Request for OMB Approval of Extension of the

Atherosclerosis Risk in Communities Study (ARIC)

Supporting Statement Part B

Table of Contents

B. Collections of Information Employing Statistical Methods

List of Tables

B. Collections of Information Employing Statistical Methods

B.1. Respondents Universe and Sampling Methods

B.1.a. Design Summary

ARIC is both a prospective epidemiologic study and a community surveillance study designed to investigate the etiology and natural history of atherosclerosis and its clinical sequela. The prospective epidemiological component, called the Cohort component, examines and follows a sample of approximately 4,000 men and women in each of four communities. The Community Surveillance component identifies from hospital and death records a sample of all MIs and CHD which occur in all age eligible residents in the community. For events occurring after 2005 the age range for MI and CHD is extended to 84 and hospitalized heart failure information is identified on all community residents over age 55.

B.1.b. Respondent Universe

ARIC is conducted in four geographically distinct communities: Forsyth County, North Carolina; Jackson, Mississippi; Minneapolis suburbs, Minnesota; and Washington County, Maryland. Each community has recruited approximately 4,000 men and women between the ages of 45 and 64 at visit 1 in 1987-89. The cohort in Jackson, Mississippi is sampled and recruited to have an all black population. The population (Table B.1.b.1) and social and economic (Table B.1.b.2) characteristics of the communities are summarized in the following tables. In the Community Surveillance component, sampling percentages are based on specific diagnostic codes. As described in greater detail in section B.2.b., the study obtains a weighted sample on all fatal and non-fatal cases of MI and CHD in <u>all</u> residents aged 35-84.

ARIC was designed to collect data in four diverse communities. This design was chosen so that data could be obtained for groups that differed by geography, race, and socioeconomic status. Each community provides information on the occurrence and trends in CHD in a unique environmental setting. The cohort samples were drawn from each community so that inferences about association between risk factors and disease can be made from diverse population groups. The diversity of the groups permits evaluation of the consistency of any observed association. Thus, it was important to maximize the diversity rather than attempt to obtain a random sample of the United States. It was also important to select communities in which identification, repeated examinations, and follow-up of a cohort would be possible and linkage between CHD occurring in the community and the cohort could be made. ARIC was not designed to select either a random or representative sample of the entire U.S. population.

Table B.1.b.1. Population Characteristics, 2000

Study Community	Total	Ages 35-84
Forsyth County, North Carolina	306,067	153,330
Jackson, Mississippi	184,256	79,398
Minneapolis suburbs, Minnesota	240,797	120,031

Washington County, Maryland	131,923	69,100
Total	**863,043**	**421,859**

Table B.1.b.2. Social and Economic Characteristics, 2000

Community	% Black	% Urban	% Educ 12+	Median Income ($)
Forsyth County, NC	26	75	82	42,097
Jackson, MS	42	100	79	30,414
Minneapolis suburbs, MN	1	100	85	56,846
Washington County, MD	9.1	57	83	51,034

B.1.c. Cohort Sampling

As described in the original submission to OMB for this study, a probability sample of each community was conducted to select persons eligible for the Cohort component of ARIC. Though the sampling techniques were not the same (list and household samples), each method is designed to yield representative samples of each community. The number of persons in each community and actual Visit 1 clinic attendance is shown in Table B.1.c.1.

Table B.1.c.1. Numerical Estimates and Sample Respondents

Study Community	Number in Community Ages 45-64	Number Sampled and Attending Clinic Visit
Forsyth County, North Carolina	50,424	4,035
Jackson, Mississippi (Black)	11,480	3,728
Minneapolis, Minnesota	36,546	4,009
Washington County, Maryland	24,146	4,020
Total	**122,596**	**15,792**

Response rates at each stage of sampling and recruitment have been determined for Visit 1. Shown in Table B.1.c.2. are the response rates for household enumeration, completion of home interview, completion of a clinic examination and overall recruitment for each of the four communities. As can be seen, the response rates vary by community.

Table B.1.c.2. Response Rates to Stages of Recruitment, Visit 1

Recruitment Stages	Forsyth County	Jackson City	Minneapolis Suburbs	Washington County
	Percent Response			
Household Enumeration	97	81	91	82
Home interview	80	81	83	92

Clinic completion	67	46	72	79
Overall recruitment	65	37	66	65

	Total Number of Participants Seen in Each Clinic:			
Participants	4,035	3,728	4,009	4,020
	Total Number of Participants Seen:			**15,792**

       The poorer response rate among the black population in Jackson was expected and noted in the original OMB submission. Because there is a serious lack of information on CHD in blacks, the NHLBI regards the inclusion of a black cohort to be essential in ARIC despite the higher non-response. Every effort has been made to minimize these non-response rates (see section B.3.). Early in ARIC, the Jackson community began using a list sample, adopting the methods used in Minneapolis and Washington County. This change, from a household sampling method, permitted effort to be put into recruitment of eligible persons and not on the inefficient listing and enumeration of ineligible households (i.e. white households).

       Response to the annual follow-up telephone call has been very high (between 94.2 and 99.9 percent). Overall participation at Visit 2 was 93 percent, at Visit 3 was 86 percent and at Visit 4 was 81 percent. Although Jackson has lower response rates for clinic visits, it has a high response rate for the annual follow-up. At contact year 16, the response rate at Jackson was 97.6%. We anticipate to maintain this level of response rates for this proposal.

Table B.1.c.3. Response to Annual Follow-up and Clinic Visits through Contact Year 19, by Field Center

| | Percent Response | | | | |
	Forsyth County	Jackson City	Minneapolis Suburbs	Washington County	Total
Contact Year 2	99.5	98.5	99.9	99.8	**99.4**
Contact Year 3	99.2	98.4	99.7	99.7	**99.3**
Contact Year 4	98.8	99.0	99.8	99.7	**99.3**
Contact Year 5	98.7	98.4	99.3	99.1	**98.9**
Contact Year 6	97.8	97.3	98.5	98.8	**98.1**
Contact Year 7	97.4	97.8	98.1	99.0	**98.1**
Contact Year 8	98.0	96.5	96.9	98.5	**97.5**
Contact Year 9	97.5	96.5	96.5	98.2	**97.2**
Contact Year 10	94.2	96.5	96.2	97.5	**96.1**
Contact Year 11	94.7	96.2	95.3	96.2	**95.6**
Contact Year 12	93.9	97.0	94.9	97.2	**95.7**
Contact Year 13	91.6	97.3	93.8	97.0	**94.9**
Contact Year 14	91.6	96.7	93.0	96.6	**94.4**
Contact Year 15	91.4	97.4	92.2	96.2	**94.2**
Contact Year 16	90.9	97.6	91.6	96.0	**93.8**
Contact Year 17	89.7	97.9	91.2	95.0	**93.3**
Contact Year 18	88.0	97.7	90.2	94.8	**92.4**
Contact Year 19	86.2	97.2	89.4	94.2	**91.5**

B.1.d. Sample Size Requirements

For community surveillance, it is required that the combined communities be large enough to detect a two percent annual changes in the incidence of definite fatal CHD and definite nonfatal MI over a nine year period. Using $\alpha=.05$ and $\beta=.20$, the populations need to be large enough to generate 1120 cases per year. This requires a combined community population, aged 35-74 of 227,000 persons. In actuality, ARIC will have a combined population, aged 35-74 of 279,000 persons. For the extended age group 75-84 years in CHD surveillance, 350 new cases per year in four communities would be needed for an estimated 3% annual change in trend assessment. For hospitalized heart failure surveillance, 950 first heart failure cases per year would be needed to for an estimated 3% annual change in trend assessment. The combined population is sufficient to ensure a statistical power of 80 percent.

Secondly, for community surveillance, it is required that each community be large enough to detect the same two percent annual decline over a 10 year period. To achieve this, each community would need to be large enough to generate 350 new cases per year. This requires each community to have a population, aged 35-74 of 48,000. In actuality, all of the ARIC communities, except Washington County, exceed this 48,000. A lower population size was permitted for Washington County because mortality from CHD is much higher in Washington County than in the United States (36 percent higher for men, either percent higher for women). Thus, a sufficient number of cases will be generated. For the extended age group 75-84 years, approximately 60-110 CHD cases are needed for estimated three percent annual change, depending on community size. For hospitalized heart failure surveillance, 190-290 cases are needed for an estimated three percent annual change, depending on community size. All of the ARIC communities have sufficient population size for the planned analyses.

The sample size requirements for the sampled cohort populations were as follows:
- Cohort populations in each community should permit calculating reasonably precise estimates of the proportion of cohort events missed by surveillance (and the number of events which surveillance falsely diagnosed as definite CHD).
- The combined cohort size should provide enough new events in three years for the prospective evaluation of the effects of risk factors.
- Cohort population in each community should be limited to the number of fasting participants in a single clinic that can examine in a three year period (six each working day = 4,000 in three years).
- A combined cohort size of 16,000 men and women aged 45-64 would be expected to include 15,086 participants free of CHD at entry (based on Framingham rates) and generate 471 CHD events in a three year follow-up. Even if incidence rates have declined since Framingham, there would be an adequate number of events for evaluating effects of the more important risk factors in three years. Additional statistical power would be obtained from using ultrasound diagnosis as a dependable variable.
- Excluding angina, 279 new CHD events would be expected in three years of follow-up. The number of events available for validating surveillance (estimated by adjusting 279 for (1) the difference between Framingham and current rates and (2) the number of new events occurring among persons with prevalent CHD at entry) is 286.
- For the four cohorts, the expected number of events for validating surveillance is 71 in each (286/4). If surveillance missed 25 percent of the cohort events, 95 percent confidence limits around this estimate would be $\pm$ 10 percent. Confidence limits around the same estimate in the combined communities would be $\pm$ 5 percent.
- In a complex epidemiologic study such as ARIC, sample size calculations based on a simple comparison of one risk factor for cases and controls will tend to underestimate the sample size required. Risk factor analyses are usually more complex using covariate adjustment and statistical modeling. Thus, it is also instructive to determine the yield from existing studies of varying sample sizes. From the Framingham Heart Study, OMB # 0925-0216 (5,209 men and women followed for 30+ years), the Honolulu Heart Study, OMB #925-0122 (8,006 men

followed for 12 years), and the Puerto Rico Heart Health Program, OMB #68-6444 (approximately 9,824 men followed for 12 years), there is experience showing that risk factor relationships can be detected with the expected number of events (471) occurring in ARIC over three years in the combined population.

- For the measurement of atherosclerosis by ultrasound, estimates of precision have been made. One measurement, the width of the atherosclerotic lesion as visualized in the artery, has a variance of 5.9 mm. Thus, to detect a difference of 1 mm between two groups at the baseline examination (i.e., men vs. women, or blacks vs. whites, or younger vs. older), each group must have at least 400 individuals (α=.05, β=.20). The cohort samples satisfy this requirement.

B.1.e. Statistical Analysis of Sample Design Effects

The analysis of ARIC will identify and test for any sample design effects. The possible design effects are as follows:

- One community uses cluster sampling of households.
- Three communities identify eligible households from list samples of individuals and then interview all members of the household. Each individual does not have an equal probability of selection. Where appropriate, analysis will take design effects into account by weighing and adjustment of sampling variances.

B.2. Procedures for the Collection of Information

The procedures for follow-up and for community surveillance were explained in detail in the previous submission to OMB. They will be briefly summarized here since they are continuing.

B.2.a.1. Cohort Follow-up

Annual follow-up of the ARIC cohort is used to maintain contact, to correct address information of cohort participants and to ascertain medical events between each contact. Follow-up contacts are targeted to occur within a month of the anniversary date of the original visit.

A telephone interview is conducted unless the participant cannot be reached by telephone. A home interview is scheduled instead. The questionnaire queries information on hospitalizations for illness or surgery, diagnoses, medical care and symptoms. The participant is asked about possible MI and heart failure diagnosis. Verification of address and phone number is made along with an update of the other information used to contact the participant. Every attempt is made to identify cohort participants who have died in advance of the annual contact through regular review of obituaries and death certificates.

B.2.a.2. Cohort Clinical Review and Diagnostic Classification

During the follow-up contact, the cohort participant may indicate that he or she has been hospitalized for a condition of interest to the study (CHD, peripheral vascular disease, cerebral vascular disease or heart failure). In these cases, the hospital record is identified and all relevant

information becomes part of the participant's study data. The participants have signed a medical release form allowing the study to access medical records, but often the hospitals will require a recent or hospital specific release form which the study staff obtains.

Similarly, during the follow-up contact it may be determined that the participant has died. In these cases, the death certificate is obtained from the Vital Statistics registrars and the place of death determined. For in-hospital deaths, the hospital record is reviewed as indicated above. For out-of-hospital deaths and decedents admitted without a pulse rate or blood pressure, the participant's family is contacted to provide information on the circumstances surrounding the death. The participant has given consent to contract family members regarding the participant's illness or death.

The information on hospitalizations and deaths is reviewed and a determination of the occurrence of CHD, peripheral vascular disease and cerebral vascular disease is made according to defined criteria. Heart failure diagnosis will be added to our list of determinations beginning with 2005 events. Cause of death is also determined.

B.2.b. Community Surveillance Component Design

The Community Surveillance study currently provides measures of the geographic and temporal variation of atherosclerosis and CHD in four U.S. communities and will suggest reasons for the observed patterns. In each community, the study currently obtains a complete enumeration and valid diagnostic classification of the fatal CHD and hospitalized MI in a sample of all residents aged 35-74. Beginning with 2005 events, this age increased to 84 and heart failure was added to the community surveillance events.

Community Surveillance data gathering procedures for hospitalized MI and heart failure are based on a review of hospital records of a sample of all age-eligible residents with either a diagnosis of MI or heart failure or one of several screening diagnoses who were discharged from any of the acute care hospitals in the area. Review and abstraction of hospital records is conducted by study personnel, including the filing and return of records.

The surveillance of CHD deaths is accomplished by the review and abstraction of a sample of all age and residence eligible death certificates with various manifestations of CHD coded as the underlying cause of death during the study period. An additional subset of death certificates is sampled from a group of related high yield ICD codes. Sources of validation for out-of-hospital death, and dead-on-arrivals include interviews with the next-of-kin, and personal physician, coroner or medical examiner reports, and hospital records. Deaths occurring in the hospital are classified by abstracting information from the medical record. All CHD deaths which are not positively classified by the diagnostic algorithm undergo review by a classification panel.

B.3. Methods to Maximize Response Rates and Deal with Non-response

Since the clinic visits are all completed, the procedures to maximize responses focus on maintaining high-response rates to the annual follow-up telephone calls. As has been shown in

Table B.1.c.3., ARIC has been successful in keeping participants active in this study.

The response rate of the annual follow up at contact year 16 was 93.8% over all centers (whereas 97.6% for Jackson). We anticipate the future response rates maintaining this level. To maintain high response to the annual telephone call the following procedures are done:

- Information on telephone numbers, addresses, and persons who would know the location of a participant were collected at ARIC Visit 1 and updated on subsequent contacts so that if a person has moved, ARIC can locate the participant.
- If the information for the annual follow-up cannot be obtained by telephone contact, then an interviewer will visit the household to obtain the information.

B.4. Tests of Procedures or Methods to be Undertaken

The new procedures or methods of data collection being undertaken are being pre-tested at our field centers by qualified staff. This is followed up by pilot testing with age-eligible volunteers and physician volunteers in our communities (for forms that requires physicians to complete them). The older procedures and methods of data collection have all been refined previously in the ARIC Study to minimize burden and improve utility.

B.5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Paul Sorlie, Ph.D.	NHLBI, DCVS, EB National Institutes of Health (301) 495 0456
Lloyd E. Chambless, Ph.D.	Department of Biostatistics, University of North Carolina School of Public Health (919) 962-3264

Contractors responsible for the data collection at the field centers:

Forsyth County, NC Gerardo Heiss, M.D., Ph.D. Principal Investigator	Department of Epidemiology, University of North Carolina School of Public Health (919) 966-3253, and Department of Comparative Medicine Bowman Gray School of Medicine
Jackson, MS Tom Mosley, M.D. Principal Investigator	Department of Medicine University of Mississippi (601) 984-5610
Minneapolis, MN Aaron Folsom, M.D. Principal Investigator	Department of Epidemiology University of Minnesota (612) 624-9950

Washington County, NC
Josef Coresh, M.D., Ph.D
Principal Investigator

Department of Epidemiology
Johns Hopkins School of Hygiene and Public Health
 (410) 955-0495