Millennium Cohort Program-VA Merged Data

Data User Guide

September 2021

Version 2.0

Department of Veterans Affairs

Version History:

Date	Version	Description	Author(s)
May 19, 2021	1.0	V1.0 final review	Nicole Stamas, MSc
May 24, 2021	1.0	V1.0 release	Nicole Stamas, MSc
September 2021	2.0	Updated details on Nosos and CAN scores, DMDC data, and 2021 transfer of MCS data to VA	Nicole Stamas, MSc

Table of Contents

1. List of abbreviations

ADUSH	Assistant Department Under Secretary for Health for Policy & Planning
BDN	Benefits Delivery Network
BIHR	Birth and Infant Health Research program (formerly Birth and Infant Health Registry)
BIRLS	Beneficiary Identification Records Locator System
CAM	Complementary and alternative medicine
CAN	Care Assessment Needs score
CCN	Community Care Network
CMS	Centers for Medicare & Medicaid Services
CSP	VA Cooperative Studies Program
DART	Data Access Request Tracker
DMDC	Defense Manpower Data Center
DoD	Department of Defense
DUA	Data Use Agreement
FBCS	Fee Basis Claims System
IAA	Interagency Agreement
III	Individually identifiable information

ISI	Insomnia Severity Index
MCP	Millennium Cohort Program
MCS	Millennium Cohort Study
MDR	Military Health System Data Repository
MISSION Act	VA Maintaining Internal Systems and Strengthening Integrated Outside Networks
MVC	Motor vehicle crashes
MVI	Master Veteran Index
NDI	National Death Index
NHANES	National Health and Examination Survey
NHRC	Naval Health Research Center
OCC	VA Office of Community Care
ORD	VA Office of Research and Development
PCL-C	PTSD Checklist – Civilian version
PDHS	Post-Deployment Health Services
PHI	Patient health information
PHQ	Patient Health Questionnaire
PIT	Program Integrity Tool

PRIME-MD	Primary Care Evaluation of Mental Disorders
SAFE	Secure Access File Exchange
SF-36V	Medical Outcomes Study Short Form 36-item for Veterans
SPA	Specific Project Agreement
SSA	Social Security Administration
USRDS	United States Renal Data System
VA	Department of Veterans Affairs
VBA	Veterans Benefits Administration
VETSNET	Veterans Service Network
VHA	Veterans Health Administration
VINCI	VA Informatics and Computing Infrastructure
VIReC	VA Information Resource Center
VistA	Veterans Health Information Systems and Technology Architecture
VSF	Vital Status File

2. List of contacts

Edward Boyko, MD, MPH	Study Chair, CSP #505	Edward.Boyko@va.gov
Amber Banerjee, MPH	Project Manager, CSP #505	Amber.Banerjee@va.gov
Elaine Hu	Data Analyst, CSP #505	Elaine.Hu@va.gov
Rudy Rull, PhD, MPH	Principal Investigator, Millennium Cohort Study	
Valerie Stander, PhD	Principal Investigator, Millennium Cohort Family Study	

3. Introduction

This Data User Guide serves as a comprehensive resource for investigators using Millennium Cohort Program (MCP)-VA merged data for research. The MCP-VA merged data refers to Millennium Cohort Program survey data linked to electronic VA and Department of Defense (DoD) data, such as personnel records, deployment dates, DoD medical records, VA medical records, and VA benefits information.

The Millennium Cohort Program is comprised of both the Millennium Cohort Study and the Millennium Cohort Family Study. The Millennium Cohort Study, launched in 2001, is the largest DoD longitudinal cohort study, enrolling military personnel while in service and following them after separation from the military and through their transition back to civilian life. The primary objective of the Millennium Cohort Study is to determine the short and long-term associations between military risk factors including service branch, deployment history, combat experience, and military occupation and the development of chronic disease (1). As of June 2020, 201,620 service members across four panels have enrolled in the Cohort and over 70% have become veterans*. Additional enrollment is currently ongoing, and follow-up of this Cohort is planned to continue through 2068. The Millennium Cohort Family Study (Family Study) is a parallel cohort of

*__Special note__: Veterans with an upper-case 'V' is used to indicate veterans who use the VA (VHA and/or VBA) while veterans with a lower-case 'v' is used to indicate the general population of separated service members regardless of VA utilization or eligibility.

the Millennium Cohort Study in which a sample of spouses of Millennium Cohort Study participants complete triennial surveys regarding the impact of military service on family well-being. The primary aim of the Family Study is to determine the long-term association between military experiences, particularly combat deployment, and the health and well-being of service members and their families (2). As of June 2020, the Family Study consisted of 9,872 spouses of Millennium Cohort participants. Family Panel 2 enrollment is currently ongoing and follow-up on the Family Cohort is planned to continue through at least the year 2032.

The following Data User Guide outlines all major components of the MCP-VA merged data, including an overview of VA data sources, Millennium Cohort Study and Family Study data, Cohort matching and identification, data transfer, data interpretation and dissemination best practices.

4. Overview of data sources

4.1. VA data sources

4.1.1. CDW

The Corporate Data Warehouse (CDW) functions as a repository of data from multiple sources within the VHA enterprise, with most data obtained from VistA, the VA electronic health record system (3). CDW is national-level, relational database that houses VHA clinical, enrollment, financial, administrative and utilization data. Raw data domains are extracted directly from VistA and not developed for research and production. Production data domains have been uploaded from VistA and cleaned and indexed for query efficiency (3). Examples of production domains commonly used for VA research include Inpatient, Outpatient, Health Factors, Lab Chemistry, Purchased Care, Allergy, and Immunization (3).

More information on CDW's VHA database including a conceptual overview of CDW architecture, uses, and basic SQL queries can be found at: https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=2287 and https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/926-notes.pdf.

Tips on matching and identifying encounters and special cases may be accessed by VA investigators in the CDW Researcher's Notebook (https://vaww.virec.research.va.gov/Notebook/Overview.htm).

Assessing alcohol using AUDIT-C in CDW: Health Factors table vs Mental Health domain

Alcohol use is assessed within VHA records using the AUDIT-C assessment. The AUDIT-C is the gold standard alcohol screen for identifying problematic drinking behaviors or active alcohol use disorders (including alcohol abuse or dependence). VHA providers conduct universal annual AUDIT-C screening on patients. The AUDIT-C is made up of 3 standard questions with answers ranging from values 0-4 (4). The possible range of AUDIT-C scores is 0-12. Among men, a score of 4 or more is the threshold to indicate a positive screen (4). Among women, a score of 3 or more indicates a positive screen (4). It is widely understood that the higher the AUDIT-C score, the more likely it is that the patient's drinking behavior is affecting overall health and safety (4).

AUDIT-C scores for Veterans can be found in both the Health Factors and Mental Health production domains of CDW. Within the Health Factors domain, screening results for Veterans who completed the AUDIT-C assessment are sometimes presented as score ranges (e.g. "AUDIT C RESULTS 3-7," "AUDIT-C SCORE >=8") rather than exact assessment score. Some provided ranges such as "AUDIT-C RESULTS 16-40" and "AUDIT C RESULTS 8-15" appear to exceed the maximum possible score of 12 for an AUDIT-C assessment. Dates and times for AUDIT-C surveys are masked in the Health Factors domain, preventing the researcher from matching surveys to participants.

CSP #505 has previously opted to use CDW's Mental Health domain to obtain AUDIT-C scores. The Mental Health domain provides variables for score per question, summary score, and raw score. Raw score in the SurveyResult table should *not* be relied on solely as the variable entries appear inconsistent and exceed maximum range, particularly among older scores. Invalid entries (unfinished, exceeding maximum score) will still appear as a raw score in the SurveyResult table. CSP #505 has found that the most reliable and consistent AUDIT-C scores may be produced by calculating the sum of individual, item-level scores provided in the Mental Health domain.

More information on assessing AUDIT-C scores can be found in the VA HSR&D cyberseminar titled "Using AUDIT-C Alcohol Screening Data in VA Research: Interpretation, Strengths, Limitations & Sources": https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=378.

More information on the Mental Health domain of CDW including data elements (including AUDIT-C), uses, and limitations can be found in the VA HSR&D cyberseminar titled "A Practical Guide to Using the CDW Mental Health Domain": https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=1190.

4.1.2. ADUSH Enrollment Files

The VHA Office of the Assistant Deputy Under Secretary for Health (ADUSH) for Policy and Planning releases a monthly file and an annual file known as the ADUSH Enrollment Files (5). The ADUSH files are used primarily to investigate health care expenditures, enrollment, and patient characteristics for policy and budget planning (5,6). The ADUSH Enrollment Files capture VHA health care utilization, VHA eligibility, demographics, cost and location information for all Veterans and some non-Veterans who have received any VA care, regardless of VHA enrollment status (5,6). Cost of inpatient, outpatient, psychological, and surgical care received within VHA are available in the ADUSH files. The files also include variables of non-VA care paid for by the VA, or VA community care. Examples of variables include VHA enrollment status, service-connected disability rating, priority group, Means Test score, and Medicaid eligibility status. Once individuals are added to the file, they are not removed from the updated yearly file unless they are known to be deceased or ineligible (6). A 2013 ADUSH data dictionary and information on enrollee classification, and ADUSH file creation can be found in the VIReC Research User Guide: ADUSH Enrollment Files (https://vaww.virec.research.va.gov/RUGs/ADUSH/RUG-ADUSH-EF-FY99-12-RA.pdf).

4.1.3. VSF: Master & Mini

The VHA Vital Status File (VSF) contains mortality data from several official data sources including VHA and VBA data, Master Veteran Index (MVI), Beneficiary Identification Records Locator Subsystem (BIRLS) Death file, Centers for Medicare and Medicaid Services (CMS) Medicare Vital Status File, and the Social Security Association (SSA) Death Master File (7). The VSF was developed as a VA project in collaboration with the National Death Index (NDI) to evaluate the completeness and validity of VA death data compared to NDI death data (7). The VSF includes individuals who have received VHA care since 1992, are enrolled in the VHA, or have received VBA benefits since 2002 (7). The VSF is split into the Master file and the Mini file. The Master file includes Veterans and non-Veterans with one row per SSN-DOB-gender match across all data sources. Approximately 40% of unique SSNs within the Master file have more than one mortality record across multiple sources (8). The Mini file includes only Veterans with one row per SSN match determined by the VA National Data Systems to be the “best” match (8). Variables included in the VSF include demographics, healthcare and benefits activity, and birth and death dates. While the Master file contains 125 total variables, the Mini file contains only 16 variables (including only “best” birth and death dates) (8).

More information on strategies utilizing VSF, identifying Veterans with Medicare, and matching a cohort may be found at: https://vaww.virec.research.va.gov/VSF/Research-Utility.htm.

More information on obtaining vital status and mortality among Veterans can be found in the VA HSR&D cyberseminar titled "Ascertaining Veterans' Vital Status" : https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/1242-notes.pdf.

4.1.4. Fee and PIT

Choice and MISSION Acts

Non-VA community care is care provided by a non-VA facility that is paid for by the VA under special circumstances and eligibility criteria (9,10). The Choice Act of 2014 allowed for Veterans to choose to receive care at a non-VA facility based on their residence or if they are unable to receive an appointment at the closest VA facility within 30 days (11). On June 6, 2019, the Choice Act of 2014 was replaced by the VA Maintaining Internal Systems and Strengthening Integrated Outside Networks (MISSION) Act (10,11). The VA MISSION Act served to expand and streamline the VA Community Care Network (CCN) to allow Veterans to access non-emergency care and schedule their own appointments at any facility within the VA network without prior authorization from the VA (10). Like the Choice Act, the Veteran must meet distance "drive time" thresholds (>= 30 minute average drive time for primary care or >=60 minute average drive time to specialty care) to a VA facility with necessary services in order to be eligible for community care under the MISSION Act (10). In addition, Veterans are eligible for community care if a VA facility cannot provide it for them within 20 days (for primary care, mental care or extended care) or 28 days (for specialty care) from the date of request or if the care is not available from their local VA facility at all (10).

Fee to PIT

The Fee Basic Claims System (FBCS), or purchased care/fee data, has historically aggregated non-VA community care claims as part of the VA Office of Community Care (OCC) (12). In FY2016, the Non-VA Program Integrity Tools (PIT) system was developed to consolidate community care claims to replace the FBCS (12,13). PIT was officially released into the CDW work environment in 2018 (13). Since 2018, there has been a data migration of claims to PIT rather than fee, resulting in a need to access both PIT and fee systems to capture community care utilization and costs dating beyond 2018. There is significant overlap in claims between the two sources, with 86-93% of claims found in both fee and PIT systems (14).

PIT Claims

The PIT system is a dynamic environment as it is updated daily with new claims once they have been processed (14). Claims within PIT have three values: accepted, denied, and rejected (14,15). If the claim is accepted, the facility delivering the care will be paid and the Veteran will be reimbursed if he/she paid for the care. Claims may be denied if the claimant did not meet the filing deadline, a VA facility with the same service was available, or the Veteran did not meet the eligibility requirements of accessing community care (15). Claims may be rejected if there is some missing information or errors and the claimant may be asked by the VA to resubmit the claim within a given time frame (15). Claims with statuses other than rejected, accepted, or denied are not available in PIT (14). Alternative claim statuses include 'suspended', 'suspended-verify', and claims that have not been given a status name (14). Both professional and institutional claims exist within the PIT system and may refer to one or more medical encounters in which both the provider and the institution billed the VA for community care services provided to a Veteran.

Methods for identifying Choice claims vs MISSION Act claims within PIT, understanding the differences between institutional and professional claims, and identifying costs and procedure codes within PIT may be found here: https://www.herc.research.va.gov/include/page.asp?id=choice-pit.

4.1.5. VETSNET

The Veterans Service Network Corporate Mini Master File (VETSNET) is a dataset that extracts VBA compensation, benefits, and pension payments to Veterans or their beneficiaries (16). VA has been working to implement VETSNET to completely replace the Compensation and Pension Benefits Delivery Network (BDN) since 1996 (16). VA now estimates that VETSNET holds 99% of all Veteran compensation claims and 94% of all Veteran pension claims (16). VETSNET also includes eligibility information such as VBA disability ratings, service and non-service-connected disabilities, military service, education, and burial data (17). Diagnosis codes, procedure codes, and eligibility codes for disability and service-connected compensation or pension are also available. The VETSNET file is a dynamic file that is updated monthly. The following link provides multiple VETSNET user guides: https://vbaw.vba.va.gov/bl/21/Systems/awards.htm.

Service-connected disability rating in ADUSH, VETSET, and CDW

A service-connected disability entitles Veterans to specific disability compensation and benefits. Service-connected disability ratings range from 0-100% by deciles. A disability rating is granted for medical conditions that develop during military service or as a result

of military service. The disability percent refers to the impairment rendered by the condition after military service ends. Service-connected disabilities are granted through a review process of a disability claim submitted by a veteran or by VA policy that automatically grants a disability for a condition that occurred in a veteran subjected to a certain exposure (e.g., agent orange). Many Veterans have multiple service-connected conditions and related disabilities. Service-connected disability ratings may change over time and can be upgraded or downgraded as a disability improves or worsens. New service-connected disabilities can occur in response to a Veteran claim or by VA policy. Disability ratings are *not strictly* additive (18). In other words, a Veteran with a 20% service-connected disability and a second 60% service-connected disability does not necessarily receive benefits at the 80% disability tier (18). First, disabilities are arranged in order of severity, then a Combined Ratings Table is used to determine combined disability rating for two or more conditions (18). For example, a Veteran with two service-connected conditions rated as 30% and 20% disability, respectively, is entitled to a 44% combined disability rating, which is rounded to 40%. All combined disability ratings are rounded to the nearest 10% (18).

A Veteran with a non-compensable, or 0%, service-connected disability indicates a Veteran with a disability that is service-connected but does not entitle them to monetary payments (19). For example, a Veteran may report hearing loss as a result of their service but does not suffer any severe ear injuries or present hearing loss symptoms deemed severe enough to entitle the Veteran to VA disability payments. A Veteran with a non-compensable service-connected disability may be entitled to other benefits such as travel allowance and no cost healthcare and prescription drugs for service-connected conditions (19). Individuals with two or more separate 0% service-connected disabilities may be entitled to compensation under the combined service-connected disability rating minimum of 10% (19). A Veteran with no documented service-connected condition should be demonstrated with a 'NULL' value within disability data.

Service-connected disability rating can be found in the ADUSH enrollment files, VETSNET and the CDW's Spatient table. ADUSH files include monthly files and an annual cumulative file comprised of one record of service-connected disability rating per person per year updating enrollees and service-connected data from each monthly file produced through the fiscal year. ADUSH contains historical disability ratings. VETSNET and CDW contain only the most current disability rating which overwrites an outdated rating, resulting in a lack of historical service-connected disability data.

Please reference variable source table for more information (see 12.1 Data sources for select MCS-VA variables for research on service members and Veterans in Appendix A).

4.1.6. Centers for Medicare and Medicaid Services (CMS)

VA has partnered with the Centers for Medicare and Medicaid Services (CMS) to acquire CMS and United States Renal Data System (USRDS) data to capture a more complete picture of healthcare utilization among Veterans. USRDS compiles information regarding individuals with chronic kidney disease and end-stage renal disease in the United States. All individuals with end-stage renal disease are eligible for Medicare, regardless of age, hence the inclusion of USRDS data within the VA/CMS data enclave. VA/CMS data, managed by VIReC for VA research use, include the following domains: Medicare, Medicaid, USRDS, Patient Assessment, Healthcare Effectiveness data and Information Set (HEDIS), Medicare Current Beneficiary Survey data (MCBS), and CMS Provider data. The Medicare and Medicaid domains include demographics, prescription information, fee claims, inpatient and long-term care information for Medicare and Medicaid enrollees. USRDS data includes demographics, treatment, transplant, and Medicare claim information for all those in the USRDS database. VIReC also hosts a VA/CMS Match Indicator File containing one record for every SSN match between VHA Cohort (Veterans enrolled in VHA) and Veterans enrolled in either Medicare or Medicaid.

The raw VA/CMS data is available exclusively for use among VA researchers and cannot be shared outside of VA. For investigators outside of the VA interested in using VA/CMS data for research, please contact the CSP #505 team to discuss steps forward.

VA researchers can access more information on available VA/CMS data and data request information at https://vaww.virec.research.va.gov/Index-VACMS.htm.

4.1.7. Risk adjustment: Nosos and CAN scores

Risk adjustment scores are used to understand a patient's clinical and care needs within a healthcare setting. The most commonly used scores at VA are Nosos and CAN risk scores.

Nosos scores

The Centers for Medicare and Medicaid (CMS) Hierarchical Condition Categories (HCC) version 21 Nosos scores are used to predict the annual VA cost for patients. "Nosos" is the Greek word for "chronic disease" (34). Nosos scores use diagnoses, age, gender, pharmacy records, priority status, and VA-computed costs to predict risk of cost. Nosos scores begin as HCC risk scores which use only diagnoses, age, and gender to compute predictive cost. Then the additional pharmacy and VA-specific variables are added into the HCC risk model to become Nosos scores. Nosos scores are rescaled so that the entire population mean score always equals 1 (34). Nosos

scores may be requested for VA research purposes by initiating a DART request for the Health Economics Resource Center (HERC) V21 and Nosos Risk Scores. For more information on the construction and origins of Nosos scores, please see the following reference:

"Risk Adjustment Tools for Learning Health Care
Systems: A Comparison of DxCG and CMS-HCC V21" (Wagner et al., 2016, *Health Services Research*; DOI: 10.1111/1475-6773.12454).

CAN scores

The Care Assessment Needs (CAN) Primer is a predictive score used to measure risk of future adverse events among Veterans to identify populations most likely to experience hospitalization, death, or a combination of hospitalization and death (35). Predictive scores for these three adverse outcomes are modeled at 90 days and 1 year, for a total of six possible outcomes (35). Variables considered for the models are derived from Veteran-specific demographics, health care utilization, medical conditions, medications, vital signs, and lab data. All Veterans included in the models have their scores for all six outcomes ranked from lowest to highest and are placed into percentiles from 0 (lowest risk) to 99 (highest risk). These CAN scores appear in increments of 5 from 0 - 90 and increments of 1 from 95 – 99 (35). CAN scores should be interpreted as an estimated probability of the adverse outcome of interest. CAN scores have been created for all Veterans receiving primary care at a VA facility, totaling approximately 5.5 million Veterans as of 2019. CAN 2.5, the most recent version of the score, was released in in FY19. CAN 2.5 scores are available for research access on CDW. For scores generated between 2013 and 2017, monthly CAN 2.5 scores are available. For scores generated in 2018 and onwards, weekly CAN 2.5 scores are available. Older versions of CAN (1.0 and 2.0) are no longer available for access (35). For more information on the creation and origins of CAN scores please use the following reference:

Wang, Li, Brian Porter, Charles Maynard, Ginger Evans, Christopher Bryson, Haili Sun, Indra Gupta, et al. 2013. "Predicting Risk of Hospitalization or Death Among Patients Receiving Primary Care in the Veterans Health Administration." *Med Care* 51: 368–373.

4.2 Millennium Cohort Program surveys

4.2.1. Millennium Cohort Study survey Panels and Waves

As of June 2020, there were 201,620 total Millennium Cohort Study participants. The MCS consists of four complete enrollment "Panels," or participant groups recruited during the same enrollment window and administered an initial baseline questionnaire. The fifth Panel began enrollment in September 2020 and is expected to conclude enrollment in late summer/fall 2021. The launch times of Panels 1,2,3,4 and 5 are 2001, 2004, 2007, 2011 and 2020, respectively. The surveys are administered in "Waves," or time periods in which a survey is distributed to participants. Surveys for each Wave are distributed every 3-5 years. Figure 1 below shows the launch years and subsequent points of follow up for Panels 1-4 between the years 2001-2019. Panel 5 is currently enrolling with a recruitment goal of 50,000 participants. The Millennium Cohort Study is expected to continue until 2068. More information on participant demographics can be found in the following section (5.1. Millennium Cohort participants). Survey response rates among Millennium Cohort participants in Panels 1-4 have historically ranged from 46% to 71% of enrolled participants. Figure 2 below displays the follow-up survey response rates among Millennium Cohort participants through Panel 4.

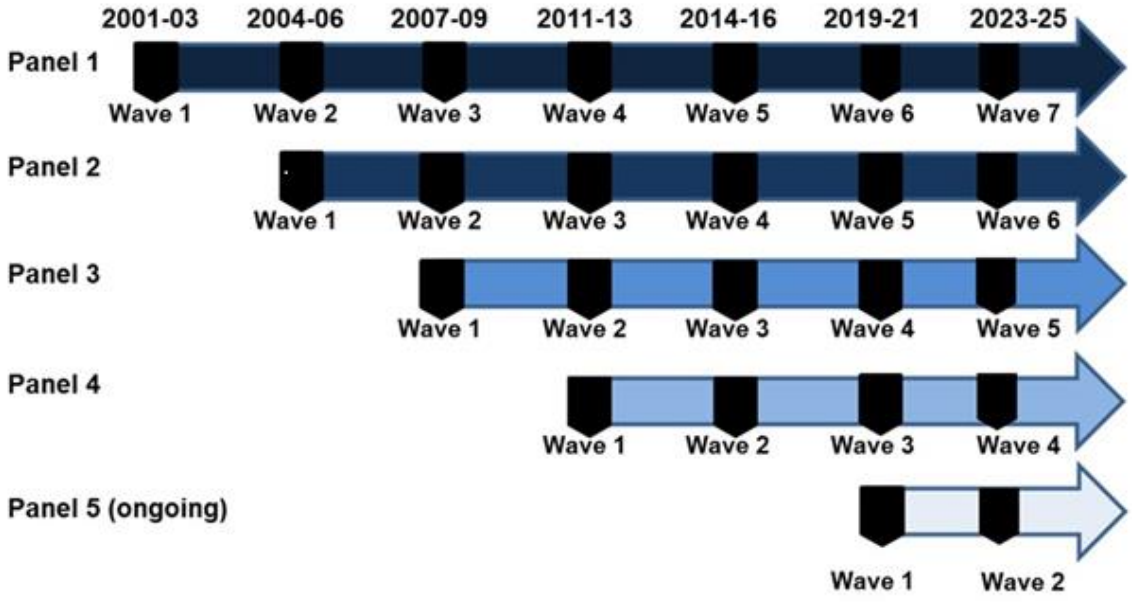

Figure 1. Panels and Waves of the Millennium Cohort Study updated 10/29/2020. Source: VA HSR&D cyberseminar "Overview of the Millennium Cohort Study: Opportunities for VA investigators" (https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/3878-notes.pdf).

Follow-up Response Rates Total # of Responders per Cycle (% follow-up rate)				
Panel (# in Panel)	**2004-2006**	**2007-2008**	**2011-2013**	**2014-2016**
Panel 1 (n=77,019)	55,021 (71%) Wave 2	54,790 (71%) Wave 3	51,678 (67%) Wave 4	51,146 (66%)
Panel 2 (n=31,110)	Enrollment of Panel 2	17,152 (55%) Wave 2	15,149 (49%) Wave 3	14,793 (48%)
Panel 3 (n=43,439)		Enrollment of Panel 3	22,071 (51%) Wave 2	19,991 (46%)
Panel 4 (n=50,052)			Enrollment of Panel 4	27,233 (54%)
Family Panel 1 (n=9,872)			Enrollment of Panel 1	6,789 (68%)

Figure 2. Number and % of total responders among Millennium Cohort participants in Panels 1-4 and Family Study Panel 1. This figure is based on the Department of Defense Millennium Cohort Study supporting documents, Part B: Collections of Information Employing Statistical Methods (https://www.reginfo.gov/public/do/DownloadDocument?objectID=82729703) .

4.2.2. Family Study survey Panels and Waves

Launched in 2011 as a parallel cohort study to the Millennium Cohort Study, the Millennium Cohort Family Study includes male and female spouses of active duty, Reserve, and National Guard personnel from the Army, Navy, Air Force, Marine Corps, and Coast Guard. Panel 1 of the Family Study consists of opposite-sex spouses of Panel 4 MCS participants (service members with two to five years of military service across branches and components). There were 28,603 Millennium Cohort Study participants with spouses eligible for the Millennium Cohort Family Study (22). Of this eligible population, 9,872 spouses responded to the Panel 1 Family Study baseline survey (22). The enrollment rate for Panel 1 of the Family Study was 34.5% (22). Among enrolled participants, 6,789 (68%) completed the first follow-up survey (see Figure 2 above). For more information on nonresponse bias, see 9.1 Potential sources of bias in MCP-VA merged data). Panel 2 of the Family Study began recruitment in December of 2020 and enrollment is ongoing.

4.2.3. Millennium Cohort Program participant consent forms

Millennium Cohort Study and Family Study participants are asked to complete a baseline questionnaire at enrollment and triennial follow-up surveys. All participants of MCS and Family Study provide written informed consent on their baseline questionnaire that also includes agreement to longitudinal follow-up (23). The consent form and Privacy Act statement may be found here:

https://www.millenniumcohort.org/files/milco/2019/consentform_and_privacyact.pdf.

4.2.4. Millennium Cohort Study survey topics

The Millennium Cohort Study baseline survey administered at enrollment is longer and collects more detailed demographic information than the follow-up surveys that inquire predominantly about health, wellbeing, and life events within the last 3 years (since the last survey cycle). Survey questions and topics have changed slightly with each survey launch. For example, questions on sleep quality were introduced to the Millennium Cohort Study surveys in 2014. The following topics are taken from the Millennium Cohort Study data dictionary updated in 2017 and represent general categories of subjects included on these surveys: demographic data, mental health, sleep measures, health conditions, pregnancy & women's health, physical activity & nutrition, complementary and alternative medicine (CAM), life experiences, alcohol consumption, tobacco use, social support measures, military-specific characteristics, deployment & environmental exposures, and injuries and motor vehicle crashes (MVC). A complete list of variables from the Millennium Cohort Study surveys can be found in the MCS data dictionary available upon request of the investigator. Creation of a complete data dictionary of all merged MCS-VA variables is in progress.

The health conditions listed on the MCS survey for participant endorsement are as follows: anemia, angina, asthma, Crohn's disease, chronic fatigue syndrome, coronary heart disease, high cholesterol requiring medication, chronic bronchitis, cirrhosis, cancer, depression, diabetes or sugar diabetes, emphysema, fibromyalgia, gallstones, heart attack, hepatitis B, hepatitis C, any other hepatitis, significant hearing loss, any other heart condition, hypertension, kidney failure requiring dialysis, infertility, kidney stones, degenerative joint disease, lupus, manic-depressive disorder, memory loss or memory impairment, migraine headaches, multiple sclerosis, neuropathy, pancreatitis, posttraumatic stress disorder, rheumatoid arthritis, acid reflux requiring medication, sleep apnea, schizophrenia or psychosis, sinusitis, stroke, seizures, traumatic brain injuries, thyroid condition other than cancer, tinnitus, ulcerative colitis or proctitis, stomach, duodenal, or peptic ulcer, bladder infection, significant vision loss, and other health condition accompanied by a free-text entry field.

4.2.5. Millennium Cohort Study standardized instruments

The major instruments used on the Millennium Cohort survey are the Patient Health Questionnaire (PHQ), the Medical Outcomes Study Short Form 36-item for Veterans (SF-36V), the Posttraumatic Stress Disorder (PTSD) Checklist – Civilian Version (PCL-C), PCL-5 (launched at Wave 6 timepoint), the Primary Care Evaluation of Mental Disorders (PRIME-MD), the CAGE questionnaire for problem drinking behavior, the Insomnia Severity Scale (ISI), the Holmes and Rahe scale for stressful life events, the Pearlin-Schooler Mastery Scale, and items derived from the National Health and Nutrition Examination Survey (NHANES) (21, 24).

On the Millennium Cohort survey, the PHQ is used to measure major depressive disorder, panic disorder, other anxiety syndrome, bulimia nervosa, alcohol misuse, and binge eating disorders (21). The PHQ yields high accuracy of 85% for classification of these disorders, a sensitivity 75%, and a specificity of 90%, with sensitivity and specificity fluctuating by condition (21). The SF-36V measures 8 defined components of physical functioning, role limitations caused by physical problems, pain, general health, vitality, social functioning, emotional functioning, and mental health (21). The SF-36V was developed specifically for military and veteran populations, and therefore is assumed to have high accuracy and consistency within the Millennium Cohort population. Using responses of the 8 components of health on the SF-36V, the Physical Component Score (PCS) and Mental Component Score (MCS) are calculated as composite scores to measure overall physical and mental functioning. The CAGE questionnaire (acronym for cut back, annoyed, guilty, eye-opener) measures alcohol misuse and problematic drinking behaviors. It has been found that there is low overall internal consistency among alcohol misuse responses between the CAGE and PHQ in the Millennium Cohort population, suggesting that participants demonstrate an inconsistent pattern of reported drinking (21). Inconsistently reported drinking may be attributed to fears of military career limitation as a result of these responses or trouble defining measurements for one “drink” (21).

4.2.6. Family Study survey topics

The Millennium Cohort Family Study aims to investigate the associations between military experiences and service member readjustment on families’ health and well-being (24). The Family Study baseline questionnaire includes survey topics aimed at capturing spouse physical health, spouse mental health and adjustment, spouses’ reports of their children’s mental/physical health and functioning, family functioning, and protective and vulnerability factors relating to family functioning as a result of military service (24). Specific survey questions on the baseline questionnaire include spouses’ demographics, physical and mental health, coping skills, life experiences, military

service (for spouses that also have military experience a.k.a. “dual military families”), marital relationship, service member’s deployment experience, return and reunion experiences post-deployment, and children’s physical, mental, and emotional health (24). Spousal and child medical encounter, pharmaceutical, and mortality data are assessed through DoD medical records and the Birth and Infant Health Research (BIHR) program (24). BIHR is used primarily by the Family Study to capture data regarding pregnancy, birth outcomes, and reproductive health among female spouses (24).

4.2.7. Family Study standardized instruments

The Family Study baseline survey uses many identical measures and standardized scales as in Millennium Cohort baseline survey to enable comparison of response within spousal dyad (22). Similar instruments used by both the Millennium Cohort Study and the Family Study include the SF-36V, PCL-C, PHQ, ISI, and CAGE questionnaire. Other standardized instruments presented on the Family Study are Adverse Childhood Experiences, Quality of Marriage Index, Family Adaptability and Cohesion Evaluation Scale, and Strengths and Difficulties Questionnaire (24). Family Study data can be linked to Millennium Cohort Study service member data, DMDC personnel data and military characteristics, military medical records (including TRICARE medical records for both the service member and spouse), and VA health and benefits records for the service member (24).

4.3. Other military data sources

The first transfer of data under the bilateral DUA from the Naval Health Research Center (NHRC) to VA CSP #505 occurred in August 2021. The transfer included Millennium Cohort Study survey data through 2016 as well as Military Health System Data Repository (MDR) inpatient, outpatient, and pharmacy data, Defense Manpower Data Center (DMDC) personnel, and DMDC separation data.

4.3.1. Military Health System Data Repository (MDR) data

Specific variables from MDR inpatient, outpatient, and pharmacy transferred from NHRC to VA can found in the MCP-VA merged data dictionary created by CSP #505. More information on MDR data can be found in the MDR data dictionary and the MDR Data User Guide.

4.3.2. Defense Manpower Data Center (DMDC) personnel data

DMDC houses information on identity, military characteristics, and various Department of Defense benefits of current and former US servicemembers. DMDC personnel data includes variables such as service branch, education level, and dates in and out of service. DMDC personnel data sent from NHRC to VA also includes monthly snapshots of servicemember activity status taken from administrative and pay files and basic active service date, 'BASD'. BASD may change over time within DMDC personnel files for reasons such as a break and return in service or corrected administrative errors. BASD also appears as survey data ('BASD_w1' – 'BASD_w5') which are filled from DMDC records. In most cases, BASD from DMDC and BASD from survey data should match. However, NHRC has discovered a few instances in which the two BASD variables may not match. BASD at baseline (BASD_w1) is taken from DMDC records at the time of survey invitation, not enrollment. Between invitation and enrollment, the servicemember's status may change and these changes will be reflected in BASD for the servicemember, but not BASD_w1. BASD survey data from waves 2 through 5 are more likely to match BASD from DMDC records as the BASD record date is pulled to match the survey date. However, discrepancies are possible as a result of updates or administrative errors.

5. Study population

MCP-VA merged data is comprised of data from Millennium Cohort Study or Family Study who have an existing VA record.

5.1. Millennium Cohort Study participants

The Millennium Cohort Study is the largest prospective cohort study among the US military population consisting of four completed enrollment Panels totaling 201,620 participants, as of 2020. Enrollment for Panel 5 of the Millennium Cohort Study closed in August 2021. Panel 1 was launched in 2001 and enrolled 77,047 (35.9% of invited service members). Those eligible for recruitment into Panel 1 included a weighted random sample of all US military rosters as of October 1, 2000 (including all service branches) acquired by DMDC records (20 MAK). Panels 1 oversampled for Reservists and National Guard service members, women, and those who had previously deployed to Bosnia, Kosovo, or the Persian Gulf (20 MAK). Panel 2 enrolled 31,110 participants (25.3% of invited service members) between 2004 and 2006 and restricted to service members with 1-2 years of military service as of October 2003 (25, 26). Panel 3 enrolled 43,440 participants (28.2% of invited service members) between 2007 and 2009and

restricted to service members with 1-3 years of military service as of October 2006 (25, 26). Panel 4 enrolled 50,052 service members between 2011 and 2019 (26). Panel 5 launched in September 2020 with a goal enrollment of 50,000 service members among the 500,000 eligible recruitment sample. Eligible participants for Panel 5 included current service members from all service branches with 1-5 years of military service while oversampling for women and married individuals (to increase the number of service members with spouses eligible for recruitment into the Family Study). Figure 3 below provides detailed information enrollment and oversampled populations for Panels 1-4.

The Millennium Cohort Study sample represents all branches of military service (Army, Navy, Coast Guard, Air Force and Marines) and all service components (active duty, National Guard, Reserves). Among Panels 1-4, 44% of participants are current or former members of the Army, 16% Navy, 29% Air Force, 9% Marine Corps, and 2% Coast Guard (26). As of October 2020, 65% of Millennium Cohort participants had ever

Panel (Group)	Enrollment Dates	Years of Service at Enrollment	Oversampled Groups	Roster Size (Date)	Total Contacted	Total Enrolled (%)
1	Jul 2001-Jun 2003	Unrestricted	Women Reserves/Guard Prior deployers*	256,400 (Oct 2000)	214,388	77,019 (36%)
2	Jun 2004-Feb 2006	1-2	Women Marine Corps	150,000 (Oct 2003)	123,001	31,110 (25%)
3	Jun 2007-Dec 2008	1-3	Women Marine Corps	200,000 (Oct 2006)	154,270	43,439 (28%)
4	Apr 2011-Apr 2013	2-5	Women Married	250,000 (Oct 2010)	247,266	50,052 (20%)

*Deployment to Southwest Asia, Bosnia, and/or Kosovo after August 1997

Figure 3. Sampling frames of Millennium Cohort Study Panels 1-4. Based on figures displayed in VA HSR&D cyberseminar "Overview of the Millennium Cohort Study: Opportunities for VA investigators.

been deployed (27). As of June 2020, 69.4% of total Millennium Cohort participants across all Panels were male (139,825) and 30.6% (61,794) were female. As of June 2020, the average age of Panel 1 participants was 52 years and the average age of Panels 2-4 was 36 years (27). While Panel 1 was a cross-section of all military service members, later Panels were restricted to those with fewer years of service, resulting in a younger age among Panels 2-4 compared to Panel 1. Enrolling service members with fewer years of service encourages the capture of participant longitudinal data

throughout the military life cycle: pre-deployment, deployment, post-deployment, and transition/readjustment period.

Millennium Cohort Study-VHA user population

Millennium Cohort participants who have ever utilized VHA or VHA-paid services by community care providers are included in available VHA records. To qualify for VHA services, a service member must have been discharged other than dishonorably from any service branch including Reserve and National Guard and must have served a minimum of 24 continuous months (unless the service member was discharged for injury or hardship). There are special circumstances that can impact VHA eligibility and VHA eligibility can change if Congress enacts reform. For example, all service members regardless of discharge status or length of service are eligible for medical support for sexual assault and PTSD. In a June 2020 pull of current Millennium Cohort participants in Panels 1-4 (N=201,620), 71% of MCS enrollees had separated from military service and 95% of those separated had a VHA record (27). The average age of MCS-VHA co-enrollees was 44 years as of June 2020. As of March 2021, the total number of MCS-VAH co-enrollees was 135,781 individuals.

Top outpatient diagnoses among MilCo participants who had at least one VHA outpatient encounter up to FY2019 included chronic PTSD, depressive disorder, low back pain, and hypertension. Top inpatient diagnoses among MilCo participants who had at least one VHA inpatient encounter up to FY2019 include PTSD, alcohol dependence, major depressive disorder, chest pain, and opioid dependence. As of June 2020, 13,465 Millennium Cohort participants had a 100% service-connected disability rating and 3,686 participants had a 0% service-connected disability rating. For more information on service-connected disability rating please refer to section 4.1.5. VETSNET.

5.2. Family Study participants

Panel 1 of the Family Study was sampled from spouses of Panel 4 Millennium Cohort participants (service members with 2-5 years of service). Panel 4 of the Millennium Cohort Study was oversampled for married and female service members to ensure a large and representative sampling pool for Panel 1 of the Family Study (24). Married Millennium Cohort participants enrolled in Panel 4 were asked at enrollment if they would like to refer their spouse. A total of 28,603 MCS service members were eligible to refer their spouse (24). In an effort to increase enrollment numbers and decrease the risk of referral bias, the sampling method was altered to enroll MCS spouses by both

Panel 4 MCS participant referral and by direct invitation to the spouse (24). Potential referral bias resulting within the Family Study is described in more detail in 9.1 Source of bias in MCP-VA merged data.

Panel 1 of the Family Study consists of 9,872 (35% of total contacted for recruitment) spouses of Millennium Cohort Panel 4 participants (27). Beginning in fall of 2020, the research team contacted 185,000 service members eligible for Millennium Cohort Study participation with spouses potentially eligible for participation in Panel 2 of the Family Study (25). Enrollment is expected to continue into 2021. Eligible participants for the Millennium Cohort Family Study include male and female spouses of service members or veterans from all service branches who are currently enrolled in the Millennium Cohort Study (22). Only heterosexual couples are included in the study and therefore results are not generalizable to single-parent families, families with couples that are gay or lesbian, or individuals that are transgender or non-binary (22). There is currently research being conducted among LGBTQ individuals in the Millennium Cohort Study.

Survey data from Family Study participants is linked to data from spousal Millennium Cohort participants and DoD and DMDC health records creating a dyadic database (28). Of the 9,872 participants in Family Study Panel 1, approximately 10% are current or former service members (28). Among Family Study Panel 1 responders, 86% are female and more than 46% are spouses of members of the U.S Army (29). As of June 2020, among all current participants in the Millennium Cohort Program, there are 260 participants that are dual enrolled in both Millennium Cohort and Family studies. These “duals” are service members or veterans who are enrolled in the Millennium Cohort Study and were married to Panel 4 Millennium Cohort participant and enrolled in the Family Study.

Family Study-VHA user population

As of March 2021, the total number of Family Study- VHA co-enrollees was 996 individuals. The average age of all living Family Study participants who had any record of VHA utilization was 38 years. There are special instances in which a spouse or child of a current or former service member may have an encounter with VA services. Such instances are rare and include circumstances in which services are not available at a DoD military treatment facility or other local facility and only available at the local VA. In addition, approximately 10% of Family Study members are former/current service members and may be eligible for VHA utilization.

6. Data transfer

6.1. Data sharing agreements

Bilateral data sharing between VA and NHRC

CSP and NHRC have a bilateral Data Use Agreement (DUA) that allows for the transfer of individual-level data between institutions. The DUA was initially executed in August 2018 and a major modification, allowing for the transfer of MCS study data to CSP, was executed in August 2020. Under the terms of the DUA, CSP #505 shares health status, healthcare, and benefits information from VA with NHRC on all Millennium Cohort participants who have VA records. Ultimately, CSP #505 seeks to host Millennium Cohort Study data on the secure VA VINCI server to allow direct access to VA investigators who are approved to work with data (see 8.2. VINCI: Towards VA investigator workspace). The first transfer of Millennium Cohort Study survey data to VA occurred in August of 2021. This data transfer also included Military Health System Data Repository (MDR) inpatient, outpatient, and pharmacy data and Defense Manpower Data Center (DMDC) personnel files and separation data.

Oversight & Collaboration between ORD-PDHS-NHRC

An Interagency Agreement (IAA) between VA's Office of Research & Development (ORD), Post Deployment Health Services (PDHS), and NHRC allows VA to support analysts at NHRC to perform Veteran-focused studies. A VA-DoD MCS working group currently led by Ed Boyko, MD, MPH, oversees the collaboration. The group meets regularly and includes representation from PDHS, ORD, CSP #505, and NHRC including NHRC analysts funded under the IAA. An Executive Committee comprised of VA investigators and leaders from CSP, PDHS, and NHRC provides oversight and guidance to CSP #505. The CSP #505 Executive Committee meets quarterly to discuss CSP #505 updates and review any project proposals seeking approval for the use of merged MCS-VA data (see 8. Investigator data access).

6.2. Data flow & security

VA data shared with NHRC includes information on physical, mental, and reproductive health; functional health; healthcare utilization; and receipt of benefits. VA data from ADUSH, VETSNET, Vital Status Files, and other CDW files may be shared with NHRC under the existing agreement. VA completes a large-scale transfer of inpatient, outpatient, and mortality data among Millennium Cohort participants to NHRC annually. NHRC provides CSP #505 study staff with Millennium Cohort and Family Study participant SSN, date of birth, gender, and name to identify study participants in VHA.

7. Matching MCP participants with VA records

Please note that that Figure 2 below shows the general strategy for matching MCP participants with VA records using personal identifiers. The numbers shown in Figure 2 are subject to change and ***not necessarily reflective of the current MCP-VA population***.

7.1. Matching using identifiers

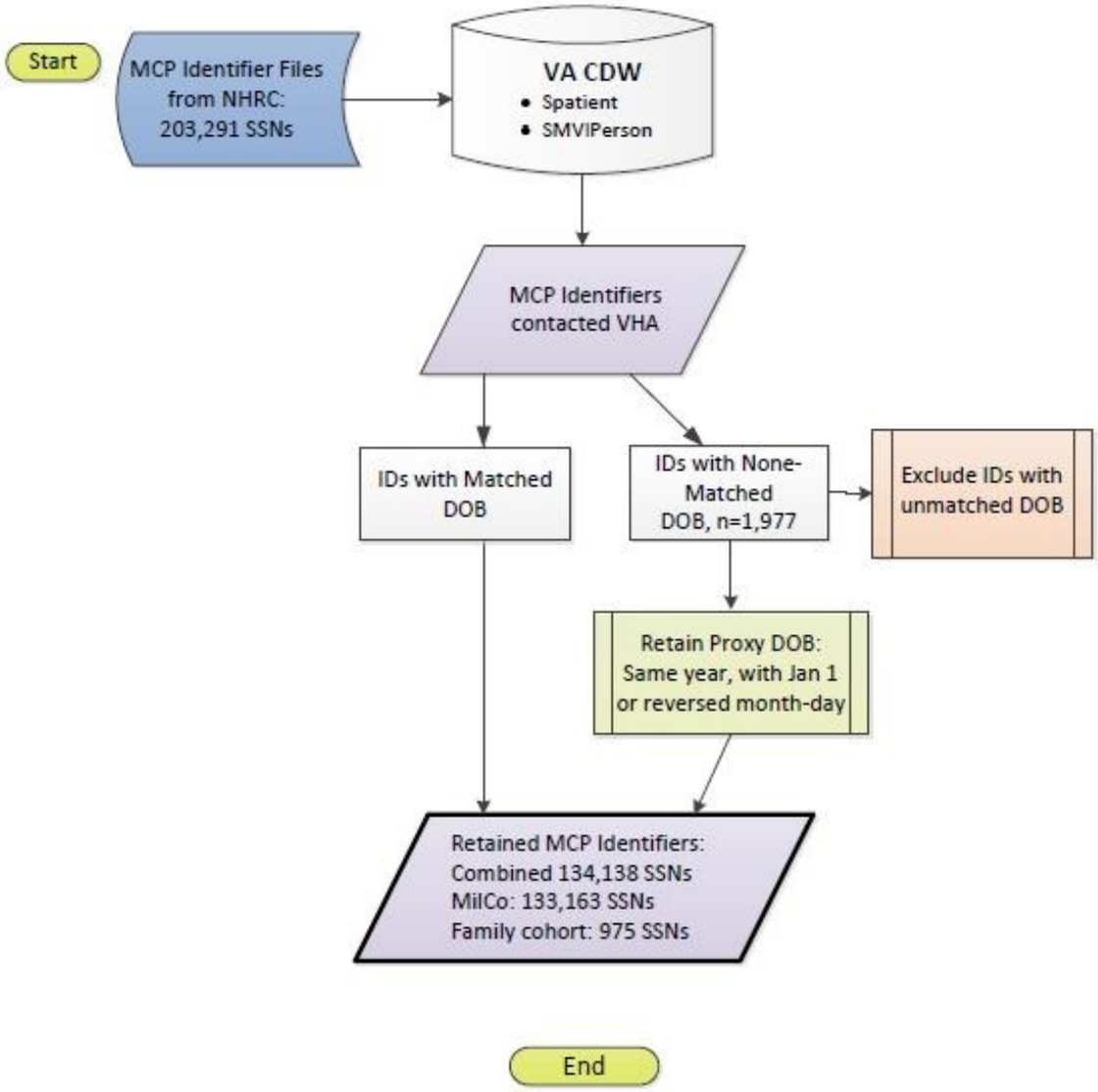

Figure 2. Flow diagram of matching Millennium Cohort and Family Cohort participants to VHA records (excludes MCS participants that do not have a VHA record). Updated 6/24/2020. Note that Veteran status and VA use among MilCo participants is constantly changing and these numbers may not be the most current.

VA data sent from VA to NHRC

VA data is transferred to NRHC using a secure data exchange method (see 6.2. Data flow & security). All VA data shared with NHRC is stripped of all identifiers except a unique study ID. All matching occurs at VA.

Millennium Cohort Program data sent from NHRC to VA

Millennium Cohort and Family Study participant identifiers are sent to VA from NHRC. Participant identifiers are used only to extract data relating to MCS participants and match these participants to a unique study ID and VA records. All matching occurs at VA. Figure 2 above outlines the steps for matching Millennium Cohort and Family Study participants to records within the VHA. This flowchart excludes MCP participants who have not had any contact with a VHA facility or VA community care or have not applied for/received VA benefits. Some common barriers to matching Millennium Cohort and Family Study participants to VHA records include missing data, erroneously coded date of birth (i.e. reversed month and day), or non-matched gender. Participants with non-matched genders are flagged, but not excluded so as to not inadvertently systematically exclude Veterans who may identify as transgender.

The remaining 133,163 Millennium Cohort Study and 975 Family Study participants are defined as Millennium Cohort Program participants who utilize the VHA and exist in the VistA records and are matched on unique study ID, name, and date of birth. These participants include 151 “duals,” MCP-VHA users who are co-enrolled in both Millennium Cohort Study and Family Study. The population numbers for MCP-VHA co-enrollees are not fixed and are expected to rise over time as more Millennium Cohort participants separate from the military and develop VA records. In a March 2021 data pull that compared current MCP-VHA users to those in March 2020, CSP #505 observed that VHA captured over 2,600 new VHA enrollees from Millennium Cohort Program participants over the course of one year, estimated to roughly 50 new MCS-VHA co-enrollees per week.

8. Investigator data access

VA investigators who are interested in using MCS-VA merged data are encouraged to contact the CSP #505 team and complete a Research Resource Request (RRR) form to outline their initial project objectives, exposures, outcomes, desired data, and funding information. The following section displays checklists to assist investigators in obtaining

project approval and access to MCS-VA merged data. An investigator *not seeking external funding* refers to an investigator who is affiliated with and fully funded by the VA or another organization and does not require external funding or grants for project support. An investigator *who is seeking external funding* refers to an investigator who will require partial or all funds from an external source. Please note that any project that requires substantial data access and retrieval will require funding allocation to CSP #505 data analysts and project administrators. For more information, please contact the CSP #505 Project Manager.

8.1. Data access checklists

Millennium Cohort Program-VA Merged Data Access Checklist

Investigator Not Seeking External Funding

- [ ] Demonstrate interest in utilizing MCS-VA merged data by contacting CSP #505 Project Manager with a description of proposed research and receive official CSP #505 Welcome Letter.
- [ ] Submit a Research Resource Request (RRR) form to Amber.Banerjee@va.gov using the instructions provided on the form. The RRR form is attached to your Welcome Letter.
- [ ] Receive an official RRR review notification (Approve, Revise & Resubmit, Denied).
- [ ] All approved RRR forms will immediately receive Investigator Packet. Complete and submit a Secondary Data Analyses Proposal using the template and instructions included in the Investigator Packet. The Secondary Data Analyses Proposal will be reviewed by the CSP #505 Executive Committee and the NHRC Scientific Review Committee (SRC).
- [ ] Receive notification of CSP #505 Executive Committee (EC) decision from CSP #505 Project Manager. If the proposal is approved, CSP #505 will forward your proposal to the NHRC SRC.
- [ ] Receive notification from CSP #505 Project Manager of NHRC SRC review decision (Approved, Revise & Resubmit, Denied).
- [ ] Obtain IRB approval from *your* institution. Please forward your IRB approval letter to CSP #505 Project Manager.
- [ ] Complete and submit the VA data intake form.

- [ ] Communicate with your Millennium Cohort Program point of contact to discuss next steps regarding data access and analyses.

Millennium Cohort Program-VA Merged Data Access Checklist

Investigator Seeking External Funding

- [ ] Demonstrate interest in utilizing MCS-VA merged data by contacting CSP #505 Project Manager with a description of proposed research and receive official CSP #505 Welcome Letter.

- [ ] Submit a Research Resource Request (RRR) form to Amber.Banerjee@va.gov and attach a copy of the Letter of Intent (LOI), pre-proposal, or 1-page project description that you are submitting to your funding agency. Drafts are acceptable. Investigators should send CSP #505 the RRR + LOI *before* submitting to a funding agency.

- [ ] Receive an official RRR review notification (Approve, Revise & Resubmit, Denied). If approved, you will be invited to submit a Secondary Data Analyses Proposal for review by the CSP #505 Executive Committee and the Naval Health Research Center (NHRC) Scientific Review Committee (SRC). This proposal may be submitted to CSP #505 *after* the completion of your funding application. CSP #505 will also provide you with a Letter of Support to demonstrate your partnership with the Millennium Cohort Program for your funding application.

- [ ] Receive Investigator Packet. After you have been notified of funding approval, complete and submit a Secondary Data Analyses Proposal using the template and instructions included in the Investigator Packet.

- [ ] Receive notification of CSP #505 Executive Committee (EC) review decision from CSP #505 Project Manager. If approved, CSP #505 will forward your proposal to the NHRC Scientific Review Committee (SRC).

- [ ] Receive notification from CSP #505 Project Manager when NHRC SRC has reviewed your application and determined your status (Approved, Revise/Resubmit, Denied).

- [ ]

Obtain IRB approval or exemption from *your* institution. Please forward your IRB approval letter to CSP #505 Project Manager.

- [ ] Complete and submit the VA data intake form.

- [ ] Communicate with your Millennium Cohort Program point of contact to discuss next steps regarding data access and analyses.

8.2. VINCI: Towards VA investigator workspace

It is the goal of CSP #505 to house Millennium Cohort Study survey data at VA within the VINCI environment to be available for use by VA investigators with project approval. Once we achieve full bilateral data sharing between VA and DoD's NHRC and DMDC, we will place the data on VINCI. Millennium Cohort Study survey data was transferred onsite to VA in August 2021. This Data User Guide will be adapted to include additional instructions for accessing and using Millennium Cohort Study survey data within VINCI.

9. Interpreting MCP-VA merged data

9.1. Potential sources of bias in MCP-VA merged data

Non-response bias/loss to follow-up

It is important to investigate non-response and characteristics of non-responders in a population-based sample to ensure representativeness, understand potential bias, and develop targeted strategies to improve responsiveness (22).

Panel 1 of the Millennium Cohort Study has historically demonstrated the highest rate of survey participation compared to other Panels. Of the 256,400 military personnel that were invited to participate in Panel 1 of the Millennium Cohort study in 2001, 76,861 eligible individuals completed the baseline questionnaire and were alive at the first follow-up timepoint (30, 31). Between June 2004 and February 2006, 54,960 (71.6%) individuals completed the first follow-up survey (30). In a 2010 paper investigating non-response from baseline to first follow-up among Panel 1 of the Millennium Cohort Study, the characteristics associated with a higher probability of response included being female, older age, higher educational achievement, being married, having officer rank,

currently serving on active duty, and self-reporting military exposures prior to 2001 (including witnessing death or chemical or biological warfare) (30). Characteristics at baseline including smoking, a history of unhealthy alcohol use, major depressive disorder, or separation from the military for reasons other than retirement were all associated with a lower probability of response to the first follow-up survey. When investigating the effect of non-response on the strength of associations between predictors and outcomes such as PTSD and eating disorder incidence, the authors found that re-analysis using inverse probability weighting for non-response did not change the strength of associations significantly compared to the complete case analysis (30).

One potential source of response bias is if participants are more likely to accept an invitation to join a study because they have better or poorer health compared to non-responders (32). A 2008 report investigated potential differences in healthcare utilization prior to enrollment among responders and non-responders of the Millennium Cohort Study (32). Using DoD inpatient and outpatient records to assess healthcare utilization within the 12 months preceding study invitation, the study found no significant differences in healthcare utilization prior to enrollment between responders and non-responders of MCS (31).

The referral sampling method of enrollment for Panel 1 of the Family Study may have resulted in limited enrollment or referral bias (22). Panel 4 Millennium Cohort participants were given only one opportunity to refer their spouse on the triennial survey, potentially resulting in a more limited enrollment pool compared to a sampling strategy that allowed service members the opportunity to refer their spouse on multiple occasions (22). There is a possibility that Millennium Cohort participants who referred their spouses may be significantly different from participants who did not refer their spouse. To mitigate this potential bias, the Family Study team expanded sampling to include both referral by MCS participant and direct invitation to eligible spouses (24).A 2017 report investigating non-response bias in the Family Study identified that the following characteristics *among Millennium Cohort participants* were correlated with higher spousal response rate: male gender, older age, white non-Hispanic ethnic identity, higher educational attainment, an income between $25,000-$74,999, dependent children, serving in the Reserves/National Guard compared to active duty, and deployment with combat experience (22). Unhealthy behaviors among Millennium Cohort service members such as smoking, higher caffeine intake, poorer sleep patterns, and higher sedentary behavior were associated with lower spousal response rate (22). Better mental health among Millennium Cohort service members, including the absence of major depressive disorder and fewer PTSD symptoms, was correlated with higher spousal response rate in the Family Study (22). Referring a spouse to the Family Study by the MCS participant was the greatest predictor of Family Study response rate compared to other Millennium Cohort participant characteristics (22).

The Millennium Cohort Study and Millennium Cohort Family Study use modified Dillman recruitment methods to maximize response rates including an introductory postcard sent prior to survey mailing, reminder mailings (postal and electronic), holiday mailings (postal and electronic), and pre- and post-incentives to maintain engagement. These contacts also include a reminder to update any address changes. Historically, the majority of Millennium Cohort Study surveys are submitted online. On average, 76% of Millennium Cohort participants in Panels 1-4 submitted web surveys (33). The Family Study survey was originally created as a web-only survey but expanded to include a paper version with the objective of reaching a wider audience and decreasing potential response bias. Both the Millennium Cohort Study and Family Study plan to transition to online-only surveys for future cycles.

External validity: Generalizability of MCS-VA data to other populations

Millennium Cohort Study participants are selected from active rosters across all service branches and service components of the US military (1). Oversampling of selected sub-groups has been a feature of the study design but differs by recruitment Panel as shown in Figure 3 on page 13. The Millennium Cohort Program has oversampled historically under-represented subgroups in military research (women, Marines, prior deployers, Reservists and National Guard), and married service members for enrollment of spouses in the Family Study (25,22). Please refer to section 5. Study population for more information.

Due to self-selected participation in the Millennium Cohort Study, there may be some limitations to generalizability outside of the Cohort (31,32). However, the Millennium Cohort baseline data provide evidence that this Cohort is a representative sample of the larger military population in terms of health status, vaccination history, and health care utilization at baseline (31,32).

10. Disseminating research & results

10.1. NHRC Command Review

All research completed using Millennium Cohort Program data must be submitted for review by the commanding officer at the NHRC prior to submission for publication or presentation. Command review can take up to 6 months for full review, feedback, and approval for dissemination but is often shorter for manuscripts and much shorter for presentations. For more information on submitting work based in the Millennium Cohort Program for NHRC Command Review, please contact the CSP #505 Project Manager.

10.2. Publications & presentations using merged MCP-VA data

Please notify CSP #505 of any accepted publications, presentations, or other media events using merged MCP-VA data. All such publications and presentations must be reported to ORD. Newsworthy or controversial reports of results should be shared with leadership prior to publication and media coverage. Please contact the CSP #505 Project Manager to report any acceptances at conferences or other media events. Please note that all presentations and media appearances displaying Millennium Cohort Study information must be submitted for NHRC Command Review and approval (see 10.1. NHRC Command Review).

11. References

1. Gray GC, Chesbrough KB, Ryan MAK, Amoroso P, Boyko EJ, Gackstetter GD, Hooper TI, Riddle JR, and the Millennium Cohort Study Group. "The Millennium Cohort Study: A 21-Year Prospective Cohort Study of 140,000 Military Personnel." *Military Medicine*. 2002; 167(6):483-488.
2. US Department of Defense. The Millennium Cohort Family Study. "Frequently Asked Questions." 2015. Retrieved from: https://www.familycohort.org/faq.
3. Department of Veterans Affairs. VHA Data Portal. "Corporate Data Warehouse (CDW)." Last updated June 4, 2019. Retrieved from: http://vaww.vhadataportal.med.va.gov/DataSources/CDW.aspx.
4. VA Center for Integrated Healthcare (CIH). The Alcohol Use Identification Test (AUDIT-C) Ver3.0. July 2013. Retrieved from https://www.mirecc.va.gov/cih-visn2/Documents/Provider_Education_Handouts/AUDIT-C_Version_3.pdf.
5. Department of Veterans Affairs. VHA Data Portal. "ADUSH Enrollment Files." Last updated May 22, 2020. Retrieved from: http://vaww.vhadataportal.med.va.gov/DataSources/ADUSHEnrollmentFile.aspx.
6. US Department of Veterans Affairs Information Resource Center. "VIReC Research User Guide: VHA Assistant Deputy Under Secretary for Health (ADUSH) Enrollment Files." 2nd Edition. Hines IL: US Department of Veterans Affairs, Health Services Research & Development Service, Information Resource Center; Sept. 2013. http://vaww.virec.research.va.gov/RUGs/ADUSH/RUG-ADUSH-EF-FY99-12-RA.pdf
7. Department of Veterans Affairs. VHA Data Portal. "Vital Status File." Last updated June 4, 2019. Retrieved from: http://vaww.vhadataportal.med.va.gov/DataSources/VitalStatusFile.aspx.
8. Maynard, Charles. VIReC Database & Methods Cyberseminar Series. "Ascertaining Veterans' Vital Status: VA Data Sources for Mortality Ascertainment and Cause of Death." March 3, 2017. Retrieved from: https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/1242-notes.pdf.

9. US Department of Veterans Affairs. Community Care. “Patient-Centered Community Care – Information for Providers.” Last updated April 8, 2020. Retrieved from: https://www.va.gov/COMMUNITYCARE/providers/info_PC3.asp.
10. US Department of Veterans Affairs Office of Community Care. “Veteran Community Care: General Information.” Created September 9, 2019. Retrieved from: https://www.va.gov/COMMUNITYCARE/docs/pubfiles/factsheets/VHA-FS_MISSION-Act.pdf.
11. US Department of Veterans Affairs Office of Public Affairs. “Veterans Access, Choice, and Accountability Act of 2014 (‘Choice Act”)”. Retrieved from: https://www.va.gov/opa/choiceact/documents/choice-act-summary.pdf.
12. US Department of Veterans Affairs Health Economics Resource Center (HERC) “VA Community Care.” Last updated April 22, 2020. Retrieved from: https://vaww.herc.research.va.gov/include/page.asp?id=va-cc.
13. US Department of Veterans Affairs Health Economics Resource Center (HERC). ”Community Care Data – Program Integrity Tool (PIT).” Retrieved from: https://www.herc.research.va.gov/include/page.asp?id=choice-pit.
14. US Department of Veterans Affairs Information Resource Center. The Researcher’s Notebook. “Community Care Data: Comparison of Claims in the CDW FBCS Domain and PIT Domain.” June 2020. 20(20):1-26.
15. US Department of Veterans Affairs Office of the Inspector General. “Non-VA Emergency Care Claims Inappropriately Denied and Rejected.” Report #18-00469. August 6, 2019.
16. US Department of Veterans Affairs Office of the Inspector General. “Veterans Benefits Association: Audit of the Veterans Service Network.” Report #: 09-03850-99. February 18, 2011.
17. Department of Veterans Affairs. VHA Data Portal. “VETSNET File." Last updated September 22, 2020. Retrieved from: http://vaww.vhadataportal.med.va.gov/DataSources/VETSNET.aspx.
18. US Department of Veterans Affairs. Veterans Benefits Administration. “Compensation.” Last updated November 26, 2019. Retrieved from: https://www.benefits.va.gov/compensation/rates-index.asp#combinedRatingsTable1.
19. US Department of Veterans Affairs. Veterans Benefits Administration. “Service-Connected Matrix.” Last updated November 15, 2019. Retrieved from: https://benefits.va.gov/benefits/derivative_sc.asp.
20. Ryan MAK, Smith TC, Smith B, Amoroso P, Boyko EJ, Gray GC, Gackstetter GD, Riddle JR, Wells TS, Gumbs G, Corbeil TE, Hooper TI. “Millennium Cohort: enrollment begins a 21-year contribution to understanding the impact of military service.” *Journal of Clinical Epidemiology.* 2007; 60:181-191.
21. Smith TC, Smith B, Jacobson IG, Corbeil TE, Ryan MAK for the Millennium Cohort Team. “Reliability of Standard Health Assessment Instruments in a Large, Population-Based Cohort Study.” *Annals of Epidemiology*. 2007; 17:525-532.
22. Corry NH, Williams CS, Battaglia M, McMaster HS and Stander VA. “Assessing and adjusting for non-response in the Millennium Cohort Family Study.” *BMC Medical Research Methodology*. 2017; 17(16):1-17.
23. Millennium Cohort Study. “Voluntary Consent to Participate in a Research Study.” https://www.millenniumcohort.org/files/milco/2019/consentform_and_privacyact.pdf.

24. Crum-Cianflone NF, Fairbank JA, Marmar CR, and Schlenger W. “The Millennium Cohort Family Study: a prospective evaluation of the health and well-being of military service members and their families.” *International Journal of Methods in Psychiatric Research*. 2014; 23(3):320-330.
25. Rivera AC, Powell TM, Boyko EK, Lee RU, Faix DJ, Luxton DD, and Rull RR, for the Millennium Cohort Study Team. “New-Onset Asthma and Combat Deployment: Findings From the Millennium Cohort Study.” *American Journal of Epidemiology*. 2018;187(10):2136-2144.
26. US Department of Defense. The Millennium Cohort Study. “Study Timeline.” 2000. Retrieved from: https://www.millenniumcohort.org/about.
27. Boyko EJ, Rull RP, Schneiderman AI. Department of Veterans Affairs Health Services Research & Development (HSR&D) Cyberseminar series. “Overview of the Millennium Cohort Study (MILCO): Opportunities for VA Investigators.” Seminar date: October 29, 2020. Retrieved from: https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=3878.
28. Porter B, Rodriguez LM, Woodall KA, Pflieger JC, Stander VA. “Alcohol misuse and separation from military service: A dyadic perspective.” *Addictive Behaviors*. 2020; 110:1-8.
29. Department of Defense Millennium Cohort Study. “Supporting Statement – Part B: Collections of Information Employing Statistical Methods.” Retrieved from https://www.reginfo.gov/public/do/DownloadDocument?objectID=82729703.
30. Littman AJ, Boyko EJ, Jacobson IG, Horton J, Gackstetter GD, Smith B, Hooper T, Wells TS, Amoroso PJ, Smith TC, the Millennium Cohort Study. “Assessing nonresponse bias at follow-up in a large prospective cohort of relatively young and mobile military service members.” *BMC Medical Research Methodology*. 2010; 10(99):1-11.
31. Wells TS, Jacobson IG, Smith TC, Spooner CH, Smith B, Reed RJ, Ryan MAK for the Millennium Cohort Study Team. “Prior health care utilization as a potential determinant of enrollment in a 21-year prospective study, the Millennium Cohort Study.” *European Journal of Epidemiology*. 2008;23:73-87.
32. Smith TC, Zamorski M, Smith B, Riddle JR, LeardMann CA, Wells TS, Engel CC, Hoge CW, Adkins J, Blaze D, for the Millennium Cohort Study Team. “The physical and mental status of a large military cohort: baseline functional health status of the Millennium Cohort.” *BMC Public Health*. 2007;7(340):1-13.
33. Faix, Dennis CDR for the Millennium Cohort Study Team. “Millennium Cohort Study.” Presented at the 2017 Millennium Cohort Strategic Board Meeting. April 25-26, 2017.
34. Wagner T, Stefos T, Morgan E, Cashy J, Shen ML, Gehlert E, Asch S, Almenoff P. Risk Adjustment: Guide to the V21 and Nosos Risk Score Programs. Technical Report 30. Menlo Park, CA. VA Palo Alto, Health Economics Resource Center. February 2016.
35. US Department of Veterans Affairs. Office of Clinical Systems Development and Evaluation. Care Assessment Needs (CAN) Primer. September 30, 2019.

Appendix A – Data sources for select variables used in MCP-VA research

Variable	CDW	ADUSH	VSF	PIT/Fee	VETSNET	MCS survey self-report	DMDC personnel files
Demographics							
Age	☑	☑	☑	☑	☑	☒	☑
Date of death	☑	☑	☑	☑	☑	☒	☑
Gender/sex	☑	☑	☑	☑	☑	☒	☑
Race/ethnicity	☑	☑	☑	☒	☒	☒	☑
Marital status	☑	☒	☒	☒	☑	☑	☑
Education	☒	☒	☒	☒	☒	☑	☑
Household income	☒	☒	☒	☒	☒	☑	☒
Height/weight	☑	☒	☒	☒	☒	☑	☒
Sexual orientation	☒	☒	☒	☒	☒	☑	☒
Military characteristics							
Period of service	☑	☒	☒	☒	☑	☒	☑
Service branch	☑	☒	☒	☒	☑	☒	☑
Occupation	☒	☒	☒	☒	☒	☑	☑
Pay grade	☒	☒	☒	☒	☑	☒	☑
Deployment experience	☒	☒	☒	☒	☒	☑	☑
Combat exposure	☒	☒	☒	☒	☒	☑	☒
VHA healthcare data							
Visit date	☑	☒	☒	☑	☒	☒	☒
Admission date	☑	☒	☒	☑	☒	☒	☒
Diagnosis ICD code	☑	☒	☒	☑	☒	☒	☒
Procedure CPT code	☑	☒	☒	☑	☒	☒	☒
Procedure ICD code	☑	☒	☒	☑	☒	☒	☒
Pharmacy	☑	☒	☒	☑	☒	☒	☒
Appointment	☑	☒	☒	☒	☒	☒	☒
Lab test	☑	☒	☒	☒	☒	☒	☒
Vital signs	☑	☒	☒	☒	☒	☒	☒
VHA enrollment							
VHA enrollment date	☑	☑	☒	☒	☒	☒	☒

Service-connection (SC) %	☑	☑	☒	☒	☑	☒	☒
VHA eligibility	☑	☑	☒	☒	☒	☒	☒
Priority group	☑	☑	☒	☒	☒	☒	☒
VBA benefits							
SC disability condition code	☒	☒	☒	☒	☑	☒	☒
Compensation amount	☑	☒	☒	☒	☑	☒	☒
Health conditions							
Anxiety	☑	☒	☒	☑	☒	☑	☒
Major depressive disorder	☑	☒	☒	☑	☒	☑	☒
PTSD	☑	☒	☒	☑	☒	☑	☒
Sleep disorder	☑	☒	☒	☑	☒	☑	☒
Physical functioning	☑	☒	☒	☒	☒	☑	☒
Condition diagnoses†	☑	☒	☒	☑	☒	☑	☒
Health risk factors							
Smoking status	☑	☒	☒	☒	☒	☑	☒
Alcohol use	☑	☒	☒	☒	☒	☑	☒
Pregnancy/miscarriage	☑	☒	☒	☑	☒	☑	☒
Homelessness status	☑	☒	☒	☒	☒	☑	☒
Adverse childhood event	☒	☒	☒	☒	☒	☑	☒

† For list of diagnoses available for endorsement on the MCS survey, please see section 4.3.2 Millennium Cohort survey topics.

Data source populations:

CDW: All Veterans who have at least one VHA encounter including community care.

ADUSH: All Veterans who have received any VA care regardless of VHA enrollment.

VSF: All Veterans who have received any VA care, are enrolled in VHA, or have received VBA benefits; includes variables from both Master and Mini files.

PIT/fee: All Veterans who have received any eligible community care paid for by the VA.

VETSNET: All Veterans who have applied for or received VBA benefits.

MCS self-report: All Millennium Cohort participants who have a non-NULL value for the relevant survey question.

DMDC personnel files: All individuals who have served in the US military.

Appendix B – Recommended Millennium Cohort Program publications

1. Crum-Cianflone NF, Fairbank JA, Marmar CR, Schlenger W. “The Millennium Cohort Family Study: a prospective evaluation of the health and well-being of military service members and their families.” *International Journal of Methods in Psychiatric Research*. 2014; 23(3):320-330.

2. Gray, GC, Chesbrough KB, Ryan MAK, Amoroso P, Boyko EJ, Gackstetter GD, Hooper TI, Riddle JR, and the Millennium Cohort Study Group. “The Millennium Cohort Study: A 21-Year Prospective Cohort Study of 140,000 Military Personnel.” *Military Medicine*. 2002; 167(6):483-488.

3. Jacobson IG, Ryan MAK, Hooper TI, Smith TC, Amoroso PJ, Boyko EJ, Gacksetter GD, Wells TS, Bell NS. “Alcohol Use and Alcohol-Related Problems Before and After Military Combat Deployment.” *JAMA*. 2008;300(6):663-675.

4. LeardMann CA, Powell TM, Smith TC, Bell MR, Smith B, Boyko EJ, Hooper TI, Gacksetter GD, Ghamsary M, Hoge CW. “Risk factors associated with suicide in current and former US military personnel.” *JAMA*. 2013;310(5):496-506.

5. Riddle JR, Smith TC, Smith B, Corbeil TE, Engel CC, Wells TS, Hoge CW, Adkins J, Zamorski M, Blazer D, for the Millennium Cohort Study Team. “Millennium Cohort: The 2001-2003 baseline prevalence of mental disorders in the U.S. military.” *Journal of Clinical Epidemiology*. 2007;60:192-201.

6. Ryan MAK, Smith TC, Smith B, Amoroso P, Boyko EJ, Gray GC, Gackstetter GD, Riddle JR, Wells TS, Gumbs G, Corbeil TE, Hooper TI. “Millennium Cohort: enrollment begins a 21-year contribution to understanding the impact of military service.” *Journal of Clinical Epidemiology*. 2007; 60:181-191.

7. Smith TC, Jacobson IG, Hooper TI, LeardMann CA, Boyko EJ, Smith B, Gackstetter GD, Wells TS, Amoroso PJ, Gray GC, Riddle JR, Ryan MAK and the Millennium Cohort Study Team. “Health impact of US military service in a large population-based military cohort: findings of the Millennium Cohort Study, 2001-2008.” *BMC Public Health*. 2011;11:69.

8. Smith TC, Smith B, Corbeil TE, Riddle JR, Ryan MAK for the Millennium Cohort Study Team. “Self-Reported Mental Health Among US Military Personnel Prior and Subsequent to the Terrorist Attacks of September 11, 2001.” *Journal of Occupational and Environmental Medicine*. 2004; 46(8):775-782.

9. Smith TC, Smith B, Jacobson IG, Corbeil TE, Ryan MAK for the Millennium Cohort Team. “Reliability of Standard Health Assessment Instruments in a Large, Population-Based Cohort Study.” *Annals of Epidemiology*. 2007; 17:525-532.

10. Smith TC, Zamorski M, Smith B, Riddle JR, LeardMann CA, Wells TS, Engel CC, Hoge CW, Adkins J, Blaze D, for the Millennium Cohort Study Team. “The physical and mental status of a large military cohort: baseline functional health status of the Millennium Cohort.” *BMC Public Health*. 2007;7(340):1-13.

For a complete list of Millennium Cohort publications please visit the Millennium Cohort Study website: https://www.millenniumcohort.org/research/publications.