This paper was commissioned for the Committee on the Well-being of Military Families. Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Committee on the Well-being of Military Families or the National Academy of Sciences, Engineering, and Medicine.

# Well-being of Military Families and a Learning Health System: The Importance of Data-Driven Decision Making

Prepared by:
**Schelomo Marmor, PhD, MPH**
MOR Health Advisory Group, LLC

Commissioned by the Committee on the Well-being of Military Families to provide background to the National Academies of Sciences, Engineering, and Medicine's consensus study on the potential of technology to further understanding of family well-being.

Submitted on: 8/14/2018

**Table of Contents**

**Executive Summary**

Information technology, the volume of data captured by electronic health records (EHRs), and the growing capacity for linkage of structured and unstructured data has created enormous opportunities for measuring outcomes and, consequently, for improving population health. This revolution in big data analytics and machine learning promises increasingly more accurate predictions and advancement of best practices in public health and medicine and their application to military readiness and family well-being.

This paper, “Well-being of Military Families and a Learning Health System: The Importance of Data-Driven Decision Making,” commissioned by the Committee on the Well-being of Military Families, is designed to help the committee better understand the processes by which data are collected, analyzed, and applied to guide decision making at the system and patient levels for military service members and their families. The paper presents opportunities for enhancing collection and analysis of data within the context of the big data and machine learning revolution.

To implement these opportunities effectively, efforts should be driven by cautious optimism, science, innovation, and individual safety. Data should be used to perform comparative effectiveness research to determine what works well for both individual patients and systems of care. This will ensure that patients receive, and systems deliver, optimal care. Summarizing and disseminating best practices in the use of big data and machine learning, including both the approach and the algorithms used to generate conclusions, is critical. Smarter big data systems can be used to prompt clinicians to provide specific, individualized treatment plans for each patient, whether based on simple demographics or on advanced biomarkers, in order to fulfill the promise of personalized medicine. Embracing the developing field of knowledge translation will enable successful efforts to propagate from small pockets of excellence to the civilian and military health systems as a whole. However, cautious optimism should be placed in how the system generates policy ideas using big data, as opaque results could impair policy decisions.

Used effectively, big data analytics and machine learning can increase the knowledge of key stakeholders on how to support better decision making on military readiness and well-being and improve health outcomes in military members and their families.

## I. Introduction

The rapid expansion of big data and computational machine learning approaches promise ever more accurate predictions of innovation and advancement of best practices in medicine. In 1965, Gordon Moore, the co-founder of Intel, posited computer chips would double in speed every two years.[1] The speed at which health data is generated suggests that best practices in healthcare and public health would easily surpass the conventional growth trajectory suggested by Moore's law, but significant challenges remain.

Innovation and advancement of best practices in medicine are often well disseminated throughout clinical and research communities. However, because of the gap between the actual and potential performance of the United States health care system, which the Institute of Medicine (IOM) termed a "quality chasm,"[2] only incremental progress in healthcare quality improvement has been seen, and it can take decades to fully benefit from translation of clinical research to improved best practices. A similar divide has been seen in the data revolution. As private sector and industry make major headway using data to inform and predict better outcomes, the healthcare system has been slow to adopt this approach. Although data has always been at the forefront of healthcare innovation and research, the use of big data harmonized across multiple layers of input and machine learning is still in its nascent stage.

Computing innovations and the proliferation of user-specific data have opened an incredible library of data and the estimated rate of internet traffic is expected to rise even further, reaching or surpassing 50% in parts of Latin America, the Middle East and Africa.[3] Although this massive collection of data points could potentially be perceived as a meaningless collection of unrelated moments in time, efforts have been and must continue to be made to leverage complex unstructured seemingly unrelated data into actionable insights.

Importantly, the United Nations posits that although big data presents the most value for global health in low-resource settings, it is also most vulnerable to fragmentation and misuse in such settings. Collaborative governance, careful analysis, and technical partnerships are needed to minimize the risks.[4] These conclusions apply equally to the military and their families and may be especially important in the early adoption of big data applications. The racial and socio-economic diversity of the military and their families further mirrors the challenges across global regions to standardize big data usage and application and the importance of establishing best practices to avoid misuse.

### *A. Organization of the Paper and Objectives*

This paper provides a brief overview of the big data field and the potential of utilizing predictive analytics to learn more about military family readiness and well-being. We define and introduce big data within the framework of a continuous learning system. Next, we present a conceptual model that incorporates the Donabedian framework[5] of healthcare quality across the three domains of structure, process, and outcomes. Within this model, we underscore data security and safety. We use a "SWOT" analysis approach to outline, identify, and evaluate the strengths, weaknesses, opportunities and threats of big data strategies for military family readiness and well-being. Finally, we summarize key findings, and offer some recommendations for future directions.

### *B. Definitions and Review of the Literature*

#### 1. Definition of terms

Big data, first defined in 2003, is the rapidly increasing volume of available data, the velocity at which data are generated, and the ways in which the data are represented.[6] A "learning health system" framework is defined by the IOM as a system in which "science, informatics, incentives, and culture are aligned for continuous improvement and innovation, with best practices seamlessly embedded in the delivery process and new knowledge captured as an integral by-product of the delivery experience."[7] The IOM has categorized characteristics of a continuously learning health system into four major groups: (1) Science and informatics, with real-time access to knowledge and digital capture of the care experience; (2) Patient-clinician-data scientist partnerships with engaged and empowered patients; (3) Incentives that are aligned for value, but with full transparency; and (4) Continuous learning culture, with a leadership-instilled culture of learning and supportive system competencies.[8]

To be successfully utilized for military family readiness and well-being, big data must be integrated into these four major categories while also following the plan, do, study, act (PDSA) principles[9] as outlined in Table 1.

**Table 1.** Big data utilization within a continuously learning health system using the IOM learning system framework integrated with Plan Do Study Act (PDSA) principles.

| Science and informatics | Patient-clinician-data scientists partnerships | Incentives | Continuous learning culture |
|---|---|---|---|
| Plan: real-time access to knowledge and digital capture of all components of the care experience for military families in data-safe environments | Do: Engage and empower military families in the data being captured with data use agreements that emphasize enhanced data security | Study: Collect meaningful data aligned with values in military families; create a fully transparent, data-safe system that avoids wasting resources and inaccurate predictions supporting poor decisions | Act: Create a leadership-instilled culture of rigorous continuous review of machine learning algorithms driven by a multidisciplinary thought team that critically evaluates policies and preserves data safety |

2. Past relevant National Academy of Science, Engineering and Medicine (NASEM) studies

Several important NASEM reports have outlined the benefits of using big data to improve healthcare for active service members and to provide insight into how future veterans will access care. For example, predictive analytics such as machine learning techniques offer the promise of finding patterns in huge datasets to improve military readiness as defined by a Booz Allen report.[10] In 2014, a report of the Military Health System (MHS) found that the MHS has access to an enormous amount of data, but a limited ability to analyze those data and to use the results to guide patient safety and policy.[11]

A problem not unique to the MHS is that, without a common set of metrics and measures, it is difficult to utilize system-wide data in a coherent fashion and arrive and conclusive policy decisions. Targeted approaches may be more useful. A 2018 NASEM report of the Veterans Affairs (VA) Mental Health Services reported that younger veterans tended to be more open to obtaining mental health care using the internet.[12] Capitalizing on this, big data and machine learning could be used to capably increase the quality of real-time care provided online. In another NASEM report, the committee recommended that the Department of Defense implement comprehensive family and patient centered evidence-based prevention programming" specifically directed to psychological health in military families, spouses, partners, and children.[13] Such targeted strategies are likely to be most successful when borne out of an effective learning health system that incorporates the PDSA cycle outlined in Table 1. Importantly, data safety must be a core competency of any big data learning system framework. In a 2017 NASEM report, the authors recommended that the government adopt the "five safes" (safe projects, safe people, safe settings, safe data, and safe outputs) to prevent privacy breaches and minimize the risk of data being handled incorrectly.[14]

## II. Conceptual Model

Dr. Avedis Donabedian suggested the best known conceptual framework for quality improvement work in medicine. The model focuses scientific study of quality across three domains: structure, process, and outcomes.[5] The National Quality Measures Clearinghouse defines structure as "a feature of a healthcare organization or clinician related to the capacity to maximize high quality care."[15] Structural elements include physical resources (eg, buildings and equipment) as well as human resources (eg, physicians, nurses, and mental health professionals). It also includes organizational protocols and/or guidelines. A process measure is defined as "a health care-related activity performed for, on behalf of, or by a patient."[15] Thus, unlike structure elements, a process measure applies on an individual patient level. Outcomes, which are defined as the "health state of a patient resulting from health care,"[15] may be clinical (eg, mortality, morbidity, and complications) or patient-reported (eg, symptoms and quality of life).

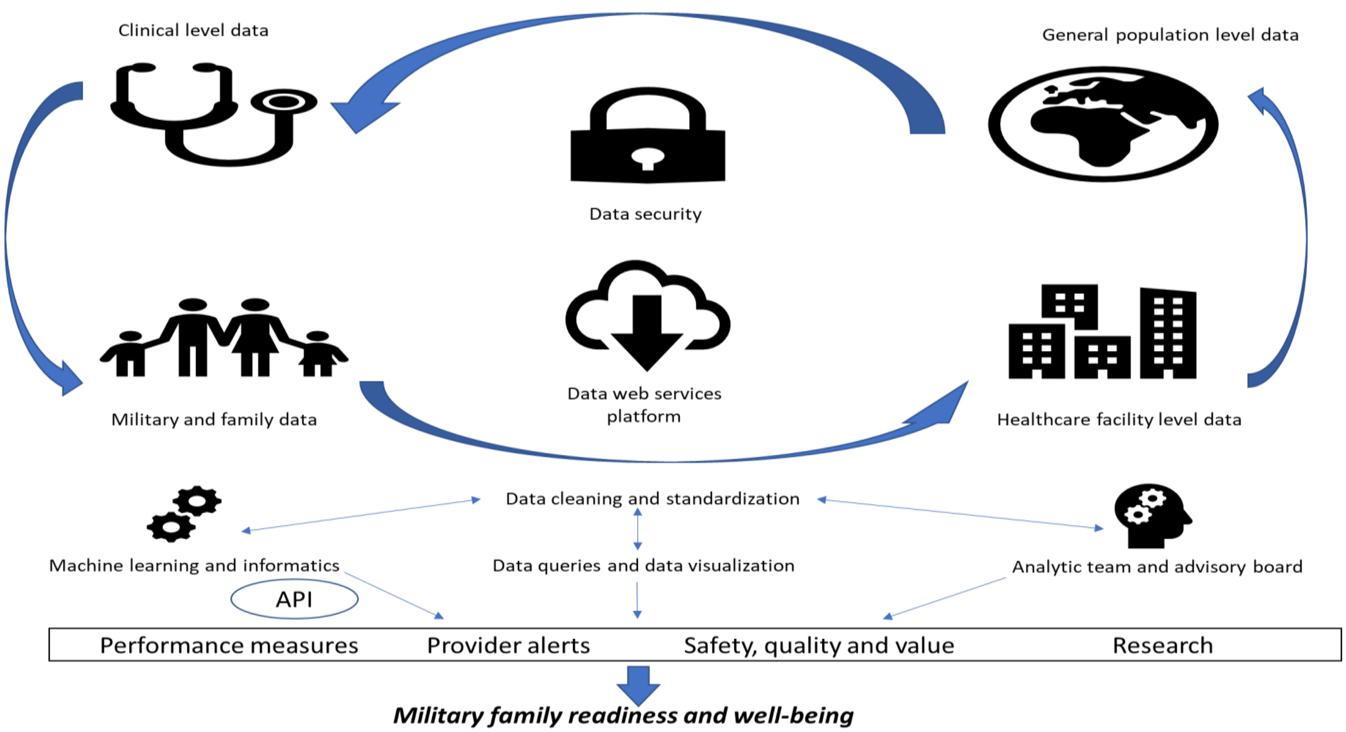
The conceptual model shown in Figure 1 illustrates how collection, analysis, and dissemination of big data can use the Donabedian framework to provide higher quality care to and improve well-being of military families. The structural elements include the infrastructure and data components from which the data points are collected, including the provider teams, active service members and their families, and health care facilities, as well as population-level data for the group as a whole. As part of the process measures, the data are transferred to a web services platform that allows for data cleaning, standardization, and visualization. Output to the end user is through application programing interfaces (APIs). This stack of informatics, which sits across many primary data sources, including EHRs and clinician notes, supplies the analytics needed to give researchers, health care providers, and patients valuable, real-time information on military family readiness and well-being.

Figure 1. Conceptual model of big data collection, analysis, and dissemination to improve military family readiness and well-being.

*Structure*

*Process*

*Outcomes*



The analytic team and advisory board is a key part of the process measures performed for or on behalf of military families. Not only do they oversee compliance with data use agreements and patient and data safety, but also prevent abuse and misuse of data to ensure that military policy makers and end users do not misinterpret big data results and analyses. The real-time information in the outcomes portion of the model provides care teams and patients with performance measures; provider alerts; safety, quality, and value notifications; and research data points, all of which can be used for future analyses.

## III. SWOT Analysis

### *A. Strengths*

In the United States, early efforts at using data to reduce costs and drive improvements in outcomes focused on administrative claims data, which was readily available through the Centers for Medicare & Medicaid Services (CMS). Administrative claims data are routinely collected for comparative effectiveness of treatments, payment reimbursements, and health services research. CMS data, in particular, are relatively easy to analyze and are replete with searchable data points. However, claims data do not always capture the nuances of comorbidities, severity, conditions present on admission, complications, patient experience, and other socio-demographic and socioeconomic factors critical to understanding health outcomes. Moreover, administrative data rarely capture the real-time health dynamics of family well-being and mental health.

Big data and predictive analytics that are grounded in a learning health system have significant potential to promote military family readiness and well-being by supporting forward-looking decisions and healthcare policies. The system of health care delivery, medical education and training, public health, civilian sector partnerships, research and development, and performance improvement included in the MHS lends itself well to the development of predictive analytics in this arena. Large amounts of clinical, socio-demographic, and socioeconomic data can be merged and harmonized in order to systematize clinical and public health best practices to become the norm. Broad military family participation, providing extensive data points from both quantitative and qualitative data sources, would facilitate optimal healthcare and maximize military family readiness and well-being.

In 1996-1997, the VA created the Veterans Health Information System and Technology Architecture (VISTA) to launch the Computerized Patient Record System (CPRS), an EHR for clinicians throughout the VA health care system.[16]Although CPRS was created at the national level, VA facilities was given some flexibility to personalize their systems. This combination of national implementation and individualized customization has been instrumental in ensuring that the system is well integrated within the VA, allows for data sharing across VA sites, and is widely accepted by end users.[17] Broad military family participation in CPRS will provide even more data points to maximize military family readiness and well-being.

Other health systems have a successful track record with using big data to improve outcomes in imaging and head injuries,[18-19] and may serve as a model for military communities. Harvesting data across multiple EHR vendors, registries, and payers allows providers with real-time tools to improve the quality and value of care, allowing for a renewed focus on preventative care and public health.[20] Private sector initiatives that trawl through patient data to provide better care for low-income Medicaid beneficiaries[21] account for socioeconomic and socio-demographic differences, which is similarly important in military communities considering the diversity of trainees, active duty members, and their families. Ultimately, big data that is generated by military service members and their families and networks is the backbone of the learning health system. Incorporating predictive analytics to learn more about military family readiness and well-being will optimize public health in the military.

*B. Weaknesses*

Despite the ability of big data and predictive analytics to surpass the gains made of analyzing claims data, gaps remain in the homogeneity, quantity, and quality of the data collected. Issues involving data inaccuracy, data missingness, and selective measurement remain substantial concerns. Robust systems allow for real-time data access, natural language processing, and digital capture of health care but are currently limited to certain medical conditions and have yet to fully integrate socioeconomic and socio-demographic factors.

The quality of any learning health system, especially one that is rooted in big data, is contingent on the data from which it is derived. Erroneous and ambiguous data points, as well as misaligned assumptions and algorithms, could lead to exponentially flawed policy conclusions. Moreover, building interoperability into the system requires thoughtfulness and foresight. Currently available EHRs have heterogeneous architectures not built for big data analyses and a learning health system framework, and are further limited by clinicians and scribes entering information that cannot easily be prepared for predictive analytics. Without a fully integrated system, the data is more difficult to harmonize.

Varying degrees of information and data-driven decisions and machine learning can yield non-causal findings that can introduce statistically "noisy" and difficult to interpret results. Although the advisory board could help set big data standards and ensure integrity of the results, military policy makers and end users could potentially misinterpret big data results and analyses, causing widespread confusion. This would magnify the skepticism of military clinical, medical, and research leaders of big data analytic results because of the possibility of false-positive conclusions. Inferred causality leading to misinformed policy decisions would further add to public scrutiny and mistrust. For example, as part of the government's National Cancer Moonshot program, a public-private partnership has been developed to help doctors expand and scale access to precision medicine for American veterans with cancer.[22] Although the technology is optimistically expected to help in the department's precision oncology program by providing information to help physicians identify precision treatment options for nearly 30 times more patients than could be previously served, it has so far shown only fair to middling results.[22-23] Policy ideas based on big data analytics should be generated and viewed cautiously as these systems are still underdeveloped.

*C. Opportunities*

The use of technology has yet to be maximized and globally integrated into data science and public health. Big data analytics could include population level surveillance, personalized education, self-management, and intervention delivery, as well as continuous quality improvement processes. Harmonization of disparate data sources will undoubtedly expand the analytic data silo framework in which health systems currently process data, allowing not only for a snapshot in time but also an ability to view individuals over the life course.

Moreover, foundational work implemented now could establish cost savings in the future as predictive analytic tools mature. This cost savings is most apparent in data safety, given the significant costs associated with the storage and necessary protections of data. Creating a learning health system in which data safety is paramount will avoid a need to spend future resources on preventing data breaches. A ground-up approach in building data systems with safety at the forefront will minimize future costs and maximize the trust and usability of the systems for the military and their families.

Cost savings can also be seen if scalable systems are built to incorporate future data sources. The advanced interconnectivity and data collection of mobile and wearable devices is still in its nascent stages, especially in relation to the collection of meaningful health data. Ensuring that the system can integrate usable data from wearables will minimize future costs associated with resource expansion once data are more readily available. Finally, integrating broad military family participation in CPRS and across the military health system will provide even more data points to maximize military family readiness and well-being.

Military families need personalized predictions about prognosis and response to treatments and a deeper understanding of the complex factors and their interactions that influence health at the level of the patient, the health system, and society. Big data initiatives allow for cluster-level queries with multi-level stratifications when evidence does not exist for a decision. For example, analysis of data with learning-enhanced approaches may be used to detect mental health issues, suicide and family well-being, and safety problems with drugs and devices, and ultimately lead to more effective methods of comparing prevention, diagnostic, and treatment options.

*D. Threats*

Although the promise of big data is enormous, the benefits gained through a big data learning health system are also replete with potential threats. Given the growth in health data assets, data collection and analysis involve user trust, transparency, and appropriate privacy settings. Any big data learning health system must strive to create safeguards that protect classified military information and have the highest levels of protection for military families. Current military regulations may need to be re-examined to maximize the benefit from research while simultaneously maintaining operational security and minimizing risk to service members and their families. At the same time, as military hospital administrators and policymakers may be ill equipped to make policy decisions based on ambiguous big data results, conclusions and policy changes drawn from big data analytics should not be made without consulting the advisory board, as discussed above.

Although a large consumer user base has made seamless transitions into data collection for their personal and work lives (eg, cloud-based sharing of files, contacts, photos, etc.), fewer consumers are comfortable sharing personalized healthcare data and information. Consumer participations and enthusiasm to share information is the lifeblood of any big data analytics program. Without data points, the system is left unprepared to predict real-time solutions; even with the appropriate data points, selective interpretation of data could lead to significant downstream effects that include public mistrust and program scrutiny.[24]

In the wake of the several data scandals,[25-26] data use agreements must involve participants in decision-making processes, set clear standards for ethical rigor, and specify sanctions for data misuse and abuse. The nefariousness of some individuals and organizations could create an enormous breach of trust for military families and erode the trust of active military. Adding to the complexity, available data are currently siloed and housed in a number of public and private systems. Data use agreements must therefore be worked out with each system and private sector data resources will have to agree to share data points with the military in order to create more powerful harmonization.

**IV. Recommendations**

First and foremost, big data is a tool that works to enhance the analytic framework. Big data has no intrinsic value without a learning health system that supports continuous PDSA feedback. A core reason that investments in big data fail to add value, though, is that most enterprises are not effective users of the data they already have. Organizations should be fluent in their ability to manage data, analyze it in ways that enhance their understanding, and then make changes in response to new insights. In addition, investment in high-end analytics tools in and of itself does not develop the core competencies necessary to establish a learning health system. Systems and organizations that have the dexterity to use data and analysis in support of operating decisions will be able to benefit most from big data.

We underscore the following four key recommendations:

1) Plan: Capture and curate real-time data securely and harmonize multi-layered data sources. These sources include structural level data from multiple EHR vendors, population health data, and clinical and healthcare facility data.
2) Do: Cultivate engaged leadership and empower military families in the data being captured with data use agreements that emphasize enhanced data security and facilitate continuous feedback.
3) Study: Confirm that the harmonized data is meaningfully aligned with values in military families by creating a fully transparent, data-safe system that avoids wasting resources and inaccurate predictions supporting poor decisions.
4) Act: Create an iterative learning health system that is guided by a leadership-instilled culture of rigorous continuous review of machine learning algorithms. This system, driven by a multidisciplinary advisory group composed of clinical providers, computer scientists, and health services researchers, must critically evaluate policies and preserve data safety and public health for military families' readiness and well-being.

## V. References


1. Moore GE. Electronics. 1965;38:114.

2. Institute of Medicine (IOM). Crossing the Quality Chasm: A New Health System for the Twenty-first Century. Washington: National Academy Press, 2001.

3. Big Data for Development: Challenges and Opportunities. UN Global Pulse, 2012. Available from: http://www.unglobalpulse.org/projects/BigDataforDevelopment. Accessed August 13, 2018.

4. Wyber R, Vaillancour S, Perry W, Mannava P, Folaranmi T, Celi LA. Big data in global health: improving health in low-and middle-income countries. Bulletin of the World Health Organization. 2015;93:203-208.

5. Donabedian A. The quality of care: how can it be assessed? JAMA. 1988;260:1743-1748.

6. Hashem IAT, Yaqooba I, Anuara NB, Mokhtara S, Gania A, Khan SU. The rise of “big data” on cloud computing: review and open research issues. Information Systems. 2015;47:98-115.

7. Institute of Medicine. Digital infrastructure for the learning health system: the foundation for continuous improvement in health and health care. In: Grossman C Powers B, McGinnis JM, eds. The Foundation for Continuous Improvement in Health and Health Care Workshop Series Summary. National Academies Press, 2011:1-311.

8. Institute of Medicine. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Washington, DC: The National Academies Press, 2013. Available from: https://doi.org/10.17226/13444. Accessed August 13, 2018.

9. Agency for Healthcare Research and Quality. Plan-Do-Study-Act (PDSA) Cycle. Available from: https://innovations.ahrq.gov/qualitytools/plan-do-study-act-pdsa-cycle. Accessed July 15, 2018.

10. Booz Allen. Report: Predictive Analytics Handbook for National Defense. Available from: https://www.boozallen.com/d/insight/thought-leadership/predictive-analytics-handbook-for-national-defense.html. Accessed August 12, 2018.

11. Military Health System. Final Report to the Secretary of Defense: Military Health System Review, 2014. Available from: http://archive.defense.gov/pubs/140930_MHS_Review_Final_Report_Main_Body.pdf. Accessed August 1, 2018.

12. National Academies of Sciences Engineering and Medicine. Evaluation of the Department of Veterans Affairs Mental Health Services [Internet]. Washington, DC: National Academies Press, 2018. Available from: https://www.nap.edu/catalog/24915. Accessed August 13, 2018.

13. National Academies of Sciences Engineering and Medicine. Preventing Psychological Disorders in Service Members and Their Families [Internet]. Washington, D.C.: National Academies Press, 2014. Available from: http://www.nap.edu/catalog/18597. Accessed August 13, 2018.

14. National Academies of Sciences, Engineering, and Medicine. Federal Statistics, Multiple Data Sources, and Privacy Protection: Next Steps. Washington, DC: The National Academies Press, 2017. Available from: https://doi.org/10.17226/24893. Accessed August 13, 2018.

15. U.S. Department of Health and Human Services. Agency for Healthcare Research and Quality. National Quality Measures Clearinghouse. Domain Framework and Inclusion Criteria. Available from: ftp://ftp.ihe.net/Quality/2011_2012_YR_5/QRPH%20Technical%20Committee/AHRQ%20National%20Quality%20Measures%20Clearinghouse%20_.pdf. Accessed August 1, 2018.

16. History of Information Technology at Veterans Affairs. Available from: https://www.oit.va.gov/about/history.cfm. Accessed August 1, 2018.

17. Agency for Healthcare Research and Quality; Jha AK. What Can the Rest of the Health Care System Learn from the VA's Quality and Safety Transformation? Available from: https://psnet.ahrq.gov/perspectives/perspective/31/what-can-the-rest-of-the-health-care-system-learn-from-the-vas-quality-and-safety-transformation. Accessed August 1, 2018.

18. Dabek F, Caban JJ. Leveraging big data to model the likelihood of developing psychological conditions after a concussion. Procedia Computer Science. 2015;53:265-273.

19. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Information Science and Systems. 2014;2(1):3.

20. Coffron MR, Opelka F. Big promise and big challenges for big health care data. Bulletin of the American College of Surgeons. April 1, 2015.

21. Alphabet spinoff Cityblock raises $20 million to help low-income Americans get health care. Available from: https://www.cnbc.com/2018/01/04/alphabet-spin-off-cityblock-raises-20m-for-low-income-health-care.html. Accessed August 1, 2018.

22. VA enlists IBM’s Watson Health technology in the war on cancer. Available from: https://www.blogs.va.gov/VAntage/28796/va-enlists-watson-health-technology-war-cancer/. Accessed August 5, 2018.

23. IBM Has a Watson Dilemma. Available from: https://www.wsj.com/articles/ibm-bet-billions-that-watson-could-improve-cancer-treatment-it-hasnt-worked-1533961147. Accessed August 12, 2018.

24. Release VA nursing homes data. USA Today. Available at: https://www.usatoday.com/story/opinion/2018/06/20/release-va-nursing-homes-data-editorials-debates/717381002/. Accessed July 31, 2018.

25. Facebook and Cambridge Analytica: What you need to know as fallout widens. New York Times. Available from: https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html. Accessed June 1, 2018

26. Target says up to 70 million more customers were hit by December data breach. Available from: https://www.washingtonpost.com/business/economy/target-says-70-million-customers-were-hit-by-dec-data-breach-more-than-first-reported/2014/01/10/0ada1026-79fe-11e3-8963-b4b654bcc9b2_story.html. Accessed June 1, 2018

**Acknowledgments**

The author thanks Shira Berman and Caitlin Bakker, MLIS, for editorial assistance in preparing the paper.