

## Repositório ISCTE-IUL

---

Deposited in *Repositório ISCTE-IUL*:

2023-07-25

Deposited version:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Pereira, A., Laureano, R. M. S., Neto, F. & Macedo, J. (2021). Simulação computacional da adesão ao rastreio da retinopatia diabética: Modelo baseado em agentes dotados de lógica fuzzy. In Rocha, A., Gonçalves, R., Penalvo, F. G., & Martins, J. (Ed.), 2021 16th Iberian Conference on Information Systems and Technologies (CISTI). Chaves: IEEE.

Further information on publisher's website:

10.23919/CISTI52073.2021.9476265

Publisher's copyright statement:

This is the peer reviewed version of the following article: Pereira, A., Laureano, R. M. S., Neto, F. & Macedo, J. (2021). Simulação computacional da adesão ao rastreio da retinopatia diabética: Modelo baseado em agentes dotados de lógica fuzzy. In Rocha, A., Gonçalves, R., Penalvo, F. G., & Martins, J. (Ed.), 2021 16th Iberian Conference on Information Systems and Technologies (CISTI). Chaves: IEEE., which has been published in final form at <https://dx.doi.org/10.23919/CISTI52073.2021.9476265>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

---

### Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

---

# Simulação computacional da adesão ao rastreio da retinopatia diabética

Modelo baseado em agentes dotados de lógica fuzzy

## *Computer simulation of diabetic retinopathy screening adherence*

*Agent based model with fuzzy logic*

Andreia Pereira

Raul M. S. Laureano

BRU-IUL e ISTAR-IUL, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal  
andreia\_marisa\_pereira@iscte-iul.pt  
raul.laureano@iscte-iul.pt

Fernando Neto

João Macedo

Escola Politécnica, Computer Engineering (POLI/PPG-EC), Universidade de Pernambuco (UPE), Brasil,  
fbln@ecomppoli.br  
jssm@ecomppoli.br

*Resumo* — A retinopatia diabética é uma das principais causas de cegueira nos países desenvolvidos. Uma vez que esta doença permanece assintomática nos estádios iniciais, é da máxima importância a realização de rastreios periódicos. No entanto, o sucesso dos programas de rastreio depende em grande parte da adesão da população alvo. Assim, o desenvolvimento de simulações computacionais que permitam compreender os fatores correlacionados com a decisão individual de aderir ou não ao rastreio, prever baixas taxas de adesão e definir estratégias que promovam o seu aumento, reveste-se de grande utilidade. O objetivo da presente investigação é demonstrar que é possível simular o comportamento de adesão ao rastreio, através de um modelo com elevado grau de abstração e escalabilidade. Para tal, foi desenvolvido um modelo baseado em agentes dotados de lógica fuzzy, calibrado e validado com recurso a dados reais de 271.867 convocatórias para rastreio na Administração Regional de Saúde do Norte. Os resultados obtidos são muito próximos dos reais, quer no grupo de treino, quer no grupo de teste, o que atesta a validade do modelo para o estudo desta problemática e seu uso como ferramenta preditiva para planeamento em saúde pública.

*Palavras Chave* – *simulação computacional, modelos baseados em agentes, lógica fuzzy, retinopatia diabética, taxa de adesão a rastreios.*

*Abstract* — Diabetic retinopathy is a major cause of blindness in developed countries. Since this disease remains asymptomatic in the early stages, periodic screening is of utmost importance. However, the success of screening programmes depends largely on the adherence of the target population. Thus, the development of computer simulations that make it possible to understand the factors correlated with the individual decision to adhere to the screening or not, to predict low adherence rates and to define strategies that promote its increase, is very useful. The aim of the present research is to demonstrate that it is possible to simulate the behavior of adherence to screening, using a model with a high degree of abstraction and scalability. To this end, a model based

on agents with fuzzy logic was developed, calibrated, and validated using real data from 271,867 calls for screening in the Northern Regional Health Administration. The results obtained are very close to the real ones, both in the training and in the test sets, which attests to the validity of the model for the study of this problem and use as a predictive tool for public health planning.

*Keywords* - *computational simulation, agent-based models, fuzzy logic, diabetic retinopathy, screening adherence rate.*

### I. INTRODUÇÃO

A retinopatia diabética (RD) é uma complicação da diabetes caracterizada pela alteração estrutural dos vasos sanguíneos da retina. Atualmente é uma das principais causas de cegueira em países desenvolvidos. Em fases iniciais, a RD é assintomática, por este motivo, o doente com diabetes não deve esperar pelo aparecimento de sintomas visuais, devendo realizar exames oftalmológicos periódicos [1]. Neste contexto, vários países implementaram rastreios de RD de base populacional [2]. Um fator crucial para a custo-eficiência deste tipo de programas é a adesão por parte dos diabéticos [3] [4] [5] [6]. No entanto, esta problemática ainda está pouco explorada em termos simulação computacional [6] [7].

Anteriormente, desenvolvemos um modelo de simulação baseado em agentes (ABM) que demonstrou o potencial da utilização deste tipo de técnicas no estudo da adesão a rastreios [7]. No entanto, se por um lado o modelo revelou uma boa capacidade de replicação da realidade e utilidade na encenação de cenários num contexto geográfico específico, reconhecemos que a sua escalabilidade era reduzida, já que assentava principalmente em comportamentos previamente observados num programa de rastreio concreto [7]. Assim, a presente investigação foca-se na seguinte questão: É possível estudar a taxa de adesão a rastreios de base populacional através de ABMs

com um elevado nível de abstração e escalabilidade? Mais especificamente, pretende-se i) demonstrar que é possível desenvolver um ABM que retrate fielmente a decisão de aderir ou não ao rastreio, recorrendo a características intrínsecas do agente e do programa de rastreio. ii) demonstrar que um ABM com as características referidas pode ser utilizado em contextos diferentes daquele onde foram recolhidos os dados para o seu desenvolvimento. Para levar a cabo estes objetivos desenvolveu-se um ABM que utiliza lógica *fuzzy* para simular a tomada de decisão.

## II. REVISÃO DA LITERATURA

### A. Modelos de simulação no contexto do rastreio da retinopatia diabética

As primeiras simulações computacionais de rastreios de RD remontam aos anos 90 e recorreram fundamentalmente a cadeias de Markov para demonstrar a custo-efetividade da implementação de programas de rastreio de base populacional [8] [9] [10] [11]. Nos anos seguintes, e com o consenso alargado sobre a custo-efetividade dos rastreios de RD de base populacional, o foco dos investigadores alterou-se para a análise de diferentes alternativas de rastreios. Nesta fase, foram desenvolvidos diversos modelos de simulação para comparar métodos de rastreio [12], analisar a custo-efetividade do recurso à telemedicina [13], [14], e os resultados da adoção de diferentes intervalos entre rastreios [15], [16]. De entre este último grupo destacam-se os trabalhos de Davies e seus colegas que desenvolveram um modelo de simulação baseado em eventos discretos, para encenar diferentes intervalos de rastreio [3] [5] [4] [17]. Uma das conclusões mais importantes desta investigação foi a constatação de que a adesão da população ao rastreio tem um papel determinante no sucesso do mesmo. Ainda assim, nos seus modelos os autores adotam uma probabilidade fixa de aderir ao rastreio, e nenhuma tentativa é feita para modelar o comportamento individual dos sujeitos [3] [4] [5] [17]. Uma tentativa de incluir o comportamento humano nestes modelos foi feita por Schmidt e Brailsford [6], ao incorporarem o HBM-Modelo de Crenças em Saúde numa arquitetura teórica que possibilita a obtenção de um resultado (comportamento – output) através da combinação de vários fatores (inputs) que influenciam a adesão ao rastreio. Cada paciente é uma entidade individual, com suas próprias características. Esta abordagem foi implementada através de atributos numéricos, para representação de várias características do diabético (número de vezes que aderiu a rastreios anteriores, percepção do seu estado de saúde geral, estágio atual da RD, informação e ansiedade face à RD e habilitações literárias). A probabilidade de participação no rastreio foi calculada como uma variável binária e o modelo utiliza exclusivamente dados artificiais, levando a que os resultados deste modelo sejam artefactos teóricos, que carecem de validação com dados reais. Por fim, esta investigação também evidenciou a dificuldade de se incorporarem variáveis qualitativas, como as utilizadas pelo HBM, surgindo, assim, a necessidade de aplicação de outro tipo de modelos de simulação. [6]

### B. Modelos baseados em agentes

Os modelos de simulação aplicados ao estudo de programas de rastreio de RD desenvolvidos até ao momento tendem a negligenciar os aspetos relacionados com o impacto do

comportamento individual [18]. Os ABM configuram uma boa alternativa para o estudo deste tipo de sistemas, frequentemente classificados como sistemas adaptativos complexos uma vez que são constituídos por redes e processos formados por agentes interativos e adaptativos [18]. De facto, num ABM o sistema é representado como um conjunto de agentes autónomos com capacidade de tomar decisões. Em cada iteração, cada agente avalia individualmente a sua situação e toma decisões com base num conjunto de regras, executando de seguida determinada ação. Mesmo um ABM simples pode exibir padrões de comportamento complexos e fornecer informações valiosas sobre a dinâmica do sistema do mundo real que ele simula [18] [19]. De entre as principais vantagens da utilização dos ABM destacam-se, a sua capacidade de detetar anomalias que não são explicadas pelas teorias clássicas, como a adoção de comportamentos que reiteradamente não resultam no melhor desfecho; a heterogeneidade e interações entre agentes; a flexibilidade e a possibilidade de acompanhar a evolução de um sistema [19]. Assim, a literatura sugere que os ABM são uma excelente opção quando se pretende simular comportamentos e processos inerentes a sistemas complexos, como os programas de rastreio de base populacional, podendo configurar uma poderosa ferramenta de apoio à decisão para os responsáveis pela implementação de medidas de saúde pública [18] [19].

### C. Lógica fuzzy

O conceito de lógica *fuzzy* foi introduzido por Lotfi Zadeh em 1965 e baseia-se na constatação de que os seres humanos tomam decisões com base em informações imprecisas, subjetivas e não numéricas [20] [21]. Assim, conjuntos *fuzzy* são entidades matemáticas que visam representar informações imprecisas e que conferem aos modelos a capacidade de reconhecer, representar, manipular, interpretar e utilizar informações vagas e/ou subjetivas [21]. Um exemplo comum de informação subjetiva e imprecisa é a designação “quente”, quando usada para descrever a temperatura de uma sala. Certamente esta designação corresponderá a registos térmicos quantitativos muito diferentes para alguém de um clima nórdico e para alguém que reside junto ao equador. No entanto, apesar da falta de precisão da designação, ela não é aleatória e, portanto, não pode ser tratada por meios estatísticos. Uma variável *fuzzy* é um input de um modelo *fuzzy* representada por um conjunto *fuzzy* e um descritor [20].

A temperatura ambiente é a variável mostrada na Fig. 1. Três conjuntos *fuzzy*: ‘quente’, ‘frio’ e ‘confortável’ foram definidos sobre uma gama de temperaturas reais. O poder da lógica *fuzzy* é permitir a sobreposição de valores [21]. Assim, uma temperatura específica no mesmo instante pode pertencer a mais do que um conjunto. E isso permite usos mais realísticos de modelos computacionais. As funções utilizadas para definir os conjuntos *fuzzy* da Fig.1 são funções triangulares. No entanto não existem restrições relativas ao tipo de função a utilizar. Para além da distribuição triangular, as distribuições normal e linear também são comumente utilizadas [22]. A implementação da lógica *fuzzy* compreende três etapas fundamentais: A fuzziificação, onde são utilizadas as funções para determinar, para cada valor mensurável, o grau de pertença a cada conjunto *fuzzy*. Por exemplo, uma temperatura de 15°C é transformada em três valores *fuzzy*: 0.67 ‘frio’, 0.33 ‘confortável’ e 0.00 ‘quente’; A segunda etapa é a aplicação do modelo linguístico, geralmente

na forma de Regras IF – THEN; finalmente, o resultado *fuzzy* é convertido novamente num valor numérico através de um processo de desfuzzificação [22].

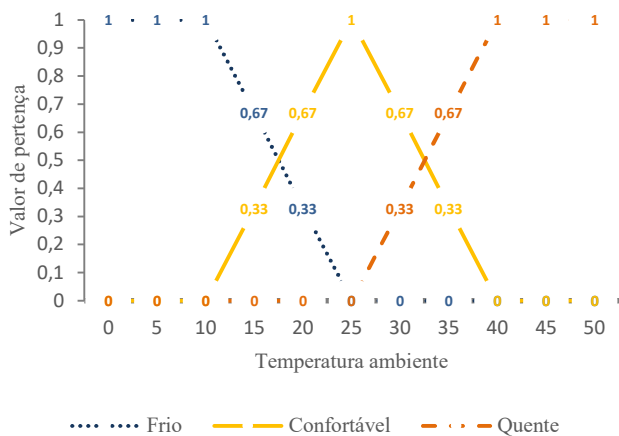


Figura 1- Funções fuzzy para a variável temperatura ambiente

### III. METODOLOGIA

#### A. Dados

Na presente investigação foram utilizados dados fornecidos pela Administração Regional de Saúde do Norte (ARSN) relativos a todas as convocatórias para rastreio de RD, entre os anos 2013 e 2018. A amostra é composta por 271,867 convocatórias para rastreio de RD, o que corresponde a 108,620 diabéticos distintos. Foram recolhidas as seguintes variáveis: Idade; Género; Local de residência; Situação profissional; Existência de contacto telefónico que possibilite o envio de lembretes; Agrupamento de Centros de saúde; Unidade de Saúde; Tipo de Unidade de Saúde; Existência ou não de Médico de Família; Motivo de isenção de Taxa Moderadora; Número de consultas na unidade de saúde nos últimos 12 meses; Tipo de Diabetes (I ou II); Índice de Massa Corporal (IMC); Níveis de glicose sanguínea (HbA1C); Mês da convocatória para rastreio; Dias decorridos entre convocatórias; Número de vezes que o diabético foi convocado; Resultado do último rastreio; e Percentagem de vezes em que o diabético aderiu a rastreios anteriores. Posteriormente, recorreu-se aos dados do Instituto Nacional de Estatística (INE) para obter as variáveis “Rendimento (Mediana)” e “Habilitações literárias”, por código postal com 7 dígitos [22], já que estas variáveis são referidas na literatura como fortemente relacionadas com a taxa de adesão [6]. Para a classificação das zonas geográficas de acordo com o grau de urbanização foram utilizados dados da Tipologia de áreas urbanas 2014 (TIPAU, 2014) disponível no portal do INE [22]. A informação referente à qualidade/estratégia de rastreio baseou-se na opinião dos peritos da ARSN e na análise dos documentos oficiais da instituição [23].

#### B. Diagrama de estados

O ABM foi desenvolvido com recurso ao software NetLogo 6.1.1., um simulador escrito em linguagem Scala. O diagrama de estados implementado (Fig. 2) contempla quatro estados

possíveis: (i) não convocado; (ii) convocado; (iii) aderiu ao rastreio; e (iv) não aderiu ao rastreio.

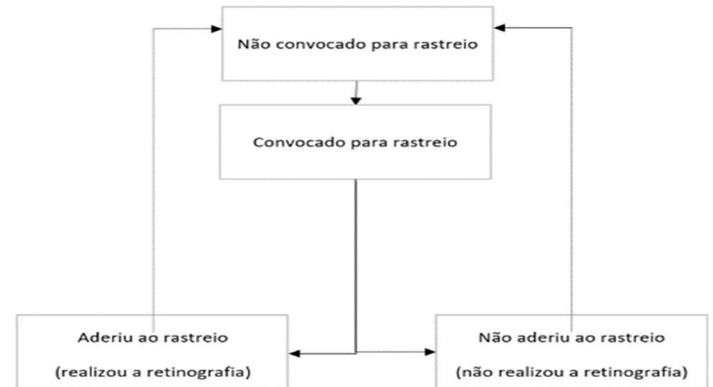


Figura 2– Diagrama de estados para a simulação da adesão ao rastreio da RD

Inicialmente todos os diabéticos assumem o estado não convocado. No início da simulação cada diabético é convocado para o rastreio através de carta enviada com quatro semanas de antecedência. Nesse momento o diabético transita provisoriamente (até à data de realização do rastreio) para o estado convocado. Na data de realização do rastreio verifica-se se o diabético compareceu, dando-se a transição para o estado aderiu ao rastreio ou para o não aderiu ao rastreio, consoante o caso. Após esta fase inicia-se um novo ciclo, em que o diabético retorna ao estado não convocado, até que decorram 52 semanas após a data da convocatória anterior (visto que se pretendem simular rastreios anuais). Nesse momento é emitida nova convocatória e o diabético assume novamente o estado de convocado, repetindo-se todo o processo.

#### C. Tomada de decisão - Lógica fuzzy

Cada vez que é emitida uma convocatória o diabético decide se adere ou não ao rastreio com base numa combinação de três componentes fuzzy, obtidas de acordo com a metodologia explicada de seguida.

Os dados da ARSN foram divididos em dois grupos: treino e teste. Os dados relativos às zonas geográficas do Tâmega e Sousa, Cavado, Douro, Trás-os-Montes e Área Metropolitana do Porto, que correspondem a 66,41% do total da população abrangida pela ARSN, foram utilizados para treino. Os dados relativos às zonas geográficas do Alto Minho, Ave e Entre Douro e Vouga, que correspondem a 33,59% da população da ARSN, foram reservados para teste. A seleção das variáveis a incluir em cada componente baseou-se na análise estatística dos dados de treino (teste t para amostras independentes, análise de variância a um fator, coeficiente de correlação linear e análise de componentes principais). Incluíram-se ainda, variáveis relativas às características do programa de rastreio identificadas na literatura como relevantes para a taxa de adesão ao programa [13], [24]. As variáveis qualitativas ordinais ou quantitativas discretas foram uniformizadas numa escala de 0 a 100. Assim, da análise de componentes principais resultaram três componentes *fuzzy*: “barreiras de acesso”, “conhecimento sobre a doença”, “qualidade/ estratégia do programa de rastreio” e foram identificadas as variáveis que compreendem cada uma destas componentes. A seleção da função representativa de cada

variável baseou-se na análise da distribuição dos dados reais e na validação através de testes estatísticos de ajustamento.

A componente “barreiras de acesso” compreende as variáveis B1, B2, B3 e B4. B1 diz respeito à percepção das barreiras de acesso em função da idade. O “difícil acesso devido à idade” é definido pela função linear que passa pelos pontos: [0, 1], [100, 0]. “Fácil acesso” é definido pela função linear que passa pelos pontos: [0, 0] e [100, 1]. A variável B2 corresponde à percepção das barreiras de acesso em função dos rendimentos. O “difícil acesso devido aos rendimentos” é definido pela distribuição normal de média 50,000 euros/ano e desvio padrão 17,000 euros/ano. A classificação “fácil acesso devido aos rendimentos” corresponde ao máximo de duas distribuições normais com médias 0 e 100,000 euros/ano respetivamente e desvios padrão 17,000 euros/ano. B3 corresponde à percepção das barreiras de acesso em função do local de rastreio. “Difícil acesso devido ao local do rastreio” é definido pela função linear que passa pelos pontos: [0, 0], [100, 1]. “Fácil acesso devido ao local do rastreio” é definido pela função linear que passa pelos pontos: [0, 1] e [100, 0]. B4 corresponde à percepção das barreiras de acesso em função do grau de urbanização do local de residência. O “difícil acesso devido ao grau de urbanização” é definido pela distribuição normal de média 0.3 e desvio padrão 0.1. A classificação “fácil acesso ao grau de urbanização” corresponde ao máximo de duas distribuições normais com médias 0 e 0.5 respetivamente e desvios padrão 0.1. A componente relativa ao conhecimento sobre a doença compreende as variáveis C1, C2 e C3. A variável C1 avalia o conhecimento sobre a doença em função da idade. O “nível de conhecimento alto devido a idade” é definido por uma distribuição normal de média 65 anos e desvio padrão 30. “Nível de conhecimento baixo devido a idade” corresponde ao máximo de duas distribuições normais com médias 18 e 100 anos respetivamente e desvios padrão 30. A variável C2 corresponde ao conhecimento sobre a doença em função das habilitações literárias. O “nível de conhecimento alto devido às habilitações literárias” é definido por uma função linear que passa pelos pontos [0, 0] e [100, 1]. “Nível de conhecimento baixo devido às habilitações literárias” é definido por uma função linear que passa pelos pontos [0, 1] e [100, 0]. A variável C3 corresponde ao conhecimento sobre a doença em função da percentagem de vezes que o agente aderiu ao rastreio anteriormente. O “nível de conhecimento alto devido à adesão prévia” é definido por uma função linear que passa pelos pontos [0, 0] e [100, 1]. “Nível de conhecimento baixo devido à adesão prévia” é definido por uma função linear que passa pelos pontos [0, 1] e [100, 0]. A componente relativa a qualidade da estratégia de rastreio compreende as variáveis E1, E2 e E3. A variável E1 corresponde à qualidade da estratégia considerando o envio de lembretes. A “elevada qualidade considerando o envio de lembretes” é definida por uma função linear que passa pelos pontos [0, 0] e [100, 1]. A “baixa qualidade considerando o envio de lembretes” é definida por uma função linear que passa pelos pontos [0, 1] e [100, 0]. A variável E2 corresponde à qualidade da estratégia considerando o tempo de espera no momento do rastreio (em minutos). A “elevada qualidade considerando o tempo de espera” é definida por uma função

linear que passa pelos pontos [0, 1] e [500, 0]. A “baixa qualidade considerando o tempo de espera” é definida por uma função linear que passa pelos pontos [0, 0] e [500, 1]. A variável E3 corresponde ao tempo (em semanas) entre o envio da convocatória e a data de realização do rastreio. A “elevada qualidade considerando a antecedência da convocatória” é definida por uma distribuição normal de média 4 e desvio padrão 2. “Baixa qualidade considerando a antecedência da convocatória” corresponde ao máximo de duas distribuições normais com médias 0 e 8 respetivamente e desvios padrão 2. Finalmente, acrescentou-se um ruído aleatório, cuja magnitude é controlada pelo parâmetro “*variability*”.

Em cada componente foram definidas regras do tipo *IF Then*, de forma que se for obtido o valor fácil/ alto/ elevado em pelo menos metade das variáveis que a integram, há uma forte probabilidade do agente aderir ao rastreio. Por conseguinte, para a componente “barreiras de acesso” foram definidas 18 regras, 9 para a componente “conhecimento sobre a doença” e 9 para a componente “qualidade/ estratégia do programa de rastreio”, resultando num total de 36 regras *IF Then*. Foi utilizado o máximo como operador de agregação, o Método de Inferência Difusa Mamdani, a desfuzificação de cada componente foi efetuada pelo método do Centro de Gravidade (COG) [21] e o resultado final corresponde à média dos resultados das três componentes.

#### IV. RESULTADOS

As simulações efetuadas visam a comparação dos resultados obtidos para cada área geográfica com as taxas de adesão reais, aferindo assim a capacidade do modelo replicar a realidade. Foi gerada uma população virtual de 1,000 diabéticos e simulado o processo de convocatória para rastreio durante 52 períodos de simulação. A população inicial de agentes foi desenhada de acordo com as características da população diabética residente em cada uma das zonas geográficas que pertencem à área de abrangência da ARSN. Uma vez que os dados reais pertencem todos à mesma Administração Regional de Saúde as características do rastreio que conseguimos apurar são semelhantes em todas as áreas geográficas (quer de treino, quer de teste), ou seja o envio de lembretes é uma prática ainda muito incipiente e o rastreio é sempre efetuado nos cuidados de saúde primários. Relativamente às outras variáveis que compreendem a componente “qualidade/estratégia de rastreio” não foi possível obter informação até ao momento. Assim, foi assumido que a convocatória é enviada com 4 semanas de antecedência em todos os locais e que o tempo de espera no dia do rastreio é sempre de 10 minutos. Realizaram-se 10 simulações. A Fig. 3 corresponde ao output gráfico da Netlogo obtido com a primeira simulação efetuada. Cada ponto corresponde a um diabético. Os pontos aglomeram-se em formas concêntricas que representam as zonas geográficas e a aproximação de cada ponto ao centro é determinada pelo escalão de rendimentos do diabético representado. A cor atribuída a cada ponto corresponde ao estado em que o diabético se encontra: cinzento não convocado ou convocado para

rastreio, mas a aguardar a data de realização do mesmo; verde aderiu ao rastreio, vermelho não aderiu.

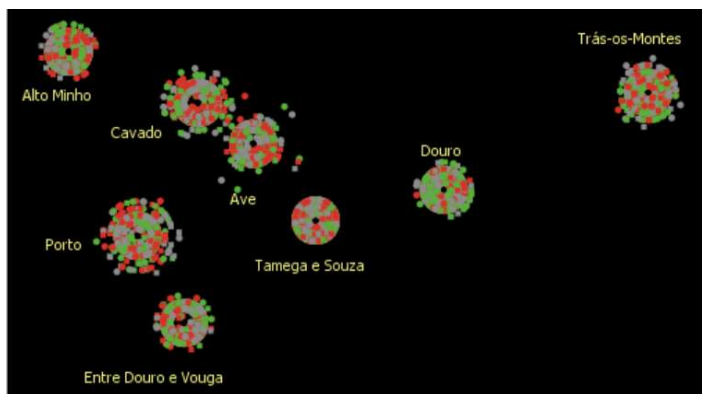


Figura 3 - Taxa de adesão por zonas geográfica – Output gráfico Netlogo

A Tabela 1 e a Fig. 4 resumem os resultados obtidos nas 10 simulações efetuadas, comparando-os com os dados reais.

TABELA 1 – TABELA RESUMO DOS RESULTADOS OBTIDOS EM 10 SIMULAÇÕES E COMPARAÇÃO COM DADOS REAIS

	NUTS II	Resultados da simulação		Resultados Reais	Diferença
		Média	Desvio Padrão		
Treino	Tâmega e Sousa	67.03%	1.13%	67.59%	0.56%
	Cávado	65.58%	1.42%	67.02%	1.44%
	Douro	65.44%	0.66%	66.86%	1.42%
	Trás-os-Montes	62.79%	0.90%	65.69%	2.90%
	Área Metrop. do Porto	62.00%	1.00%	63.38%	1.38%
Teste	Alto Minho	70.92%	0.81%	74.10%	3.18%
	Ave	70.78%	0.87%	72.09%	1.31%
	Entre Douro e Vouga	65.40%	0.93%	66.25%	0.85%

COMPARAÇÃO ENTRE OS RESULTADOS SIMULADOS E OS RESULTADOS REAIS

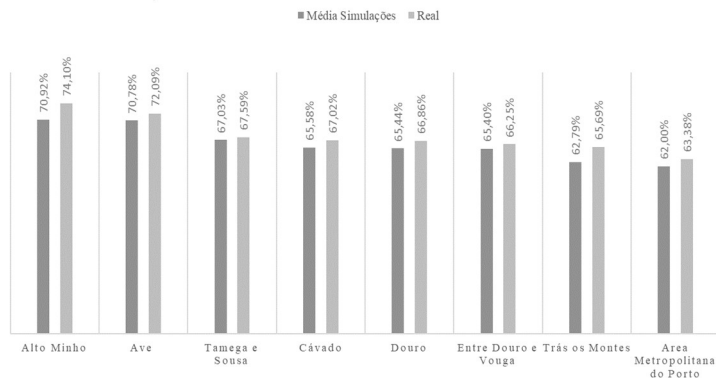


Figura 4 – Gráfico dos resultados obtidos em 10 simulações e comparação com dados reais

Como se pode observar, os resultados obtidos com as simulações estão ligeiramente abaixo da taxa adesão real em todas as regiões geográficas, registando-se a menor diferença no Tâmega e Sousa

(0.56%) e a maior no Alto Minho (3.18%). No entanto, consideramos os resultados obtidos muito satisfatórios já que o modelo consegue capturar eficazmente as nuances entre as diferentes regiões em termos de adesão ao rastreio. É também de registar que o ajuste à realidade é semelhante nos grupos de treino e de teste, levando a crer que o modelo tem capacidade preditiva em novos contextos. De seguida, executamos um novo conjunto de simulações considerando alterações na estratégia de rastreio. De acordo com o modelo, a alteração do local de rastreio dos Cuidados de Saúde Primários para os Hospitais conduziria a uma diminuição de 9.17% na taxa de adesão global da ARSN (de 65.05% para 55.88%), o que vai ao encontro da literatura que defende a importância de adotar estratégias de proximidade [24]. Já o envio de lembretes (SMS+ telefonema) conduziria a um aumento de 2.14% na taxa de adesão global (de 65.05% para 67.93%). No entanto, estes resultados carecem ainda de validação com dados reais, pelo que remetemos para trabalhos futuros o aprofundamento desta análise.

## V. CONCLUSÕES

Os resultados obtidos com o modelo desenvolvido indicam que é possível prever a taxa de adesão ao rastreio da retinopatia diabética utilizando dados demográficos e económico-sociais da população alvo, e informação relativa à estratégia de rastreio. Assim, consideramos que as técnicas utilizadas têm um grande potencial para conferir aos modelos a escalabilidade e abstração necessária para a sua ampla utilização. No entanto, uma limitação à presente investigação foi o facto de não termos acesso a dados de locais com diferentes estratégias de rastreio. De futuro pretendemos melhorar a validação do modelo, testando-o com dados provenientes de outros locais onde quer as características populacionais, quer a estratégia de rastreio difiram substancialmente das encontradas no grupo de treino. Também planeamos testar o modelo com dados provenientes de outros rastreios de base populacional e aferir a sua capacidade de replicar as taxas de adesão reais.

## AGRADECIMENTOS

Esta investigação é parcialmente financiada por fundos nacionais através da FCT - Fundação para a Ciência e Tecnologia, I.P., no âmbito do projeto FCT UIDB / 04466/2020.

Agradecemos também ao ISCTE-IUL, pelo apoio no âmbito do projeto HOPE e à Administração Regional de Saúde do Norte pela colaboração e cedência de dados, devidamente encriptados e anonimizados.

## REFERÊNCIAS BIBLIOGRÁFICA

- [1] World Health Organization, “Prevention of blindness from diabetes mellitus,” 2005.
- [2] World Health Organization/International Foundation Europe, “The Saint Vincent Declaration on diabetes care and research in Europe,” *Diabet Med* 1990, 1989.
- [3] R. Davies e C. Canning, “Discrete Event Simulation to evaluate screening for diabetic eye disease,” *Simulation*, pp. 209-216, 1996.

- [4] R. Davies, S. Brailsford, P. Roderick, C. Canning e D. Crabbe, "Using simulation modelling for evaluating screening services for diabetic retinopathy," *Journal of the Operational Research Society*, vol. 51, pp. 476-484, 2000.
- [5] R. Davies, P. Roderick, C. Canning e S. Brailsford, "The evaluation of screening policies for diabetic retinopathy using simulation," *Diabetes Medicine*, vol. 19, pp. 762-770, 2002.
- [6] S. Brailsford e B. Schmidt, "Towards incorporating human behaviour in models of health care systems: an approach using discrete event simulation," *European journal of operational research*, vol. 150, pp. 19-31, 2003.
- [7] A. Pereira, R. Laureano e F. Neto, "Adesão ao Rastreio da Retinopatia Diabética: um Modelo de Simulação Baseado em Agentes," em *CAPSI*, 2020.
- [8] J. Javitt, L. Aiello, Y. Chiang, F. Ferris, J. Canner e S. Greenfield, "Preventive eye care in people with diabetes is cost-saving to the Federal-Government: Implications for health-care reform," *Diabetes Care*, vol. 17, n° 8, pp. 909-917, 1994.
- [9] J. Javitt e L. Aiello, "Cost-effectiveness of detecting and treating diabetic retinopathy," *Annals of Internal Medicine*, vol. 125, n° 11, pp. 939-939, 1996.
- [10] A. J. Palmer, C. Weiss, P. P. Sendi, K. Neeser, A. Brandt, G. Singh, H. Wenzel e G. A. Spinaz, "The cost-effectiveness of different management strategies for Type I diabetes: a Swiss perspective," *Diabetologia*, n° 43, pp. 13-26, 2000.
- [11] D. Vetrini, C. A. Kiire, I. P. Burgess, S. P. Harding, P. C. Kayange, K. Kalua, G. Msukwa, N. A. Beare e J. Madan, "Incremental cost-effectiveness of screening and laser treatment for diabetic retinopathy and macular edema in Malawi," *PLoS ONE*, pp. 1-14, 2018.
- [12] D. Maberley, H. Walker, A. Koushik e A. Cruess, "screening for diabetic retinopathy in James Bay Ontario: a cost-effectiveness analysis.," *CMAJ*, vol. 168, pp. 160-164, 2003.
- [13] D. B. Rein, J. S. Wittenborn, X. Zhang, B. A. Allaire, M. S. Song, R. Klein e J. B. Saaddine, "The cost-effectiveness of three screening alternatives for people with no or early diabetic retinopathy," *Health services research*, pp. 1534-1561, 2011.
- [14] J. D. Whited, 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, A. W. Darkins e S.-E. Bursell, "A modeled economic analysis of a digital teleophthalmology system as used by three federal healthcare agencies for detecting proliferative diabetic retinopathy," *Telemedicine an e-health*, vol. 11, n° 6, pp. 641-651, 2005.
- [15] T. E. Day, N. Ravi, H. Xian e A. Brugh, "An agent-based modelling template for a cohort of veterans with diabetic retinopathy," *PLoS ONE*, vol. 8, 2013.
- [16] S. Vijan, T. P. Hofer e R. A. Hayward, "Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus," *Journal of the American Medical Association*, vol. 283, n° 7, pp. 889-896, 2000.
- [17] R. Davies e S. Brailsford, "Screening for diabetic retinopathy," em *Handbook of OR/MS Applications in Health Care*, M. B. a. W. P. Sainfort F, Ed., 2004, pp. 493.
- [18] U. Siebert, O. Alagoz, A. Bayoumi, B. Jahn, D. Owens, D. Cohen e K. Kuntz, "State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force," *Value in Health*, vol. 15, pp. 812-820, 2012.
- [19] D. . A. Marshall, L. Burgos-Liz, M. J. IJerman, W. Crown, W. V. Padula, P. K. Wong, K. S. Pasupathy, M. K. Higashi e N. D. Osgood, "Selecting a Dynamic Simulation Modeling Method for HealthCare Delivery Research—Part 2 Report of the ISPOR DynamicSimulation Modeling Emerging Good Practices Task Force," *Value in health*, n° 147-160, 2015.
- [20] R. R. Yager e Z. A. Lotfi A., An introduction to fuzzy logic applications in intelligent systems., vol. Vol. 165., Springer Science & Business Media, 2012.
- [21] H. Singh, M. M. Gupta, T. Meitzler, Z.-G. Hou, K. K. Garg, A. M. G. Solo e L. A. Zadeh, "Real-life applications of fuzzy logic.," *Advances in Fuzzy Systems*, 2013.
- [22] Instituto Nacional de Estatística, [Online]. Available: [https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_base\\_dados&contexto=bd&selTab=tab2&xlang=pt](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_base_dados&contexto=bd&selTab=tab2&xlang=pt).
- [23] Administração Regional de Saúde do Norte, "Relatórios Anuais de Actividades," 2007-2018.
- [24] European Association for the Study of Diabetes., "Screening for Diabetic Retinopathy in Europe –Progress Since 2011," Manchester, 2016.