Disposition of the Air Force Health Study: Interim Letter Report

Committee on the Disposition of the Air Force Health Study

ISBN: 0-309-65583-8, 22 pages, 8 1/2 x 11, (2006)

Disposition of the Air Force Health Study

Interim Letter Report

Committee on the Disposition of the Air Force Health Study
Board on Population Health and Public Health Practices

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

Support for this project was provided by the US Department of Veterans Affairs. This study was supported by Contract No. V101(93)P-2154 between the National Academy of Sciences and US Department of Veterans Affairs. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

For more information about the Institute of Medicine, visit the IOM home page at: **www.iom.edu.**

Printed in the United States of America.

The serpent has been a symbol of long life, healing, and knowledge among almost all cultures and religions since the beginning of recorded history. The serpent adopted as a logotype by the Institute of Medicine is a relief carving from ancient Greece, now held by the Staatliche Museen in Berlin.

"Knowing is not enough; we must apply.
Willing is not enough; we must do."
—Goethe

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Advising the Nation. Improving Health.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

COMMITTEE ON THE DISPOSTION OF THE AIR FORCE HEALTH STUDY (Interim Letter Report)

DAVID J. TOLLERUD, (*Chair*), Professor and Chair, Department of Environmental & Occupational Health Sciences, University of Louisville, Louisville, Kentucky

DAN G. BLAZER, Professor of Psychiatry, Duke University Medical Center, Duke University, Durham, North Carolina

ELLEN WRIGHT CLAYTON, Professor, Center for Genetics and Health Policy, Vanderbilt University, Nashville, Tennesse

MANNING FEINLEIB, Professor, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland

MARK S. GOLDBERG, Associate Professor, Department of Medicine, McGill University, Quebec, Canada

SUSAN E. HANKINSON, Assistant Professor, Harvard School of Public Health, Boston, Massachusetts

DAVID A. KALMAN, Professor and Chair, Department of Environmental Health, University of Washington, Shoreline, Washington

DEJURAN RICHARDSON, Professor of Mathematics, Lake Forest College, Chicago, Illinois

REGINA SANTELLA, Professor, Mailman School of Public Health, Columbia University, New York City, New York

CONSULTANTS TO THE COMMITTEE

JAMES G. HODGE, JR., Associate Professor, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland

VICTOR PONTES, Senior Data Manager, The Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts

STAFF

David A. Butler, Study Director
Amy O'Connor, Research Associate
Rose Marie Martinez, Director, Board on Population Health and Public Health Practices
Sonia J. Cheruvillil, Senior Program Assistant
Michael Saulle, Staff Intern

REVIEWERS

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the NRC's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

Christine B. Ambrosone, Director, Epidemiology, Roswell Park Cancer Institute
Tee L. Guidotti, Professor and Chair, Occupational & Environmental Health, George Washington University
Stephen W. Lagakos, Professor and Chair, Biostatistics, Harvard School of Public Health
Kyle Steenland, Professor, Environmental and Occupational Health, Emory University

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by **M. Donald Whorton**, WorkCare, Inc and **Stephen E. Fienberg,** Carnegie Mellon University**.** Appointed by the National Research Council and Institute of Medicine, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

TABLE OF CONTENTS

BACKGROUND

Public Law 108-183, the Veterans Benefits Act of 2003, directed the Secretary of Veterans Affairs to contract with the National Academy of Sciences (NAS) to address several questions regarding the appropriate disposition of the Air Force Health Study (AFHS). The AFHS is an epidemiologic study of Air Force personnel who were responsible for conducting aerial spray missions of herbicides during the Vietnam era—called *Ranch Hands* because the spray program was designated *Operation Ranch Hand*—and a matched cohort of comparison subjects who performed similar duties in Southeast Asia during the same time period but who were not involved with herbicide spraying. The study's first of six cycles of physical examinations were conducted in 1982 on 1,046 Ranch Hands and 1,223 comparison subjects (Michalek, 2005).

Section 602(c) of P.L. 108-183 charged the National Academies to evaluate

> (1) The scientific merit of retaining and maintaining the medical records, other study data, and laboratory specimens collected in the course of the Air Force Health Study after the currently scheduled termination date of the study in 2006.
>
> (2) Whether or not any obstacles exist to retaining and maintaining the medical records, other study data, and laboratory specimens referred to in paragraph (1), including privacy concerns.
>
> (3) The advisability of providing independent oversight of the medical records, other study data, and laboratory specimens referred to in paragraph (1), and of any further study of such records, data, and specimens, and, if so, the mechanism for providing such oversight.
>
> (4) The advisability of extending the Air Force Health Study, including the potential value and relevance of extending the study, the potential cost of extending the study, and the federal or nonfederal entity best suited to continue the study if extended.
>
> (5) The advisability of making the laboratory specimens of the Air Force Health Study available for independent research, including the potential value and relevance of such research, and the potential cost of such research.

In response to a request by the Secretary of Veterans Affairs, the Institute of Medicine (IOM) of the National Academies constituted the Committee on the Disposition of the Air Force Health Study to tackle these issues.

PURPOSE OF THE INTERIM LETTER REPORT

The purpose of this interim letter report is to address a component of the charge that bears on the conduct of the Air Force Health Study in the time leading up to its scheduled termination date of September 30, 2006. Specifically, the committee applies the information it learned in a May 2005 site visit to the AFHS research facility to the question of "whether or not any obstacles exist to retaining and maintaining the medical records, other study data, and laboratory specimens" collected in the course of the study. The committee believes that it is important to offer findings, conclusions, and recommendations on this topic in advance of its final report to allow for their timely consideration, as the availability of the investigators most familiar with the

AFHS data assets and the funding to support them cannot be assured after the end of fiscal year 2006.

All remaining components of the charge will be addressed in the committee's final report. The committee had not made final determinations on these at the time this report was completed and none should be inferred from the material presented here.

INFORMATION GATHERING BY THE COMMITTEE

As of August 31, 2005, the committee had conducted three meetings. During the first meeting, the committee received its charge from Dr. Mark Brown, the director of Environmental Agents Service of the Department of Veterans Affairs (VA). The committee was also briefed on the study design, protocol, and results of the AFHS by the study's then principal investigator, Dr. Joel E. Michalek. The second and third meetings included workshop sessions with presentations from experts in the conduct of longitudinal epidemiologic studies and the management and dissemination of epidemiologic data and biospecimens. Representatives of veterans service organizations also presented information for the consideration of the committee.

In addition, the committee has engaged in an extensive effort to collect other information on topics relevant to its charge. As part of this effort, four members of the committee—Drs. Blazer, Hankinson, Kalman, and Richardson—conducted a site visit to the AFHS research facility at the Brooks City-Base, San Antonio, Texas, on May 27, 2005. Accompanying them were IOM study staff and Mr. Victor Pontes, a consultant to the committee on issues related to SAS datasets.[1] The intent of the visit was to evaluate the state of the documentation of the study's data assets, examine how they were stored, and assess the ease of access to them. Working groups were established to focus on the electronically stored data and the specimens. The observations of these working groups are summarized below.

The committee's information gathering has been greatly aided by the AFHS staff, who have been helpful in answering the committee's many questions. The committee thanks them for their continuing cooperation.

Findings of the Data Working Group

The AFHS database is vast. It comprises electronically stored datasets containing the information collected from subjects during the six cycles of in-person physical examinations, as well as other data generated in the course of various analyses. The database also includes a number of materials originally collected in hard copy. These include paper originals of cycle physical exam reports and completed questionnaires; medical records from the subjects' physicians, dentists, and other health providers; X-rays and other diagnostic imagery; lists of medications taken; military administrative records such as duty station orders, flight records, performance reports, awards and decorations, and discharge documents; vital status records such as birth and death certificates; limited information on the subjects' spouse(s) and children; research reference materials; and the study's reports and papers. All hard copies of materials

[1] SAS is a software database management and analysis system. AFHS uses SAS to electronically store and analyze data collected over the course of the study.

have been scanned and the images stored in electronic files in Portable Document Format (PDF). The PDF files containing a particular subject's materials are stored in a directory labeled with the subject's name and study ID number (which the AFHS refers to as a *case number*). All told, there are over 8.8 million PDF image files (Michalek, 2005). In addition to these materials, there are ECG strips and approximately 3,000 SVHS video tapes containing high-resolution images of participants' teeth.

The data working group (Drs. Blazer and Richardson) began by asking the AFHS investigators to execute a SAS program to randomly draw the case numbers of five Ranch Hand veterans and five comparison subjects. These case numbers were used as a starting point for a series of exercises intended to elucidate the ease with which data could be accessed and analyzed. The AFHS investigators carried out the exercises and explained the steps they took to answer the questions that were posed. The working group did not examine the hard copy documentation, scanned images, or any other files or output that contained personal identification information on any study subject.

The AFHS investigators maintain what will be referred to in this report as *master* data files for each of the cycles that are used as the starting point for all analyses.[2] Within a cycle, there are separate master files for various components of the cycle's data gathering effort: responses to questionnaires, results of particular physical exams, and the like. Data in these files are stored by case number.[3] Working files are created from these masters and used for specific tasks. Working files are saved for potential future reference (for example, if it was necessary to reexamine the steps that lead to a particular result) but are not used in other analyses.

The amount, detail, and quality of documentation of the data vary by the cycle in which the data were collected. Data from the most recent cycle (Cycle 6, collected in 2002) appear to be well documented and the data dictionary for this cycle contains many desirable features: it exists as a searchable PDF, and it contains thorough descriptions of the variables and how information gathering changed from Cycle 5. Documentation for earlier cycles is less complete and is not necessarily in printed or electronically stored form. In one exercise, for example, investigators referred to handwritten annotations in their data dictionary to determine the meaning of coded responses to a questionnaire item.

AFHS reports and papers focus on analyses of a particular cycle's data; little longitudinal (across cycles) analysis has been done to date. When the working group asked the investigators to show the steps they would take to perform a longitudinal analysis, the working group identified several characteristics of the database that complicate such analyses. Data from early cycles are in different file formats than later data, requiring the analyst to be familiar with this fact and able to write code that accommodates it. The data location for a particular piece of information—for example, the questionnaire or the master data file where a variable can be found—may change between cycles. Variable names for the same piece of information sometimes change between cycles. Data formatting—whether responses are coded as a numeric value versus an alphanumeric character—may be inconsistent between cycles. Differences

[2] The original form in which data are delivered to the AFHS is referred to as a *raw* data file. A raw data file is quality-control checked against paper copy or other alternative documentation and corrected where necessary before a master data file is created. Summary variables derived from collected information—such as body mass index—are also incorporated into master data files.

[3] There are separate files that relate case numbers to participant names, military records, and other information that does not vary between cycles.

sometimes exist in how data are coded in identically labeled variables: for example, the year of the subject's birth (DOBYY) is given as the last two digits of the year in the Cycle 4 database, but it is given by all four digits in the Cycle 6 database. Technology (sensitivity or limit of detection or quantitation in lab tests, for example) has changed over time; it is therefore possible that an observation coded as below the limit in one cycle may have a value associated with it in a later cycle. And, missing data codes, error codes, and codes for specific outcomes are not uniformly documented or necessarily consistent between cycles.

Because the data may vary in so many and subtle ways, it is necessary to carefully consult the variable name and data dictionaries for each cycle where such documents exist—and to know where such information may be found in their absence—in order to carry out analyses.

Findings of the Specimens Working Group

The specimens working group (Drs. Hankinson and Kalman) interviewed program staff and conducted exercises to evaluate access to and preservation of whole blood, serum, urine, and semen specimens.

Specimens are stored at -70°C in a total of 23 freezers in a single room in a building that also houses other AFHS facilities. This storage temperature was the accepted state of the art in the 1980s.[4] In general, specimens exist in lots, reflecting the manner in which they were received; in some cases several vials are stored together in a zipper-sealed plastic bag, and in other cases they are held together with rubber bands. Labels on stored samples are not uniform but in general contain identifiers including the subject's case number, last and first name, and middle initial; and the sample type (e.g., SER3), collection time, and date. Some aliquots also indicate the subject's age.

Multiple steps are required to locate particular specimens. First, the subject's case number and the first four letters of his last name are keyed in to a master (computer-based) database that identifies the box in which the sample is stored and the freezer in which that box is located. However, boxes have been moved over time, and locations stored in this database have not been updated and are thus not reliable. The AFHS investigators recently completed a physical inventory of all boxes in the repository and were able to locate all boxes listed in the database. To locate a particular box, one must refer to the hard copy list generated in the physical inventory to determine the current freezer and shelf location of a particular specimen. For example, the database may indicate that subject X's serum specimen from Cycle 2 is located in freezer 19, drawer 2, box 4, but the inventory hard copy might indicate that the actual location of the specimen is freezer 14, drawer 2, box 4.[5] Despite the intricacy of the inventory records, staff were able to locate specimens requested by the working group with little difficulty.

Neither the existing database nor the hard copy reinventory provide data regarding the number of aliquots or volumes stored. It appears that these data were not recorded consistently. For those cycles for which quantities were recorded, the current volumes in the record reflect the

[4] The practice in 2005 is to store biologic samples at least -80°C; liquid nitrogen storage—which maintains samples at still lower temperatures (below -130°C)—is used in some studies.
[5] The working group was told that drawer and box numbers were maintained in moves.

original quantities received. These have not been updated to indicate removal from the original amounts received.

No documentation exists for the thermal history of specimens. However, the physical condition of the freezers appears to be adequate. All had external labels indicating that they had been inspected at 6-month intervals for at least the past 3 years. The freezers have been monitored for failure throughout the course of the study—in earlier years by physical inspection and later through electronic monitoring and alarm systems. Three people are on call at all times to move samples in case of a freezer failure. Empty freezers kept at -70°C are available for immediate transfer of samples in case of freezer breakdown. In the earlier years of the study, there was one instance where specimens were compromised by an equipment failure. This occurred in 1986 when 656 urine samples were lost because of a freezer outage.

Retrieval and inspection of three specimens from randomly selected subjects showed no signs of leakage or major thaws. The appearance of the urine specimen examined by the working group was open to interpretation: the position of the fluid in the vial was consistent with either minor thawing and refreezing, or the initial freezing of the specimen occurring in two phases.

Some residuals from serum samples sent to CDC for dioxin analysis were returned to the AFHS repository. These samples are kept separate from others, and they are marked as returns in an inventory document. With this exception, all samples in the repository appear to have remained frozen since collection. Over the course of the study, specimens have been sent to Duke University, the University of Cincinnati, and the University of Virginia for various analyses. None of these were returned to the inventory. During the Cycle 5 exams 313 adipose tissue samples were collected; they are being held by a collaborating investigator at the University of California, Davis. The committee does not have any information on the condition and storage circumstances of these samples.

During the committee's first meeting, the AFHS investigators mentioned that they plan to conduct a reassay experiment to evaluate the analytical viability of the specimens. The plan calls for analyzing specimens from a small sample of subjects over the six study cycles. These samples will be assayed for endpoints for which historical values are available. Results are expected to generally indicate the current condition of preserved specimens, although their interpretation is not straightforward, and the suitability of these samples for specific analytes will need to be determined in some cases.

Summary Observations of the Working Groups

The AFHS has been in existence for approximately 25 years. The working groups found that data and specimens have been maintained at a level typical of most long-term epidemiological research. Over the course of the study, the best practices in epidemiologic investigations and specimen storage have changed, and the technology for managing and analyzing data and samples has advanced. The AFHS has evolved in response to these advances. Study personnel were not tasked with updating and harmonizing the system over time or rendering it accessible to outside researchers, and there was no particular incentive to expend time and funds on efforts to do so. The observations offered here on the state of the data and specimens should therefore not be viewed as a criticism of the work of the AFHS staff.

FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

On the basis of the information gathered in its meetings to date, the site visit to the AFHS research facility, and review of relevant literature, the committee offers the following findings, conclusions, and recommendations.

The committee finds that the medical records, other study data, and laboratory specimens collected in the course of the AFHS have been properly maintained. However, they are not currently organized and documented in a manner that allows them to be easily understood, evaluated, managed, or analyzed by persons outside of the AFHS.

There are several possible options for the disposition of the medical records and other study data collected in the course of the AFHS. These include continuation of the study, making the data assets available to outside researchers, and rendering them to the National Archives for final disposition. The committee will address this issue in its final report.

However, whatever recommendations may be offered and whatever policy decision may be taken in response, the committee believes that there is merit in creating a more complete and uniform accounting of the AFHS medical records and other study data. Such action will allow better inventorying of the assets should they be archived and will preserve the possibility of making further meaningful use of the data. The size and complexity of the AFHS database make issues of organization, storage, and accessibility especially salient.

The AFHS data assets may be subject to records retention statutes or regulations and may therefore need to be appraised for possible retention. It thus makes sense that the data be in a form that is comprehensible to people who are not already familiar with them so that reasoned decisions can be made. Any future uses of the data—should such options be pursued—would be facilitated by having them documented completely and in a uniform manner. This would be especially helpful in studies that examine changes in health over time and across examination cycles.

Similarly, there are two primary options for the disposition of the laboratory specimens collected by the AFHS: disposal or maintenance for possible future use. Regardless of what decision is made concerning their disposition, there is value in fully documenting and reorganizing the laboratory specimens prior to the study's presently scheduled termination date. This would be helpful in auditing their proper disposal and would be vital to facilitating any possible future use.

As already noted, the AFHS is currently scheduled to terminate on September 30, 2006, and the availability of the investigators most familiar with the data assets and the funding to support them cannot be assured thereafter. Any actions to be taken with respect to the data and specimens must therefore be initiated as soon as possible and should be completed by the end of current funding.

The committee concludes that the present state of the documentation and organization of the AFHS medical records, other study data, and laboratory specimens is an obstacle to retaining and maintaining these materials after the currently scheduled termination date of the study.

The committee therefore recommends that action be taken prior to the currently scheduled termination date of the AFHS to reorganize and document the study's medical records, other study data, and laboratory specimens in a form and format that allows them to be easily understood, evaluated, managed, or analyzed by persons outside of the AFHS.

Specific Recommendations Regarding Medical Records and Other Study Data

The committee recommends that the following actions regarding the documentation and reorganization of medical records and other study data take place before the currently scheduled termination date of the Air Force Health Study:

- Create a comprehensive inventory of master data files, organized by examination cycle. This inventory should include the file name and type (flat file,[6] SAS database, and so on); a brief description of its contents; and the name, column location, and length; variable type (character or numeric); data codes; and description of each variable stored in the file. The *Variable Name Dictionary for the Air Force Health Study, 1992 Questionnaire and Analyses* (Michalek, 2000c) is an example of such a document.[7]
- Create a comprehensive inventory of the variables contained in the master data files, organized by examination cycle and by questionnaire, physical examination report, or other data intake instrument. The *Data Dictionary for Physical and Psychological Examination and Laboratory Data, Air Force Health Study, Cycle 6* (SAIC, 2003) is an example of such a document.[8] The contents of this Cycle 6 data dictionary include

    1. an annotated version of the data collection forms, comprising the variables and codes used in the associated database;
    2. a synopsis of the variable names and their descriptions;
    3. the summary variables created and codes used;
    4. the number of study subjects examined for each test;
    5. changes in the database structure from the previous exam; and
    6. notes on data comparability between cycles.

    In addition, the laboratory results section of the dictionary contains descriptions of assays, units of measurement and normal ranges, and data codes.

    The committee recommends that such information be compiled for all variables in all exam cycles and that it include notation of whether any attributes of a variable have changed over the course of the study.
- Create a master data codebook containing the name of every data variable represented anywhere in the AFHS database—that is, at any exam cycle—along with a brief description of the variable, the master data file(s) in which it was stored, and its pertinent attributes. This

[6] A flat file contains records that are stored without structured relationships or formatting. The simplified form allows data to be used by a variety of applications and minimizes the possibility of data loss due to software obsolescence.

[7] Variable name dictionaries also exist for the Cycle 2 (Michalek, 2000a) and Cycle 3 (Michalek, 2000b) physical exams.

[8] There is a separate data dictionary for the Cycle 6 "Health Interval" and "1982 Baseline" questionnaires (NORC, 2003).

codebook would be derived from the documents outlined above and would constitute a one-stop-shopping distillation of database contents. It should make clear during which exam cycles a particular piece of information was gathered and the variable name(s) associated with that information over the course of the study. A master identification table with rows containing variable descriptions, columns representing each of the six exam cycles, and the intersections listing the variable name used in that cycle, would be a useful adjunct to this effort.

- Create a document describing the contents, format, and location of the AFHS collection of materials that have been scanned into PDF image files—subjects' medical records, diagnostic imagery, military and vital status records, and the like—and explaining the collection's organizational structure. This document will serve as a directory to these data.

The committee recommends that these documents be in a form and format that facilitates easy access to their contents; searchable electronic files with paper backup for archival purposes would accomplish this.

As the committee notes, some of these documents already exist in some form. Indeed, the documentation available for later exam cycles already exhibits many of the features described above. The concern is that the information for all cycles be compiled into easily identified, definitive reference documents with uniform information content. Doing so will greatly reduce the amount of archaeological investigation that those unfamiliar with the AFHS database must do to understand it.

In addition, the committee recommends that an overall plan be developed and implemented for archiving the medical records and other study data of the AFHS.[9] Ease of accessibility to the data should be a primary consideration in this effort. The committee notes that federal regulations addressing the preparation of electronic records for transmittal to the National Archives (36 C.F.R. § 1228.270) contain specific information about file formats and media that are appropriate for long-term storage.

Specific Recommendations Regarding Laboratory Specimens

The committee recommends that the following actions regarding the documentation and reorganization of laboratory specimens take place before the currently scheduled termination date of the Air Force Health Study:

- Reinventory all laboratory specimens held by the AFHS, verifying their location and ascertaining the number and volume of aliquots and type of sample. Carry out a visual inspection of specimen condition while performing this activity, and identify any problematic samples.
- Update and create a single specimen database that includes case number, exam cycle, specimen type, and freezer location.

[9] The committee's recommendations constitute elements of such a plan.

- Compile all information regarding specimen history (receipt, realiquoting, freeze-thaw cycles, dispersal, and the like) into a single reference database. Where data gaps are present, note their existence.
- Compile all protocols regarding receipt, maintenance, dispersal, and return of specimens for all cycles into a single reference document.
- Document the status of all laboratory specimens sent to outside investigators. Ensure that extant specimens are either disposed of using current best practices or reintegrated into the inventory, clearly marked with their history.
- Perform the currently planned reassay to aid in the evaluation of specimen stability and condition.

The committee notes that AFHS investigators informed the working group in May 2005 that some of these actions were already planned. Specifically, the working group was told that the AFHS was undertaking a reinventory and physical reorganization of the specimens in the final year of the study. The committee is in agreement with this effort as it will reconcile the inventories and create a new database that will provide the current location of all specimens and indicate the volumes and types of biological material in storage for each individual.

The planned reorganization will collect all specimens for a study subject in one or sequential boxes. Currently, specimens are stored in freezers by cycle and specimen type, not by subject. This endeavor will greatly decrease the effort associated with pulling the samples. However, if a request is for a single type of sample (e.g., urine), many more boxes will have to be pulled than if samples were separated by specimen type. Combining the samples from various cycles will necessitate very careful pulling of requested samples in the future so that not only the correct sample type but also the correct collection time period for an individual is selected. As already noted, it was not clear from the working group's visit that each specimen vial has the exam cycle or date of collection on the label. **Thus, the committee recommends that—if this element of the planned reorganization is carried out—the exam cycle or date of collection be added to any vials that are not already marked with such information.**

Other Remarks and Recommendations

It is difficult to estimate the potential cost of implementing the recommendations offered here. Such costs would depend on several factors, including the ease with which existing documentation can be adapted to satisfy the recommendations,[10] whether the AFHS personnel perform the work themselves or contract it out, and the amount of time available to accomplish the tasks once the decision is made to undertake them. As this report notes, at least some of the actions recommended by the committee were already planned by the AFHS investigators and may be underway. Other work may well fall under existing budget items and should thus not represent an incremental cost to the study. The AFHS FY 2006 budget (DTIC, 2005) includes

[10] The committee reviewed a great deal of AFHS electronic and paper material but could not perform a detailed examination of all of it. The difficulty of determining the extent to which appropriate records currently exist underscores the need for the documentation actions it proposes.

allocations to perform documentation and organization of the data assets in anticipation of their future disposition:

> Continue to process and document examination data. Continue archiving previous cycles' examination data and digitize and archive the Cycle 6 data as received. Conduct medical records coding and verification of examination database and Cycles 1 through 6 coding. … Prepare for and complete transition or turnover of cycles' examination data and digitize and archive the Cycle 6 data as received.

A total of $1,612,000 is assigned to these and other data analysis and support tasks.[11] Whether or not existing funds are sufficient to accomplish the recommended actions, the committee believes that it is incumbent on the Air Force, as the custodian of the AFHS research materials, to insure their proper documentation and organization for both historic reasons and for possible future use. **Therefore, should the available AFHS program funds not be sufficient to accomplish the actions elucidated above, the committee recommends that supplemental funding be provided to carry out such work in a complete and timely manner.**

REFERENCES

DTIC (Defense Technical Information Center). 2005. *Approved FY2006 Presidential Budget Request. RDT&E Budget Item Justification Sheet, Ranch Hand II Epidemiology Study, PE number: 0605306F*. [Online]. Available: http://www.dtic.mil/descriptivesum/Y2006/AirForce/0605306f.pdf [accessed August 31, 2005].

Michalek JE. 2000a. *Variable Name Dictionary for the Air Force Health Study. 1985 Physical Exam Questionnaire and Analyses*. Brooks AFB, TX: Air Force Research Laboratory. AFRL-HE-BR-SR-2001-0003. [Online]. Available: http://handle.dtic.mil/100.2/ADA387644. [accessed October 30, 2005].

Michalek JE. 2000b. *Variable Name Dictionary for the Air Force Health Study. 1987 Analyses Data Sets for 1987 Report and Dioxin Report*. Brooks AFB, TX: Air Force Research Laboratory. AFRL-HE-BR-SR-2000-0003. [Online]. Available: http://handle.dtic.mil/100.2/ADA380762. [accessed October 30, 2005].

Michalek JE. 2000c. *Variable Name Dictionary for the Air Force Health Study. 1992 Questionnaire and Analyses*. Brooks AFB, TX: Human Effectiveness Directorate, Human Systems Wing. AFRL-HE-BR-SR-2000-0008. [Online]. Available: http://handle.dtic.mil/100.2/ADA387728. [accessed August 31, 2005].

Michalek JE. 2005. *The Air Force Health Study: An Overview and Update*. PowerPoint presentation before the Committee on the Disposition of the Air Force Health Study, January 5, 2005. Washington, DC.

[11] The other tasks specified in the line item were support of the annual mortality analysis, conduct of data analysis for journals and reports to Congress, and continued maintenance of study's local area network. An additional $1,677,000 was allocated for data analysis under a separate line item.

NORC (National Opinion Research Center). 2003. *Air Force Health Study 2002-2003 Follow-Up Examinations. Data Dictionary: Health Interval Questionnaire; 1982 Baseline Questionnaire.* Chicago, IL: NORC, University of Chicago.

SAIC (Science Applications International Corporation). 2003. *Data Dictionary for Physical and Psychological Examination and Laboratory Data Air Force Health Study, Cycle 6*. Prepared for the US Air Force Human Systems Wing/YASP by SAIC in conjunction with the National Opinion Research Center and Scripps Clinic.