Leveraging the Air Force Health Services Data Warehouse for Transformational Healthcare Research: *An Action Agenda*

October 2011

Prepared for

Prepared by

CHIDS Center for Health Information and Decision Systems

ROBERT H. SMITH SCHOOL OF BUSINESS
AT THE UNIVERSITY OF MARYLAND

Leveraging the Air Force Health Services Data Warehouse for Transformational Healthcare Research: *An Action Agenda*

Authors

Ritu Agarwal
University of Maryland

David Carnahan
Air Force Medical Service

Kenyon Crowley
University of Maryland

Albert Bonnema
Air Force Medical Service

Ahmed Calvo
Health Resources and Services Administration

Seth Eisen
Veterans Health Administration

Iain Sanderson
Medical University of South Carolina

Leveraging the Air Force Health Services Data Warehouse for Transformational Healthcare Research: *An Action Agenda*

Acknowledgements

This whitepaper represents a team effort in which individuals made contributions in addition to the authors. We gratefully acknowledge the assistance of the HSDW Research Committee members and other contributors listed below. Their thoughtful and insightful commentary added tremendous value and this whitepaper would not have been possible without their input. We also acknowledge Lt. Gen. (Dr.) Charles B. Green, Surgeon General of the Air Force, Headquarters U.S. Air Force, for the vision to translate the HSDW into an extramural research tool to advance the nation's health priorities.

Elaine Collier
National Institutes of Health

Kenneth Mandl
Harvard Medical School

JD Whitlock
Air Force Medical Service

Jesus Zarate
Air Force Medical Service

Leveraging the Air Force Health Services Data Warehouse for Transformational Healthcare Research: *An Action Agenda*

Executive Summary

The nation is poised at the brink of transformation in healthcare. Rising concerns about the quality and cost of healthcare have spurred concerted efforts to devise strategies that can be effective in improving patient safety, reducing medical errors, and driving systemic efficiencies throughout the healthcare value chain. The Institute of Medicine (IOM) has issued a national call for a "learning healthcare system" that is constructed on a foundation of systematic evidence, and predicated on the use of proven best practices. The US Department of Defense (DoD) is the owner and curator of a unique data resource that can be leveraged to address the IOM's healthcare transformation goals. The U.S. Air Force Health Services Data Warehouse (HSDW), which has been developed by the U.S. Air Force Medical Service (AFMS), consolidates clinical and administrative data from across the DoD for over 9 million beneficiaries from one of the nation's largest federally owned, integrated health systems through an interconnected set of clinical information systems.

The Air Force Office of the Surgeon General articulated a vision for making the HSDW a public resource for conducting research that can improve health services and healthcare not only for the military population, but for the entire nation. In May 2011 the Air Force Medical Service, in collaboration with the National Institutes of Health (NIH) National Center for Research Resources (NCRR) and Clinical and Translational Science Awards (CTSA) Consortium hosted a workshop to bring together stakeholders from the military, government, academic, and industry communities to engage in a dialog about how to optimally utilize this rare data resource. This white paper summarizes the deliberations of a research committee convened after the workshop to develop a near-term research agenda for the HSDW. It represents the collective input and priorities of different stakeholder communities: DoD, VA, HRSA, NIH, and Academia.

In selecting research priorities the committee sought to balance the needs of the military and other stakeholders, while simultaneously being mindful of research priorities for the nation as a whole. Three areas of research priority are suggested. These include: 1) Condition/disease specific research focused on traumatic brain injury (TBI), post-traumatic stress disorder (PTSD), and obesity; 2) Health services delivery research spanning the patient-centered medical home (PCMH) model, disease management programs, and care coordination across care transitions; and 3) data mining and knowledge discovery to generate novel hypotheses and find hither-to unexpected relationships in the data.

To accelerate the utilization of the HSDW by the scientific and practice communities and operationalize the research agenda, the committee identified three areas that require immediate and on-going attention. These are the development of a policy and governance framework, extending the HSDW to include data linkages with the VA, HRSA, and other federal and civilian health data repositories, and investigating the development of a data delivery model based on a third-party managed and curated de-identified data repository.

Leveraging the Air Force Health Services Data Warehouse for Transformational Healthcare Research: *An Action Agenda*

Background

The nation is poised at the brink of transformation in healthcare. Rising concerns about the quality and cost of healthcare have spurred concerted efforts to devise strategies that can be effective in improving patient safety, reducing medical errors, and driving systemic efficiencies throughout the healthcare value chain. Central to addressing the challenges facing our nation's healthcare system and realizing the vision of healthcare improvement is evidence-based practice, where clinical and administrative decision making is constructed on a foundation of systematic evidence, and predicated on the use of proven best practices. Key enablers of this vision are developments in information, communication, and storage technologies that support the capture, processing and analysis of data, and offer the potential of radically altering the practice and delivery of healthcare and medical research.

There is growing national and global interest in the use of large clinical and administrative data repositories in healthcare, and the number of available databases is rising. For example, the cancer Biomedical Informatics Grid (caBIG) project at the National Cancer Institute is a "virtual network of interconnected data, individuals, and organizations that work together to redefine how cancer research is conducted," and provides a robust and shareable infrastructure and an extensive set of tools to enable discovery in cancer research and care[1]. The Biomedical Translational Research Information Systems (BTRIS) project at the National Institutes of Health (NIH) consolidates clinical research data across three NIH institutes to facilitate data reuse and knowledge discovery[2]. The Veterans Health Administration's (VA) Office of Research and Development conducts and supports research for healthcare discovery and innovation to improve the lives of veterans and the nation's population. The VA has developed an extensive set of clinical and administrative data resources and provides opportunities for internal VA researchers and external investigators to propose and execute research studies[3]. Stanford University is consolidating clinical information on over 1.3 million pediatric and adult patients served by the Stanford Medical Center since 1995 into a standards-based informatics platform, Stanford Translational Research Integrated Database Environment (STRIDE), to more effectively address the needs of clinical researchers[4].

[1] caBIG® Community Website. (2011). National Cancer Intitute. Retrieved from https://cabig.nci.nih.gov/. Accessed July 20, 2011.
[2] Biomedical Translational Research Information System (BTRIS). (2011). National Institutes of Health. Retrieved from http://btris.nih.gov/. Accessed July 20, 2011.
[3] VA Office of Research and Development (ORD). (2011). United States Department of Veteran's Affairs. Retrieved from http://www.research.va.gov/. Accessed July 20, 2011.
[4] Lowe, H.J., Ferris, T.A., Hernandez, P.M., Weber, S.C. (2009). STRIDE: An Integrated Standards-Based Translational Research Informatics Platform. *AMIA Annual Symposium Proc*., 2009: 391–395.

The breakthroughs enabled through analyses of health data repositories span the gamut from clinical discoveries (e.g., the long-term effects of HIV/AIDS treatments), to cost effectiveness analyses (e.g., optimal medical therapy with or without percutaneous coronary intervention (PCI)), to improved understanding of healthcare utilization by populations (e.g., the relative need for physician-based medical services among users of complementary and alternative medicine and users of conventional medicine.) Robust collections of health data are critical to addressing the Institute of Medicine's (IOM) priorities that emphasize the need for a "learning healthcare system" where evidence-based decision making is used for continuous improvement in effectiveness, safety, quality, and efficiency of the healthcare system.

The US Department of Defense (DoD) is the owner and curator of a unique data resource that can be leveraged to address the IOM's healthcare transformation goals. The U.S. Air Force Health Services Data Warehouse (HSDW), which has been developed by the Air Force Medical Service (AFMS) and consolidates data from across the DoD, one of the nation's largest federally owned, integrated health systems through an interconnected set of clinical information systems. In FY 2008, the number of beneficiaries eligible for DoD medical care was 9.4 million. The HSDW merges data across the spectrum of inpatient and outpatient care, and uses the Military's AHLTA EHR as the foundation for capturing clinical data in the outpatient setting. Further, the HSDW integrates direct care and purchased care data in the same repository, which amplifies the research value for cost-effectiveness studies. The HSDW has been characterized as an "undiscovered lab" for conducting innovative and path breaking biomedical and health services utilization research[5].

The Air Force Office of the Surgeon General articulated a vision for making the HSDW a public resource for conducting research that can improve health services and healthcare not only for the military population, but for the entire nation. In May 2011 the Air Force Medical Service, in collaboration with the National Institutes of Health (NIH) National Center for Research Resources (NCRR) and Clinical and Translational Science Awards (CTSA) Consortium hosted a workshop to bring together stakeholders from the military, government, academic, and industry communities to engage in a dialog about how to optimally utilize this rare data resource. Workshop participants included representatives from HRSA, the VA, FDA, AHRQ, Indian Health Service, the Military Health System, major academic research institutions, and companies in the health information technology and healthcare spaces. Workshop discussions focused on research domains and questions that would be valuable from the perspective of the different stakeholders. At the conclusion of the workshop a research committee was formed to propose a set of near-term research priorities that address the needs of the military and the entire nation, and to develop a set of recommendations for driving the research agenda forward. This white paper summarizes the deliberations of the research committee. It represents the collective input and priorities of different stakeholder communities: DoD, VA, HRSA, NIH, and Academia.

[5] Gimbel, R.W., Pangaro, L., Barbour, G. (2010). America's "Undiscovered" Laboratory for Health Services Research. *Medical Care*, 48(8): 751-756.

The HSDW Resource

The HSDW consists of over 50TB of data aggregated from over 400 sites via 140 different data feeds (see Appendix A for more detail on HSDW data). It currently produces over 160 different daily or monthly data products for the AFMS, Army, Navy, Coast Guard, CDC, and others. The HSDW contains significant breadth and depth of data on over 9 million beneficiaries, approximately half of which should be considered covered lives. Because it contains both direct care and purchased care data in the same repository, it represents a unique dataset that is of enormous potential value to clinical research. And, in many ways, the Military Health System beneficiary base represented in the HSDW is a microcosm of the general American population. Men and women are nearly equally represented, there are large numbers of children and there is an expanding pool of aging beneficiaries. The rare combination of biological, clinical, administrative claims, and outcomes data in a single collection makes the HSDW a unique and powerful tool to support health services research. Further, a number of components have been architected and are integrated or are under development with the HSDW to facilitate usefulness and scalability, including:[6]

1. Informatica – Data integration software that can flexibly acquire, transport, and transform data.
2. SAS/SAS BI- a business intelligence and data mining software that can present and analyze data from simple reporting to complex statistical and predictive modeling.
3. Massively Parallel Processing (MPP) database- a database management system designed for very large volumes of data, both loading and querying in complex, high volume processes.
4. Enterprise Service Bus (ESB) - a Services Oriented Architecture (SOA) messaging coordinator designed to nimbly route data between information systems and transactional applications in low latency, high volume, transactional ecosystems; the ESB decouples applications by removing point-to-point interfaces creating a plug-and-play operational architecture removing many system based dependencies.
5. Master Data Management (MDM) – a custom or commercial product that maintains an accurate source of record for the reference data set; MDM enhances data quality at the point of data capture as well as the secondary analytical layers.
6. Natural Language Processing (NLP) & Text Analytics – an unstructured data miner that categorizes text by parsing the digital written language into component parts and making those parts easy to analyze through pattern searches, clustering, and other analytical means; NLP can be enhanced with a recognizable taxonomy or ontology managed through MDM.
7. Rules Engine- software that records, manages, and makes accessible for both application and user the business/clinical rules that define healthcare process and measurement.
8. DeID- de-identification software that masks Protected Health Information (PHI) allowing data consumers to research and perform analytical studies of AFMS patients while ensuring patient anonymity. De-identified, anonymized, limited, and custom privacy algorithms are all achievable outputs for selecting cohorts and research contextualized data.

[6] SPAWAR. (2011, May). *Research via the Air Force Health Services Data Warehouse: An Executive Summary on Enabling Medical Research*.

9. Data Federation & Virtualization: software that allows for the inclusion, mapping, and integration of remotely located data sets while maintaining both HIPAA privacy and non-custodial ownership or storage of the data. Remote data sets may include:
    a. Other healthcare institutions data
    b. Federal health sources (CMS, FDA, DEA, ONC, VA, etc.)
    c. Bio-repositories
    d. Bio-surveillance or syndromic surveillance systems
    e. Genotype sources
    f. Primary biological sequence search databases

The combination of these data tools and in conjunction with the large amount of rich longitudinal data has yielded an asset poised to provide a clinical and research environment capable of advancing clinical discovery, surveillance, evidence-based medicine and a range of benefits for scientists and policy makers across multiple disciplines.

Recommended Research Priorities

The committee acknowledged that given the richness of the data captured within it, the HSDW can be an enabler of a wide variety of research efforts, with distinct objectives and research questions. It was therefore challenging to identify a handful of areas that represent fruitful near-term opportunities. In selecting research priorities the committee sought to balance the needs of the military and other stakeholders, while simultaneously being mindful of research priorities for the nation as a whole. Three areas of research are recommended as near-term opportunities that will be of significant value for the military and have a high potential for impact in the national healthcare discourse.

1. Condition/Disease-specific research

The nation in general and the military in particular is grappling with a set of specific diseases, medical conditions, and injuries that contribute disproportionately to mortality and healthcare expenditures, and have long term implications for the health of individuals, their families, and the communities in which they live. The HSDW provides access to longitudinal data across a large population, supporting the construction of a variety of cohorts that can be used to address a range of important questions related to these conditions/diseases, including predictive models of risk for early diagnosis and prevention and the effectiveness of alternate therapies and medical procedures.

Initial focus is recommended on the following diseases/conditions:

a) ***Traumatic brain injury (TBI)***: The CDC estimates that 1.7m people each year sustain a TBI; of which 52,000 die. Direct medical and indirect expenditures

related to TBI cost the nation an estimated $60 billion in 2000[7]. In the U.S. Military, reported incidence of TBI for the time period 2000 to Q1 2010 is 178,876 cases[8]. The occurrence of TBI among members of the military has seen a sharp increase as a result of the wars in Iraq and Afghanistan; the DoD and Veteran's Brain Injury Center estimate that 22% of combat casualties from these conflicts are TBI. Because TBI is often difficult to detect and diagnose, this is a conservative estimate. An additional 60-80% of military personnel with other blast injuries may also have TBI[9].

The health implications of TBI are substantial. In addition to increases in mortality and morbidity, TBI has been associated with severe cognitive, physical, and emotional consequences. Cognitive consequences include memory loss, spatial disorientation, and problems with communication. Physical consequences may be manifest in the form of speech impairments, seizures, headaches and migraines, and balance problems. Finally, the emotional effects of TBI, which can range from impulsive behavior to agitation to depression, are a cause of considerable concern for the families of TBI patients. TBI also has serious adverse ramifications for the mental health of patients; those who have experienced TBI are at greater risk for psychiatric disorders compared to the general population.

The HSDW as a link to better TBI care at the bedside

The health services data warehouse will allow better linkages in translational research. For example, a service member that participates in genetic research may find that they have the genetic predisposition to develop seizures in the setting of traumatic brain injury. This information will be linked to his point of care data, which will inform his clinicians that they should treat him with seizure prophylaxis should he be unfortunate enough to experience traumatic brain injury. Having all the data in an organized, relational structure will advance the discovery of associations, and reduce the amount of time it will take to apply such discoveries at the bedside.

In addition to its classification as an important public health problem, given the higher likelihood of military personnel experiencing a TBI, it is of substantial

[7] Injury Prevention & Control: Traumatic Brain Injury. (2011). CDC. Retrieved from http://www.cdc.gov/TraumaticBrainInjury/statistics.html. Accessed August 16, 2011.
[8] Fisher, H. (2010). U.S. Military Casualty Statistics: Operation New Dawn, Operation Iraqi Freedom, and Operation Enduring Freedom, *Congressional Research Service* RS22452.
[9] Traumatic Injury and PTSD. (2011). United States Department of Veteran's Affairs National Center for PTSD. Retrieved from http://www.ptsd.va.gov/professional/pages/traumatic-brain-injury-ptsd.asp. Accessed July 20, 2011.

interest to the DoD and Veteran's Administration. TBI symptoms span a broad spectrum and because mild TBI symptoms are often common to other medical conditions, an accurate diagnosis is complex and challenging. Greater understanding and improved clinical guidelines and tools are needed for effective diagnosis of TBI. The HSDW can be used to address multiple research questions related to diagnosis such as differences in patient health indicators before and after a TBI diagnosis for populations with an accurate versus an inaccurate diagnosis; or the relative effectiveness of different diagnosis methods for accurate identification of the condition. Such research may yield important screening tools that could be deployed across the military.

TBI treatment also represents an important area for research that can usefully leverage the data in the HSDW. The data warehouse captures detailed longitudinal information about patients' health indicators and the treatments they receive. For the subset of patients with a TBI diagnosis, the data can be analyzed to, for example, understand the effectiveness of approaches to rehabilitation such as alternative forms of physical, occupational, and speech therapies, or general best practices in TBI treatment that are associated with the most positive long-term outcomes.

b) ***Post-Traumatic Stress Disorder (PTSD)***: Post-traumatic stress disorder, defined by the American Psychiatric Association[10] as an anxiety disorder that is manifest as a result of exposure to an extreme traumatic event, has been characterized as a "common" condition – an estimated 6.8% of Americans are expected to experience PTSD[11]. PTSD belongs to a general classification of mental disorders that are among the five most costly conditions in the nation. According to estimates derived from AHRQ's Medical Expenditure Panel Survey (MEPS), the costs of mental healthcare have steadily increased in the decade from 1996-2006. In 2006, total expenditure on mental healthcare was $57.5 billion for the 36.2 million Americans seeking treatment, resulting in an average expenditure per person of $1,591. According to World Health Organization estimates, the burden of disability as measured in disability adjusted life years (DALYs) is highest for neuropsychiatric disorders that contribute nearly twice as many DALYs as cancers and cardiovascular diseases.

The military population, both active duty and veterans, are considered "at-risk" groups for PTSD and have a higher than average incidence of this condition.

[10] American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR), 4th edition, Arlington, VA: American Psychiatric Publishing, Inc. 2000.

[11] Kessler, R.C., Berglund, P., Delmer, O., Jin, R., Merikangas, K.R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6): 593-602.

Although reliable estimates are not easily available, a recent analysis chronicles 21,784 newly reported cases of PTSD among those not deployed for the decade 2000-2010, and 66,935 new cases for personnel deployed during the same time period[12]. Exposure to stressors during combat is a leading cause of PTSD among military personnel: between 10-18% of troops deployed in Operation Enduring Freedom and Operation Iraqi Freedom are expected to experience PTSD upon return. The long-term adverse health effects of PTSD are substantial: PTSD patients are at increased risk for chronic illnesses such as cardiovascular diseases, arthritis, and certain cancers[13]. PTSD has also been associated with greater than average rates of obesity[14], increased likelihood of suicide, the fourth leading cause of death in the United States, and functional impairment[15]. Socially, PTSD has a variety of negative implications for patients, their families, and society at large, including domestic problems, parenting issues, absenteeism from work, lost productivity, and greater propensity for homelessness. The prevalence of PTSD places a substantial economic burden on the military and on society. A RAND study estimated the societal cost of PTSD and depression for approximately 1.6 million service members to range from $4.0 billion to $6.2 billion[16].

Although the DoD, the VA, and the medical research community have made considerable strides in understanding and improving the diagnosis and treatment of PTSD, and taken various actions such as the development of clinical practice guidelines, significant gaps in health service delivery and quality of care remain[17]. A large number of military personnel suffering from PTSD do not seek care for a variety of reasons ranging from fear of stigmatization and negative career consequences to availability of and access to providers. The treatment of PTSD also poses challenges. Providers do not consistently follow clinical practice guidelines, limiting the delivery of evidence-based care and quality improvements. Patients frequently fail to adhere to treatment regimens. There are unanswered questions and newly emerging evidence

[12] Fisher, H. (2010). U.S. Military Casualty Statistics: Operation New Dawn, Operation Iraqi Freedom, and Operation Enduring Freedom, Congressional Research Service RS22452.

[13] Harrison, J.P., Satterwhite, L.F., Ruday, W. (2010). The Financial Impact of Post-Traumatic Stress Disorder on Returning Military Personnel. *J Health Care Finance,* 36(4): 65-74.

[14] Vieweg, WV, Fernandez, A, Julius, D, Satterwhite, L, Benesek, J, Feuer, S, Oldham, R, Pandurangi, AK. (2006). Body Mass Index Relates to Males with Posttraumatic Stress Disorder. *Journal of the National Medical Association*, 98(4): 580–586.

[15] Thomas, JL, Wilk, JE, Riviere, LA, McGurk, D, Castro, CA, Hoge, CW. (2010). Prevalence of Mental Health Problems and Functional Impairment Among Active Component and National Guard Soldiers 3 and 12 Months Following Combat in Iraq. *Arch Gen Psychiatry.* 67(6):614-623.

[16] RAND. (2008). Invisible Wounds: Mental Health and Cognitive Care Needs of America's Returning Veterans (RB No. 9336-CCF).

[17] Fisher, H. (2010). U.S. Military Casualty Statistics: Operation New Dawn, Operation Iraqi Freedom, and Operation Enduring Freedom, Congressional Research Service RS22452.

related to the effectiveness of different pharmacological therapies. The FDA has approved only two pharmacologic agents for the treatment of PTSD, both belonging to the SSRI class of drugs. However, these treatments are effective only in about 50% of the patients and require continued medication to maintain positive treatment effects. Limited consensus exists related to second line treatments for PTSD[18], and the evidence base for the effectiveness of alternatives to the standard SSRIs is mixed. For example, a recent randomized controlled trial (RCT) showed that among patients with military-related PTSD with SRI-resistant symptoms, 6 month treatment with the second-generation antipsychotic risperidone compared with placebo did not reduce PTSD symptoms[19].

The HSDW is a unique resource that includes data on a substantial population that has been diagnosed with PTSD. It can be effectively leveraged to gain novel and breakthrough insights into the risk factors, diagnosis, and treatment of PTSD. In addition to studies related directly to patient outcomes, investigations can also focus on the effectiveness of different provider education strategies and their implications for patient outcomes such as adherence to medication, improvement in PTSD symptoms, etc.

c) ***Obesity:*** The economic consequences of obesity have been described as "staggering:" in 2008, these costs were estimated at $147 billion[20]. A 2010 report by the Society of Actuaries[21] reviewed over 500 articles on obesity and its relation with mortality and morbidity and, based on published evidence, concluded that obesity has a negative impact on health, mortality, and related costs. This study placed the annual costs attributable to overweight and obesity in the US and Canada at $300 billion. Overweight and obesity increases the risk of multiple adverse health conditions including type 2 diabetes, coronary heart disease, and certain cancers. Obesity has also been found to have a reciprocal relationship with mental health conditions, specifically, depression. A systematic review and meta-analysis of longitudinal studies confirms that obesity increases the risk of depression, and depression is predictive of

[18] Cukor, J, Olden, M, Lee, F, Difede, J. (2010). Evidence-based treatments for PTSD, new directions, and special challenges. *Ann. N.Y. Acad. Sci*., 1208: 82-89.

[19] Krystal, JH, Rosenheck, RA, Cramer, JA, Vessicchio, JC, Jones, KM, Vetrees, JE, Horney, RA, Huand, GD, Stock, C. , (2011). Adjunctive Risperidone Treatment for Antidepressant-Resistant Symptoms of Chronic Military Service–Related PTSD: A Randomized Trial. *JAMA,* 306(5):493-502.

[20] Finkelstein, EA, Trogdon, JG, Cohen, JW, and Dietz, W. (2009). Annual medical spending attributable to obesity: Payer- and service specific estimates. *Health Affairs,* 28(5): 822-831.

[21] Behan, DF and Cox, SH. (2010, December). Obesity and its Relation to Mortality and Morbidity Costs. Society of Actuaries. Retrieved from http://www.soa.org/files/pdf/research-2011-obesity-relation-mortality.pdf. Accessed July 30, 2011.

obesity[22]. The US Surgeon General has identified obesity as a condition of vital national concern: more than 1/3 of US adults and 17% of US children are currently obese, constituting what has been termed a "national obesity epidemic.". Obesity is widely regarded as an important indicator of population health that nations across the globe are seeking to proactively manage.

The obesity epidemic poses two critical challenges for the U.S. Military. One, the national prevalence of obesity considerably limits the pool of potential candidates for military recruitment. A recent report by Mission: Readiness, a non-profit national security organization led by 200 retired senior military leaders, concluded that being overweight or obese was the primary medical reason for applicants failing to qualify for military service[23]. The inability to find qualified applicants poses grave threats for national security. Two, although it may commonly be believed that because of stringent fitness requirements and extensive physical activity military members are at lower risk of being overweight and obese, surprisingly, this is not the case. In January 2011, the Medical Surveillance Monthly Report (MSMR) chronicled the rates and trends of overweight related diagnoses among active duty members. Alarmingly, during the thirteen year period spanning 1998-2010, the percent of active duty members with overweight related diagnoses during at least one outpatient encounter steadily increased, to an overall total of 382,448 members. The study found evidence that the median duration of active service of members after an overweight-related diagnosis was 15 months shorter than a matched control sample, and "…clearly documents the significant impacts on the active force of "nutritional casualties." Such casualties are not directly related to combat service or military-specific activities; nonetheless, they have large and costly impacts on the health, fitness, sense of well-being, and military operational capabilities of the active force." Across the wider DoD beneficiary pool, which includes family members and retired service members, the impact and cost of obesity is pronounced. A 2005 healthcare survey of all DoD beneficiaries found that, based on self-reported height and weight, approximately two-thirds of all MHS beneficiaries are overweight or obese[24].

The wealth of data available in the HSDW for active duty and retired military members and their families represents an important resource for understanding

[22] Luppino, F., de Witt LM, Bouvy, PF, Stijnen, T, Cuijpers, P., Bennix, BWJH, Zitman, FG. (2010). Overweight, Obesity, and Depression. *Arch Gen Psychiatry*, 67(3): 220-229.
[23] Too Fat to Fight: Retired Military Leaders Want Junk Food out of America's Schools. (2010). *Mission: Readiness*. Retrieved from http://www.missionreadiness.org/2010/too-fat-to-fight/. Accessed August 20, 2011.
[24] Overweight in the Military Issue Brief, Health Care Survey of DoD Beneficiaries. (2005). *Health Program Analysis and Education* Directorate. Retrieved from www.tricare.mil/survey/hcsurvey/issue-briefs/issuebriefCY05Q1.pdf. Accessed August 16, 2011.

the causes and consequences of obesity, as well as strategies for mitigation. For example, as the military implements nutritional and fitness interventions (a recent experiment is exploring the efficacy of yoga in the fight against obesity[25]), their effects on reducing the magnitude and incidence of obesity can be studied. The data support a detailed analysis of co-morbidities associated with overweight and obese diagnoses, as well as potential investigations into the relationship between various self-reported nutritional and fitness actions and the likelihood of an overweight or obese diagnoses. With increasing use of health management technologies such as personal health records that empower patients for health self-management, the effects of these technologies on alleviating obesity needs to be studied. Finally, risk factors for obesity such as length of active duty or other stressors, family history, and various demographic factors can be examined.

2. Health services delivery research

The pressures on the nation's healthcare system and national resources have made it abundantly clear that traditional approaches to health services need to be reexamined and augmented with innovations. There is significant interest in engaging patients as focal players in the management of their health and well-being[26], as well as making providers more accountable for the quality outcomes they deliver[27]. Recent federal legislation in the form of the HITECH Act and Affordable Care Act is requiring providers and healthcare organizations to implement new technologies, processes, and mechanisms for ensuring better coordination of care and accountability across the healthcare spectrum. This research domain is broadly focused on understanding the value of new models of care delivery and health management that would improve quality and reduce costs.

Three target areas are recommended for initial investigation.

a) ***Patient centered medical home (PCMH):*** It is well-established that the current model of primary health care in the U.S. must be revised and rebuilt. Costs are ever-increasing and the current reimbursements models account for a fraction of what is needed to provide comprehensive care[28]. In efforts to curb costs and increase quality in primary care, the patient-centered medical home model advocates a physician-led team approach to care with a focus on enhanced

[25] U.S. Military Experiments with Yoga to Cure Our Most Preventable National Security Issue. (2010). *Secure Nation* Retrieved from http://www.securenation.org/u-s-military-experiments-with-yoga-to-cure-our-most-preventable-national-security-issue/. Accessed August 16, 2011.
[26] Krist AH, Woolf SH. (2011). A vision for patient-centered health information systems. *JAMA,* 305:300-301.
[27] Physician Quality Reporting System. (2011). CMS. Retrieved from https://www.cms.gov/PQRS/Downloads/SODF_March22-2011_v03-15-2011_Final.pdf. Accessed July 18, 2011.
[28] Baron, R. J. (2010). What's Keeping Us So Busy in Primary Care? A Snapshot from One Practice. *New England Journal of Medicine*, 362: 1632-163.

access, coordination, continuity of care, quality of care, and payment reform[29]. This model is being tested in various forms through state and national pilot programs throughout the country. To achieve these goals, practitioners and their care teams must fundamentally reform their practice, especially how they manage information and communication. Incorporating supportive technologies is a vital component of reforms to increase access, quality, and coordination. Some have suggested that the notion of the PCMH needs to be expanded and broadened beyond just the primary care-provider team to include other dimensions of care such as mental and behavioral health and dental health. Such a setting may be characterized as a "Health Home" where health is viewed as a larger construct.

Services in the military have invested resources over many years developing and running medical home programs and are actively exploring new implementations of the PCMH concept leveraging enabling technologies such as personal health records and secure patient-provider electronic communication. The outcomes associated with such interventions are urgently in need of research. The HSDW captures data on a range of patient characteristics and health outcomes that can be investigated in pre- and post-designs to examine the effects of the PCMH on health indicators. As the Air Force continues its roll out of the PCMH initiative across multiple sites, there is an opportunity to explore PCMH outcomes in a natural experiment. When coupled with prospective field research, such studies could provide valuable insights into best practices not only for the military, but for the nation as a whole.

b) ***Disease management programs***: The concept of disease management is central to the IOM's goals of a "systems approach" to healthcare[30]. Fundamentally, disease management programs seek to coordinate healthcare interventions and communications to drive improvements in health self-management among populations with chronic conditions such as diabetes and hypertension[31]. They involve numerous components such as evidence-based practice guidelines, procedures for measuring outcomes and on-going evaluation, collaborative practice models, and extensive feedback loops. The evidence regarding the effectiveness of disease management programs across a range of diseases and medical conditions is mixed and researchers have voiced the need for additional studies of return on investment to guide the optimal allocation of scarce

[29] Joint Principles of the Patient-Centered Medical Home. (2007). *American Academy of Family Physicians American Academy of Pediatrics American College of Physicians American Osteopathic Association*, 1-3.
[30] IOM (Institute of Medicine). (2011). *Engineering a learning healthcare system: A look at the future: Workshop summary.* Washington, DC: The National Academies Press.
[31] CCA Disease Management Definition. (2011). Care Continuum Alliance. Accessed at http://www.carecontinuum.org/dm_definition.asp. Retrieved August 16, 2011.

resources. For example, a 2007 RAND report summarizing existence evidence on the effectiveness of disease management noted that although such interventions resulted in better disease control for certain conditions such as congestive heart failure, diabetes, and depression, evidence regarding their effectiveness in promoting healthier behaviors was inconclusive[32]. Likewise, a review of 27 studies published in 2009 concluded that it was difficult to reach definitive conclusions about the value of asthma related disease management programs for adults[33].

The U.S. Military, through the TRICARE system, has implemented a number of disease management programs for a range of conditions including heart failure, diabetes, depression, and asthma[34]. Various pilots have been launched at Military Treatment Facilities (MTFs) across the nation at different points in time, such as one for asthma in 2006, diabetes in 2007, and chronic obstructive pulmonary disease (COPD) in 2009[35]. The HSDW can be utilized to perform retrospective and prospective analyses on the effectiveness of these programs both in terms of their effects on health outcomes as well as cost-effectiveness compared with cohorts not enrolled in disease management programs. The insights gained from such studies will be extremely valuable for guiding national resource allocation decisions.

c) ***Coordination across care transitions***: The current system of healthcare delivery in the United States is fragmented, with limited coordination across different episodes of care and care settings. Patient safety and healthcare quality are often compromised as a consequence of inadequate communication and information exchange, resulting in adverse outcomes for population health such as frequent re-hospitalizations and failure to see improvements in patient health indicators. The economic ramifications of poor coordination across care transitions are also substantial: AHRQ reports[36] that American hospitals spent $31 billion in 2006 on more than 4 million patient stays that could potentially have been prevented with effective ambulatory care. To respond to this

[32] Mattke, S *et al.* (2007). Evidence for the effect of disease management: is $1 billion a year a good investment?. *American Journal of Managed Care,* 13(12): 670–676.
[33] Maciejewski ML, et al (2009). Adult asthma disease management: an analysis of studies, approaches, outcomes, and methods. *Respir Care,* 54(7): 878–886.
[34] Humana Military's Disease Management Home Page. (2011). *Humana's Military Services*. Retrieved from http://www.humana-military.com/south/bene/health-wellness/Disease-Management/DiseaseManagement.asp. Accessed August 15, 2011.
[35] Arday, DR, Dorn, P. (2010). MHS Disease Management Program: Current Results and Future Direction. Presentation made at 2010 Military Health System Conference.
[36] *More Than 4 Million Potentially Preventable Admissions Cost Hospitals Nearly $31 Billion*. (2009). AHRQ News and Numbers, April 29, 2009. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from http://www.ahrq.gov/news/nn/nn042909.htm.

challenge, a number of efforts to improve care transitions have been launched on a national scale. CMS initiated the Care Transitions initiative as part of the Medicare Quality Improvement Program[37], and this initiative has funded a number of pilot studies and trials.

With a population of patients that is fundamentally mobile as service members get assigned to different locations, the military environment offers a unique opportunity to study transitions in care. Patients are frequently reassigned to different primary care providers and may move between direct care at an MTF and a network provider, while providers themselves transition between locations and serve different patient groups. The technologies and systems used for healthcare information management also vary across military facilities. The HSDW captures information on patient and provider assignments, in addition to clinical and outcomes data. A number of important research questions can be fruitfully explored, including, what coordination mechanisms and data exchange capabilities are comparatively more effective, what are the consequences of effective transitions, what are the exact characteristics of information that needs to be exchanged (e.g., timeliness, depth) across transitions to balance patient privacy and optimal healthcare delivery? Future work could also explore the related issue of care transitions for military members and their families across their life-spans by combining data from the DoD and VA.

3. Data mining and knowledge discovery

The availability of clinical and administrative health data in electronic repositories has created exciting new opportunities for the progress of science. Advances in statistics and artificial intelligence now offer powerful capabilities to discover hither-to unknown relationships and associations in large data sets. Generally referred to as data mining or knowledge discovery in databases, the application of such techniques across various domains such as the retail industry and financial services has generated significant value. Data mining is defined as "…the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner[38]." The traditional model of clinical research has largely focused on hypothesis testing (i.e., a deductive approach to science), however, the research community is increasingly acknowledging the usefulness of inductive approaches for health services and clinical research that can potentially produce breakthrough insights. The new knowledge discovered through application of these techniques can then serve as the basis for generating novel hypotheses about risk factors for diseases, the effectiveness of different interventions, factors influencing adherence and

[37] Care Transitions Homepage. (2011). TMF Health Quality Institute. Retrieved from http://caretransitions.tmf.org/CareTransitionsHome/tabid/778/Default.aspx. Accessed August 3, 2011.
[38] Hand D, Mannila H, Smyth P. *Principles of Data Mining*. MIT, 2001.

persistence, etc. that can be subsequently tested and validated. Advancing the science of predictive modeling is an important priority for health researchers.

The scientific community has developed a wide array of data mining algorithms for classification (placing data into categories), clustering (discovering collections of objects that are similar to each other, such as patient records), and association (finding relationships among variables) that offer considerable promise in healthcare and biomedicine[39]. There is growing evidence to suggest that knowledge discovery tools can have an impact on patient safety, quality of care, cost reduction, and efficiency improvements in healthcare. For example, studies demonstrate that artificial neural networks can be utilized to predict length of stay for patients with spinal cord injuries[40], early detection of heart disease[41], and adverse drug events[42].

There are several opportunities for extending this body of work and we urge investigators to propose projects that would construct new models for data mining, perform initial tests of the models in simulated environments, and subsequently validate them using data from the HSDW. To illustrate, McCabe et al. report the development of predictive tools constructed from simulated treatment data to predict errors of omission in clinical patent data related to the diagnosis of type 2 diabetes[43]. The HSDW provides a useful repository for finer-grained calibration and validation of the approach.

We also recommend that investigators explore, test, and validate novel approaches to data analysis by applying the current and stretch capabilities of the technology. One promising area of exploration is the application of natural language processing (NLP) and associated text processing tools to the unstructured data captured in the HSDW in the form of clinical notes. Examples of NLP applications exist in the literature. Chen et al.[44] applied text mining (an NLP method) and statistical methods to an analysis of biomedical literature and clinical narratives from patient records to identify drug-disease associations. Luther et al.[45] analyzed clinical notes

[39] Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang JF, Hua L. (2011). Data Mining in Healthcare and Biomedicine: A Survey of the Literature. *J Med Syst.* 2011 May 3. [Epub ahead of print]

[40] Kraft MR, DeSouza KC, Androwich A. (2002). Data Mining in Healthcare Information Systems: Case Study of a Veteran's Administration Spinal Cord Injury Population. *Proceedings of 36th HICSS.* IEEE Computer Society. 2002.

[41] Cheng TH, Wei CP, Tseng VS. (2006). Feature Selection for Medical Data Mining: Comparisons of Expert Judgment and Automatic Approaches. *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems* (CBMS'06).

[42] Wilson A, Thabane L, Holbrook A (2004). Application of data mining techniques in pharmacovigilance. *British Journal of Clinical Pharmacology*. 57(2): 127-134.

[43] McCabe RM, Adomavicius G, Johnson PE, Ramsey G, Rund E, Rush WA, O'Connor PJ, Sperl-Hillen J-A. (2008, Aug). Using Data Mining to Predict Errors in Chronic Disease Care. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 3: Performance and Tools)*. Henriksen K, Battles JB, Keyes MA, et al., editors. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008 Aug.

[44] Chen ES, Hripcsak G, Xu H, Markatou M, Friedman C. (2008). Automated Acquisition of Disease-Drug Knowledge from Biomedical and Clinical Documents: An Initial Study. *J Am Med Inform Assoc*. 15(1):87-98.

[45] Luther S, Trembley M, Foulis P, Berndt D, French D, Rubenstein L. Using Knowledge Discovery in Databases to Identify and Characterize Fall Related Injuries. *AcademyHealth. Meeting* (2005: Boston, Mass.).

in the VA's electronic medical record and applied data mining tools to build models of key phases that predict a Fall-related injury.

The U.S. Military and the VA are actively exploring ways in which emerging tools and techniques can be leveraged to generate knowledge from medical data. As part of its effort to leverage knowledge assets more effectively, the DoD has awarded an SBIR grant to explore the application of data mining methods to medical diagnostics[46]. The VA supports data mining projects through its Health Services Research and Development (HSR&D) mechanism. It is recommended that the research community propose innovative and scientifically meritorious knowledge discovery projects that would explore and mine the data in the HSDW. Projects that seek to address the specific diseases/conditions and health services interventions described in this whitepaper are of greater priority in the near term.

Operationalizing the Research Agenda

This white paper has described an agenda for leveraging the data resources available in the HSDW to conduct leading-edge research that can address the healthcare challenges facing the military and the nation. Beyond identifying domains in which scientific endeavor is needed, there are additional components that need to be institutionalized prior to and on-going with the research activities. We briefly summarize those below, offering recommendations as appropriate.

1. **Policy and governance framework**

   The data warehouse cannot become a robust asset for investigators without the implementation of policies and governance mechanisms that will be used to safeguard this valuable resource, manage access to the data, protect the privacy of patients and the confidentiality of data. A recommended policy and governance framework is being made available in a related white paper being published soon. Researchers interested in utilizing the data warehouse are encouraged to review this document.

2. **Extending the capabilities of the HSDW**

   Public health priorities can be more effectively addressed by linking military health databases with the resources of federal agencies and external organizations. We recommend increased collaboration between the DoD, VA, HRSA, and other federal agencies for data sharing and linking. Such linkages would permit the investigation of a richer set of research questions such as exploring relationships between mental health issues encountered during active duty and homelessness as a veteran, long-term outcomes of therapies and interventions as individuals transition from the military to the VA, and the development of public-health surveillance models from military and civilian data that enable prediction of disease outbreaks. Potential

[46] Military Health Data Mining Algorithms Library (M-HDML). (2011). KBSI. Retrieved from http://www.kbsi.com/Research/MHDML.htm. Accessed August 16, 2011.

approaches suggested for initiating linkages include the launch of pilot studies where data linkage is done for a limited research question in an exploratory mode, and the creation of patient registries for specific diseases and conditions.

3. **Custodial responsibility for the HSDW**

   Although the military is the owner of the data in the HSDW, much of the research that is envisioned and recommended in this white paper is likely to use de-identified data. Access to it can be facilitated by exploring alternative models of data management. We recommend that the model of a third-party curated and managed de-identified data repository be actively explored. For example, an academic institution can be charged with the responsibility of managing de-identified data and making it available for the research community. An exemplar of such an approach is available in i2b2 (Informatics for Integrating Biology & the Bedside), a National Center for Biomedical Computing funded by NIH and based at Partners HealthCare System in Boston, MA. Data in i2b2 comes from hospitals affiliated with Harvard Medical School and is available for use by the research community through appropriate protocols. I2b2 has developed a vibrant community of clinical researchers, software developers, and other professionals that are actively exploring how to leverage patient data to improve healthcare practice and delivery. A strategy similar to that used in i2b2 may be usefully applied to the HSDW.

Conclusion

The Military and the nation are actively seeking solutions to the healthcare challenges facing our society and countries across the globe. There is a pressing need to find approaches that will improve patient safety and quality of care while simultaneously driving down the costs of healthcare delivery. There is also an urgent need to transfer knowledge from medical research to medical practice in a more timely fashion. Against this backdrop, enhanced utilization of existing healthcare data assets represents a promising avenue for breakthrough innovation in clinical and health services research. The HSDW can be a valuable resource in advancing discovery in biomedical research and driving the nation towards a learning healthcare system. The scientific community can leverage the data in this repository to provide novel and innovative insights into safer and more effective healthcare delivery. We recommend that policies and structures be instituted that would expedite use of the HSDW by biomedical and health informatics researchers.

How to Get Involved

Over the coming months, opportunities to participate in pilot research opportunities will be made available. To receive updates on the continued evolution of the initiative including funding announcements, send an email to chids@rhsmith.umd.edu.

Appendix A: HSDW Data Elements

The following table describes the data in the HSDW. Although the HSDW does not include all data in the Military's Clinical Data Repository (CDR), additional desired CDR data fields will be made available in the future.

Military Treatment Facility Service Events	Civilian Network	Occupational
• Appointments • Hospitalizations • Laboratory • Radiology • Pharmacy • Enrollment • Lab, Rad, Rx Orders • Referrals • Allergies • Immunizations • MEPRS (Cost Accounting) • Pathology	• Professional Claims • Institutional Claims • VA Catchment • Soc Security Death	• Medical Readiness • Disability • Eye Glass Rx • Dental Status • Deployments • Agreements for Environmental Exposures • Safety Center Data
EMR	**Health Behavior and Risks**	**Measures**
• Vital Signs • Wellness Reminders • Historical Procedures	• Self-Reported Health Assessment • Fitness • Weight/BMI • Deployment Surveys • Adjusted Clinical Groups (ACGs)	• HEDIS-10 **+ 7** • AHRQ Patient Safety QI • AHRQ Inpatient QI • AHRQ Pediatric QI • AHRQ Prevention QI • Framingham CRF • Behavioral Risk Factor Surveillance System • Public Health Reportable Events • CMS Mortality Measures • PTSD Surveillance

This whitepaper brought to you by

AFMS works in close coordination with the Assistant Secretary of Defense for Health Affairs, the major air command surgeons, the Departments of the Army, Navy, and other government agencies to deliver medical service for just over 2 million eligible beneficiaries. As part of a dedicated team of health professionals, AFMS helps lead the way to a strong, healthier tomorrow by providing Trusted Care, Anywhere.

Prepared in partnership with

CHIDS Center for Health Information and Decision Systems

ROBERT H. SMITH SCHOOL OF BUSINESS
AT THE UNIVERSITY OF MARYLAND

CHIDS is an academic research center in the Robert H. Smith School of Business. It works in collaboration with industry and federal, state, and local government affiliates, and is designed to research, analyze, and recommend solutions to challenges surrounding the introduction and integration of information and decision technologies into the healthcare system. The research at CHIDS seeks to understand how digital technologies can be more effectively deployed to address outcomes such as patient safety, healthcare quality, efficiency in healthcare delivery, and a reduction in health disparities. CHIDS offers the benefit of a world-class research staff and renowned scholars in the economic, social, behavioral, and managerial aspects of technology implementation, adoption, assimilation, and return on investment. CHIDS serves as a focal point for thought leadership around the topic of health information and decision systems.

Suggested citation: Agarwal, R., Carnahan, D., Crowley, P.K., Bonnema, A., Calvo, A., Eisen, S., & Sanderson, I. (2011, October). Leveraging the Air Force HSDW for Transformational Healthcare Research: *An Action Agenda (Air Force Medical Service Whitepaper).*