

ARTÍCULO ORIGINAL

Confiabilidad y reproducibilidad de la escala de fototipos de Fitzpatrick antes y después de un ejercicio de estandarización clínica

Guillermo Sánchez, John Nova

Oficina de Docencia e Investigación, Centro Dermatológico Federico Lleras Acosta, Bogotá D.C., Colombia.

Introducción. La escala de fototipos de Fitzpatrick permite conocer la sensibilidad de la piel frente a la luz ultravioleta; una estimación confiable del fototipo permite establecer el verdadero riesgo de cáncer de piel de acuerdo con esta característica.

Objetivo. Establecer si existen diferencias en la concordancia intraobservador e interobservador de dos dermatólogos que evalúan el fototipo utilizando la escala de Fitzpatrick, antes y después de un proceso de estandarización clínica.

Materiales y métodos. Se realizó un estudio analítico de concordancia intraobservador e interobservador. La escala de fototipos de Fitzpatrick se estandarizó mediante la metodología de grupos focales. Para conocer el acuerdo intra e interobservador se utilizó el estadístico kappa ponderado. El efecto de la estandarización se midió mediante un contraste de hipótesis de igualdad de coeficientes kappa utilizando el estadístico de Wald, a través de la metodología de mínimos cuadrados ponderados.

Resultados. Se incluyeron 155 pacientes mayores de 15 años evaluados en cuatro oportunidades, por dos observadores independientes. En la fase de pre-estandarización, el kappa ponderado interobservador basal fue de 0,31 y de 0,40 a las seis semanas. El kappa intraobservador A fue de 0,47 y el intraobservador B fue de 0,51. Después del proceso de estandarización, se obtuvo un kappa ponderado interobservador basal de 0,77, y de 0,82 a las seis semanas. Los coeficientes kappa intraobservador A y B fueron 0,78 y 0,82, respectivamente. Se establecieron diferencias estadísticamente significativas entre los coeficientes antes y después de la estandarización ($p=0,000$ en todas las comparaciones).

Conclusiones. La escala de fototipos de Fitzpatrick posterior a un ejercicio de estandarización arroja resultados confiables, reproducibles y estables en el tiempo.

Palabras clave: reproducibilidad de resultados, dermatología, piel, pigmentación de la piel, neoplasias cutáneas, radiación solar/efectos adversos.

Reliability and reproducibility of the Fitzpatrick phototype scale for skin sensitivity to ultraviolet light

Introduction. The Fitzpatrick phototype scale has been used to determine skin sensitivity to ultraviolet light. The reliability of this scale in estimating sensitivity permits risk evaluation of skin cancer based on phototype.

Objective. Reliability and changes in intra and inter-observer concordance was determined for the Fitzpatrick phototype scale after the assessment methods for establishing the phototype were standardized.

Materials and methods. An analytical study of intra and inter-observer concordance was performed. The Fitzpatrick phototype scale was standardized using focus group methodology. To determine intra and inter-observer agreement, the weighted kappa statistical method was applied. The standardization effect was measured using the equal kappa contrast hypothesis and Wald test for dependent measurements. The phototype scale was applied to 155 patients over 15 years of age who were assessed four times by two independent observers. The sample was drawn from patients of the Centro Dermatológico Federico Lleras Acosta.

Results. During the pre-standardization phase, the baseline and six-week inter-observer weighted kappa were 0.31 and 0.40, respectively. The intra-observer kappa values for observers

A and B were 0.47 and 0.51, respectively. After the standardization process, the baseline and six-week inter-observer weighted kappa values were 0.77, and 0.82, respectively. Intra-observer kappa coefficients for observers A and B were 0.78 and 0.82. Statistically significant differences were found between coefficients before and after standardization ($p < 0.001$) in all comparisons.

Conclusion. Following a standardization exercise, the Fitzpatrick phototype scale yielded reliable, reproducible and consistent results.

Key words: reproducibility of results, dermatology, skin, skin pigmentation, skin neoplasm, solar radiation/adverse effects.

El concepto de reactividad solar fue introducido en 1975 por la necesidad específica de clasificar a las personas de piel blanca con el fin de seleccionar de forma correcta la dosis adecuada de luz ultravioleta A que debía ser aplicada para el tratamiento de la psoriasis. Fitzpatrick (1) planteó una clasificación basada en una entrevista personal explorando la historia de quemadura solar y la capacidad de broncearse, buscando una aproximación estimada de la tolerancia cutánea a la luz ultravioleta.

Algunos estudios que han evaluado factores de riesgo en cáncer de piel sugieren que la sensibilidad de la piel frente a la radiación ultravioleta podría ser un factor en la susceptibilidad del individuo a desarrollar la enfermedad (2,3). Se ha encontrado que los sujetos con mayor riesgo son los que poseen fototipo I y II (2,4). En la práctica clínica, el mayor desacuerdo entre los observadores se encuentra en el fototipo II y III, en los cuales paradójicamente se ha establecido el punto de corte para determinar riesgo de cáncer de piel Vs. fototipo.

Aunque la escala de Fitzpatrick es una herramienta de uso diario, económica, de amplia difusión y utilidad en dermatología, está basada en una evaluación subjetiva de la historia de quemadura y bronceado por efecto de la luz solar (5), situación que puede ocasionar variaciones en la confiabilidad de la evaluación. Es importante conocer la confiabilidad de esta escala, la

estabilidad del acuerdo en la observación hecha por el mismo individuo y el comportamiento en la observación hecha por dos individuos diferentes. El objetivo de este estudio fue determinar la confiabilidad y reproducibilidad de la escala de Fitzpatrick y cómo se modifica la concordancia intraobservador e interobservador cuando se estandariza y unifica la metodología de evaluación del fototipo.

Materiales y métodos

Se realizó un estudio analítico, para evaluar la concordancia intraobservador e interobservador y la confiabilidad test-retest de una escala, antes y después de un ejercicio de estandarización clínica.

Se incluyeron individuos mayores de 15 años que consultaron al Centro Nacional de Dermatología de Bogotá, Colombia; se excluyeron los pacientes con alteraciones asociadas a hipersensibilidad lumínica, fotodermatopatía o enfermedad de Hansen, que podrían modificar la sensibilidad cutánea a la luz, y aquéllos con alteraciones cognitivas que impidieran una adecuada colaboración.

El tamaño de la muestra fue calculado con una confiabilidad del 95%, probabilidad estimada de desacuerdo interobservador de 0,4, amplitud máxima del intervalo de confianza de 0,20, probabilidad de desacuerdo intraobservador de 0,10, para una muestra calculada de 153 individuos (6).

Procedimiento

Las observaciones fueron realizadas a cada paciente por dos dermatólogos de manera independiente y en cuatro momentos diferentes. Dos mediciones antes de la estandarización y dos después de la estandarización con intervalos de seis semanas, respectivamente. El intervalo fue

Correspondencia:

Guillermo Sánchez, Oficina de Docencia e Investigación, Centro Dermatológico Federico Lleras Acosta, Avenida 1ª N° 13-A-61, Bogotá, D.C., Colombia.
Teléfonos: 242 8160 y 242 8130, extensiones 115-145
epidemiologia@dermatologia.gov.co

Recibido: 24/04/08; aceptado:03/09/08

definido con base en las características del instrumento y en lo recomendado por la literatura para este tipo de mediciones (7).

Estandarización

Se utilizó la metodología de grupos focales (8-10) compuestos por 8 dermatólogos, en los cuales se revisó la escala de Fitzpatrick, se evaluaron las dificultades en su uso y la forma propuesta para su aplicación estandarizada en el centro en el que se efectuó el estudio. A partir de la escala se obtuvo un instrumento estandarizado para aplicar la escala de Fitzpatrick bajo parámetros específicos, de tal forma que los pacientes fueran interrogados de forma unificada.

Análisis

La evaluación del acuerdo intraobservador e interobservador se llevó a cabo mediante el estadístico kappa ponderado con sus respectivos intervalos de confianza (11,12). Se definió un valor α de 0,05, para hipótesis a dos colas. La interpretación de los resultados se realizó desde el punto de vista clínico con base en la escala propuesta por Landis y Koch (13), calificando como un acuerdo bueno si el resultado del estadístico kappa se encontraba entre 0,61 a 0,8.

Para conocer el efecto de la estandarización en el acuerdo obtenido, se planteó un contraste de hipótesis de igualdad de coeficientes kappa con medidas dependientes, mediante el estadístico de Wald, mediante la metodología de mínimos cuadrados ponderados (14). Se efectuaron las siguientes comparaciones entre coeficientes kappa:

- kappa basal y a las seis semanas

- kappa basal después de estandarización y seis semanas después de estandarización
- Kappa antes de estandarización y kappa después de estandarización.

En todos los casos se manejó una hipótesis nula de igualdad de coeficientes kappa. La comparación de coeficientes se efectuó mediante el módulo de datos categóricos PROC CATMOD (14) del paquete estadístico SAS®.

Consideraciones éticas

El presente estudio cumplió con la normatividad nacional e internacional para investigación en humanos. Cada uno de los pacientes incluidos aceptó participar mediante consentimiento informado; el proyecto fue aprobado y supervisado por el Comité de Ética independiente del Centro Dermatológico Federico Lleras Acosta E.S.E.

Resultados

Se incluyeron 167 pacientes, de los cuales, 155 (92%) completaron las cuatro evaluaciones, 102 mujeres (66%) y 53 hombres (34%). La edad osciló entre los 15 y los 67 años, con promedio y mediana de 32 años, desviación estándar de 14 años, con 50% de la población entre los 15 y los 55 años.

Fase antes de la estandarización

La medición basal de los dos observadores fue diferente en todos los fototipos, con predominio del II al V. En la evaluación realizada a las seis semanas, se documentaron diferencias en los fototipos del I al V, con predominio del III al V (cuadro 1).

Cuadro 1. Distribución por fototipos. Fase antes de la estandarización.

Fototipo	Basal		Seis semanas	
	Observador A	Observador B	Observador A	Observador B
	%			
I	1,29	1,94	0	2,58
II	31,61	14,19	23,87	24,52
III	53,95	40	58,06	37,42
IV	11,61	27,74	15,48	26,45
V	1,29	14,84	1,29	7,74
VI	0,65	1,29	1,29	1,29

Evaluación de concordancia

Los coeficientes kappa ponderados de la fase inicial tanto intraobservador como interobservador, oscilaron entre 0,31 y 0,51, considerados según la clasificación propuesta por Landis y Koch entre acuerdo débil y moderado (cuadro 2).

Grupos focales y proceso de estandarización

Se tomó como base la escala de fototipos propuesta por Fitzpatrick, se definieron las preguntas, se aclararon los términos y se construyó un instrumento unificado para obtener la evaluación del fototipo. Para un ajuste final, se aplicó a un grupo de 30 pacientes y posteriormente fue implementado (Anexo 1. Escala estandarizada de Fitzpatrick).

Fase después de la estandarización

La evaluación basal después de la estandarización mostró diferencias en los fototipos III y V. En la

evaluación realizada a las seis semanas después de la estandarización, se documentaron diferencias inferiores a dos puntos porcentuales entre los observadores A y B en los fototipos II al V, mientras que en los fototipos I y VI la correlación absoluta fue del 100% (cuadro 3).

Evaluación de concordancia

Los coeficientes kappa ponderados después de la estandarización tanto interobservador como intraobservador oscilaron entre 0,77 y 0,82, clasificados según Landis y Koch como acuerdo bueno (cuadro 4).

Concordancia antes y después de la estandarización

Para establecer el efecto de la estandarización sobre el acuerdo intraobservador e interobservador en la evaluación del fototipo a través de la escala de Fitzpatrick, se contrastaron los coeficientes

Cuadro 2. Coeficiente kappa ponderado. Fase antes de la estandarización.

Evaluación	Kappa ponderado	Error estándar	Intervalo de confianza	
Interobservador basal	0,31	0,04	0,26	0,33
Interobservador 6 semanas	0,40	0,05	0,34	0,46
Intraobservador A	0,47	0,06	0,45	0,50
Intraobservador B	0,51	0,04	0,44	0,55

Cuadro 3. Distribución por fototipos. Fase después de la estandarización.

Fototipo	Basal		Seis semanas	
	Observador A	Observador B	Observador A	Observador B
	%			
I	5,81	5,81	5,81	5,81
II	23,87	23,23	23,87	25,16
III	50,32	43,87	43,23	41,94
IV	14,84	13,55	17,42	16,77
V	3,87	12,26	8,39	9,03
VI	1,29	1,29	1,29	1,29

Cuadro 4. Coeficiente kappa ponderado. Fase después de la estandarización.

Evaluación	Kappa ponderado	Error estándar	Intervalo de confianza	
Interobservador basal	0,77	0,03	0,69	0,83
Interobservador seis semanas	0,82	0,03	0,75	0,88
Intraobservador A	0,78	0,03	0,70	0,85
Intraobservador B	0,82	0,03	0,76	0,88

kappa ponderados intraobservador antes y después de la estandarización, así como los coeficientes kappa ponderados interobservador (basal y seis semanas) antes y después de la estandarización (cuadro 5).

Se establecieron diferencias significativas entre los coeficientes kappa ponderados antes y después del proceso de estandarización en todas las comparaciones realizadas, pasando en todos los casos de un acuerdo débil o moderado a un acuerdo bueno, según la clasificación de Landis y Koch.

Discusión

La escala de fototipos propuesta en 1975 por Fitzpatrick (1) con el objetivo fundamental de seleccionar de forma adecuada la dosis inicial de radiación ultravioleta A para el manejo de los pacientes con psoriasis, se ha empleado en algunos estudios que han evaluado el papel del fototipo como un factor de riesgo para desarrollar cáncer de piel (2,4). Sin embargo, algunos autores han cuestionado la utilidad y confiabilidad de la escala (5,15,16).

Luego de una revisión ampliada de la literatura buscando estudios que hubieran evaluado validez y confiabilidad de la escala, se detectaron y revisaron 126 artículos relacionados en los que se evidenció que, a pesar de su amplia aplicación desde 1975, no había sido sometida a un proceso que permitiera conocer la confiabilidad y reproducibilidad de los resultados.

Nosotros evaluamos la confiabilidad de la escala de Fitzpatrick entre dos observadores independientes con intervalos de seis semanas presentando el acuerdo más allá del azar intraobservador e interobservador. El acuerdo interobservador basal y a las seis semanas, según la clasificación de Landis y Koch, mostró

resultados de acuerdo débil (kappa de 0,3 y 0,4), haciendo evidente una baja concordancia en las evaluaciones de dos dermatólogos que evalúan el fototipo basados en la escala de Fitzpatrick.

Más llamativo aún es el hallazgo en la evaluación intraobservador, en la cual se documentaron acuerdos más allá del azar moderados (observador A: kappa=0,47; observador B: kappa=0,52), lo cual sugiere que el evaluador tiende a modificar la forma de elaborar su pregunta frente a la fotosensibilidad del paciente originando estimaciones diferentes del fototipo en el mismo paciente.

Pensamos que, aunque la escala de Fitzpatrick está claramente definida en cada una de sus categorías, es susceptible de modificación por el clínico, produciendo bajas estimaciones de confiabilidad. Bajo este supuesto, desarrollamos un ejercicio de estandarización partiendo de la metodología de grupos focales (8-10), con el equipo de clínicos de nuestra institución, en el que se construyó un instrumento estandarizado basado en la escala de Fitzpatrick.

Los resultados obtenidos luego de este proceso fueron dramáticamente superiores y mostraron un acuerdo basal interobservador de 0,76 y un acuerdo a las seis semanas de la medición basal de 0,82. De igual forma, evaluamos la confiabilidad intraobservador y encontramos acuerdos más allá del azar de 0,77 y de 0,82 en los observadores A y B, respectivamente. Estos resultados sugieren que la escala de Fitzpatrick, aplicada de forma estandarizada como un instrumento definido y utilizado de forma sistemática al momento de evaluar la fotosensibilidad de los pacientes, mejora sustancialmente los acuerdos intraobservador e interobservador, con resultados que son estables en el tiempo, produciendo estimaciones confiables y reproducibles de la medición de la foto-sensibilidad de nuestros pacientes.

Cuadro 5. Comparación de coeficientes kappa ponderados antes y después de la estandarización.

Evaluación	Kappa previo	Kappa posterior	Wald	Valor p
Interobservador basal	0,31	0,77	60,2	0,000
Interobservador 6 semanas	0,40	0,82	40,7	0,000
Intraobservador A	0,47	0,78	18,5	0,000
Intraobservador B	0,51	0,82	27,3	0,000

El coeficiente de concordancia kappa es una herramienta útil que permite conocer el grado de acuerdo intraobservador e interobservador (11,12). Al momento de identificar niveles bajos de acuerdo, es posible efectuar procesos de ajuste y mejoramiento de la calidad que den lugar a estimaciones precisas y confiables. Una vez realizados estos cambios, se pueden obtener nuevas mediciones de confiabilidad con el propósito de conocer el efecto que sobre el grado de acuerdo produjo la intervención.

En este sentido, algunos autores se han limitado a hacer comparaciones descriptivas entre coeficientes kappa (15,16), que no permiten incluir en el análisis el tema de inferencia, comprobación y rechazo de hipótesis estadísticas que refuercen las hipótesis clínicas planteadas. El presente estudio, basado en los desarrollos de Barnhart y Williamson (14), aplicó pruebas de comprobación de hipótesis que permitieron establecer el efecto positivo de la estandarización sobre el grado de acuerdo intraobservador e interobservador en la evaluación del fototipo aplicando la escala de Fitzpatrick.

De acuerdo con estos resultados, la escala de Fitzpatrick utilizada bajo una metodología estandarizada de aplicación, recobra importancia y vigencia, ya que además de ser un método de bajo costo, amplia difusión y aplicabilidad en dermatología, produce evaluaciones del fototipo que muestran valores altos de acuerdo más allá del azar, garantizando confiabilidad y reproducibilidad de las observaciones.

En relación con lo anterior, este instrumento de medición de fotosensibilidad basado en la escala de Fitzpatrick puede ser utilizado por los dermatólogos y clínicos en general, que deseen conocer el grado de sensibilidad de la piel del paciente frente a la radiación ultravioleta; de igual forma, es de gran utilidad en los estudios que buscan establecer el fototipo como un factor de riesgo para diferentes afecciones cutáneas, ya que garantiza la confiabilidad de las mediciones.

La evaluación del fototipo mediante la escala de Fitzpatrick no se basa en el color de la piel sino en la sensibilidad cutánea a la luz y su forma de reaccionar frente a ella. Sin embargo, un paciente

que luzca una piel morena o demasiado blanca puede sesgar la evaluación del médico. El método estandarizado propuesto dio origen a un instrumento estandarizado a partir de la escala de Fitzpatrick, que luego podría ser aplicado de forma autoadministrada, obteniendo estimaciones libres del sesgo del observador. Se deben desarrollar estudios que evalúen la confiabilidad de la medición autoadministrada comparada con la medición administrada por el dermatólogo, ya que puede ser una manera confiable y eficiente de conocer el fototipo, con ahorro de tiempo del clínico.

Nuestros resultados apuntan a que la intervención realizada (estandarización en la aplicación de la escala de fototipos de Fitzpatrick) mejora sustancialmente el proceso de evaluación del fototipo y permite obtener de la escala mediciones confiables, reproducibles y estables en el tiempo.

Agradecimientos

Agradecemos el apoyo técnico científico aportado por el Instituto de Investigaciones clínicas de la Facultad de Medicina de la Universidad Nacional de Colombia en cabeza de Ricardo Sánchez y Javier Eslava. De igual forma agradecemos el apoyo recibido por Luisa Porras de Quintana, dermatóloga.

Conflicto de intereses

Los autores del presente estudio no reportaron conflicto de intereses.

Financiación

Este proyecto fue patrocinado por el Centro Dermatológico Federico Lleras Acosta E.S.E., Bogotá, Colombia.

Referencias

1. **Fitzpatrick TB.** Soleil et peau. *J Med Esthet.* 1975;2: 33-4.
2. **Walter U, Kron M, Sander S, Sebastian G, Sander R, Peter RU, et al.** Risk and protective factors for sporadic basal cell carcinoma: results of a two centre case control study in southern Germany. Clinical actinic elastosis may be a protective factor. *Br J Dermatol.* 2004;151:170-8.
3. **Lear JT, Tan BB, Smith AG, Bowers W, Jones PW, Heagerty AH, et al.** Risk factors for basal cell carci-

- noma in the UK: case control study in 806 patients. *J R Soc Med.* 1997;90:371-4.
4. **Loria D, Matos E.** Risk factors for cutaneous melanoma: a case-control study in Argentina. *Int J Dermatol.* 2001;40:108-14.
 5. **Youn JI, Oh JK, Kim BK, Suh DH, Chung JH, Oh SJ, et al.** Relationship between skin phototype and MED in Korean, brown skin. *Photodermatol Photoimmunol Photomed.* 1997;13:208-11.
 6. **Machin D, Campbell M, Fayers P, Pinol A.** Sample size tables for clinical studies. Second edition. Oxford: Blackwell Science; 1997.
 7. **Streiner DL, Norman GR.** Health measurement scales. A practical guide to their development and use. Second edition. Oxford: Oxford University Press; 2000. p. 104-27.
 8. **Strauss A, Corbin J.** Basics of qualitative research. Techniques and procedures for developing grounded theory. London: Sage; 1998.
 9. **Walker D, Myrick F.** Grounded theory: An exploration of process and procedure. *Qual Health Res.* 2006;16:547-59.
 10. **Shaw VN.** Research with participants in problem experience: challenge and strategies. *Qual Health Res.* 2005;15:841-54.
 11. **Sim J, Wright CC.** The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther.* 2005;85:257-68.
 12. **Kramer MS, Feinstein AR.** Clinical biostatistics. LIV. The biostatistics of concordance. *Clin Pharmacol Ther.* 1981;29:111-23.
 13. **Landis JR, Koch GG.** The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-74
 14. **Barnhart H, Williamson J.** Weighted least squares approach for comparing correlated kappa. *Biometrics.* 2002;58:1012-9.
 15. **Venkataram MN, Haitham AA.** Correlating skin phototype and minimum erythema dose in Arab skin. *Int J Dermatol.* 2003;42:191-2.
 16. **Stern RS, Momtaz K.** Skin typing for assessment of skin cancer risk and acute response to UV-B and oral methoxalen photochemotherapy. *Arch Dermatol.* 1984;120:869-73.

Anexo 1. Instrumento estandarizado para la evaluación del fototipo basado en la escala de Fitzpatrick

Pregunta:

¿Qué sucede con la piel de su espalda o sus hombros, al día siguiente después de haberse expuesto al sol más de una hora, sin usar protector solar?

Recomendación:

Tenga en cuenta que ponerse rojo equivale a quemarse o arderser y que broncearse equivale a que la piel se ponga más oscura o morena.

Señale con una "X" la opción que más se ajuste a su piel.

1. Siempre me pongo rojo (siempre me quemó).
Nunca se me oscurece la piel (nunca me bronceo).
2. Casi siempre me pongo rojo.
Casi nunca se me oscurece la piel.
3. Algunas veces me pongo rojo.
Casi siempre se me oscurece la piel.
4. Casi nunca me pongo rojo.
Siempre se me oscurece la piel.
5. Nunca me pongo rojo.
Siempre se me oscurece la piel (raza morena).
6. Nunca me pongo rojo.
Siempre se oscurece la piel intensamente (raza negra).