Cost-Effectiveness Analysis of a Personalized, Teleretinal-Inclusive Screening Policy for Diabetic Retinopathy via Markov Modeling

Poria Dorali[1], BS; Zahed Shahmoradi[2], PhD; Christina Y. Weng[3,4], MD, MBA; Taewoo Lee[5], PhD

[1] Department of Industrial Engineering, University of Houston, Houston, Texas

[2] Center for Health Services Research, Department of Management, Policy, and Community Health, UTHealth School of Public Health, Houston, Texas

[3] Department of Ophthalmology, Ben Taub Hospital, Houston, Texas

[4] Department of Ophthalmology, Baylor College of Medicine, Houston, Texas

[5] Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania

Financial support: This work is supported by NSF Grants CMMI#1907933 and CMMI#1908244. The sponsor had no role in the design or conduct of this research.

Conflict of interests: No conflicting relationship exists for any author.

Running head: Cost-Effectiveness of Personalized DR Screening

Corresponding author: Christina Y. Weng, MD, MBA
Phone: 7137986100,
Email: Christina.Weng@bcm.edu
Address for reprints: Department of Ophthalmology, Baylor College of Medicine, 1977 Butler Blvd., Houston, TX 77030, United States

Abbreviations/Acronyms: DR – Diabetic retinopathy, QALY – Quality-adjusted life year, ICER – Incremental cost-effectiveness ratio, ADA – American Diabetes Association, AAO – American Academy of Ophthalmology, TRI – Teleretinal imaging, AI – Artificial intelligence, POMDP – Partially observable Markov Decision Process, DME – Diabetic macular edema, NIN-DR - Non-intervention-needed DR, IN-DR – Intervention-needed DR, PT – Post-treatment, BL – Blindness, DE – Death, NPDR – Non-proliferative DR, PDR – Proliferative DR, WW – Wait and watch, CS – Clinical screening, WTP – Willingness to pay, HHS – Harris Health System, EMR – Electronic medical record, IRB – Institutional review board, n – Number.

Key Words: Markov, Modeling, Simulation, Cost-Effectiveness, Personalization

Abstract

Purpose: While teleretinal imaging has proved effective in increasing population-level screening for diabetic retinopathy (DR), there is a lack of quantitative understanding on how to incorporate teleretinal imaging into existing screening guidelines. We develop a mathematical model to determine personalized DR screening recommendations that utilize teleretinal imaging and evaluate the cost-effectiveness of the personalized screening policy.

Design: A partially observable Markov decision process is employed to determine personalized screening recommendations based on patient compliance, willingness to pay, and A1C level. Deterministic sensitivity analysis was conducted to evaluate the impact of patient-specific factors on personalized screening policy. The cost-effectiveness of identified screening policies was evaluated via hidden-Markov chain Monte Carlo simulation on a data-based hypothetical cohort.

Participants: Screening policies were simulated for a hypothetical cohort of 500000 patients with parameters based on literature and electronic medical records (EMR) of 2457 patients who received teleretinal imaging in 2013–2020 from the Harris Health System (Harris County, Texas).

Methods and Intervention: Population-based mathematical modeling study. Interventions included dilated fundus examinations referred to as clinical screening, teleretinal imaging, and wait-and-watch recommendations.

Main Outcome Measures: Personalized screening recommendations based on patient-specific factors. Accumulated quality-adjusted life years (QALYs) and cost (USD) per

patient under different screening policies. Incremental cost-effectiveness ratio (ICER) to compare different policies.

Results: For the base cohort, on average, teleretinal imaging was recommended 86.69% of the time over each patient's lifetime. The model-based personalized policy dominated other standardized policies, generating more QALY gains and cost savings for at least 57% of the base cohort. Similar outcomes were observed in sensitivity analyses of the base cohort and the Harris Health-specific cohort and rural population scenario analysis.

Conclusions: A mathematical model was developed as a decision support tool to identify a personalized screening policy that incorporates both teleretinal imaging and clinical screening and adapts to patient characteristics. Compared to current standardized policies, the model-based policy significantly reduces costs while performing comparably, if not better, in terms of QALY gain. A personalized approach to DR screening has significant potential benefits that warrant further exploration.

Introduction

Diabetic retinopathy (DR) is the leading cause of blindness in working-age US adults.[1,2] Over 60% of patients with type II diabetes mellitus and almost all patients with type I diabetes mellitus will develop DR within 20 years of diagnosis.[1] Among adults over the age of 45 diagnosed with diabetes, 8.6% have DR while 4.1% have experienced vision loss due to DR.[3] Currently, over 37 million US adults have diabetes and over 96 million US adults have prediabetes.[2] Given the epidemic of diabetes, there is a growing concern about the associated increase in DR cases and potential vision loss.

Timely screening for DR is one of the most cost-effective tools for mitigating DR-related vision loss[4–6]; studies show that up to 98% of DR-related vision loss cases can be prevented by early detection and treatment.[1,7] Currently, the American Diabetes Association (ADA) and American Academy of Ophthalmology (AAO) recommend annual comprehensive eye screening examinations for every diabetic patient,[1,7,8] based on the 7-field stereoscopic color fundus photography.[9] However, only 50–65% of diabetic patients are screened on a yearly basis in the US,[10] with even lower compliance for patients with limited access or low socioeconomic status.[11–12]

Recently, teleretinal imaging (TRI) has emerged as an effective alternative to conventional DR screening. TRI involves images captured via single-field nonmydriatic monochromatic digital photography typically obtained in a non-eye care setting and remotely graded.[9] Studies have illustrated the potential of TRI to alleviate barriers to DR screening due to its lower cost and improved accessibility.[13,14] The Centers for Medicare and Medicaid Services recognizes TRI as an acceptable and reimbursable service.[15] Large-scale healthcare systems in the US, such as the Veterans Health Administration, have

provided TRI services since the early 2000s.[16] TRI is expected to be even more widely used as its accuracy and convenience further improve with artificial intelligence (AI)-based interpretation.[17]

However, there is a lack of quantitative understanding on how often TRI screening should be recommended to diabetic patients with different characteristics. Currently, many TRI programs recommend annual TRI follow-ups, although optimal intervals of TRI may vary based on the patient's compliance, socioeconomic status, severity of DR, age, duration of disease, and glycemic control. With the emergence of more accessible TRI-based screening examinations, development of a personalized, patient-specific screening approach may be advantageous for patients, providers, and society.

In this study, we develop a proof-of-concept mathematical model based on a previously validated Markov model to generate personalized DR screening recommendations. Unlike previous studies that examined standardized screening guidelines with fixed screening modality and interval,[16,18] we examine the usefulness of the model-based, personalized policy in exploiting dynamic interaction of TRI and traditional screening based on patient-specific factors. The aim of this study is to evaluate the cost-effectiveness of the personalized screening policy and demonstrate the potential of this modeling framework as a clinical decision support tool that can be used in both primary and eye care settings.

Methodology

POMDP model

We develop a partially observable Markov decision process (POMDP) model to identify personalized DR screening policies. The POMDP model is built on a previously validated Markov-based DR natural history model.[6] Markov models have been widely used in ophthalmology for predicting disease progression and analyzing cost-effectiveness of various screening and intervention policies: DR screening and treatment,[16] diabetic macular edema (DME) treatment,[19] and glaucoma screening.[20] Other simulation models such as decision trees have also been recently used for cost-effectiveness analyses in ophthalmology, including those of injection treatments for DR and DME[21] and AI-based screening for retinopathy of prematurity.[22]

Figure 1 shows the Markov model with DR-related health states and transitions. The health states in the model include non-intervention-needed DR (NIN-DR), intervention-needed DR (IN-DR), post-treatment (PT), blindness (BL), and death (DE). The NIN-DR state is defined to include no DR and non-proliferative DR (NPDR) while the IN-DR state is defined as proliferative DR (PDR). Note that unless the patient is in the PT, BL, or DE state, the model often cannot specify exactly which health state the patient belongs to. Thus, the POMDP model uses the notion of a "belief" state: instead of assuming a specific, fully observable disease state as in traditional Markov models, a belief state probabilistically represents a patient's health state. For example, a patient's health state can be represented as 90% in NIN-DR and 10% in IN-DR. The belief state structure is particularly relevant in our study because asymptomatic earlier stages of DR, potential inaccuracy of TRI outcomes, and patient non-compliance can make the patient's state uncertain or only

partially observable.[23] The belief state is updated semi-annually via Bayesian inference based on the patient's compliance, recent screening outcomes and accuracy, and DR natural progression. More details about the update mechanism and examples can be found in Appendix S1 (available online at www.ophthalmologyretina.org).

For each 6-month time period and for each belief state, the model recommends one of the three actions—wait and watch (WW), TRI, or traditional screening (referred to as clinical screening or CS)—until a patient progresses to PT, BL, or DE and ceases screening. Whenever a TRI outcome indicates IN-DR, the patient is referred for follow-up CS immediately. Whenever a CS outcome identifies IN-DR, the model immediately sends the patient for treatment, i.e., the PT state.

The POMDP model chooses recommendations such that the total accumulated discounted reward is maximized, where the reward is defined as the quality-adjusted life years (QALYs) multiplied by the willingness to pay (WTP) factor minus out-of-pocket costs (USD) associated with each state-recommendation pair. A yearly discount factor of 3% was used as is traditionally recommended in healthcare modeling.[24] The WTP value was treated as a patient-specific parameter that changes based on socioeconomic status.

Data used

A summary of the parameters for the base case POMDP model can be found in Table 1. Values for TRI accuracy, screening and treatment costs, and TRI and follow-up CS compliance rates are based on the Harris Health System (HHS) TRI program data from the time period of 2013–2020. Institutional Review Board (IRB) approval was obtained to view and analyze patient medical records. HHS is the largest safety-net hospital system in Harris County which encompasses the Houston metropolitan area and represents the

third most populous county in the US. In total, 2457 patient profiles were collected from the HHS electronic medical record (EMR) database, including race/ethnicity, age, zip-code location, A1C-level, insurance status, TRI screening outcomes over time, and TRI adherence rate. This patient population was composed of 59.23% Hispanic, 22.71% African American, 5.63% Asian, 3.91% white and 8.52% other. The average age of this cohort was 54.62 years at the time of first DR screening. The average A1C level was 7.91% and initial prevalence of IN-DR was 2.32%. A classification decision tree identified a patient's race to be a significant factor for yearly TRI compliance rate, with the following average race-specific compliance rates: 68% for Hispanic patients, 65% African American, 84% Asian, 68% White and 73% other.

While the previous Markov models for DR progression assume annual transition probabilities,[16,18] we use semi-annual transition probabilities found via a cycle-length conversion technique[28] to allow the model to generate more frequent recommendations, if needed. The same conversion technique was used to produce the semi-annual discount rate and age-dependent mortality rates.[29] Annual values for QALYs and direct blindness costs[26] were divided in half and allocated at each semi-annual time period as needed. All screening costs were per-visit costs while treatment costs were one-time costs estimated by a weighted average of the costs of different types of DR treatments taken by the HHS cohort.[13] For CS and follow-up CS, an additional $21.54 was added per visit to account for travel costs and lost wages.[13] The semi-annual model was validated via simulation where both the semi-annual and previously validated annual Markov models generated a similar number of blindness cases and patient lifetimes for the same patient cohort. Finally, a

baseline WTP value of $10000 per QALY was determined to represent patients at a safety-net system most of whom are medically underserved with budgetary restrictions.

Base case simulation and sensitivity analysis

We used hidden-Markov chain Monte Carlo simulation to evaluate the cost-effectiveness of the POMDP-based screening policy for 500000 hypothetical patients generated from the HHS data, set to start first DR screening at age 40 within the simulation. This hypothetical patient population forms the base case simulation cohort to represent the maximum number of potential diabetic patients in Harris County eligible for the TRI service at HHS. Each patient was tracked up until death or the age of 99 and accumulated QALY gains and out-of-pocket costs were recorded. The initial health condition of the hypothetical cohorts was probabilistically assigned based on DR prevalence in the literature.[18]

We conducted deterministic one-way sensitivity analysis on the POMDP-based DR screening policies. The scenarios selected included (1) increasing A1C level to 13%, (2) decreasing WTP to $5000, (3) CS compliance rates of 20%, and (4) CS compliance rate of 50%, which were chosen to represent realistic scenarios where the model can be useful for making personalizing recommendations for (1) patients with poor glycemic control, (2) patients with severe financial barriers, and (3)-(4) patients with significantly different CS compliance behaviors. For each scenario, the updated POMDP model regenerated a new personalized screening policy, which was then evaluated under the same hidden-Markov chain Monte Carlo simulation framework for the same cohort of 500000 patients.

Cost-effectiveness analysis

The personalized policy was benchmarked against other standardized policies: annual CS, annual TRI, biennial CS and WW. All standardized policies were simulated using the hidden-Markov chain Monte Carlo simulation framework for the same hypothetical patient cohort with base case parameters. Incremental cost-effectiveness ratio (ICER) analysis was conducted to examine the cost-effectiveness of the POMDP-based policy against each of the standardized policies. Specifically, we examined the difference in direct patient costs between the model-based policy and each standardized policy divided by the difference in QALY gains, that is, the average incremental cost for 1 additional QALY. To visualize the cohort-level policy comparisons, we divided each cohort of 500000 patients into 200 subgroups of 2500 patients. Sensitivity analysis was conducted by comparing the personalized policy to the standardized policies under the following cohort cases: 13% A1C level and 20% and 50% CS compliance rates.

To further demonstrate the impact of the personalized policy in a clinical setting, we created a new cohort of 500000 patients that mimics the composition of the patient cohort extracted from the HHS EMR data proportionally in terms of race, TRI compliance rate, A1C level at the beginning of DR screening, and age at the beginning of DR screening. Each hypothetical patient retained a unique disease and screening trajectory. The personalized policy was compared to each of the standardized policies for this HHS-specific cohort. Cost per QALY was examined for each policy across different age groups of this cohort.

As an extension to the cost-effectiveness analysis, a hypothetical rural cohort was constructed to examine the generalizability of the model to populations with limited

geographic access to eye care. Geographic access to care and transportation burdens are typically considered factors that drive poor compliance with CS-based screening.[30,31] To model this cohort, we assumed annual CS compliance rate to be randomly generated between 12% and 45% based on previous studies,[32,33] and CS-based transportation costs were increased to $60.49 per visit.[34] For a rural scenario with TRI service assumed to be available, TRI compliance rate was set to either 40%, 60%, or 80%. IN-DR prevalence was set to 2.3% based on previous studies.[35] All other parameters such as the starting age, A1C, and WTP values were set identical to those of the base case cohort analysis. The POMDP model was run for the rural cohort to generate the scenario-specific screening policies which were used for the cost-effectiveness analysis.

Results

POMDP-based recommendations

The POMDP model generated threshold-based screening recommendations based on the Bayesian-updated probability that the patient is in the IN-DR state. Table 2 displays the threshold-based recommendations under five selected age levels. Each recommendation region represents the range of probabilities of being in the IN-DR state for which the model identifies either WW, TRI, or CS to be optimal. Once the model was run over each patient's lifetime, the number of each recommendation type was recorded; on average, TRI, CS, and WW were recommended 86.7%, 9.6%, and 3.7% of the time, respectively, over the lifetime of the base case cohort. The mean age of death was 72.76 (95% CI 72.74 to 72.78). For the base case cohort, between the ages of 40 to 73, WW was never recommended. As patients aged, the threshold between TRI and CS (i.e., the upper endpoint for the TRI recommendation region) increased, which in turn increased the proportion of TRI recommendations (from 89.9% at age 50 to 98% at age 70; see Table 2). WW became included as a potential recommendation once a patient reached 74 years of age and became the only available recommendation at ages 84 and older.

Sensitivity analysis

Sensitivity analysis considered four additional hypothetical cohorts; Table 2 shows the simulation results. First, for the cohort with an increased A1C level of 13%, the TRI-to-CS thresholds were found lower than those in the base case (7% A1C) for age groups of 50 and 60. As a result, the proportion of TRI recommendations decreased by 22% overall compared to the base case patients, while CS recommendations increased overall by

25.5%. Decreasing WTP from $10000 to $5000 led to small increases in the TRI-to-CS thresholds, which increased the proportion of TRI recommendations by 1.1% overall. Notably, decreasing CS compliance from the base value 35% to 20% removed CS as an optimal recommendation altogether while greatly increasing the overall proportion of TRI recommendations, from 86.7% to 96.5%. When CS compliance increased from 35% to 50%, the TRI-to-CS thresholds slightly decreased at ages 50, 60 and 70, resulting in a slight increase in CS recommendations.

Cost-effectiveness analysis

We examined the cost-effectiveness of the POMDP-based personalized policy against four alternative standardized policies: annual CS, annual TRI, biennial CS, and WW. Table 3 displays the cost-effectiveness of the personalized policy against each of the standardized policies under the base case cohort as well as other sensitivity analysis cases: 13% A1C level, and 20% and 50% CS compliance rates, along with the ICER values when available. Each comparison reflects analysis based on the cohort of 500000 patients divided into 200 subgroups of 2500 patients. For the base case comparison, the personalized policy was found dominant for most patients (57% vs Annual CS, 66% vs Annual TRI, 80% vs Biennial CS, and 100% vs WW). Similar outcomes were found across sensitivity analysis, except for the 50% CS compliance case where POMDP policy was dominant to ACS for 43% of the cohort. Overall, the personalized policy provided more QALY gains compared to all other standardized policies for most of the simulated cohort. Except for the comparison to annual CS in the 50% CS compliance case, the personalized policy resulted in at least 92.5% of each simulated cohort experiencing cost savings. In 10 out of 16 test cases, the POMDP-based policy provided cost savings for the entire

simulated cohort. Additionally, the ICER value, whenever applicable, was low for the personalized policy with the maximum of $3074.87.

Figure 2 displays the differences in accumulated costs and QALYs between the personalized policy and each of the standardized policies for the HHS-specific cohort. Each data point in Figure 2 represents a subgroup of 2500 patients. Most data points were found to have positive x-axis values and negative y-axis values across all policy comparisons (i.e., within the fourth quadrant), which indicates that the personalized policy was dominant for most patients in terms of cost savings and QALY gains for each comparison. The personalized policy performed best against WW, followed by biennial CS, annual TRI and annual CS. The contour for each comparison represents the 95% confidence region for each set of 200 data points. Figure 3 examines the out-of-pocket cost paid per QALY for the HHS-specific cohort separated by age group (40 to 50, number of patients (n) = 159009; 51 to 60, n = 191525; 61 to 70, n = 124431; 71 to 80, n = 22703; 80+, n = 2332). The 95% confidence intervals were within 0.9% of the reported means for patients within ages 40-70, 2.6% within ages 71-80 and 10% for 80+. Across all age groups, the personalized policy produced the least costs paid per QALY (40 to 50, $847; 51 to 60, $576; 61 to 70, $401; 71 to 80, $241; 80+, $106). The benefit produced by the personalized policy was more pronounced for younger patients. The maximum cost differences were $652, $364, $188, $41and $59 for the 40-to-50, 51-to-60, 61-to-70, 71-to-80, and 80+ age groups, respectively.

Table 4 shows the cost-effectiveness outcomes for the rural cohort. The POMDP-based, personalized policy was found dominant in terms of both QALY gain and cost savings for most of the simulated cohorts (at least 54% vs Annual CS, 63% vs Annual TRI, 82% vs

Biennial CS, and 100% vs WW). The dominance of the personalized policy against the annual CS policy increased from 54% to 70% as TRI compliance increased from 40% to 80%. The dominance against the annual TRI policy increased from 72% to 73% as TRI compliance increased from 40% to 60%, but decreased from 73% to 63% when TRI compliance increased to 80% due to the increased similarity between the personalized policy and annual TRI. Compared to the base case cohort representing urban safety-net system patients, the personalized policy relied more on TRI-based screening for the rural cohorts with 60% and 80% TRI compliance.

Discussion

In this study, we developed a proof-of-concept mathematical model via POMDP for personalized DR screening recommendations that utilize both TRI and traditional screening exams. Personalization of screening recommendations that include TRI is particularly important for ethnic minorities who are known to be generally less compliant with traditional screening exams as well as patients with limited access to ophthalmic care.[10,11,14] The POMDP model utilizes the DR natural history model to periodically update the risk of intervention-needed DR which helps prevent delayed detection through timely TRI-based screening recommendations. The model-based policy was shown to produce QALYs higher than or comparable to those achieved by other standardized screening policies at substantially lower costs. This suggests that integration of TRI into existing screening policies in a personalized manner may be beneficial and a model-based policy may have utility as a decision support tool.

The findings of this study are consistent with previous studies on the cost-effectiveness of TRI screening in that TRI-based screening policies are more cost-effective than standardized screening that only utilizes traditional eye clinic-based screening (Tables 3 and 4).[16,18] Our analysis shows diminishing health and cost-saving benefits of DR screening as a patient ages, which is also consistent with the findings in the cost-effectiveness study for DR screening at the Veterans Health Administration.[16] Our study shows that cost-effectiveness can further improve when the utilization of TRI is personalized based on patient-specific factors and compliance behavior, compared to the standardized TRI policies evaluated in the literature.[16,18]

297 The model-based policy shows that more frequent DR screening is beneficial for
298 patients at higher risk of sight-threatening DR, which agrees with state-dependent
299 screening frequency proposed by the International Council of Ophthalmology
300 guidelines.[36] In our analysis, the severity-based screening policy is further refined by
301 enabling TRI recommendations, where TRI screening is more frequently recommended to
302 patients who are younger or less compliant with clinical screening. Also, the modeling
303 framework offers a generalization to the existing state-based guidelines by accounting for
304 DR state uncertainty and patients' non-compliance via the concept of belief state.
305 While the POMDP model identified screening recommendations autonomously, the
306 model always provided recommendations in the order of WW, TRI, then CS as the
307 patient's probability of being in the IN-DR state increased, forming threshold-based
308 decisions. This suggests that the model considers CS for the most at-risk patients while
309 recommending TRI or WW for those less at risk.
310 The most prominent factor affecting personalized recommendations was low CS
311 compliance. For example, the model did not directly recommend CS for a patient group
312 with 20% CS compliance rate. Instead, CS only occurred after a positive TRI screening
313 outcome. Another factor that increased TRI recommendation frequency was decreased
314 WTP (from $10000 to $5000). In this case, the TRI-to-CS threshold value increased (Table
315 2), which implies the model will wait for a higher risk of IN-DR to make a direct CS
316 recommendation. These observations reinforce the importance of TRI availability in
317 settings where patients have poor screening compliance and limited access to care. When
318 A1C was increased from 7% to 13%, the number of CS recommendations increased
319 substantially, demonstrating how the model can adapt recommendations to increased risk

of developing IN-DR. Interestingly, when CS compliance increased to 50% the model made little shift in recommendations. We believe this was because the model became more "confident" that the patient would attend CS appointments and thus continued to rely on TRI.

The POMDP model provides a flexible framework that can adapt to a patient's unique characteristics and behaviors. In practice, the modeling framework can be utilized by both primary and eye care professionals with updated parameters to match the underlying cohort; for example, historical compliance rates, recent screening results, and A1C levels can be retrieved from the patient's medical record or upon an initial screening of a patient and updated as needed. Additionally, the modeling framework can accommodate different DR state definitions without fundamental changes in the model. For example, more specific states such as no DR, moderate, mild, and severe NPDR, and PDR can be used instead of the two states in the current model and states associated with DME can be added as well, with appropriate state transition probabilities.

While the base POMDP model was calibrated based on the data from a large urban safety-net system, the modeling framework is generalizable to different cohorts in other locations by inputting specific system-level information (e.g., TRI sensitivity and specificity) and cohort-specific variables such as each patient's age, DR status, A1C level, and, importantly, factors that affect screening compliance. To demonstrate this generalizability, scenario analysis was conducted on a hypothetical cohort of rural diabetic patients with limited geographic access to DR screening. CS compliance rates in the rural population were assumed to be lower than in the urban population based on previously published literature demonstrating that rural patients generally experience

343 greater transportation barriers such as increased transportation costs and driving
344 distance as well as financial burden compared to other subgroups.[30,31,37-39]

345 For the rural cohort, the POMDP model identifies TRI as a preferred screening
346 recommendation for most patients and the TRI-based personalized screening policy was
347 found cost-effective, dominating other standardized screening policies in terms of health
348 benefit and cost savings (Table 4). As TRI became more accessible to rural patients, as
349 shown through the 60% and 80% TRI compliance scenarios, the cost-effectiveness
350 dominance of the personalized policy increased against annual CS (Table 4). This
351 observation reinforces the importance of a careful design and implementation of a TRI
352 program in a rural area to assure a high participation rate. The POMDP model can capture
353 differences in economic level and compliance between different rural cohorts by adjusting
354 the WTP factor and compliance parameters, respectively, as needed.

355 In general, model-based scenario analysis generates important insights for
356 policymakers. For example, once cost-effectiveness of TRI-based personalized screening
357 policy is shown via the POMDP model and scenario analysis, findings can also guide and
358 optimize subsequent decisions such as the number of additional facilities across the rural
359 area and specific locations of them. Additional, properly located TRI facilities could
360 minimize transportation burden and increase screening compliance,[30,35] which in turn
361 further improves cost-effectiveness of the screening program.

362 Note that the POMDP modeling framework serves as an adjunctive clinical decision
363 support tool. A care provider is involved in the decision-making process, and final
364 recommendations will be based on the interaction between the provider and the model.
365 For example, if TRI is recommended by the model but CS is deemed more appropriate

(e.g., if PDR was already found on a previous TRI exam), the provider can manually enter the CS recommendation into the model to override the model recommendation, based on which the model then updates the patient's projected DR state and progression. A workflow schematic of the model-based screening decision-making process can be found in Figure S4 (available online at www.ophthalmologyretina.org).

There are a few limitations to this study. First, we assume a single-eye model, as is commonly found in ophthalmology literature[24] to reduce model complexity. Second, the parameters within our simulation analysis such as patient compliance rates and A1C level were assumed to be constant over time for each patient. To address this in practice, the model can be run periodically in an adaptive manner whenever such parameters need to be updated. Third, the model used a simplified grading system consisting of two states: NIN-DR and IN-DR, where IN-DR corresponds to PDR and NIN-DR includes all other DR states. While the two-state system was defined based on the current referral point used at the HHS TRI program (referral to CS occurs when PDR is found on TRI) and the model was validated, the model could be modified to have a more specific state system that includes no DR, mild, moderate, and severe NPDR, and PDR. DME-related states were not considered because the current HHS TRI program does not support DME screening. Lastly, this model focuses on a patient-level perspective as opposed to a societal perspective that might account for capital and overhead costs associated with screening infrastructure.

Important future studies are warranted to implement and evaluate this modeling framework in a clinical setting. The current model utilizes patient information including age, past DR state trajectory, TRI image readings, A1C, WTP, and socio-demographic

389 factors that affect compliance behavior. In the future, the model will be further refined to
390 benefit from more specific patient information, which will lead to even more personalized
391 recommendations. For example, as AI-based TRI interpretation and its interaction with
392 the EMR system improve, the model could be modified to directly incorporate digital
393 fundus photograph data into the automated risk update and screening decision-making
394 process. Development of a new natural history model that includes more risk factors will
395 also enable the model to accommodate more specific patient information. As our rural
396 scenario analysis shows the potential cost-effectiveness of a personalized screening policy
397 utilizing TRI, important operational and logistics decisions need to be examined such as
398 the locations of TRI facilities[40] and design of an effective mobile TRI-based screening
399 problem to increase TRI screening uptake in rural areas.

400 In conclusion, the POMDP model was developed to evaluate personalized screening
401 recommendations incorporating TRI for patients with limited access to eye care. The
402 model was first tested for cohorts generated based on the data from a large urban safety-
403 net system. We also studied the applicability of the model to rural patients with poor
404 compliance with DR screening and the cost-effectiveness of TRI-based personalized
405 screening. We envision two potential applications of the proposed modeling framework: it
406 can serve as a clinical decision support tool for care providers to promote personalized,
407 patient-centered decisions; it can also provide evidence-based support for the integration
408 of TRI into traditional DR screening guidelines. Based on our analysis, the model-based
409 screening policy not only significantly reduces the average per-patient screening cost
410 compared to other standardized policies, but also performs similarly if not superiorly in
411 terms of QALY gains. While validation studies are needed to assess the effectiveness and

412 feasibility of our proposed model in a clinical setting, the potential benefits of an
413 individualized approach to DR screening are significant and warrant future exploration.

References

1. Fong DS, Aiello L, Gardner TW, et al. Diabetic Retinopathy. *Diabetes Care.* 2003; 26(suppl1):s99–s102.

2. Centers for Disease Control and Prevention. *National Diabetes Statistics Report website.* https://www.cdc.gov/diabetes/data/statistics-report/index.html; Accessed October 15, 2021.

3. Cha AE, Villarroel MA, and Vahratian A. Eye disorders and vision loss among US adults aged 45 and over with diagnosed diabetes, 2016–2017. *NCHS Data Brief*, 344; 2019.

4. Ferris I. How effective are treatments for diabetic retinopathy?. *JAMA.* 1993;269(10): 1290-1291.

5. Javitt JC, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. *Annals of Internal Medicine.* 1996;124(1 Pt 2):164-169.

6. Vijan S, Hofer TP, Hayward RA. Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. *JAMA.* 2000;283(7):889–896.

7. Solomon SD, Chew E, Duh EJ, et al. Diabetic retinopathy: a position statement by the American Diabetes Association. *Diabetes Care.* 2017;40(3):412–418.

8. Flaxel CJ, Adelman RA, Bailey ST, et al. Diabetic retinopathy preferred practice pattern®. *Ophthalmology*. 2020;127(1):66–145.

9. Walton OB, Garoon RB, Weng CY, et al. Evaluation of automated teleretinal screening program for diabetic retinopathy. *JAMA Ophthalmology.* 2016;134(2):204–209.

10. MacLennan PA, McGwin G, Heckemeyer C, et al. Eye care use among a high-risk diabetic population seen in a public hospital's clinics. *JAMA Ophthalmology.* 2014;132(2):162–167.

11. Paz SH, Varma R, Klein R, et al. Noncompliance with vision care guidelines in Latinos with type 2 diabetes mellitus: the Los Angeles Latino Eye Study. *Ophthalmology.* 2006;113(8):1372–1377.

12. Eppley SE, Mansberger SL, Ramanathan S, et al. Characteristics associated with adherence to annual dilated eye examinations among US patients with diagnosed diabetes. *Ophthalmology.* 2019;126(11):1492–1499.

13. Garoon RB, Lin WV, Young AK, et al. Cost savings analysis for a diabetic retinopathy teleretinal screening program using an activity-based costing approach. *Ophthalmology Retina.* 2018;2(9):906–913.

14. Daskivich LP, Vasquez C, Martinez J, et al. Implementation and evaluation of a large-scale teleretinal diabetic retinopathy screening program in the Los Angeles County Department of Health Services. *JAMA Internal Medicine.* 2017;177(5):642-649.

15. Centers for Medicare & Medicaid Services. *Billing and coding: Remote imaging of the retina to screen for retinal diseases*, 2022. https://www.cms.gov/medicare-coverage-database/view/article.aspx?articleid=58914; Accessed 17th September 2022.

16. Kirkizlar E, Serban N, Sisson JA, et al. Evaluation of telemedicine for screening of diabetic retinopathy in the Veterans Health Administration. *Ophthalmology.* 2013;120(12):2604–2610.

17. Date RC, Jesudasen SJ, Weng CY, et al. Applications of deep learning and artificial intelligence in retina. *International Ophthalmology Clinics*. 2019;59(1):39–57.

18. Aoki N, Dunn K, Fukui T, et al. Cost-effectiveness analysis of telemedicine to evaluate diabetic retinopathy in a prison population. *Diabetes Care.* 2004;27(5):1095–1101.

19. Pesonen M, Kankaanpää E, Vottonen P. Cost-effectiveness of dexamethasone and triamcinolone for the treatment of diabetic macular edema in Finland: A Markov-model. *Acta Ophthalmologica*. 2021;99(7):e1146–e1153.

20. Tang, J, Liang, Y, O'Neill, C, et al. Cost-effectiveness and cost-utility of population-based glaucoma screening in China: a decision-analytic Markov model. *The Lancet Global Health*. 2019;7(7),e968–e978.

21. Hutton DW, Stein JD, Glassman AR, et al. Five-year cost-effectiveness of intravitreous ranibizumab therapy vs panretinal photocoagulation for treating proliferative diabetic retinopathy: a secondary analysis of a randomized clinical trial. *JAMA Ophthalmology*. 2019;137(12):1424–1432.

22. Morrison SL, Dukhovny D, Chan RVP, et al. Cost-effectiveness of artificial intelligence–based retinopathy of prematurity screening. *JAMA Ophthalmology*. 2022;140(4):401–409.

23. Bennett CC and Hauser K. Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach. *Artificial Intelligence in Medicine*. 2013;57(1):9–19.

24. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second Panel on Cost-Effectiveness in Health and Medicine. JAMA. 2016;316(10):1093–1103.

25. Date RC, Shen KL, Shah BM, et al. Accuracy of detection and grading of diabetic retinopathy and diabetic macular edema using teleretinal screening. *Ophthalmology Retina.* 2019;3(4):343–349.

26. Moshfeghi AA, Lanitis T, Kropat G, et al. Social cost of blindness due to AMD and diabetic retinopathy in the United States in 2020. *Ophthalmic Surgery, Lasers and Imaging Retina.* 2020;51(4):S6–S14.

27. Chamberlain PD. *Factors associated with follow-up examination following positive teleretinal screening for diabetic retinopathy in an urban healthcare system*. Dissertation. Baylor College of Medicine; 2018.

28. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Medical Decision Making.* 1993;13(4):322–338.

29. Social Security Administration. *Actuarial Life Table.* https://www.ssa.gov/oact/STATS/ table4c6.html; Accessed October 15, 2021.

30. Liu Y, Zupan NJ, Shiyanbola OO, et al. Factors influencing patient adherence with diabetic eye screening in rural communities: a qualitative study. *PLoS One*, 13(11):e0206742, 2018.

31. Soares RR, Parikh D, Shields CN, et al. Geographic access disparities to clinical trials in diabetic eye disease in the United States. *Ophthalmology Retina*. 2021;5(9):879–887.

32. Maberley D, Walker H, Koushik A, et al. Screening for diabetic retinopathy in James Bay, Ontario: a cost-effectiveness analysis. *CMAJ*. 2003;168(2):160–164.

33. Coon P, Zulkowski K. Adherence to American Diabetes Association standards of care by rural health care providers. *Diabetes Care*. 2002;25(12):2224–2229.

34. Richardson DR, Fry II R, Krasnow DO. Cost-savings analysis of telemedicine use for ophthalmic screening in a rural Appalachian health clinic. *West Virginia Medical Journal*. 2013;109(4):52–55.

35. Nathoo N, Ng M, Rudnisky CJ, et al. The prevalence of diabetic retinopathy as identified by teleophthalmology in rural Alberta. *Canadian Journal of Ophthalmology*. 2010;45(1):28–32.

36. Wong TY, Sun J, Kawasaki R, et al. Guidelines on diabetic eye care: The International Council of Ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology*. 2018;125(10), 1608–1622.

37. Harrington RA, Califf RM, Balamurugan A, et al. Call to action: rural health: a presidential advisory from the American Heart Association and American Stroke Association. *Circulation*. 2020;141(10):e615–644.

38. Hale NL, Bennett KJ, Probst JC. Diabetes care and outcomes: disparities across rural America. *Journal of Community Health*. 2010;35(4):365–374.

551 39. Fathy C, Patel S, Sternberg Jr P, et al. Disparities in adherence to screening guidelines for
552 diabetic retinopathy in the United States: a comprehensive review and guide for future
553 directions. *Seminars in Ophthalmology*. 2016;31(4):364–377.
554
555 40. Dorali P, Limongi RGC, Weng CY, and Lee T. Location-based analysis on optimizing
556 teleretinal imaging accessibility in a large safety-net system. *Investigative Ophthalmology &*
557 *Visual Science*. 2021;62(8):2652–2652.

Figure captions

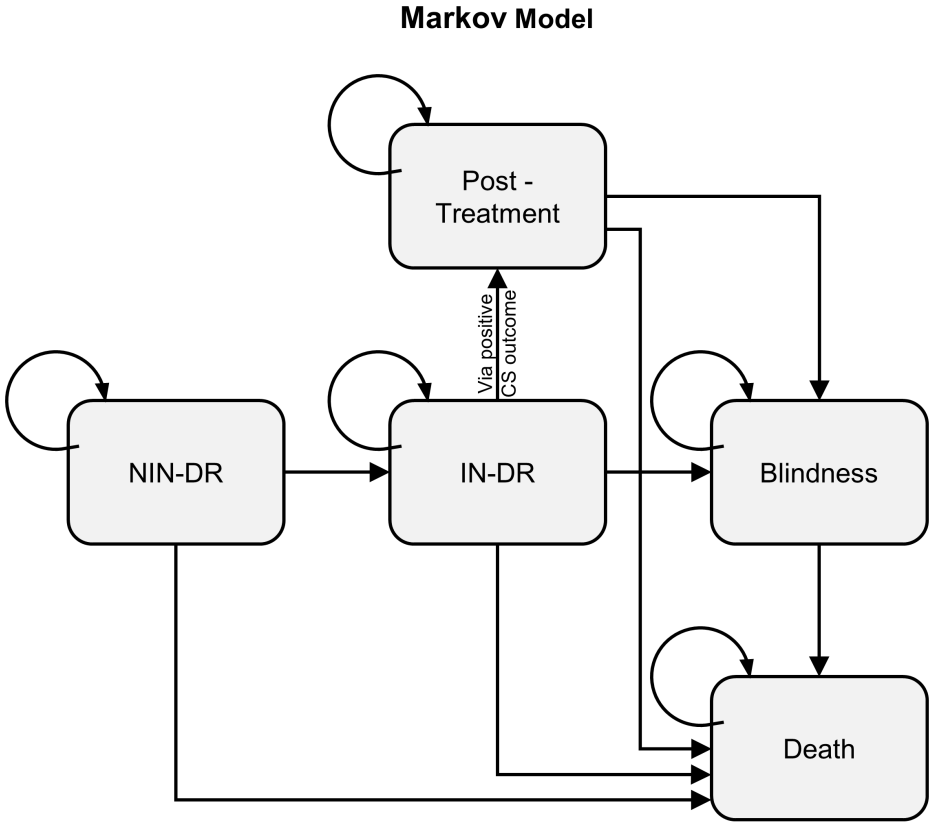

Figure 1. Markov model defining the 5 diabetic retinopathy (DR) health states: non-intervention-needed DR (NIN-DR), intervention-needed DR (IN-DR), Post-Treatment, Blindness, and Death. The transition from IN-DR to Post-Treatment only occurs when a patient receives a positive clinical screening (CS) outcome.

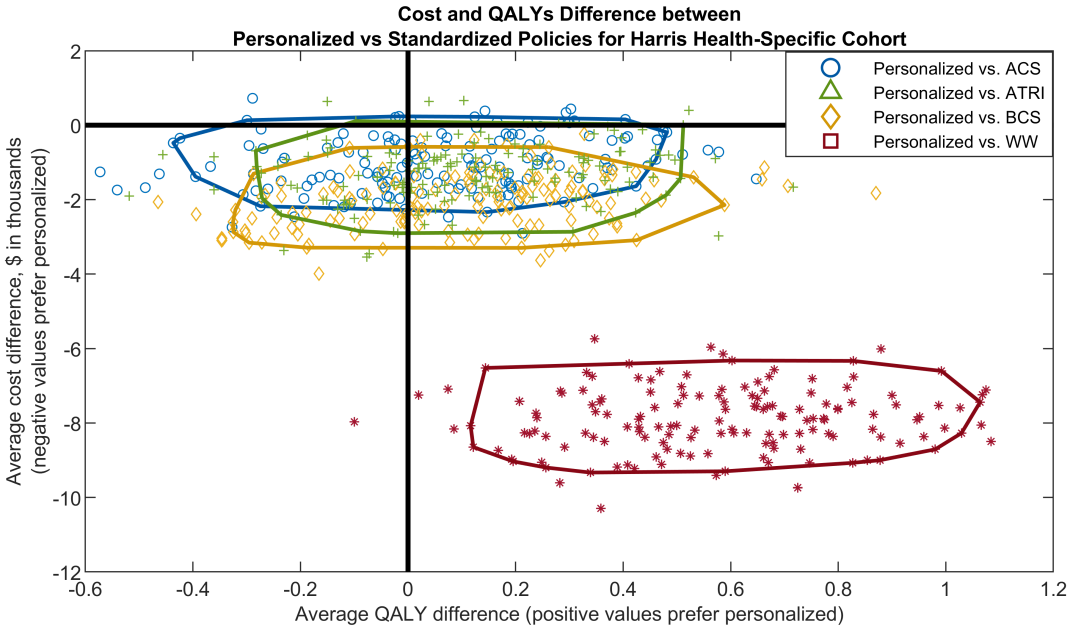

Figure 2. Each of the 200 data points represents a subgroup of 2500 patients where the personalized screening policy is compared to selected standardized screening policies: annual clinical screening (ACS), annual teleretinal imaging (ATRI), biennial clinical screening (BCS), and wait and watch (WW). Incremental cost-effectiveness is represented by the average difference in total quality-adjusted life years (QALYs) gained and the average costs spent in each data point. Data points to the right of the y-axis represent instances where the personalized policy provides more QALYs. Data points below the x-axis values represent instances where the personalized policy is less costly. The contour for each comparison encompasses the 95% confidence region for the set of 200 data points.

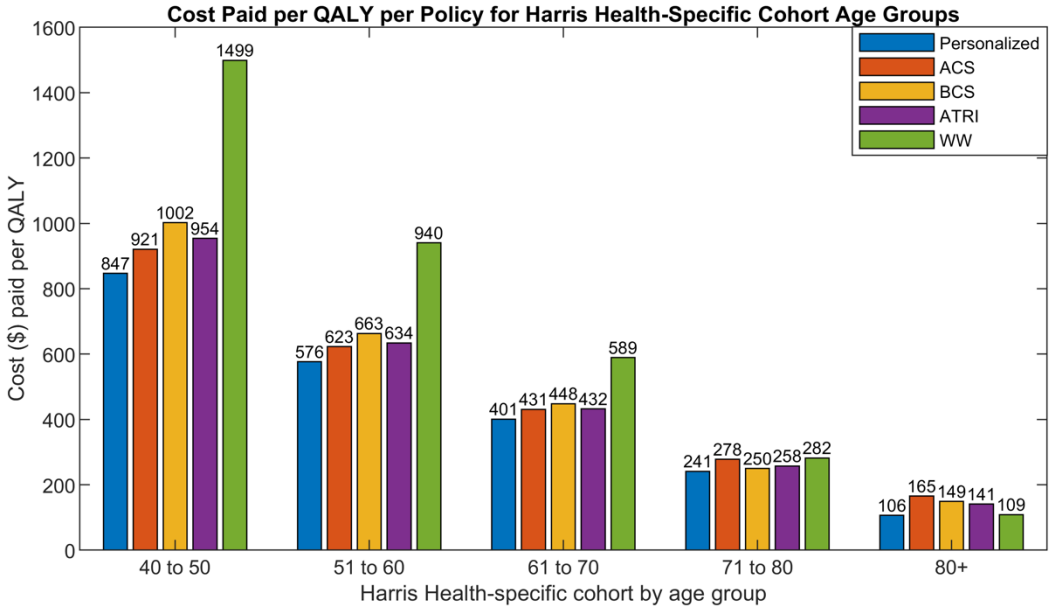

Figure 3. The out-of-pocket cost paid per QALY for each screening policy for different age groups (40 to 50, number of patients (n) = 159009; 51 to 60, n = 191525; 61 to 70, n = 124431; 71 to 80, n = 22703; 80+, n = 2332). Values are shown for the personalized policy, annual clinical screening (ACS), annual teleretinal imaging (ATRI), biennial clinical screening (BCS), and wait and watch (WW). The 95% confidence interval of the cost paid per QALY for each policy within each age group is represented by the vertical line at the top of each bar shown.