

Data-Driven Markov Decision Process Approximations for Personalized Hypertension Treatment Planning

Greggory J. Schell, PhD, Wesley J. Marrero, BS, Mariel S. Lavieri, PhD,
Jeremy B. Sussman, MD, MS, Rodney A. Hayward, MD

Background: Markov decision process (MDP) models are powerful tools. They enable the derivation of optimal treatment policies but may incur long computational times and generate decision rules that are challenging to interpret by physicians. **Methods:** In an effort to improve usability and interpretability, we examined whether Poisson regression can approximate optimal hypertension treatment policies derived by an MDP for maximizing a patient's expected discounted quality-adjusted life years. **Results:** We found that our Poisson approximation to the optimal treatment policy matched the optimal policy in 99% of cases. This high accuracy translates to nearly identical health outcomes for patients. Furthermore, the Poisson approximation results in 104

additional quality-adjusted life years per 1000 patients compared to the Seventh Joint National Committee's treatment guidelines for hypertension. The comparative health performance of the Poisson approximation was robust to the cardiovascular disease risk calculator used and calculator calibration error. **Limitations:** Our results are based on Markov chain modeling. **Conclusions:** Poisson model approximation for blood pressure treatment planning has high fidelity to optimal MDP treatment policies, which can improve usability and enhance transparency of more personalized treatment policies. **Key words:** decision rules; provider decision making; decision support techniques; Markov models. (*MDM Policy & Practice* 2016;1:1-9)

Received 28 January 2016 from the Center for Naval Analyses, Arlington, Virginia (GJS); Industrial and Operations Engineering (WJM, MSL) and Department of Internal Medicine (JBS, RAH), University of Michigan, Ann Arbor, Michigan; VA Ann Arbor Center for Clinical Management Research, Ann Arbor, Michigan (JBS, RAH). Presented in part at the 2015 Institute for Operations Research and the Management Sciences (INFORMS) Annual Meeting. Financial support for this study was provided in part (for JBS and RAH) by grants from the Department of Veterans Affairs HSR&D (IIR 06-253 and CDA 13-021) and the Michigan Center for Diabetes Translational Research (NIDDK of The National Institutes of Health [P60 DK-20572]). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Revision accepted for publication 3 August 2016.

The online supplemental appendix for this article are available on the *Medical Decision Making Policy & Practice* Web site at <http://mmp.sagepub.com/supplemental>.

Address correspondence to Greggory J. Schell, Center for Naval Analyses, 3003 Washington Blvd, Arlington, VA 22201, USA; telephone: 717-818-4898; e-mail: schellg@cna.org.

© The Author(s) 2016

Reprints and permission:

<http://www.sagepub.com/journalsPermissions.nav>

DOI: 10.1177/2381468316674214

Optimal medical decision making is challenged by uncertainty regarding both treatment effects and patient risks in the absence of treatment.^{1,2} Such uncertainties create even greater problems for situations in which decisions must be made sequentially over time, such as making medication recommendations for chronic conditions. Markov decision processes (MDPs) are well-established tools for aiding such sequential decision making.³⁻⁷ MDP models are frequently used in operations engineering and have potential applications in medicine, in which they can enable the decision maker to develop personalized treatment plans to optimize a patient's health outcome based on the best available information, with due consideration for parameter uncertainty. However, using MDPs can be difficult for real-world clinical practice. They often require specialized software, such as MATLAB or CPLEX,⁸ that are unavailable to most clinicians and hospitals. Using an MDP model often requires long computational times, especially when the medical problem studied has high dimensional state (many clinical variables) and action



This Creative Commons Non Commercial CC-BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 3.0 License (<http://www.creativecommons.org/licenses/by-nc/3.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

spaces (the possible decisions that could be made), which make look-up tables infeasible.⁹ Moreover, the decision rules obtained with an MDP may be complex and not easily interpretable. Together, these limitations also create a black box for clinicians, limiting acceptance and making communicating the results of the MDP to patients difficult. Therefore, a tool that harnesses the improved decision making capabilities of an MDP model (without the burden of specialized software, long computation time, and complexity) would prove useful in the management of patients with chronic disease requiring serial decision-making.

One such chronic disease is hypertension, or high blood pressure (BP), which is a major risk factor for cardiovascular disease (CVD; consisting of heart attack, angina, stroke, and other heart and blood vessel diseases), the leading cause of death in the United States and worldwide.^{10–12} We have created and previously reported on hypertension treatment policies using an MDP model with the goal of maximizing a patient's expected discounted quality-adjusted life years (QALYs), and we found that it saved significantly more QALYs than the Seventh Joint National Committee's (JNC7's) proposed treatment policy.¹³ This was accomplished by combining the power of MDP operations research modeling with a benefit-based tailored treatment strategy,^{14–16} which simultaneously considers all known factors that affect an individual's untreated risk and the treatment's relative effects on those risks.

In this article, we develop a novel framework that approximates the optimal hypertension dosage decision from the MDP model using Poisson regression and generalized linear mixed effect modeling. We then compare the performance and accuracy of our approximation to the original MDP model.

METHODS

All methods are described in more depth in the supplemental appendix.

Data Source and Study Design

We sampled 100,000 simulated patients from the Third National Health and Nutrition Examination Survey (NHANES III),¹⁷ a survey of U.S. patients, to generate a representative sample of the U.S. population. NHANES III was chosen because BP was treated much less aggressively in the early 1990s,

allowing for easier estimation of untreated BPs. Our data set included adults between the ages of 40 and 84 years that were examined over a 10-year horizon. The QALY MDP for CVD was solved for all 100,000 patients over a 10-year planning horizon (assuming a 0.001 disutility per prescribed medication) to obtain the MDP treatment policy for each patient. We also applied the current hypertension treatment guidelines in the United States, known as JNC7,¹⁸ to each patient in the sample.

Markov Decision Process Formulation

We modeled the process of sequentially determining hypertension treatment medications over a planning horizon as a discrete-time, finite horizon MDP formulation.¹⁹ The objective of the MDP was to determine the optimal treatment strategy π^* for a single patient that maximizes his/her expected discounted QALYs over the planning horizon, $t = 1, \dots, T$. We considered a 10-year planning horizon with annual treatment decisions. The MDP formulation was characterized by four features: state space, action space, state transition probabilities, and rewards. Table A1 summarizes the inputs and data sources for the model.

State Space

The MDP utilizes a state-space representation of a single patient to fully describe his/her characteristics at each time period $t = 1, \dots, T$ in the planning horizon of length T . The state s_t consists of demographic information, clinical observations, and the patient's health state. The demographic information includes the patient's age, sex, smoking status, and diabetes status. The clinical observations are measurements of the patient's untreated systolic blood pressure (SBP), high-density lipoprotein (HDL), total cholesterol (TC), and if the patient has left ventricular hypertrophy as determined by electrocardiogram. Last, there are 10 mutually exclusive patient health states: 1) healthy (no history of coronary heart disease (CHD) or stroke); 2) history of CHD but no CHD event this period; 3) history of stroke but no stroke this period; 4) history of CHD and stroke but no adverse event this period; 5) survived a CHD event this period; 6) survived a stroke this period; 7) death from a non-CVD-related cause; 8) death from CHD event this period; 9) death from stroke this period; and 10) dead.

Action Space

The physician makes annual treatment decisions $a_t^\pi(s_t)$ in accordance with policy π , which depends on the patient's state s_t at each decision period t . We assumed the physician is able to prescribe between 0 and 5 antihypertensive medications at full dosage. These antihypertensive medications may include thiazide diuretics, beta-blockers, calcium channel blockers, angiotensin-converting enzyme inhibitors, and angiotensin II receptor antagonists. However, we constrain the set of possible actions by a minimum allowable SBP threshold $b_{min}(s_t)$, which may depend on the patient's state. The minimum allowable SBP threshold prevents the physician from prescribing medication that would, in expectation, lower the patient's SBP below the threshold, so as to prevent treatment harm from excessive medication. We used a threshold of 120/90 mmHg. For example, a patient with an untreated SBP of 130 mmHg taking one medication would expect a BP reduction of 6.7 mmHg and a BP reduction of 12.73 mmHg if taking two medications.²⁰ With a threshold of 120 mmHg SBP, the patient could be prescribed one medication (resulting in a new SBP of 123.3 mmHg), but could not be prescribed two medications (since the new SBP of 117.27 mmHg is below 120 mmHg).

Transition Probabilities

Before treatment decisions are made at the beginning of each decision period, the patient's state s_t is mapped to a pretreatment, one-period likelihood of CHD, $r^{\text{CHD}}(s_t)$, and stroke, $r^{\text{Stroke}}(s_t)$, event using the risk calculators found.²¹ After a treatment decision is made at the beginning of each decision period, we assume the antihypertensive medication has a near-immediate effect on the patient's SBP. The expected change in SBP from treatment is calculated using the equations published in previous studies.²² Given the expected change in SBP, we compute the patient's posttreatment SBP (posttreatment SBP = pretreatment SBP – expected change in SBP). Using the patient's posttreatment SBP, we determine the post-state s_t^π , which includes all the other elements of s_t .

We further utilize the expected change in SBP to compute the relative risk reduction (RRR) for CHD, $RRR_{a_t^\pi}^{\text{CHD}}(s_t^\pi)$, and stroke, $RRR_{a_t^\pi}^{\text{Stroke}}(s_t^\pi)$, from treatment using equations from previously published studies.²² These equations utilize age- and event-specific risk slopes (see Table A2) and the reduction in SBP from treatment relative to a 20 mmHg

reduction to determine the RRR. Multiplying the patient's pretreatment risk by the RRR, we computed the posttreatment risk for CHD, $r_{a_t^\pi}^{\text{CHD}}(s_t^\pi)$, and stroke, $r_{a_t^\pi}^{\text{Stroke}}(s_t^\pi)$. If the patient has a history of CHD and/or stroke, we multiply the patient's CHD and stroke odds by a scaling factor ζ^{CHD} and/or ζ^{Stroke} .^{23,24} We used a scaling factor of 3 for CHD and stroke odds. With fatality likelihoods for CHD events, $\rho^{\text{CHD}}(s_t^\pi)$, and strokes, $\rho^{\text{Stroke}}(s_t^\pi)$,^{25–27} and total mortality likelihood, $\phi(s_t^\pi)$, we compute the probability of transitioning from one health state $h=1, \dots, 10$ to another health state $h'=1, \dots, 10$. The fatality likelihoods are presented in Table A3. We assume only one type of event may happen in a decision period, for example, a patient cannot have both a CHD event and a stroke. In order to guarantee valid probability distributions, we prioritize death from stroke over other outcomes and adjust the health state transition probabilities accordingly.

In addition to the computing health state transition probabilities, we model the dynamics of the other elements of the patient's state s_t over the planning horizon. To forecast the patient's untreated SBP, HDL, and TC, we utilized linear regression on the NHANES III data. We considered the following candidate predictors for each linear regression model: intercept term, age, squared age, sex, smoking status, diabetes status, race (white, black, Hispanic), history of CHD, and history of stroke. Using Akaike information criterion (AIC), the best models for SBP, HDL, and TC were determined. Table A4 presents the coefficients for each model. We applied the linear regression model to each patient to forecast that patient's SBP, HDL, and TC over the next 10 years. We adjusted the forecasts by applying the difference between the linear regression fitted value and the observed value in the NHANES data to all forecasted values (effectively changing the intercept term for each risk factor regression model).

The health state transition probabilities and linear regression models jointly determine the state transition probabilities $P(s_{t+1}|s_t, a_t^\pi(s_t))$, which depend on the treatment decision a_t^π .

Rewards

Our objective is to maximize the patient's expected discounted QALYs, that is, find the policy π^* , which maximizes Z :

$$Z = \max_{\pi} E^{\pi} \left[\sum_{t=1}^T \lambda^t \cdot R_t^{\pi}(s_t, a_t^{\pi}(s_t)) \right],$$

where $\lambda \in (0,1]$ is the discount factor and $R_t^\pi(s_t, a_t^\pi(s_t))$ is the reward resulting from treatment $a_t^\pi(s_t)$ when the patient is in state s_t at decision t . We used a discount factor of 0.97. The reward is equal to the quality of life associated with the patient's state $q(s_t)$ minus the burden from the treatment $d_{a_t^\pi}$. We assume the quality of life weight for each state s_t depends only on the patient's health state $h=1, \dots, 10$. Quality of life weights come from previously published literature.^{28–30} Table A5 reports the weights used.

Poisson Model Development

We split the full sample into a training set of 20,000 patients, a threshold set of 20,000 patients, and a testing set of 60,000 patients. Using the training set, we parameterized a Poisson regression model using linear mixed-effects modeling to account for the correlation of the sequential measurements taken to each patient during their annual examination, such as BP and cholesterol readings. The optimal medication count at each decision period from the MDP policy was used as the response variable, while demographic information and risk factor data were used as predictor variables. Poisson regression is a regression model where the outcome variable is a count variable (e.g., number of medications). By expressing the expected value of the outcome variable as an exponential function of a linear combination of predictor variables, Poisson regression guarantees that the predicted outcome is nonnegative—in contrast to traditional regression that would allow for predictions of negative medications. Linear mixed-effects modeling combines fixed effects (e.g., SBP, age, race) with random effects (e.g., patient). In biostatistics, the fixed effects represent population averages while the random effects represent patient-specific effects. Since our data contain multiple readings per patient, we require a random effect to control for the within patient correlation of outcome variables.

We considered two Poisson regression models: 1) full model and 2) risk-only. The full Poisson model used an initial candidate set of predictors of age, sex, smoking status, diabetes status, pretreatment SBP, diastolic blood pressure (DBP), HDL, TC, 5-year CVD risk (heart attack plus stroke) as computed by the Framingham risk calculator,²¹ and an interaction term for each variable with CVD risk. The risk-only Poisson model just considered 5-year

CVD risk, determined using the Framingham risk calculator, as a predictor.²¹ The final set of predictors and regression coefficients of the full Poisson model were determined using the AIC.

Since the optimal medication count comes from solving a finite-horizon discounted MDP, the optimal policy is time-dependent. However, since age is a state variable of the MDP, which increases at the same rate as time, the optimal policy can be viewed as approximately age-dependent. Therefore, the time-dependence of the optimal policy is approximately accounted for by including age as a predictor in the Poisson regression models.

Number of Medications to Prescribe: The Rounding Threshold

Within the models, we determined the number of medications to prescribe under the Poisson policies examining a sequence of rounding thresholds (0 to 1 by 0.1 increments). The sequence of thresholds was first applied to each parameterized Poisson regression function in the threshold set of 20,000 patients. If the Poisson regression fitted value was less than or equal to the rounding threshold, we rounded down to the nearest integer. For each rounding threshold, we computed the expected QALYs per 1000 patients. The rounding threshold that maximized QALYs was set as the preferred threshold.

Feasibility Adjustments

Because many believe hypertension is especially dangerous at higher levels, we treated all patients whose BP was $>150/90$ mmHg. Also, since there is no evidence that BP treatment is beneficial at especially low levels, our models do not allow treatment if the BP is below 120/90 mmHg.

Model Evaluation

Given the preferred threshold found in the threshold set ($n = 20,000$), we evaluate the performance of the Poisson policy model in the testing set ($n = 60,000$). We computed the expected QALYs saved and expected number of CVD events prevented (when compared to no treatment) under the two Poisson policies, the optimal MDP policy, and the JNC7 guidelines. We also compared the number of medications prescribed by the Poisson and MDP

Table 1 Summary Statistics of Population

Variable	Training ($n = 20,000$)	Testing ($n = 60,000$)
Age, median (IQR)	60.0 (22.0)	60.0 (22.0)
Male (%)	47.0%	47.1%
Diabetic (%)	8.7%	8.2%
Smoker (%)	23.6%	23.7%
SBP (mmHg), median (IQR)	133.0 (27.7)	133.1 (27.7)
DBP (mmHg), median (IQR)	77.1 (14.5)	77.3 (14.4)
HDL (mg/dL), median (IQR)	49.4 (20.2)	49.4 (20.0)
TC (mg/dL), median (IQR)	216.4 (58.7)	216.4 (57.8)
5-Year % CVD risk, median (IQR)	7.17% (11.03%)	7.23% (11.02%)

Note: IQR = interquartile range; SBP = systolic blood pressure; DBP = diastolic blood pressure; HDL = high-density lipoprotein; TC = total cholesterol.

Table 2 Comparison of MDP and Poisson Treatment Policies

Compared to MDP, Percentage of Decision Periods in Which	Full Poisson Model	Risk-Only Poisson Model
Equal meds prescribed (%)	99.6	91.6
Fewer meds prescribed (%)	0.0	8.1
More meds prescribed (%)	0.4	0.3

Note: MDP = Markov decision process.

policies to determine the rate at which the policies matched.

Sensitivity Analyses

To evaluate the robustness of the Poisson regression approach to approximating optimal MDP treatment policies, we also investigated how the Poisson policies performed when using a different CVD risk calculator, the ASCVD calculator.³¹ This analysis allowed us to determine if the Poisson policy would be effective in other countries where the risk calculators are calibrated to a different population. We also evaluated the treatment policies under scenarios of CVD risk calculator miscalibration. Given the reliance on CVD risk as a predictor variable in both Poisson policies and the optimal MDP policy, it is important to evaluate the performance of these treatment policies when the risk computed by the calculator is $\pm 25\%$ off from the patient's true CVD risk.

RESULTS

Comparison of Sample Populations

Table 1 summarizes the patient characteristics of the training and testing set. We found that

randomization was successful (i.e., no clinically significant differences between the populations).

Poisson Regression Analysis

We applied the Poisson policies to the testing set and compared the number of medications prescribed against the optimal MDP policy for each patient. Table 2 provides the percentage of decision periods where the Poisson policies exactly matched, prescribed fewer medications, and prescribed more medications than the optimal MDP policy. We found that the full Poisson model has 99% accuracy and a very low mean squared error (0.006), while the risk-only Poisson model (the model just considering the 5-year CVD risk as computed with the Framingham risk calculator) has 92% accuracy and a higher mean squared error (0.859). Further evidence of the accuracy of the Poisson policies is plotted in Figure 1, which shows the distribution of the number of medications for both Poisson policies and the optimal MDP policy. The frequency plotted is the number of times a patient was prescribed that number of medications for 1 year, based on the testing population of 60,000 patients followed for 10 years. The full Poisson policy and MDP policy do not have substantively different distributions. On the other

Table 3 Health Performance of Treatment Policies

Treatment Policy	Expected Number of CVD Events Prevented per 1000 Patients ^a		Expected Discounted QALYs Saved per 1000 Patients ^b	
	Total	Reduction Compared to MDP	Total	Reduction Compared to MDP
MDP	47.4	Ref	271.9	Ref
Full Poisson model	47.4	<0.1	271.9	<0.1
Risk-only Poisson model	46.4	1.0	266.5	5.4
JNC7	29.5	17.9	162.7	109.2

Note: CVD = cardiovascular disease; QALY = quality-adjusted life year; MDP = Markov decision process; JNC7 = Seventh Joint National Committee.

a. CVD events prevented when compared to no treatment.

b. QALYs saved when compared to no treatment.

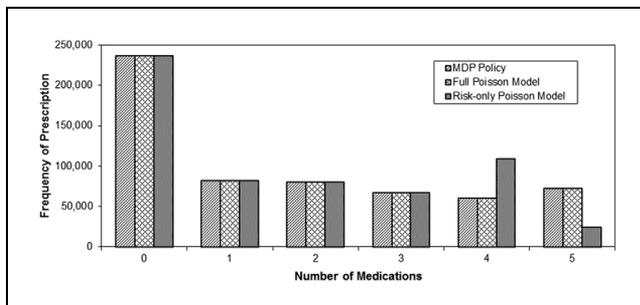


Figure 1 Number of medications prescribed for MDP and Poisson treatment policies

hand, the risk-only Poisson policy differs significantly when the optimal number of medications to prescribe is 4 or 5.

Results in Table 2 and Figure 1 only examine discordant prescribing decisions, but a much more important issue is how different prescribing decisions, using the Poisson policies instead of the optimal MDP policy, affects health outcomes for patients. Table 3 provides a comparison of the expected QALYs and expected number of CVD events when prescribing decisions are based on the two Poisson policies, the MDP policy and JNC7. We found that using the full Poisson policy results in nearly identical health performance for patients compared to the optimal MDP policy. Furthermore, the reduction in accuracy to 92% for the risk-only Poisson policy only reduced CVD events prevented by 1.0 and QALYs saved by 5.4 per 1000 patients (10,000 patient-years), compared to the MDP policy. In contrast, the full Poisson policy prevents 17.9 fewer events and saves 109.2 more QALYs per 1000 patients when compared to the JNC7 treatment policy.

Table A7 in the supplemental appendix presents the regression coefficients and *P* values of the full and risk-only Poisson models. The regression coefficients of the final set of predictor variables included in the full Poisson model were obtained using the AIC method. Based on these final regression models, we found that the optimal rounding threshold is 0, that is, always round the fitted value from the Poisson regression down to the nearest integer.

Sensitivity Analyses

To determine if the Poisson policies perform well when CVD risk is calculated based on a different population, we evaluated the performance of the four treatment policies when the ASCVD risk calculator is used to estimate CVD risk instead of the Framingham equations (see Table 4). We found that the improvement over JNC7 remained substantial for both CVD events prevented (17 fewer CVD events per 1000 patients) and QALYs gained for the full Poisson model and MDP policies when using the ASCVD calculator. The risk-only Poisson model continued to perform only slightly worse than the optimal MDP policy.

We next examined the effects of risk calculator miscalibration, that is, the calculated CVD risk systematically over- or underestimates the patient's true CVD risk. Table 4 reports the performance of the treatment policies under overestimation of CVD risk (+25% calibration error) and underestimation of CVD risk (−25% calibration error). For both over- and underestimation of true CVD risk, the Poisson policies yield higher expected CVD events prevented (22.2 and 13.5 fewer events per 1000 patients, respectively) and QALYs gained (136.2 and 81.4 more QALYs per 1000 patients, respectively) than JNC7.

Table 4 Health Performance of Treatment Policies Under ASCVD and Framingham Risk Calculator Miscalibration

Risk Calculator	Treatment Policy	Expected Number of CVD Events Prevented per 1000 Patients		Expected Discounted QALYs Saved per 1000 Patients	
		Total	Reduction Compared to MDP	Total	Reduction Compared to MDP
ASCVD	MDP	47.8	Ref	185.3	Ref
	Full Poisson	47.7	0.1	185.1	0.2
	Risk-only Poisson	46.8	1.0	181.3	4.0
	JNC7	30.6	17.2	115.7	69.6
+25% (Framingham calculator overestimates true CVD risk)	MDP	58.5	Ref	336.0	Ref
	Full Poisson	58.5	<0.1	336.0	<0.1
	Risk-only Poisson	57.7	0.8	331.3	4.7
	JNC7	36.3	22.2	199.8	136.2
-25% (Framingham calculator overestimates true CVD risk)	MDP	35.9	Ref	204.8	Ref
	Full Poisson	35.9	<0.1	204.8	<0.1
	Risk-only Poisson	34.8	1.1	199.3	5.5
	JNC7	22.4	13.5	123.4	81.4

Note: CVD = cardiovascular disease; QALY = quality-adjusted life year; MDP = Markov decision process; JNC7 = Seventh Joint National Committee.

The differences between MDP and the full Poisson model remain negligible and the differences between MDP and the risk-only Poisson model remain small.

DISCUSSION

Our analyses point the way for how decision analysis can become much more practical for daily clinical use, by finding that Poisson regression can be used effectively to approximate a fully optimized MDP model for the prescription of BP medications. The full Poisson model, which includes interactions effects, resulted in negligible differences in the treatment decisions and health outcomes compared to the optimal MDP treatment policy. The risk-only Poisson model (including only 5-year CVD risk as a factor) reduced CVD complications almost as much as the full Poisson model. Furthermore, both Poisson models dominated JNC7 treatment guidelines in QALYs and CVD events. We also found that the performance of the MDP and Poisson models was robust to the CVD risk calculator used (Framingham or ASCVD) and to calculator miscalibration ($\pm 25\%$). While not directly studied here, our QALY maximizing decisions likely lead to more cost-effective decisions than JNC7 given the high medical cost of CHD events and strokes and the low cost of BP medications.

More broadly, our study revealed that policy approximation can offer very comparable performance to MDP optimized treatment policies without the drawbacks of MDP models. MDP models generally require specialized software and lengthy computational times that would prohibit their use in many clinical settings. In this study, computational times were short (within seconds) but we required software capable of implementing recursive algorithms in order to solve the MDP. Approximations to these complex models limit these drawbacks by generating simple equations that can be solved nearly instantly using standard software or a calculator. With the advent of big data, decision support tools can develop approximations to the optimal treatment policy by training statistical models on very large data sets of patients. By doing so, the computational burden is removed from the end-user (i.e., the clinician or facility) and frontloaded to the decision support developer who has greater computational capabilities and access to the required software.

Furthermore, policy approximation improves the interpretability of the optimal treatment decision rules for clinicians. For Poisson regression, the coefficients of the predictors provide insight into how sensitive the optimal number of medications is to changes in patient risk factors. This insight may

translate to improved trust of the decision support tool and increase its utilization and acceptance in clinical practice. This is particularly true for the risk-only Poisson model since there is only one regression coefficient to interpret, whereas the full Poisson model may have multicollinearity, which prevents direct interpretation of the sign and magnitude of the coefficients. For the risk-only Poisson model, the intercept coefficient is 1.167 and the coefficient for 5-year CV risk is 1.006. The positive sign of the risk coefficient implies that increases in CV risk lead to increases in the number of medication prescribed. This matches physician intuition and guidelines, which bolsters the trustworthiness of the Poisson regression. The intercept coefficient can be interpreted to mean the risk-only Poisson model would prescribe at least 3 medications to patients (before applying the minimum SBP threshold constraint). The risk coefficient implies that the patient's pretreatment 5-year CV risk would need to be at least 22% (44%) to recommend 4 (5) BP medications. The full list of the Poisson models' coefficients is available in the supplemental appendix.

In general, our policy approximation methodology could be applied to any MDP formulation developed with a large structured dataset. For instance, the optimal prescription planning derived with MDP formulations for patients with type 2 diabetes²⁰ or for patients with heart disease⁷ could be approximated using Poisson regression, as long as there is enough data to fit the model. Our methodology can also be applied to other objective functions, including maximizing cost-effectiveness and minimizing the number of CVD events, and application domains.

Our study's limitations derive mainly from questions about the inputs in our models. Most notably, the precise details of who benefits from BP lowering are debatable. Our modeling choices, however, are based on the best available clinical trial data from an individual-level meta-analysis³² and fit the SPRINT Trial well, which found treatment benefit to 120 mmHg when using 3 to 4 BP medications in very high CVD risk patients.¹⁵ We chose to base our comparison on the JNC7,¹⁸ because American practitioners often know it well, and JNC8 was not a consensus guideline and is unlikely to be relevant for much longer. Efforts are already underway to update these guidelines, and our results suggest that our benefit-based tailored treatment approach should at least be strongly considered in deliberations of how to update BP guidelines post-SPRINT.¹⁴

Critics of using CVD risk in clinical decision making have pointed to risk prediction tools'

well-known susceptibility to calibration problems when developed in one population and applied to different populations.³³ Our results, however, demonstrate that a benefit-based treatment approach remains superior to traditional treatment target-based approaches, such as JNC7, even in the presence of substantial calibration problems. Of course, it is always preferable to use a risk prediction tool that is well calibrated to the patient population being treated, and in the era of the modern electronic medical records, it is becoming much more feasible for integrated health systems to develop or recalibrate risk tools to be optimally fitted to their own population.³⁴ Furthermore, critics have pointed out that risk prediction tools with similar discrimination, as measured by the area under the receiver operating curve, often given substantively different risk estimates for individual patients. Our study demonstrates that whether the Framingham and ASCVD tool was used,^{21,31} the Poisson models prevented many more CVD events and saved many more QALYs compared to the JNC7 approach.

CONCLUSION

We developed a Poisson regression model to approximate the optimal treatment decisions from an MDP model. We found that our policy approximation enables fast, easily interpretable, and comparable decision support without a need for specialized software. Our application to hypertension treatment planning for patients at risk for CVD indicates high fidelity to optimal treatment policies and high performance. Treatment policy approximation resulted in more QALYs and fewer CVD events than current clinical practice. Furthermore, the improvement in health outcomes over JNC7 was robust to which risk calculator was used, as well as to systematic calculator error. While this study specifically addressed hypertension treatment, we believe the methodology of using Poisson regression to approximate fully optimized MDP policies has the potential to reduce computational time, improve acceptability by clinicians, and maintain health-optimization performance for treatment decisions for chronic diseases more generally.

REFERENCES

1. Denton BT, Alagoz O, Holder A, Lee EK. Medical decision making: open research challenges. *IIE Trans Healthcare Syst Eng.* 2011;1(3):161–7.

2. Zhang J, Mason JE, Denton BT, Pierskalla WP. Disease prevention, detection, and treatment. In: *Encyclopedia of Operations Research and Management Science*. 3rd ed. Berlin, Germany: Springer; 2013. p. 437–47.
3. Brandeau ML, Sainfort F, Pierskalla WP, eds. *Operations Research and Health Care: A Handbook of Methods and Applications*. Berlin, Germany: Springer; 2004.
4. Schaefer AJ, Bailey MD, Shechter SM, Roberts MS. Modeling medical treatment using Markov decision processes. In: Brandeau ML, Sainfort F, Pierskalla WP, eds. *Operations Research and Health Care: A Handbook of Methods and Applications*. Berlin, Germany: Springer; 2004. p. 593–612.
5. Alagoz O, Hsu H, Schaefer AJ, Roberts MS. Markov decision processes: a tool for sequential decision making under uncertainty. *Med Decis Making*. 2010;30(4):474–83.
6. He M, Zhao L, Powell WB. Optimal control of dosage decisions in controlled ovarian hyperstimulation. *Ann Oper Res*. 2010;178(1):223–45.
7. Hauskrecht M, Fraser H. Planning treatment of ischemic heart disease with partially observable Markov decision processes. *Artif Intell Med*. 2000;18:221–44.
8. Bate P, Robert G, Gallivan S, et al. The development and implementation of NHS treatment centres as an organisational innovation. Available from: <http://www.sdo.lshtm.ac.uk/files/project/45-final-report.pdf>
9. Topaloglu H. Computation and dynamic programming. In: Smith C, ed. *Wiley Encyclopedia of Operations Research and Management Science*. New York, NY: Wiley; 2011.
10. Kochanek KD, Xu J, Murphy SL, Miniño AM, Kung HC. Deaths: final data for 2009. *Natl Vital Stat Rep*. 2011;60(3):1–116.
11. Gillespie CD, Hurvitz KA. Prevalence of hypertension and controlled hypertension—United States, 2007–2010. *MMWR Suppl*. 2013;62(3):144–8.
12. Santulli G. Epidemiology of cardiovascular disease in the 21st century: updated numbers and updated facts. *J Cardiovasc Dis*. 2013;1(1):1–2.
13. Schell GJ. Personalized medicine in chronic disease management (PhD thesis), University of Michigan; 2015. Available from: <https://deepblue.lib.umich.edu/handle/2027.42/111447>
14. Sussman J, Vijan S, Hayward R. Using benefit-based tailored treatment to improve the use of antihypertensive medications. *Circulation*. 2013;128(21):2309–17.
15. Hayward R, Krumholz H, Zulman M, Timbie J, Vijan S. Optimizing statin treatment for primary prevention of coronary artery disease. *Ann Intern Med*. 2010;152(2):69–77.
16. Hayward RA. Moneyball, gambling, and the new cholesterol guidelines. *Circ Cardiovasc Qual Outcomes*. 2014;7(2):311–4.
17. National Center for Health Statistics. Third National Health and Nutrition Examination Survey (NHANES III), 1988–1994, data files. Available from: <http://www.cdc.gov/nchs/nhanes/nh3data.htm>
18. Chobanian AV, Bakris GL, Black HR, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA*. 2013;289(19):2560–72.
19. Puterman M. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: John Wiley; 2009.
20. Mason JE, Denton BT, Shah DN, Smith SA. Optimizing the simultaneous management of blood pressure and cholesterol for type 2 diabetes patients. *Eur J Oper Res*. 2014;233(3):727–38.
21. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J*. 1991;121(1 Pt 2):293–8.
22. Law MR, Morris JK, Wald NJ. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *BMJ*. 2009;338:b1665.
23. Rifai N, Warnick G, Dominiczak M. *Handbook of Lipoprotein Testing*. Washington, DC: American Association for Clinical Chemistry; 2000.
24. Burn J, Dennis M, Bamford J, Sandercock P, Wade D, Warlow C. Long-term risk of recurrent stroke after a first-ever stroke. The Oxfordshire Community Stroke Project. *Stroke*. 1994;25(2):333–7 [Erratum appears in *Stroke*. 1994;25(9):1887].
25. Lloyd-Jones D, Adams R, Carnethon M, et al. Heart disease and stroke statistics—2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2009;119(3):e21–e181.
26. Arias E. United States life tables, 2004. *Natl Vital Stat Rep*. 2007;56(9):1–39.
27. Heron M. Deaths: leading causes for 2004. *Natl Vital Stat Rep*. 2007;56(5):1–95.
28. Fryback DG, Dasbach EJ, Klein R, et al. The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors. *Med Decis Making*. 1993;13(2):89–102.
29. Pignone M, Earnshaw S, Pletcher MJ, Tice JA. Aspirin for the primary prevention of cardiovascular disease in women: a cost-utility analysis. *Arch Intern Med*. 2007;167(3):290–5.
30. Pignone M, Earnshaw S, Tice JA, Pletcher MJ. Aspirin, statins, or both drugs for the primary prevention of coronary heart disease events in men: a cost-utility analysis. *Ann Intern Med*. 2006;144(5):326–36.
31. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129(25 Suppl. 1):S49–S76.
32. Drazen JM, Morrissey S, Campion EW, Jarcho JA. A SPRINT to the finish. *N Engl J Med*. 2015;373(22):2174–5.
33. Ridker PM, Wilson PW. A trial-based approach to statin guidelines. *JAMA*. 2013;310(11):1123–4.
34. Kennedy EH, Wiitala WL, Hayward RA, Sussman JB. Improved cardiovascular risk prediction using nonparametric regression and electronic health record data. *Med Care*. 2013;51(3):251–8.