

# Does Machine Learning Improve Prediction of VA Primary Care Reliance?

Edwin S. Wong, PhD; Linnaea Schuttner, MD, MS; and Ashok Reddy, MD, MSc

**T**he Veterans Affairs (VA) Health Care System is among the largest integrated healthcare systems in the United States. The VA provides comprehensive health services to eligible veterans, and most of the 8.3 million VA enrollees in 2017 were exempt from co-payments or other cost sharing.<sup>1</sup> VA enrollees are unique because, in addition to access to VA health services, more than 80% are dually enrolled in at least 1 other source of private or public insurance. Although most users of VA are exempt from cost sharing, many veterans 65 years or older supplement VA care with Medicare services. Extensive research has examined factors associated with greater dual use of VA and Medicare outpatient services. Dual VA-Medicare enrollees who were more reliant on VA outpatient care were younger and unmarried, lived closer to VA facilities, and were exempt from VA co-payments.<sup>2-4</sup>

Previous research examining determinants of VA utilization and dual use has largely been descriptive and has not explored approaches to generate accurate predictions. More recent research outside VA has applied newer machine learning algorithms, demonstrating promise in producing predictions of future health outcomes, utilization, and costs with greater accuracy compared with more traditional approaches.<sup>5</sup> Notably, several studies have examined prediction of costs within the context of improving risk adjustment, finding that supervised learning methods outperform such traditional methods as parametric regression.<sup>6-8</sup> Machine learning models are being increasingly applied within VA but have not been routinely used for the purpose of predicting reliance on VA health services. These models often have practical challenges, including greater computational complexity and lack of interpretability relative to traditional statistical models. Thus, it is important to understand trade-offs between potential improvements in model performance and these practical challenges.

Improving the ability to correctly predict patients' expected utilization and costs is important for at least 2 reasons. First, more accurate predictions can help determine the adequacy of budget and staffing resources needed to provide care for a patient population. This is particularly important for the VA, which is funded annually through a budget appropriation from Congress. This appropriation is determined, in part, from the Enrollee Health Care Projection Model,

## ABSTRACT

**OBJECTIVES:** The Veterans Affairs (VA) Health Care System is among the largest integrated health systems in the United States. Many VA enrollees are dual users of Medicare, and little research has examined methods to most accurately predict which veterans will be mostly reliant on VA services in the future. This study examined whether machine learning methods can better predict future reliance on VA primary care compared with traditional statistical methods.

**STUDY DESIGN:** Observational study of 83,143 VA patients dually enrolled in fee-for-service Medicare using VA and Medicare administrative databases and the 2012 Survey of Healthcare Experiences of Patients.

**METHODS:** The primary outcome was a dichotomous measure denoting whether patients obtained more than 50% of all primary care visits (VA + Medicare) from VA. We compared the performance of 6 candidate models—logistic regression, elastic net regression, decision trees, random forest, gradient boosting machine, and neural network—in predicting 2013 reliance as a function of 61 patient characteristics observed in 2012. We measured performance using the cross-validated area under the receiver operating characteristic (AUROC) metric.

**RESULTS:** Overall, 72.9% and 74.5% of veterans were mostly VA reliant in 2012 and 2013, respectively. All models had similar average AUROCs, ranging from 0.873 to 0.892. The best-performing model used gradient boosting machine, which exhibited modestly higher AUROC and similar variance compared with standard logistic regression.

**CONCLUSIONS:** The modest gains in performance from the best-performing model, gradient boosting machine, are unlikely to outweigh inherent drawbacks, including computational complexity and limited interpretability compared with traditional logistic regression.

*Am J Manag Care.* 2020;26(1):40-44

which includes veterans' expected reliance on the VA for care as a key input.<sup>9,10</sup> At the individual clinic level, predictions of expected VA primary care reliance are helpful in determining staffing distributions, including ancillary services such as social work and integrated behavioral health that are needed by more frequent users of VA.<sup>11</sup>

More accurate predictions of VA reliance can also help better identify veterans expected to use outpatient services outside the VA. This is significant in light of VA's current emphasis on competing for veterans as customers given increasing healthcare options available to VA enrollees.<sup>12</sup> In addition, accurate reliance predictions may inform care coordination efforts within VA's patient-centered medical home.<sup>13</sup> Receipt of fragmented care across health systems is often associated with poor clinical outcomes and increased hospitalizations.<sup>14-17</sup> As current policies continue to support veterans' access to care outside the VA, accurate predictions of reliance will help identify patients who need further support to mediate the risks of care fragmentation. To address these policy-planning priorities, we examined whether supervised machine learning methods could improve the ability to predict future reliance on VA primary care among VA-Medicare dual enrollees.

## METHODS

### Data Sources

The primary data source was the VA Corporate Data Warehouse (CDW), which houses information on utilization of VA health services, demographic data, and *International Classification of Diseases, Ninth Revision* diagnosis codes.<sup>18</sup> We linked CDW data with Medicare enrollment and claims data to ascertain enrollment in fee-for-service Medicare and utilization. We supplemented VA and Medicare administrative data with survey responses capturing patient experiences with VA outpatient care, as measured using the 2012 VA Survey of Healthcare Experiences of Patients (SHEP) Outpatient Module, which is adapted from the Consumer Assessment of Healthcare Providers and Systems.<sup>19</sup> SHEP is managed by the VA Office of Reporting, Analytics, Performance, Improvement and Deployment, which routinely administers the survey to a random sample of veterans who obtained outpatient care at VA outpatient facilities. Finally, we obtained measures of local area health resources in veterans' residence counties, as reported in the Area Health Resource File.<sup>20</sup>

### Study Sample

Using data from the 2012 outpatient SHEP, we identified 222,072 respondents enrolled in VA, of whom 90,819 were also continuously enrolled in fee-for-service Medicare in fiscal year (FY) 2012 and FY 2013. Consistent with prior studies,<sup>2-4</sup> we examined veterans dually enrolled in VA and fee-for-service Medicare because we were unable to capture utilization from other non-VA sources. We then excluded 7182 veterans with no primary care utilization in FY

## TAKEAWAY POINTS

Reliance on the Veterans Affairs Health Care System for outpatient services is a key metric for informing resource allocation and current operational priorities, including efforts to compete for veteran demand amid increasing healthcare options.

- ▶ The application of each of 6 candidate models for predicting future reliance exhibited very good performance based on previous thresholds.
- ▶ Each of the 5 machine learning models considered did not outperform, or modestly outperformed, a traditional regression model.
- ▶ Modest gains in performance of machine learning models are unlikely to outweigh practical limitations including computational complexity and difficulty in interpreting model composition.

2012 or FY 2013 whose reliance, by definition, was undefined. After excluding 494 veterans with missing covariates, the final study sample consisted of 83,143 dual VA-fee-for-service Medicare enrollees.

### Outcome Measures

Our primary outcome was veterans' reliance on VA primary care in FY 2013. We measured reliance as the number of primary care visits received in VA divided by total primary care visits (VA + Medicare). We identified primary care visits in VA and Medicare administrative data through the identification of records in which the provider of care was classified in a specialty considered primary care, including general internal medicine, family medicine, and geriatrics.<sup>21</sup> We further required records to contain at least 1 primary care evaluation and management Current Procedural Terminology code.<sup>22</sup> We dichotomized the continuous reliance measure to indicate whether veterans were mostly reliant on VA primary care (ie, VA reliance >0.5). The rationale for this binary definition is that prior research found that most VA enrollees were completely reliant on 1 system.<sup>3,23</sup>

### Predictor Variables

We predicted reliance on VA primary care as a function of 61 variables measured in FY 2012, the 12 months prior to the assessment of reliance. The predictor variables encompassed 5 broad categories routinely available to researchers examining VA utilization using administrative data and informed by prior studies examining determinants of VA utilization: patient demographics, access to VA care, physical and mental comorbidities, characteristics of veterans' residence areas, and experiences with VA care (**Table 1**).<sup>2-4</sup> Access to care included several administrative measures, as conceptualized in prior research.<sup>24</sup> Comorbidities included 20 indicator variables denoting the presence of conditions in FY 2012, as specified by the validated Gagne et al index.<sup>25</sup> Patient experience variables included 23 measures from survey data, elicited on a 4-point Likert scale. These variables included measures capturing perceptions of the availability of VA outpatient care, assessment of VA facilities, and satisfaction with providers and overall VA care. All patient experience measures were defined as binary variables denoting whether an experience dimension was rated in the top 2 categories (eg, always or usually able to obtain immediate care from VA when needed compared with sometimes or never).

## METHODS

**TABLE 1.** Selected Descriptive Statistics Stratified by Reliance Classification

	Overall (N = 83,143)	Medicare Reliant (n = 25,803)	VA Reliant (n = 57,340)
Age in years, mean (SD)	70.7 (11.1)	75.3 (9.8)	69.0 (11.1)
Male, %	96.6	97.7	96.2
Race/ethnicity, %			
Non-Hispanic white	78.3	84.2	76.1
Non-Hispanic black	11.5	7.3	13.1
Hispanic	4.4	3.3	4.7
Other/unknown	5.8	5.2	6.1
Married, %	65.7	78.5	61.0
High school education or greater, %	87.1	88.0	86.8
Selected comorbidities, %			
Alcohol abuse	5.2	2.1	6.4
Chronic obstructive pulmonary disease	17.2	13.9	18.4
Complicated diabetes	10.0	7.2	11.0
Heart failure	7.6	5.7	8.3
Hypertension	70.6	65.0	72.6
Liver disease	1.9	0.9	2.3
Psychosis	12.9	7.2	15.0
Exempt from VA co-payments, %	76.3	85.6	72.9
Resides <20 miles from nearest VA, %	79.6	82.7	78.6
Urban, %	63.1	66.7	61.8
County-level hospital beds/1000, mean (SD)	3.22 (2.38)	3.16 (2.25)	3.24 (2.43)
County-level PCPs/1000, mean (SD)	0.70 (0.28)	0.71 (0.28)	0.70 (0.29)
Mostly VA reliant in prior year, %	74.5	24.5	93.1
Perceptions of access to VA, %			
Immediate care	33.5	20.6	38.3
Routine care	58.9	61.3	58.1
Specialists	46.6	36.4	50.4
Overall VA rating, %	58.5	56.7	59.2

PCP indicates primary care provider; VA, Veterans Affairs.

### Assessing Candidate Machine Learning Algorithms

We compared the performance of several candidate supervised machine learning algorithms in classifying veterans' future reliance on VA primary care. First, we applied standard logistic regression, which is a traditional approach using parametric regression to model binary outcomes. Next, we applied elastic net regression, which is a regularization approach that combines penalty terms for (1) shrinking regression coefficients to 0, as done in LASSO regression, and (2) shrinking coefficients toward 0, as done in Ridge regression. Third, we considered decision trees, which repeatedly partition the predictor space into subspaces producing the biggest performance improvement. Fourth, we applied random forests, which extend decision trees, by constructing multiple regression trees using bootstrapped training sets, splitting over a subset of predictors, and averaging predictions across trees. Fifth, we examined a gradient boosting machine, which sequentially constructs a regression tree using information from previously grown trees and generates an overall prediction using a weighted average of

predictions across trees. Finally, we applied an artificial neural network, which involves constructing a directed graph composed of a network of nodes connected using weights that are updated using predefined rules.

To estimate the performance of the 6 algorithms, we conducted 10-fold cross-validation repeated 3 times, which is a common approach to protect against model overfit.<sup>26</sup> We prespecified a grid of tuning parameters specific to the algorithm that controlled the configuration of a model (eg, maximum number of splits for tree-based models). For each combination of tuning parameters, we estimated model parameters and calculated performance metrics for each of the 30 training samples. For each algorithm, we identified the set of tuning parameters with the highest mean performance metric. We compared model performance using the area under the receiver operating characteristic curve (AUROC), sensitivity, and specificity.<sup>27-30</sup> All models applied sampling weights to account for nonresponse of veterans who were offered the 2012 SHEP. We conducted statistical analyses using RStudio version 0.99.891 (RStudio Inc; Boston, Massachusetts). Implementation of logistic regression, elastic net regression, decision trees, random forests, gradient boosting machine, and neural networks applied the *glm*,<sup>31</sup> *glmnet*,<sup>32</sup> *rpart*,<sup>33</sup> *randomForest*,<sup>34</sup> *gbm*,<sup>35</sup> and *nnet*<sup>36</sup> packages, respectively. Assessment of candidate models was facilitated using the *caret* package.<sup>37</sup>

## RESULTS

### Descriptive Statistics

The characteristics of dual VA-Medicare enrollees were similar to those found in prior studies.<sup>2-4</sup> Specifically, the mean age was 70.7 years, and the majority of patients were male (96.6%), white (78.3%), married (65.7%), and exempt from co-payments for VA health services (76.3%) (Table 1). Compared with patients mostly reliant on VA care (ie, receiving >50% of primary care from the VA in FY 2012), Medicare-reliant patients were older and more likely to be white, married, or reside in urban areas. Medicare-reliant patients were also less likely to be reliant on VA services in the prior year, 2011. Differences between groups were observed for several physical and mental comorbidities, with Medicare-reliant patients having a lower prevalence of heart failure, chronic obstructive pulmonary disease, and psychosis. Also, Medicare-reliant patients reported substantially worse experiences with VA services, including a lower proportion perceiving an ability to access immediate care or specialty care in a timely manner.

## Prediction Model Performance

The mean (SD) AUROC estimated using logistic regression across 30 validation samples was 0.890 (0.004) (Table 2). Regularized regression using elastic net produced similar predictive performance but exhibited greater variance (AUROC, 0.891; SD = 0.005). Of the 3 tree-based algorithms, decision trees performed the worst (AUROC, 0.873; SD = 0.005), followed by random forest (AUROC, 0.889; SD = 0.003) and gradient boosting machine (AUROC, 0.892; SD = 0.004). The latter exhibited the best performance of all algorithms considered.

Performance was also similar across the 6 models using other evaluation metrics. Mean (SD) specificity ranged from 0.724 (0.007) using decision trees to 0.759 (0.008) for gradient boosting machine. Specificity was modestly lower for logistic regression (mean [SD] = 0.752 [0.007]) compared with the best-performing model. Mean (SD) sensitivity ranged from 0.916 (0.003) for decision tree to 0.922 (0.002) for random forest. Similarly, sensitivity for logistic regression (mean [SD] = 0.920 [0.002]) was only modestly lower compared with the best-performing machine learning model.

## DISCUSSION

Prior study findings have established that dual use of VA and non-VA services is common among VA enrollees. Most studies, however, are solely descriptive and do not have the objective of accurate prediction of future VA reliance. In this study, we explored whether several candidate machine learning models could better predict which patients would be highly reliant on VA primary care services compared with logistic regression, which is commonly used for prediction in health services research. The predictive model using logistic regression exhibited very good performance, as defined using AUROC thresholds from a prior study. However, average performance, as measured by AUROC, was similar across all 6 models considered and varied by less than 2 percentage points. Gradient boosting machine exhibited the best performance of all 6 models considered. However, the predictive performance was less than 1 percentage point better and exhibited modestly greater variance compared with logistic regression. Similarities in model performance were observed for 2 alternative metrics, specificity and sensitivity, with the best-performing models only modestly outperforming logistic regression.

Evidence on how to more accurately predict VA reliance is important for policy planning because the VA is publicly funded through an annual Congressional appropriation, which is informed by budget projection models that include veterans' expected reliance on VA healthcare as an input. Research in this study is also consistent with the explicit operational goal of focusing resources efficiently within the VA FY 2018-2024 Strategic Plan.<sup>12</sup> Specifically, future VA efforts will use VA funding requests to "eliminate spending without sacrificing the outcomes for veterans." Clinically, reliance prediction models have utility by identifying veterans who are at risk of receiving fragmented care from multiple health systems. For these veterans, creating services to improve care coordination can avoid

**TABLE 2.** Comparison of Performance of Reliance Classification Models

	Mean	Interquartile Range
<b>AUROC</b>		
Logistic regression	0.891	0.889-0.893
Elastic net	0.891	0.889-0.894
Decision tree	0.873	0.871-0.876
Random forest	0.889	0.888-0.891
Gradient boosting machine	0.892	0.889-0.895
Neural network	0.885	0.883-0.887
<b>Specificity</b>		
Logistic regression	0.752	0.748-0.758
Elastic net	0.753	0.748-0.756
Decision tree	0.721	0.713-0.728
Random forest	0.742	0.737-0.747
Gradient boosting machine	0.759	0.752-0.764
Neural network	0.746	0.750-0.760
<b>Sensitivity</b>		
Logistic regression	0.920	0.918-0.923
Elastic net	0.920	0.919-0.922
Decision tree	0.913	0.910-0.916
Random forest	0.922	0.920-0.924
Gradient boosting machine	0.918	0.916-0.920
Neural network	0.919	0.916-0.920

AUROC indicates area under receiver operating characteristic.

duplication of services, reduce waste, and improve patient safety by mitigating the loss of clinical information needed by providers.<sup>38-40</sup>

Stakeholders determining the best model to apply when predicting future VA reliance should consider a number of factors other than overall performance. These considerations include the model assumptions, computational complexity, variance in model performance, and intended use. Notably, more complex models such as random forest and gradient boosting machine tend to exhibit greater performance, on average, but may overfit data, resulting in higher variance and less reliable predictions when applying models in the real world. In addition, more complex models are generally more computationally intensive, require more time to train, and may be less desirable if stakeholders require that models be frequently updated to incorporate newly collected data. If interpretability is an important feature, more complex models may be less desirable despite better overall performance because model parameters that reflect effect of individual predictors, analogous to regression coefficients, are often not available. For the prediction of future VA primary care reliance using the 61 variables considered in this study, the practical limitations of more complex machine learning models likely outweigh the modest gains in predictive performance relative to traditional logistic regression, which achieved very good performance. Because model performance varies with the specific set of predictor variables chosen, the optimal prediction algorithm may differ for an alternative set of data.

## METHODS

### Limitations

This study has a number of limitations. First, our study sample consisted of a sample of VA-Medicare dual enrollees who responded to a routinely administered VA patient experience survey to capture a wider set of predictor variables not available in VA administrative data. Because these patients were recent users of the VA, and veterans remain enrolled until death regardless of utilization, the study sample may not generalize to the population of all VA enrollees. Second, we were able to measure utilization of primary care services from VA and Medicare using a unique data-sharing agreement available to VA researchers. However, we were unable to capture utilization of primary care services through other payers. Third, data used in this study measured VA primary care reliance prior to the expansion of community care options in 2014.<sup>41</sup> To date, the uptake of these community care options for primary care services among VA-Medicare dual enrollees is unknown.

## CONCLUSIONS

Knowledge of reliance on VA services is important for policy planning and care delivery, but there has been limited research on improving prediction of this important outcome. We examined several candidate methods for predicting reliance as a function of a comprehensive set of patient variables. We found that more complex machine learning methods exhibited modest improvements in predictive performance compared with traditional logistic regression. These gains in performance are unlikely to outweigh other practical limitations of more complex machine learning models that are more computationally intensive and difficult to interpret. ■

**Author Affiliations:** Center for Veteran-Centered and Value-Driven Care, VA Puget Sound Health Care System (ESW, LS, AR), Seattle, WA; Department of Health Services (ESW) and Division of General Internal Medicine, Department of Medicine (LS, AR), University of Washington, Seattle, WA.

**Source of Funding:** Dr Wong was supported by a Veterans Affairs Health Services Research and Development Career Development Award (CDA 13-024). The views expressed are those of the authors and do not necessarily reflect the views of the Department of Veterans Affairs and the University of Washington.

**Author Disclosures:** The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

**Authorship Information:** Concept and design (ESW, AR); acquisition of data (ESW); analysis and interpretation of data (ESW, LS, AR); drafting of the manuscript (ESW, LS, AR); critical revision of the manuscript for important intellectual content (LS, AR); statistical analysis (ESW); and obtaining funding (ESW).

**Address Correspondence to:** Edwin S. Wong, PhD, Center for Veteran-Centered and Value-Driven Care, VA Puget Sound Health Care System, 1660 S Columbian Way, HSR&D MS-152, Seattle, WA 98108. Email: edwin.wong@va.gov.

## REFERENCES

- Huang G, Kim S, Muz B, Gasper J. 2017 survey of veteran enrollees' health and use of health care. US Department of Veterans Affairs website. [va.gov/HEALTHPOLICYPLANNING/SOE2017/VA\\_Enrollees\\_Report\\_Data\\_Findings\\_Report2.pdf](http://va.gov/HEALTHPOLICYPLANNING/SOE2017/VA_Enrollees_Report_Data_Findings_Report2.pdf). Published April 2018. Accessed June 1, 2019.
- Hynes DM, Koelling K, Stoupe K, et al. Veterans' access to and use of Medicare and Veterans Affairs health care. *Med Care*. 2007;45(3):214-223. doi: 10.1097/01.mlr.0000244657.90074.b7.
- Liu CF, Manning WG, Burgess JF Jr, et al. Reliance on Veterans Affairs outpatient care by Medicare-eligible veterans. *Med Care*. 2011;49(10):911-917. doi: 10.1097/MLR.0b013e31822396c5.
- Petersen LA, Byrne MM, Daw CN, Hasche J, Reis B, Pietz K. Relationship between clinical conditions and use of Veterans Affairs health care among Medicare-enrolled veterans. *Health Serv Res*. 2010;45(3):762-791. doi: 10.1111/j.1475-6773.2010.01107.x.
- Morid MA, Kawamoto K, Ault T, Dorius J, Abdelrahman S. Supervised learning methods for predicting health-care costs: systematic literature review and empirical evaluation. *AMIA Annu Symp Proc*. 2018;2017:1312-1321.

- Shrestha A, Bergquist S, Montz E, Rose S. Mental health risk adjustment with clinical categories and machine learning. *Health Serv Res*. 2018;53(suppl 1):3189-3206. doi: 10.1111/1475-6773.12818.
- Park S, Basu A. Alternative evaluation metrics for risk adjustment methods. *Health Econ*. 2018;27(6):984-1010. doi: 10.1002/hec.3657.
- Rose S. A machine learning framework for plan payment risk adjustment. *Health Serv Res*. 2016;51(6):2358-2374. doi: 10.1111/1475-6773.12464.
- Harris KM, Galasso JP, Eibner C. Review and evaluation of the VA Enrollee Health Care Projection Model. RAND Corporation website. [rand.org/pubs/monographs/MG596.html](http://rand.org/pubs/monographs/MG596.html). Published 2008. Accessed June 1, 2019.
- VA uses a projection model to develop most of its health care budget estimate to inform the president's budget request. US Government Accountability Office website. [gao.gov/assets/320/315324.pdf](http://gao.gov/assets/320/315324.pdf). Published January 2011. Accessed June 1, 2019.
- Hebert PL, Batten AS, Gunnink E, et al. Reliance on Medicare providers by veterans after becoming age-eligible for Medicare is associated with the use of more outpatient services. *Health Serv Res*. 2018;53(suppl 3):5159-5180. doi: 10.1111/1475-6773.13033.
- Department of Veterans Affairs FY 2018-2024 strategic plan. US Department of Veterans Affairs website. [va.gov/oei/docs/VA2018-2024strategicPlan.pdf](http://va.gov/oei/docs/VA2018-2024strategicPlan.pdf). Updated May 31, 2019. Accessed June 1, 2019.
- Nelson KM, Helfrich C, Sun H, et al. Implementation of the patient-centered medical home in the Veterans Health Administration: associations with patient satisfaction, quality of care, staff burnout, and hospital and emergency department use. *JAMA Intern Med*. 2014;174(8):1350-1358. doi: 10.1001/jamainternmed.2014.2488.
- Kizer KW. Veterans and the Affordable Care Act. *JAMA*. 2012;307(8):789-790. doi: 10.1001/jama.2012.1196.
- Gellad WF. The Veterans Choice Act and dual health system use. *J Gen Intern Med*. 2016;31(2):153-154. doi: 10.1007/s11606-015-3492-2.
- Wolinsky FD, An H, Liu L, Miller TR, Rosenthal GE. Exploring the association of dual use of the VHA and Medicare with mortality: separating the contributions of inpatient and outpatient services. *BMC Health Serv Res*. 2007;7:70. doi: 10.1186/1472-6963-7-70.
- West AN, Charlton ME, Vaughan-Sarrazin M. Dual use of VA and non-VA hospitals by veterans with multiple hospitalizations. *BMC Health Serv Res*. 2015;15:431. doi: 10.1186/s12913-015-1069-8.
- Corporate Data Warehouse (CDW). Data.gov website. [catalog.data.gov/dataset/corporate-data-warehouse-cdw](http://catalog.data.gov/dataset/corporate-data-warehouse-cdw). Updated July 2, 2019. Accessed December 5, 2019.
- Wright SM, Craig T, Campbell S, Schaefer J, Humble C. Patient satisfaction of female and male users of Veterans Health Administration services. *J Gen Intern Med*. 2006;21(suppl 3):S26-S32. doi: 10.1111/j.1525-1497.2006.00371.x.
- Area Health Resources Files. Health Resources & Services Administration website. [data.hrsa.gov/topics/health-workforce/ahrif](http://data.hrsa.gov/topics/health-workforce/ahrif). Updated July 31, 2019. Accessed December 5, 2019.
- Liu CF, Batten A, Wong ES, Fihn SD, Hebert PL. Fee-for-service Medicare-enrolled elderly veterans are increasingly voting with their feet to use more VA and less Medicare, 2003-2014. *Health Serv Res*. 2018;53(suppl 3):5140-5158. doi: 10.1111/1475-6773.13029.
- Burgess JF Jr, Maciejewski ML, Bryson CL, et al. Importance of health system context for evaluating utilization patterns across systems. *Health Econ*. 2011;20(2):239-251. doi: 10.1002/hec.1588.
- Liu CF, Burgess JF Jr, Manning WG, Maciejewski ML. Beta-binomial regression and bimodal utilization. *Health Serv Res*. 2013;48(5):1769-1778. doi: 10.1111/1475-6773.12055.
- Fortney JC, Burgess JF Jr, Bosworth HB, Booth BM, Kaboli PJ. A re-conceptualization of access for 21st century healthcare. *J Gen Intern Med*. 2011;26(suppl 2):639-647. doi: 10.1007/s11606-011-1806-6.
- Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol*. 2011;64(7):749-759. doi: 10.1016/j.jclinepi.2010.10.004.
- Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J Mach Learn Res*. 2010;11:2079-2107.
- Vihinen M. How to evaluate performance of prediction methods? measures and their interpretation in variation effect analysis. *BMC Genomics*. 2012;13(suppl 4):S2. doi: 10.1186/1471-2164-13-S4-S2.
- McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-282. doi: 10.11613/BM.2012.031.
- Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*. 2007;115(5):654-657. doi: 10.1161/CIRCULATIONAHA.105.594929.
- Burman P. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing method. *Biometrika*. 1989;76(3):503-514. doi: 10.2307/2336116.
- Fitting generalized linear models. ETH Zürich website. [stat.ethz.ch/R-manual/R-devel/library/stats/html/glm.html](http://stat.ethz.ch/R-manual/R-devel/library/stats/html/glm.html). Accessed June 1, 2019.
- Friedman J, Hastie T, Tibshirani R, Simon N, Narasimhan B, Qian J. Package 'glmnet'. Comprehensive R Archive Network website. [cran.r-project.org/web/packages/glmnet/glmnet.pdf](http://cran.r-project.org/web/packages/glmnet/glmnet.pdf). Published November 15, 2019. Accessed December 5, 2019.
- Therneau T, Atkinson B, Ripley B. Package 'rpart'. Comprehensive R Archive Network website. [cran.r-project.org/web/packages/rpart/rpart.pdf](http://cran.r-project.org/web/packages/rpart/rpart.pdf). Published April 12, 2019. Accessed June 1, 2019.
- Liaw A, Wiener M. Package 'randomForest'. Comprehensive R Archive Network website. [cran.r-project.org/web/packages/randomForest/randomForest.pdf](http://cran.r-project.org/web/packages/randomForest/randomForest.pdf). Published March 25, 2018. Accessed June 1, 2019.
- Greenwell B, Boehmke B, Cunningham J; GBM Developers. Package 'gbm'. Comprehensive R Archive Network website. [cran.r-project.org/web/packages/gbm/gbm.pdf](http://cran.r-project.org/web/packages/gbm/gbm.pdf). Published January 14, 2019. Accessed June 1, 2019.
- Ripley B, Venables W. Package 'nnet'. Comprehensive R Archive Network website. [cran.r-project.org/web/packages/nnet/nnet.pdf](http://cran.r-project.org/web/packages/nnet/nnet.pdf). Published February 2, 2016. Accessed June 1, 2019.
- Kuhn M. Package 'caret'. Comprehensive R Archive Network website. [cran.r-project.org/web/packages/caret/caret.pdf](http://cran.r-project.org/web/packages/caret/caret.pdf). Published April 27, 2019. Accessed June 1, 2019.
- Bodenheimer T. Coordinating care—a perilous journey through the health care system. *N Engl J Med*. 2008;358(10):1064-1071. doi: 10.1056/NEJMrp0706165.
- McWilliams JM. Cost containment and the tale of care coordination. *N Engl J Med*. 2016;375(23):2218-2220. doi: 10.1056/NEJMp160821.
- Hempstead K, Delia D, Cantor JC, Nguyen T, Brenner J. The fragmentation of hospital use among a cohort of high utilizers: implications for emerging care coordination strategies for patients with multiple chronic conditions. *Med Care*. 2014;52(suppl 3):S67-S74. doi: 10.1097/MLR.0000000000000049.
- Vanneman ME, Harris AHS, Asch SM, Scott WJ, Murrell SS, Wagner TH. Iraq and Afghanistan veterans' use of Veterans Health Administration and purchased care before and after Veterans Choice Program implementation. *Med Care*. 2017;55(suppl 1):S37-S44. doi: 10.1097/MLR.0000000000000678.

Visit [ajmc.com/link/4435](http://ajmc.com/link/4435) to download PDF