AFRL-SA-WP-SR-2015-0005

The Impact of Environment and Occupation on the Health and Safety of Active Duty Air Force Members – Database Development and De-Identification

Monica U. Selent, Lt Col, USAF, BSC;
Ryan S. Mayes, PhD; Roger A. Erich, PhD;
Melinda Eaton, Lt Col, USAF, BSC

April 2015

Final Report
for August 2013 to August 2014

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2065, 22 Apr 2015

STINFO COPY

Air Force Research Laboratory
711th Human Performance Wing
U.S. Air Force School of Aerospace Medicine
Public Health and Preventive Medicine Dept
Epidemiology Consult Service
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

NOTICE AND SIGNATURE PAGE

AFRL-SA-WP-SR-2015-0005 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//	//SIGNATURE//
MELINDA EATON, LT COL, USAF, BSC	PAUL A. SJOBERG, COL USAF, BSC
Chief, Epidemiology Consult Service	Chair, Public Health & Preventive Medicine Dept

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
1 Apr 2015

2. REPORT TYPE
Special Report

3. DATES COVERED *(From – To)*
August 2013 – August 2014

4. TITLE AND SUBTITLE

The Impact of Environment and Occupation on the Health and Safety of Active Duty Air Force Members – Database Development and De-Identification

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Monica U. Selent, Lt Col, USAF, BSC
Ryan S. Mayes, PhD
Roger Erich, PhD
Melinda Eaton, Lt Col, USAF, BSC

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
USAF School of Aerospace Medicine
Public Health and Preventive Medicine Department
Epidemiology Consult Service
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-SA-WP-SR-2015-0005

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-2065, 22 Apr 2015

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The overall goal of this project is to analyze environmental and occupational factors that affect risk-taking behaviors and health outcomes among active duty U.S. Air Force personnel. Due to the complexity of this database, additional data cleaning, database linking, and de-identification procedures were performed. With the complete database, investigators will be able to explore the general hypothesis that environmental and occupational factors influence both health outcomes and risk-taking behavior of service members in certain career fields, with identification of high-risk career fields for targeted interventions. Additionally, low-risk career fields could be identified for potential protective factors. The result of this project is a comprehensive database that was formatted, cleaned, and de-identified that will enable investigators to highlight environmental and occupational factors specific to the Air Force that impact health and safety outcomes. New variables were created to capture subject characteristics from the Preventive Health Assessment data and to explore rating chain relationships between subjects and their effect on the health and safety of active duty Air Force personnel. In addition, descriptive analyses were performed on the complete data. Phase II database analysis will continue for this project. The purpose of this report is to document the data cleaning, database linking, and de-identification steps taken to build the database prior to analysis.

15. SUBJECT TERMS
Database development, occupational health, environmental factors, military personnel

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Lt Col Melinda Eaton
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	SAR	17	

This page intentionally left blank.

TABLE OF CONTENTS

Section **Page**

LIST OF TABLES

1.0 SUMMARY

The overall goal of this study is to investigate how environmental and occupational factors affect risk-taking behaviors and health outcomes among U.S. Air Force personnel. Due to the nature of their occupations, active duty Air Force members can face numerous hazards on a daily basis. These environmental and occupational hazards may directly influence an individual's physical and mental health. Previous literature has linked occupational factors to risk tasking and health and safety outcomes for service members in distal, indirect ways, such as in diseases or injuries acquired while not directly performing job duties and responsibilities, or because of stressors in the workplace. These conditions can be due to increased risk-taking behaviors or to medical issues influenced by, but not directly related to, an occupational or environmental exposure.

While the nature of the occupations of these service members can be direct sources of stress, these jobs also have the potential to provide beneficial aspects including social support or camaraderie within the organization or career field. The focus of Phase II of this project was to analyze the developed database through examination of how environmental and occupational factors affect risk-taking behaviors and health outcomes among U.S. Air Force personnel. However, prior to analysis, it was necessary to perform additional work, including data cleaning, database linking, de-identification, and incorporation of additional variables. Once we completed these final database changes, preliminary work, such as descriptive analyses, commenced. From the analyses planned in Phase II, we can identify high-risk career fields for targeted interventions and low-risk career fields for potential protective factors. Based on those factors, results from this study may be used to develop policy recommendations aimed at improving these outcomes for active duty Air Force members. The purpose of this report is to document the data cleaning, database linking, and de-identification steps taken to build the database prior to analysis.

2.0 BACKGROUND

Occupational safety is typically defined by potential environmental and occupational risk factors as well as incidents and accidents that occur on the job. However, job assignments and associated stress levels of individual service members can have major implications for safety off the clock as well. Individual workplaces in the Air Force have been examined for environmental and occupational risks; however, a broad Air Force perspective of all occupations and workplaces has not yet been conducted. The overall goal of this study is to investigate how environmental and occupational factors affect health outcomes and risk-taking behaviors among U.S. Air Force (USAF) personnel.

An Aerospace Medicine team (typically Bioenvironmental Engineering, Public Health, and flight surgeons) uses nationally identified sources, such as Air Force Occupational Safety and Health standards and Occupational Safety and Health Administration expanded standards, to determine potential environmental and occupational exposures to evaluate all workplaces on Air Force installations. Once these exposures are identified, preventive measures are implemented and documented for high-risk workplaces (AF Forms 2755 and 2766). Aerospace Medicine routinely conducts visits to each workplace to ensure that these preventive measures are utilized as well as investigate any potential mishaps or any reported accidental exposures. Any occupationally related injury or illness is reported to the Air Force Safety Center. Each base

Bioenvironmental Engineering and Public Health office maintains records for the high-risk workplaces on their installation; however, information for these workplaces have not yet been examined or analyzed at an Air Force level.

In addition to direct environmental and occupational hazards, members of the military report higher psychological strain than the general population and significant work stress [1,2]. These stressors may manifest in health and safety outcomes for service members in distal, indirect ways, such as in diseases (e.g., depression [3]) or injuries acquired while not directly performing job duties and responsibilities. These outcomes can be due to increased risk-taking behaviors like drug use [4] and smoking [5] or to medical issues, such as poor diet [5] and obesity [6]. While the occupations of service members can be a source of stress, these jobs also have the potential to protect against stress and resulting issues [7].

3.0 METHODS

The purpose of this project was to analyze a database, created in Phase I, to examine how environmental and occupational factors affect risk-taking behaviors and health outcomes among USAF personnel. Initial analysis work completed involved the exploration of the general hypothesis that environmental and occupational factors influence both health outcomes and risk-taking behavior of service members in certain career fields, with high-risk career fields identified for targeted interventions and low-risk career fields identified for potential protective factors. However, due to the nature of these data, additional cleaning and organization were necessary to prepare for analysis.

Phase I (Database Development) utilized the skills of a database manager to complete data preparation on data from six distinct data sources: Air Force Personnel Center (AFPC), Air Force Safety Center (AFSC), Standard Ambulatory Data Record (SADR), Standard Inpatient Data Record (SIDR), Air Force Reportable Event Surveillance System (AFRESS), and Preventive Health Assessments (PHAs). Phase II (Database Analysis) brought on an epidemiologist/biostatistician to finalize the data preparation and initiate analysis.

Data were maintained at the U. S. Air Force School of Aerospace Medicine, Epidemiology Consult Service at Wright-Patterson Air Force Base, in building 840, on existing computers that required appropriate access. An Institutional Review Board evaluation was conducted to review the protocol and ensure that the project did not meet the definition of Human Subjects Research. A waiver of consent was granted since it was not practical or feasible to obtain informed consent for the large number of records (514,446 unique subjects) included in this database.

4.0 RESULTS

This database only includes active duty Air Force (ADAF) members, approximately 300,000 per year, for the 5-year period from 1 January 2006 to 31 December 2010. Throughout the study period, there were 514,446 distinct subjects; many subjects were in the dataset for multiple years. The number of subjects, inclusion/exclusion criteria, and age range were determined by the data sources; no sub-sampling of the data was employed. There were no specific inclusion or exclusion criteria; therefore, the age range is 17-70 years old and the male to female ratio is approximately three to one.

Since the purpose of the database was to link all records from multiple databases into one searchable database, no new data were collected. All existing data had been collected as part of routine surveillance and clinical care from multiple data sources. Records were linked between datasets by using Social Security numbers (SSNs) and AFPC monthly Import Dates. For the data from sources other than AFPC, applicable record dates were converted to the AFPC Import Date if they fell within the designated monthly interval. To avoid duplication of demographic and occupational data, AFPC information was treated as the gold standard and all duplicate data were removed from the other six databases. Demographic data elements consisted of date of birth (DOB), age, gender, ethnicity, race, and marital status. Occupational data elements included Primary Air Force Specialty Code (PAFSC), Duty Air Force Specialty Code (DAFSC), skill level, rank, date of rank, duty status, education level, installation, and organizational structure. Outcome data elements included on-duty safety incidents (e.g., vehicle accidents, falls, sports injuries, lacerations, etc.), high-risk sexual behavior (e.g., unprotected sexual intercourse, diagnosis of sexually transmitted diseases [STDs]), and physical/mental health issues, such as high blood pressure or mental disorders. See Table 1 for outcome data elements and data source.

Table 1. Outcome Data Elements and Data Source

Potential Outcome	Variables	Data Source
Occupational Injury	Subject SSN, DOB, rank, gender, diagnosis, acute or chronic injury, date of report, number of duty days lost, location of injury	**AFSC**
Occupational Illness	Subject SSN, DOB, rank, gender, diagnosis, acute or chronic illness, date of report, number of duty days lost/duration of illness, location of illness	**AFSC**
Alcohol Use/ Tobacco Use	Subject SSN, DOB, rank, gender, encounter date, answers to alcohol and tobacco use questions (Section 8-Tobacco Use; Section 9-Alcohol Use)	**PHA**
High-Risk Sexual Activity	Subject SSN, DOB, rank, gender, encounter date, answers to sexual activity questions (Section 12-Reproductive Health Issues)	**PHA**
Sexually Transmitted Disease	Subject SSN, DOB, sponsor pay grade, gender, ICD-9 codes, date of diagnosis	**SADR/SIDR and AFRESS**
High Blood Pressure	Subject SSN, DOB, sponsor pay grade, gender, ICD-9 codes, date of diagnosis	**SADR/SIDR**
Mental Disorder	Subject SSN, DOB, sponsor pay grade, gender, ICD-9 codes, date of diagnosis	**SADR/SIDR**
Demographic Data	Subject SSN, DOB, grade (rank), date of rank, gender, PAFSC, DAFSC, duty location, ethnic designator, race, unit, marital status	**AFPC**

Note: ICD-9 = International Classification of Diseases, Ninth Revision.

4.1 Additional Data Cleaning

To prepare data from the six sources for linking and analysis, there were additional cleaning and organizing steps to complete. The work conducted on each separate database is explained in the paragraphs to follow. Note the SADR and SIDR databases needed no additional work during this phase. Once the study team accomplished data cleaning and organization, the next step consisted of putting all the data together into a usable form for analysis.

Additional cleaning procedures conducted for the AFPC database involved removing Air Force Academy cadet records (143,227) and retiree records (577). The ethnicity and race variables were also problematic (20,393 SSNs/1,005,915 records for ethnicity and 11,501 SSNs/562,734 records for race). As these subjects progressed through the study, their race/ethnic codes would change one or more times. To remedy this issue, the latest race/ethnic code, available for a particular subject in the study, was used to recode all of the previous ethnic or race codes for that particular subject. This method relied on the assumption that the latest code had the highest probability of being correct, since it was most likely updated by the subject's request to AFPC. Another issue involved a missing PAFSC, DAFSC, or both. If only one of these codes were available, the available code in the affected record replaced the missing code (158,723 records). If they were both missing, they both remained blank (127,618 records).

Sixty-two SSNs with multiple genders remained in the database. Dr. Lamar Pierce, co-investigator from Washington University of St. Louis, provided results from a gender determination algorithm that assigned probabilities of being female based on the subject's first name and birth year. Corrections were made based on these results.

To capture the organizational structure for each unit present in the database, new variables Squadron, Group, Wing, Numbered Air Force, Other, and MAJCOM [major command] were created. These new variables were filled in with the appropriate command structure applicable to the Unit and Import Date of the AFPC record. Note that only the units available in the original Unit variable were used to fill in the newly created organizational variables. In addition, many units changed organizational structure during the study. Appropriate coding of the organizational structure was used to identify these changes by modifying the unit name to reflect the new structure. For example, the 16th Special Operations Squadron belonged to the 16th Operations Group, the 1st Special Operations Group, and the 27th Operations Group, in that order, during the study timeframe. The unit was named the 16th Special Operations Squadron, the 16th Special Operations Squadron 1st, and the 16th Special Operations Squadron 27th, respectively.

Another variable, Unit Category, was created to group similar units together based upon mission type. For example, the 391st Fighter Squadron was placed in the Fighter category and the 366th Communications Squadron was placed in the Communications category. Some of the other categories include, but are not limited to, Acquisition, Air Base, Airlift, Civil Engineer, Intelligence Surveillance and Reconnaissance, and All Others. In all, there are 67 distinct unit categories.

We also created a new variable identifying those who spent time as a prisoner, were under Security Forces custody, were being investigated by the USAF Office of Special Investigation, were in legal trouble (determined by Duty Status Code or Duty Title), or other undetermined legal, judicial, criminal, or punitive-type categories. In total, 3,363 SSNs (97,322 records) were coded to reflect this status. If the subject met any of the criteria above, the study team identified all of the subject's records present in the study with this code regardless of when this status was attained. If all of an individual's records in the study contained the applicable Duty Status Code or Duty Title, this individual was removed from the study data (836 SSNs, 9,951 records).

Additional cleaning procedures conducted for the safety data involved eliminating 14 records that were present in both the Safety-Injury and Safety-Illness data. From the Safety-Injury database, 4 heat exhaustion related records were removed and from the Safety-Illness data, 10 records were removed that were not heat exhaustion related. There were also 154 cadet records removed from the Safety-Injury database.

Preparing the AFRESS database required only one adjustment. Originally, the Import Date was determined by the Case Created Date, which corresponds to the date the record was generated in AFRESS. To better align the record with the actual event time, the Case Date of Onset was used to determine which Import Date to assign. All Import Dates were adjusted to reflect this change. As a result, 606 records were dropped, since the Case Date of Onset was earlier than January 2006. Additional records, with Case Date of Onset occurring within the study timeframe, were obtained from AFRESS. After cleaning these new records, 461 records were added to the study database.

Finally, the PHA database was prepared for analysis. After removing all invalid records (Guard, Reserve, other branches of service, dependents, etc.), the main issue with this database was the presence of more than one PHA completed in the same calendar year (CY) by the same subject (up to six PHAs in CY). Therefore, only the latest PHA per CY for each subject was retained.

4.2 Final Database Preparation and Linking

To bring together all of the individual databases, a linking mechanism was created. Since the AFPC database was the main focal point of all the records, a link value was created for each available record. The first step involved randomly choosing numbers between 1 and 40,200,000 and eliminating duplicate values. Then, the study team added these random numbers to the AFPC database as a new variable. Note that these numbers were not ordered in such a way that personnel records could be identified. Now, these random numbers uniquely identify all of the possible SSN and Import Date combinations available in the AFPC database.

Before assigning this link variable to the other databases, the databases were modified to make analysis more efficient. In doing so, the linking process became more efficient, and the study team can easily update any part of the database, if needed.

Originally, the format of the PHA database contained one question response value per record. Therefore, there were up to 43 records corresponding to one completed PHA per subject and Import Date. To simplify this data structure, all records with question responses applicable to a distinct, completed PHA were combined into one record. After this modification, the PHA and AFPC databases were linked together with the SSN and Import Date combination, and the link variable was added to the PHA database.

The SADR database required several modification steps before assigning the link variable. The original SADR data structure consisted of one record per subject per visit with up to five medical diagnosis ICD-9 codes. There may have been several visits per day and/or several visits per month for a study subject. To align these data with the monthly structure of the AFPC database, each day's worth of visits for a subject were combined into one record (duplicate ICD-9 codes were removed from these records). For these newly created records, there were up to 52 distinct ICD-9 codes per record. A time-ordered visit day number was also assigned to each newly created subject's record to capture if there was more than one visit day within a given Import Date interval (time from the previous Import Date up to and including the next Import Date). After these modifications, the SADR and AFPC databases were linked together with the SSN and Import Date combination, and the link variable was added to the SADR database.

For the SIDR database, we employed the same procedures used for the SADR database. Each transformed SIDR record contained up to 10 distinct ICD-9 codes.

Preparation of the AFRESS database involved combining records of subjects with two records within a particular Import Date interval (described previously in the SADR description). The study team created another variable to capture the second ICD-9 code if applicable. Therefore, each record in the modified AFRESS database represents a subject's reportable event(s) within the Import Date interval. After this modification, the AFRESS and AFPC databases were linked together with the SSN and Import Date combination, and the link variable was added to the AFRESS database.

For the safety databases, illness and injury, a new variable was added to each database to capture the number of mishaps or illness issues for each subject within a particular Import Date interval. The maximum amount of records per subject within the monthly timeframe was two, and the study team noted that no records were combined for these databases. After this modification, each safety database and the AFPC database were linked together with the SSN and Import Date combination, and the link variable was added to each of the safety databases.

By incorporating the link variable across the study databases, the research team now has a complete, relational database. The next step involves reducing the potential for disclosure of the subject's personal and medical information. To mitigate this issue, a limited database was created as described in the following section.

4.3 Creating a Limited Dataset

The first step in this process was to eliminate all variables, per the Health Insurance Portability and Accountability Act guidelines, within each database that were unnecessary for analysis or could potentially aid in the subject's identification. Variables including name, medical record numbers, and all dates, with the exception of the AFPC Import Date and Date of Rank, were removed. SSNs were recoded as randomly assigned subject identification numbers to further mask these data, and the Date of Rank was recoded to a calculated time in grade based on the difference between Date of Rank and the AFPC Import Date.

Since location information was retained for analysis, the study database is considered limited. To mask the particular locations, random codes were used in place of the location name and type. The subject's particular unit information also has the potential for identification. All units were masked by a random code. Masking these variables was essential since there were low counts within particular units and locations (653 distinct units and 1,226 distinct locations had cell counts of one for a particular monthly AFPC Import Date).

The subject's rank and specialty codes were also sources of identification. Several PAFSCs and DAFSCs directly identified the subject's position. These specialty codes were recoded as 888 to group them as Command/Staff. High-ranking individuals were also prone to identification due to the small number of these subjects within the ADAF. For officers, each of the ranks above Colonel were recoded as Colonel, and, for enlisted personnel, Chief Master

Sergeant (CMSgt) was recoded as Senior Master Sergeant (SMSgt). Table 2 summarizes the work completed for the AFPC database.

Table 2. AFPC Database Modifications

Variable	Modification
Rank	Converted CMSgt to SMSgt & all General Officers to Colonel
Installation	Coded with random 4-digit codes
Installation Kind	Coded with random 3-digit codes
Unit	Coded with random 4-digit codes
SSN	Coded with random 6-digit codes & renamed as Subject ID
3-Digit PAFSC and DAFSC	Coded Commanders, Generals, CMSAF, First Sergeants, etc. with 888 identified as Command/Staff
PAFSC & DAFSC Skill Level	Converted G to zero to mask General Officers
Date of Rank	Converted this date to Time in Grade
Aviation Service Code Date	Removed variable
Name	Removed variable
Medical Dates of Care	Converted to AFPC Import Date or removed
Medical Record Numbers	Removed variables

4.4 Final Dataset Descriptions

After completing the de-identification process, removing invalid records, and deleting unnecessary variables, the study team completed the database-building phase. Table 3 describes the record counts and variable counts within each final database.

Table 3. Study Database Summary

Database	Records	Variables
AFPC	20,063,016	39
SADR	3,248,834	55
SIDR	44,870	13
AFRESS	21,950	4
PHA	708,088	45
SAFETY-INJURY	7,827	28
SAFETY-ILLNESS	897	30

The result of this process is a rich, multi-functional database that allows analysis of numerous research topics. For instance, the study team can either look at all of the ICD-9 codes together or select a specific subset to focus on its presence within the study.

4.5 Outcome Variables for STDs and Mental Disorders

To determine whether a subject was diagnosed with an STD or mental disorder at some point during this study, the study team developed groups of ICD-9 codes that would indicate either one of these outcomes.

For the STDs, the ICD-9 code list contained 152 unique codes identifying specific diseases such as syphilis, gonorrhea, trichomoniasis, and human papillomavirus. These codes are searchable in AFRESS, SADR, and SIDR data to identify subjects diagnosed with an STD during a particular month and year in this study.

For the mental disorders, the ICD-9 code list contained 511 unique codes identifying specific disorders such as schizophrenia, bipolar disorder, depressive affective disorder, and post-traumatic stress disorder. As with STD coding, researchers can search for these codes in SADR and SIDR to identify subjects diagnosed with a mental disorder.

4.6 Comparison of Study Data to Air Force Almanac Data

To determine that our study data accurately represented the ADAF from 2006 – 2010, our summary counts were compared to those published in Air Force Magazine's Annual USAF Almanac (2006 – 2008) and the USAF Statistical Digest (2009 – 2010). Since the almanac reported data available from September 30th of each year, the study team analyzed demographic data from records in September of each study year. Counts were compared within each rank (broken down by gender). The study team also looked at education levels, marital status, average age, and two-digit DAFSC breakdown for both officers and enlisted personnel.

For counts within each rank, the numbers closely agree. However, there are higher counts in the study data when compared to almanac data for senior ranks in both officer and enlisted. For some of these subjects, there is an AFPC record for September, but no record for October, which indicates that subject's retirement or discharge from active duty. When data are collected for the almanac, these subjects may not be included in the counts.

When comparing education levels, marital status, and average age, the study counts closely mirrored the almanac numbers. In 2009, the almanac reported no enlisted personnel with a PhD or professional degree, while our data suggested there were 23 subjects with this type of degree.

For counts within the two-digit breakdown of officer and enlisted career fields, there were a few instances where our counts differed by 10% or more. Enlisted DAFSC counts agreed for all of the two-digit career fields except for the 1T and 2P groups in 2008, 2009, and 2010. Study counts were rechecked with no discrepancies found. It is interesting to note that the PAFSC numbers matched up better with the almanac counts for these career fields. Officer DAFSCs were more problematic. For 2006, the 16, 34, 35, and 37 career fields had large differences over 10%. For 2007, the 16, 35, and 37 career fields were over 10%. For 2008, 16 and 35 were over 10%. For 2009 and 2010, only the 16 career field had a large difference. All of these differences indicated there were more counts in our data compared to the published data. Again, the study counts were rechecked with no discrepancies found. It is interesting to note that the almanac summary contained an "Other" category with 1,448 subjects for 2006, 1,552 for 2007, and 464 for 2008. In addition, a "Commander and Director" category contained 1,305 for 2006, 1,303 for 2007, and 2,454 for 2008. The USAF Statistical Digest contained a category "Unknown" with counts of 2,126 for 2009 and 2,447 for 2010. These categories may account for the differences.

4.7 Creation of Rating Chain Variable

Discussion among the investigative team led to inclusion of an additional variable, within the AFPC data, that would allow analysis of the leadership structure and its effect on the health and safety of subjects in the study. For each subject, there may be an assigned rater (if one exists) who evaluates their performance. Most often, this is the direct supervisor. These data were obtained from AFPC, cleaned, and formed into a relational data table for linking to the appropriate AFPC record. When completed, 16,701,196 records (83.2% match) were assigned a rater who was also a subject within the study.

5.0 DISCUSSION

The multiple steps to de-identify personnel in the database ensured removal of personal identifiers to the maximum extent reasonably possible without compromising the integrity of the analysis. The completion of the database and the modifications described previously allowed the study team to continue with the data analysis phase of the project (Phase II). Data analysis will include an examination of occupational illness and injury as reported to the Air Force Safety Center. The study team will utilize regression analysis to identify which career fields are associated with high risk of occupational illness and injury. Once these high-risk populations have been identified, multivariable regression analyses will be performed to identify specific environmental and occupational exposures that directly relate to an increase in occupational injury and illness in these populations. Next, researchers will conduct detailed analyses that will combine the results of the regression analysis with potential outcomes of interest. The primary methodological focus will be on predicting both positive and negative health and safety outcomes for ADAF members, employing two primary approaches: regression analysis and hazard function analysis. Lastly, the study team will look at common variables between models produced above. This will allow identification of key drivers of risk for ADAF members. From this, results may be utilized to form policy recommendations that may allow the USAF to reduce risks for its Airmen. For instance, if the study team finds that particular career fields are a common element in negative health outcomes, recommended screening or prevention programs may be targeted to those particular career fields. If it is found that peer support partially mitigates risk-taking behaviors, a potential recommendation would be creating and promoting peer counseling and support groups.

The completed database will also provide the opportunity to describe the current utilization of military medical care by ADAF members through examination of direct outcomes that occur because of an occupational injury or illness, as well as indirect outcomes that may manifest because of risk-taking behavior or additional occupational stressors. The study team expects to identify the most current medical and personnel data that are available for the purposes of this study. Through the phases of the study outlined previously, the study team will be able to characterize the occupational experience of high-risk career fields with respect to illness and injury. In addition, the study team will be able to identify demographic variables associated with these occupational injuries and illnesses, as well as occupational stressors that may increase an individual's risk of injury or illness.

6.0 CONCLUSIONS

This study will allow the development of pathways toward occupation-related human performance improvement by tailoring specific counseling and/or prevention programs that may be implemented to reduce the stress and stress-related outcomes experienced in specific occupations within the Air Force community, in both garrison and deployed environments. The results of this study can be used to develop prevention strategies that can be presented to Air Force leaders as policy recommendations to ensure that Air Force members are able to operate efficiently and ensure full mission capabilities. The identified policy recommendations will be routed through the Air Force Surgeon General's office upon completion of the study.

7.0 REFERENCES

1. Bridger RS, Brasher K, Dew A, Kilminster S. Occupational stress and strain in the Royal Navy 2007. Occup Med (Lond). 2008; 58(8):534-539.
2. Pflanz SE, Sonnek S. Work stress in the military: prevalence, causes and relationship to emotional health. Mil Med. 2002; 167(11):877-882.
3. Mausner-Dorsch H, Eaton WW. Psychosocial work environment and depression: epidemiologic assessment of the demand-control model. Am J Public Health. 2000; 90(11): 1765-1770.
4. Reed PL, Storr CL, Anthony JC. Drug dependence enviromics: job strain in the work environment and risk of becoming drug-dependent. Am J Epidemiol. 2006; 163(5):404-411.
5. Hellerstedt WL, Jeffery TW. The association of job strain and health behaviours in men and women. Int J Epidemiol. 1997; 26(3):575-583.
6. Schulte PA, Wagner GR, Ostry A, Blanciforti LA, Cutlip RG, et al. Work, obesity, and occupational safety and health. Am J Public Health. 2007; 97(3):428-436.
7. Grunberg L, Moore S, Anderson-Connolly R, Greenberg E. Work stress and self-reported alcohol use: the moderating role of escapist reasons for drinking. J Occup Health Psychol. 1999; 4(1):29-36.

LIST OF ABBREVIATIONS AND ACRONYMS

ADAF	active duty Air Force
AFPC	Air Force Personnel Center
AFRESS	Air Force Reportable Events Surveillance System
AFSC	Air Force Safety Center
CY	calendar year
DAFSC	Duty Air Force Specialty Code
DOB	data of birth
ICD-9	International Classification of Diseases, Ninth Revision
PAFSC	Primary Air Force Specialty Code
PHA	Preventive Health Assessment
SADR	Standard Ambulatory Data Record
SIDR	Standard Inpatient Data Record
SSN	Social Security number
STD	sexually transmitted disease
USAF	U. S. Air Force