

Concordancias intra e interobservador entre oftalmólogos y médicos y enfermeras de Atención Primaria sobre 210 retinografías tomadas a personas con diagnóstico de diabetes en Centros de Salud del SES.

Verónica Martín Galán, Sonia Becerra Becerra, Galo A Sánchez Robles. *Concordancias intra e interobservador entre oftalmólogos y médicos y enfermeras de Atención Primaria sobre 210 retinografías tomadas a personas con diagnóstico de diabetes en Centros de Salud del SES*. Web evalmed.es, 5-mar-2021. Disponible en: <http://evalmed.es/2021/03/05/concordancias-entre-oftalmologos-y-profesionales-de-ap-sobre-210-retinografias/>

INTRODUCCIÓN

La concordancia intra-observador de los diagnósticos de las retinografías enviadas vía telemedicina desde los centros de salud a los servicios de oftalmología no ha sido evaluada en el SES en los oftalmólogos que las categorizan. Tampoco la concordancia inter-observador de cada oftalmólogo respecto a cada uno de los demás tomados sucesivamente como referentes. El presente estudio pretende estimar los resultados de ambas comparaciones. Y más allá de esto, a condición de que ambas comparaciones ofrezcan resultados aceptables, pretende estimar estas mismas medidas frente a una muestra de profesionales sanitarios de atención primaria. Nuestros resultados podrían servir para sugerir ampliar o no el papel diagnóstico de los profesionales de atención primaria, con o sin un determinado curso de formación específica en el diagnóstico de retinopatía diabética mediante retinografía.

OBJETIVO

Análisis de la concordancia intra e inter-observador entre oftalmólogos y médicos y enfermeras de Atención Primaria sobre 210 retinografías tomadas a personas con diagnóstico de diabetes en Centros de Salud del SES.

DISEÑO

Estudio trasversal descriptivo sobre las categorizaciones realizadas por los profesionales participantes tras la observación de las retinografías anonimizadas que se les presentan, mediante un cuestionario con las 6 categorías diagnósticas predefinidas en el cribado que se realiza actualmente en el SES.

MATERIAL Y MÉTODOS

El banco retinográfico del SES para este estudio está formado por los registros anonimizados de todas las 12.152 retinografías de 3 campos oculares realizadas en Centros de Salud del SES desde enero 2015 a abril 2019, en el “Programa de Detección Precoz de Retinopatía Diabética en Atención Primaria (del SES)”, que hayan sido estimadas por los oftalmólogos en una de las 6 categorías diagnósticas:

1. Imágenes no valorables
2. Imágenes normales (sin hallazgos patológicos)
3. Hallazgos compatibles con Retinopatía Diabética leve-moderada
4. Hallazgos compatibles con Retinopatía Diabética moderada-severa
5. Hallazgos compatibles con Retinopatía Diabética proliferativa
6. Hallazgos compatibles con Edema de Mácula

A los efectos de este estudio, la categoría estimada por los oftalmólogos se consideró únicamente como orientativa, tomándose como un medio instrumental para obtener una variabilidad en la selección de la muestra que garantizara la mayor discriminación en la gravedad e intensidad de la enfermedad.

Del banco retinográfico se tomó un conjunto de 174 retinografías, constituido por una selección aleatoria de 29 de cada una de las seis categorías. Para ello se aleatorizaron los códigos de todas las retinografías del banco, y se situaron en el orden aparecido en un listado, del que se tomaron las 29 primeras de cada categoría, para obtener las $29 \times 6 = 174$.

Con las 174 del conjunto y 36 repetidas (siendo éstas las primeras 6 obtenidas aleatoriamente de cada una de las seis categorías) se construyó la muestra de exposición de 210 retinografías, cuyo orden de aparición también se asignó aleatoriamente.

La muestra de exposición de 210 retinografías se presentó en ese mismo orden a los participantes en sucesivas tandas de 30 un día a la semana durante siete semanas consecutivas. La autora principal, VMG, facilitó desde la Coordinación de Telemedicina la exposición de las imágenes retinográficas en las pantallas los ordenadores de sus respectivos puestos de trabajo, además de un cuaderno digital de recogida de datos con las 6 categorías diagnósticas, de modo que el profesional marcaba la opción que estimaba durante la visualización.

SIGNIFICACIÓN ESTADÍSTICA PARA EL ANÁLISIS Y TAMAÑO DE LA MUESTRA

Esperando en los participantes un Porcentaje de Coincidencia intra-observador e inter-observador del 58% al 82% (con promedio en 70%), y asumiendo un error alfa del 5%, harían falta un mínimo de 56 participantes, que no incluyen a los oftalmólogos de referencia.

MUESTRA DE PARTICIPANTES VOLUNTARIOS

Tras una solicitud para participar voluntariamente a todos los oftalmólogos que valoran retinografías diabéticas en los hospitales del SES, y, anónimamente, a los médicos y enfermeras/os de varios centros de salud del SES, se consiguió una muestra de conveniencia de voluntarios hasta alcanzar el tamaño mínimo calculado para este estudio, que estuvo compuesta por: **a)** once oftalmólogos de los hospitales del SES que venían interpretando las retinografías que se les envían desde los centros de salud, vía telemedicina, dentro del *Programa de Detección Precoz de RD en AP del SES*; **b)** ocho Residentes de Medicina de Familia, siete Residentes de Enfermería, nueve Médicos y ocho Enfermeras/os de Atención Primaria, todos los cuales Sí recibieron previamente el *“Curso de cribado de retinopatía diabética mediante retinografía”*; y **c)** nueve Residentes de Medicina de Familia, nueve Residentes de Enfermería, siete Médicos y ocho Enfermeras/os de Atención Primaria, todos los cuales NO recibieron previamente el *“Curso de cribado de retinopatía diabética mediante retinografía”*.

ESTADÍSTICOS PARA LAS MEDICIONES EN LAS VARIABLES

Los fenómenos que se miden son la concordancia-discordancia en las observaciones de la misma imagen retinográfica: **a)** entre dos observadores independientes en el mismo momento, denominada inter-observador; y **b)** del mismo observador en dos momentos distintos, denominada intra-observador.

Los estadísticos utilizados para la estimación de cada concordancia-discordancia individual son: Porcentaje de Coincidencia, Índice kappa de Cohen, Coeficiente de Consistencia y V de Cramer.

Las mediciones individuales son las siguientes:

1) Concordancia inter-observador entre cada oftalmólogo tomado como referente y cada uno del resto de oftalmólogos, al categorizar las 210 retinografías en: a) “Sí-Patológica, NO-Patológica o NO-Valorable”; b) “Sí-Patológica y [Leve-Moderada, Moderada-Severa o Proliferativa]”; c) “Sí-Patológica y [Sí-Proliferativa o NO-Proliferativa]”; d) Sí-Patológica y [Sí-Edema Macular o NO-Edema Macular”.

2) Concordancia intra-observador entre cada participante respecto a sí mismo, al categorizar en “Sí-Patológica, NO-Patológica o NO-Valorable” las 36 retinografías repetidas y aleatoriamente repartidas entre las 210 observadas.

3) Concordancia inter-observador entre cada oftalmólogo tomado como referente y cada uno del resto de participantes no oftalmólogos, al categorizar las 210 retinografías en “Sí-Patológica, NO-Patológica o NO-Valorable”.

Tras reunir a los participantes individuales en 9 subgrupos, se calcula la media de cada estadístico de concordancia respecto al primer oftalmólogo, tomado como referente. Esto mismo se hace respecto al segundo, al tercero y así sucesivamente hasta el décimo-primer, obteniéndose 11 medias distintas de cada uno de los 9 subgrupos. Asumiendo que esas 11 medias forman parte de una distribución poblacional de medias, calculamos “la MEDIA y desviación estándar de las 11 medias” de cada uno de los 9 subgrupos, como estimación más aproximada de nuestros estadísticos de “concordancia respecto a un **OFTALMÓLOGO REFERENTE PROMEDIO**”. Estas “MEDIAS de las 11 medias”, además de informarnos de sus respectivos grados de concordancia, nos permiten las comparaciones estadísticas entre sí.

RESULTADOS

Las exposiciones de las 210 imágenes retinográficas se llevaron a cabo desde el 13 de enero de 2020 hasta el 1 de marzo de 2020. La obtención de los datos y las posteriores operaciones y relaciones con ellos se realizaron desde el 10 de marzo hasta el 10 abril de 2020. Los resultados de las relaciones se muestran detalladamente en los **suplementos**, porque su amplitud aconseja mostrar aquí abajo sólo un resumen que mantenga todo el rigor inferencial.

A) CONCORDANCIAS DE TODOS LOS PARTICIPANTES

1º Concordancias INTER-observador tras categorizar cada una de las 210 retinografías como: “Sí-Patológica”, “NO-Patológica”, o “NO-Valorable”:

1) Las concordancias inter-observador por grupos y subgrupos se muestran resumidamente en la **Tabla Conc inter, G+SG, %, k – 1a**.

Tabla Conc inter, G+SG, %, k - 1a: Concordancias inter-observador MEDIAS de cada Grupo y Subgrupo de participantes respecto al "OFTALMÓLOGO REFERENTE PROMEDIO" (media de los 11 oftalmólogos tomados cada uno de ellos sucesivamente como referentes), al categorizar cada una de las 210 retinografías como "Sí-Patológica, NO-Patológica, o NO-Valorable".		
Participantes agrupados en Grupos y Subgrupos	Respecto al OFTALMÓLOGO REFERENTE PROMEDIO	
	Porcentaje en las que coinciden (Desviación Estándar)	Índice kappa de Cohen (Desviación Estándar)
Grupo de oftalmólogos (GO), n= 11	78% (DE 6%)	0,63 (DE 0,09)
Grupo de intervención (GI), n= 32	74% (DE 7%)	0,56 (DE 0,11)
Grupo de intervención (GI), Subgrupo MED, n= 9	71% (DE 10%)	0,51 (DE 0,15)
Grupo de intervención (GI), Subgrupo MIR, n= 8	75% (DE 6%)	0,59 (DE 0,15)
Grupo de intervención (GI), Subgrupo ENF, n= 8	76% (DE 5%)	0,60 (DE 0,07)
Grupo de intervención (GI), Subgrupo EIR, n= 7	72% (DE 7%)	0,54 (DE 0,10)
Grupo de control (GC), n= 35	70% (DE 9%)	0,48 (DE 0,14)
Grupo de control (GC), Subgrupo MED, n= 7	66% (DE 7%)	0,43 (DE 0,10)
Grupo de control (GC), Subgrupo MIR, n= 7	71% (DE 8%)	0,49 (DE 0,10)
Grupo de control (GC), Subgrupo ENF, n= 8	72% (DE 7%)	0,51 (DE 0,12)
Grupo de control (GC), Subgrupo EIR, n= 11	70% (DE 11%)	0,49 (DE 0,17)

Abreviaturas: GO: grupo de oftalmólogos; GI: grupo de intervención que recibe curso formación; GC: grupo de control sin curso de formación; OFT: oftalmólogo; MED: médico; MIR: médico interno residente; ENF: enfermera; EIR: enfermera interna residente; DE: desviación estándar.

El caso real que mejor representa a las Medias del Grupo de oftalmólogos (GO), n= 11

"Sí-Patológica, NO-Patológica o NO-Valorable"		Referente:	OFT	11	GO	
Comparador	Observadas	Sí	No	NO-Valorable	Total	
OFT	Sí-Patológica	104	2	1	107	
4	NO-Patológica	21	23	6	50	
GO	NO-Valorable	9	6	38	53	
	Total	134	31	45	210	
Conciden en		Cfte Consistencia	Máx Cfte Consist		V de Cramer	Índice Kappa de Cohen
165 de 210		78,6%	0,67	0,82	82,1%	0,64
						0,63

El caso real que mejor representa a las Medias del Grupo de intervención (GI), n= 32

"Sí-Patológica, NO-Patológica o NO-Valorable"		Referente:	OFT	1	GO	
Comparador	Observadas	Sí	No	NO-Valorable	Total	
mir	Sí-Patológica	100	14	14	128	
49	NO-Patológica	9	20	12	41	
GI	NO-Valorable	3	2	36	41	
	Total	112	36	62	210	
Conciden en		Cfte Consistencia	Máx Cfte Consist		V de Cramer	Índice Kappa de Cohen
156 de 210		74,3%	0,62	0,82	75,9%	0,56
						0,56

El caso real que mejor representa a las Medias del Grupo de control (GC), n= 35

"Sí-Patológica, NO-Patológica o NO-Valorable"		Referente:	OFT	1	GO	
Comparador	Observadas	Sí	No	NO-Valorable	Total	
mir	Sí-Patológica	83	18	7	108	
48	NO-Patológica	2	10	1	13	
GC	NO-Valorable	27	8	54	89	
	Total	112	36	62	210	
Conciden en		Cfte Consistencia	Máx Cfte Consist		V de Cramer	Índice Kappa de Cohen
147 de 210		70,0%	0,58	0,82	70,7%	0,50
						0,49

2) Las diferencias en las concordancias inter-observador por Grupos se muestran resumidamente en la **Tabla DifConc inter, G, % – 2a** y **Tabla DifConc inter, G, k – 2b**.

En el Porcentaje de coincidencias no se encuentran diferencias estadísticamente significativas entre el Grupo de Oftalmólogos y el Grupo de Intervención. Y sí se encuentran entre el Grupo de Oftalmólogos y el Grupo de Control.

Tabla DifConc inter, G, % - 2a: Diferencias de Medias por Grupo entre los Porcentajes de Coincidencias (Desviación Estándar) respecto al OFTALMÓLOGO REFERENTE PROMEDIO, en la categorización de las 210 retinografías como: "SI-Patológica, NO-Patológica, o NO-Valorable".

Diferencia de Medias (IC 95%) entre la Fila menos la Columna	Grupo de oftalmólogos (GO), OFT; Coinciden en un 78% (DE 6%), n= 11	Grupo de intervención (GI), MED+MIR+ENF+EIR; Coinciden en un 74% (DE 7%), n= 32	Grupo de control (GC), MED+ENF+MIR+EIR; Coinciden en un 70% (DE 9%), n= 35
Grupo de oftalmólogos (GO), OFT; Coinciden en un 78% (DE 6%), n= 11		4% (-1% a 10%)	8% (2% a 14%)
Grupo de intervención (GI), MED+MIR+ENF+EIR; Coinciden en un 74% (DE 7%), n= 32			4% (-3% a 8%)
Grupo de control (GC), MED+ENF+MIR+EIR; Coinciden en un 70% (DE 9%), n= 35			

Abreviaturas: GO: grupo de oftalmólogos; GI: grupo de intervención que recibe curso formación; GC: grupo de control sin curso de formación; OFT: oftalmólogo; MED: médico; MIR: médico interno residente; ENF: enfermera; EIR: enfermera interna residente; DE: desviación estándar.

En el Índice kappa de Cohen no se encuentran diferencias estadísticamente significativas entre el Grupo de Oftalmólogos y el Grupo de Intervención. Y sí se encuentran entre el Grupo de Oftalmólogos y el Grupo de Control, y entre el Grupo de Intervención y el Grupo de Control.

Tabla DifConc inter, G, k - 2b: Diferencias de Medias por Grupo entre los Índices kappa de Cohen (Desviación Estándar) respecto al OFTALMÓLOGO REFERENTE PROMEDIO, en la categorización de las 210 retinografías como: "SI-Patológica, NO-Patológica, o NO-Valorable".

Diferencia de Medias (IC 95%) entre la Fila menos la Columna	Grupo de oftalmólogos (GO), kappa Cohen 0,63 (DE 0,09), n= 11	Grupo de intervención (GI), MED+MIR+ENF+EIR; kappa Cohen 0,56 (DE 0,11), n= 32	Grupo de control (GC), MED+ENF+MIR+EIR; kappa Cohen 0,48 (DE 0,14), n= 35
Grupo de oftalmólogos (GO), kappa Cohen 0,63 (DE 0,09), n= 11		0,07 (-0,01 a 0,15)	0,15 (0,05 a 0,25)
Grupo de intervención (GI), MED+MIR+ENF+EIR; kappa Cohen 0,56 (DE 0,11), n= 32			0,07 (0,01 a 0,14)
Grupo de control (GC), MED+ENF+MIR+EIR; kappa Cohen 0,48 (DE 0,14), n= 35			

Abreviaturas: GO: grupo de oftalmólogos; GI: grupo de intervención que recibe curso formación; GC: grupo de control sin curso de formación; OFT: oftalmólogo; MED: médico; MIR: médico interno residente; ENF: enfermera; EIR: enfermera interna residente; DE: desviación estándar.

Mostramos con más detalle estos resultados por grupos y todos los resultados por subgrupos en el **suplemento 1**.

2º Concordancias INTRA-observador tras categorizar como: "Sí-Patológica, "NO-Patológica", o "NO-Valorable" cada participante respecto a sí mismo 36 retinografías repetidas aleatoriamente dentro de las 210 observadas.

1) Las concordancias INTRA-observador por grupos y subgrupos se muestran resumidamente en la **Tabla Conc INTRA, G+SG, %, k – 3a**.

Tabla Conc INTRA, G+SG, %, k - 3a: Concordancias INTRA-observador MEDIAS de cada Grupo y Subgrupo de participantes respecto al "SÍ MISMOS" al categorizar en "SÍ-Patológica, NO-Patológica o NO-Valorable" las 36 retinografías repetidas y repartidas aleatoriamente entre las 210 observadas.

Participantes agrupados en Grupos y Subgrupos	Respecto al SÍ MISMOS	
	Porcentaje en las que coinciden (Desviación Estándar)	Índice kappa de Cohen (Desviación Estándar)
Grupo de oftalmólogos (GO), n= 11	80% (DE 6%)	0,63 (DE 0,09)
Grupo de intervención (GI); n= 32	76% (DE 8%)	0,55 (DE 0,14)
Grupo de intervención (GI), Subgrupo MED, n= 9	77% (DE 9%)	0,51 (DE 0,18)
Grupo de intervención (GI), Subgrupo MIR, n= 8	75% (DE 7%)	0,53 (DE 0,11)
Grupo de intervención (GI), Subgrupo ENF, n= 8	78% (DE 7%)	0,59 (DE 0,13)
Grupo de intervención (GI), Subgrupo EIR, n= 7	75% (DE 8%)	0,59 (DE 0,10)
Grupo de control (GC), n= 35	73% (DE 9%)	0,48 (DE 0,12)
Grupo de control (GC), Subgrupo MED, n= 7	69% (DE 10%)	0,42 (DE 0,14)
Grupo de control (GC), Subgrupo MIR, n= 7	77% (DE 8%)	0,51 (DE 0,10)
Grupo de control (GC), Subgrupo ENF, n= 8	75% (DE 5%)	0,46 (DE 0,10)
Grupo de control (GC), Subgrupo EIR, n= 11	70% (DE 10%)	0,50 (DE 0,13)

Abreviaturas: GO: grupo de oftalmólogos; GI: grupo de intervención que recibe curso formación; GC: grupo de control sin curso de formación; OFT: oftalmólogo; MED: médico; MIR: médico interno residente; ENF: enfermera; EIR: enfermera interna residente; DE: desviación estándar.

El caso real que mejor representa a las Medias del Grupo de oftalmólogos (GO), n= 11

"SÍ-Patológica, NO-Patológica o NO-Valorable"		Referente:	OFT	8	GO	
Comparador	Observadas		Sí	No	NO-Valorable	Total
OFT	SÍ-Patológica		5	2	0	7
8	NO-Patológica		1	20	1	22
GO	NO-Valorable		1	2	4	7
	Total		7	24	5	36
Concide consigo misma/o en		Coef Consistencia	Máx Coef Consist		V de Cramer	Índice Kappa de Cohen
29 de 36 80,6%		0,67	0,82	82%	0,64	0,63

El caso real que mejor representa a las Medias del Grupo de intervención (GI), n= 32

"SÍ-Patológica, NO-Patológica o NO-Valorable"		Referente:	ENF	23	GI	
Comparador	Observadas		Sí	No	NO-Valorable	Total
ENF	SÍ-Patológica		3	1	1	5
23	NO-Patológica		2	18	3	23
GI	NO-Valorable		1	1	6	8
	Total		6	20	10	36
Concide consigo misma/o en		Coef Consistencia	Máx Coef Consist		V de Cramer	Índice Kappa de Cohen
27 de 36 75,0%		0,61	0,82	74%	0,54	0,55

El caso real que mejor representa a las Medias del Grupo de control (GC), n= 35

"Sí-Patológica, NO-Patológica o NO-Valorable"		Referente:	mir	45	GC	
Comparador	Observadas	Sí	No	NO-Valorable	Total	
mir	Sí-Patológica	3	5	0	8	
45	NO-Patológica	1	18	1	20	
GC	NO-Valorable	1	2	5	8	
	Total	5	25	6	36	
Concide consigo misma/o en		Coef Consistencia	Máx Coef Consist		V de Cramer	Índice Kappa de Cohen
26 de 36	72,2%	0,61	0,82	74%	0,54	0,49

2) Las diferencias en las concordancias INTRA-observador por Grupos se muestran resumidamente en la **Tabla DifConc INTRA, G, % – 3a** y **Tabla DifConc INTRA, G, k – 3b**.

En el Porcentaje de coincidencias no se encuentran diferencias estadísticamente significativas entre el Grupo de Oftalmólogos y el Grupo de Intervención. Y sí se encuentran entre el Grupo de Oftalmólogos y el Grupo de Control.

Tabla DifConc INTRA, G, % - 3a: .Diferencias entre las Medias de los 3 Grupos en los % de Coincidencias de cada participante consigo mismo en 36 retinografías repetidas y repartidas aleatoriamente entre las 210 observadas.

Diferencia de Medias (IC 95%) entre la Fila menos la Columna	GO, OFT; Coinciden consigo en 80% (DE 6%), n= 11	GI, MED+MIR+ENF+EIR; Coinciden consigo en en 76% (DE 8%), n= 32	GC, MED+ENF+MIR+EIR; Coinciden consigo en 73% (DE 9%), n= 35
GO, OFT; Coinciden consigo en 80% (DE 6%), n= 11		4% (-2% a 9%)	7% (0,5% a 13%)
GI, MED+MIR+ENF+EIR; Coinciden consigo en en 76% (DE 8%), n= 32			3% (-1% a 7%)
GC, MED+ENF+MIR+EIR; Coinciden consigo en 73% (DE 9%), n= 35			

Abreviaturas: GO: grupo de oftalmólogos; GI: grupo de intervención que recibe curso formación; GC: grupo de control sin curso de formación; OFT: oftalmólogo; MED: médico; MIR: médico interno residente; ENF: enfermera; EIR: enfermera interna residente; DE: desviación estándar.

En el Índice kappa de Cohen no se encuentran diferencias estadísticamente significativas entre el Grupo de Oftalmólogos y el Grupo de Intervención. Y sí se encuentran entre el Grupo de Oftalmólogos y el Grupo de Control, y entre el Grupo de Intervención y el Grupo de Control.

Tabla DifConc INTRA, G, k - 3b: Diferencias entre las Medias de los 3 Grupos en los Índices kappa de Cohen de cada participante consigo mismo en 36 retinografías repetidas aleatoriamente dentro de las 210 observadas.

Diferencia de Medias (IC 95%) entre la Fila menos la Columna	GO, OFT; kappa de Cohen consigo 0,63 (DE 0,09), n= 11	GI, MED+MIR+ENF+EIR; kappa de Cohen consigo en 0,55 (DE 0,14), n= 32	GC, MED+ENF+MIR+EIR; kappa de Cohen consigo 0,48 (DE 0,12), n= 35
GO, OFT; kappa de Cohen consigo 0,63 (DE 0,09), n= 11		0,07 (-0,02 a 0,17)	0,15 (0,07 a 0,23)
GI, MED+MIR+ENF+EIR; kappa de Cohen consigo en 0,55 (DE 0,14), n= 32			0,08 (0,02 a 0,14)
GC, MED+ENF+MIR+EIR; kappa de Cohen consigo 0,48 (DE 0,12), n= 35			

Abreviaturas: GO: grupo de oftalmólogos; GI: grupo de intervención que recibe curso formación; GC: grupo de control sin curso de formación; OFT: oftalmólogo; MED: médico; MIR: médico interno residente; ENF: enfermera; EIR: enfermera interna residente; DE: desviación estándar.

Mostramos con más detalle estos resultados por grupos y todos los resultados por subgrupos en el **suplemento 2**.

B) CONCORDANCIAS INTER-OBSERVADOR ADICIONALES SÓLO ENTRE OFTALMÓLOGOS DENTRO DE LAS “Sí-Patológicas”.

Concordancias al categorizar las retinografías como se especifica en cada línea.

Cuando en la línea se indica como condición previa Sí-Patológica, como por ejemplo en “Sí-Patológica y [Leve-Moderada, Moderada-Severa o Proliferativa]”, entonces se incluye sólo la concordancia entre las que ambos observadores consideran simultáneamente como Sí-Patológicas, tras observar las 210 retinografías. Por tanto, no se incluyen las discordancias entre las consideradas por un observador como leve-moderada, moderada-severa o proliferativa cuando el otro las considera no patológicas o no valorables, y viceversa.

Las concordancias se muestran resumidamente en la **Tabla Conc inter, OFT, %, k – 4a**.

Tabla Conc inter, OFT, %, k - 4a: Concordancias inter-observador MEDIAS del Grupo de Oftalmólogos respecto al "OFTALMÓLOGO REFERENTE PROMEDIO" (media de los 11 oftalmólogos tomados cada uno de ellos sucesivamente como referentes), al categorizar cada una de las 210 retinografías como se especifica en cada línea.

Los 11 oftalmólogos agrupados en el Grupo de oftalmólogos (GO)	Respecto al OFTALMÓLOGO REFERENTE PROMEDIO	
	Porcentaje en las que coinciden (Desviación Estándar)	Índice kappa de Cohen (Desviación Estándar)
Sí-Patológica, NO-Patológica o NO-Valorable	78% (DE 6%)	0,63 (DE 0,09)
Sí-Patológica y [Leve-Moderada, Moderada-Severa o Proliferativa] (*)	68% (DE 11%)	0,40 (DE 0,12)
Sí-Patológica y [Sí-Proliferativa o NO-Proliferativa] (*)	86% (DE 6%)	0,39 (DE 0,15)
Sí-Patológica y [Sí-EdemaMacular o NO-EdemaMacular] (*) (**)	78% (DE 6%)	0,49 (DE 0,13)

Abreviaturas: GO: grupo de oftalmólogos; OFT: oftalmólogo; DE: desviación estándar.

(*) Cuando en la línea se indica como condición previa Sí-Patológica. entonces se incluye sólo la concordancia entre las que ambos observadores consideran simultáneamente como Sí-Patológicas, tras observar las 210 retinografías. Por tanto, no se incluyen las discordancias entre las consideradas por un observador como leve, moderada o proliferativa cuando el otro las considera no patológicas o no valorables, y viceversa.

(**) Se ha eliminado al oftalmólogo nº 7 porque no rellenó este apartado.

Mostramos con más detalle estos resultados en los **suplementos 3a, 3b y 3c**.

DISCUSIÓN

Algunos autores han investigado el grado de acuerdo en el diagnóstico de la retinopatía diabética mediante imágenes retinográficas tomando a un oftalmólogo como “estándar de oro” (gold standard), obteniendo los valores de sensibilidad, especificidad, valor predictivo positivo (VPP) y valor predictivo negativo de los profesionales (generalmente de atención primaria) que funcionan como grupo de comparación (1-4). Sin embargo, esta operación no produciría un resultado materialmente correcto si la asunción de un oftalmólogo como “estándar de oro” fuera errónea, porque varios oftalmólogos clasificaran sus observaciones distintamente, y esto se tradujera en que un mismo comparador obtuviera distintos valores predictivos positivos y negativos con dos o más oftalmólogos referentes. Y, aunque algunos autores han observado diferencias de categorización entre dos oftalmólogos independientes (5), la mayoría han utilizado más de dos oftalmólogos como referentes, dando la idea de que sus categorizaciones son iguales (6-8).

Nosotros hemos contado con once oftalmólogos y hemos comprobado que cada uno tomado como referente produce un variable grado de acuerdo con cada uno de los otros al clasificar las retinografías observadas, razón por la cual hemos obtenido las medias (con sus respectivas desviaciones estándar) de los estadísticos que hemos utilizado para estimar el grado de concordancia como grado de acuerdo. Estos han sido el porcentaje de coincidencia, el índice kappa de Cohen, la V de Cramer y el Coeficiente de Consistencia.

El porcentaje de coincidencia es fácil de comunicar porque incluso los lectores no sanitarios entienden su significado, pero no nos da idea de la distribución de las no coincidencias.

El índice kappa de Cohen nos indica la concordancia no debida al azar entre dos observadores independientes en variables nominales, y es fácil de comunicar porque oscila entre -1 (discordancia perfecta), pasando por 0 (concordancia nula), hasta 1 (concordancia perfecta). Usualmente se toman los siguientes valores de kappa, según propusieron Landis y Koch (9), como grados de acuerdo descriptivos: no acuerdo (< 0), insignificante (0 a 0,2), bajo (0,2 a 0,4), moderado (0,4 a 0,6), bueno (0,6 a 0,8), muy bueno (0,8 a 1). Una de las limitaciones de kappa es que puede haber más sesgos cuantas más opciones haya desde las tres categorías en adelante, sobre todo si las categorías nominales pueden concebirse también como ordinales, como sucede cuando el desacuerdo entre lo normal y lo grave puede tener más peso que entre lo normal y lo leve, razón por la cual algunos autores utilizan una kappa ponderada, o kappa corregida tras la asignación de pesos, generalmente por el canónico método lineal o el cuadrático (5,10). Otra limitación es que, a igualdad de coincidencias, cuando una categoría tiene una prevalencia mucho más baja que la complementaria, el índice kappa es mucho más bajo que cuando las prevalencias son similares. Este es el caso que se observa en las Sí-Patológica y [Sí-Proliferativa o NO-Proliferativa], en el que, a igualdad de coincidencias (86 de 98), kappa es 0,39 cuando surge de “coinciden en 5+81= 86 de 98”, debido a que la prevalencia de las Sí-Proliferativa es baja respecto a su complementario, y kappa sube a 0,76 cuando con el mismo porcentaje de coincidencias, surge de “coinciden en 43+43 = 86 de 98”, que hubiera surgido si las prevalencias hubieran sido similares. Ciertamente, hubiéramos podido introducir la corrección por prevalencia denominada PABAK (11), pero hemos preferido no hacer ninguna corrección matemática del fenómeno observado, porque, a medida que nos alejamos de la realidad material observada, estas correcciones en las fórmulas (que pretenden representar la realidad en el mundo de las formas, de las esencias), pueden no haber retenido la realidad cuando descienden del mundo de las formas (de las esencias) para producirse operatoriamente en el mundo material. Hemos preferido mostrar los ejemplos reales más representativos para poder estimar las posibles consecuencias prácticas de las discordancias.

EJEMPLO REAL

Sí-Patológica y [Sí-Proliferativa o NO-Proliferativa]	Referente:	OFT	11	GO	Observó Comparador como patológ
Comparador	Observadas	Sí-Pat y [Sí-Prol]	Sí-Pat y [NO-Prol]	Total	
OFT	Sí-Pat y [Sí-Prol]	5	9	14	15
10	Sí-Pat y [NO-Prol]	3	81	84	97
GO	Total	8	90	98	112
Observó Ref como patológ		9	100	109	
Grado de acuerdo sólo entre las que ambos han observado simultáneamente como patológicas					
Conciden en	Coef Consistencia	Máx Coef Consist	V de Cramer	Índice kappa Cohen	
86 de 98	88%	0,38	0,71	53,7%	0,41
El ejemplo real que más se parece a la media de las medias de los oftalmólogos.					

EJEMPLO SIMULADO

Sí-Patológica y [Sí-Proliferativa o NO-Proliferativa]	Referente:	OFT	XXXX		
Comparador	Observadas	Sí-Pat y [Sí-Prol]	Sí-Pat y [NO-Prol]	Total	
OFT	Sí-Pat y [Sí-Prol]	43	9	52	
YYYY	Sí-Pat y [NO-Prol]	3	43	46	
	Total	46	52	98	
Grado de acuerdo sólo entre las que ambos han observado simultáneamente como patológicas					
Conciden en	Coef Consistencia	Máx Coef Consist	V de Cramer	Índice kappa Cohen	
86 de 98	88%	0,61	0,71	85,7%	0,76
Un ejemplo simulado en el que mantenemos el mismo porcentaje de coincidencias, pero igualando las prevalencias de ambas opciones, para ver cómo afecta al índice kappa de Cohen.					

La V de Cramer y el Coeficiente de Consistencia se basan en la distribución ji cuadrado para estimar la distancia entre dos observadores al clasificar en categorías nominales, la primera va de 0 (inconsistencia) a 1 (consistencia total), con lo que es fácil de comunicar, pero la segunda va de 0 hasta a un valor máximo de 0,71 si la tabla es 2x2, o 0,82 si ésta es 3x3, por lo que es más difícil de comunicar e interpretar.

Si bien los resultados que hemos obtenido mediante estos cuatro estadísticos en las distintas variables medidas nos dan una buena estimación numérica del grado de acuerdo en forma de concordancia y consistencia, este número no nos revela meridianamente la significación o relevancia clínica de los desacuerdos. Por esta razón, debajo de los resultados agregados nosotros hemos mostrado las tablas con las categorizaciones reales de dos observadores que más se parecen a las medias de las principales variables, porque éstas son las que mejor representan significación práctica (relevancia clínica) de los acuerdos y desacuerdos, incluso mejor que los estadísticos corregidos por pesos de gravedad asignados.

Refiriéndonos sólo a los valores numéricos, en nuestro estudio hemos observado unas medias en los cuatro estadísticos de concordancia moderada. Las comparaciones con otros estudios previos son arriesgadas, porque los diseños distan de cada uno de ellos en varias operaciones, por lo que establecen relaciones lo suficientemente distintas como para que los números tengan distintos significados teóricos y prácticos.

Al analizar las diferencias entre los grupos, no hemos encontrado diferencias estadísticamente significativas entre las medias del grupo de oftalmólogos y las medias del grupo de intervención (que había recibido el curso de formación). Y sí hemos encontrado diferencias entre el grupo de oftalmólogo y el grupo de control, es decir el que no había recibido el curso de formación, funcionando como grupo de espera para recibirla si los resultados son favorables al grupo que sí la había recibido.

Sin que sea nuestra intención trascender los conceptos del campo categorial de las ciencias con las que estamos operando (estadística, además de las ciencias fundamentales de la oftalmología), por el riesgo que ello tiene de incurrir inadvertidamente en una de las habituales formas de *filosofía espontánea* en la que incurren los científicos cuando salen de su particular campo categorial, decimos que con los resultados que hemos encontrado, el grupo de intervención, mediante una formación específica previa y protocolo en continuo, puede participar como primera línea del programa de cribado de retinopatía diabética mediante retinografías "Sí patológica/ No patológica". Simultáneamente, nos parece conveniente que los oftalmólogos que establezcan sospechas diagnósticas relativas a la retinopatía diabética mediante retinografías consensúen con acuerdo los atributos de las categorías diagnósticas dentro del grupo de "Sí patológicas", para disminuir en lo posible su variabilidad diagnóstica.

CONCLUSIONES

Tras participar en un curso de formación específica de cribado de retinopatía diabética, los médicos y enfermeras de Atención Primaria de nuestra muestra, pueden participar como primera línea del programa de cribado de retinopatía diabética mediante retinografías "Sí patológica/ No patológica", al haber obtenido una concordancia no inferior que los oftalmólogos respecto al mismo OFTALMÓLOGO REFERENTE PROMEDIO.

Se ha observado una asociación entre la mejora en la concordancia intra-observador y la participación en el mencionado curso de formación específica.

Parece conveniente que los oftalmólogos que establezcan sospechas diagnósticas relativas a la retinopatía diabética mediante retinografías consensúen con acuerdo los atributos de todas las categorías diagnósticas dentro del grupo de “Sí patológicas”, para disminuir en lo posible su variabilidad diagnóstica.

LIMITACIONES: No puede garantizarse que los resultados de este estudio sean extrapolables a otras instituciones sanitarias que disten en sus características del “Programa de Detección Precoz de Retinopatía Diabética en Atención Primaria (del SES)”.

RECURSOS DE LOS QUE SE HA DISPUESTO: Curso de formación de cribado de retinopatía diabética mediante retinografía (construido con material del SES por Ana Sánchez de la Morena y Verónica Martín Galán, e impartido por Ana Sánchez de la Morena), hoja de cálculo Excel® (Excel for Windows, Illinois, USA), calculadoras de la web evalmed.es.

ASPECTOS ÉTICOS:

1º Conflictos de intereses: Los tres autores declaran no tener conflictos de intereses económicos y no económicos, según los cinco apartados de los que consta el Formulario del Comité Internacional de Editores de Revistas Médicas para la declaración de potenciales conflictos de intereses. Las opiniones expresadas en este trabajo son responsabilidad de los autores, por lo que no reflejan necesariamente el punto de vista de los organismos en los que trabajan.

2º Financiación: Ninguna financiación a los autores y participantes ajena a sus salarios del SES.

3º Consentimiento informado: Dado que no se trata de un ensayo clínico, sino de un estudio epidemiológico transversal con los registros de pacientes anonimizados, nuestra investigación observacional de ámbito epidemiológico, está exenta de la obligatoriedad de consentimiento informado. La autora principal, VMG, garantizará la anonimización del banco retinográfico del SES mediante la asignación de un código numérico no asociable a la identificación del paciente, eliminando seguidamente los datos personales de carácter personal, para obtener el registro anonimizado sobre el que se podrá realizar toda la investigación.

AGRADECIMIENTOS:

A Ana Sánchez de la Morena por la construcción y por la impartición del *Curso de formación de cribado de retinopatía diabética mediante retinografía*.

A Raúl Garzón Álvarez por el soporte técnico informático.

A todos los miembros del **Grupo de Trabajo Retinografía Diabética del SES (ReDiSES)** por sus contribuciones para haber podido llevar a cabo este estudio, que está formado en orden alfabético por:

OFTALMÓLOGOS: Luciano Bermudez Uriá, Ana Fernández Ragaz, Eva García Suárez, María Jerez Fidalgo, Alberto Vicente López Masegosa, Manuel Pacheco Guerrero, Alicia Plaza Díaz, Consuelo Rubio Sánchez, Juan Ruiz González, Ana Sánchez de la Morena, Manuel Valencia Terrón. **MÉDICOS Y ENFERMERAS DE ATENCIÓN PRIMARIA:** Cristina Alcarranza Saucedo, Vicente Alonso González, Alexis Arjona Bravo, Alberto Barneto Clavijo, Inés Blanco Regodón, María del Carmen Calderón Gil, Ramón Candel Motellón, Gabriela Candilejo Fortes, Francisco Carramiñana Barrera, Adriana Carrasco Fernández, María Yolanda Castaño Blanco, Victoriano Chavero Carrasco, Marina Corchuelo Merino, María Carmen Costa Fernández, Beatriz Cruz Bernet, Concha Cruz Cánovas, Sara de Augusto Gil, Marta de Miguel Enrique, Iván Domínguez Fidalgo, María Luisa Franco Chamorro, María Josefa Gajardo Barrena, María Victoria Galindo Cantillo, Andrés José Gaspar Hernández, Carmen María Gil Martínez, Rebeca Gil García, Ángel F González Manero, Elena González Moreno, María Eugenia González Martín, María Isabel Guerra Puerta, Yolanda Guerrero Viñuela, Manuela Gutiérrez Fernández, Pedro Gutiérrez Moraño, Dimas Igual Fraile, Beatriz Jiménez Muñoz, Carmen Jiménez Fernández, Celia Jiménez Jiménez, Ángel Lara Barea, Paula Lloria Cascales, Manuel Jesús López Nieto, Noelia López Moreno, Noelia López Velázquez, Silvia López Galván, Elena Lorenzo Conde, Lidia Márquez Gómez, Adriana Martín Morcuende, Raquel Martín Jiménez, María Dolores Martín Gil, Cristina Martínez Martínez, Teresa Megía Rivera, Alejandro Mora Ortega, Inés Muñoz Galán, Verónica Murillo de Llanos, José Miguel Parejo Muñoz, Javier Parra Sánchez, María del Pilar Pascual Montero, Leticia Quintanilla Sellers, Naiara Reverter Velasco, Fernando Rubén Romero Angulo, María Romero Muñoz, María Luz Rubio Sánchez, Luis Sánchez Solís, María José Sánchez Blázquez, María del Pilar Sánchez

Marcos, Irene Santos Domínguez, Laura Santos Lucio, Antonia Soriano Lara, Félix Suárez González, Antonio Zapata Buide.

REFERENCIAS

1. Andonegui J, Berástegui L, Serrano L, et al. Concordancia en el estudio de retinografías en diabéticos: oftalmólogos vs médicos de familia. Arch Soc Esp Oftalmol 2008 Sep;83(9):527-531.
2. Vargas-Sánchez C, Maldonado-Valenzuela JJ, Pérez-Durillo FT, et al. Cribado de retinopatía diabética mediante retinografía midriática en atención primaria. Salud Publica Mex. 2011 May-Jun;53(3):212-219.
3. Silva PS, Cavallerano JD, Tolson AM, et al. Real-Time Ultrawide Field Image Evaluation of Retinopathy in a Diabetes Telemedicine Program. Diabetes Care. 2015 Jun 1. pii: dc150161.
4. Alonso Porcel C, Martínez Ibán M, Arboleya Álvarez L, Suárez Gil P, Sánchez Rodríguez LM. Cribado de retinopatía diabética en atención primaria. Concordancia diagnóstica entre médicos de familia y oftalmólogos. Semergen. 2016 sEP;42(6):357-362.
5. Molina-Fernández E, Valero-Moll MS, Pedregal-González M, Díaz-Rodríguez E, Sánchez-Ramos JL, Soriano-Villegas JM. Variabilidad interobservador en el diagnóstico y clasificación de la retinopatía diabética mediante biomicroscopía. Arch Soc Esp Oftalmol. 2008 Jan;83(1):23-28.
6. Sender Palacios MJ, Monserrat Bagur S, Badia Llach X, et al. Cámara de retina no midriática: estudio de coste-efectividad en la detección temprana de la retinopatía diabética. Med Clin (Barc). 2003 Oct 11;121(12):446-52.
7. Sender Palacios MJ, Vernet Vernet M, Maseras Bové M, et al. Oftalmopatía en la diabetes mellitus: detección desde la Atención Primaria de salud. Aten Primaria. 2011 Jan;43(1):41-8.
8. Rodríguez Villa S, Suárez Muñiz MT, De Dios Del Valle R, Alonso Álvarez C. Retinopathy diabetic screening by non-mydratic retinography: concordance between primary care physicians, nurses and ophthalmologists. Enferm Clin. 2018 Jan - Feb;28(1):44-48.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159-74.
10. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull. 1968 Oct;70(4):213-20.
11. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993 May;46(5):423-9.