Million Veteran Program (MVP)

Presentation for: Research Advisory Committee on Gulf War Veterans’ Illnesses
Presented by: Sumitra Muralidhar, Director, MVP
Date of briefing: January 27, 2022
Presentation Overview

- MVP Overview
- Cohort Description
- Data from MVP Biospecimens
- Accessing MVP Data
- MVP Research
  - Genomics of Gulf War Illness (CSP2006/MVP029)
- Looking Ahead
MVP Overview

• MVP is a national VA research program, launched in 2011, designed to advance precision health care by learning how genes, lifestyle, and military experiences and exposures affect health and illness
  – Establish a comprehensive, diverse cohort of at least one million Veterans
  – Provide broad access to the data for scientific discovery
  – Establish pipelines to translate discoveries to the clinic to improve the health of Veterans

• MVP is one of the world’s largest research programs of its kind based in a healthcare system with over 864,000 Veterans enrolled (as of Dec. 2021)
MVP Enrollment

Enrollment Sites

- >60 Main Enrollment Sites
- >17 Million Total Mailings
- >6 Million Individuals Contacted

MVP On-Line

The future of medicine is in your genes

The Million Veteran Program (MVP) is a national research program to learn how genes, lifestyle, and military exposures affect health and illness. Since launching in 2011, over 850,000 Veterans have joined MVP, making it one of the world’s largest programs on genetics and health.

Joining MVP is easy

We make it easy to become an MVP partner. To join online, Veterans are asked to:

- Securely sign in or create an account
- Complete the consent process and allow access to health records
- Conveniently schedule a visit to provide a blood sample
- Complete surveys about health, lifestyle, and military exposure

Join today
The enrollment process:

- **Provide consent**, including allowing **access to health records** *(online or in-person)*

- **Give blood specimen** *(in-person or via a new at-home self-administered blood collection kit for on-line enrollees)*

- **Complete surveys**, including a baseline and lifestyle survey *(online or paper-based)*
• 91% male, 9% female
• 8.2% Hispanic
• 17.3% Black (MVP has the largest cohort of people of African ancestry of any research program in the world)
## Cohort Summary (cont.)

### Service Era

- September 2001 or Later: 2.40%
- August 1990 to August 2001 (Including Gulf War): 10.25%
- May 1975 to July 1990: 12.24%
- August 1964 to April 1975 (Vietnam Era): 37.83%
- February 1955 to July 1964: 5.46%
- July 1950 to January 1955 (Korean War): 4.35%
- January 1947 to June 1950: 0.18%
- December 1941 to December 1946 (WWII): 1.98%
- November 1941 or Earlier: 0.04%
- Multiple Service Eras: 22.05%
- Missing Service Eras: 3.22%

### Age

- Average age is 67

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-29 years</td>
<td>0.64%</td>
</tr>
<tr>
<td>30-39 years</td>
<td>5.88%</td>
</tr>
<tr>
<td>40-49 years</td>
<td>6.81%</td>
</tr>
<tr>
<td>50-59 years</td>
<td>12.24%</td>
</tr>
<tr>
<td>60-69 years</td>
<td>21.89%</td>
</tr>
<tr>
<td>70-79 years</td>
<td>35.82%</td>
</tr>
<tr>
<td>80-89 years</td>
<td>11.68%</td>
</tr>
<tr>
<td>90-99 years</td>
<td>4.70%</td>
</tr>
<tr>
<td>100+ years</td>
<td>0.33%</td>
</tr>
</tbody>
</table>

(N=850,000)
Women’s Campaign

- **WHAT:** MVP’s 1st digital campaign, designed to enroll women Veterans

- **WHY:** Women are 9% of all Veterans in MVP (~80,000) and reflect the 9.1% in VHA
  - Historically, women have been underrepresented in biomedical research

- **HOW:** Multipronged digital communication approach

- **RESULTS:** Between March 2021 and September 2021, 4,333 women Veterans enrolled; 20% of total enrollees during that period

- **CONCLUSION:** Focused digital campaigns can increase underrepresented populations in MVP
Top Ten Diseases/Conditions

- Hypertension: 73.4%
- Skin condition: 51.6%
- Cataracts: 49.8%
- Depression: 45.6%
- Acid Reflux/GERD: 43.7%
- Ostearthritis: 41.7%
- Hearing loss: 41.6%
- High Cholesterol: 40.9%
- Diabetes: 38.0%
- Sleep apnea: 36.4%
Launch a mental health precision medicine initiative to identify and validate brain and mental health biomarkers in coordination with MVP *(Part of VHA’s response to Hannon Act Section 305)*

- Additional sites will be added for enrollment
- Recruitment in VA Mental Health and Substance Use Disorder clinics across VHA
- New MIND survey developed
- Goal of 50,000 enrollees in 5 years
VA Phenomics Library

**CIPHER**

Centralized Interactive Phenomics Resource

(> 2000 Phenotypes)

Overview

The VA Phenomics Library (VAPheLib) is a shared knowledgebase of VA EHR-based data on behalf of the Office of Research and Development (ORD). This effort is part of an initiative to enable researchers to share their work. VAPheLib is led by a collaboration between the

Mission

To provide an encyclopedia of VHA EMR based phenotyping through integration of research community

Objectives

1. To provide a knowledgebase framework to collect, store/archive and share phenotypic data
2. To expedite VA science by enabling phenotype reusability and scalability across VA
3. To build a platform to encourage and enhance collaboration and communication
MVP Biospecimen Data Overview

**GOAL:** Generate the maximum amount of data from biospecimens to enhance scientific discovery

- Baseline genetic data profile (genotype) generated for all participants
  - Data from 650,000 samples are currently provided to approved researchers
- Genetic data for specific ethnic groups (Blacks, Hispanics and Asians) using a customized analytic tool currently underway for ~200K participants
- Whole genome sequences have been generated on ~140,000 samples
  - Processing underway
- Other data such as proteomics and metabolomics are being piloted
### MVP Scientific Computing Infrastructure

<table>
<thead>
<tr>
<th></th>
<th>VA Servers</th>
<th>Department of Energy (DOE) Servers</th>
<th>U of Chicago Cloud (VA Data Commons)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong># of projects</strong></td>
<td>30+</td>
<td>4 (new projects under review)</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Accessibility</strong></td>
<td>VA only</td>
<td>VA &amp; DOE only</td>
<td>VA and non-VA</td>
</tr>
<tr>
<td><strong>Currently in use</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No – in pilot testing</td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>Coded (No direct identifiers)</td>
<td>Coded (No direct identifiers)</td>
<td>Deidentified</td>
</tr>
</tbody>
</table>
MVP Summary Data Access in NIH’s dbGaP

15 publications
96 analyses

7 phenotype categories
166 authorized requests for access

Search “VA Million Veteran Program”
• Computing infrastructure within the VA meets VA IT Privacy and Security requirements; DOE and the University of Chicago VA Data Commons have an approved VA authority to operate (ATO)
• Biospecimen (blood sample) and data (surveys) collected are labeled using a code instead of identifiable information
• Crosswalk to identity of participant is held by few authorized staff
• Researchers access only coded data (no direct identifiers such as SSN, name, date of birth, street address)
• Researchers sign rules of behavior and analyze data in a central, secure computing system
• No data leaves the system; only summary results can be taken out
<table>
<thead>
<tr>
<th>Current Research Areas Using MVP data</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD</td>
</tr>
<tr>
<td>Multi-substance use</td>
</tr>
<tr>
<td>Schizophrenia and bipolar disorder</td>
</tr>
<tr>
<td>Cardiovascular disease and cardiometabolic disorders</td>
</tr>
<tr>
<td>Diabetes and complications</td>
</tr>
<tr>
<td>Age related macular degeneration</td>
</tr>
<tr>
<td>Gulf War Illness</td>
</tr>
<tr>
<td>Breast cancer</td>
</tr>
<tr>
<td>Prostate cancer</td>
</tr>
<tr>
<td>Lung cancer</td>
</tr>
</tbody>
</table>

**COVID-19 was added in 2020**

- **30+ Research projects**
- **75+ Publications**
- **500+ Researchers and affiliates**
- **4 patents filed**
COVID-19 Research in MVP

- **New data collection**
  - COVID-19 Survey to 700,000+ MVP participants by mail and on-line
    - ~ 250,000 completed surveys received
    - Assessing experience with COVID-19, including preventive measures, symptoms, complications, hospitalization, treatment received, impact on routine care

- **Science projects**
  - Genetic basis of infection, severity, complications death, *vaccine responsiveness*, *post-COVID symptoms*
  - Response to treatments
  - Disease mechanisms

Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19
Partnership with the Department of Energy

GOAL: Use DOE’s high-performance computing capabilities and expertise in Artificial Intelligence (AI) and VA’s big data (EHR and MVP) for compute-intensive data science projects

• VA clinical data (24M Veterans) and MVP genetic and survey data at the Oak Ridge National Lab

• Research projects focused on developing AI-driven models for risk prediction
  – Suicide (jointly with VHA clinical operations to improve the REACHVET model)
  – Metastatic prostate cancer
  – Cardiovascular disease

• Largest genetic association study of ~30 million genetic variants with 2000 health and disease characteristics on the SUMMIT supercomputer
CSP 2006/MVP029: Genomics of Gulf War Illness

• Purpose – To Understand genetic contributions to GWI by Resurveying MVP participants in the military during 1990–1991 (GW cohort)
• Characterize GW cohort by deployment, exposures, symptoms, and conditions, including GWI and exclusionary conditions
• Perform Genome Wide Association Study of GWI
• Perform gene x environment analysis
• Perform related genomic, phenomic, and other –omic analyses
• Published Design paper in Brain Sciences Special Issue on 6/25/2021

GWI = Gulf War Illness
MVP = Million Veteran Program
GW = Gulf War

The CSP 2006/MVP029 GW Veteran Cohort

Mailed survey questionnaire to 109,976 eligible individuals

45,169 individuals returned survey questionnaire

41,077 individuals

Missing entire GWI symptom responses (n=4,092)

Missing deployment status (n=299)

40,778 individuals

Individuals deployed (n=14,103)

Individuals did not deploy (n=26,675)
CSP2006/MVP029 Products to date

• GWI Case Definition Standardization
  – Research case definition – not an update to any clinical case definition
    • Goal: identify those who express GWI symptoms to look for exposures, epigenetic markers, or genetic markers that differ between the GWI cases and non-cases
  – Several case definitions are generally accepted
  – Need computer code to label GWI for the 40,000 Veterans in CSP2006
  – Collaboration between CSP585 and CSP2006/MVP029
• 10 preliminary analyses presented as posters or presentations at MVP science meetings in 2019, 2020, and 2021
• 3 presentations at public meetings in 2021 (American Society for Human Genetics, Boston University Seminar)

GWI = Gulf War Illness
CSP = Cooperative Studies Program
MVP = Million Veteran Program
CSP2006/MVP029 Presentations to date

- Titles of Abstracts/Presentations
  - Re-Contact Efforts of Gulf War Era Million Veteran Program Enrollees
  - Genomics of Gulf War Illness in Veterans: Results of a Pilot study
  - Challenges of assessing phenotype of Gulf War Illness using electronic medical records
  - Methods for writing replicable algorithms for IOM-recommended Gulf War Illness indicators using self-reported survey questions
  - GWAS of chronic fatigue syndrome, irritable bowel syndrome, and myalgia/myositis comparing Gulf War veterans with other Million Veteran Program participants
  - Characteristics of deployed vs. non-deployed Gulf War-era Veterans
  - Gulf War Illness in CSP2006/MVP029: An aging cohort
  - Deployment Exposures and Characteristics of 1990-91 Gulf War Veterans (GWV) in the CSP 2006 Cohort
  - Multi-trait genome-wide association study of Gulf War Illness in the Million Veteran Program
  - Modeling the longitudinal changes of ancestry diversity in the MVP cohort: systematic comparison of methods
  - Modeling the longitudinal changes of ancestry diversity in the MVP
  - Genomics of Gulf War Illness in Veterans (CSP 2006) – An Update
CSP 2006/MVP029 Next Steps

• Analyze and publish/disseminate findings
• Expand analyses to address other important questions related to GWI using the CSP 2006 cohort.
  – Healthcare utilization of GWVs with and without GWI and other conditions
  – Apply different case definitions
  – Explore overlap between GWI and other CMI (FM, ME/CFS, IBS)
• Explore the experience of GW era veterans with COVID (*in discussion with the MVP/CSP Program Offices*)
  – Are GWVs with GWI at greater risk of COVID? More severe COVID? Long COVID?
  – Are GWVs with GWI at greater risk of COVID vaccine adverse events?
  – Do GWVs with GWI derive the same benefit from COVID vaccines as those without GWI?
  – Is there an association between GW deployment and COVID risk or COVID vaccination efficacy or risk
• Topics for future proposals
  – GWV experience with neurologic conditions
  – Suicidality in GWVs

CMI = Chronic Multisymptom Illness
FM = Fibromyalgia
ME/CFS = Myalgic Encephalomyelitis/Chronic Fatigue Syndrome
IBS = Irritable Bowel Syndrome
MVP: Looking Ahead

- Reaching the 1 million enrollees milestone
- Increasing the diversity of enrollees (gender, age, race, ethnicity, medical condition) through focused campaigns
- Expanding data access broadly to researchers within and outside the VA
  - Making increasing types and numbers of data available
- Increasing the breadth and scope of scientific projects and scientific products in all relevant areas of health, including mental health and effect of military exposures
- Establishing pipelines to move scientific discoveries into the clinic to improve Veterans’ health and well-being
“As a three-time cancer survivor, I believe genetic research is important because it can help show us in the future how to take care of our patients more unique to who they are as an individual.”

Missina Schallus, Navy Veteran
VA Center for Women Veterans