PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Dynamic Treatment Selection and Modification for Personalized Blood Pressure Therapy using a Markov Decision Process Model: A Cost-effectiveness Analysis
AUTHORS	Choi, Sung Eun; Brandeau, Margaret; Basu, Sanjay

VERSION 1 – REVIEW

REVIEWER	Muge Capan, PhD Christiana Care Health System, USA
REVIEW RETURNED	09-Jul-2017

GENERAL COMMENTS	The authors present a discrete-time finite-horizon Markov Decision Process (MDP) approach to identify personalized blood pressure treatment policies. This manuscript is contributing to an important clinical domain: Fine-tuning the blood pressure treatment regime using medication for different populations has been receiving increasing attention with significant implications for chronic conditions as well as acute deterioration episodes. The authors provide a clear description of the model components and present the MDP model results regarding initiation and continuation of blood pressure treatment by comparing the age-based policies derived from the MDP model with current U.S. blood pressure treatment guidelines. The paper would greatly benefit from implementing following feedback. General Comments □ Throughout the paper, the literature discussion sections lack a focus on previously applied analytical models on blood pressure treatment and key findings. When existing studies are discussed, please discuss their findings and contribution and how your approach differs from previous work/ or how it contributes to this field: (i) from a clinical perspective, and (ii) from an IE/methodology perspective. □ A major concern with the methodology arises due to the inconsistent use of terminology and need for clarification of the personalization aspect in modeling approach which will be discussed further below. □ Some of the model assumptions are relatively strong, such as the cycle length of 1 month, and treatment/dose-based actions including the introduced treatment dose scale, and require further discussion with clear clinical explanation and potentially an expansion of the presented sensitivity analysis. □ Certain clinically important aspects are not considered in the modeling approach, such as the impact of compliance/adherence to treatment which would benefit from further discussion.

	□ Discussion of the clinical implications of the findings would benefit from more structured discussion on implication for different acute and chronic conditions. 2 Terminology Authors use the terms dynamic Markov Decision Process (DMDP) and MDP models interchangeably to describe their methodology. MDP models are dynamic models due to their nature, they are used to capture sequential decision making problems under uncertainty. In this context, the use of “dynamic” is redundant. I recommend revising the terminology consistently throughout the manuscript. State space The presented MDP model has seven fully observable states which are based on CVD and MI experience and death. Based on Figure 1 and description of state space in the text, state space does not include any additional patient-level information, such as patient demographics. In this discrete-time MDP model states change in 1-month intervals when the decisions are made about treatment initiation and/or continuation. □ What is the reasoning behind the cycle length? Is it realistic to assume that patients will be seen by a doctor every month since the target population is relatively heterogeneous and includes different types of chronic and acute conditions which may require blood pressure treatment? □ It is unclear why the state description includes the following: “The state space S, comprising the states at time t (s_t), consisted of demographic information (age, sex, race/ethnicity), the patient’s CVD-related covariates (age-, sex- and race dependent chronic kidney disease status, type 2 diabetes status, lipid profile, and tobacco smoking status), and the patient’s health state”. It seems that these components impact the state transition probabilities and costs/rewards rather than being part of the state space. Clarification would help. Action space Action space includes the following at any given 1-month decision epoch: stop medication, continue current medication dose/type, and increase/change medication. The treatment dose scale includes number of drugs and ½ dose increments which require further discussion. □ How were the ½ dose increment selected? Was this discussed with clinicans? The dose selection may depend on the population, their SBP baseline, and treatment purpose. Is this a clinically reasonable assumption? □ Do the authors suggest that only the number of drugs that a patient is currently on

	matters and not the combination of which drugs are being prescribed and used? Why does the action space not include decrease of the dose? Following statement would benefit from further clarification “Treatment dose levels were capped at full doses of 4 different medicines, given evidence of no incremental 3 benefit and substantial harm from side effects when escalating therapy beyond 4 full doses” How would action space differ for patients with certain comorbidities or if the treatment is prescribed for a certain chronic condition rather than an acute treatment of blood pressure – which would significantly, impact the action choice since not every blood pressure treatment has the same target SBP levels? State transition probabilities Please clarify the following statement “We back-calculated the pre-treatment blood pressure for those NHANES subjects reporting current blood pressure treatment, using a previously-published procedure” Are the state transition probabilities non-stationary and based on the age of the patient? Clarification would help. Results Results show that “MDP strategy prescribed 2.34 (95% CI: 2.33-2.34) medication doses per person, versus 2.22 (95% CI: 2.21-2.22) under the JNC8”, and “Treated individuals experienced 19.03 (95% CI: 19.01-19.05) and 18.97 (95% CI: 18.95- 18.98) total discounted QALYs per person under the DMDP and JNC8 strategies,” The difference in both QALYs and dose per person seems to be negligible. Did the authors test if the difference is statistically significantly different? What could be the reason behind the results of MDP being so similar to the existing guidelines? Personalization aspects It is stated that the main difference between the MDP policy and the existing guidelines is the fact that the MDP model recommends earlier intensive treatment and lower doses at later stages in life preventing more intensive treatment later in life and saving QALYs. While the age-based recommendation related to the treatment intensity is a starting point for personalization, the demographics, comorbidities and any other personalization aspect mentioned in the paper in not reflected in the results. The personalization contribution of the modeling approach need clear reformulation and potential modifications to the model to highlight the difference between MDP policies and existing guidelines not just based on age, but other characteristics of the populations. In the current version of the manuscript, sensitivity analysis focuses on the treatment-disutility and disutility from CVD events including stroke and MI. In this context sensitivity analysis expanded beyond just QALYs and disutility function would help.

REVIEWER	Mario Naranjo, MD Albert Einstein Medical Center, Philadelphia, PA.
REVIEW RETURNED	21-Jul-2017

GENERAL COMMENTS	I really enjoyed reading this study with an important and useful new strategy to focus on cost effectiveness in the treatment of hypertension. The new JNC-8 has lots of controversy not only because of the new, generalized and liberal goals, but because of the interesting on opposite results of the SPRINT trial. We clearly see multiple new studies coming soon to help us understand what are the best goals and treatments for our patients. Taking all these facts together, the authors have integrated the newest and most available information the help us understand that this new method could help not only saving money, but also QALYs. This new dynamic approach will definitely aid the decision-making to personalize blood pressure treatment and improve overall population health compared to current and not well accepted blood pressure treatment guidelines.

REVIEWER	Jinsung Yoon University of California, Los Angeles, CA, USA No Competing Interest
REVIEW RETURNED	02-Aug-2017

GENERAL COMMENTS	In summary, the paper is well written, and authors understand the problem very well. Furthermore, I am surprised that authors understand the limitations of the works very accurately. It is straightforward to comprehend because their purpose to write the paper is clear. These are my concerns, and you may need to address to the revision. 1. Simulated patients - I know that it is tough to obtain a large number of real patients data. However, using real-world medical data to justify your work is very important, and it increases the quality of paper much higher. - Even if you justified your sampling process by some tests, you only show that for each univariate distribution but not for the multi-variate distribution. - Therefore, I would highly recommend that you should obtain some actual patient datasets and compare your simulated datasets with multivariate statistical tests. It will verify that your simulated patients come from actual distribution. (We can make many different data sets that can fit for the univariate statistical test, but not fit for the multivariate statistical test). 2. Gain is too small. - I think the QALY gain 0.06 is too small. Even if you justified that 95% CI is also higher than 0; it does not say that your gain is high enough to be used. - Furthermore, the more than 1,000 dollars can be substantial per a year. However, I think it is computed for the entire lifetime. In that point of view, I am not sure that this has the significant impact to be published in BMJ. 3. Other machine learning methods.

	- MDP is a standard Machine learning method to do the decision making. However, this is not the only model that ML has. We have many reinforcement learning methods, PoMDP, HMM, etc. - I think you need to explain why you select MDP instead of other methods. The intuition is needed. - For instance, you assume that your states consist of 7 known states. However, in reality, the patient states are decided by both observed features and non-observed features. Therefore, PoMDP or HMM would be more general methods without assumptions and fitted to this problem. Please justify why you use MDP instead of other methods. 4. Etc. - I think the action set is too small. There are various medications; however, you treat them all the same. (Only divide them into the number of medications and the dosage) I think in the future, you should consider the differences between medications. - You learned your models based on the high-quality meta-analytic data. You also said that some of them are randomized trials. Are all of them are randomized trials? Otherwise, you need to explain how to deal with treatment selection bias. - As it can be seen in Figure 1, JNC8 and Intensive JNC8 is interpretable. However, your model is not interpretable. In other words, your MDP model is a black box and doctors cannot justify why I should use this medication dosage. To convince doctors, you should visualize your MDP model. I know this is hard to make like JNC8 decision tree; however, you need to provide the reason to doctors that why MDP make this decision.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1

The authors present a discrete-time finite-horizon Markov Decision Process (MDP) approach to identify personalized blood pressure treatment policies. This manuscript is contributing to an important clinical domain: Fine-tuning the blood pressure treatment regime using medication for different populations has been receiving increasing attention with significant implications for chronic conditions as well as acute deterioration episodes. The authors provide a clear description of the model components and present the MDP model results regarding initiation and continuation of blood pressure treatment by comparing the age-based policies derived from the MDP model with current U.S. blood pressure treatment guidelines. The paper would greatly benefit from implementing following feedback.

General Comments

1. Throughout the paper, the literature discussion sections lack a focus on previously applied analytical models on blood pressure treatment and key findings. When existing studies are discussed, please discuss their findings and contribution and how your approach differs from previous work/ or how it contributes to this field: (i) from a clinical perspective, and (ii) from an IE/methodology perspective.

Response: We now have added a discussion on page 5 of how a prior study of evaluating blood pressure treatment using MDP assesses its impact on certain cohorts of patients, but not at the population level. Our study advances this prior work by incorporating large-scale meta-analytic and network meta-analytic data sources, providing a full life-course simulation in order to evaluate the long-term impact of treatment, and using population-representative data to assess the overall national implications of personalized blood pressure treatment selection. We have also edited the discussion on page 14 to highlight the fact that while previous studies have suggested that blood pressure therapy should involve patient risk calculations for a binary treatment decision (whether to treat or not) using a Markov simulation model, our study proposes a systematic and dynamic approach to incorporate patient heterogeneity in risk factors and responses to treatment (probabilistically), and recommends detailed treatment suggestions (doses and types of medication). From a clinical perspective, our analysis advances the literature beyond the decision of whether to treat high blood pressure because both type and dosage of medication are critical decision points. We also highlight this point in the Discussion section.

2. A major concern with the methodology arises due to the inconsistent use of terminology and need for clarification of the personalization aspect in modeling approach which will be discussed further below.

Response: We have removed the term “dynamic" and now use the term “Markov Decision Process-based treatment (MDPT)” instead throughout the manuscript. Also, we have now edited the material on pages 7 and 11 and have added eFigure 5 in the Appendix to explain in more detail the personalization aspect of the modeling approach. In particular, we now clarify that the MDP does not simply rely on age as a single factor for recommending treatment; rather, the MDP incorporates multiple factors as we now further explain on page 7 (i.e., it is more personalized to multiple covariates). One of our key results is that treatment differs from current guidelines based on the age of the patient (with younger patients treated more intensively by the MDP than by current guidelines). We have added further results by sex and race/ethnicity on page 11 and in the new eFigure5 to demonstrate how the personalized treatment regimen determined by the MDP produces differences in recommended treatment across several population demographic features. Additionally, we ensure in our revision that our use of terms for personalization versus population-level results are consistent and defined.

3. Some of the model assumptions are relatively strong, such as the cycle length of 1 month, and treatment/dose-based actions including the introduced treatment dose scale, and require further discussion with clear clinical explanation and potentially an expansion of the presented sensitivity analysis.

Response: We have added further discussion of cycle length on page 6. The cycle length of one month was chosen to be in accordance with current clinical guidelines, which indicate that clinicians should adjust treatment on roughly a 1-month time period. However, even shorter periods with more rapid adjustment would not produce dramatic differences in our results since the QALYs and costs accumulate from blood pressure treatment over the long term and therefore small within-month QALY and cost differences are subsumed by much larger long-term QALY and cost outcomes. The 1-month time step also corresponds to the timing of observations for SBP reduction in randomized trials used to inform the model, and captures the rapid response in blood pressure to antihypertension medications. Regarding the treatment dose scale, we now note in Appendix eText 1 that treatment dose scale and medications used in our model were based on a meta-analysis of 147 randomized studies, and our choice of dose scale corresponds to the dose scales in the randomized trials that serve as our input data.1 To our knowledge, this meta-analytic data provides the most reliable relationship between antihypertensive treatment dose levels and systolic blood pressure reduction. For this reason, we implement the treatment dose scales used in that study.

4. Certain clinically important aspects are not considered in the modeling approach, such as the impact of compliance/adherence to treatment which would benefit from further discussion.

Response: We have added discussion on page 15-16 that our purpose in this model is to compare how a change in guidelines from the current guidelines to the MDP-based strategy would affect overall population-level outcomes under ideal treatment conditions. Adding in a compliance/adherence parameter would simply linearly scale the outcomes to the proportion of patients who adhere, unless we have further data from [not yet extant] randomized trials suggesting that patients would adhere differently to treatment regimens found using the MDP-based approach than to regimens identified under the current guidelines. We suggest that such randomized trials and real-world studies should be performed, now that a proof-of-concept model is available, to gather empirical data comparing adherence rates and observed outcomes between the MDP and current guideline approach.

5. Discussion of the clinical implications of the findings would benefit from more structured discussion on implication for different acute and chronic conditions.

Response: We have added material to our Discussion section on page 15 to highlight that the cumulative QALY benefit from the MDP approach accrues primarily from long-term chronic outcomes of high blood pressure leading to long-term sequelae of MIs and strokes. We now highlight in Appendix eFigure 7 the portion of the QALYs that come from acute MI/stroke versus the portion that come from post-CVD-event disability.

Terminology

1. Authors use the terms dynamic Markov Decision Process (DMDP) and MDP models interchangeably to describe their methodology. MDP models are dynamic models due to their nature, they are used to capture sequential decision making problems under uncertainty. In this context, the use of “dynamic” is redundant. I recommend revising the terminology consistently throughout the manuscript.

Response: We have removed the word “dynamic" and now use the term “Markov Decision Process-based treatment (MDPT)” instead throughout the manuscript.

State space

The presented MDP model has seven fully observable states which are based on CVD and MI experience and death. Based on Figure 1 and description of state space in the text, state space does not include any additional patient-level information, such as patient demographics. In this discrete-time MDP model states change in 1-month intervals when the decisions are made about treatment initiation and/or continuation.

1. What is the reasoning behind the cycle length? Is it realistic to assume that patients will be seen by a doctor every month since the target population is relatively heterogeneous and includes different types of chronic and acute conditions which may require blood pressure treatment?

Response: We have added a statement on page 6 to clarify why we assumed a 1-month cycle length. The cycle length of one month was chosen to be in accordance with current clinical guidelines which indicate that clinicians should adjust treatment on roughly a 1-month time period. The senior author, who is a practicing clinician, follows current practice guidelines in which most clinicians see patients monthly during up-titration of blood pressure medications. The 1-month time step also corresponds to the timing of observations for SBP reduction in randomized trials used to inform the model, and captures the rapid response in blood pressure to antihypertension medications.

2. It is unclear why the state description includes the following: “The state space S, comprising the states at time t (s_t), consisted of demographic information (age, sex, race/ethnicity), the patient’s CVD-related covariates (age-, sex- and race dependent chronic kidney disease status, type 2 diabetes status, lipid profile, and tobacco smoking status), and the patient’s health state”. It seems that these components impact the state transition probabilities and costs/rewards rather than being part of the state space. Clarification would help.

Response: We have edited the exposition on page 7 to clarify that state transition probabilities to the next state, P(st+1| st, at), are determined by the state (st) the patient is currently in and the action taken at time t. The current state of the patient is not unidimensional, but multidimensional: the patient is represented by a set of covariates, such as demographic information and CVD-related covariates.

Action space
Action space includes the following at any given 1-month decision epoch: stop medication, continue current medication dose/type, and increase/change medication. The treatment dose scale includes number of drugs and ½ dose increments which require further discussion.

1. How were the ½ dose increment selected? Was this discussed with clinicians? The dose selection may depend on the population, their SBP baseline, and treatment purpose. Is this a clinically reasonable assumption?

Response: We now note in Appendix eText 1 that the treatment dose scale and medications used in our model were based on a meta-analysis of 147 randomized studies, and therefore our choice of dose scale corresponds to the randomized trials that serve as our input data.1 To our knowledge, this meta-analytic data provides the most reliable relationship between antihypertensive treatment dose levels and systolic blood pressure reduction. Hence, we implemented the treatment dose scales used in that study.

2. Do the authors suggest that only the number of drugs that a patient is currently on matters and not the combination of which drugs are being prescribed and used?

Response: We have added material to Appendix eText2 stating that the blood pressure lowering effect of the drugs was estimated as a function of treatment dose and pretreatment blood pressure, and was not dependent on types of drugs, except for beta-blockers, concordant with the current literature on blood pressure treatment choice. Beta-blockers have been shown to be more effective than other antihypertensive drugs included in our study in lowering blood pressure among patients with previous history of CVD.1 In our model, beta-blockers tended to be favored over other types of drugs for patients with CVD history due to their added benefit. Also, we highlight on page 8 and in Appendix eTable 1 that the probability of severe adverse events was determined by blood pressure drug choice and dosage.

3. Why does the action space not include decrease of the dose?

Response: We have now added a statement to Appendix eText 1 to clarify that we did not include decreases in dosage in our action space in order to mimic how current clinical practice (including the protocol in randomized trials) is conducted. Usually blood pressure medication is prescribed, with increases in dosage if necessary, until the patient’s blood pressure meets the target blood pressure goal. Once the target blood pressure is reached, the patient no longer changes medication dosage and typically stays on the same dosage for life.

4. Following statement would benefit from further clarification “Treatment dose levels were capped at full doses of 4 different medicines, given evidence of no incremental 3 benefit and substantial harm from side effects when escalating therapy beyond 4 full doses”

Response: We now have edited the statement on page 7 and the statement in Appendix eText 2 to indicate that current data suggest that increasing blood pressure medications beyond 4 full doses has been found to only increase side effects/adverse effects while not providing any incremental benefit for patients.

5. How would action space differ for patients with certain comorbidities or if the treatment is prescribed for a certain chronic condition rather than an acute treatment of blood pressure – which would significantly, impact the action choice since not every blood pressure treatment has the same target SBP levels?

Response: There is not a clinical separation between acute and chronic treatment of blood pressure. All blood pressure treatment outside of the setting of the rare condition of acute hypertensive emergency (which is treated in a hospital, not an outpatient setting) is designed to reach the same target SBP levels, dependent on patient characteristics, to avoid chronic sequelae of elevated blood pressure. Our paper focuses on the key issue of chronic blood pressure elevation, which is the vast majority of blood pressure treatment.

State transition probabilities
1. Please clarify the following statement “We back-calculated the pretreatment blood pressure for those NHANES subjects reporting current blood pressure treatment, using a previously-published procedure”

Response: We edited the statement on page 8 to clarify that because some NHANES participants were taking antihypertensive medications at the time of the survey, we estimated untreated blood pressure of those NHANES participants using a previously published back-calculation method. This allows us to accurately sample baseline pretreatment blood pressure levels among the US population.

2. Are the state transition probabilities non-stationary and based on the age of the patient? Clarification would help.

Response: We have added a statement on page 7 to clarify that the state transition probabilities are non-stationary and depend on patient characteristics – both the patient’s demographic information (age, sex, race/ethnicity) and CVD-related covariates. For example, the probability that a person transitioned from healthy to either MI or stroke was based on equations previously validated in several diverse cohorts, and was a function of patient characteristics as highlighted on page 7 and in Appendix eText 3.

Results
Results show that “MDP strategy prescribed 2.34 (95% CI: 2.33-2.34) medication doses per person, versus 2.22 (95% CI: 2.21-2.22) under the JNC8”, and “Treated individuals experienced 19.03 (95% CI: 19.01-19.05) and 18.97 (95% CI: 18.95- 18.98) total discounted QALYs per person under the DMDP and JNC8 strategies,” The difference in both QALYs and dose per person seems to be negligible. Did the authors test if the difference is statistically significantly different? What could be the reason behind the results of MDP being so similar to the existing guidelines?

Response: The result of a small gain in QALYs is expected in the case of blood pressure treatment because only a small subset of patients experience a CVD event and the QALY differences among the subset of patients who are treated differently and then also get a CVD event are resultingly small. This is typical in CVD intervention research. We follow current cost-effectiveness guidelines in capturing the overall QALYs and costs across the entire population, discounted over the entire life-course, to avoid inflated estimates of QALY and cost comparisons. We also follow current International Society for Pharmacoeconomics and Outcomes Research (ISPOR) guidelines that indicate that statistical significance at arbitrary p-value levels should not be tested using simulations, since the width of confidence intervals depends on the choice of simulation iterations and sampling rather than on an empirical distribution.

Personalization aspects
It is stated that the main difference between the MDP policy and the existing guidelines is the fact that the MDP model recommends earlier intensive treatment and lower doses at later stages in life preventing more intensive treatment later in life and saving QALYs. While the age-based recommendation related to the treatment intensity is a starting point for personalization, the demographics, comorbidities and any other personalization aspect mentioned in the paper in not reflected in the results. The personalization contribution of the modeling approach need clear reformulation and potential modifications to the model to highlight the difference between MDP policies and existing guidelines not just based on age, but other characteristics of the populations. In the current version of the manuscript, sensitivity analysis focuses on the treatment-disutility and disutility from CVD events including stroke and MI. In this context sensitivity analysis expanded beyond just QALYs and disutility function would help.

Response: We have edited the material on page 7 and page 11 and added Appendix eFigure 5 to explain in more detail how the MDP performs optimization to choose the most beneficial treatment option accounting for a patient's characteristics (both demographic and CVD-related covariates) by estimating the expected quality-adjusted life year (QALY) gains from averted CVD events or side effects that are specific to that patient. eFigure 5 shows how treatment dose levels varied by age, sex, and race/ethnicity, not just by age (used for stratification of the US population in the NHANES data). The increases in treatment dosages resulting from MDP optimization are not only highlighted by age, but also by race based on patient risks in those populations.

Reviewer #2

I really enjoyed reading this study with an important and useful new strategy to focus on cost effectiveness in the treatment of hypertension. The new JNC-8 has lots of controversy not only because of the new, generalized and liberal goals, but because of the interesting on opposite results of the SPRINT trial. We clearly see multiple new studies coming soon to help us understand what are the best goals and treatments for our patients. Taking all these facts together, the authors have integrated the newest and most available information the help us understand that this new method could help not only saving money, but also QALYs. This new dynamic approach will definitely aid the decision-making to personalize blood pressure treatment and improve overall population health compared to current and not well accepted blood pressure treatment guidelines.

Response: Thank you.

Reviewer #3

In summary, the paper is well written, and authors understand the problem very well. Furthermore, I am surprised that authors understand the limitations of the works very accurately. It is straightforward to comprehend because their purpose to write the paper is clear.

These are my concerns, and you may need to address to the revision.

1. Simulated patients: I know that it is tough to obtain a large number of real patients data. However, using real-world medical data to justify your work is very important, and it increases the quality of paper much higher. Even if you justified your sampling process by some tests, you only show that for each univariate distribution but not for the multi-variate distribution. Therefore, I would highly recommend that you should obtain some actual patient datasets and compare your simulated datasets with multivariate statistical tests. It will verify that your simulated patients come from actual distribution. (We can make many different data sets that can fit for the univariate statistical test, but not fit for the multivariate statistical test).

Response: We have now added discussion to Appendix eText1 with a detailed explanation of how we created the simulated population using National Health and Nutrition Examination Survey (NHANES) data, which is a nationally-representative study to assess the health and nutritional status of the US population. NHANES data includes demographic, socioeconomic, and health-related information from real individuals, and hence constitutes 'real patient data'. We constructed a nationally representative sample population of the US by first constructing a population distributed among demographic groups per US Census statistics, and then assigned CVD risk factors to each individual using repeated Monte Carlo sampling. The sampling procedure took into account the probability distributions of each risk factor in NHANES, while also accounting for the correlation between risk factors and the NHANES sample weights. To account for correlation between these CVD risk factors, we used multivariate sampling with copulas (i.e., a method in which the covariance matrix among risk factors, not just the marginal distributions, is taken into account during sampling)
In Appendix eFigures 3-4, we show how well the disease incidence estimates of the simulated population compare to estimates for the current US population.

2. Gain is too small.
I think the QALY gain 0.06 is too small. Even if you justified that 95% CI is also higher than 0; it does not say that your gain is high enough to be used. Furthermore, the more than 1,000 dollars can be substantial per a year. However, I think it is computed for the entire lifetime. In that point of view, I am not sure that this has the significant impact to be published in BMJ.

Response: The results of a small gain in QALYs is expected in the case of blood pressure treatment because only a small subset of patients experience a CVD event and the QALY differences among the subset of patients who are treated differently and then also get a CVD event are resultingly small. This is typical in CVD intervention research. We follow current cost-effectiveness guidelines in capturing the overall QALYs and costs across the entire population, discounted over the entire life-course, to avoid inflated estimates of QALY and cost comparisons. We also follow current International Society for Pharmacoeconomics and Outcomes Research (ISPOR) guidelines that indicate that statistical significance at arbitrary p-value levels should not be tested using simulations, since the width of confidence intervals depends on the choice of simulation iterations and sampling rather than on an empirical distribution. Note also that the MDP-based treatment strategy is cost-saving compared to the JNC8 guidelines.

3. Other machine learning methods.
MDP is a standard Machine learning method to do the decision making. However, this is not the only model that ML has. We have many reinforcement learning methods, PoMDP, HMM, etc. I think you need to explain why you select MDP instead of other methods. The intuition is needed. For instance, you assume that your states consist of 7 known states. However, in reality, the patient states are decided by both observed features and non-observed features. Therefore, PoMDP or HMM would be more general methods without assumptions and fitted to this problem. Please justify why you use MDP instead of other methods.

Response: We have added discussion on page 16 to explain why we chose to use a Markov Decision Process (MDP) rather than other reinforcement learning methods. The goal of this study was to use dynamic programming to find the optimal actions (treatment decisions) at each decision epoch based on a patient's characteristics to maximize the expected QALYs experienced over the patient's simulated lifetime. While a Hidden Markov Model (HMM) can be useful to model disease progression with unobserved (hidden) states due to its ability to recognize hidden states from sequences of observations, state transitions in HMMs are not controlled, often have a large number of unstructured parameters, and cannot express dependencies between unobserved states, which would be required when modeling CVD health states.

MDPs and Partially Observable Markov Decision Processes (POMDPs) are both useful for calculating optimal decisions from sequences of observations. The difference between these two approaches is full observability of the states. Perfect knowledge about a patient's states under MDP is a strong assumption, but we chose to use MDP since it is often tractable to solve (exact solution) and relatively easy to specify, and because CVD health states are objectively observable by clinicians under nearly all circumstances (unlike with some other diseases). POMDPs are more complex than MDPs and are often computationally intractable to solve, and thus we did not choose to use a POMDP. In the Discussion section, we mention the potential use of POMDP in future research.

4. Etc.

i) I think the action set is too small. There are various medications; however, you treat them all the same. (Only divide them into the number of medications and the dosage) I think in the future, you should consider the differences between medications.

Response: We have clarified that the action set is meant to reflect current clinical knowledge of blood pressure treatment, in which there are limited differences between medications in their effect on SBP (outside of beta-blockers, as discussed below) with the only substantial differences between medications being adverse events. We have added a statement in Appendix eText2 that the blood pressure lowering effect of the drugs was estimated as a function of treatment dose and pretreatment blood pressure, and was not dependent on types of drugs, except for beta-blockers. Beta-blockers have been shown to be more effective than other antihypertensive drugs included in our study in lowering blood pressure among patients with previous history of CVD.1 In our model, beta-blockers tended to be favored over other types of drugs for patients with CVD history due to their added benefit. Also, we highlight on page 8 and in Appendix eTable 1 that the probability of severe adverse events was determined by blood pressure drug choice and dosage.

ii) You learned your models based on the high-quality meta-analytic data. You also said that some of them are randomized trials. Are all of them are randomized trials? Otherwise, you need to explain how to deal with treatment selection bias.

Response: We now clarify on page A5 in Appendix eText 2 that we used a meta-analysis of 147 randomized placebo controlled trials of blood pressure-lowering drugs in fixed dose.1 In other words, all of the data are from randomized trials.

iii) As it can be seen in Figure 1, JNC8 and Intensive JNC8 is interpretable. However, your model is not interpretable. In other words, your MDP model is a black box and doctors cannot justify why I should use this medication dosage. To convince doctors, you should visualize your MDP model. I know this is hard to make like JNC8 decision tree; however, you need to provide the reason to doctors that why MDP make this decision.

Response: We now have added more explanation in Appendix eText1 and eFigure 2 to better illustrate how MDP chooses the optimal actions. We highlight the limitations of MDP modeling in the Discussion section: existing clinical guidelines are typically easy for clinicians to interpret because they involve univariate decisions; to make use of the MDP-based approach, clinicians would need to shift conceptually from a univariate to a multivariate decision process as well as accept computationally complex "black box" results (as they do currently with some imaging and pathological diagnostic guidelines).

Reference

1. 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.

VERSION 2 – REVIEW

REVIEWER	Muge Capan, PhD Christiana Care Health System
REVIEW RETURNED	30-Aug-2017

GENERAL COMMENTS	I would like to thank the authors for responding to reviewers' comments and revising the manuscript. My general comments regarding the literature section, contribution of the study from a clinical and IE/methodology perspective, MDP cycle length, state and action space clarification, and consistent use of terminology are addressed in the revised manuscript. The manuscript would benefit from implementing following minor edits. • Personalization: It is stated that "By treating based on individual risks, the MDPT strategy tended to treat patients more intensively earlier in life and less intensively at older ages", thus the main difference between the MDPT and existing guidelines is earlier intensive treatment and lower doses at later stages in life preventing more intensive treatment later in life and saving QALYs. The personalization contribution of the modeling approach would benefit from targeted simulation and discussion of further personalization aspects mentioned earlier in the paper (gender, race, comorbidities,..) but not reflected in the results. • Clinical implications: In the revised Discussion section on Pg 15 it is added that "From a clinical perspective, our analysis advances the literature beyond the decision of whether to treat high blood pressure because both type and dosage of medication are critical decision points." This statement indicates that existing literature mainly focused on a binary treat/don't treat type of decision rather than type and dosage which may not necessarily be correct. It would help to give examples from existing blood pressure treatment decision literature/systematic-reviews to support this statement.

REVIEWER	Jinsung Yoon University of California, Los Angeles, CA, USA No Competing Interest
REVIEW RETURNED	31-Aug-2017

GENERAL COMMENTS	1. Sampling process - I understand that you follow the guidelines to sample the simulated data. Furthermore, you used MC sampling and try to use multivariate sampling with copula functions. However, what I want to see is the "multivariate statistical tests" that can verify that your generated samples are not statistically different with the real patient data. Furthermore, eFigure 3-4 verify your results and not verify your sampling. 2. Okay. I understand this point. - However, I concerned that even if this small gain is common in CVD event, it is not enough impacts to replace the guidelines that are verified for a long time. - The proposed model has less interpretability than the guidelines. Therefore, I am not sure the advantages (small performance gain) is larger than the disadvantages (less interpretability). 3. I can see the intuitions that you use MDP. - However, MDP model needs a strong assumption (perfect knowledge about the patient's states) and it is not often true. Therefore, POMDP or HMM are more practical methods to modeling this problem. - Even if HMM and POMDP are more complex and hard to solve, it can be applied with much weaker assumption. Therefore, I think it is worth to add the results with POMDP and HMM in addition to the MDP. (Not a future work, but in this paper) - I agreed with less interpretability of the HMM. However, if the performance gain of HMM much larger, it may be worth than MDP.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Muge Capan, PhD
Institution and Country: Christiana Care Health System, USA
Competing Interests: None declared

I would like to thank the authors for responding to reviewers' comments and revising the manuscript. My general comments regarding the literature section, contribution of the study from a clinical and IE/methodology perspective, MDP cycle length, state and action space clarification, and consistent use of terminology are addressed in the revised manuscript. The manuscript would benefit from implementing following minor edits.

1) Personalization: It is stated that "By treating based on individual risks, the MDPT strategy tended to treat patients more intensively earlier in life and less intensively at older ages", thus the main

difference between the MDPT and existing guidelines is earlier intensive treatment and lower doses at later stages in life preventing more intensive treatment later in life and saving QALYs. The personalization contribution of the modeling approach would benefit from targeted simulation and discussion of further personalization aspects mentioned earlier in the paper (gender, race, comorbidities,..) but not reflected in the results.

Response: In addition to the previous edits on page 5 and 11 in the main text and eFigure 5 in Appendix, we have now added further clarifications to page 7 in the main text and eFigure 3 in Appendix to reveal the other aspects of personalization included in the model. We specifically note that while one of our key results is that treatment differs from current guidelines based on the age of the patient (with younger patients treated more intensively by the MDP than by current guidelines), the personalized treatment regimen is also determined by other patient features, including sex, race, and CVD-related covariates, such as co-morbid hyperlipidemia (eFigure3 in the Appendix).

2) Clinical implications: In the revised Discussion section on Pg 15 it is added that “From a clinical perspective, our analysis advances the literature beyond the decision of whether to treat high blood pressure because both type and dosage of medication are critical decision points.” This statement indicates that existing literature mainly focused on a binary treat/don’t treat type of decision rather than type and dosage which may not necessarily be correct. It would help to give examples from existing blood pressure treatment decision literature/systematic-reviews to support this statement.

On page 15, we have changed the term “literature" to “current guidelines”, which recommend a treat-to-target approach for large categories of people. We now cite the key current guidelines for blood pressure treatment decision-making to support this claim.

Reviewer: 3
Reviewer Name: Jinsung Yoon
Institution and Country: University of California, Los Angeles, CA, USA
Competing Interests: None

1) Sampling process
- I understand that you follow the guidelines to sample the simulated data. Furthermore, you used MC sampling and try to use multivariate sampling with copula functions. However, what I want to see is the "multivariate statistical tests" that can verify that your generated samples are not statistically different with the real patient data. Furthermore, eFigure 3-4 verify your results and not verify your sampling.

Response: We have now added results from statistical tests comparing the simulated population to the nationally-representative weighted NHANES population on page A5 and A30 (eTable 13) in the Appendix. These results reveal that the generated samples are not statistically different from real patient data.

2) Okay. I understand this point.
- However, I concerned that even if this small gain is common in CVD event, it is not enough impacts to replace the guidelines that are verified for a long time.
- The proposed model has less interpretability than the guidelines. Therefore, I am not sure the advantages (small performance gain) is larger than the disadvantages (less interpretability).

Response: We have now added further discussion on page 16-17 in the limitations section to note that one of the largest ongoing debates in personalized medicine is the fact that model-based personalization may offer incremental gains beyond large standardized guidelines, but at the risk of

producing a “black box”, and that a larger discussion must take place in the literature about the benefits and risks of such model-based personalization before such measures are considered for practice. We specifically suggest that the next logical step for research is to perform a head-to-head randomized pilot trial to compare usability, interpretability, patient and provider reactions, and patient safety of current guidelines versus our personalized blood pressure treatment selection tool. We also note that the result of a small gain in QALYs is expected in the case of blood pressure treatment because only a small subset of patients experience a CVD event; this gain, although small, is cost-saving when judged from a societal perspective as recommended by current cost-effectiveness analysis guidelines.

3) I can see the intuitions that you use MDP.
- However, MDP model needs a strong assumption (perfect knowledge about the patient's states) and it is not often true. Therefore, POMDP or HMM are more practical methods to modeling this problem.
- Even if HMM and POMDP are more complex and hard to solve, it can be applied with much weaker assumption. Therefore, I think it is worth to add the results with POMDP and HMM in addition to the MDP. (Not a future work, but in this paper)
- I agreed with less interpretability of the HMM. However, if the performance gain of HMM much larger, it may be worth than MDP

Response: In addition to our previous response below,
“We have added discussion on page 16 to explain why we chose to use a Markov Decision Process (MDP) rather than other reinforcement learning methods. The goal of this study was to use dynamic programming to find the optimal actions (treatment decisions) at each decision epoch based on a patient’s characteristics to maximize the expected QALYs experienced over the patient’s simulated lifetime. While a Hidden Markov Model (HMM) can be useful to model disease progression with unobserved (hidden) states due to its ability to recognize hidden states from sequences of observations, state transitions in HMMs are not controlled, often have a large number of unstructured parameters, and cannot express dependencies between unobserved states, which would be required when modeling CVD health states.
MDPs and Partially Observable Markov Decision Processes (POMDPs) are both useful for calculating optimal decisions from sequences of observations. The difference between these two approaches is full observability of the states. Perfect knowledge about a patient’s states under MDP is a strong assumption, but we chose to use MDP since it is often tractable to solve (exact solution) and relatively easy to specify, and because CVD health states are objectively observable by clinicians under nearly all circumstances (unlike with some other diseases).”

We have cited further literature on page 17 to demonstrate that this is the case, as there is strong evidence for this assumption in the cardiovascular disease literature. POMDPs are more complex than MDPs and are often computationally intractable to solve, and thus we did not choose to use a POMDP. In the Discussion section, we mention that potential evaluation of POMDP would be a future research and would pose an entirely new investigation/enterprise that would merit its own paper.

VERSION 3 – REVIEW

REVIEWER	Muge Capan, PhD Christiana Care Health System USA
REVIEW RETURNED	08-Oct-2017

GENERAL COMMENTS	I would like to thank the authors for their responses and thorough revision following the reviewers' recommendations.

REVIEWER	Jinsung Yoon University of California, Los Angeles
REVIEW RETURNED	10-Oct-2017

GENERAL COMMENTS	1. Sampling process. - I think you mentioned eTable 2 and not eTable 13. - However, eTable2 shows that each variable is not statistically different (Uni-variate analysis) because you compare each variable separately (the evidence is that p-value is different for each variable). - Therefore, it is not a multi-variate statistical test. Making data with not statistically different (verified by uni-variate statistical tests) is easy. However, the important thing is the data is not statistically different (verified by multi-variate) - I think I said this three times in a row ("multi-variate statistical test") but it is not addressed unfortunately on three rebuttals. 2. 3. Okay. I understand this point.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Muge Capan, PhD
Institution and Country: Christiana Care Health System, USA
Competing Interests: None declared

Comment: I would like to thank the authors for their responses and thorough revision following the reviewers' recommendations.

Response: Thank you.

Reviewer: 3
Reviewer Name: Jinsung Yoon
Institution and Country: University of California, Los Angeles, USA
Competing Interests: Not declared

1. Sampling process.
- I think you mentioned eTable 2 and not eTable 13.
- However, eTable2 shows that each variable is not statistically different (Uni-variate analysis) because you compare each variable separately (the evidence is that p-value is different for each variable).

- Therefore, it is not a multi-variate statistical test. Making data with not statistically different (verified by uni-variate statistical tests) is easy. However, the important thing is the data is not statistically different (verified by multi-variate)
- I think I said this three times in a row ("multi-variate statistical test") but it is not addressed unfortunately on three rebuttals.

Response: We have edited the statement on page A5 in the Appendix that we used a multivariate statistical test (MANOVA) comparing the simulated population to the nationally-representative weighted NHANES population, and cited a literature. Also, we added results and summary outputs from a MANOVA testing on page A25 (eTable 2) in the Appendix. These results reveal that the generated samples are not statistically different from real population data.

2. 3. Okay. I understand this point.

Response: Thank you.