

# Comparação da Diferenciação Oftalmológica no Rastreamento da Retinopatia Diabética

André Diogo Barata, MD;<sup>1,2</sup> Rita Couceiro, MD, FEBO;<sup>2,3</sup> Sara Vaz-Pereira, MD<sup>1,2</sup>

<sup>1</sup> Serviço de Oftalmologia, Hospital de Santa Maria, Lisboa, Portugal

<sup>2</sup> Clínica Universitária de Oftalmologia, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal

<sup>3</sup> Serviço de Oftalmologia, Hospital Vila Franca de Xira, Portugal

## RESUMO

**Objetivo:** Comparar a concordância na classificação da retinopatia diabética (RD) entre observadores com diferentes níveis de diferenciação oftalmológica.

**Métodos:** Estudo prospetivo de 122 indivíduos diabéticos avaliados em rastreio de RD durante 6 semanas com retinografia não-midriática 45° com aquisição de 2 campos (centrados no disco ótico e mácula). A classificação incluiu o grau de retinopatia (R– 0: ausente, 1: ligeira, 2: pré-proliferativa, 3: proliferativa, 4: não-classificável (NC)), maculopatia (M– 0: ausente, 1: presente, 2: NC), e fotocoagulação (P– 0: ausente, 1: presente, 2: NC), adaptando os critérios do rastreio britânico e foi realizada por 3 observadores com diferentes níveis de diferenciação (O1– interno do 2º ano, O2– interno do 4º ano, O3– assistente). Tratamento dos dados efetuado com o software STATA 12.0 e a análise da concordância através do cálculo dos kappa (k) para um intervalo de confiança de 95%.

**Resultados:** A idade média da população foi de 64.1±10.9 anos, 54.1% eram do sexo feminino e 94.3% tinham diabetes mellitus tipo 2. Existiu concordância absoluta entre os 3 observadores em 128 olhos (52.5%). Em R a concordância global e entre os 3 observadores foi razoável (k=0.377) e moderada em R0 (k=0.455). A maior concordância ocorreu entre O2 e O3 (k=0.420) e a menor entre O1 e O3 (k=0.347). Em M, a concordância global foi razoável (k=0.299), sendo maior na presença de maculopatia (M1, k=0.507) e menor quando NC (M2, k=0.237). Em P, a concordância entre os 3 observadores foi perfeita (k=0.85).

**Conclusões:** Este estudo demonstrou uma concordância razoável entre os 3 observadores na classificação da RD por retinografia não-midriática, apesar dos seus diferentes níveis de diferenciação. Como expectável, verificou-se em R uma maior concordância entre os observadores mais experientes. As classificações menos concordantes ocorreram no “não classificável”.

**Palavras-chave:** diabetes mellitus, retinopatia diabética, rastreio, retinografia não-midriática.

## ABSTRACT

**Objectives:** To assess the agreement in the diabetic retinopathy (DR) classification between observers with distinct levels of clinical differentiation.

**Methods:** Prospective study that included 122 diabetics screened for DR using 45-degree non-mydratic color fundus photographs and acquisition of 2 images (optic disc and macula-centered). The classification included the degree of retinopathy (R– 0: none, 1: background, 2: pre-proliferative, 3: proliferative, 4: ungradable (UG)), maculopathy (M– 0: none, 1: present, 2: UG), and photocoagulation (P– 0: none, 1: present, 2: UG), was adapted from the British screening programme and was performed by 3 independent observers with different levels of differentiation (O1– 2<sup>nd</sup> year resident, O2– 4<sup>th</sup> year resident, O3– consultant). Statistical analyses were performed using STATA software version 12.0 and the simple kappa (k) was used to assess agreement.

**Results:** Mean age was 64.1±10.9 years, 54.1% were female and 94.3% had type 2 diabetes. There was overall exact agreement in 128 eyes (52.5%) and regarding R there was also an overall fair agreement (k=0.377). R0 had a stronger agreement (k=0.455). Comparing pairs of observers, the highest agreement was between O2 and O3 (k=0.420) e the lowest between O1 and O3 (k=0.347). The overall M agreement was also fair (k=0.299), but higher in the presence of maculopathy (M1, k=0.507) and lower when UG (M2, k=0.237). In relation to P, overall agreement was perfect (k=0.85).

**Conclusions:** There was a reasonable agreement in the DR classification between observers with distinct levels of differentiation, when using a non-mydratic fundus photographs. As expected agreement was higher between more experienced observers. The less concordant classification, occurred in the “ungradable” parameter.

**Key-words:** diabetes mellitus, diabetic retinopathy, screening, agreement, non-mydratic color fundus photographs

---

## INTRODUÇÃO

A diabetes mellitus (DM) é uma doença metabólica crónica cuja incidência e prevalência está a aumentar globalmente, tornando-a numa das grandes pandemias do século XXI<sup>1-4</sup>.

De acordo com o estudo PREVADIAB 2013, a prevalência da DM na população adulta (20-79 anos) em Portugal cresceu de 11.7% para 13% entre 2009 e 2013, equivalendo atualmente a mais de 1 milhão de adultos, sendo a prevalência da doença diagnosticada de 7.3% e da não-diagnosticada de 5.7%<sup>5</sup>.

A retinopatia diabética (RD) é uma das complicações microvasculares mais comuns da DM<sup>6-7</sup> e também uma das principais causas de cegueira na população em idade activa<sup>8-9</sup>.

Estima-se que o risco de cegueira seja 25 vezes superior comparado com os indivíduos não diabéticos, levando a RD no seu estágio proliferativo a perda visual irreversível em 50% dos indivíduos 5 anos após o seu diagnóstico<sup>10</sup>.

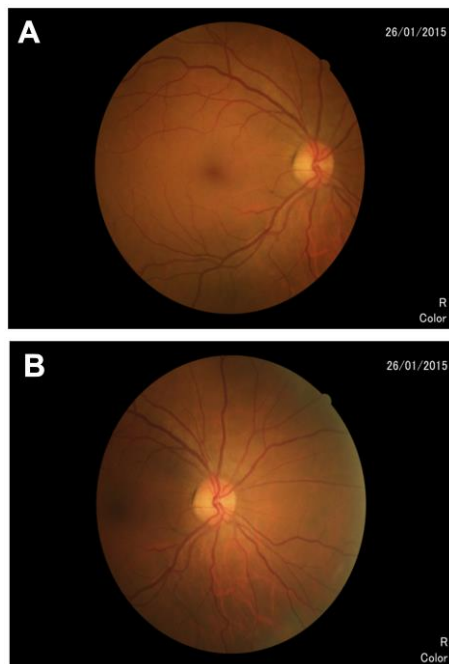
O impacto da disfunção ocular de causa diabética é globalmente reconhecido tendo vários países instituído programas de rastreio para uma deteção e tratamento precoces<sup>10</sup>, pelo que a necessidade de capacitância dos serviços no rastreio da doença ganha importância.

Este trabalho visa comparar a concordância na classificação da RD entre observadores com diferentes níveis de experiência e diferenciação oftalmológica, aspetos que poderão ter interesse na gestão de recursos e estratégias dos Serviços de Oftalmologia na abordagem ao rastreio da RD.

## MATERIAL E MÉTODOS

Estudo prospetivo e comparativo com base na população avaliada em consulta de rastreio de RD, num período de 6 semanas (5 de Janeiro de 2014 a 13 de Fevereiro de 2015), no Serviço de Oftalmologia do Hospital de Santa Maria. A origem de referenciação dos doentes é, em primeiro lugar, os cuidados de saúde primários (unidades de saúde familiar) da área deste hospital, seguida da consulta de diabetes do serviço de Endocrinologia do mesmo hospital.

Foram incluídos 122 indivíduos (244 olhos) que realizaram retinografia não-midriática de 45° (Canon CR-2 PLUS AF Digital) com aquisição de imagens com 2 campos, centrada na mácula (figura 1A) e centrada no disco ótico (figura 1B).



Figuras 1A e 1B – Retinografias centradas na mácula (campo 1) e no disco ótico (campo 2)

As retinografias foram classificadas de acordo com o grau de retinopatia (R – 0: ausente, 1: ligeira, 2: pré-proliferativa, 3: proliferativa, 4: não classificável (NC)), maculopatia (M – 0: ausente, 1: presente, 2: NC), e fotocoagulação (P – 0: ausente, 1: presente, 2: NC), adaptando os critérios do rastreio britânico de RD (National Health System Diabetic Eye Screening Programme)<sup>11</sup>.

A classificação das retinografias foi realizada por 3 observadores independentes com diferentes níveis de diferenciação (O1 – interno da formação específica (IFE) do 2º ano, O2 – IFE do 4º ano, O3 – assistente hospitalar).

O tratamento dos dados foi efetuado com o software STATA versão 12.0 (StataCorp. 2012. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP) e a análise da concordância entre observadores avaliada através do cálculo dos Kappa (k) de Fleiss e, e entre observadores múltiplos (pares) através do k de Cohen<sup>12</sup>, para um intervalo de confiança de 95%.

A concordância foi classificada em razoável se k entre 0.21-0.40, moderada se k entre 0.41-0.60, substancial se entre 0.61-0.80 e perfeita se entre k entre 0.81-1.00, de acordo com a literatura existente<sup>13</sup>.

## RESULTADOS

A idade média da população estudada foi de 64.1±10.9 anos (mediana, máxima, mínima de respetivamente, 65, 84 e 25 anos), 54.1% eram do sexo feminino e 94.3% tinham DM tipo 2.

Existiu concordância absoluta na classificação dos 3 parâmetros (R, M e P) entre os 3 observadores em 128 olhos (52.5%,  $p < 0.05$ ).

No parâmetro R, a concordância global entre os 3 observadores (tabela 1) foi razoável ( $k=0.377$ ), tendo sido mais elevada na ausência de RD (R0,  $k=0.455$ ) e mais baixa quando não-classificável (R4,  $k=0.284$ ). Nas classificações R1, R2 e R3 manteve-se razoável ( $k=0.296, 0.379, 0.315$ , respetivamente). Relativamente à comparação entre pares de observadores (tabela 2), a maior concordância ocorreu entre O2 e O3 ( $k=0.420$ ) e a menor entre O1 e O3 ( $k=0.347$ ). Analisando os resultados por subtipo classificado, verificou-se uma maior concordância entre O1 e O2 em R0 e R4 ( $k=0.571$  e  $0.524$ , respetivamente), ao passo que em R1, R2 e R3 verificou-se maior concordância entre O2 e O3 ( $k=0.558, 0.638$  e  $0.658$ , respetivamente).

Tabela 1 – Concordância global entre os 3 observadores

Variável	Subtipo	Kappa
R	R0	0,455
	R1	0,296
	R2	0,379
	R3	0,315
	R4	0,284
	<b>Combinado</b>	<b>0,377</b>
M	M0	0,287
	M1	0,507
	M2	0,237
	<b>Combinado</b>	<b>0,299</b>
P	P0	0,848
	P1	0,897
	P2	0,813
	<b>Combinado</b>	<b>0,851</b>

Tabela 2 – Concordância por subtipo e por par de observadores

Variável	Subtipo	Observadores	Kappa
R	R0	O1 vs O2	0.571
		O1 vs O3	0.443
		O2 vs O3	0.439
	R1	O1 vs O2	0.188
		O1 vs O3	0.152
		O2 vs O3	0.558
	R2	O1 vs O2	0.188
		O1 vs O3	0.161
		O2 vs O3	0.638
	R3	O1 vs O2	0.604
		O1 vs O3	0.592
		O2 vs O3	0.658
R4	O1 vs O2	0.524	
	O1 vs O3	0.327	
	O2 vs O3	0.208	
<b>Combinado</b>	O1 vs O2	0.414	
	O1 vs O3	0.347	
	O2 vs O3	0.420	
M	M0	O1 vs O2	0.459
		O1 vs O3	0.323
		O2 vs O3	0.261
	M1	O1 vs O2	0.575
		O1 vs O3	0.622
		O2 vs O3	0.323
	M2	O1 vs O2	0.432
		O1 vs O3	0.245
		O2 vs O3	0.245
	<b>Combinado</b>	O1 vs O2	0.472
		O1 vs O3	0.332
		O2 vs O3	0.262
P	P0	O1 vs O2	0.804
		O1 vs O3	0.897
		O2 vs O3	0.838
	P1	O1 vs O2	0.853
		O1 vs O3	0.853
		O2 vs O3	1.000
P2	O1 vs O2	0.753	
	O1 vs O3	0.939	
	O2 vs O3	0.707	
<b>Combinado</b>	O1 vs O2	0.807	
	O1 vs O3	0.898	
	O2 vs O3	0.840	

No parâmetro M, a concordância global (tabela 1) obtida entre os 3 observadores foi razoável (k=0.299), sendo, no entanto, maior na presença de maculopatia (M1, k=0.507) e

menor quando NC (M2, k=0.237). Por sua vez, quando comparando os resultados entre observadores, a maior concordância global ocorreu entre O1 e O2 (k=0.472). Por

subtipo (tabela 2), a concordância mais elevada verificou-se entre O1 e O2 na ausência de maculopatia e em NC (M0,  $k=0.459$ ; M2,  $k=0.432$ ) e entre O2 e O3 em M1 ( $k=0.622$ ).

Por último, relativamente ao parâmetro P, a concordância entre os 3 observadores foi perfeita ( $k=0.85$ ). Comparando pares de observadores, a maior concordância verificou-se entre O1 e O3 ( $k=0.898$ ) e a menor entre O2 e O3 ( $k=0.840$ ).

## DISCUSSÃO

Os resultados demonstraram, globalmente, uma concordância razoável entre os 3 observadores na avaliação dos achados retinográficos (R), maculopatia (M) e uma concordância perfeita relativamente à fotocoagulação laser (P).

É de notar que relativamente ao estágio de retinopatia (R) existiu uma concordância superior, moderada, na ausência de RD (R0), verificando-se a menor concordância em imagens classificadas como NC (R4). Também na análise de maculopatia (M), a menor concordância foi observada na classificação NC, o mesmo se passando no parâmetro P.

Estes resultados parecem apontar para um maior acordo na identificação da ausência de doença/exclusão dos achados retinográficos e uma menor concordância em torno das imagens tidas como não-classificáveis. A explicação para tal facto poderá advir da subjetividade inter-pares ser maior quando, por exemplo, devido a opacidade de meios ou pupilas mais mióticas da população mais idosa, as imagens apresentarem uma menor qualidade técnica (menor luminosidade, menor nitidez, periferias menos visualizáveis), constituindo um fator confundente, pelo que se deve discutir a pertinência de realizar um rastreio na ausência de midríase farmacológica, tendo em conta as características populacionais da maioria da população diabética. Por sua vez, é lícito concluir que o especialista tenderá a ser mais exigente na classificação das imagens, explicando a maior concordância entre internos nas imagens NC.

Analisando os resultados tendo em conta o grau de diferenciação oftalmológica, relativamente a R, verificou-se, como seria expectável, uma concordância crescente com o aumento da experiência oftalmológica, sendo maior a concordância verificada entre o interno mais sénior (O2) e o especialista (O3). Analisando os resultados por subtipo R, assistiu-se a uma maior concordância nas classificações

entre internos (O1 e O2) em R0 e R4, mas uma maior concordância entre o interno mais velho (O2) e o especialista (O3) com o aumento da gravidade das alterações (R1 a R3). No que diz respeito a M, verificou-se novamente uma maior concordância entre O2 e O3 na presença de maculopatia (M1), o mesmo se verificando na presença de fotocoagulação (P1). De referir que voltou a ocorrer uma maior concordância quando NC entre os dois internos, tanto em M como em P.

Os autores consideram que os resultados encontrados são expectáveis, na medida em que se compreende que, com a experiência e diferenciação oftalmológica crescente, alterações minor do fundo ocular passam menos despercebidas, existindo uma maior concordância também na identificação dos estádios de gravidade crescente. Não se verificou a tendência para uma classificação mais grave nos avaliadores menos experientes. Contudo, os resultados de concordância ligeira a moderada encontrados nos parâmetros R e M urgem a necessidade de procurar soluções e alternativas para aumentar a concordância e fiabilidade nas classificações de forma a melhorar o encaminhamento para tratamento, quando necessário, destes doentes. Por sua vez, o presente estudo não mostra evidência de que avaliadores menos experientes tenham resultados substancialmente discordantes face aos mais experientes.

Métodos automatizados para o estadiamento e classificação de retinografias têm sido recentemente discutidos na literatura, com resultados variáveis<sup>14-17</sup>. Estudos comparativos com vários equipamentos demonstraram uma elevada sensibilidade, mas pior especificidade que a classificação humana de imagens, mas sistemas como o Retmarker ou o EyeArt, mostraram uma sensibilidade e especificidade aceitáveis, tornando-os custo-efetivos face a rastreios manuais isolados<sup>15</sup>. Estes sistemas automatizados têm interesse especialmente nos países em desenvolvimento, reduzindo custos potenciais, e também em países desenvolvidos, particularmente em localizações com acesso remoto aos cuidados de saúde<sup>15</sup>.

## CONCLUSÕES

Este estudo demonstrou em termos globais uma concordância razoável entre os 3 observadores na classificação da RD por retinografia não-midriática, apesar dos seus diferentes níveis de diferenciação.

Como seria expectável na maior parte dos casos verificou-se uma maior concordância entre os observadores mais experientes.

As classificações menos concordantes ocorreram na atribuição “não classificável”, possivelmente dada a maior exigência do especialista na classificação das imagens.

Como limitações do nosso trabalho apontamos o número reduzido de doentes, a classificação da RD se ter baseado em apenas uma observação de imagens, e a avaliação dessas mesmas imagens ter sido realizada pelos observadores em diferentes dispositivos (ecrãs) com resolução distinta, o que poderá influenciar a perceção de detalhes e achados mais ténues.

Apesar destas limitações consideramos o presente estudo relevante ao procurar determinar a concordância na classificação da RD entre observadores com diferenciação oftalmológica distinta de modo a melhorar o rastreio destes doentes.

## REFERÊNCIAS

1. Henriques J, Vaz-Pereira S, Nascimento J, Rosa PC. [Diabetic eye disease]. *Acta Med Port.* 2015;28:107-13.
2. King H, Aubert RE, Herman WH. Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections. *Diabetes Care.* 1998;21:1414-31.
3. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care.* 2004;27:1047-53.
4. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract.* 2010;87:4-14.
5. Correia LG, Boavida JM, Almeida JP, et al. Diabetes: factos e números 2013 – Relatório Anual do Observatório Nacional da Diabetes 11/2013. Lisboa: OND; 2013.
6. Stratton IM, Adler AI, Neil HA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. *BMJ.* 2000;321:405-12.
7. Klein R. Hyperglycemia and microvascular and macrovascular disease in diabetes. *Diabetes Care.* 1995;18:258-68.
8. Kempner JH, O'Colmain BJ, Leske MC, et al. The prevalence of diabetic retinopathy among adults in the United States. *Arch Ophthalmol.* 2004;122:552-63.
9. Moss SE, Klein R, Klein BE. The 14-year incidence of visual loss in a diabetic population. *Ophthalmology.* 1998;105:998-1003.
10. Ling R, Ramsewak V, Taylor D, Jacob J. Longitudinal study of a cohort of people with diabetes screened by the Exeter Diabetic Retinopathy Screening Programme. *Eye.* 2002;16:140-5.
11. National Health System Diabetic Eye Screening Programme, v 1.4, 1st November 2012.
12. Fleiss JL, Levin B, Paik MC (2003). *Statistical Methods for Rates and Proportions* (3th ed.). New York: Wiley.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
14. Kapetanakis VV, Rudnicka AR, Liew G, et al. A study of whether automated Diabetic Retinopathy Image Assessment could replace manual grading steps in the English National Screening Programme. *J Med Screen.* 2015;22:112-8.
15. Tufail A, Rudisill C, Egan C, et al. Automated Diabetic Retinopathy Image Assessment Software - Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders. *Ophthalmology.* 2017;124:343-351.
16. Kirkizlar E, Serban N, Sisson JA, et al. Evaluation of telemedicine for screening of diabetic retinopathy in the Veterans Health Administration. *Ophthalmology.* 2013;120:2604–2610.
17. Abramoff MD, Folk JC, Han DP, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol.* 2013;131:351-7.

---

## CONTACTO

Sara Vaz-Pereira  
Serviço de Oftalmologia, Hospital de Santa Maria  
Av. Prof. Egas Moniz  
1649-035 Lisboa  
E-mail: saravazpereira@gmail.com