

## BIOCHEMICAL EPIDEMIOLOGY IN COMMUNITY-BASED STUDIES: PRACTICAL LESSONS FROM A STUDY OF T-CELL SUBSETS

LINDA MORRIS BROWN,<sup>1,\*</sup> DAVID J. TOLLERUD,<sup>2,†</sup> LINDA M. POTTERN,<sup>1</sup>  
JEFFREY W. CLARK,<sup>1,‡</sup> RONALD KASE,<sup>3</sup> WILLIAM A. BLATTNER<sup>1</sup>  
and ROBERT N. HOOVER<sup>1</sup>

<sup>1</sup>National Cancer Institute, Bethesda, MD 20892, <sup>2</sup>University of Cincinnati Medical Center, Cincinnati, OH 45221 and <sup>3</sup>Westat, Inc., Rockville, MD 20850, U.S.A.

(Received in revised form 7 October 1988)

**Abstract**—Elaborate laboratory tests are increasingly being incorporated into traditional epidemiologic research designs, a concept commonly termed biochemical epidemiology. Some of the issues encountered are illustrated by a recent population-based survey of healthy individuals in the Washington, D.C. area designed to examine the effects of demographic characteristics, lifestyle, and medical conditions on peripheral blood T-cell subsets. The study was conducted in three phases: selection of households by random digit dialing (Phase I); telephone interviews (Phase II); and self-administered questionnaires and phlebotomy (Phase III). Although this design facilitated the selection of the final study population, it influenced the participation rates by offering opportunities for nonresponse at each phase. Race was the strongest determinant of response rate despite the use of highly-trained, racially-matched telephone interviewers and repeated attempts at refusal conversion. Also discussed are issues of confidentiality, and logistics of biologic specimen collection and handling. The difficulties encountered in this survey are examined, with suggestions for future population-based investigations involving biochemical epidemiology.

Epidemiology    Biochemical epidemiology    Population-based studies  
Community-based studies    T-cell subsets

### INTRODUCTION

Technological advances in cellular immunology, genetics, and molecular biology have opened a new era of investigation for chronic disease epidemiologists: biochemical epidemiology [1]. The incorporation of laboratory measurements into traditional, rigorous epidemiologic designs should greatly enhance investigations of disease causation and environmental effects on human populations [2]. It is now possible to carry out population-based investigations to test hypo-

theses generated in the laboratory. However, the theoretical advantages of using a population-based sample needs to be weighed against the practical constraints of collecting biologic specimens from a representative sample of the general population. Population-based studies present significant logistical and methodological difficulties in population selection and enrollment, biologic specimen handling, and confidentiality and quality control which are not encountered in studies utilizing nonrepresentative groups such as blood donors, laboratory "volunteers", and hospitalized patients. The practical realities that epidemiologists must face in conducting and interpreting population-based biochemical epidemiologic studies have received relatively little attention.

The present study was designed to examine the effects of demographic characteristics, lifestyle, and medical conditions on peripheral

\*Address all correspondence to: Dr L. M. Brown, Epidemiology and Biostatistics Program, Executive Plaza North, Room 415C, National Cancer Institute, Bethesda, MD 20892, U.S.A.

†Present address: Channing Laboratory, Brigham and Women's/Beth Israel Department of Medicine, Boston, MA 02115, U.S.A.

‡Present address: Hospital Division of Oncology, Brown University, Providence, RI 02912, U.S.A.

blood T-cell subsets in a sample of healthy adult smokers and nonsmokers. Previous results of T-cell subsets in "normal" subjects have largely come from small or highly selected populations such as laboratory "volunteers" or blood donors. Since T-cell subsets might be affected by characteristics of the population (i.e. age, sex, race, health status, and use of cigarettes, and drugs) we attempted to choose a random sample of healthy subjects from the general population stratified on age, race, sex, and smoking status as a way to obtain a representative study sample with less opportunity for selection bias.

In this paper, we present the study protocol as an example of a population-based investigation in which delicate biologic specimens are collected and preserved for sophisticated, state of the art laboratory analysis. In addition to the usual preparations for a population-based epidemiologic study, this investigation required logistical planning and preparation for specimen collection, handling and storage, and careful coordination of laboratory and field activities. The difficulties encountered in such an effort are discussed, with suggestions for future investigations.

#### STUDY PROTOCOL

From May 1984 through September 1984, a population-based survey of the Washington, D.C. area was conducted in three phases: Phase I: selection of households and potential study subjects by random digit dialing (RDD); Phase II: telephone interviews and selection of study subjects medically eligible for phlebotomy; and Phase III: self-administered interviews, phlebotomy, and selection of final study sample. The study protocol was designed to obtain questionnaire data and blood specimens on a stratified random sample of approximately 300 "healthy" adults aged 20-69 years from the Washington, D.C. area. Subjects who met exclusionary criteria (medical conditions or lifestyle characteristics suspected to significantly alter peripheral blood T-cell subsets [3]) in either Phase II or Phase III were not included in this sample.

The selection of the study population involved stratification by age, sex, and race to ensure a sufficient number of subjects in each strata. In addition, subjects were stratified by smoking status to facilitate investigation of the role of cigarette smoking. Data on smoking and T-cell subsets are limited and controversial [4-6].

After reviewing Washington, D.C. Standard Metropolitan Statistical Area (SMSA) census data and national statistics on cigarette smoking, it was determined that the RDD effort would have to be doubled in order to enroll a sufficient number of older black women smokers. Likewise, the population size of other minority groups in the Washington, D.C. area were judged too small to allow adequate RDD sampling for a study of this magnitude. Therefore, for financial and logistical reasons, black smokers and members of other racial groups were not included in the pool of potential study subjects. It was decided *a priori* that analysis of cigarette smoking effects would be limited to whites, whereas analysis of demographic and other factors would be performed on non-smoking blacks and whites.

#### Phase I

RDD using the Waksberg method [7] was used to select a random sample of households based on telephone exchanges which approximated the Washington, D.C. SMSA. To reduce the chance of selecting two study subjects from the same household, thus minimizing the possibility of intraclass correlation, households were given a sex designation based on the last digit of their telephone number. Only males were eligible to participate in male-designated households and only females in female-designated households [8].

A 5 minute household screening questionnaire was administered from a telephone center by trained interviewers to ascertain eligible study subjects aged 20-69 years. Obtained were the name, sex, age, and race of each person in the household and the smoking status of those aged 14 years or older. A total of 3888 households completed the household screening questionnaire for a response rate of 83.3%. Table 1 presents a comparison of basic demographic characteristics for households who responded to the telephone screening compared to 1980 Census data for the Washington, D.C. SMSA [9]. The percentages from the telephone survey by race and sex, age group, and number of persons in the household were similar to those estimated from Census data.

The goal of the recruitment effort was to obtain valid questionnaire and blood data for ten respondents in each of 30 age-sex-race-smoking strata. Sampling fractions were established for each of these strata by applying national smoking rates by age, race, and sex to

Table 1.  
Phase I te  
for th

Race and  
White  
White  
Black  
Black  
Not as  
Total

Age group  
< 15  
15-24  
25-34  
35-44  
45-59  
60-69  
70+  
Not as  
Total

Number  
in house  
1  
2  
3  
4  
5  
6+  
Total

the p  
D.C.  
rates  
perce  
medi  
oped  
were  
subj  
eligi  
men  
app:  
tere  
age,  
hou  
of t  
Pha

Phu  
I  
ave  
wa  
of  
sel  
po  
th  
a

Table 1. Comparison of demographic characteristics of the Phase I telephoned household sample with 1980 Census data for the Maryland-Washington, D.C.-Virginia SMSA

	Study		Census (%)
	No. of persons	(%)	
<i>Race and sex</i>			
White male	3308	34.0	34.6
White female	3413	35.1	36.3
Black male	1400	14.4	13.6
Black female	1585	16.3	15.5
Not ascertained	19	0.2	—
Total	9725	100.0	100.0
<i>Age group</i>			
<15	1845	19.0	21.3
15-24	1739	17.9	18.7
25-34	1878	19.3	19.5
35-44	1640	16.9	13.9
45-59	1420	14.6	15.4
60-69	724	7.4	6.5
70+	462	4.7	4.7
Not ascertained	17	0.2	—
Total	9725	100.0	100.0
<i>Number of persons in household</i>			
	<i>No. of households</i>		
1	869	22.3	25.6
2	1221	31.4	28.7
3	737	19.0	17.6
4	636	16.3	15.4
5	302	7.8	7.6
6+	123	3.2	5.1
Total	3888	100.0	100.0

the population distribution of the Washington, D.C. SMSA and applying estimated response rates for each phase of the study and estimated percentages of the population expected to be medically eligible. Sampling worksheets developed on the basis of these sampling fractions were used by the interviewers to select the study subjects as a stratified random sample of all eligible household members. All household members greater than 19 years of age and of the appropriate household sex designation were entered on the sampling worksheets according to age, sex, race, and smoking status. The 1043 household members entered on a "white space" of the worksheets were selected to participate in Phase II of the study.

#### Phase II

If the respondent selected in Phase I was available, a 25 minute telephone questionnaire was administered immediately after completion of the household screening questionnaire. If the selected respondent was not available, an appointment was made to call back and administer the questionnaire. A total of 849 subjects out of a possible 1043 (81.4%) completed the tele-

phone questionnaire. The percentages of subjects responding and not responding did not vary significantly by age or sex. However, a significantly smaller percentage of blacks than whites (73.2 vs 85.5%,  $\chi^2 = 21.125$ ,  $p < 0.0001$ ) answered the questionnaire. The race-sex specific response rates were 82.9, 88.1, 71.5, and 74.5% for white males, white females, black males, and black females, respectively.

The purpose of the questionnaire was twofold: (1) to collect basic demographic, occupational, smoking, beverage, and medical information; and (2) to determine medical eligibility for participation in Phase III of the study. Background information was collected on sex, ethnicity, ancestry, education, number of years in the Washington, D.C. area, marital status, date of birth, weight, and height. Both usual and current occupation were ascertained and coded using 1980 Census codes [10]. The codes for usual occupation were then grouped into five categories: professional and managerial, technical, saleswork and clerical, skilled and semiskilled, and unskilled to be used as a measure of socioeconomic status (SES). Detailed information was collected on the intensity and duration of use of cigarettes, cigars, pipes, and chewing tobacco. Measures of intensity and duration were also collected for the consumption of caffeinated and decaffeinated coffee, tea, cola, wine, beer, and hard liquor. The medical section included information on pregnancy history (females only), medical exclusionary conditions (listed below) which have been shown to alter T-cell subsets [3], and other conditions and medications which might affect immune status.

Two hundred forty six (29.0%) of the respondents who completed the telephone questionnaire met at least one of the medical exclusionary criteria. Subjects were considered to be medically ineligible for participation in Phase III of the study if they had: (1) been hospitalized overnight or undergone surgery requiring general anesthesia during the past 3 months; (2) been treated by a doctor for an allergy (either shots or a prescription) during the past 3 months; (3) taken steroid medication prescribed by a doctor during the past year; (4) received blood transfusion since 1975; (5) been told by a doctor that they had systemic lupus erythematosus (SLE), rheumatoid arthritis, hemophilia or a serious bleeding disorder, or cancer; or (6) were pregnant now or in the past 3 months. Since many individuals met more than one of these exclusionary criteria, Table 2

Table 2. Reasons for medical ineligibility—Phase II

Reason	Number of subjects
Steroid use	72
Allergy	64
Rheumatoid arthritis	41
Hospitalization	37
Blood transfusion	36
Cancer	34
Anesthesia	25
Pregnant	17
Bleeding disorder	5
SLE	4

\*The total is not provided in the table because it exceeds the number excluded for at least one reason by about 36%. Persons reporting more than one exclusionary criteria were enumerated in each appropriate medical reason category.

presents the number of individuals who reported each medical reason.

### Phase III

The 603 Phase II respondents determined to be medically eligible were sent a letter inviting them to complete a short questionnaire and have a blood specimen taken by a skilled nurse at a mobile van parked at a hospital in their area. Subjects were told that they would be paid \$25.00 for giving blood and an additional \$10.00 for transportation. Also mailed to the subject was a medication history form to record the names of all medications (both prescription and nonprescription) taken in the past 30 days. A follow-up telephone call was made to each subject to answer any questions and to schedule an appointment for blood drawing at one of 11 designated locations. A bonus of \$15.00 was offered to respondents who came to the van on Sundays.

The number of study subjects who agreed to give blood was 401 or 66.5% of the 603 considered medically eligible. The percentages of those agreeing to give blood were similar to those refusing by age group, sex, education, marital

status, and whether or not they were currently employed. Blacks gave blood significantly less often than whites (50.8 vs 73.8%,  $\chi^2 = 30.992$ ,  $p < 0.0001$ ). White males gave blood most often (75.1%) followed by white females (72.3%), black males (57.3%), and black females (45.1%). There was a significant difference in response rates by usual job classified in one of six SES categories when the total study population was compared ( $\chi^2 = 18.821$ ,  $p = 0.0021$ , but not when whites ( $\chi^2 = 6.544$ ,  $p = 0.2568$ ) and blacks ( $\chi^2 = 3.135$ ,  $p = 0.6757$ ) were considered separately (Table 3).

The final participation rate for the study is the product of the response rates achieved at each of the three phases. For example, the final participation rate for all subjects combined would be 83.3% (Phase I)  $\times$  81.4% (Phase II)  $\times$  66.5% (Phase III) or 45.1%. The final participation rates for white males, white females, black males, and black females were 52, 53, 34, and 28%, respectively.

In the van, study subjects submitted the medication questionnaire, signed a study consent form, and completed a 15 minute two-part self-administered questionnaire. The first part of the questionnaire obtained recent smoking, alcohol, and dietary habits and recent medical history. The medical questions were asked to discover whether any of the study subjects previously determined to be eligible for phlebotomy now met any of the medical exclusionary criteria screened for in Phase II. The second part of the questionnaire collected personal lifestyle information concerning illicit drug use and homosexual activity. This information was used to exclude from the final study population subjects whose medical conditions or lifestyle practices might affect T-cell subsets. We did not ask subjects whether they had AIDS nor did we test blood to see if it was HIV+ because of the extremely sensitive nature of the question and the potential legal ramifications.

Table 3. Distribution by occupation and race for Phase III-eligible respondents who consented to phlebotomy

Occupation	Total (N = 401)		Whites (N = 304)		Blacks (N = 97)	
	N	(%) <sup>a</sup>	N	(%) <sup>a</sup>	N	(%) <sup>a</sup>
Professional/management	167	(71.1)	142	(75.1)	25	(54.4)
Technical/sales	50	(69.4)	36	(73.5)	14	(39.1)
Saleswork/clerical	97	(64.2)	69	(71.1)	28	(51.8)
Skilled/semiskilled	30	(83.3)	24	(88.9)	6	(66.7)
Unskilled	49	(53.8)	29	(69.1)	20	(40.8)
Unemployed	8	(44.4)	4	(50.0)	4	(40.0)

<sup>a</sup>Percentages participating in each occupation group.

In order to protect the confidentiality of the respondent, the second part of the questionnaire was separated from the main questionnaire and sealed in a plain envelope upon completion. The only identifying information was a coded identification number. Because of the potentially incriminating nature of the second part of the questionnaire, a certificate of confidentiality providing protection against prosecution for study subjects supplying information concerning illicit drug use was obtained through the National Institute on Drug Abuse (NIDA).

Blood was drawn by a registered nurse with extensive phlebotomy experience for routine clinical laboratory testing (blood chemistries, a complete blood count with differential, and a cortisol test) and for testing on the fluorescent antibody cell sorter (FACS) to determine T-cell subsets. Subjects were instructed to lie down and extend their arm, palm up, and straight at the elbow. A tourniquet was applied and the skin was swabbed with prepodyne and alcohol and then dried. A butterfly needle was inserted into a vein and secured with a small piece of tape. Approximately 100 ml of blood were drawn, (60 using a syringe prepared with 3 ml of preservative free heparin and the remainder using a vacutainer holder and adapter).

In addition to drawing the blood, the nurse was responsible for filling out the laboratory background and receipt forms, labelling and storing the tubes of blood, and preparing the blood for transport to the laboratory. A small tube of anticoagulated blood for hematologic analysis required refrigeration until laboratory testing 12-18 hours later. Tubes for serum collection required centrifugation soon after phlebotomy (using a small table-top centrifuge installed in the van) and subsequent refrigeration. All tubes were carefully packed with

freezer packs in styrofoam containers for transport to local laboratories. Heparinized blood for immunologic analyses was mixed with stabilizing media in transfer packs and maintained at room temperature. The packs were transported to a local laboratory in styrofoam containers.

Of the 401 subjects who came to the van and had blood drawn, one subject was excluded from further analysis due to unsuccessful phlebotomy and three due to laboratory errors resulting in loss of routine hematology data. Laboratory tests were conducted on the blood of the 397 remaining subjects. Twenty-three of the study subjects who completed phlebotomy were considered ineligible because they met previously described exclusionary criteria during the interval between telephone interview and phlebotomy (hospitalization, 3 subjects; allergy, 1 subject; steroids, 2 subjects; and pregnant, 2 subjects) or they reported homosexual activity (11 subjects) or intravenous drug use (4 subjects) in the confidential portion of the questionnaire. Thus, the final study population consisted of 374 subjects with valid blood results (Table 4).

*Quality control procedures*

Quality control procedures began with the careful hiring and training of staff. Field activities were conducted by Westat Inc., an experienced survey research firm and were supervised by the NCI principal investigators. Staff from Westat's telephone center monitored 10% of all telephone work as well as reviewed all sampling worksheets and questions on medical eligibility. All questionnaires for the first part of the self-administered interview were carefully edited and data retrieval was completed either at the time of phlebotomy or through the telephone center. There was no editing or data retrieval for the questions concerning sexual preference and

Table 4. Number of study subjects with valid blood results and percentage of Phase III eligibles by age group, race, cigarette smoking status and sex

Age Group	White						Black				Total	
	Nonsmoker		Smoker		Nonsmoker		Nonsmoker					
	Male n (n/N%) <sup>a</sup>	Female n (n/N%)	Male n (n/N%)	Female n (n/N%)	Male n (n/N%)	Female n (n/N%)	Male n (n/N%)	Female n (n/N%)	n	(n/N%)		
20-29	12 (60.0)	16 (72.7)	11 (57.9)	8 (72.7)	17 (60.7)	8 (33.3)	72	(58.1)				
30-39	23 (74.2)	16 (64.0)	9 (47.4)	19 (67.9)	8 (36.4)	14 (53.9)	89	(58.9)				
40-49	25 (86.2)	15 (71.4)	12 (63.2)	10 (83.3)	9 (47.4)	10 (52.6)	81	(68.1)				
50-59	23 (82.1)	16 (69.6)	12 (66.7)	10 (50.0)	7 (70.0)	8 (42.1)	76	(64.4)				
60-69	14 (70.0)	14 (70.0)	9 (64.3)	8 (61.5)	5 (50.0)	6 (42.9)	56	(61.5)				
Total	97 (75.8)	77 (69.4)	53 (59.6)	55 (65.5)	46 (51.7)	46 (45.1)	374	(62.0)				

<sup>a</sup>n = number of subjects with valid blood results; N = Total number of subjects who answered telephone questionnaire and were eligible for phlebotomy.

illicit drug use because of their sensitive nature. Operations in the mobile van were supervised by the Westat study manager and NCI investigators every day for the first week and periodically thereafter. A field notebook was kept which recorded the monthly van schedule, the daily schedule of respondents, and the total number of blood specimen tubes and transfer packs sent to the laboratories each day. For each respondent, a subject checklist monitored each step of data collection and a phlebotomy checklist recorded the number of tubes of blood drawn and any associated problems. Daily inventories were maintained by the laboratories receiving the blood and compared with records maintained by Westat.

#### DISCUSSION

The goal of this study was to apply elaborate laboratory assays of cellular immunology to specimens obtained from a carefully selected population-based, stratified random sample from a large metropolitan area. The study illustrates three major methodologic concerns which are encountered in any study of this type: (1) population selection and recruitment; (2) logistics of proper specimen collection and handling; and (3) confidentiality and quality control of the information collected.

Random digit dialing provided an efficient method for selection of households in an area where 97% of households were estimated to have a telephone [11]. The Phase I household screening questionnaire administered by telephone to respondents allowed immediate ascertainment of eligible household members. The response rate for this phase of the study was approximately 83%. Based on the results of the subsequent phases, it is likely that this rate was higher for whites than for blacks, although the race-specific rates are unknown. This response rate to a telephone screener is lower than that achieved in a number of other studies, but is similar to several more recent efforts in urban populations in the north and east [8]. The poor response rates in these urban areas suggest that population-based studies might be conducted among more "cooperative" populations if their use can fulfill study goals.

The 81% response rate for the Phase II telephone questionnaire was achieved through the use of experienced, carefully trained interviewers. Racially matched interviewers were used for refusal conversion attempts in difficult

cases. In spite of these efforts, significantly more blacks than whites declined to answer the questionnaire. Socioeconomic and cultural factors (including a distrust of programs sponsored by the Federal Government) may account for the apparent hesitancy of black subjects to participate in this research survey. Identification of these determinants could be important to studies designed to address health issues in the black community.

It was anticipated that the phlebotomy phase of the study would present the greatest challenge to achieving a high overall response rate. Use of a mobile van parked at multiple locations minimized the burden on the respondents while capitalizing on the efficiency of a centrally located facility for interviews, phlebotomy, and processing. Hospitals were selected as parking locations for the van because they tended to be well-known landmarks and were easily accessible. In addition, medical support was available had there been any unforeseen emergencies during the course of phlebotomy. Study subjects were seen by prearranged appointments, and reminder cards were sent several days in advance of the appointment. A mobile telephone in the van was used to give last minute reminders, to give further encouragement to "no-shows", and to reschedule appointments.

Remuneration of study subjects for the inconvenience involved was considered a necessary adjunct to attract healthy subjects to enroll in the phlebotomy phase of the study. Blacks were significantly more likely than whites to decline phlebotomy. However, when blacks and whites were considered separately, participation did not appear to be related to socioeconomic status as characterized by occupation or education. It is not known whether offering more remuneration would have achieved a substantially higher phlebotomy response rate.

Our study design utilized three phases to facilitate the selection of the final study population. The response rate for each phase of the study, particularly among whites, was in the range generally viewed as acceptable by epidemiologists. This is somewhat deceptive, however, since the overall participation rate is the product of all three rates (approx. 50% for whites). It is not known whether the resulting total participation rate using alternative designs involving only two or even a single phase would be meaningfully higher.

The trade-off in reducing the number of phases is an increase in the complexity of the

survey with an accompanying increase in cost and opportunity for error. For example, a reduction to a single phase would have required a dwelling-based sampling scheme with in-person canvassing and simultaneous progression to interview and phlebotomy of eligibles. This would have required either the use of persons with skills in each of these areas, or the simultaneous use of several persons in teams. Also, it would have required initially visiting around 4000 households. Multiple return visits would have been necessary for residences where either no one was home at the time of the initial visit or the eligible person was not at home. In addition, it is likely that this approach would have made the use of the mobile van impractical, forcing phlebotomy and initial specimen processing to be carried out in the less-controlled home environment. In the current circumstances, the additional complexities would have been substantial, and the cost prohibitive.

These difficulties in participation are undoubtedly not unique to this investigation, since most studies involving biologic specimen collection from members of the general population conspicuously avoid documenting cooperation rates; and those that do, report rates comparable to our own [12]. Furthermore, response problems appear to be increasing over time as documented in well-established, periodic national surveys. For example, the participation rates for the National Health and Nutrition Examination Surveys (NHANES) initiated in the 1970's were approx. 74% compared to approx. 90% for the Health Examination Surveys (HES) conducted throughout the 1960's [13]. This drop was due to nonresponse at the interview and examination phases since the screener response rates remained near 100%. An important question that epidemiologists must face is whether low response rates in population-based studies are more meaningful than higher response rates for less rigorously selected nonrepresentative populations. While it can be argued that a population-based study with a low participation rate may actually not provide a representative sample of that population, the mere ability to calculate response rates is a significant improvement over series of "volunteers", blood donors, and hospitalized patients where the source population is unknown.

For this study, it was essential to collect accurate data on homosexual activity and illicit drug use in order to identify individuals with potential immunologic abnormalities. Several

measures were taken to assure confidentiality. A "Certificate of Confidentiality" was obtained which stated official assurances by the U.S. Government that all information provided by study participants was confidential and exempt from legal proceedings. In addition, all sensitive information was collected on self-administered questionnaires with detachable face sheets containing all personal identifier data. The face sheets were physically separated from the remainder of the questionnaire in the presence of the study subject and kept in a secured locked file with access limited to the principal investigators. Although the sensitive questions were answered by all subjects, it was impossible to judge the accuracy of the responses.

The logistics of biologic specimen collection, handling, and processing present a special challenge for studies involving biochemical epidemiology, particularly those investigations seeking to apply state of the art laboratory techniques requiring delicate cellular materials. Close cooperation and communication between the laboratory and survey teams is required, since specimen procurement and laboratory processing must be well coordinated. If specimens cannot be preserved, time-consuming laboratory analyses may be the rate-limiting step in the study, with field schedules dependent on laboratory availability.

This paper only deals with problems encountered in biochemical epidemiologic investigations among members of the general population. Collecting biochemical data from groups of diseased persons, such as participants in case-control studies, presents an array of additional problems including: the validity of collecting biologic specimens from diseased persons, obtaining permission from physicians, and scheduling biologic specimen collections so they will not interfere with treatment regimens. In the current study, three types of samples were collected, each requiring different handling and processing procedures. Detailed protocols were developed for specimen transport and storage, and for documenting receipt of samples by the various participating laboratories. Specimen and document tracking systems were devised to record completion of each phase of processing and analysis.

Biochemical epidemiology requires an unprecedented degree of cooperation and understanding between laboratory scientists and epidemiologists. The Principal Investigator(s) must be thoroughly familiar with the laboratory

assays and specimen handling requirements, as well as the techniques of classical epidemiology, data management, and statistical analysis. Conclusions must be tempered by assessing the strengths and limitations of both laboratory and epidemiological methods used to address the specific study hypotheses. If this balance can be achieved, biochemical epidemiology may provide a powerful tool for exploring biological relationships and biological risk factors in population-based studies.

*Acknowledgements*—The authors wish to thank Barbara Greenberg and Linda Lannom of Westat, Inc. for their contributions in the conduct of the study, and Nancy Bayless of Westat, Inc. for computer support.

#### REFERENCES

1. Glass RI. New prospects for epidemiologic investigations. *Science* 1986; 234: 951-955.
2. Hoover RN. Hormonal, infectious, and nutritional aspects of cancer of the female reproductive tract. *Biochem Molec Epidemiol Cancer* 1986; 32: 313-322.
3. Schlossman SF, Reinherz EL. Human T-cell subsets in health and disease. *Springer Semin Immunopathol* 1984; 7: 9-18.
4. Ginns LC, Goldenheim PD, Miller LG, Burton RC, Gillick L, Colvin RB, Goldstein G, Kung PC, Hurwitz C, Kazemi H. T-lymphocyte subsets in smoking and lung cancer. Analysis by monoclonal antibodies and flow cytometry. *Am Rev Resp Dis* 1982; 126: 265-269.
5. Miller LG, Golstein G, Murphy M, Ginns LC. Reversible alterations in immunoregulatory T cells in smoking. Analysis by monoclonal antibodies and flow cytometry. *Chest* 1982; 82: 526-529.
6. Burton RC, Ferguson P, Gray M, Hall J, Hayes M, Smart YC. Effects of age, gender, and cigarette smoking on human immunoregulatory T-cell subsets: Establishment of normal ranges and comparison with patients with colorectal cancer and multiple sclerosis. *Diag Immun* 1983; 1: 216-233.
7. Waksberg J. Sampling methods for random digit dialing. *J Am Stat Assoc* 1978; 73: 40-46.
8. Hartge P, Brinton LA, Rosenthal JF, Cahill JI, Hoover RN, Waksberg J. Random digit dialing in selecting a population-based control group. *Am J Epidemiol* 1984; 120: 825-833.
9. U.S. Department of Commerce. **1980 Census of Population. Characteristics of the Population.** Washington, D.C.: U.S. Government Printing Office; 1983: Chap. C, Part 10, pp. 10-13, 10-84.
10. U.S. Department of Commerce. **1980 Census of Population Code Book. Alphabetical Index of Industries and Occupations.** Washington, D.C.: U.S. Government Printing Office; 1981.
11. U.S. Department of Commerce. **Census of Population and Housing, 1980. Summary Tape File 3C. SMSA:** Washington; D.C.-MD-VA.
12. Rawls WE, Clarke A, Smith KO, Docherty JJ, Gilman SC, Graham S. Specific antibodies to herpes simplex virus type 2 among women with cervical cancer. *CSH Conf on Cell Proliferation* 1980; 7: 117-133.
13. Winn D. National Center for Health Statistics. Personal communication 14 October 1987.