

### Million Veteran Program

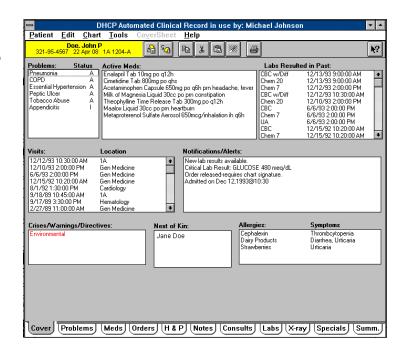
Ronald M. Przygodzki, MD

Acting Director, Biomedical Laboratories Research & Development Associate Director, Genomic Medicine Program US Department of Veterans Affairs Washington, DC



# Population Research in the VA Healthcare System

- VA ideal setting for nested large-scale population research
  - Stable and altruistic Veteran population of 8 million using the system each year
  - Outstanding electronic medical record; fully integrated; data reaching back as far as 20 years; access to CMS and NDI data
  - Intramural research program with diverse clinician and doctoral expertise
  - >70% of investigators are clinicians linking research with Veterans' needs
  - An example of team science and big data within VA infrastructure, involving administrative, technical, ethical and scientific challenges.



## Million Veteran Program (MVP)

- Enroll up to one million users of the VHA into an observational mega-cohort
  - Blood collection for storage in biorepository for future research
  - Collect self-reported health and lifestyle information
  - Access to electronic medical record
  - Ability to recontact participants





### Veteran Consultation Projects – Genomic Medicine Research

- 83% said program should be done
- 71% said they would participate
- 61% also said they would:
  - Allow medical records from non-VA health care to be added to the database
  - Have follow up exams over time
- Most respondents were willing to participate under both opt-in (80%) and opt-out (69%) models
- Stronger preferences for the opt-in approach were expressed among younger Veterans and Hispanic Veterans

#### ARTICLE

### Veterans' attitudes regarding a database for genomic research

David Kaufman, PhD<sup>1</sup>, Juli Murphy, MS<sup>1</sup>, Lori Erby, PhD<sup>2</sup>, Kathy Hudson, PhD<sup>1</sup>, and Joan Scott, MS, CGC<sup>1</sup>

Purpose: Large cohort studies to investigate interactions between genes, environment, and lifestyle require large representative samples of the population. The Department of Veterans Affairs health care system is uniquely positioned to carry out such research, with a large patient population and a sophisticated system of electronic medical records. As the Veterans Affairs considers establishing a large database of genetic information and medical records for research purposes, a survey of 931 Veterans Affairs patients was carried out to measure their willingness to participate, what their concerns would be, and their preferences about

Some link genomic and other exposure data to electronic medical records, which can streamline data collection but may limit the information available to that recorded within a particular medical records system. Others perform standardized health exams and collect medical histories at baseline and follow-up intervals.<sup>7-0</sup> This approach may lead to more complete data, but also may require more time and expense. All such studies face the challenge of accurately and efficiently measuring participants' environmental exposures.

Genetics in Medicine, Vol 11:329-337, May 2009

©American College of Medical Genetics and Genomics

#### **ORIGINAL RESEARCH ARTICLE**

Genetics inMedicine

## Preferences for opt-in and opt-out enrollment and consent models in biobank research: a national survey of Veterans Administration patients

David Kaufman, PhD1, Juli Bollinger, MS1, Rachel Dvoskin, PhD1 and Joan Scott, MS, GCG2

Purpose: In 2006, the Department of Veterans Affairs launched the Genomic Medicine Program with the goal of using genomic information to personalize and improve health care for veterans. As tep toward this goal is the Million Veteran Program, which aims to enroll a million veterans in a longitudinal cohort study and establish a database with genomic, lifestly, emilitary-exposure, and health information. Before the launch of the Million Veteran Program, a survey of Department of Veterans Affairs patients was conducted to measure preferences for opt-in and opt-out models of enrollment and consent.

Methods: An online survey was conducted with a random sample of 451 veteran. The survey described the proposed Million Veteran Program database and asked respondents about the acceptability of opt-in and opt-out models of enrollment. The study examined differences in responses among demographic groups and relationships between beliefes about each model and willingeness to participate.

Results: Most respondents were willing to participate under both opt-in (80%) and opt-out (69%) models. Nearly 80% said they would be comfortable providing access to residual clinical samples for research. At least half of respondents did not strongly favor one model over the other; of those who expressed a preference, significantly more people said they would participate in a study using opt-in methods. Stronger preferences for the opt-in approach were expressed among younger patients and Hispanic patients.

Conclusion: Support for the study and willingness to participate were high for both enrollment models. The use of an opt-out model could impede recruitment of certain demographic groups, including Hispanic patients and patients under the age of 55 years.

Genet Med 2012:14(9):787-794

Key Words: biobank; genetics; informed consent; opt-out; recruitment: research ethics

## Research Challenges – Standardization and Uniformity of Processes

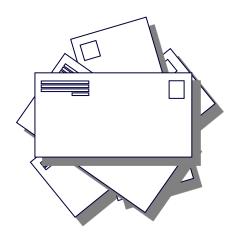
- Recruitment
- Consent & HIPAA
- Self reported data collections
- Sample collections
- Analytics
- Access
- Data sharing and repurposing
- A partnership with Veterans

### Distribution of MVP Sites



### **MVP** Recruitment and Enrollment

- Invitational Mailing/Appointment Mailing
  - Invitation letter, Baseline Survey, MVP Brochure
  - Appointment letter, Informed consent language
- Walk-in recruitment
- Study visit
  - Informed consent/HIPAA
  - Blood collection
- Thank You Mailing
  - Thank you letter, Lifestyle Survey
- Specimens sent daily
  - Receive & process 400-600/d
  - Stretch 1000/d





## VA Central Biorepository





### **MVP** Governance Structure

- VA Central Office Operations (GMP/CSP)
  - Oversees day-to-day operations at CSP coordinating centers (e.g., funding/resources, policy guidance, decision making)
  - VA Central IRB (internal)
    - VA Office of Research Oversight (external)
  - Provides coordination and facilitates communication among different groups
- MVP Steering Committee
  - Mission: In balance with Central Office (Operations), provides additional guidance and ideas to expand the Million Veteran Program. In coordination with the MVP Subcommittees, oversees organizational activities and makes recommendations to guide the Million Veteran Program
- MVP Subcommittees (dedicated to specific areas)
- Genomic Medicine Program Advisory Committee

## MVP Recruitment to Date

Invitation mailings sent	2.78 Million
Expressed interest by mail	13.9%
Optout	13.5%
Completed Baseline Surveys	461,709
Consented Veterans	362,662
Specimens in Lab	360,200
Unscheduled (proportion)	48.3%
Upcoming appointments	5,400
Call volume	Over 120,000 inbound

### Characteristics of MVP Enrollees

Age:	<50 yrs	13.8%	<u>Branch</u>	: A
	50-69 yrs	58.7%		Na
	≥70 yrs	27.5%		Ai
6		04.00/		M
<u>Sex</u> :	Male	91.9%		Na
	Female	8.1%		Co
Race:	White	77.5%		[0
	African-Amer 18.5%		Era:	th
	[other]	3.3%		7/
	[Native]	0.7%		2/
mala a tata a 110 a a a a a		E 40/		-, 8/
Ethnicit	<u>y</u> : Hispanic	5.4%		
				5
				8/

50.6% Army 21.2% lavy ir Force 16.7% 10.1% /larines lat'l Guard 0.5% Coast Guard 0.8% 0.1% other] hru 6/1950 5.3% 7/50-1/55 5.0% 2/55-7/64 7.1% 6/64-4/75 41.0% 5/75-7/90 11.8% 3/90-current 10.0% [multiple] 19.8%

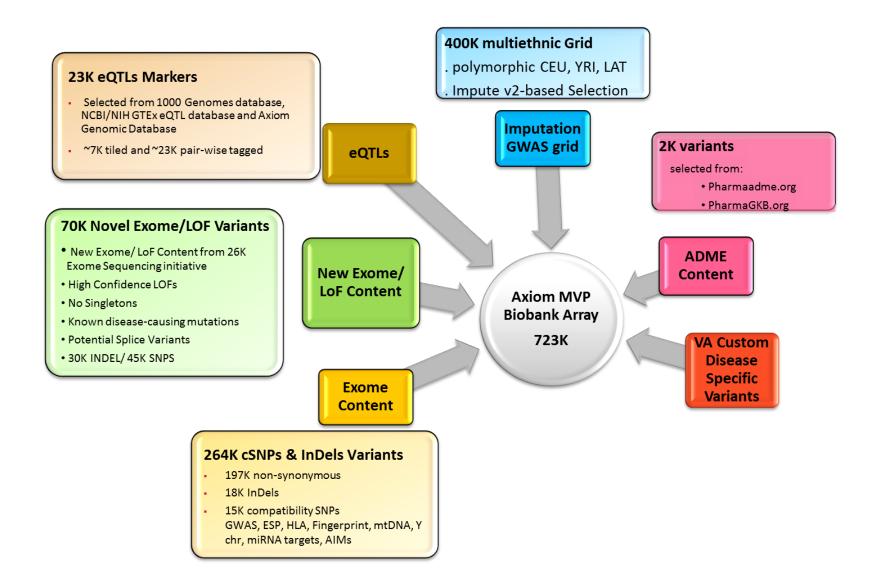
*Note: based on N*≤275,806 *enrollees* 

## Sample Breakout by Analysis Type (FY12-14)

	SNP Genotype	Exome	Genome
MVP	408,113	18,611	612
CSP 572 (functional disability			
of SZ and BPI) cases	9,356	9,356	285
CSP 575* (PTSD)	52,238	2,366	N/A
ALS			999

<sup>\*</sup>estimated number of PTSD cases from MVP cohort, based on self-reported survey data

## **Axiom MVP Biobank Array**



## MVP Genomic Analyses: Whole Exome Sequencing

- Contracts started 2013
- 50x depth coverage across 95% exome regions (~5-15 GB data/sample)
  - Personalis—Illumina
  - Claritas Genomics—Ion Torrent platform
- Data sent back to VA= raw data and annotated (fastq, BAM files, .vcf, quality scores and other metadata)
- Schizophrenia, Bipolar disorder, oversampled (>50%) for MVP African
   Americans





## MVP Genomic Analyses: Whole Genome Sequencing

- Contracts started 2012
- 30x depth coverage over 90% genome (~300GB data/sample)
  - Personalis—Illumina platform
  - Claritas Genomics—Ion Torrent platform
- Data sent back to VA as raw data and annotated (fastq, BAM files, .vcf, quality scores, some additional metadata)
- Samples = ALS, Schizophrenia, Bipolar disorder, and Exceptionally aged MVPers (95+)

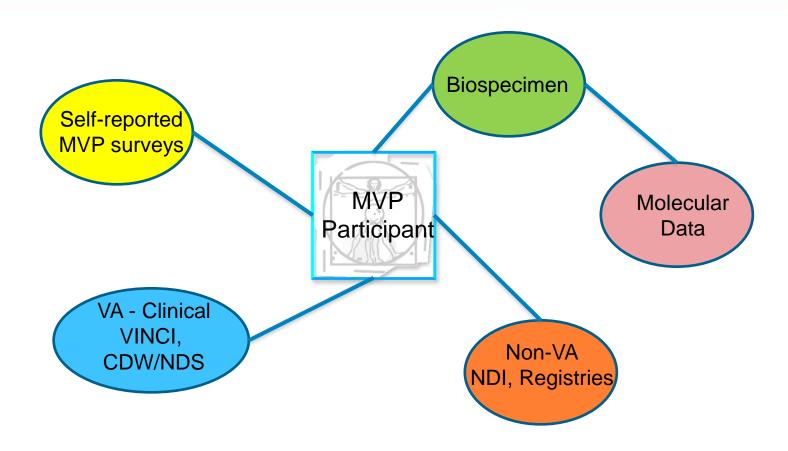




### Ongoing and Upcoming Sequence Data Analysis

- Whole Genome and Exome Sequence Data
  - Boston VA and Palo Alto VA
  - Perform quality check on variant data
  - Concordance with SNP arrays
    - ti/tv, het/hom, private variants, novel variants, missingness rates, gender, ethnicity
    - % un/mapped reads; coverage
    - % of exonic, intronic, intergenic variants
  - Aid in analysis of case/control studies
    - ALS; Schizophrenia/Bipolar; Exceptionally Aged MVPers (95+)
  - Development and testing of Analysis Pipeline VAPAHCS/Stanford
  - Cross-comparison of data derived from the Illumina and Ion Torrent platforms

### MVP Data Universe



### Other Data Sources

#### **MVP Data**

- Self-Reported Survey Data:
  - Lifestyle Survey Data (Personal Information, Well-Being, Activity, Health, Military Experience, Dietary Intake, Medication, Habits)
  - Baseline Survey (Health, Military Experiences, family medical history)
- Genetic Data
  - Genotype data
  - Sequence data

### **Other Data**

- VA Healthcare System Data
- Other Data
  - National Death Index (NDI)
  - Centers for Medicare and Medicaid Services (CMS)
  - State Mortality Data

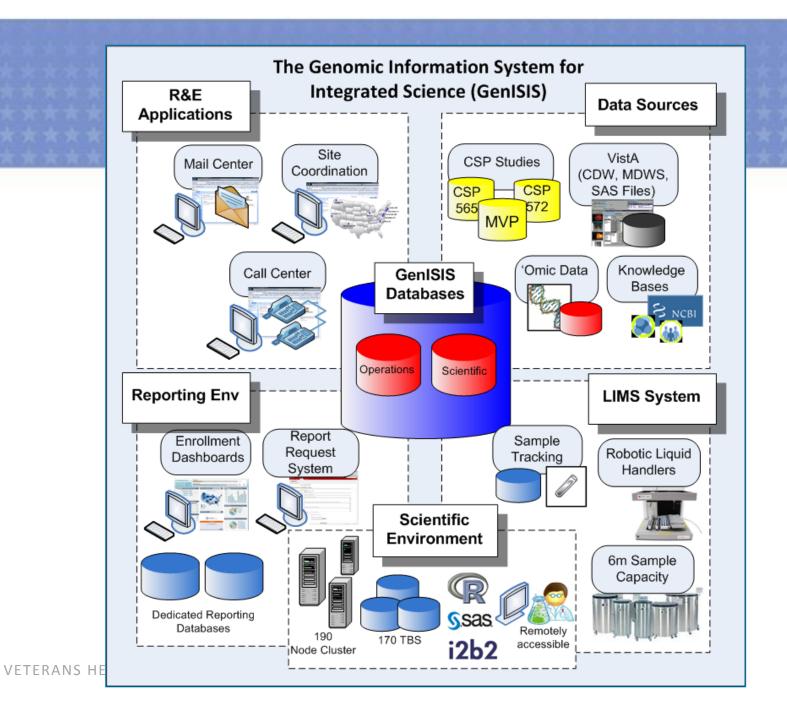
### VA Data Sources

- Corporate Data Warehouse Databases
- National Patient Care Databases
- Vital Status
- Decision Support System
- National Data Extract
- Beneficiary Identification Records Locator (BIRLS) death file
- New England VISN-1 Pharmacy files
- Outpatient Clinic File (OPC)
- Patient Treatment File (PTF)
- Inpatient and Outpatient Hospitalizations



Special
Data
Access w/
Data
Steward

- Clinic Inpatient and Outpatient Visits
- Diagnosis (ICD-9) codes
- Procedure (CPT) codes
- Pharmacy data and laboratory data
  - Pharmacy Benefit Management (PBM) system database
- OEF/OIF and OND Roster
- VA Clinical Assessment Reporting and Tracking (CART)
- Veterans Affairs Surgical Quality Improvement Program (VASQIP)
- Veterans Affairs Central Cancer Registry (VACCR)



### GenISIS Secure Computing Environment

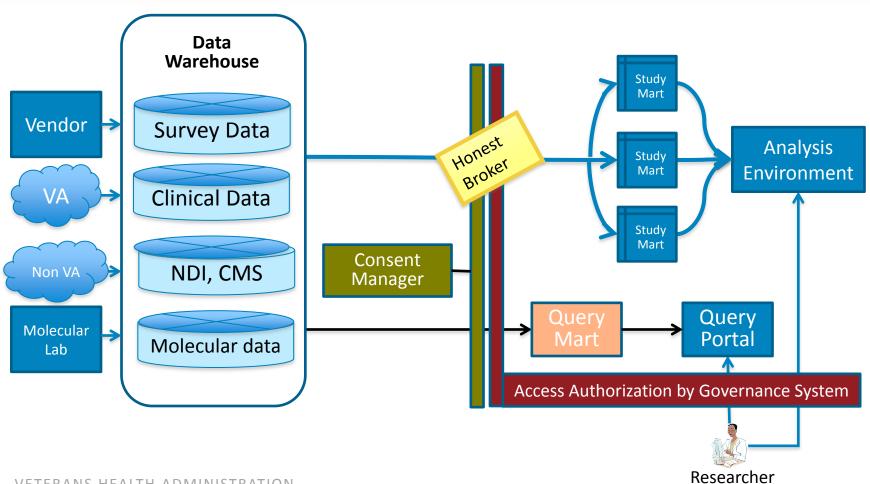
- Provides a secure computing environment with analytical tools
- Allows users to submit and manage their HPC jobs from a web dashboard
- Users authenticated by their VA network accounts
- Default 5 TB space allocated to each project, plus additional scratch space
- Analytic software available (PLINK, R/Bioconductor, SAS, Matlab, Perl/Bioperl, JMP Genomics and others) and will be updated based on user needs

### VINCI: Resource for Clinical Data

- Veterans Informatics and Computing Infrastructure (VINCI) securely hosts select data from national VA databases
- Provides data to credentialed VA investigators with appropriate approvals
- Data updated nightly for many clinical domains
- Provides services and tools for data provisioning, curating, NLP, analytics and data services, annotation and chart review, feasibility determination, and application development
  - Established an enclave with the MVP crosswalk within VINCI



### System Architecture



### **MVP Phenotyping Activities**

### **Complex Phenotypes**

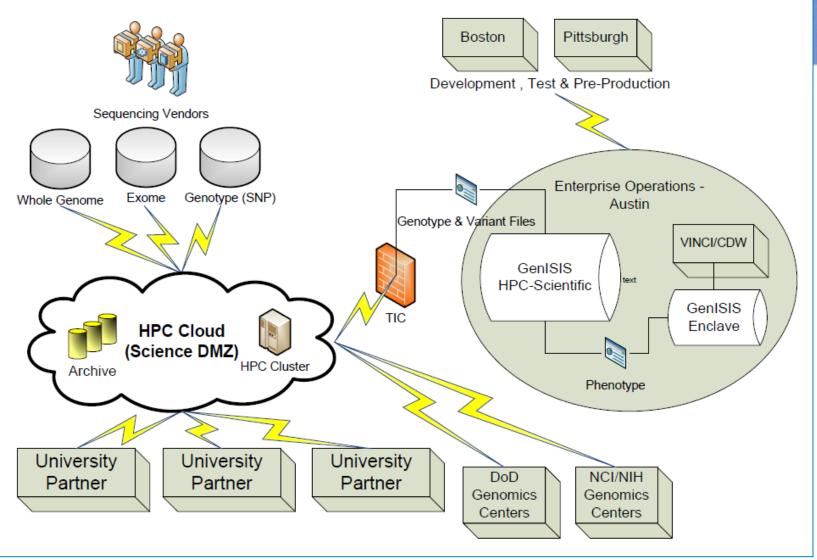
- Disease
  - Myocardial infarction (MI)
  - Stroke
  - Unstable angina with revascularization
  - Acute congestive heart failure
  - Death from cardiovascular disease
  - Vascular procedure
  - Posttraumatic stress disorder (PTSD)
  - Schizophrenia
  - Bipolar disorder
  - Traumatic brain injury
  - Depression
  - Vascular dementia
  - Cognitive impairment
  - Type 2 diabetes mellitus
- Other
  - Creatinine trajectory
  - Glucose trajectory

#### **Core Variables**

- Demographics
  - Age
  - Sex
  - Race
- Laboratory values
  - Total cholesterol
  - HDL, LDL
  - Albumin
  - Serum creatinine
  - Triglycerides
- Medications
- Other characteristics
  - Blood pressure
  - Height/weight/BMI
  - Smoking
  - Alcohol consumption
  - Combat exposure

## **Algorithm Development Validation Methods**

### MVP "To-Be" Architecture



### MVP Data Access Roll-Out

- Alpha Test Projects (Ongoing)
  - CSP572 (Genetics of functional disability in Schizophrenia and Bipolar Disorder); ~9500 cases deeply phenotyped; enrollment completed; Controls from MVP
    - genotyping and exome sequencing to be completed in FY15; data analysis in FY16
  - CSP575B (Genetics of PTSD in Veterans; Cases and Controls with self-reported combat exposure from MVP; ~10,000 cases
    - deep phenotyping ongoing; QC of genotype data ongoing; data analysis in FY 16
- Beta Test Projects
  - RFA for analysis of 200K genotyped dataset

### Acknowledgments

#### VA Central Office

Timothy O'Leary, M.D., Ph.D.; Sumitra Muralidhar, Ph.D.; Ronald Przygodzki, M.D.; Grant Huang, Ph.D., M.P.H.; James Breeling, M.D.; Jennifer Moser, Ph.D., Kendra Schaa, Kristina Hill, M.P.H.

#### MVP Steering Committee and Subcommittees

Steering Committee: John Concato, M.D., M.S., M.P.H., Co-Chair; J. Michael Gaziano, M.D., M.P.H., Co-Chair; Rayan Al Jurdi, M.D.; James L. Breeling, M.D.; Louis Fiore, M.D., M.P.H., ex-officio; John B. Harley, M.D., Ph.D.; Elizabeth Hauser, Ph.D.; Grant Huang, M.P.H., Ph.D. Ex-Officio; Phil Tsao, Ph.D.; Janet D. Crow, ORD Liaison; Philip D. Harvey, Ph.D.; Timothy J. O'Leary, M.D., Ph.D., Ex-Officio; Stephen G. Waxman, M.D., Ph.D.; Peter W. F. Wilson, M.D.

Access Policy Subcommittee: John, Concato, M.D., M.S., M.P.H., (Chair); Sumitra Muralidhar, Ph.D. (VACO Liaison); Michael Hart, M.D.; Fred Wright, M.D.; John (Tom) Callaghan M.D., Ph.D.; J. Michael Gaziano, M.D., M.P.H.; Saiju Pyarajan, Ph.D.; Carl Grunfeld, M.D., Ph.D.

*Chemical Analysis Subcommittee*: Philip Tsao, Ph.D. (Chair); Steven Schichman, M.D., Ph.D.; Saiju Pyarajan, Ph.D.; Chris Corless, M.D., Ph.D.; Robert Lundsten; Jenny Moser, Ph.D.; Sunil Ahuja, M.D.

Recruitment Subcommittee: Rayan Al Jurdi, M.D. (Chair); Jennifer Deen, Hermes Florez, M.D., Ph.D., M.P.H.; Amy Kilbourne, Ph.D., M.P.H.; Frank Lederle, M.D.; Kendra Schaa, ScM, C.G.C. (ORD Liaison); Ralph Heussner; Stephen Herring.

Phenotyping and Epidemiology Subcommittee: J. Michael Gaziano, M.D., M.P.H. (Chair); Jenny Moser, Ph.D.; Seth Eisen, M.D., M.Sc.; David Gagnon, M.D., Ph.D.; Saiju Pyarajan, Ph.D.; Jonathan Nebeker, M.D., M.S.; Denise Hynes, Ph.D., M.P.H., R.N.; Matthew Samore, M.D.

#### Boston Coordinating Center (MAVERIC)

J. Michael Gaziano, M.D., M.P.H.; Stacey Whitbourne, Ph.D., Jennifer Deen; Colleen Shannon, M.P.H.; Kelly Cho, Ph.D.; Xuan-Mai Nguyen, Ph.D.; David Gagnon, M.D., Ph.D.; Donald Pratt; Lindsay Hansen, M.P.A.; Keri Hannagan, M.P.H.; Derrick Morin

#### West Haven Coordinating Center (CERC)

John Concato, M.D., M.P.H.; Mihaela Aslan, Ph.D.; Peter Guarino, Ph.D.; Daniel Anderson; Rene LaFleur; Lesley Mancini, M.B.A.; John Russo; Cindy Cushing; Vales Jean-Paul; Fred Sayward; Louis Piscitelli; Donna Cavaliere

#### VA Central Biorepository - Boston

Mary Brophy, M.D., M.P.H.; Donald Humphries, Ph.D.; Robert Lundsten; Geetha Sugumaran, Ph.D.; Deborah Govan; Christine Govan

#### Genomic Information System for Integrated Sciences (GenISIS) - Boston

Louis Fiore, M.D., M.P.H.; Saiju Pyarajan, Ph.D.; John Gargas; Tony Chen, Ph.D.; Edmund Peirce; Paul Hseih, Ph.D.; Luis Selva, Ph.D.; Beth Katcher; Robert Hall, M.P.H.; Weijia Mo; Felicia Fitzmeyer; Anahit Aleksanyan

#### MVP Information Center - Canandaigua, NY

Annie Correa; Melissa Juhl; George Barzac III; Liam Cerveney; Anne Van Patten; Jonathan Martinez; Eileen Loh-Fontier; Jessica Marino

#### MVP Local Site Investigators

VA HEALTHCARE SYSTEM UPSTATE NEW YORK: Laurence Kaminsky, Ph.D. NEW MEXICO VA HEALTHCARE SYSTEM: Jose Canive, M.D. ATLANTA VA MEDICAL CENTER: Faroog Amin, M.D. VA MARYLAND HEALTH CARE SYSTEM: Alan Shuldiner, M.D.; Miriam Smyth, Ph.D. BAY PINES VA HEALTHCARE SYSTEM: Rachel McArdle, Ph.D.; Theodore Strickland, M.D., M.P.H., F.C.A.P. EDITH NOURSE ROGERS MEMORIAL VETERANS HOSPITAL: John Wells, Ph.D. BIRMINGHAM VA MEDICAL CENTER: Louis Dell'Italia, M.D.; Jasvinder Singh, M.D., M.P.H. VA BOSTON HEALTHCARE SYSTEM: Ildiko Halasz, M.D. VA WESTERN NEW YORK HEALTHCARE SYSTEM: Junzhe Xu, M.D.; Ali A. El Solh, M.D., M.P.H. CENTRAL TEXAS VETERANS HEALTH CARE SYSTEM: Keith Young, Ph.D. RALPH H. JOHNSON VA MEDICAL CENTER: Mark Hamner, M.D. CINCINNATI VA MEDICAL CENTER: John B. Harley, M.D., Ph.D. LOUIS STOKES CLEVELAND VA MEDICAL CENTER: Eric Konicki, M.D.; Curtis Donskey, M.D. WM. JENNINGS BRYAN DORN VA MEDICAL CENTER: Kathlyn Sue Haddock, R.N., Ph.D. VA NORTH TEXAS HEALTH CARE SYSTEM: Padmashri Rastogi, M.D. VA EASTERN COLORADO HEALTH CARE SYSTEM: Robert Keith, M.D. DURHAM VA MEDICAL CENTER: William Yancy, M.D. N. FL/S. GA VETERANS HEALTH SYSTEM: Peruvemba Sriram, M.D. HAMPTON VA MEDICAL CENTER: Marinell Mumford, Ph.D.; Pran Iruvanti, D.O., Ph.D. EDWARD HINES, JR. VA HOSPITAL: Salvador Gutierrez, M.D. MICHAEL E. DEBAKEY VA MEDICAL CENTER: Rayan Al Jurdi, M.D.; Laura Marsh, M.D. RICHARD ROUDEBUSH VA MEDICAL CENTER: John Callaghan, M.D., Ph.D.; Thomas Sharp, M.D. KANSAS CITY VA MEDICAL CENTER: Prashant Pandya, D.O.; Thomas Demark, M.D. VA EASTERN KANSAS HEALTH CARE SYSTEM: Mary Oehlert, Ph.D. CENTRAL ARKANSAS VETERANS HEALTH CARE SYSTEM: K. David Straub, M.D., Ph.D.; Sue Theus, Ph.D. VA LOMA LINDA HEALTHCARE SYSTEM: Ronald Fernando, M.D. VA LONG BEACH HEALTHCARE SYSTEM: Timothy Morgan, M.D. VA GREATER LOS ANGELES HEALTH CARE SYSTEM: Agnes Wallbom, M.D., M.S. WILLIAM S. MIDDLETON MEMORIAL VETERANS HOSPITAL: Robert Striker, M.D., Ph.D. VA MAINE HEALTHCARE SYSTEM: Ray Lash, M.D., F.A.C.P. VA MEDICAL CENTER MANCHESTER: Nora Ratcliffe, M.D. VA NEW YORK HARBOR HEALTHCARE SYSTEM: Scott Sherman, M.D., M.P.H. MEMPHIS VA MEDICAL CENTER: Richard Childress, M.D.; Marshall Elam, M.D., Ph.D. MIAMI VA HEALTH CARE SYSTEM: Hermes Florez, M.D., Ph.D. CLEMENT J. ZABLOCKI VA MEDICAL CENTER: Jeff Whittle, M.D., M.P.H. MINNEAPOLIS VA HEALTH CARE SYSTEM: Frank A. Lederle, M.D. VA TENNESSEE VALLEY HEALTHCARE SYSTEM: Adriana Hung, M.D., M.P.H.; Jeffrey Smith, M.D., Ph.D. CENTRAL WESTERN MASSACHUSETTS HEALTHCARE SYSTEM: Kristin Mattocks, Ph.D., M.P.H. ORLANDO VA MEDICAL CENTER: Kenneth Goldberg, M.D.; Adam Golden, M.D. VA PALO ALTO HEALTH CARE SYSTEM: Philip Tsao, Ph.D. PHILADELPHIA VA MEDICAL CENTER: Darshana Jhala, M.D.; Kyong-Mi Chang, M.D. PHOENIX VA HEALTH CARE SYSTEM: Samuel Aguayo, M.D. VA PITTSBURGH HEALTH CARE SYSTEM: Elif Sonel, M.D.; Gretchen Haas, Ph.D. PORTLAND VA MEDICAL CENTER: David Cohen, M.D. RICHMOND VA MEDICAL CENTER: Michael Godschalk, M.D. W.G. (BILL) HEFNER VA MEDICAL CENTER: Robin Hurley, M.D. VA SALT LAKE CITY HEALTH CARE SYSTEM: Laurence Meyer, M.D., Ph.D.; Vickie Venne, M.S. SOUTH TEXAS VETERANS HEALTH CARE SYSTEM: Sunil Ahuja, M.D.; Jacqueline Pugh, M.D. VA SAN DIEGO HEALTHCARE SYSTEM: Gwen Anderson, Ph.D., R.N. VA CARIBBEAN HEALTHCARE SYSTEM: Carlos Rosado-Rodriguez, M.D.; William Rodriguez, M.D. VA PUGET SOUND HEALTH CARE SYSTEM: Edward Boyko, M.D.; Karin Nelson, M.D. OVERTON BROOKS VA MEDICAL CENTER: Ronald Washburn, M.D. ST. LOUIS VA HEALTH CARE SYSTEM: Michael Rauchman, M.D. JAMES A. HALEY VETERANS' HOSPITAL: Stephen Mastorides, M.D.; Lauren Deland, R.N., M.P.H. SOUTHERN ARIZONA VA HEALTH CARE SYSTEM: Ronald Schifman, M.D.; Stephen Thomson, M.D. TUSCALOOSA VA MEDICAL CENTER: Lori Davis, M.D.; Patricia Pilkinton, M.D. WASHINGTON DC VA MEDICAL CENTER: Ayman Fanous, M.D. VA CONNECTICUT HEALTH CARE SYSTEM: Daniel Federman, M.D. WHITE RIVER JUNCTION VA MEDICAL CENTER: Brooks Robey, M.D.

## Thank you!

