some participants, especially of the directors and staff of child day care centers, was whether individual data would be released to any regulatory agency or to others. To allay this concern, a Certificate of Confidentiality for the study was obtained from the National Institute of Mental Health. This Certificate provided legal protection of the privacy of the individual data. Under this Certificate, the study researchers cannot and will not release any individual data to anyone, including the courts, without written permission of the individual.

To encourage participation, both monetary and non-monetary incentives were offered to participants. Participating families and child day care centers received $100 to cover their costs of providing food and other samples. If the children were to be videotaped for about 2 h, an additional incentive payment of $50 was furnished to the participating household; a $25 gift certificate for a book or other appropriate item for the classroom was provided to child care centers. At each visit to homes or centers, field staff brought small age-appropriate gifts for the participating children. Field staff encouraged participants to realize that they were performing important research, and that their participation was valuable. Participants were given a project T-shirt and pen. All participants received a framed certificate, acknowledging their contributions, at the conclusion of field sampling.

To enhance response rates in the study, user-friendly materials and brochures were developed. Letters and statements of endorsement were obtained from child care organizations, such as the National Head Start Organization, and from past pilot study participants. Press

releases prepared by the U.S. EPA describing the study were used in the selected areas, and EPA's principal investigator provided radio interviews. Prior to personal contact with centers and parents, introductory letters and brochures were sent to them by overnight courier. Multiple follow-up calls and personal visits were made by study staff to potential participants. Throughout, the study staff tried to develop a sense of a research partnership between centers, teachers, parents, and researchers.

For the initial telephone screening of potential participants, scripts were developed for interviewers, so that the screening information could be entered directly into the computer. Written consent forms for participation and for possible future contact were developed.

4.5 Recommendations

Study recruitment required far more effort and time than initially anticipated. In the future, similar studies should allocate more time and staff resources to the recruitment of participants. Recruitment should begin at least four months prior to field sampling. In addition, the problem with participant recruitment was exacerbated by the requirement that no contact could be made with subjects during the 2000 Census, which meant that some participant recruiting had to occur during the field activity phase of the study.

Overall, the recruitment methods worked well. However, several participants indicated that they should receive greater compensation for performing data collection activities that they found burdensome. In addition, increased monetary incentives should help to increase the response rates and participant cooperation.

Recommendations to improve day care center participation in future studies of this type include the following:

- Increase the compensation to day care centers, both to the center director and to the individual classroom teachers.
- Prepare a special document that would contain information to ease the concerns of the center directors. This information would address privacy issues and guarantees, compensation for time spent on the project activities, a description of day care recruitment procedures and study activities, and the assistance that would be provided by study staff.
- Design and implement a study web site that would explain the study and also provide a means for participants to ask questions.
- Increase the staff and resources for the project recruitment team, so that more intensive recruitment activities, such as follow-up visit to the day care centers, can be conducted.
- Increase the compensation to day care parents.
- Conduct additional in-depth staff training on subject recruitment and data collection activities.

- Have at least two or three staff members attend meetings with parents at the center. This would ensure full attention by the staff to all participants and minimize parents' waiting time.
- Minimize participant burden as much as possible.

Although the telephone recruiting worked very well, the advance mailings were not very effective, as about 65% of the mailed packages were returned as undeliverable. Recommendations to improve participation for stay-at-home participants in future studies of this type include the following:

- Increase the compensation to the parents.
- Mail the study brochure and introductory letter to the potential participant immediately after their initial telephone screening is completed.
- Minimize participant burden as much as possible.