Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

UTILIZING SIMULATION TO UPDATE ROUTINE DIABETIC RETINOPATHY SCREENING POLICIES

Poria Dorali
Rosangel Limongi
Fariha Kabir Torsha

Department of Industrial Engineering
University of Houston
4222 Martin Luther King Blvd
Houston, TX 77204, USA

Christina Y. Weng

Department of Ophthalmology
Baylor College of Medicine
1977 Butler Blvd
Houston, TX 77030, USA

Taewoo Lee

Department of Industrial Engineering
University of Pittsburgh
1008 Benedum Hall
Pittsburgh, PA 15261, USA

ABSTRACT

Diabetic retinopathy (DR) is the leading cause of blindness for working-age US adults. While comprehensive screening examinations detect most early-stage DR cases, only 50–60% of diabetic patients adhere to the current annual screening guidelines. Recently, teleretinal imaging (TRI) has emerged as an accessible screening tool for patients with limited access. However, there exists no well-established guideline that incorporates TRI-based screening for such patients. We develop a Monte Carlo simulation model to replicate a safety-net system patient population using electronic medical record data from the Harris Health System (Houston, TX) and examine cost and health benefits of various TRI-based screening policies. We conduct sensitivity analysis to study the impact of patient-specific factors including age, A1C level, and screening adherence on screening policy performance. Our findings support TRI-based screening for patients with limited access and highlight the significant role of patient-specific factors in determining cohort-level screening policies.

1 INTRODUCTION

Diabetic retinopathy (DR) is the leading cause of blindness for working-age adults in the US (CDC 2022). It is estimated that over 60% of patients with type II diabetes and over 90% of patients with type I diabetes will develop DR within 20 years of diagnosis (Fong et al. 2003). Furthermore, as DR is asymptomatic in its early stages, patients may not realize the risk of developing advanced stages of DR, which are irreversible and four times as likely to cause blindness as early stages of the disease (Wykoff et al. 2021). With over 37 million diabetic cases and over 96 million prediabetic cases in the US (CDC 2022), there is a growing concern about the potential impact of DR in the near future.

Routine comprehensive screening examinations have proved effective in detecting early stages of DR (Javitt et al. 1994) and can prevent up to 98% of DR-related vision loss (Ferris 1993; Solomon et al. 2017).

The American Diabetes Association (ADA) and American Academy of Ophthalmology (AAO) recommend every diabetic patient receive annual dilated screening examinations (Flaxel et al. 2020; Solomon et al. 2017) via 7-field stereoscopic color fundus photography (Walton et al. 2016). However, on average only 50–65% of patients are adherent to this guideline (Eppley et al. 2019; MacLennan et al. 2014), with adherence rates as low as 35% for patients with limited access to care (Paz et al. 2006).

Recently, teleretinal imaging (TRI) has emerged as a viable and affordable alternative to the traditional eye clinic-based screening. Via TRI, patients have digital images of their retinas taken by primary care providers, which are then graded remotely (Walton et al. 2016); those deemed to meet the threshold for advanced stages of DR are then referred for the traditional screening exam at eye clinics (henceforth called clinic-based screening or CS). Despite lower sensitivity and specificity of TRI compared to CS, TRI has been shown to be advantageous for patients with low adherence and/or limited access to care due to its affordability, efficiency, and accessibility (Daskivich et al. 2017; Garoon et al. 2018; Walton et al. 2016).

However, there is little quantitative understanding about how TRI-based screening exams should be recommended to patients with limited access to address important trade-offs and promote the cost-effective use of the emerging screening technique. In particular, despite the improved access and affordability, TRI recommendations should be carefully coordinated in conjunction with the traditional CS exams due to its relatively low accuracy. Furthermore, health and cost benefits from TRI might vary depending on various patient-specific factors such as adherence behavior, health condition, and age.

In this paper, we develop a Monte Carlo simulation model that examines the health and cost benefits of various routine CS and TRI-based DR screening policies at different time intervals for patients with type II diabetes. The simulation model is built on a previously validated discrete-time Markov chain-based natural history model (Aoki et al. 2004; Kirkizlar et al. 2013; Vijan et al. 2000; Lee et al. 2021) and calibrated based on the electronic medical record (EMR) database of type II diabetes patients who were screened during the time period of 2013–2021 through an existing TRI program at the Harris Health System (HHS) in Harris County, TX, USA. As the model is built with the data from the largest safety-net hospital system in the third most populous county in the US that encompasses the Houston metropolitan area, there is significant diversity in the patient cohorts who are eligible for the TRI program.

In particular, our work aims to identify patient subgroups who would truly benefit from TRI in terms of health benefits and cost savings and how often TRI screening exams should be recommended to different subgroups. These findings provide important insights into cohort-based screening policies with respect to both the type (TRI or CS) and timing of screening. We consider a variety of patient subgroups based on the HHS EMR data with varying adherence rates, ages, and A1C levels. This study is uniquely positioned in that it provides recommendations on a cohort-level policy as opposed to one-size-fits-all policies or even individual-level recommendations found in literature (Flaxel et al. 2020; Lee et al. 2021; Solomon et al. 2017). While the former may fail to address patient-specific needs for DR screening, the latter may be challenging to implement due to the lack of patient-specific data and substantial resource burdens (e.g., a recommender system calibrated for the individual patient) (Lamoureux et al. 2017; Ricciardi and Boccia 2017). The cohort-level policy proposed in this paper attempts to address these concerns by providing various routine screening policies to reasonably stratified patient subgroups. The rest of this paper is structured as follows. Section 2 provides a brief literature review, Section 3 illustrates details of the simulation model and data sources, Section 4 presents our results based on the EMR data to examine health and cost benefits of various screening policies for the current patient population at HHS and sensitivity analysis to examine the impact of patient-specific factors on screening policies, and Section 5 discusses concluding remarks.

2 LITERATURE REVIEW

Markov models have been widely used in healthcare literature to model disease states and progression; e.g., liver cancer stages (Kay 1986), HIV stages (Longini Jr et al. 1989), and diabetes prevalence forecasts (Honeycutt et al. 2003). Simulation and sequential decision making techniques have utilized underlying

Markov models to examine the performance and cost-effectiveness of various screening or intervention policies; e.g., statin treatment policies (Mason et al. 2012), diabetes management (Zhang et al. 2019), and cancer screening (Ayer et al. 2012). Recent studies also demonstrate the importance of incorporating patient adherence within the model for developing personalized screening or intervention policies (Ayer et al. 2016; Mason et al. 2012).

Monte Carlo simulation based on the underlying Markov chain for DR natural history has been widely used for screening policy evaluation. Javitt et al. (1994) used Monte Carlo simulation to study the cost-effectiveness of CS-based DR screening from a societal perspective. Vijan et al. (2000) evaluated the cost-effectiveness of CS-based DR screening for patient groups stratified by age and A1C level. Aoki et al. (2004) used simulation to show the cost-effectiveness of a teleopthalmology system for diabetic eye diseases in a prison population and Kirkizlar et al. (2013) analyzed cost saving of telemedicine in DR screening for patients in the Veterans Health Administration. Whited et al. (2005) conducted economic analysis of digital teleopthalmology systems versus CS in federal healthcare agencies over a 12-month period. Rein et al. (2011) utilized simulation to evaluate the benefit of biennial CS, telemedicine screening, and self referral for low-risk patients. Swan et al. (2020) developed a microsimulation model to evaluate the benefits of DR screening in primary care and care coordination. Lastly, Fuller et al. (2022) evaluated the 5-year cost-effectiveness of an automated retinal imaging system for low-income populations.

Unlike the previous studies, to the best of our knowledge, this paper is the first to examine lifetime health and cost benefits of TRI screening from a patient cohort-level decision-making perspective in a limited access setting. In particular, the simulation model was built on the EMR database at a large safety-net system. This study uniquely considers a variety of screening policies that are varied by both screening type and interval. Our sensitivity analysis also examines risk (A1C level) and age-based screening policies, which were not studied previously. Additionally, our sensitivity analysis uniquely considers both CS and TRI adherence rates, each stratified by risk (A1C level). Our contributions are as follows:

1. We develop a discrete-time Monte Carlo simulation model based on the underlying Markov chain for DR natural history to recreate a large, highly heterogeneous urban safety-net system population. To the best of our knowledge, our study is the first to generate a diabetic population of a safety-net system from real screening data and patient EMRs to examine the realistic, lifetime impact of both CS and TRI-based screening policies.
2. We conduct sensitivity analysis based on patient adherence, age, and A1C level to identify cohort-specific routine screening policies. In addition to fixed-interval policies, we also examine the performance of a risk and age-based screening policy.

3 METHODOLOGY

The Monte Carlo simulation model was built using MATLAB R2021a. The simulation model considered a large number of diabetic patients, where patient parameters and characteristics were randomly sampled from the HHS EMR data and each patient's health state trajectory was individually tracked and recorded. Patient characteristics such as starting age, initial DR state, A1C level, and race were probabilistically assigned using fixed random seed values (this ensures a consistent heterogeneous patient cohort for each tested policy). We considered 100 replications for each simulated scenario. Lifetime-based analysis was chosen so as to study the development and long-term impact of DR and associated vision loss from both health and cost perspectives. Simulation outcomes collected for each patient include total costs spent (USD), accumulated quality adjusted life-years (QALYs), whether and when the patient became blind or received treatment, and time of death. Costs and QALYs were discounted over time. The effectiveness of DR screening was quantified by total QALY gains as a result of a screening policy for the underlying patient cohort. As this study considers patient-level recommendations and adherence, all cost analysis was based on out-of-pocket costs incurred by each simulated patient.

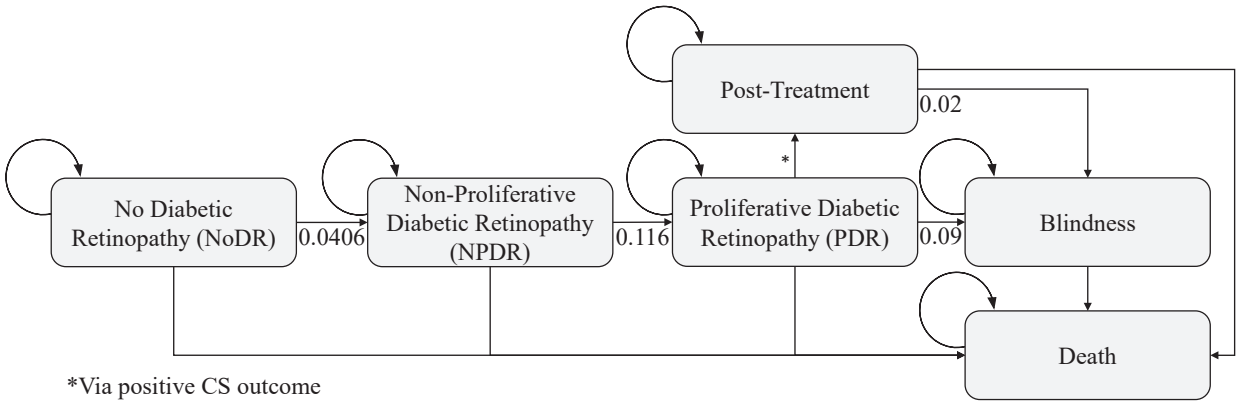


Figure 1: The Markov chain with all possible states and state transitions. Values on the arcs represent annual natural state transition probabilities for patients with 7% A1C. Death probabilities are defined by state and age-based mortality rates. Transitions to post-treatment occur only upon a positive CS.

3.1 DR Markov Model

We utilized a discrete-time, non-stationary Markov model to represent the progression of DR and patient health status. At each time epoch, a patient is in one of the six states: no diabetic retinopathy (NoDR), non-proliferative diabetic retinopathy (NPDR), proliferative diabetic retinopathy (PDR), post-treatment, blindness and death. Figure 1 shows the health states and possible transitions; the values shown represent the annual transition probabilities for patients with 7% A1C based on the natural disease progression in literature (Aoki et al. 2004; Kirkizlar et al. 2013; Moss et al. 1998; Vijan et al. 2000; Younis et al. 2003). As A1C level increases to 9%, 11%, and 13%, the transition probability from NoDR to NPDR increases to 0.0761, 0.1425, and 0.2669, respectively (UKPDS Group et al. 1998; Vijan et al. 2000). Note that transition probabilities are non-stationary because of age-specific mortality rates (Social Security Administration 2019) and health state-specific mortality multipliers (Vijan et al. 2000). We assumed that a patient enters the post-treatment state only if PDR is detected via CS (see Section 3.3 for more details).

3.2 Policy Evaluation

Seven different screening policies were considered: Annual CS (ACS), Annual TRI (ATRI), Semi-annual CS (SCS), Semi-annual TRI (STRI), Biennial CS (BCS), Biennial TRI (BTRI), and No Screening (NS). In Section 4.3, motivated by our sensitivity analysis results, we also consider a dynamic screening policy where screening intervals vary based on age and risk of developing PDR.

3.3 Simulation Model

The simulation model was based on our collaboration with HHS where the simulated patient population was created to mimic a large, urban diabetic population eligible for TRI screening. Patient data were collected from the EMRs of 2,456 Type II diabetic patients who received TRI screening exams from 2013 to 2021 at HHS; race, age, zip-code, lab values including A1C, insurance status, and TRI screening outcomes were collected. The cohort was comprised of 57.74% Hispanic, 24.76% Black, 8.79% white, 5.25% Asian, and 3.46% other. Analysis of the EMR data indicated that 62.75% of the patients had no DR, 34.93% NPDR, and 2.32% PDR. The average A1C level was 7.91% and the average age at first screening was 54.62.

The implementation of TRI in the model was based on the TRI screening program jointly operated by the Baylor College of Medicine (BCM) and HHS since 2013, which is one of the largest TRI programs in

the US with over 160,000 eyes screened (Walton et al. 2016). TRI false negative and false positive rates were 0.02 and 0.151, respectively (Date et al. 2019; Walton et al. 2016). Patient adherence rate for TRI recommendation was calculated based on whether or not a patient complied with a primary care provider's TRI recommendation, available from the EMR data. Our statistical analysis showed that race was the only patient characteristic associated with TRI adherence, which we then used to generate adherence rate for *de novo* hypothetical patients. Across all 2,456 patients considered, the average TRI adherence rate was 68.4%. Adherence to CS recommendation was set to 35% based on a previous DR screening study (Paz et al. 2006). Adherence to *follow-up* CS recommendation (i.e., after a positive TRI), on the other hand, was set to 54.9% based on our BCM-HHS TRI program data (Chamberlain 2018).

A synthetic population of 50,000 type II diabetes patients was generated to represent the size of diabetic patient population currently served by HHS. Each generated patient was uniquely assigned a race, starting age, A1C level, initial DR state, and TRI adherence rate, sampled from the EMR data of the highly heterogeneous population at HHS. The patient cohort was uniquely generated in each simulation replication. More details about the base case cohort can be found in Section 4.1.

The simulation model used discrete time epochs at 6-month interval and kept track of each patient until the age of 99 or death, whichever came first. A 3% discount rate was used for QALYs and costs collected over time (Sanders et al. 2016). Furthermore, the annual DR state transitions and age-specific mortality rates were converted into semi-annual values using a cycle-length changing technique (Sonnenberg and Beck 1993). We verified our semi-annual model against the previously validated natural history models (Aoki et al. 2004; Kirkizlar et al. 2013; Vijan et al. 2000) by comparing the proportion of the simulated cohort in each state between the two models after running them for the same 50,000 patients over 10 years under the NS policy (i.e., natural DR progression). The percentage point differences in the proportions between the two models (100 replications) were $0.92\pm0.11\%$ (95% CI), $1.53\pm0.23\%$, $2.31\pm0.24\%$, $1.51\pm0.19\%$ and $1.11\pm0.17\%$ for the No DR, NPDR, PDR, Blindness, and Death states, respectively.

At the beginning of each time epoch, a specific DR screening recommendation was provided to a patient based on the policy being evaluated. A patient's health transition was composed of two parts: a mortality check and a DR progression check. The mortality check part determined whether or not a patient died during the current time epoch based on the patient's mortality rate. Mortality rates were determined by multiplying the age-specific base mortality rate (Social Security Administration 2019) by the state-specific mortality multiplier (1.36, 1.76, 1.76, and 2.34 for NPDR, PDR, post-treatment, and blindness, respectively (Aoki et al. 2004; Kirkizlar et al. 2013)). At each time epoch, if a randomly generated number in $[0, 1]$ was less than the patient-specific mortality rate, the patient entered death. The DR progression check part determined if a patient transitioned to the next DR state or blindness (See Figure 1), which was implemented in a similar fashion via random number generation.

At the end of each time epoch, QALYs and costs were collected. State-specific QALYs (0.94, 0.87, 0.83, 0.83, and 0.71 annually for NoDR, NPDR, PDR, post-treatment, and blindness, respectively (Kirkizlar et al. 2013)) were collected and the half-cycle correction method was employed (Sonnenberg and Beck 1993). Costs included screening costs ($27.35 for TRI and $124.69 for CS), treatment cost of $1297.79, and blindness-related direct annual cost of $4944.00 (Moshfeghi et al. 2020). Note that treatment cost was determined by the weighted average of costs for three interventions: photocoagulation, pars plana virectomy (PPV) and endolaser or PPV with membrane peel and intraocular tamponade (Garoon et al. 2018). Additionally, $21.53 was added to the cost of CS to account for travel and lost wages (Garoon et al. 2018); this type of cost was not considered for TRI because TRI can be implemented as a part of primary care visits (our data indicates that most diabetic patients visit primary care physicians every 3-6 months).

At each time epoch, one of the three recommendations was made: no screening (NS), TRI or CS. Figure 2 shows a detailed overview of the DR screening process during each 6-month time period. Unlike TRI or CS, whenever NS is recommended to a patient, a screening outcome is not observed. As mentioned, our simulation model mimics the HHS TRI program where TRI is utilized as a tool to refer patients for follow-up CS exams when they are believed to be in an advanced DR state (Walton et al. 2016). That

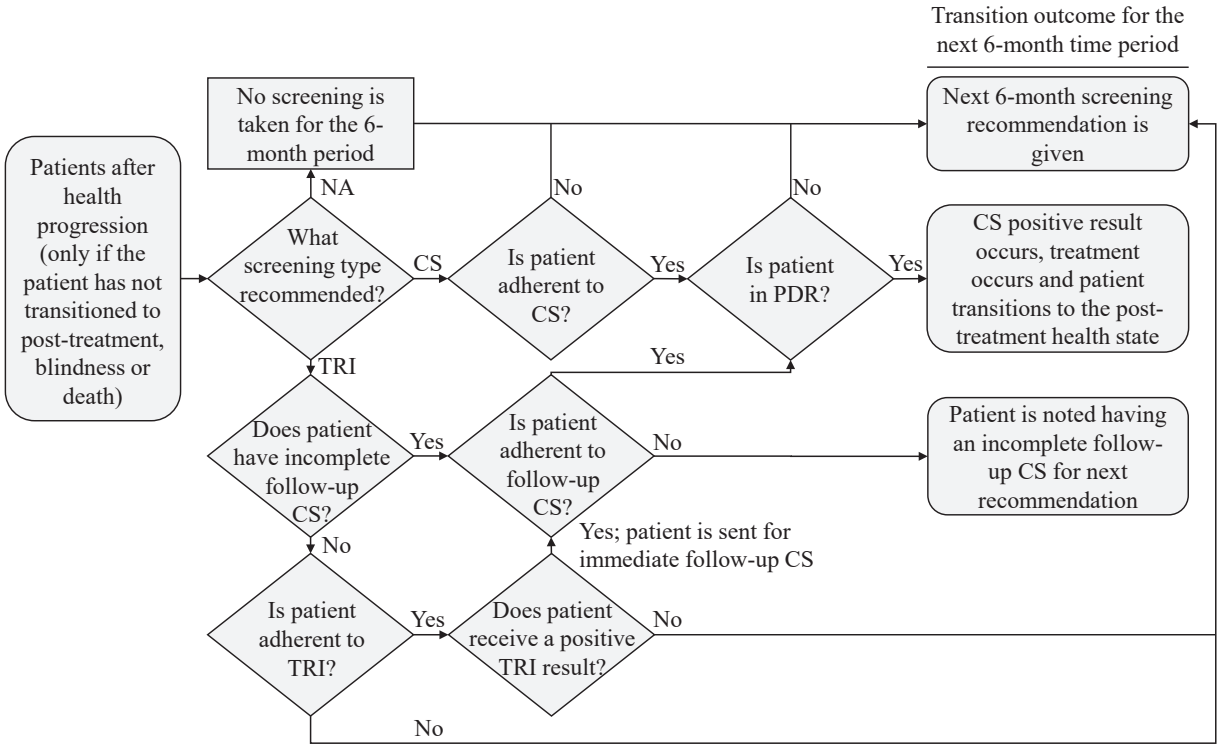


Figure 2: Overview of DR screening pathways and outcomes during a single 6-month time period.

is, upon a positive TRI outcome, a patient is immediately sent for follow-up CS, which must be carried out before any other screening occurs; from the modeling perspective, the patient is flagged as "follow-up CS incomplete" and is recommended to complete this at every subsequent time epoch unless the patient has already transitioned to blindness or death (see Figure 2). If a patient attends follow-up CS and a false positive outcome occurs, the patient will still continue to the next 6-month screening recommendation.

4 EXPERIMENTAL RESULTS

In this section, we first present base case analysis to examine health and cost benefits of various routine screening policies for the HHS diabetic patient population. We then present sensitivity analysis to examine the screening policies for patient subgroups varied by age, A1C level, and adherence rates.

4.1 Base Case Analysis

The simulation model was first applied to a base cohort constructed from the EMR data of the 2,456 patients who received regular TRI screening exams between 2013-2021 and was run for 100 replications. Synthetic patients were generated by sampling various patient parameters from the EMR data such as race, age, initial DR state, A1C level and race-based TRI adherence to match the composition of the real patient population at HHS. While multiple patients with the same characteristics may be generated, each of their trajectories is uniquely determined by probabilistic adherence and state transition. The base cohort closely resembled the population from the EMR data. The difference between the base cohort across all replications and the EMR data was as follows. The percentage point difference in the proportion of each race was less than 0.65$\pm$0.47% (95% CI), the difference in average age was 0.1$\pm$0.01 year (95% CI), and the difference in average A1C level was 0.51$\pm$0.01 (95% CI). For the proportions of initial health states, the percentage point differences were 3.52$\pm$0.05% (95% CI), 2.91$\pm$0.04%, and 0.61$\pm$0.01% for NoDR, NPDR, and PDR, respectively.

Table 1: Base case cohort simulation outcome comparison between various routine policies: Annual CS (ACS), Annual TRI (ATRI), Semi-annual CS (SCS), Semi-annual TRI (STRI), Biennial CS (BCS), Biennial TRI (BTRI), and No Screening (NS). Values represent the average values over 100 simulation replications.

	ACS	ATRI	SCS	STRI	BCS	BTRI	NS
Total QALYs gains (thous.)	12.68	14.59	15.09	15.68	9.50	12.62	0.00
Per-patient screening costs	1,993	2,027	2,764	2,457	1,405	1,651	0
Average costs paid per QALY	585.8	527.6	563.5	516.3	642.1	569.1	838.2
Patients treated (thous.)	24.78	27.56	28.24	29.27	19.83	24.60	0.00
Blindness cases averted (thous.)	9.53	10.66	10.94	11.35	7.57	9.44	0.00
Blindness years averted	2.20	2.50	2.58	2.69	1.70	2.17	0.00
Average age of death	76.64	76.69	76.70	76.72	76.57	76.65	76.31

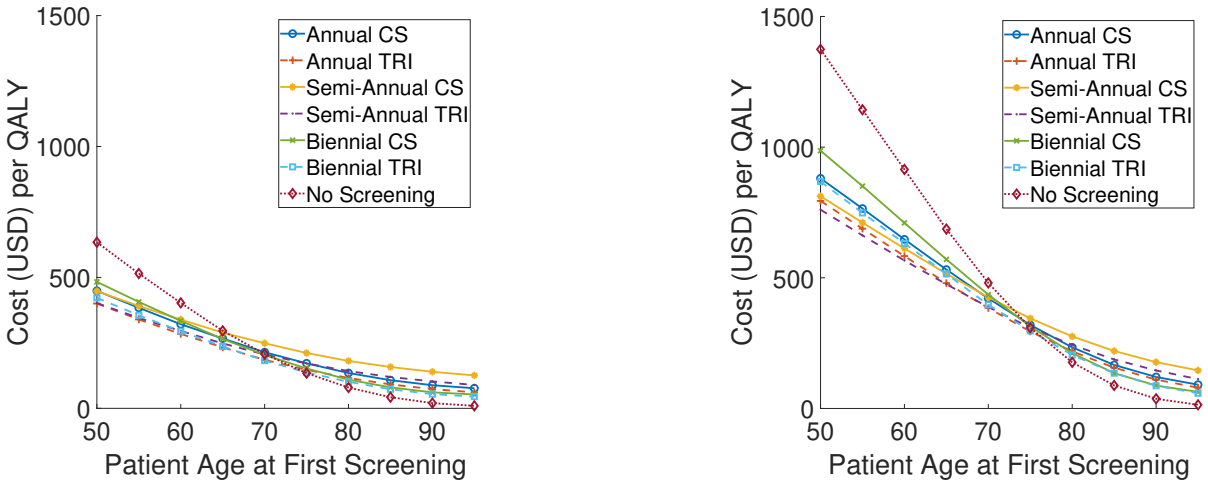
Table 1 compares simulation outcomes between different routine policies for the base case cohort. Each value represents the average over 100 replications and all values reported had a relative standard deviation within $\pm 1\%$. Furthermore, the 95% confidence interval bounds were within $\pm 2\%$ of the recorded averages. In all outputs besides average costs paid per QALY and average age of death, the value shown represents the difference between each policy and NS. Overall, STRI was the best performing policy in terms of both total QALY gains and cost savings (costs per QALY gain), followed by ATRI. The differences between the policies were found statistically significant via paired-sample t-tests (p-values < 0.01). As a general trend, the more patients a policy was able to treat, the more blindness cases were prevented and the more positive health outcomes occurred. As expected, when the screening interval decreased there was an increase in health benefits, i.e., QALY gains, total blindness cases averted, and average blindness years saved.

Interestingly, patients experienced more cost savings (less costs per QALY gain), on average, as screening frequency increased; while this may appear counterintuitive, the large direct annual costs associated with blindness imply that even a single year of blindness prevented could produce more cost savings than a lifetime of screening. Another noteworthy finding is that more frequent screening exams with the less accurate TRI provided more health benefits than accurate CS exams with less or even equal frequency (e.g., see STRI vs. SCS); we believe this is because TRI is more accessible for the base cohort (TRI adherence rate is significantly higher than CS adherence for most patients). In fact, base cohort patients under the STRI policy on average attended 19.85 TRI exams over their lifetimes, with an average of 3.03 *follow-up* CS exams. In contrast, base cohort patients under the SCS policy on average attended 10.13 CS exams over their lifetime. The increased number of overall screening exams under STRI is largely attributed to improved adherence. Furthermore, as TRI is aggressive in sending patients for *follow-up* CS (its false positive rate is 15.1%), TRI can help capture advanced DR in a more timely manner compared to less accessible CS. For more adherent patients, on the other hand, frequent screening may lead to unnecessary spending with little health benefits, as shown by our sensitivity analysis results in the next subsection.

4.2 Sensitivity Analysis

Sensitivity analysis was conducted to examine the impact of age, A1C level, and adherence rates on the performance of screening policies. All results shown in this section represent the averages over 100 replications. We first considered cohorts varied by age groups from age 50 to 95 by 5-year increments as well as risk levels represented by 7% and 13% A1C levels, where a higher A1C level is known to increase risk of developing advanced DR (UKPDS Group et al. 1998; Vijan et al. 2000). The A1C level of each patient in each cohort was assumed to be static to isolate the analysis from the impact of uncertain changes in A1C level over time. Our second analysis considered cohorts varied by CS and TRI adherence rates both from 20% to 90% by 5% increments as well as the two risk levels previously defined.

Figure 3 shows the average per-patient cost per QALY gain achieved by different policies at different ages for (a) low and (b) high-risk patients. We notice a clear shift in the types of policies that offered the least cost per QALY as a patient's age increases. For low-risk patients, ATRI was found to produce the least cost per QALY for patients aged 50–68, BTRI for patients age 68–74 and NS for age 74 and above. That is, while TRI was chosen for screening, the screening frequency decreased as patient age increased and eventually no screening was recommended as age exceeded a certain threshold. This suggests that the benefits of screening diminish as patients age, which can be attributed to increased mortality rates and decreased probability to develop long-term blindness. For high-risk patients, similarly, STRI produced the least cost per QALY for patients aged 50–68, ATRI for age 68–76, and NS for age 76 and above. Note that TRI is recommended more frequently and the age threshold at which no screening starts to produce the least cost per QALY is delayed for these higher-risk patients; i.e., the model captured the increased risk for this cohort and assessed the performance of the policies accordingly. For both low-risk and high-risk cohorts, TRI-based screening policies produced less cost per QALY than their CS-based counterparts.



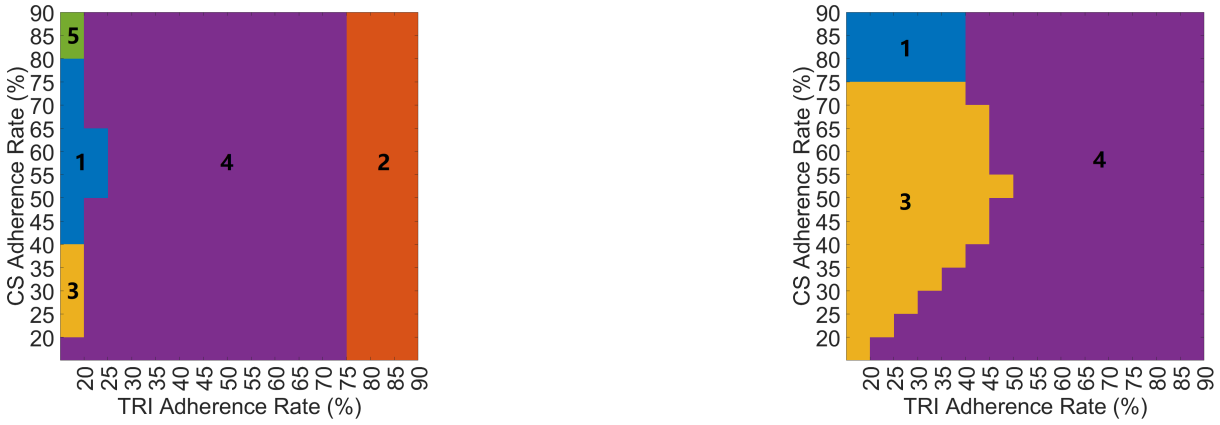
(a) Low-risk patients (7% A1C)

(b) High-risk patients (13% A1C)

Figure 3: Comparison of average costs per QALY between various screening policies for patient cohorts stratified by age and A1C level.

Next, to examine the impact of both patient adherence behavior and DR risk level, we conducted three-way sensitivity analysis where both TRI and CS adherence rates were varied from 20% to 90% by 5% intervals for both low and high-risk patients. Figure 4 shows heatmaps that indicate the screening policy with the least cost per QALY for each cohort represented by TRI adherence, CS adherence, and A1C level. Note that only regular CS adherence is varied on the y-axis; follow-up CS was set to 54.9% or equal to the regular CS adherence, whichever was greater. Overall, Figure 4 suggests that there exists an adherence-based threshold for screening policy recommendation. For example, for a low-risk patient (Figure 4a) with a fixed CS adherence rate of 35%, as TRI adherence rate increases from 20% to 30%, the screening policy with the least cost per QALY changes from SCS to STRI. Similarly, the policy switches from STRI to ATRI when the patient's TRI adherence rate reaches 80%.

For low-risk patients (Figure 4a), our analysis finds strong support for TRI-based policies, i.e., STRI and ATRI, even for patients with higher CS adherence rates. Interestingly, we observe a non-monotonic behavior in the heatmap where the least cost/QALY policy changes from STRI to ACS and then changes back to STRI as CS adherence goes up. This is because increased CS adherence also increases *follow-up* CS adherence upon a positive TRI. Since TRI is quite aggressive in sending patients for follow-up CS (with a false positive rate of 15.1%), once follow-up CS adherence is sufficiently high, it may be plausible



(a) Low-risk patients (7% A1C)

(b) High-risk patients (13% A1C)

Figure 4: Heatmaps of policies that provide the minimum average costs (USD) per QALY across 100 replications for patient cohorts stratified by adherence rates and A1C level. Each section corresponds to the following policy: 1 - Annual CS (ACS), 2 - Annual TRI (ATRI), 3 - Semi-annual CS (SCS), 4 - Semi-annual TRI (STRI), and 5 - Biennial CS (BCS).

to rely on TRI first and refer the patients for follow-up CS as needed. For high-risk patients, Figure 4b shows that the need for CS-based exams increases especially for patients with low TRI adherence rates. Furthermore, ATRI (i.e., less frequent TRI exams) is never recommended. Similar to low-risk patients, for those who are highly adherent to CS (hence follow-up CS), STRI produces less cost than CS-based policies (e.g., see 45% TRI-adherent and 80% CS-adherent). Overall, Figure 4 indicates the feasibility of patient cohort-specific screening policy based on A1C level and adherence rates.

4.3 Age and Risk-based Screening Policy

In addition to the routine DR screening policies considered, an age and risk-based DR screening policy was developed motivated by the outcomes from our age and A1C-based sensitivity analysis. Patients with A1C levels classified as 7% were assigned semi-annual TRI, annual TRI, biennial TRI, and NS for entering age of $<$50, 50–68, 68–74, and $>$74, respectively. Patients with higher A1C levels were assigned semi-annual TRI, annual TRI, and NS for entering age of $<$68, 68–76, or $>$76, respectively. The dynamic policy performed very similarly to semi-annual TRI for the base cohort, providing marginally less costs per QALY ($514.77 vs. $516.30) while resulting in 150 additional cases of blindness on average. We note that there may be other policies that could be varied by not only age and A1C but also other patient-specific characteristics. Personalized DR screening guideline via such dynamic policies would be beneficial from individual-level perspective and remains our future work. For example, preliminary results based on the partially-observable Markov decision process are available in Lee et al. (2021).

5 DISCUSSION AND CONCLUSION

In this study, we utilized a Monte Carlo simulation model to evaluate different CS and TRI-based policies at varying screening intervals. Of particular interest to this study was identifying the screening policy that should be recommended to a large urban safety-net system population, mostly consisting of under/uninsured patients with low CS screening adherence. We recreated the safety-net population in our base case simulation model based on the EMR data of type II diabetic patients at HHS and their existing TRI program.

Analysis of the base case cohort indicated the semi-annual TRI screening policy as most favorable in terms of both health benefit and cost savings, followed by annual TRI-based screening. Overall, our results

support the utilization of TRI-based policies at varying intervals depending on risk level and age. In fact, we notice from the EMR data that physicians often recommend TRI exams every 6 months for patients who are at high risk or exhibit poor adherence or diabetic control. While these decisions are currently made in an ad-hoc manner, our results reinforce the role of the simulation model as a clinical decision support tool that identifies cohort-specific screening recommendations based on various characteristics.

Our sensitivity analysis examined how policies should be recommended to patients with varying age, A1C level, and adherence rates. Age, A1C level and adherence were all found to be significant factors in determining best screening policies in terms of both health benefit and cost savings. More frequent screening was found beneficial for younger patients, whereas older patients experienced diminished benefits of screening. For patients with sufficiently high TRI and CS adherence rates, increasing TRI frequency from annual to semi-annual screening did not generate additional health benefits and cost savings. We believe this indicates potential TRI overutilization issues for those who are already highly adherent. Patients with increased risk of developing advanced DR (i.e., increased A1C level) benefitted more from frequent TRI screenings than low risk patients. Finally, an age and risk-based screening policy was found to perform similarly to the routine semi-annual TRI policy, which demonstrates the potential benefits of dynamic screening policies and warrants further exploration.

This study has a few limitations. First, the model is a single-eye model, i.e., each patient within the model represents only one eye instead of two eyes, as is common in current modeling studies in the ophthalmology community (Karakosta et al. 2012). In a study of a large safety-net population with similar demographics to ours, 77% of patients screened were found to have the same level of DR in both eyes (Varma et al. 2004), justifying the choice of the single-eye perspective in this study. Second, ungradable TRI images were not considered in this study. Note that the BCM-HHS TRI program data indicated only 2% of the TRI images were ungradable, with each individual patient having at least one eye successfully graded (Date et al. 2019). Finally, we assume that A1C level and adherence are static over time. Our data did not indicate any predictable changes in these parameters over the study period, and thus the model was assumed non-anticipatory of future changes in such parameters when examining a screening policy. If changes in these parameters are observed, the model can be periodically updated and rerun.

Overall, this simulation study demonstrates the benefit of utilizing TRI-based policies for safety-net system patient populations with limited access to eye care and low adherence to eye clinic-based screening recommendations. Our findings support cohort-specific routine screening policies where benefits of TRI-based screening are determined by patient age, A1C level, and adherence. As this modeling framework is highly adaptable, TRI programs can utilize their own patient data within the model to evaluate the impact of different DR screening policies on their specific cohort. In the future, as health information system continues to improve, the routine screening policies proposed in this study could eventually evolve into purely personalized screening policies with little compromise in implementation efficiency (Lee et al. 2021). This simulation model can be used for evaluating the performance of such personalized screening policies and serve as a clinical decision support tool.

ACKNOWLEDGMENTS

This work was supported in part by DHHS Grant #1W1CMS331751-01-00.

REFERENCES

Aoki, N., K. Dunn, T. Fukui, J. R. Beck, W. J. Schull, and H. K. Li. 2004. "Cost-Effectiveness Analysis of Telemedicine to Evaluate Diabetic Retinopathy in a Prison Population". *Diabetes Care* 27(5):1095–1101.

Ayer, T., O. Alagoz, and N. K. Stout. 2012. "OR Forum - A POMDP Approach to Personalize Mammography Screening Decisions". *Operations Research* 60(5):1019–1034.

Ayer, T., O. Alagoz, N. K. Stout, and E. S. Burnside. 2016. "Heterogeneity in Women's Adherence and Its Role in Optimal Breast Cancer Screening Policies". *Management Science* 62(5):1339–1362.

CDC 2022. "National Diabetes Statistics Report Website". https://www.cdc.gov/diabetes/data/statistics-report/index.html, accessed 10th January 2022.

Chamberlain, P. D. 2018. "Factors Associated with Follow-up Examination Following Positive Teleretinal Screening for Diabetic Retinopathy in an Urban Healthcare System". *Texas Medical Center Dissertations* AAI10786152.

Daskivich, L. P., C. Vasquez, C. Martinez, C.-H. Tseng, and C. M. Mangione. 2017. "Implementation and Evaluation of a Large-Scale Teleretinal Diabetic Retinopathy Screening Program in the Los Angeles County Department of Health Services". *JAMA Internal Medicine* 177(5):642–649.

Date, R. C., K. L. Shen, B. M. Shah, M. A. Sigalos-Rivera, Y. I. Chu, and C. Y. Weng. 2019. "Accuracy of Detection and Grading of Diabetic Retinopathy and Diabetic Macular Edema Using Teleretinal Screening". *Ophthalmology Retina* 3(4):343–349.

Eppley, S. E., S. L. Mansberger, S. Ramanathan, and E. A. Lowry. 2019. "Characteristics Associated with Adherence to Annual Dilated Eye Examinations among US Patients with Diagnosed Diabetes". *Ophthalmology* 126(11):1492–1499.

Ferris, F. L. 1993. "How Effective are Treatments for Diabetic Retinopathy?". *Journal of the American Medical Association* 269(10):1290–1291.

Flaxel, C. J., R. A. Adelman, S. T. Bailey, A. Fawzi, J. I. Lim, G. A. Vemulakonda, and G.-S. Ying. 2020. "Diabetic Retinopathy Preferred Practice Pattern®". *Ophthalmology* 127(1):66–145.

Fong, D. S., L. Aiello, T. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, I. Ferris, Fredrick L., R. Klein, and American Diabetes Association. 2003. "Diabetic Retinopathy". *Diabetes Care* 26(1):226–229.

Fuller, S. D., J. Hu, J. C. Liu, E. Gibson, M. Gregory, J. Kuo, and R. Rajagopal. 2022. "Five-Year Cost-Effectiveness Modeling of Primary Care-Based, Nonmydriatic Automated Retinal Image Analysis Screening Among Low-Income Patients with Diabetes". *Journal of Diabetes Science and Technology* 16(2):415–427.

Garoon, R. B., W. V. Lin, A. K. Young, A. G. Yeh, Y. I. Chu, and C. Y. Weng. 2018. "Cost Savings Analysis for a Diabetic Retinopathy Teleretinal Screening Program Using an Activity-Based Costing Approach". *Ophthalmology Retina* 2(9):906–913.

Honeycutt, A. A., J. P. Boyle, K. R. Broglio, T. J. Thompson, T. J. Hoerger, L. S. Geiss, and K. Venkat Narayan. 2003. "A Dynamic Markov Model for Forecasting Diabetes Prevalence in the United States through 2050". *Health Care Management Science* 6(3):155–164.

Javitt, J. C., L. P. Aiello, Y. Chiang, F. L. Ferris III, J. K. Canner, and S. Greenfield. 1994. "Preventive Eye Care in People with Diabetes is Cost-Saving to the Federal Government: Implications for Health-Care Reform". *Diabetes Care* 17(8):909–917.

Karakosta, A., M. Vassilaki, S. Plainis, N. H. Elfadl, M. Tsilimbaris, and J. Moschandreas. 2012. "Choice of Analytic Approach for Eye-Specific Outcomes: One Eye or Two?". *American Journal of Ophthalmology* 153(3):571–579.

Kay, R. 1986. "A Markov Model for Analysing Cancer Markers and Disease States in Survival Studies". *Biometrics* 42(4):855–865.

Kirkizlar, E., N. Serban, J. A. Sisson, J. L. Swann, C. S. Barnes, and M. D. Williams. 2013. "Evaluation of Telemedicine for Screening of Diabetic Retinopathy in the Veterans Health Administration". *Ophthalmology* 120(12):2604–2610.

Lamoureux, E. L., H. Taylor, and T. Y. Wong. 2017. "Frequency of Evidence-Based Screening for Diabetic Retinopathy". *The New England Journal of Medicine* 377(2):194–195.

Lee, T., P. Dorali, Z. Shahmoradi, and C. Y. Weng. 2021. "Developing Behavior-Based Diabetic Retinopathy Screening Guidelines". *Investigative Ophthalmology & Visual Science* 62(8):1141–1141.

Longini Jr, I. M., W. S. Clark, R. H. Byers, J. W. Ward, W. W. Darrow, G. F. Lemp, and H. W. Hethcote. 1989. "Statistical Analysis of the Stages of HIV Infection Using a Markov Model". *Statistics in Medicine* 8(7):831–843.

MacLennan, P. A., G. McGwin, C. Heckemeyer, V. R. Lolley, S. Hullett, J. Saaddine, S. S. Shrestha, and C. Owsley. 2014. "Eye Care use among a High-Risk Diabetic Population Seen in a Public Hospital's Clinics". *JAMA Ophthalmology* 132(2):162–167.

Mason, J. E., D. A. England, B. T. Denton, S. A. Smith, M. Kurt, and N. D. Shah. 2012. "Optimizing Statin Treatment Decisions for Diabetes Patients in the Presence of Uncertain Future Adherence". *Medical Decision Making* 32(1):154–166.

Moshfeghi, A. A., T. Lanitis, G. Kropat, A. Kuznik, A. Gibson, H. Feng, and J. Prenner. 2020. "Social Cost of Blindness Due to AMD and Diabetic Retinopathy in the United States in 2020". *Ophthalmic Surgery, Lasers and Imaging Retina* 51(4):S6–S14.

Moss, S. E., R. Klein, and B. E. Klein. 1998. "The 14-Year Incidence of Visual Loss in a Diabetic Population". *Ophthalmology* 105(6):998–1003.

Paz, S. H., R. Varma, R. Klein, J. Wu, S. P. Azen, Los Angeles Latino Eye Study Group et al. 2006. "Noncompliance with Vision Care Guidelines in Latinos with Type 2 Diabetes Mellitus: The Los Angeles Latino Eye Study". *Ophthalmology* 113(8):1372–1377.

Rein, D. B., J. S. Wittenborn, X. Zhang, B. A. Allaire, M. S. Song, R. Klein, J. B. Saaddine, and V. C.-E. S. Group. 2011. "The Cost-Effectiveness of Three Screening Alternatives for People with Diabetes with No or Early Diabetic Retinopathy". *Health Services Research* 46(5):1534–1561.

Ricciardi, W., and S. Boccia. 2017. "New Challenges of Public Health: Bringing the Future of Personalised Healthcare into Focus". *European Journal of Public Health* 27(4):36–39.

Sanders, G. D., P. J. Neumann, A. Basu, D. W. Brock, D. Feeny, M. Krahn, K. M. Kuntz, D. O. Meltzer, D. K. Owens, L. A. Prosser et al. 2016. "Recommendations for Conduct, Methodological Practices, and Reporting of Cost-Effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine". *Journal of the American Medical Association* 316(10):1093–1103.

Social Security Administration 2019. "Actuarial Life Table, 2019". https://www.ssa.gov/oact/STATS/table4c6.html, accessed 12th January 2022.

Solomon, S. D., E. Chew, E. J. Duh, L. Sobrin, J. K. Sun, B. L. VanderBeek, C. C. Wykoff, and T. W. Gardner. 2017. "Diabetic Retinopathy: A Position Statement by the American Diabetes Association". *Diabetes Care* 40(3):412–418.

Sonnenberg, F. A., and J. R. Beck. 1993. "Markov Models in Medical Decision Making: A Practical Guide". *Medical Decision Making* 13(4):322–338.

Swan, B., S. Nambiar, P. Koutouan, M. E. Mayorga, J. Ivy, E. P. Fitts, and S. Fransen. 2020. "Evaluating Diabetic Retinopathy Screening Interventions in a Microsimulation Model". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 944–955. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

UKPDS Group et al. 1998. "Intensive Blood-Glucose Control with Sulphonylureas or Insulin Compared with Conventional Treatment and Risk of Complications in Patients with Type 2 Diabetes (UKPDS 33)". *The Lancet* 352(9131):837–853.

Varma, R., M. Torres, F. Peña, R. Klein, S. P. Azen, Los Angeles Latino Eye Study Group et al. 2004. "Prevalence of Diabetic Retinopathy in Adult Latinos: The Los Angeles Latino Eye Study". *Ophthalmology* 111(7):1298–1306.

Vijan, S., T. P. Hofer, and R. A. Hayward. 2000. "Cost-Utility Analysis of Screening Intervals for Diabetic Retinopathy in Patients with Type 2 Diabetes Mellitus". *Journal of the American Medical Association* 283(7):889–896.

Walton, O. B., R. B. Garoon, C. Y. Weng, J. Gross, A. K. Young, K. A. Camero, H. Jin, P. E. Carvounis, R. E. Coffee, and Y. I. Chu. 2016. "Evaluation of Automated Teleretinal Screening Program for Diabetic Retinopathy". *JAMA Ophthalmology* 134(2):204–209.

Whited, J. D., S. K. Datta, L. M. Aiello, L. P. Aiello, J. D. Cavallerano, P. R. Conlin, M. B. Horton, R. A. Vigersky, R. K. Poropatich, P. Challa et al. 2005. "A Modeled Economic Analysis of a Digital Teleophthalmology System as Used by Three Federal Healthcare Agencies for Detecting Proliferative Diabetic Retinopathy". *Telemedicine Journal & E-Health* 11(6):641–651.

Wykoff, C. C., R. N. Khurana, Q. D. Nguyen, S. P. Kelly, F. Lum, R. Hall, I. M. Abbass, A. M. Abolian, I. Stoilov, T. M. To et al. 2021. "Risk of Blindness among Patients with Diabetes and Newly Diagnosed Diabetic Retinopathy". *Diabetes Care* 44(3):748–756.

Younis, N., D. M. Broadbent, J. P. Vora, and S. P. Harding. 2003. "Incidence of Sight-Threatening Retinopathy in Patients with Type 2 Diabetes in the Liverpool Diabetic Eye Study: A Cohort Study". *The Lancet* 361(9353):195–200.

Zhang, Y., H. Wu, B. T. Denton, J. R. Wilson, and J. M. Lobo. 2019. "Probabilistic Sensitivity Analysis on Markov Models with Uncertain Transition Probabilities: An Application in Evaluating Treatment Decisions for Type 2 Diabetes". *Health Care Management Science* 22(1):34–52.

AUTHOR BIOGRAPHIES

PORIA DORALI is a PhD candidate in Industrial Engineering at the University of Houston. He holds a BS in Industrial Engineering from the University of Houston. His research interests include simulation and mathematical modeling in medical decision making and healthcare operations. His email is pdorali@uh.edu.

ROSANGEL LIMONGI is a graduate student pursuing her MS degree in Industrial Engineering at the University of Houston. She holds a bachelor's degree in Industrial Engineering from the University of Houston. Her research interests include optimization and data-driven modeling and analysis in healthcare policy making. Her email is rglimongicifelli@uh.edu.

FARIHA KABIR TORSHA Fariha Kabir Torsha is a PhD student in Industrial Engineering at the University of Houston. She holds a BSc in Industrial and Production Engineering from Bangladesh University of Engineering and Technology. Her research interests include inverse optimization and data-driven predictive modeling in healthcare applications. Her email is ftorsha@uh.edu.

CHRISTINA Y. WENG is an Associate Professor in the Department of Ophthalmology at the Baylor College of Medicine and Staff Physician in the Department of Ophthalmology at Ben Taub Hospital. She holds an MD from the University of Michigan Medical School and an MBA from University of Michigan Ross School of Business. Her research interests include diabetic retinopathy, surgical outcomes, and cost-efficiency analysis. Her email address is Christina.Weng@bcm.edu.

TAEWOO LEE is an Assistant Professor of Industrial Engineering at the University of Pittsburgh. He holds a PhD in Operations Research from the University of Toronto. His research interests include mathematical modeling and optimization in healthcare operations and medical decision making. His email address is taewoo.lee@pitt.edu.