TABLE 2-8.

Military Status of Participants of the Followup Examination by Group

Military Status	Group			
	Ranch Hand		Comparison	
	Number	Percent	Number	Percent
Active Duty	89	8.8	118	9.1
Retired	553	54.4	683	52.8
Separated	313	30.8	420	32.5
Reserve Forces	55	5.4	65	5.0
Deceased[a]	6	0.6	7	0.5
		p=0.90		

[a]Died after the followup examination.

These data reflected the overall equivalence of the two groups in social and behavioral characteristics. The differences observed when these data were contrasted to similar data at Baseline might have reflected differences in data collection methods or slight changes in the cohorts rather than changes in behavior among group members.

LONGITUDINAL LOSSES AND GAINS

A total of 2,269 Ranch Hands and Comparisons was fully compliant with the Baseline study. The study population of 2,309 for the followup included a loss of 159 participants and the addition of 199 individuals.

Loss to the followup occurred either because the participant was deceased, refused to participate, or was unlocatable. The loss to followup was 7 percent in both the Ranch Hand and Comparison groups. Of the 69 Comparisons lost to the followup study due to refusal or inability to locate, 17 were replaced. For the remaining 52, no replacement who satisfied the health status matching criterion and was willing to participate was identified from the candidate replacements. The categories of these individuals are provided in Table 2-10. A total of 199 new participants were recruited into the study based on the selection methodology used. Information on the new participants is provided in Table 2-10.

TABLE 2-9.

Risk-Taking Behavior of Participants of the Followup Examination by Group

Activity	Group								
	Ranch Hand				Comparison				
	Yes	Percent	No	Percent	Yes	Percent	No	Percent	p-Value
Scuba Diving	103	10.1	913	89.9	160	12.4	1,133	87.6	0.09
Auto, Boat, or Motorcycle Racing	131	12.9	885	87.1	135	10.4	1,158	89.6	0.07
Acrobatic Flying	43	4.2	973	95.8	43	3.3	1,250	96.7	0.25
Sky Diving	22	2.2	994	97.8	32	2.5	1,261	97.5	0.62
Hang Gliding	11	1.1	1,005	98.9	14	1.1	1,279	98.9	1.00
Mountain Climbing	82	8.1	934	91.9	102	7.9	1,191	92.1	0.86
Surfboard Riding	81	8.0	935	92.0	91	7.0	1,202	93.0	0.40
Long-Distance Sailing	54	5.3	962	94.7	55	4.3	1,238	95.7	0.23
Fast Downhill Skiing*	170	16.7	846	83.3	184	14.2	1,108	85.8	0.10
				p=0.10					

*One Comparison was unwilling to respond.

TABLE 2-10.

Losses/Gains of Participants Between the Baseline and Followup Examinations

Losses

Number	Category
10	Ranch Hands Deceased
59	Ranch Hand Refusals
5	Ranch Hands Unlocatable
74	**Total Ranch Hands Lost**
16	Comparisons Deceased
55	Comparison Refusals
14	Comparisons Unlocatable
85	**Total Comparisons Lost**

Gains

Number	Category
39	Ranch Hands Partially Compliant at Baseline
6	Newly Verified or Located Ranch Hands
45	**Total Ranch Hands Added to Study**
61	Partially Compliant Original Comparisons at Baseline
32	Partially Compliant Replacement Comparisons at Baseline
11	Newly Selected Original Comparisons (For Newly Verified Ranch Hands)
16	Replacements for Compliant Comparisons Who Refused Followup
10	Noncompliant Original Comparisons Who Agreed to Attend Followup
11	Noncompliant Replacement Comparisons Who Agreed to Attend Followup
1	Original Comparison Not Locatable at Baseline but Found at Followup
3	Replacement Comparisons Not Locatable at Baseline but Found at Followup
9	Replacement Comparisons Not Contacted at Baseline
154	**Total Comparisons Added to Study**

SUMMARY

Participants were recruited for the first followup in accordance with the Study Protocol. All participants (Ranch Hands and Comparisons) who were contacted for enrollment at Baseline were recruited for this phase of the study. Newly verified and located Ranch Hands, since Baseline, and their respective Comparisons were invited to join the study. Due to refusals among the Comparisons, replacements from the previously uncontacted Comparisons were selected for enrollment. The replacements were matched to the refusing Comparisons on self-perception of health; health status data were obtained in the telephone survey.

Personal characteristics of the two groups were compared, based on data obtained from the followup questionnaire. Contrasts of age, educational background, religious preference, current military status, and income revealed no significant differences between the Ranch Hand and Comparison groups. Significantly more Ranch Hands smoked cigarettes at the time of the followup examination than did Comparisons, although there were no significant differences found for past history of cigarettes, cigars, or pipe use or for recent or past use of marijuana. A much higher percentage of both groups reported smoking marijuana at some time in the past at the followup than at Baseline. This difference was most likely due to a greater sense of confidentiality generated by the random response techniques used in 1985. The use of alcohol since the Baseline examination was not significantly different between the two groups. The difference in the risk-taking behavior patterns of the Ranch Hands and the Comparisons was marginally significant. Slightly more Ranch Hands than Comparisons raced motor vehicles, and more Comparisons were scuba divers.

The followup study population included the loss of 159 participants (74 Ranch Hands and 85 Comparisons) who were fully compliant at Baseline and the addition of 199 participants (45 Ranch Hands and 154 Comparisons). The 199 newly examined study subjects consisted of 132 participants (39 Ranch Hands, 61 Original Comparisons, and 32 replacement Comparisons) who were partially compliant at Baseline, 21 participants (10 Originals and 11 replacements) who refused at Baseline, and 46 participants (6 Ranch Hands, 12 Originals, and 28 replacements) who were new to the study.

Thus, the study population for the first followup of the AFHS consisted of 2,309 individuals: 1,016 who had been associated with Operation Ranch Hand and 1,293 Comparisons.

CHAPTER 2

REFERENCES

1. Greenberg, B.G., A-L.A. Abdul-Ela, W.R. Simmons, and D.G. Horvitz. 1969. The unrelated question randomized response model: Theoretical framework. J. Am. Stat. Assoc. 64(326):520-539.

CHAPTER 3

QUESTIONNAIRE METHODOLOGY

This chapter discusses the development and the implementation of the questionnaires used in the study: the participant interval questionnaire, the spouse interval questionnaire, the Baseline participant and spouse questionnaires, and the telephone survey of previously uncontacted Comparisons.

The participant interval questionnaire was designed to capture the study participant's health history in the 3 years since his participation in the Baseline study. Data collection was comparable to the Baseline effort: The questionnaire was very similar, and it was administered using the same face-to-face methodology to virtually the same population. In the Baseline study, interviews were conducted in the participants' homes and the followup interview was conducted at the physical examination site. The revised methodology was more efficient and better subject to quality control.

The spouse interval questionnaire collected reproductive data similar to those collected at Baseline from spouses for the interval since Baseline. The spouse interval questionnaires were mailed to the spouses to be self-administered, or were completed in La Jolla, California, if the spouse accompanied the participant to the physical examination site. Analysis of the spouse data is not included in this report.

Since some study subjects refused to participate in 1982 and other participants were new to the study, Baseline questionnaires were administered to these new participants and their spouses. The same procedures used at Baseline were used to administer the Baseline questionnaires in the homes of these individuals.

The elements of each questionnaire are identified in Table B-1 of Appendix B. Questionnaire development and administration and scheduling of participants were conducted by the National Opinion Research Center (NORC), a social science research center at the University of Chicago.

QUESTIONNAIRE DEVELOPMENT

The goal of questionnaire development was to maintain to the maximum extent possible the question wordings, context, and procedures that were used in the 1982 Baseline study. The largest task of questionnaire development was asking for interval histories on crucial questionnaire items to update the information provided by the 1982 Baseline questionnaires. For the participant interval questionnaire, new questions were also developed on risk factors for skin cancer, since the Baseline Morbidity Report[2,3] found Ranch Hands to have an excess of nonmelanoma skin cancer. Because the chemical constituents of Herbicide Orange had not previously been associated with skin cancer in the literature, no questions had been included in the Baseline participant questionnaire to collect information on risk factors for this condition.

New questions were added to determine personality type, since Type A behavior is associated with coronary heart disease.[4] The Jenkins Activity Scale was administered to collect these data. Enhancements were also made to improve data collection for birth defects, smoking habits, and drinking habits. A copy of the participant interval questionnaire is provided in Appendix B.

An information sheet containing a computer-generated summary of key respondent answers to the Baseline survey was used to provide bounded recall for participants. Even when given a precise "starting date," respondents frequently repeat information given earlier, neglect to report new information because they thought they had previously reported it, and otherwise misplace events in time or forget them completely. The best means of preventing such errors is through the use of bounded recall, in which the respondent is reminded of information he has already reported and new information is sought with reference to an updated information sheet. Among the data elements included were date of birth, highest educational degree, military status at last interview, marital status at last interview, and name of spouse.

The questionnaire was pretested on 8 ineligible individuals who had been interviewed during Baseline, and on 10 men who participated in the pretest examination.

INTERVIEWER TRAINING

Twelve interviewers were recruited and trained by NORC's field management and Chicago office staffs in May 1985 to administer the interval questionnaires. The onsite NORC interview staff was not informed of the exposure status of any study participant either before or after contract completion. The site supervisor reported to the Project Director in Chicago on a weekly basis, and quarterly visits were made to the site by the Director. The site supervisor observed a sample of interviews, at least one per interviewer per week, and reviewed and edited interview questionnaires before shipping them to Chicago for further processing.

In early 1985, personal interviewers were recruited to conduct Baseline interviews for new participants in their homes. The interviewers were trained in the Chicago NORC office, using questionnaires and procedures established for the Baseline survey. They were supervised by an assistant survey director in the NORC office, who edited each completed questionnaire and talked with each interviewer weekly.

TELEPHONE SURVEY

The telephone survey of uncontacted Comparisons was intended to gather data on the general health status of the 7,963 replacement candidates for the active Comparison group. The sample consisted of men who served in C-130 units in Southeast Asia between 1962 and 1971, but who did not participate actively in the Baseline phase of the study. A total of 7,411 cases (93%) was completed by NORC computer-assisted telephone interviewers. The telephone survey was conducted prior to the scheduling of the physical examinations.

The key question asked was, "Compared to other people your age, would you say that your health is...excellent, good, fair, poor?" Other questions asked about current medications, severity of illness or injury during the last 6 months, and income. Locating and refusal conversion algorithms similar to the Baseline data collection efforts were used.

The data from the telephone survey of uncontacted Comparisons were used to select a replacement whose self-reported health status matched that of the noncompliant Comparison. If a willing replacement was not found by this method, the perception of health status variable was dichotomized into excellent/good versus fair/poor, and a new replacement was selected from the Comparison set. If this second attempt at identifying a suitable replacement failed, no replacement was made. The selection procedure is provided in Figure 3-1. In this example, the first randomly ordered Comparison was contacted but refused to participate. In the second attempt, the Comparison was deceased. The third Comparison volunteered to participate in the morbidity study.

SCHEDULING OF PARTICIPANTS

NORC recruited and trained four schedulers to perform the initial contacts with study subjects. Their training included background information on the details and purpose of the study, simulation of the actual scheduling of calls, documentation of results, and conversion of refusals. An initial letter was sent by the Air Force to each study subject, informing him of the upcoming interval physical examination. The NORC scheduler then followed this letter with a call to attempt to schedule the participant.

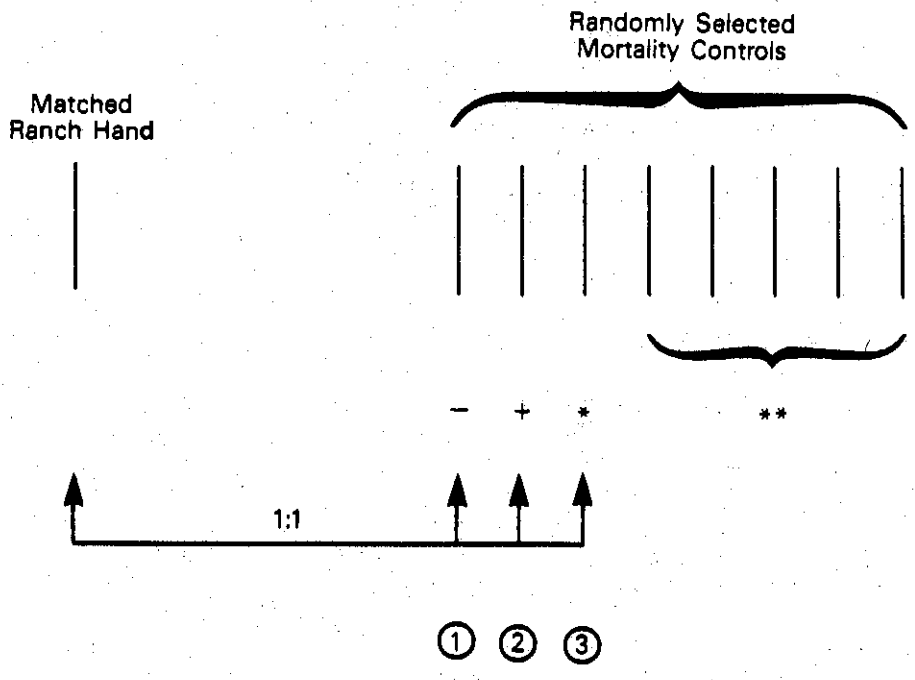
Refusals occurred at a number of steps in the scheduling process. A team of conversion specialists was assigned to contact refusing study subjects and attempt conversions. Help in conversion was also received from individuals in the U.S. Air Force School of Aerospace Medicine and the Ranch Hand Association. Many more participants were scheduled, but due to "no-shows" at the examination site, and passive refusals who rescheduled frequently, the final figure stood at 2,309.

The Baseline interviewer contacted the potential study participant by telephone for scheduling the in-home interview. Toward the end of the physical examination phase, the Baseline questionnaire was administered at the examination site by one of the interviewers who had been trained in administering that questionnaire. Of the 106 participant Baseline questionnaires administered during the first followup, 21 had to be conducted at the examination site.

The supervisor of the Baseline interviewers conducted the locating efforts for new and interval participants. Procedures similar to those used in 1982 were followed: a postal search, followed by a local telephone directory search, a motor vehicle registration search, and personal locating efforts in the area of last known residence when appropriate. The Air Force also provided locating support through its records.

DATA COLLECTION

Upon arrival at the Scripps Clinic and Research Foundation (SCRF), the participant received a schedule including the time and place for the interval interview, and a race-matched interviewer was appointed to conduct the



– Unwilling
+ Deceased
* Volunteered
** Replacement Candidates

Figure 3-1.
Selection Procedure for the Questionnaire, Physical Examination, and Followup Study

interview. Because of scheduling problems and the unavailability of a Black interviewer, 65 of the 143 Black study participants were interviewed by whites.

As in all of the personal interviews for the AFHS, interviewers were required to ask questions exactly as written, were not allowed to interpret questions or inject personal commentary, and were not allowed to skip between sections of the questionnaire. They were also instructed to probe "don't know" answers at least once. During the interview, medical record release forms were signed. The respondent was also asked to give the current name and address for each former spouse listed in the questionnaire, so that spouse questionnaires could be mailed to these individuals.

The spouse interval survey was mailed to current spouses at the time the study subject was at the SCRF. Two NORC Chicago telephone interviewers were trained to prompt refusing spouses to return the questionnaire, or to administer the spouse interview by telephone as part of the refusal conversion effort. If the spouse also traveled to La Jolla, the questionnaire was completed under the supervision of a site interviewer. Of the 1,898 completed spouse interval questionnaires, 1,066 were returned by mail, 348 were completed by telephone, and 484 were completed in La Jolla.

DATA PROCESSING

All completed interviews were sent to the NORC Chicago office following editing by the site supervisor, who retrieved missing data from study subjects while they were still onsite; any further retrieval of critical items was conducted from the Chicago office through telephone contacts. Critical items were those for which missing data were unacceptable.

The questionnaires were coded for data entry by a staff of five coders who received a week of training on the various AFHS instruments. Data entry was programmed to provide value and range checks as the data were being entered, to perform logic checks and arithmetic checks, to flag important missing items, and to verify the key entry of 10 percent of each questionnaire. Then the data were run through an automated cleaning program to detect a wide range of data errors that were corrected by pulling the hard copy questionnaires and reviewing each situation on a case-by-case basis. No changes were ever made in the hard copy data; corrections were entered into the data tape, and the tape was run against the cleaning program until no errors were detected.

CHAPTER 3

REFERENCES

1. Vitaliano, P., and F. Urbach. 1980. The relative importance of risk factors in nonmelanoma carcinoma. Arch. Dermatol. 116:454-456.

2. Stern, R.S., and K. Momtaz. 1984. Skin typing for assessment of skin cancer risk and acute response to UV-B and oral methoxalen photochemotherapy. Arch. Dermatol. 120:869-873.

3. Scotto, J., and T.R. Fears. 1978. Skin cancer epidemiology: Research needs. Natl. Cancer Inst. Monogr. 50:169-177.

4. Jenkins, C.D., R.H. Rosenman, and M. Friedman. 1967. Development of an objective psychological test for the determination of the coronary-prone behavior pattern in employed men. J. Chronic Dis. 20:371-379.

CHAPTER 4

PHYSICAL EXAMINATION METHODOLOGY

The first followup examination was provided to four categories of individuals: those who had taken the Baseline questionnaire and Baseline physical examination; those who had been invited to the Baseline events but chose not to participate, only took the questionnaire, or were unlocatable; those Comparisons who had not been invited previously, but who were selected as replacements for Baseline Comparisons noncompliant to this followup examination; and the six newly identified Ranch Hands. As noted in the Baseline Report, all potential study participants were verified as eligible for the AFHS following a detailed review of military personnel records. Replacement individuals were carefully selected, by matching data on the self-perception of health from the noncompliant Comparison (obtained from the telephone survey) with those of the replacement candidate (see Chapter 3 for details).

The followup examination differed logistically from the Baseline examination in one significant way: All structured interval questionnaires were administered at the examination site as contrasted to the in-home interviews conducted at Baseline. The followup examination consisted of the following major elements:

- Interval Questionnaire
- Combat Experience Questionnaire
- Review-of-Systems Questionnaire
- Psychological Testing
- Physical Examination
- Specialized Testing, e.g., Doppler Arterial Studies
- Laboratory Testing
- Psychological and Medical Outbriefings.

Details of the above examination elements were carefully prescribed by the Air Force and set forth as contractual requirements. Clinical innovations or variations were neither desired nor authorized; all proposed examination procedural changes were reviewed in detail by Air Force technical and contractual personnel. An important objective of the technical review was to ensure that bias was not created by any procedural change. The requirement to maintain blind examinations was particularly stringent: The clinical staff was prohibited from knowing or seeking information as to the group identity (Ranch Hand, Comparison) of any participant. At the end of the examination, each participant was asked to note on the critique form whether such information was sought by any member of the clinical or paramedical staff.

EXAMINATION CONTENT

Examination content was designed by the Air Force to emphasize detection of medical endpoints suspected of being associated with exposure to phenoxy herbicides, chlorophenols, or dioxin. In addition, findings in the Baseline examination were used by the Air Force to direct changes in the followup examination (e.g., abnormal pulses at Baseline suggested the need for Doppler measurements at the followup). The general content of the physical examination and psychological test battery is shown in Table 4-1, and the complete laboratory test series is displayed in Table 4-2.

Quality control requirements for both laboratory testing and clinical procedures were extensive. Although details are provided in Chapter 6, the following categories provide an overview of the extent of the quality emphasis. For laboratory testing, single reagent lots and control standards were used when practical, duplicate specimens were routinely and blindly retested, testing overlaps were mandatory when test reagents required change, and fast initial response cumulative statistical techniques (FIR CUSUM) were used to detect rapidly any subtle test drift over time. In addition, 50 specimens from the Baseline serum bank were retested to assess the comparability of laboratory methods. The SCRF clinical team was carefully instructed to assure clinical quality. The quality control elements included: a pretest of the examination process; detailed clinical inspection techniques by SCRF, Science Applications International Corporation (SAIC), and Air Force physicians and personnel; preprinted mark-sense examination forms; clinical quality assurance meetings to detect and correct problems; and blindness of exposure status at the examination. In addition, participant rapport-building techniques were added to boost participation in future followup studies, such as participant critique forms and recreational opportunities afforded to the accompanying family members.

CONDUCT OF EXAMINATIONS

All examinations were conducted at SCRF, La Jolla, California, from May 1985 to March 1986. Except for weeks with national holidays, two groups of participants, averaging about 32 per group, were examined weekly. Midway through the study, NORC recruiters noted that a number of participants refused the examination because of weekday business commitments or because of single-parent responsibilities. Consequently, two special weekend examinations were arranged late in the examination cycle, and many of the former refusals were then able to attend. The examination was identical to the regular 2 1/2-day process, except that it was compressed into 2 days by reducing the number of participants in a group.

The logistics effort required in contacting, transporting, and examining 2,309 study members was formidable. Preexamination contacts consisted of the telephone health survey, telephone recruitment to the examination if necessary, and calls by either the NORC scheduling specialists or by the travel agent to arrange transportation and determine whether special requirements existed (e.g., wheelchair assistance, weekend examination schedule). Once scheduling was reasonably firm, the SAIC logistics coordinator sent each participant a detailed information package outlining dietary requirements, inbriefing schedules, important telephone numbers, a request for medical records, and local maps designating examination-site eating and recreational facilities.

TABLE 4-1.

Elements of the Followup Physical Examination

Elements	Remarks
General Physical Examination	Internist
Neurological Examination	Neurologist
Dermatological Examination	Dermatologist
Electrocardiogram	Resting, 4-Hour Fasting and Nicotine Abstinence
Doppler Peripheral Arterial Blood Flow Studies	4-Hour Nicotine Abstinence
Chest X Ray	
Immunological Studies	50% Random Sample
Skin Test Studies	75% Sample
Psychological Evaluation: Minnesota Multiphasic Personality Inventory (MMPI) Cornell Medical Index Halstead-Reitan Battery	
Patient Outbriefing and Discussion of Individual Results	Medical Diagnostician, Internist, and Ph.D. Psychologist

TABLE 4-2.

Laboratory Test Procedures of the Followup Physical Examination

Clinical Laboratory

Fasting Glucose
Blood Urea Nitrogen (BUN)
Cholesterol
HDL Cholesterol
Triglyceride
Serum Glutamic-Oxaloacetic Transaminase (SGOT)
Serum Glutamic-Pyruvic Transaminase (SGPT)
Gamma-Glutamyl Transpeptidase (GGTP)
Alkaline Phosphatase
Lactic Dehydrogenase (LDH)
Thyroid Stimulating Hormone (TSH)
Initial Cortisol
2-Hour Cortisol

Prothrombin Time
Quantitative Immunoglobulins
Complete Blood Count (CBC)
Leuteinizing Hormone (LH)

2-Hour Postprandial Glucose
Creative Phosphokinase (CPK)
Total Bilirubin
Direct Bilirubin
Total Protein
Protein Electrophoresis
Routine Urinalysis
T_3 % Uptake
T_4
Testosterone
Hepatitis B Surface Antigen
Hepatitis B Surface Antibody
Follicle Stimulating Hormone (FSH)

Rapid Plasma Reagin (RPR)
Porphyrins (Mayo Clinic)
Sedimentation Rate

Immunological Laboratory

Cell Surface (Phenotype) Analyses
Lymphocyte Mitogen Stimulation Assays
Mixed Lymphocyte Culture (MLC)
Natural Killer Cell Assay by Specific Cellular Cytotoxicity Using K-562 Target Cells

Natural Killer Cell Assay (Using Interferon) by Specific Cellular Cytotoxicity Using K-562 Target Cells

The logistical flow of the entire examination process was complex. Figures 4-1 and 4-2 outline participant flow for the first 2 examination days. As depicted in these figures, each group of participants (generally containing equal numbers of Ranch Hands and Comparisons) was transported early in the morning to SCRF on the first 2 days in a fasting state; tobacco, alcohol, and coffee abstinence were also required. Following initial inbriefing and blood draw on the first day, each participant was randomly assigned to the examination group or to the psychological testing group. On the second day, these groups were reversed. After randomization, each member was given an individualized 3-day schedule outlining his medical, interviewing, and laboratory appointments. The schedule carefully noted the specific required periods of fasting and tobacco abstinence (see Figures 4-1 and 4-2 for generalized periods in relation to ECG and Doppler testing). Each individual was reminded of the fact that all aspects of the examination were strictly voluntary, and that refusals would be honored without question. Both general and specific consent forms (e.g., skin biopsy), approved by the Air Force, were explained in detail.

In contrast to the Baseline examination, great reliance was placed upon each individual to find the appropriate clinic area at his scheduled time. This approach had great appeal to this self-reliant population as evidenced by critique feedback. Throughout the examination day, generous time was provided for waiting-room activities, i.e., renewal of past friendships, discussions of the Vietnam War, consumption of refreshments when permitted, and completion of paperwork. Day 3 of the examination was largely spent in finishing up the specialty examinations and receiving the outbriefings from a psychologist and medical diagnostician. Only upon completion of these important debriefings were the participants paid their stipend, reimbursed for travel expenses, and transported to the airport.

As noted previously, the SCRF clinical team was hand-picked for participation in this project. In total, 15 board-certified physicians in internal medicine, neurology, and dermatology participated in the general, specialty, and diagnostic examination. To reduce observer variability, turnover in the clinical or paramedical staffs was minimized during the 9 months of examinations. One SCRF physician served as the Project Medical Director, responsible for the scheduling, conduct, and quality control of the examinations. All examining physicians were introduced to the mark-sense examination forms during the pretest examination. The layout of the form was designed to parallel the flow of the clinical examination so as to minimize recording errors. Because data transcription was not permitted, each physician was responsible for filling in the bubbled form. To a large extent, these mark-sense forms and subsequent quality control were the primary reason for a remarkably clean data set. Two examples of the mark-sense forms are presented as Figures 4-3 and 4-4; a complete set of forms is provided in Appendix C.

For the first followup, the special testing included Doppler tests, delayed hypersensitivity skin tests, and immunological tests. Doppler measurements were obtained on all participants by highly experienced technicians; results were recorded and Polaroid photographs were taken of representative oscilloscope displays. As previously noted, considerable emphasis was placed upon tobacco abstinence prior to Doppler evaluations. Skin tests for four antigens were administered in a standardized manner: Candida (1:1,000 weight/volume, 0.1 ml intradermal), mumps (2 complement-fixing units), Trichophyton (1:1,000 weight/volume, 0.1 ml intradermal), and

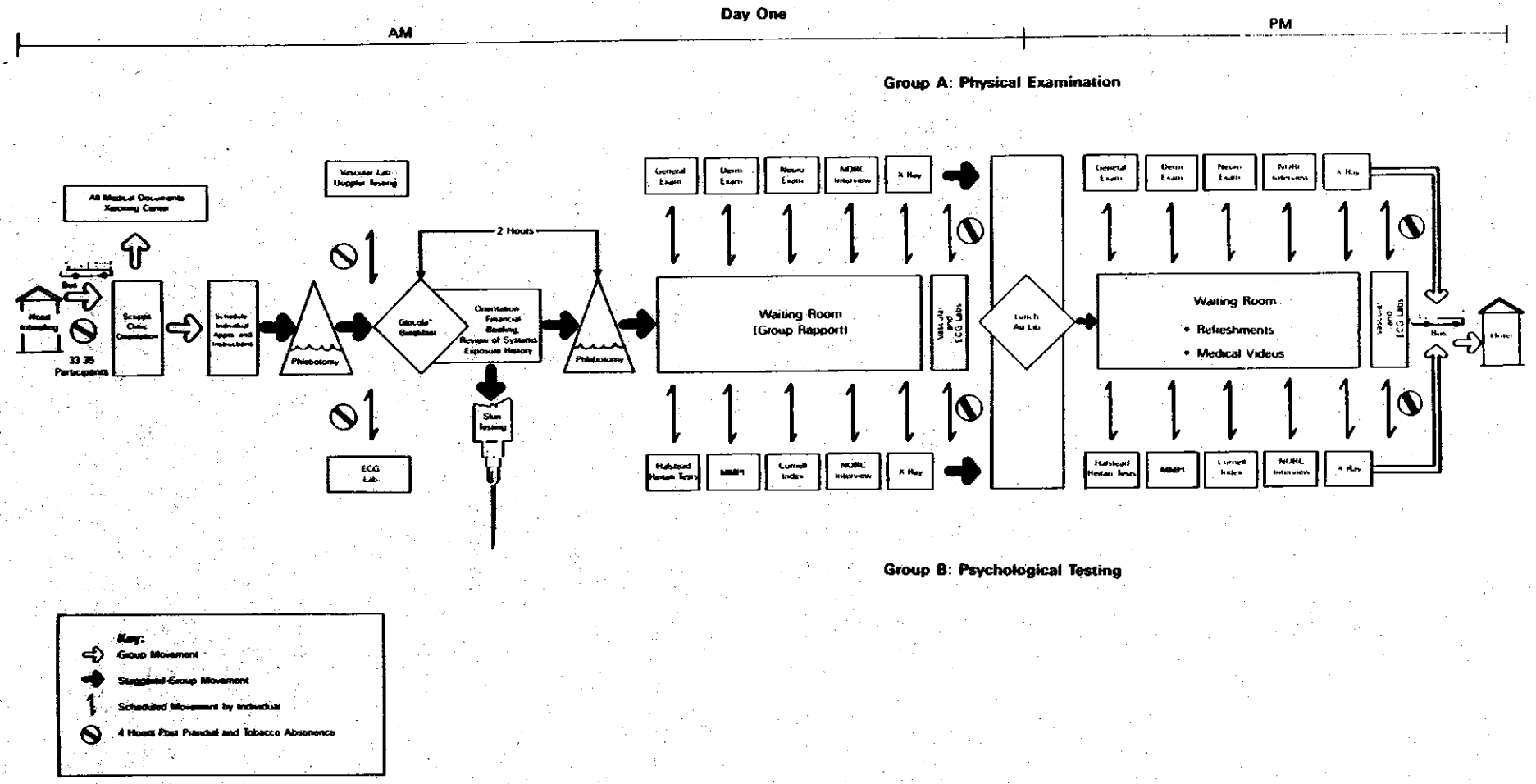


Figure 4-1.
Flow Diagram of Day One Followup Interview and Physical Examination

CHAPTER 5

STUDY SELECTION AND PARTICIPATION BIAS

INTRODUCTION AND BASELINE SUMMARY

The Protocol

During the design phase, the authors of the Protocol anticipated that loss to followup would pose the greatest threat to the validity of the study. In particular, they expected differential compliance with relatively more Ranch Hands self-selecting themselves into the study than Comparisons and with health differences of unknown character between noncompliant Ranch Hands and noncompliant Comparisons. As a partial correction, the study design specified that noncompliant Comparisons would be replaced by Comparisons having the same values of the matching variables and the same health perception. In this way, the replacement Comparisons would serve as surrogates for those Comparisons who refused to participate. This, in turn, would tend to reduce the bias due to noncompliance in the Comparison group and would have the added advantage of maintaining this group's sample size.

The Comparison in each randomized matched set who happened to be first asked to participate in the Baseline questionnaire and physical examination was identified as the Original Comparison for his respective Ranch Hand (in accordance with the Protocol). If the Original Comparison was noncompliant, that is, he refused to take the Baseline questionnaire or physical examination, his replacement was called a replacement Comparison. Replacement Comparisons were so distinguished to satisfy the Protocol requirement that they be contrasted with the noncompliant Comparisons, also called refusals, they replaced. No corresponding replacement strategy for the Ranch Hands was possible since all Ranch Hands had been identified and invited to participate.

The Protocol further specified that the replacements would be statistically compared with the noncompliant Original Comparisons to determine the extent to which the replacement strategy was being realized. The statistical contrast of replacements and refusals was to be based on responses to a noncompliance telephone questionnaire administered to refusals and to their potential replacements. This questionnaire assessed self-perception of health, days lost from work due to illness, and medication use, and was to serve as the basis for the health matching called for in the Protocol. Although the Protocol was not explicit on this point, it implied that the decision to include or exclude the replacements from the study would be based only on this contrast.

The Baseline Replacement Operation

The health-matching questions (identical to the noncompliance questionnaire) were, in fact, not administered to any potential replacement

Comparison before selection at Baseline, although questions regarding self-perception of health, medication use, and work loss were asked as part of the Baseline questionnaire after entry into the study. The noncompliance telephone questionnaire was offered to noncompliant study participants, but only 79 completed the telephone questionnaire, and of these only 57 were actually replaced. Replacements were, therefore, not health matched to refusals at Baseline. Rather, they were matched only on the basic matching variables: date of birth, race, and occupation. The statistical contrast of refusals and their replacements was not performed at Baseline.

During the scheduling operation at Baseline, two untoward events occurred that led to the identification of two additional categories of Comparisons, shifted Comparisons and Air Force-interviewed replacements. First, 212 of the Original Comparisons were discovered to be ineligible for participation in the study due to errors in the data base regarding their unit of assignment in Southeast Asia. These men had not served in Southeast Asia but, due to a duplication of codes, were mistakenly included in the Comparison population. They were deleted from the study.

This resulted in another Comparison in each previously randomized match set being first asked to participate in the study. These new Original Comparisons were figuratively called "shifted" Comparisons, labeled S in the Baseline Report, to describe the effective movement of these Comparisons in each matched set to fill the space left by the removed ineligible Original Comparison. The eligible Original Comparisons were labeled O in the Baseline report. Shifted Comparisons are more accurately referred to here as shifted Original Comparisons to emphasize that they are not replacement Comparisons and that they are the legitimate Original Comparisons for their respective Ranch Hands. Shifted Original Comparisons are not replacement Comparisons because their invitation to participate in the study was not the result of a previous refusal of another Comparison in their respective matched sets. Shifted Original Comparisons were identified to reflect concern that the process by which Comparisons were determined ineligible may not have distributed ineligible Comparisons uniformly.

Second, 30 replacement Comparisons were interviewed by Air Force staff rather than by the contractor. These replacements were labeled A. All other replacement Comparisons, labeled R, were simply called "replacements."

The removal of ineligible Comparisons from the study caused a pause in the scheduling operation that delayed the scheduling of the shifted Original and replacement Comparisons relative to that of the Original Comparisons. This scheduling delay is apparent in Figures V-3 and V-4 in the Baseline Report. Some study investigators speculated that this scheduling slip might cause shifted Original Comparisons and replacement Comparisons to self-select differently from Original Comparisons. Statistical analyses in Chapter V of the Baseline Report and further unpublished analyses following the release of the Baseline Report investigated the effect of this scheduling problem.

The Baseline Selection Bias Analyses

Since replacements were not health matched at Baseline to their corresponding noncompliant Comparisons and since differential scheduling opportunity may have created self-selection biases, statistical contrasts of the various Comparison groups were done at Baseline. In particular, the Comparisons labeled O, S, R, and A were contrasted on the basis of self-perception of health, medication use, work loss, and five clinical variables.

The results of these analyses suggested to some investigators that shifted Original Comparisons were not statistically distinguishable from Original Comparisons and that shifted Original Comparisons were not statistically different from replacements, but that replacement Comparisons appeared to be statistically different from Original Comparisons. The 30 Air Force-interviewed replacement Comparisons were not statistically distinguishable from other replacement Comparisons and were not investigated further as a group. Since opinions differed among Air Force principal investigators and statisticians, a management decision was reached to use only the Original Comparisons in the primary analyses and to contrast Ranch Hands with all Comparisons in the appendix of the Baseline Report. The reader is referred to Chapter V of the Baseline Report for additional detail. In retrospect, the concern with statistical distinguishability between replacement Comparisons and Original Comparisons is difficult to justify, since the only valid question regarding the replacements is their similarity to the refusals whom they replaced.

The Baseline Compliance Bias Analyses

Telephone questionnaire data obtained from the 57 noncompliant Comparisons, who were replaced, were not analyzed in the Baseline Report. Instead, compliance bias was analyzed by contrasting partially compliant with fully compliant participants, with adjustment for group (Ranch Hands, O, S, R, A). These analyses were based on data from the Baseline questionnaire regarding self-perception of health, medication use, work loss, anger, anxiety, erosion, depression, liver ailments, miscarriages, and acne. Results suggested that partially compliant participants were statistically different from fully compliant participants for some of these variables. Based on these results, calculations were presented to suggest that the noncompliance bias could produce an error in relative risk of 25 percent, either overestimating or underestimating the risk, and a spurious mean shift of up to 8 percent in either direction.

THE FIRST FOLLOWUP SCHEDULING AND REPLACEMENT OPERATION

Matching of replacements to noncompliant Comparisons on the basis of health status was initiated with the first followup scheduling operation. This was accomplished by administering a short telephone questionnaire to all previously uncontacted Comparisons and then using their health status responses to select from among the Comparisons in a matched set the first one who was similar to the refusal regarding self-perception of health. In addition, NORC was required to schedule replacements within 5 working days of a confirmed refusal. These features were intended to correct the described Baseline scheduling deficiencies and to bring the study into Protocol compliance regarding health matching of replacements.

To further minimize the possibility of scheduling bias, the entire study population was partitioned into 79 groups; these groups were then randomly scheduled for an examination time. In this way, no single group would be favored a priori for a certain scheduling period. The groupings, consisting of approximately 32 participants, corresponded to the examination groups established at Baseline. Group integrity was maintained to enhance study compliance and comradery. Study participants were given the option to remain with their group or to reschedule their examination at a time more convenient to them.

FIRST FOLLOWUP COMPLIANCE

Eighty-five percent (1,016/1,191) of the Ranch Hands and 81 percent (955/1,176) of the Original Comparisons participated in the first followup examination and questionnaire process. Of 288 replacements, 267, or 93 percent, chose to attend the first followup examination; additionally, 71 new replacements participated in the followup, yielding a total sample size of 338 replacements at followup. These counts and others are summarized in Table 5-1. In Table 5-1 and subsequently in this report, the shifted Original Comparisons were combined with the Original Comparisons, and the Air Force replacements were combined with the replacement Comparisons.

TABLE 5-1.

Baseline Versus First Followup Sample Sizes

	Group		
		Comparison	
Participation	Ranch Hand	Original	Replacement
Baseline Only	74	64	21
Baseline and Followup	971	872	267
Followup Only	45	83	71

Although fully compliant at Baseline, 74 Ranch Hands, 64 Original Comparisons, and 21 replacement Comparisons chose not to participate in the first followup examination. In the interim, 10 of the 74 Ranch Hands and 16 of the 85 Comparisons died. An additional 5 of the 74 Ranch Hands and 14 of the 85 Comparisons were unlocatable during the scheduling operation. There were 56 of 59 remaining Ranch Hands and 50 of 55 remaining Comparisons who refused to participate in the first followup, although they were alive and locatable during scheduling, and responded to the noncompliance telephone questionnaire, giving their reported health status and reason for nonparticipation. The 3 remaining Ranch Hands and 5 Comparisons refused to participate in the telephone survey. Reasons for nonparticipation given in the telephone survey are summarized in Table 5-2. The totals in Table 5-2 do not correspond to Table 5-1 because some participants gave more than one reason for nonparticipation.

Of the 56 living locatable Ranch Hands and the 50 Comparisons who took the noncompliance telephone questionnaire, only 35 Ranch Hands and 42 Comparisons responded to the question regarding health status. The reported health status of these 77 nonparticipants is summarized in Table 5-3.

TABLE 5-2.

Reasons for Nonparticipation in the First Followup of 56 Ranch Hands and 50 Comparisons Who Were Fully Compliant at Baseline*

Reason	Group			
	Ranch Hand		Comparison	
	Number	Percent	Number	Percent
Fear of Physical	0	0	2	4
Job Commitment	13	17	9	16
Dissatisfaction with USAF	10	13	9	16
No Time or Interest	7	9	6	11
Travel Distance, Family	13	17	12	21
Confidentiality	0	0	1	2
Health Reasons	8	11	3	5
Passive Refusal	11	15	6	11
Dissatisfaction with Baseline	5	7	2	4
Financial Hardship	3	4	0	0
Other	5	7	7	12
Total	75		57	

*Some participants gave more than one reason for nonparticipation.

TABLE 5-3.

Reported Health Status of 35 Ranch Hands and 42 Comparisons Fully Compliant at Baseline and Noncompliant at First Followup

Reported Health Status	Group			
	Ranch Hand		Comparison	
	Number	Percent	Number	Percent
Excellent	5	14	10	24
Good	22	63	22	52
Fair	6	17	8	19
Poor	2	6	2	5
Total	35		42	

p=0.72

Among the individuals responding to the health status question, there was no statistically significant difference between noncompliant Ranch Hands and Comparisons regarding reported health (p=0.72).

Further detail regarding the 45 Ranch Hands, 83 Originals, and 71 replacements newly examined at followup is shown in Table 5-4, which gives the Baseline status of these participants. Taking the questionnaire but not the physical examination at Baseline were 39 of the 45 Ranch Hands newly examined at followup. Five of the 45 Ranch Hands who were identified too late to be invited at Baseline were simply described as having had "no action" taken.

TABLE 5-4.

Baseline Status of Newly Examined Participants

	Group		
		Comparisons	
Baseline Status	Ranch Hand	Original	Replacement
Interview Only, Refused Physical Examination	39	61	32
No Interview, No Physical Examination	0	10	11
Unlocatable	0	1	3
No Action	5	11	16
Proxy	1	0	0
New to Study	0	0	9
Total	45	83	71

Of the 71 newly examined replacements, 43 (32+11) were either partially compliant at Baseline or were at least contacted at Baseline and, therefore, identified as replacements, although not health matched to a noncompliant Comparison. The remaining 28 newly examined replacements were not previously contacted. Of these, 14 were health-matched replacements and 2 were replacements added to the study in August 1985 after completion of the Baseline physical examination. Thus, of the 71 replacements who took the physical examination for the first time at followup, only 14 were new health-matched replacements. All 71 replacements may be regarded as new to the study, even though 43 had been previously contacted at Baseline and knew that they were potential study participants. The 28 replacements who had not been previously contacted may be regarded as new in a more restrictive sense since they did not know of their potential involvement in this study before they were recruited for the first followup examination. This set of 71 replacement Comparisons and the subset of 28 are distinguished from each other using

the unrestricted and restricted definitions of "new" to provide data regarding changes in replacement self-selection, an issue explored later in this chapter.

FACTORS KNOWN OR SUSPECTED TO INFLUENCE STUDY PARTICIPATION

A multitude of factors may be considered to influence self-selection. These may be broadly classified as health, logistic, operational, publicity, or demographic factors. The Baseline Report contains a list of specific factors within each of these categories. For example, health factors are thought to include self-perception of health as well as demonstrable health indicators, such as medication use and work days lost due to illness or injury. Logistic factors are thought to include distance to the examination site, reluctance to spend time away from family or job, income, and occupation. Demographic factors might include flying status, age, race, or military duty status (active, retired, separated). Operational factors include any aspect of study operation that may cause differential compliance, such as differential treatment of participants during scheduling, physical examination, interview, or debriefing. Publicity factors have to do with national attitudes and media presentations regarding the Agent Orange issue, the Vietnam war, veteran health care, or health care in general. Additionally, these considerations may affect people differently and, in particular, may influence Ranch Hands differently than Comparisons.

The decision to volunteer for this study is admittedly complex, making statistical assessment of compliance bias difficult and necessarily crude in that many of the factors contributing to self-selection cannot be measured directly. Instead, compliance bias was investigated at first followup, as in the Baseline Report. Specifically, it was investigated with respect to self-perception of health, medication use, daily aspirin use, work days lost due to illness or injury, and income in comparing partially compliant with fully compliant participants. In other selection bias assessments, such as statistical contrasts of Original and shifted Original Comparisons, these same factors and 26 variables taken from the physical examination and psychometric testing were analyzed.

THE TELEPHONE SURVEY

In April 1985, all previously uncontacted living Comparisons were identified for telephone contact to assess their current health. This health status information was necessary for the matching of replacements to noncompliant Comparisons. From a total of 9,982 available Comparisons, 7,963 were included in the telephone survey. The 2,019 nonselected Comparisons included 488 deceased, as of 1 August 1985, and 1,531 who had been previously contacted. The group of 1,531 previously contacted Comparisons comprised all Comparisons who were fully compliant, partially compliant, or noncompliant at Baseline.

The survey questionnaire is shown in Appendix D. In brief, it queried the respondent regarding self-perception of health (excellent, good, fair, poor), current prescribed medication use (yes, no), work days lost due to illness or injury, special health care needs (wheelchair, nurse, or other special equipment), and income (less than $20,000, $20,000 to $40,000, or more than $40,000). If the respondent indicated that he was taking

prescribed medication, he was asked to identify the illness for which the medication was prescribed. If work days were lost due to illness or injury, the respondent was asked to identify the causing illness or injury. If special health care or equipment was needed, he was asked to specify the illness or condition requiring the special care. He was further asked to distinguish conditions requiring special care from those that were previously identified in response to the medication and days lost from work questions. The telephone interview was accomplished via CATI.

Of the 7,963 cases fielded, 7,411 telephone surveys were actually completed. The nature of the 552 noncompletions is summarized in Table 5-5.

TABLE 5-5.

Summary of Reasons for Noncompleted Telephone Interviews

Reason	Number	Percent of 7,963
Deceased	26	0.3
Active Refusal	93	1.2
Passive Refusal	242	3.0
Unlocatable	190	2.4
Ineligible	1	0.0
Total	552	6.9

Several questionnaires that could not be administered by telephone were accomplished by mail; these numbered 540 out of the 7,411 completed. Summaries of the responses to each of the five questions are shown in Table 5-6.

Of the 1,271 respondents who reported that they had lost work days due to illness or injury, 550 (43%) lost 1 to 5 days, 197 (15%) lost between 6 and 10 days, and 524 (41%) lost more than 10 days. The maximum number of days reported lost was 965. The 56 respondents who reported more than 180 days lost misinterpreted the question; it referred only to the past 6 months.

The telephone interviewer reported whether the respondent was friendly, cooperative but not interested, impatient, or hostile. The association between the interviewer's remark and the self-reported health of the respondent was investigated. The results are displayed in Table 5-7. The association between the interviewer's remark and reported health status is statistically significant ($p=0.02$), with hostile repondents reporting poorer health than friendly, cooperative, or impatient respondents.

Other analyses of these data, not shown here, demonstrated significant associations between health perception and income ($p=0.001$), rank ($p=0.001$), age ($p=0.001$), medication use ($p=0.001$), and need for special health care ($p=0.001$). Positive health perception increased with income and rank and

TABLE 5-6.

Summary of Results to the Telephone Questionnaire

Self-Assessment of Health Compared to Others Same Age

Response	Number	Percent
Excellent	2,882	38.89
Good	3,306	44.61
Fair	972	13.11
Poor	245	3.31
Do Not Know	3	0.04
Missing	3	0.04
Total	7,411	100.00

Taking Medication for Current Illness

Response	Number	Percent
Yes	2,129	28.73
No	5,277	71.20
Refused	1	0.01
Missing	4	0.05
Total	7,411	100.00

Illness or Injury Absence From Job During Last 6 Months

Response	Number	Percent
Yes	1,271	17.15
No	6,135	82.78
Refused	3	0.04
Missing	2	0.03
Total	7,411	100.00

TABLE 5-6. (continued)

Summary of Results to the Telephone Questionnaire

Need Assistance in Daily Activities

Response	Number	Percent
Yes	114	1.54
No	7,291	98.38
Refused	4	0.05
Missing	2	0.03
Total	7,411	100.00

Earned Income From Any Job During 1984

Response	Number	Percent
Yes	6,636	89.54
No	755	10.19
Refused	17	0.23
Missing	3	0.04
Total	7,411	100.00

Income Level

Response	Number	Percent
Less than $20,000	2,015	27.19
$20,000-$40,000	3,034	40.94
More than $40,000	1,411	19.04
Not Applicable	774	10.44
Refused	161	2.17
Do Not Know	9	0.12
Missing	7	0.10
Total	7,411	100.00

TABLE 5-7.

Contrast of Interviewer's Remark from Telephone Interviews and Reported Health Status

Remark	Reported Health Status									
	Excellent		Good		Fair		Poor		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Friendly	2,209	39	2,476	44	730	13	191	3	5,606	76
Cooperative	622	38	755	46	229	14	44	3	1,650	22
Impatient	42	40	48	45	10	9	6	6	106	1
Hostile	9	21	27	63	3	7	4	9	43	0
Total	2,882	39	3,306	45	972	13	245	3	7,405	

p=0.02

decreased with age, medication use, and special health care. Further, there was no significant association between health perception and the duration of the telephone interview (p=0.17) or the time of day of the interview (p=0.98). There was no significant health-by-duration-by-time of day interaction (p=0.77).

These data were also used to assess the self-reported health of 773 Original Comparisons (excluding shifted Original Comparisons) fully compliant at Baseline relative to the reported health of the 7,411 previously uncontacted Comparisons who completed the telephone survey. The self-reported health status of the Original Comparisons from the Baseline questionnaire was contrasted with that of the previously uncontacted Comparisons on a three-category scale (excellent, good, fair/poor) with an adjustment for date of birth (born during or before 1942, born after 1942). The results are displayed in Table 5-8. Previously uncontacted Comparisons who completed the survey are indicated by T (telephone); Original Comparisons are labeled O. Data are missing for 12 Original Comparisons and 16 telephone-surveyed Comparisons.

There was no statistically significant difference between these groups regarding health perception after adjustment for age (p=0.14), and this equivalence did not change with age (p=0.80). Additionally, there was a statistically significant age effect (p=0.001), as expected. These results suggested that the Original Comparisons were representative of the entire Comparison cohort with respect to health perception.

TABLE 5-8.

Self-Reported Health of Previously Uncontacted Comparisons, in 1986, Versus Self-Reported Health Status of Original Comparisons at Baseline

		Health Perception						
		Excellent		Good		Fair/Poor		
Age	Group*	Number	Percent	Number	Percent	Number	Percent	Total
Born >1942	T	1,847	39	2,003	43	837	18	4,687
	0	203	39	239	46	83	16	525
Born ≤1942	T	1,034	38	1,298	48	376	14	2,708
	0	91	39	120	51	25	11	236

*T = previously uncontacted Comparisons
0 = Original Comparisons.

REPLACEMENT COMPARISONS VERSUS THE NONCOMPLIANT COMPARISONS THEY REPLACED

Baseline Replacement

These analyses are refinements of the analyses in Chapter V of the Baseline Report. Of 288 Comparisons replaced at Baseline, only 57 responded to the short noncompliance telephone questionnaire shown in the appendix. These 57 comprised 38 Original Comparisons and 19 replacements. As in the followup telephone survey, the short noncompliance telephone questionnaire queried respondents on health status, work days lost due to illness, medication use, and income level. In accordance with the Protocol, replacements were statistically contrasted with the noncompliant Comparisons they replaced based on their reported health status (excellent, good, fair, poor), medication use (yes, no), and income level (less than $20,000, $20,000 to $40,000, more than $40,000). This contrast, with adjustment for group membership (Original, replacement) of the noncompliant Comparison, is shown in Table 5-9.

There was no significant difference between the reported health patterns in the upper and lower panels of Table 5-9. When these two tables were merged, no statistically significant difference was found between the health status of noncompliant Comparisons and their non-health-matched replacements ($p=0.99$). It is noteworthy that 53 percent of Original and replacement noncompliant Comparisons were matched, by chance, perfectly to their replacements on the basis of reported health status. Only 7 percent (4/57) were mismatched by two categories and one replacement was mismatched by three categories.

These same groups were contrasted on medication use; the results are shown in Table 5-10.

TABLE 5-9.

Noncompliant Original Comparisons and Replacement Comparisons Versus Their Baseline Replacements: Reported Health Status at Baseline

Group	Health Status	Health Status of Replacements: Excellent	Good	Fair	Poor	Total
Noncompliant Original Comparison	Excellent	13	4	2	0	19
	Good	9	7	0	0	16
	Fair	1	1	0	0	2
	Poor	1	0	0	0	1
Total		24	12	2	0	38
Noncompliant Replacement	Excellent	7	5	0	0	12
	Good	3	3	0	0	6
	Fair	1	0	0	0	1
	Poor	0	0	0	0	0
Total		11	8	0	0	19

TABLE 5-10.

Noncompliant Original Comparisons and Replacement Comparisons Versus Their Baseline Replacements: Medication Use at Baseline

Group	Medication Use	Medication Use of Replacements: Yes	No	Total
Noncompliant Original Comparison	Yes	0	4	4
	No	3	31	34
Total		3	35	38
Noncompliant Replacement	Yes	0	1	1
	No	1	17	18
Total		1	18	19

Due to sparseness these data were not analyzed. It is interesting to note, however, that there was 82 percent agreement in the upper panel of Table 5-9 (31/38) and 89 percent in the lower panel (17/19), with 84 percent agreement in the combined table (48/57), close to expected within group percentages of 83 and 90 percent, respectively, due purely to chance.

Work loss was not analyzed due to slight differences between the way the work loss question was worded in the noncompliance telephone and telephone survey questionnaires.

The contrast regarding income level is shown in Table 5-11.

TABLE 5-11.

Noncompliant Original Comparisons and Replacement Comparisons Versus Their Baseline Replacements: Income at Baseline

Group	Income Level	Income Level of Replacements (in thousands) <$20	$20-$40	>$40	Total
Noncompliant Original Comparison	<$20	1	3	0	4
	$20-$40	6	6	3	15
	>$40	0	7	6	13
Total		7	16	9	32*
Noncompliant Replacement	<$20	0	0	2	2
	$20-$40	1	7	0	8
	>$40	1	3	5	9
Total		2	10	7	19

*Six noncompliant Original Comparisons were unwilling to respond.

The patterns of income matching in the first and second panels of Table 5-11 were not significantly different ($p>0.10$). In the combined table, replacements reported significantly lower income than the Comparisons they replaced ($p<0.05$) although 49 percent (25/51) were perfectly categorically matched.

These analyses suggested that the Baseline replacements were very similar to the noncompliant Comparisons they replaced regarding reported health status, medication use, and income. These analyses were also pertinent to the question of whether there was selection bias due to noncompliance in the Comparison group. The predominantly negative findings suggested that there was little or no Comparison selection bias. These

results suggested that the upper-bound bias calculations reported in Chapter V of the Baseline Report are overestimates of reality. However, lack of clinical data for the noncompliant Comparisons precluded refining those Baseline bias calculations at this time. Accordingly, the Baseline selection bias calculations may be viewed as crude bounds to an unknown bias that must await future data for proper recalculation.

First Followup Replacement

Replacements were matched to noncompliant Comparisons at first followup on the basis of the matching variables--date of birth, race, and occupation--and self-reported health status (excellent, good, fair, poor), as recorded in the telephone survey. This was accomplished by recording the self-reported health status of the noncompliant Comparison during the attempt to schedule and matching that status against those of the other Comparisons in the same matched set. A Comparison in a matched set was considered to replace a noncompliant Comparison if he had the same health status as that recorded for the noncompliant Comparison during the attempt to schedule him. If no willing Comparison reporting the same health status could be found in the matched set, health status was dichotomized to excellent or good versus fair or poor. A willing Comparison with the same health status as the refusal on the dichotomized scale was then accepted as a replacement. If no willing Comparison could be found using the dichotomized scale, attempts to find a replacement were terminated.

During this process, 14 Comparisons were health matched to noncompliant Comparisons. The results are summarized in Table 5-12.

TABLE 5-12.

Health Status of Refusals and Their Matched Replacements

Replacement's Health	Refusal's Health				
	Excellent	Good	Fair	Poor	Total
Excellent	1	2	0	0	3
Good	5	6	0	0	11
Fair	0	0	0	0	0
Poor	0	0	0	0	0
Total	6	8	0	0	14

All refusals reported good or excellent health. This implied that bias due to noncompliance in the Comparison group could possibly bias the study away from finding an herbicide effect. The inclusion of health-matched replacements tended to correct for this by replacing healthy noncompliant Comparisons with healthy replacement Comparisons. The relatively small number of new health-matched replacements minimized the actual effect of this bias "correction," however.

SCHEDULING AT FIRST FOLLOWUP

The schedulers were required to find and schedule a willing health-matched replacement within 5 working days of a confirmed refusal to correct scheduling differences experienced at Baseline. This constraint proved impractical to implement since Comparisons would vacillate, forcing a series of repeated telephone calls. Rather than terminate the process at 5 days, as required by the contract, the schedulers continued their recruiting attempts, sometimes for several months. Hence, new health-matched replacements were brought into the study much later than other participants.

The percent completing the physical examination by calendar date is plotted in Figure 5-1 for all Ranch Hands, Original Comparisons, and all Comparisons.

The corresponding plot for Ranch Hands, Original Comparisons, old replacements, and the 28 restricted new replacement Comparisons is shown in Figure 5-2.

Additionally, schedulers experienced reticence and vacillation with other Comparisons being scheduled for the first time. In particular, as a group, the 71 unrestricted new replacement Comparisons were also scheduled later than other participants. Figure 5-3 shows the percent of Ranch Hands, Original Comparisons, "old" Comparisons, and the 71 unrestricted newly examined replacement Comparisons completing the physical examination by calendar date.

During the scheduling for the 1987 followup examination, schedulers will attempt to schedule health-matched replacements within 15 working days of a refusal.

NEW REPLACEMENTS VERSUS OLD REPLACEMENTS

Another statistical issue of concern is the homogeneity of the replacement Comparisons. The validity of the study might be compromised if, for example, newly admitted replacements had self-selected themselves into the study differently than previously admitted replacements. This kind of difference may occur due to changes in public opinion regarding the Agent Orange issue, the national political climate, changes in national opinion regarding health care, changes in the location of the examination site, or a combination of these and other factors. This issue was addressed by comparing new with old replacements on a variety of endpoints with adjustment for the matching variables. Blacks were deleted from the analyses.

Two separate series of analyses were performed, one for each of the two kinds of new replacements (unrestricted and restricted) defined earlier. First, unrestricted new replacements were identified as the 71 replacements who were examined for the first time at first followup, regardless of their compliance at Baseline. Second, analyses were restricted to the 28 replacements who were examined for the first time and who had never been contacted before the first followup; these were called restricted new Comparisons. In each of the two series of new replacement analyses, all replacements not satisfying the definition of "new" are included by referring to them as "old" replacements. All "old" replacements were at least contacted at Baseline and were fully compliant at first followup.

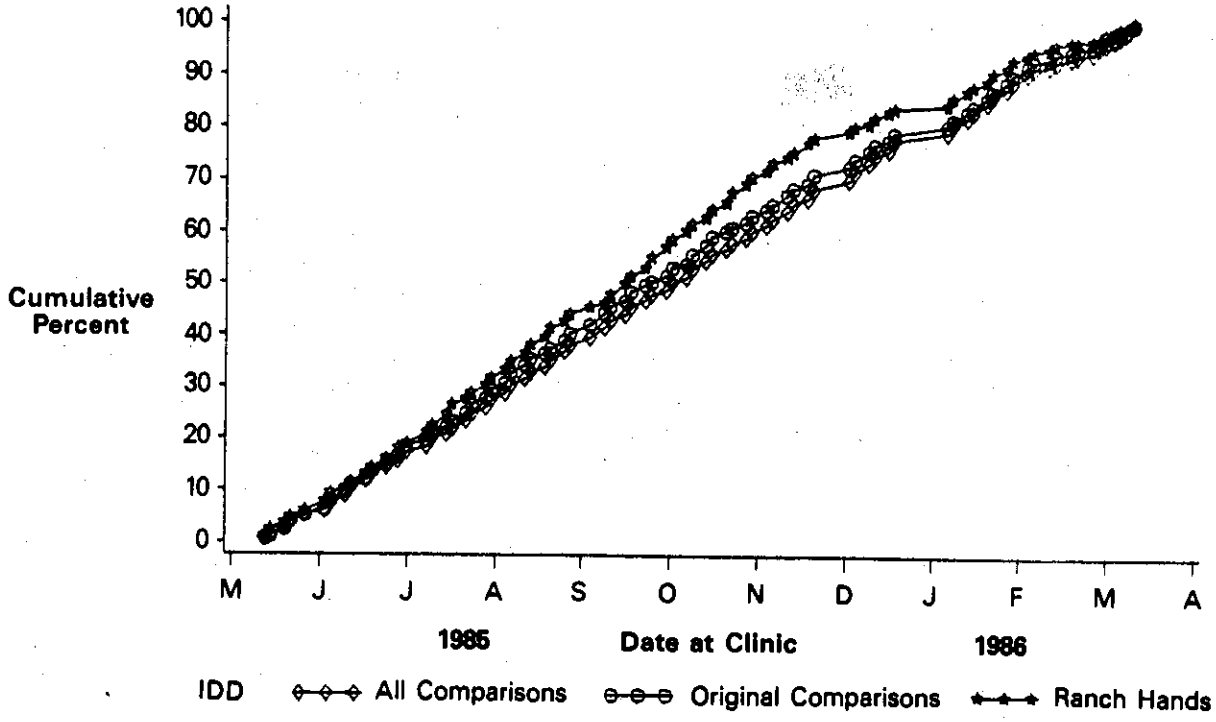


Figure 5-1.
Percent Completed Physical Examination by Calendar Date for All Comparisons

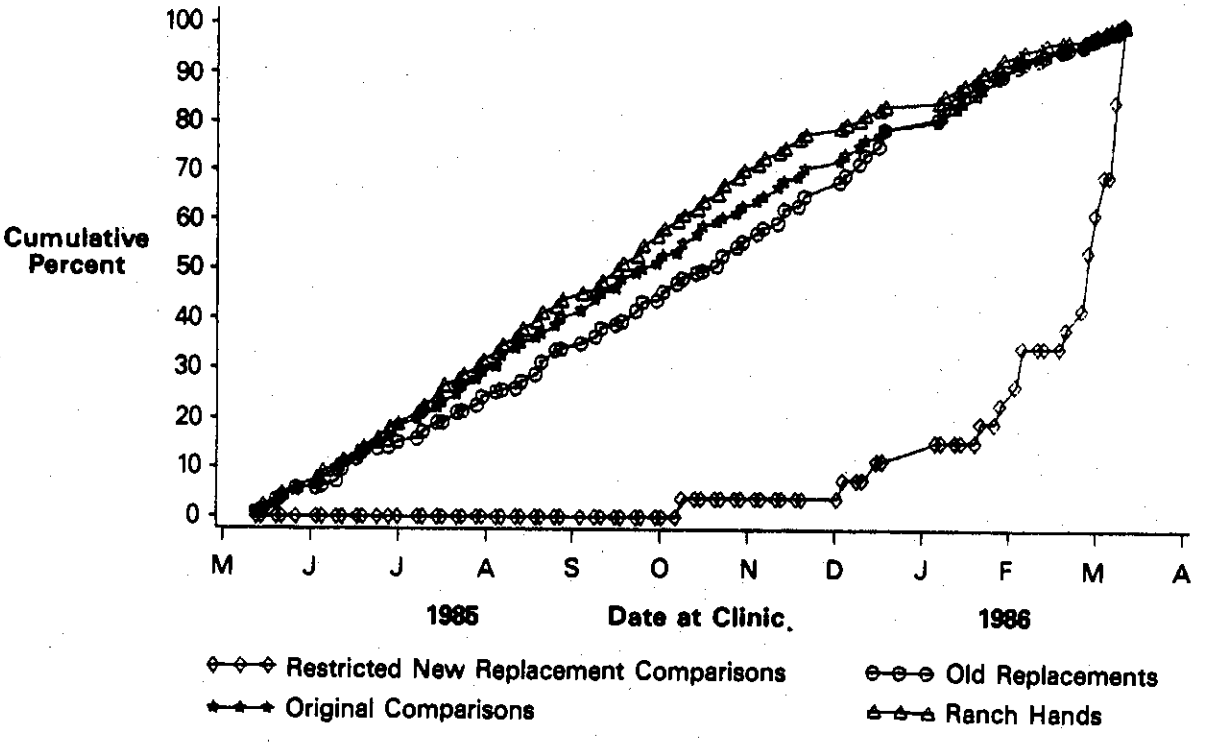


Figure 5-2.
Percent Completed Physical by Calendar Date

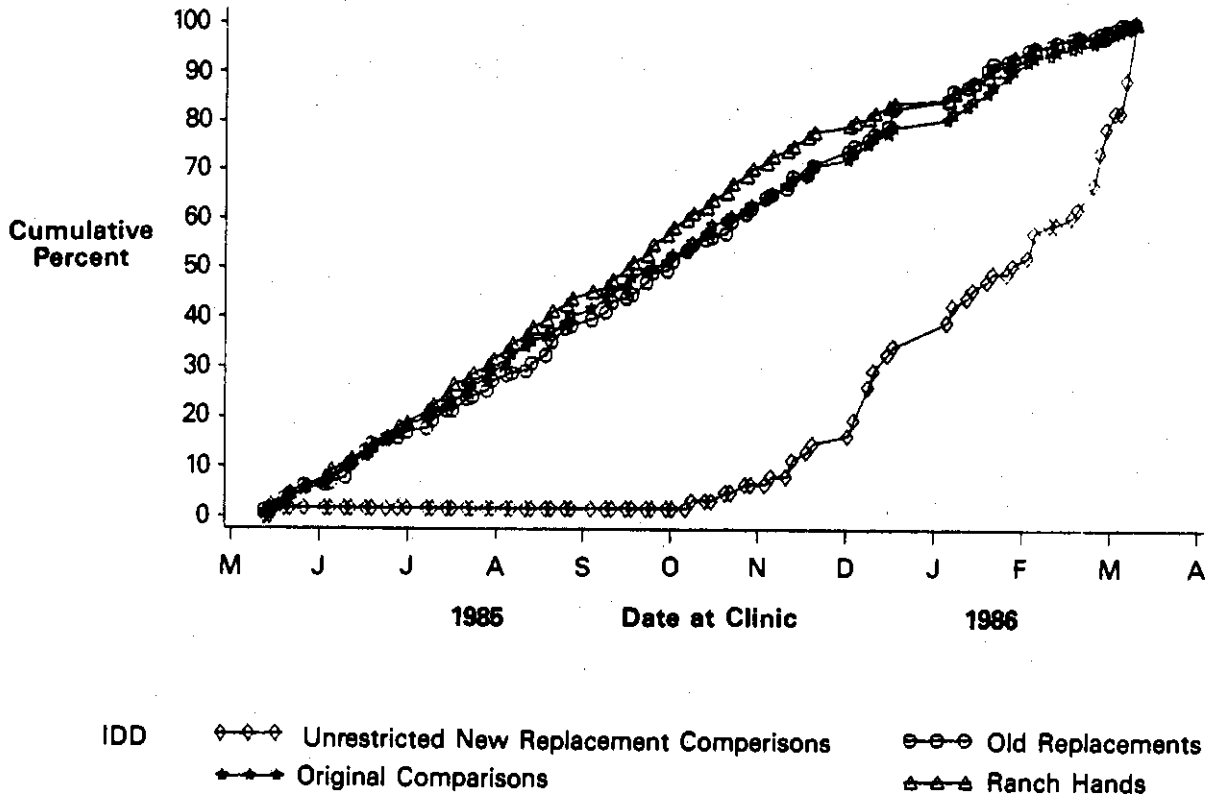


Figure 5-3.
Percent Completed Physical Examination by Calendar Date for Unrestricted New and Old Replacement Comparisons

In each of the two series of analyses, new and old replacement Comparisons were contrasted on health perception (excellent, good, fair, or poor), medication use (yes, no), work loss (yes, no), and daily use of aspirin (yes, no). Blacks were deleted from all analyses. New and old replacements were then contrasted on 20 clinical determinations from the first followup examination. Table 5-13 shows two cross-classifications of 313 nonblack replacements, from a total of 338 replacements fully compliant at first followup, by group (old, new) and reported health status.

In the unrestricted sense, the reported health status of new and old replacements differed significantly ($p=0.04$), with new replacements reporting more fair or poor health than old replacements. In the restricted sense, the difference between new and old replacements was statistically significant ($p=0.001$), with new replacements tending to declare themselves of fair or poor health more often than old replacements.

The same groups were contrasted on medication use; the results are shown in Table 5-14. The difference between old and new Comparisons under the unrestricted definition was not statistically significant (p=0.16) as regards medication use. The difference between old and new Comparisons under the restricted definition was, however, statistically significant (p=0.003). This difference was due to the higher reported medication use of the 26 non-black new replacements not previously contacted.

New and old replacements were contrasted on work loss due to illness; the results are shown in Table 5-15.

TABLE 5-13.

Reported Health Status of Nonblack New and Old Replacements, According to Two Definitions of "New"

	Unrestricted				Restricted			
	Old		New		Old		New	
Health	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Excellent	142	56	30	49	161	56	11	42
Good	91	36	20	33	103	36	8	31
Fair/Poor	19	8	11	18	23	8	7	27
Total	252		61		287		26	
	p=0.04				p=0.001			

TABLE 5-14.

Reported Medication Use of Nonblack New and Old Replacements, According to Two Definitions of "New"

	Unrestricted				Restricted			
	Old		New		Old		New	
Medication	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Yes	30	12	12	20	33	11	9	35
No	222	88	49	80	254	89	17	65
Total	252		61		287		26	
	p=0.16				p=0.003			

TABLE 5-15.

Reported Work Loss of Nonblack New and Old Replacements, According to Two Definitions of "New"

	Unrestricted				Restricted			
	Old		New		Old		New	
Work Loss	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Yes	47	19	12	20	54	19	5	19
No	205	81	49	80	233	81	21	81
Total	252		61		287		26	
		p=0.99				p=0.99		

The difference between new and old replacements regarding work loss under the unrestricted or restricted definition was not statistically significant (p=0.99 and p=0.99, respectively).

Results of a similar contrast on daily aspirin usage are shown in Table 5-16. The difference between new and old replacements regarding daily use of aspirin under the unrestricted or the restricted definition was not statistically significant (p=0.99 and p=0.75, respectively).

It is noteworthy that the differences for general health and medication use did not occur for work loss and daily aspirin usage, suggesting that some participants may have over-reported when asked less specific questions about their health.

New and old replacement Comparisons were also compared on 20 clinical and psychometric variables measured during the physical examination and psychological testing. These 20 variables are a subset from 26 selected from among an entire collection of nearly 200 endpoints in this study by requiring near statistical independence within and between organ systems. Variables selection was accomplished by screening the correlation matrices of variables as an entire set and separately within each organ system, including examining partial correlations between single variables and linear combinations of other variables within organ systems. Identified first were 10 variables with pairwise correlations less than 0.10 in absolute value. This was followed by identification of 16 additional variables with pairwise correlations between 0.10 and 0.20 in absolute value, making a total of 26 variables. These variable selection screens were accomplished on Baseline data for 1,154 nonblack fully compliant Comparisons subsequent to publication of the Baseline Report. The complete set of 26 dependent variables selected as

TABLE 5-16.

Reported Daily Aspirin Usage of Nonblack New and Old Replacements, According to Two Definitions of "New"

	Unrestricted				Restricted			
	Old		New		Old		New	
Aspirin Usage	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Yes	182	73	44	72	206	72	20	77
No	69	27	17	28	80	28	6	23
Total	251		61		286		26	
	p=0.99				p=0.75			

nearly statistically independent is shown in Table 5-17. The Baseline correlation matrix of these 26 variables as determined on the entire Comparison data set is shown in Table D-1 of Appendix D. It is recognized that relative statistical independence of these variables does not imply biological independence of these variables.

These 26 variables were intended to serve as the basis for statistical contrasts of Original Comparisons, shifted Original Comparisons, and replacement Comparisons in the decision regarding the inclusion of shifted Original Comparisons and replacement Comparisons in the primary analyses. Generically, the analyses first compared two groups on each of the 26 variables with adjustment for rank (officer, enlisted), age at Baseline (40 or under, over 40), occupation (officer flyer, officer nonflying, enlisted flyer, enlisted groundcrew), and race (Black, nonblack). Blacks were deleted from the analysis. The total number of significant differences on the first set of 10 dependent variables was used as the basis for a decision regarding group difference. These 10 analyses were assumed to be 10 independent repetitions of a Bernoulli trial with probability of 0.05 of success under the null hypothesis that there were no group differences for any of the 10 variables. The probability of observing three or more successes in 10 independent repetitions of a Bernoulli trial, with probability of 0.05 of success, is 0.012. The entire set of 26 analyses was then assessed to test the hypothesis of group equality. The probability of 4 or more successes in 26 independent repetitions of a Bernoulli trial, with probability of 0.05 of success, is 0.039. These 2 critical values, both probabilities below 0.05, were used to assess the analyses on the 10 and on the 26 selected variables.

TABLE 5-17. Twenty-Six Dependent Variables Selected as Nearly Statistically Independent With the Use of Baseline Data

Variables Having Pairwise Absolute Correlations Less Than 0.10

- Total Bilirubin (TBILI)
- Diastolic Blood Pressure (DBP)
- White Blood Cell Count (WBC)
- Skin Index (SKIN)
- MMPI Depression Scale (MMPID)
- Blood Urea Nitrogen (BUN)
- Urine Specific Gravity (USG)
- Pulse Index (PULSE)
- Nerve Conduction Velocity Above the Elbow (NCVE)
- Semen Count (SEMEN)

Variables Having Pairwise Absolute Correlations Greater Than 0.10 and Less Than 0.20

- Red Blood Cell Count (RBC)
- FEV1/FVC (PULM)
- Glucose (GLUC)
- Electrocardiogram (ECG)
- Platelet Count (PLAT)
- Full IQ (IQ)
- Central Nervous System Index (CNS)
- Nerve Conduction Velocity Above the Ankle (NCVA)
- Cholesterol (CHOL)
- Alkaline Phosphatase (ALKPHOS)
- Coproporphyrins (COPRO)
- Delta-Aminolevulinic Acid (ALA)
- Thyroid T_4 (T4)
- Testosterone (TEST)
- Sedimentation Rate (SED)
- Gamma-Glutamyl Transpeptidase (GGTP)

The statistical issue of how to account for the many interactions in the 26 separate analyses was not resolved during or since the first application of this method. Only the group main effect was regarded as the basis for determining whether a particular analysis was a success.

At first followup, only 20 of the 26 variables were measured. The six variables not measured were the two-nerve conduction velocities (NCVE, NCVA), semen count (SEMEN), FEV1/FVC (PULM), full IQ (IQ), and delta-aminolevulinic acid (ALA). New and old replacements were contrasted on each of the remaining 20 variables via the general linear model and log-linear model. The variables--skin index (SKIN), pulse index (PULSE), electrocardiogram (ECG), and central nervous system index (CNS)--were analyzed as dichotomous variables, with each being scored abnormal if any of its components were abnormal. All others were analyzed as continuous variables. The correlation matrix of the 20 variables, based on 1,210 nonblack Comparisons fully compliant at first followup, on first followup data is shown in Table D-2 of Appendix D.

The results of these analyses contrasting new versus old replacements with "new" following the unrestrictive definition and Blacks removed from the analyses are shown in Table 5-18. There were 61 nonblack new replacements and 251 nonblack old replacements. In some analyses, the dependent variable was transformed to better approximate normality. Unadjusted means are presented when there is a significant interaction involving group.

The probability of observing 2 or more successes in 8 independent repetitions of a Bernoulli trial, with probability of 0.05 of success, is 0.057. In view of the results for the first 8 dependent variables in Table 5-18, new and old replacements appeared to be statistically indistinguishable. The probability of observing 3 or more successes in 20 independent repetitions of a Bernoulli trial, with probability 0.05 of success, is 0.075; the probability of 4 or more is 0.016. Recognizing the slight correlations between the dependent variables in the lower panel of Table 5-18, and the results of the analyses, new and old replacements again appeared to be statistically indistinguishable.

The same analyses were conducted to contrast new and old replacement Comparisons, with "new" defined in the restrictive sense. The results are shown in Table 5-19, with the same notations as Table 5-18.

The same binominal critical values, 2 for the first panel and 4 for the entire set of 20 analyses, and the results shown in Table 5-18 indicated that there was no statistical difference between the 26 nonblack new replacements and the 287 nonblack old replacements.

The negative findings shown in Tables 5-18 and 5-19 suggested very strongly that there has been no change in the way replacements self-select for entry into this study.

ORIGINAL COMPARISONS VERSUS SHIFTED ORIGINAL COMPARISONS

The removal of ineligible Comparisons early in the Baseline scheduling operation resulted in the exclusion of approximately 18 percent of all Comparisons from the study. Since some of these ineligibles had been randomized as Original Comparisons, some previously randomized Comparisons were allocated to the positions vacated by the removed original Comparisons and, thus, were referred to as shifted Original Comparisons.

TABLE 5-18.

Summary Results of Unrestricted New Versus Old Nonblack Replacements Contrasted on 20 Variables

Variable (Transformation)	Replacement Group Means* (Percent Abnormal) Old	New	p-Value	Significant Interactions
Variables With Absolute Pairwise Correlations Less Than 0.10				
TBILI (LOG)	0.76	0.76	NS	
DBP (SQRT)	79.17	79.51	NS	
WBC (LOG)	7.06	7.13	NS	
SKIN	(54.0)	(49.2)	NS	
MMPID (LOG)	56.21	57.19	NS	
BUN (SQRT)	14.15	13.79	NS	
USG	1.014	1.014	NS	
PULSE	(16.7)	(11.5)		GRP*OCC, GRP*AGE
Variables With Absolute Pairwise Correlation Between 0.10 and 0.20				
RBC	5.00	5.00		GRP*OCC*AGE
GLUC (LOG)	109.31	101.33	NS	
ECG	(15.5)	(13.1)	NS	
PLAT (SQRT)	269.5	275.0	NS	
CNS	(2.8)	(5.0)	NS	
CHOL (SQRT)	212.7	208.8	NS	
ALKPHOS (LOG)	87.9	87.10		GRP*OCC
COPRO (SQRT)	116.9	122.6	0.03	
T4	7.51	7.94	NS	
TEST (SQRT)	601.4	605.3	NS	
SED (LOG)	4.17	4.93		GRP*OCC*AGE
GGTP (LOG)	31.06	29.77		GRP*AGE

*All means are expressed in original units.

NS: Not significant (p>0.05)
LOG: Analysis performed on logarithmic scale.
SQRT: Analysis performed on square root scale.

GRP: Group
OCC: Occupation
AGE: Birth year (Age)

TABLE 5-19.

Summary Results of Restricted New Versus Old Nonblack Replacements Contrasted on 20 Variables

Variable (Transformation)	Replacement Group Means* (Percent Abnormal) Old	New	p-Value	Significant Interactions
Variables With Absolute Pairwise Correlations Less Than 0.10				
TBILI (LOG)	0.76	0.75	NS	
DBP (SQRT)	79.44	76.98	NS	
WBC (LOG)	7.01	7.91	NS	
SKIN	(52.3)	(61.5)	NS	
MMPID (LOG)	56.11	59.73	NS	
BUN (SQRT)	14.02	14.75	NS	
USG	1.014	1.013	NS	
PULSE	(15.3)	(19.2)	NS	
Dependent Variables With Absolute Pairwise Correlation Between 0.10 and 0.20				
RBC	5.01	4.90	NS	
GLUC (LOG)	108.8	95.86	0.007	
ECG	(14.3)	(23.1)		GRP*AGE
PLAT (SQRT)	270.5	271.56	NS	
CNS	(2.8)	(7.7)	NS	
CHOL (SQRT)	212.5	205.6	NS	
ALKPHOS (LOG)	87.75	87.72	NS	
COPRO (SQRT)	117.8	120.5	NS	
T4	7.56	8.00	NS	
TEST (SQRT)	601.2	612.6	NS	
SED (LOG)	4.15	6.37	0.03	
GGTP (LOG)	31.23	26.41	NS	

*All means are expressed in original units.

NS: Not significant (p>0.05).
LOG: Analysis performed on logarithmic scale.
SQRT: Analysis performed on square root scale.

Fully compliant Original and shifted Original Comparisons were compared in the Baseline Report with respect to reported health status, medication use, and work loss. Group differences for health status were significant ($p=0.001$) but were not so for medication use or for work loss; the shifted Original Comparisons tended to report themselves in poorer health than the Original Comparisons but were statistically equivalent to the Originals regarding medication use and work loss.

Fully compliant Original and shifted Original Comparisons were contrasted at first followup on reported health status, work loss, medication use, and daily use of aspirin. As in the Baseline Report, these analyses were done for only nonblack Comparisons.

The results of the contrast of Original and shifted Original Comparisons on reported health status are shown in Table 5-20. Here, health status is evaluated on a three-category scale (excellent, good, fair/poor).

The group difference between Original and shifted Original nonblack Comparisons regarding reported health status was not significant ($p=0.30$).

The results of the contrast of Original versus shifted Original Comparisons on medication use are shown in Table 5-21. The group difference between Original and shifted Original nonblack Comparisons regarding medication use was not significant ($p=0.68$).

The results of the contrast on work loss are shown in Table 5-22. The group difference between nonblack Original and shifted Original Comparisons regarding work loss was not significant ($p=0.82$).

The results of the contrast on daily aspirin usage are shown in Table 5-23. The group difference between Original and shifted Original nonblack Comparisons regarding daily aspirin usage was not significant ($p=0.98$).

Fully compliant Original and shifted Original nonblack Comparisons were also contrasted on each of the full set of 26 nearly uncorrelated variables shown in Table 5-17 on Baseline data. The results are shown in Table 5-24.

Sedimentation rate (SED) was analyzed as a categorical variable with values low (0-1), medium (2-3), and high (3-4). The percents of Original Comparisons within these categories were 35.8, 33.1, and 31.1 percent, respectively; the shifted Original Comparison percents were 30.8, 36.3, and 32.9, respectively. The probability of observing 3 or more successes in 10 independent repetitions of a Bernoulli trial, with a probability of 0.05 of success, is 0.0115. The probability of observing 2 or more is 0.0861. Based on these critical values and the results shown in the upper panel of Table 5-24, there appeared to be no statistical difference between Original Comparisons and shifted Original Comparisons.

The probability of observing 4 or more successes in 26 independent repetitions of a Bernoulli trial is 0.039. The probability of observing at most 2 successes in 26 independent repetitions of a Bernoulli trial, with probability 0.05 of success, is 0.86. Based on these critical values and the known slight correlation of the 16 dependent variables in the second panel of Table 5-19, these results suggested that Original and shifted Original Comparisons are not statistically distinguishable.

TABLE 5-20.

Reported Health Status of Fully Compliant Original and Shifted Original Nonblack Comparisons: First Followup

Reported Health	Original Comparison Group				Total	p-Value
	Original		Shifted Original			
	Number	Percent	Number	Percent		
Excellent	387	52	76	51	463	0.30
Good	307	41	68	45	375	
Fair/Poor	53	7	6	4	59	
Total	747		150		897	

TABLE 5-21.

Medication Use of Fully Compliant Original and Shifted Original Nonblack Comparisons: First Followup

Medication Use	Original Comparison Group				Total	p-Value
	Original		Shifted Original			
	Number	Percent	Number	Percent		
Yes	102	14	23	15	125	0.68
No	645	86	127	85	772	
Total	747		150		897	

TABLE 5-22.

Work Loss of Fully Compliant Original and Shifted Original Nonblack Comparisons: First Followup

	Original Comparison Group					
	Original		Shifted Original			
Work Loss	Number	Percent	Number	Percent	Total	p-Value
No	631	83	116	82	747	0.82
Yes	125	17	25	18	150	
Total	756		141		897	

TABLE 5-23.

Daily Aspirin Use of Fully Compliant Original and Shifted Original Nonblack Comparisons: First Followup

	Original Comparison Group					
	Original		Shifted Original			
Daily Aspirin Use	Number	Percent	Number	Percent	Total	p-Value
Yes	529	71	107	71	636	0.98
No	218	29	43	29	261	
Total	747		150		897	

TABLE 5-24.

Summary Results of Original Versus Shifted Original Nonblack Comparisons on 26 Variables at Baseline

Variable (Transformation)	Original Comparison Group Means* (Percent Abnormal) Original	Shifted Original	p-Value	Significant Interactions
Variables With Absolute Pairwise Correlations Less Than 0.10				
TBILI	0.61	0.61		GRP*OCC*AGE
DBP	80.46	78.95	NS	
WBC	7.52	7.18	NS	
SKIN	(37.5)	(43.8)	NS	
MMPID	56.25	58.40	NS	
BUN	14.26	13.76	NS	
USG	1.0209	1.0205	NS	
PULSE	(10.7)	(8.9)	NS	
NCVE	56.26	55.88	NS	
SEMEN (LOG)	77.4	72.8	NS	
Variables With Absolute Pairwise Correlation Between 0.10 and 0.20				
RBC	5.20	5.18	NS	
PULM	0.80	0.81	NS	
GLUC (LOG)	97.4	94.5	NS	
ECG	(27.6)	(26.7)	NS	
PLAT	270.6	269.9	NS	
IQ	108.6	108.4	NS	
CNS	(23.7)	(31.5)	0.02	
NCVA	48.17	47.59	0.01	
CHOL	220.7	213.1	NS	
ALKPHOS	7.84	7.60	NS	
COPRO (LOG)	31.1	30.4	NS	
ALA	2,497.0	2,505.3	NS	
T4	8.42	8.35	NS	
TEST	634.6	634.3	NS	
SED	given in text		NS	
GGTP (LOG)	38.43	35.53	NS	

*All means are expressed in original units.

Taken together, the results displayed in Table 5-24 very strongly suggested that Original and shifted Original Comparisons did not differ statistically at Baseline.

These analyses were repeated on the 20 available variables at the first followup. The results are shown in Table 5-25.

The results in the first and second panels of Table 5-25 and the binomial critical values given above suggested that no statistical difference was present between the Original and shifted Original Comparisons.

A single multivariate linear regression analysis was done on the 20 dependent variables shown in Table 5-25; no significant interactions involving group (Original, shifted Original) were noted and the group effect was not significant ($p=0.28$). Taken together, these analyses strongly suggested that there was also no statistical difference between Original and shifted Original Comparisons at first followup.

PARTIALLY COMPLIANT VERSUS FULLY COMPLIANT PARTICIPANTS

Ideally, compliance bias should be assessed by comparing the health of noncompliant and fully compliant participants with adjustment for group (Ranch Hand, Comparison) and the matching variables. The only information available on the noncompliant participants, however, is their responses to the health status questions, if they were willing to answer them, during the telephone conversation in which they refused to participate in the study. Noncompliant Comparisons were contrasted with their Baseline replacements (see noncompliance telephone questionnaire data, Tables 5-9 to 5-12). In addition, as in the Baseline Report, selection bias was studied by contrasting partially compliant with fully compliant participants with adjustment for group (Ranch Hand, Comparison). Taking the Baseline questionnaire at followup but refusing to take the physical examination or followup questionnaire were 9 Ranch Hands and 30 Comparisons who were either nonlocatable or noncompliant at Baseline. These 39 men were the only partially compliant participants at first followup. Their Baseline compliance is summarized in Table 5-26.

One of these individuals, a Ranch Hand with no interview, no physical, and no telephone interview, was Black. The label "no action" indicates that these individuals were not contacted because the Baseline contract expired. Individuals labeled "new Comparisons" were added to the study after the Baseline examination but before start of the first followup.

Data from these 39 partially compliant participants were statistically compared with similar data from fully compliant participants with adjustment for group (Ranch Hand, Comparison). This is shown in Table 5-27. Endpoints evaluated were reported health, medication use, and work loss. These analyses are similar to those reported in Table V-15 of the Baseline Report. Reported health status was collapsed to two categories (excellent, good/fair/poor) due to sparse data. One Black participant, a Ranch Hand, was deleted from these analyses.

The health versus compliance association in these data was of borderline statistical significance ($p=0.08$), with partially compliant participants tending to report themselves in better health than fully compliant

TABLE 5-25.

Summary Results of Original Versus Shifted Original Nonblack Comparisons on 20 Variables: First Followup

Variable (Transformation)	Original Comparison Group Means* (Percent Abnormal) Original	Shifted Original	p-Value	Significant Interactions
Variables With Absolute Pairwise Correlations Less Than 0.10				
TBILI (LOG)	0.75	0.73		GRP*OCC*AGE
DBP (SQRT)	80.0	79.60	NS	
WBC (LOG)	6.88	6.92		GRP*AGE
SKIN	(49.7)	(42.1)	NS	
MMPID (LOG)	56.2	55.1	NS	
BUN (SQRT)	14.8	14.04	NS	
USG	1.015	1.015	NS	
PULSE	(16.7)	(16.4)	NS	
Variables With Absolute Pairwise Correlation Between 0.10 and 0.20				
RBC	4.97	4.95	NS	
GLUC (LOG)	111.8	111.6	NS	
ECG	(15.3)	(11.9)	NS	
PLAT (SQRT)	263.2	271.9	NS	
CNS	(2.6)	(2.3)	NS	
CHOL (SQRT)	219.5	214.1	NS	
ALKPHOS (LOG)	89.76	85.53	NS	
COPRO (SQRT)	115.4	114.9	NS	
T4	7.58	7.58	NS	
TEST (SQRT)	576.6	559.0		GRP*OCC, GRP*AGE
SED (LOG)	5.11	4.91	NS	
GGTP (LOG)	32.39	29.77	NS	

*All means are expressed in original units.

TABLE 5-26.

Baseline Compliance Status of 39 Partially Compliant Participants: First Followup

Baseline Compliance	Group	
	Ranch Hand	Comparison
No Interview, No Physical, No Telephone Interview	3	23
No Interview, No Physical, Telephone Interview	2	1
New Comparison	0	3
No Action	4	3
Total	9	30

TABLE 5-27.

Reported Health of Partially Compliant Versus Fully Compliant Nonblack Participants

Compliance Status	Reported Health	Group				
		Ranch Hands		Comparisons		
		Number	Percent	Number	Percent	Total
Full	Excellent	473	43	635	57	1,108
	Good/Fair/Poor	482	46	575	54	1,057
Total		955		1,210		2,165
Partial	Excellent	5	20	20	80	25
	Good/Fair/Poor	3	23	10	77	13
Total		8		30		38

participants; 66 percent of partially compliant participants reported excellent health while only 51 percent of fully compliant participants reported excellent health. This association did not change with group ($p=0.91$).

The data on medication use and compliance status demonstrated no association ($p=0.57$), and this equivalence did not change with group ($p=0.79$). These data are shown in Table 5-28.

As shown in Table 5-29, the work loss-by-compliance association in these data was significant ($p=0.03$), with 84 percent of fully compliant participants reporting work loss and 95 percent of partially compliant participants reporting work loss.

These data are sparse and are not considered supportive or nonsupportive of the compliance bias calculations presented in the Baseline Report. The conclusions of the Baseline Report regarding the potential effects of compliance bias should be regarded as conservative overestimates, but worthy of consideration in inference formulations until more data become available.

CONCLUSIONS

These predominantly negative findings suggest that there has been no change in the way replacements self-select for entry into this study and, due to the obvious scheduling differences between new and old replacements, that no additional bias has been introduced at followup by scheduling differences. These data also strongly suggest that shifted Original Comparisons are not statistically distinguishable from Original Comparisons, either at Baseline or at first followup. This interpretation is also equivalent to the conclusion that no additional bias was introduced by scheduling differences between Original Comparisons and shifted Original Comparisons at Baseline. Available data on noncompliant Comparisons and their replacements suggest that, although replacements were not health-matched to refusals at Baseline, they are remarkably similar to refusals with respect to reported health, medication use, and income level. This result also supports a conclusion that there has been little, if any, selection bias due to nonparticipation in the Comparison group. This conclusion supports the use of the total Comparison group for all of the main analyses in the body of this report. Data regarding the few partially compliant participants at first followup are not sufficient to confirm or deny compliance bias calculations published in the Baseline Report.