



Investigación en
Educación Médica

<http://riem.facmed.unam.mx>



ARTÍCULO ORIGINAL

Proceso de validación de un instrumento para medir la lectura crítica de informes de investigación médica



Héctor Cobos Aguilar^{a,*}, Patricia Pérez Cortés^b, Héctor de la Garza Quintanilla^b
y Carlos Enrique Ochoa Castro^c

^a Coordinación del internado de pregrado, Departamento de Ciencias Clínicas, División de Ciencias de la Salud, Universidad de Monterrey, San Pedro Garza García, Nuevo León, México

^b Departamento de Ciencias Clínicas, División de Ciencias de la Salud, Universidad de Monterrey, San Pedro Garza García, Nuevo León, México

^c Residencia de medicina interna, Hospital Christus Muguerza Alta Especialidad, Monterrey, Nuevo León, México

Recibido el 22 de octubre de 2014; aceptado el 5 de enero de 2015

Disponible en Internet el 6 de agosto de 2015

PALABRAS CLAVE

Instrumento;
Validez;
Consistencia;
Lectura crítica;
Internado

Resumen

Introducción: La lectura crítica de investigación es una herramienta de aprendizaje que debe ser manejada por profesores y alumnos. Las estrategias educativas deben enfocarse al aprendizaje activo del alumno. Instrumentos válidos y confiables deben elaborarse para medir los avances en este campo.

Objetivo: Construir, validar y determinar la consistencia de un instrumento para evaluar la lectura crítica.

Método: Instrumento: se revisaron artículos publicados sobre construcción de instrumentos, encuestas, casos y controles, pruebas diagnósticas, ensayo clínico aleatorizado, cohortes, seguimiento y metaanálisis. Se seleccionaron finalmente 8 artículos con los respectivos diseños. Se realizaron resúmenes y se derivaron enunciados y reactivos que exploraban la validez, la consistencia y los diseños, los estadísticos, resultados y discusión. Los reactivos se sustentaban en los indicadores de la lectura crítica: interpretación, enjuiciamiento y propuestas. Se realizó validación por técnica de Delphi en 2 rondas con 6 doctores o másteres en ciencias o educación, con experiencia en estos instrumentos. Finalmente se equilibró con 96 reactivos, 16 por diseño y 32 por indicador, con 48 respuestas falsas y verdaderas. Se calificó en forma ciega y con un programa para evitar errores de captura. Se obtuvieron medianas para el análisis. Se aplicó prueba piloto con 2 grupos de internos de pregrado (neófito G₁ e intervenido G₂) y se determinó la consistencia por Kuder-Richardson. Se compararon extremos con U de Mann-Whitney. Se determinó también el nivel de azar.

* Autor para correspondencia: Calle 8, n.º 430, Col. Villa Azul. San Nicolás de los Garza, N.L. CP: 66420. Tel.: +01(81)63 76 85 04.
Correo electrónico: hector.cobos@udem.edu (H. Cobos Aguilar).

La revisión por pares es responsabilidad de la Universidad Nacional Autónoma de México.

Resultados: Se obtuvo una consistencia de 0.79, aunque una validez concurrente inadecuada (Spearman = 0.31). Hubo diferencias globales en las medianas grupales G₁ vs. G₂ (11 vs. 29) así como en respuestas por azar (70% vs. 8%) y una comparación de extremos significativa ($p < 0.0019$). **Conclusiones:** Es necesario construir y actualizar instrumentos de medición para medir el desarrollo de esta habilidad compleja trascendental en la crítica de la información médica publicada.

Derechos Reservados © 2015 Universidad Nacional Autónoma de México, Facultad de Medicina. Este es un artículo de acceso abierto distribuido bajo los términos de la Licencia Creative Commons CC BY-NC-ND 4.0.

KEYWORDS

Instrument;
Validity;
Consistency;
Critical reading;
Internship

Process of validation of an instrument to evaluate the critical reading in medical research papers

Abstract

Introduction: Critical appraisal is a learning tool that should be familiar to professors and students. Educational strategies should focus on the student's active learning. Valid and reliable tools should be constructed in order to measure progress in this field.

Objective: To construct, validate and determine the consistency of a tool for evaluating critical reading.

Method: Tool: a review was conducted on published articles on the construction of evaluation tools, surveys, cases and controls, diagnostic tests, randomised clinical trials, cohorts, follow-up and meta-analyses. Eight articles on each design were finally selected. They were summarised and statements and items were developed, by examining the document validity, consistency, design, statistics, results, and the discussion. The items were based on critical reading indicators: interpretation, judgment, and proposals. Validation was obtained using the Delphi technique in 2 rounds that included six Doctors or Masters in Science or Education, previously experienced in the use of these tools. It was balanced with 96 items, 16 per design and 32 per indicator, with 48 "true or false" answers. Grading was blind and using a software program to avoid capture errors. Medians were obtained for analysis. A pilot test was applied to two groups of student interns (neophytes G₁ and intervened G₂) and consistency was established with the Kuder-Richardson formula. Extremes were compared with the Mann-Whitney U test. The level of randomness was also determined.

Results: A consistency of 0.79 was obtained, although concurrent validity was inadequate (Spearman = 0.31). There were overall differences in the group median values, G₁ vs. G₂ (11 vs. 29) as well as in random answers (70% vs. 8%), and in comparison of extremes ($P < .0019$).

Conclusions: It is necessary to construct and update tools for measuring the development of this complex and transcendental ability required for the critical appraisal of published medical information.

All Rights Reserved © 2015 Universidad Nacional Autónoma de México, Facultad de Medicina. This is an open access item distributed under the Creative Commons CC License BY-NC-ND 4.0.

Introducción

La lectura crítica de investigación continúa siendo una herramienta de aprendizaje muy poderosa necesaria para evaluar la información médica publicada en el medio científico, pero no se encuentra destacada en los planes de estudio de las universidades o de las instituciones de salud que forman a su personal relegando la necesidad de formación en este campo¹. Se han publicado diferentes reportes para medir su desarrollo, encontrando, en forma persistente, resultados que se ubican en niveles de azar, a través de los años².

Sin embargo se observan resultados promisorios posteriores a estrategias educativas de participación que muestran los beneficios a su exposición en diferentes ámbitos, en los que la crítica es el eje de aprendizaje a través

de la lectura crítica en sus 3 componentes (interpretación, enjuiciamiento y propuestas), ampliamente referidos en trabajos previos^{3,4}. A pesar de la necesidad de desarrollar esta habilidad, que la sitúa en un papel curricular predominante, se soslaya su implementación.

Cuando la investigación se encuentra en los planes de estudio, en los cursos correspondientes se continúan utilizando estrategias de aprendizaje no sustentadas en la elaboración del conocimiento y en la crítica de la experiencia, como propone Viniegra⁵.

Por otro lado, los diseños metodológicos se fortalecen a través de los años y a los 4 diseños básicos que propuso originalmente la Universidad de Mac Master para iniciar la lectura crítica de diseños metodológicos es necesario ahora agregar la construcción de instrumentos, los estudios de seguimiento y los metaanálisis para transitar apenas por la superficie de

la medicina basada en evidencias, aspecto fundamental para un ejercicio clínico adecuado⁶.

El manejo apropiado de los diseños referidos antes es ya imprescindible en los estudiantes de medicina y en general para todo el personal de salud. Por ello es necesario abordarlos a través de la lectura crítica, que enfatiza aspectos metodológicos como la validez, la consistencia, la fortaleza de los diseños seleccionados, el muestreo, la determinación del tamaño muestral, el control de los sesgos, los estadísticos pertinentes, el análisis de los resultados presentados, la discusión y la pertinencia en el medio en el que se llevan a cabo, entre otros, y que le confieren o no validez a los reportes de investigación factual.

Evaluar el conocimiento así como el desarrollo de esta habilidad compleja en el personal de salud, de cualquier nivel educativo, nos obliga a construir instrumentos válidos y confiables para tal fin. Además estos instrumentos se deben actualizar a medida que los profesores avanzan en el camino interminable de la crítica, con nuevos diseños para otorgar a los alumnos una mirada completa de los mismos que le permitan evaluar los reportes o enfrentar sus propias preguntas de investigación^{7,8}.

Es necesario que los instrumentos elaborados tengan validez intrínseca (teórica, de constructo, de contenido) y extrínseca (predictiva y concurrente) así como una consistencia adecuada para poder utilizarlos y obtener resultados con los que se puedan tomar decisiones adecuadas en el aprendizaje de los alumnos^{9,10}. Diversos metaanálisis, preponderantemente en los cursos de medicina basada en evidencias, muestran un reporte muy bajo y deficiente de las características psicométricas de los mismos^{11,12}.

No existen muchos reportes en la literatura sobre el proceso de la construcción y validación de instrumentos de lectura crítica, a pesar del amplio uso de la medicina basada en evidencias, y con frecuencia se relacionan a preguntas que exploran memoria o indagaciones generales de habilidades, en cuestionarios cortos de hasta 5 reactivos¹³⁻¹⁵.

El objetivo del estudio es reportar el proceso de construcción, validación y consistencia de un instrumento para evaluar el desarrollo de la lectura crítica de informes de investigación con los diseños indispensables al inicio del internado de pregrado.

Método

Diseño: construcción, validación y determinación de la consistencia de un instrumento para evaluar el desarrollo de la lectura crítica de informes de investigación factual en pregrado en medicina.

Población: se aplicó un instrumento a médicos internos de pregrado en una institución de seguridad social en la que el curso de investigación es parte fundamental del plan de estudios. Un grupo iniciaba apenas el internado (novato), sin el curso de investigación (G_1 , $n=23$, promedio 84.4, DE: 3.3) y el otro, después de 6 meses de iniciado el internado, lo había ya cursado (intervenido, G_2 , $n=24$, promedio 83.6, DE 5.4). Este último grupo no tuvo contacto previo con el instrumento en estudio. Los alumnos de ambos grupos provenían de distintas universidades, 2 locales privadas y una privada del occidente del país. Los promedios reportados fueron de la carrera, en una escala de 0-100.

Proceso de validación del instrumento

El plan de construcción del instrumento fue asignar un artículo de investigación por cada tipo de diseño (8 en total), a partir del cual se generaron una serie de tallos, con 96 reactivos, 12 para cada artículo.

Selección de artículos

Se revisaron 50 artículos en una primera ronda y 22 en una segunda, publicados en revistas indizadas, metodológicamente consistentes, relacionados con problemas del entorno geográfico, a fin de indagar mejor las distintas secciones, con los siguientes diseños: construcción de instrumentos, encuestas, casos y controles, pruebas diagnósticas, ensayos clínicos aleatorizados, cohortes, seguimiento y metaanálisis. Finalmente se seleccionaron 8, uno por diseño, que representaban las características más relevantes de los mismos.

Elaboración del instrumento

De los 8 seleccionados se elaboró un resumen, no mayor de media cuartilla, con las características más importantes del diseño, selección, tamaño muestral, sesgos, estadísticos usados, resultados más importantes, la significación y discusión. A cada resumen se agregaron tallos relacionados con los aspectos que otorgaban rigor metodológico como: validez, consistencia, pertinencia del diseño, muestreo, tamaño muestral, estadísticos, resultados, significación y discusión. Los enunciados incluían reactivos que exploraban la validez o no de cada uno de los aspectos referidos. Los reactivos exploraban los componentes de la lectura crítica: interpretar (el diseño usado, los implícitos del estudio, el muestreo, etc.), enjuiciar (la pertinencia de un diseño o de un estadístico sobre otro, de una técnica o no de aleatorización, etc.), y proponer (con propuestas que superaban a las del autor en los aspectos previos).

Se elaboraron los resúmenes inicialmente con 141 reactivos y se enviaron a 6 validadores para una primera ronda. Estos fueron médicos especialistas (3 pediatras, 2 internistas y un médico familiar), 2 tenían doctorado en ciencias, 2 tesis de doctorado en educación y 2 con maestría en educación, con experiencia en la construcción de estos instrumentos, con publicaciones en diferentes revistas y eran profesores de investigación en maestrías, posgrado o pregrado.

Todos ellos recibieron un formato en el que se invitaba a revisar la pertinencia de los resúmenes elaborados, revisando aspectos de sintaxis, aspectos metodológicos distintivos de los diseños explorados, los tallos incluidos así como los componentes elaborados y sus posibles respuestas a cada uno de ellos. Se anexa un ejemplo de resumen, tallo y reactivos ([anexo 1](#)).

Se les invitó de manera independiente. La revisión fue anónima e independiente, en 2 rondas, por vía electrónica (técnica Delphi). Se les pidió que evaluaran la validez teórica (lectura crítica y componentes), la validez de constructo relacionada con la elaboración de los resúmenes, los enunciados y los reactivos y la validez de contenido (los 3 componentes y los 8 diseños). Asimismo se les solicitó

Tabla 1 Comparación de resultados, por componentes, entre un grupo inicial e intermedio en lectura crítica de informes de investigación

	Componente	Máximo correctas	Novatos (n = 23)	Intervenidos (n = 24)	p*
			(Medianas)		
1.	Interpretar	32	1	4	NS
2.	Juicio	32	5	12	0.001
3.	Propuestas	32	7	14	0.001
4.	Global	96	11	29	0.001

* U de Mann-Whitney.

contestar los enunciados como «Cierto» o «Falso» de acuerdo a su experiencia.

Se tomaron como adecuados aquellos en los que hubiera un acuerdo en las respuestas de 5/6 o 6/6. En la primera ronda hubo un acuerdo de 70 reactivos, que ascendieron a 100 en la segunda. Se realizaron las modificaciones sugeridas a resúmenes y reactivos.

Posteriormente se elaboró una hoja de respuestas para los alumnos con instrucciones precisas. Cada respuesta se contestaba como «Cierto», «Falso» o «No sé», ya que no se sustentan en el recuerdo de datos y se agrega la posibilidad de determinar mejor la respuesta por azar. Las respuestas correctas (cierta o falsa) sumaban un punto, las incorrectas restaban un punto y las «no sé» no restaban ni sumaban puntos.

Se diseñó un sistema computarizado de captura en programa Excel para minimizar errores y obtener los resultados (medianas) de manera automática por indicador, global y por diseño. La captura fue ciega.

El instrumento se aplicó en una muestra piloto de 32 alumnos internos de pregrado, en una sede diferente, sin realizarse observaciones al mismo. Se aplicó finalmente a G₁ y G₂ para determinar la consistencia y se compararon los extremos con la prueba de U de Mann-Whitney para buscar diferencias significativas. Los promedios se compararon con *t* de Student.

En todos los casos se solicitó autorización escrita a las autoridades educativas correspondientes para su aplicación y a los alumnos se les pidió su consentimiento verbal para su resolución. No hubo algún beneficio por participar o represalia por negarse a ello. La duración promedio de la resolución fue de 45 min.

Al mismo tiempo y con 3 semanas de diferencia se aplicó otro instrumento válido y consistente utilizado hasta ese momento para determinar la validez concurrente a través de la *S* de Spearman, aunque el instrumento era diferente¹⁶. Las respuestas de los alumnos se clasificaron en diversas categorías (azar, muy bajo, bajo, intermedio, alto, muy alto) una vez determinado el azar, de acuerdo a las recomendaciones de Pérez-Padilla y Viniegra¹⁷.

Resultados

En su versión final el instrumento quedó conformado y equilibrado con 96 reactivos, 36 para cada componente, 12 para cada diseño y 48 respuestas correctas y falsas.

Se determinó la consistencia mediante la fórmula 21 de Kuder-Richardson con un resultado de 0.79. No hubo diferencia en ambos grupos en promedios de la carrera (*t*: NS). Se obtuvo una *S* de Spearman de 0.37, (*p* = 0.12) cuando se compararon los resultados obtenidos por instrumento construido actualmente y el otro aplicado 3 semanas después.

Se compararon los grupos novato e intervenido (G₁ y G₂) en los 3 indicadores y resultados globales (G₁ Mdn: 11 y G₂ Mdn: 29), así como en los 8 diseños referidos antes. En todos los indicadores de la lectura crítica y por diseño de artículo, hubo diferencias estadísticamente significativas a favor del G₂, excepto en interpretación y los diseños de metaanálisis y seguimiento (tablas 1 y 2).

Se compararon a su vez las medianas globales obtenidas por los alumnos en los percentiles 25 y 75 en cada grupo y en los 2 grupos, con resultados significativos para el G₁ (Mdn P₂₅ 4 vs. P₇₅ 18, *p* < 0.002) y el G₂ (Mdn P₂₅ 22 vs. P₇₅ 37,

Tabla 2 Comparación de resultados, por diseño, entre un grupo inicial e intermedio en lectura crítica de informes de investigación

	Diseño	Máximo correctas	Novatos (n = 23)	Intervenidos (n = 24)	p*
			(Medianas)		
1.	Ensayo clínico	12	1	4	0.008
2.	Casos controles	12	0	2	0.006
3.	Pruebas diagnósticas	12	1	4	0.001
4.	Cohortes	12	2	4	0.006
5.	Metaanálisis	12	1	3	NS
6.	Seguimiento	12	1	2	NS
7.	Instrumentos	12	1	4	0.03
8.	Encuestas	12	3	6	0.004

* U de Mann-Whitney.

Tabla 3 Comparación de frecuencias entre los 2 grupos, en niveles de los resultados obtenidos por los estudiantes, descartando el azar

Nivel/respuestas		Novato (n = 23)	Intervenido (n = 24)
Muy alto	81-96		
Alto	65-80		
Intermedio	49-64		2 (0.08)
Bajo	33-48	1 (0.04)	7 (0.30)
Muy bajo	17-32	6 (0.26)	13 (0.54)
Azar	16 o más	16 (0.70)	2 (0.08)

$p < 0.002$) y al combinarse ambos grupos (Mdn P_{25} 4 vs. P_{75} 37, $p < 0.001$).

Cuando se comparan las respuestas, descartando el azar, en ambos grupos y en las categorías respectivas, se observó una frecuencia más alta en el azar en el G_1 (tabla 3).

Discusión

Se presenta el proceso de construcción, validación y consistencia de un instrumento para medir el desarrollo de la lectura crítica de informes de investigación factual.

La fortaleza del estudio se relaciona con la referencia del proceso en la elaboración de estos instrumentos, desde el acopio y búsqueda de los diseños que se desea incluir –considerados más adecuados– para que el médico enfrente la lectura de los mismos en artículos publicados; ya que en diversos reportes se refiere la falta de información, si no ausencia, de su validación y consistencia¹⁸.

Otros aspectos que le confieren fortaleza fue el consenso alto de los jueces en la primera ronda y el manejo de la técnica de Delphi por profesores e investigadores con experiencia en el campo y publicaciones de investigación factual y aún con experiencia en el aprendizaje de investigación a través de estrategias de participación y la lectura crítica como sustento.

La no diferencia entre los promedios refleja que la lectura crítica no se encuentra desarrollada antes de un curso dirigido a tal fin.

Cuando se utiliza el instrumento para comparar a 2 grupos, uno de los cuales iniciaba el internado de pregrado, con la mayoría de los alumnos que no habían llevado un curso de investigación, y otro que había sido ya intervenido con los mismos, encontramos diferencias significativas en los indicadores de juicio y propuestas así como en los valores globales y en casi todos los diseños excepto en metaanálisis y seguimiento. Lo anterior denota que el instrumento, además de los datos de consistencia referida, discrimina entre alumnos que ya han tenido experiencia, a través del curso de lectura crítica, contra aquellos que se enfrentan por primera vez a este tipo de instrumentos, que indagan más el aprendizaje de habilidades complejas como es la lectura crítica y sus componentes. Sin embargo hacen falta más estudios para fundamentar el hallazgo.

Además la no diferencia en los diseños referidos fortalecen la validez del instrumento pues esos fueron los de reciente introducción y, como hemos visto en estudios previos, cuando no se lleva a cabo un proceso continuo de

reflexión y crítica es difícil que los alumnos obtengan resultados adecuados.

Cuando se comparan estos grupos, descartando el azar, observamos también las cualidades del instrumento para discriminar alumnos con experiencia en lectura crítica.

La mayoría del grupo de novatos queda en el azar y algunos alumnos se encuentran en niveles más altos. Fortaleciendo las cualidades del instrumento, los alumnos que obtienen medianas más altas se encontraban expuestos durante la carrera a la lectura crítica. En estudios previos se ha mostrado que esta ventaja es útil cuando los alumnos avanzados se convierten en docentes¹⁹.

La validez concurrente fue baja pues fueron instrumentos no del todo iguales, aunque el proceso de validación intrínseca le confiere consistencia.

Limitaciones del estudio: debido a la construcción del instrumento, con opciones de C y F no se obtienen otras mediciones psicométricas comunes en los de opción de respuesta 1 de 5, como el análisis factorial pero la validez y consistencia se acercan más a la mirada teórica más cualitativa de la participación.

La elaboración de instrumentos válidos y consistentes debe ser una tarea de los profesores que se apartan de la visión tradicional de la educación y que realizan indagaciones sobre el aprendizaje de sus alumnos. En nuestro caso esta es la cuarta versión de un instrumento para evaluar el desarrollo de la lectura crítica, y en cada una la experiencia y el propio avance de los profesores en esta actividad, que ha sido continua en los últimos 20 años, les confiere aspectos que son relevantes^{20,21}. En esta versión se incluyeron los diseños de metaanálisis y de seguimiento que actualmente aparecen con mayor frecuencia en la literatura médica y que todo el personal de salud debe conocer y criticar al estar dotados de esta herramienta metodológica, pues la generación de artículos médicos es exponencial y se deben evaluar con rigor metodológico.

Aún más, con el ejercicio de la medicina basada en evidencias se analizan con mayor frecuencia estos diseños mediante estrategias educativas diversas, con artículos publicados, esenciales en el médico para el aprendizaje de la clínica y las decisiones médicas, a través de revisiones sistemáticas que muestran su efectividad²²⁻²⁵. También se han desarrollado instrumentos a fin de confirmar las habilidades docentes para el desarrollo de esta habilidad²⁶.

No se trata de elaborar instrumentos continuamente por una exigencia administrativa, sino de desafiar el avance de los profesores e incorporarlos en su ejercicio docente para dotar a los alumnos mediante estrategias participativas de esta poderosa herramienta de aprendizaje.

La investigación es una actividad primordial en medicina y el desarrollo de la lectura crítica debe ser un aspecto prioritario en nuestros estudiantes. Los instrumentos válidos y confiables son relevantes para constatar su avance.

Conclusiones

La elaboración de instrumentos válidos y consistentes debe ser una actividad permanente del quehacer docente. El que presentamos aquí exhibe estas cualidades, aunque con las limitaciones referidas antes, lo que servirá para tomar decisiones pertinentes relacionadas con el aprendizaje de los alumnos en esta compleja habilidad. Pueden observarse sutilezas (diseños e indicadores) que no son detectadas con otros instrumentos de medición sustentados en otras miradas educativas que favorecen más el recuerdo que la crítica.

Responsabilidades éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad de los datos. Los autores declaran que en este artículo no aparecen datos de pacientes.

Derecho a la privacidad y consentimiento informado. Los autores declaran que en este artículo no aparecen datos de pacientes.

Financiación

Ninguna.

Contribución de los autores

HCA: elaboración de proyecto, elaboración del instrumento, análisis, elaboración final del manuscrito.

PPC: elaboración de proyecto, validación del instrumento, análisis, elaboración final del manuscrito.

HDLGQ: validación del instrumento, análisis, elaboración final del manuscrito.

CEOC: elaboración de programa de captura, análisis de los datos y colaboración con la elaboración final del manuscrito.

Conflicto de intereses

Ninguno.

Presentaciones previas

Ninguna.

Agradecimientos

Ninguno.

Anexo 1. Ejemplo de resumen, tallo y reactivos

Papel de fumador activo y pasivo en el cáncer de pulmón

El objetivo del estudio es identificar la relación del cáncer de pulmón con el fumador pasivo y activo, entre otros factores. Se estudiaron 385 pacientes con diagnóstico de cáncer pulmonar, confirmado por patología, en hospitales de tercer nivel. Un grupo de 898 individuos sanos, de la comunidad, se estudió con una encuesta multietápica, aleatoria, con estratos por edad. A ambos grupos (relación 1:2) se les aplicó un cuestionario por 2 entrevistadores entrenados que desconocían el motivo del estudio. Contenía preguntas sobre exposición al tabaco, activa o pasiva, número de cigarrillos al día y edad de inicio (si fumó). Los datos fueron capturados 2 veces y analizados en SPSS. Se utilizó la relación de momios (RM), en cálculo bivariado, con un intervalo de confianza (IC) del 95%.

Cuadro 1 Resultados de la exposición

	Factor	RM	IC
Fuma:	Nunca	1	
	Sí	4	2.9-5.5
Exfumador		3.2	2.2-4.5
Número de cigarros	Fuma 2-4 cigarros/día	1.4	0.9-2.4
	Fuma 5-10 cigarros/día	4.4	2.9-6.6
	Fuma pasivo casa	1.8	1.3-2.6
Edad inicio/años	20 a 65	3.4	2.4-5
	15 a 19	4	2.7-5.8
	5 a 14	5.3	3.5-7.9

En relación con el diseño del estudio

1. Busca causalidad.

2. Es longitudinal.

Son procedimientos que le incrementan validez al estudio

3. El procedimiento para determinar el número de casos.

4. La relación casos: controles.

5. El uso de IC.

De acuerdo a los resultados

6. Fumar 2 a 4 cigarrillos al día indica la aparición incipiente de cáncer pulmonar.

7. El fumador pasivo tiene tanto riesgo de tener cáncer como aquel que fuma 2 a 4 cigarrillos al día.

8. En relación con los IC existe el mismo riesgo de tener cáncer independientemente de la edad de inicio al fumar.

Son propuestas al autor

9. Utilizar un muestreo por conveniencia para los controles.

10. Realizar un análisis multivariado.

Referencias

1. Sánchez MM. La medicina basada en evidencias y la escuela de medicina. *Gac Méd Méx.* 2004;140:314-6.

2. Carranza LS, Arce HRM, Leyva SRI, Leyva GFA. Lectura crítica de informes de investigación clínica en estudiantes de pregrado. *Rev Med Inst Mex Seguro Soc.* 2009;47:431–8.
3. Coomarasamy A, Taylor R, Khan KS. A systematic review of postgraduate teaching in evidence-based medicine and critical appraisal. *Med Teach.* 2003;25:77–81.
4. Carranza LS, Arce HRM, González GP. Evaluación de una estrategia educativa para mejorar la aptitud para la lectura crítica de textos de investigación clínica en residentes de ginecoobstetricia de segundo año. *Ginecol Obstet Mex.* 2007;75:661–6.
5. Viniegra VL. La experiencia reflexiva y la educación. *Rev Invest Clin.* 2008;60:133–56.
6. Calva MJJ, Ponce de León RS, Vargas VF. Cómo leer revistas médicas. *Rev Invest Clin.* 1988;40:67–105.
7. Mansfield L. The reading, writing, and arithmetic of the medical literature, part 3: Critical appraisal of primary research. *Ann Allergy Asthma Immunol.* 2006;96:7–16.
8. Mansfield L. The reading, writing, and arithmetic of the medical literature, part 2: Critical evaluation of statistical reporting. *Ann Allergy Asthma Immunol.* 2005;95:315–22.
9. García GAJ, Ramos OG, Díaz de León PMA, Olvera-Chávez A. Instrumentos de evaluación. *Rev Mex Anest.* 2007;30:158–64.
10. Brink Y, Louw QA. Clinical Instruments: Reliability and validity critical appraisal. *J Eval Clin Prac.* 2012;18:1126–32.
11. Shaneyfelt T, Baum K, Bell D, Feldstein D, Houston T, Kaatz S, et al. Instruments for Evaluating Education in Evidence-Based Practice. *JAMA.* 2006;296:1116–27.
12. Jordan JE, Osborne RO, Buchbinder R. Critical appraisal of health literacy indices revealed variable underlying constructs, narrow content and psychometric weaknesses. *J Clin Epidemiol.* 2011;64:366–79.
13. Feldstein DA, Maenner MJ, Srisurichan R, Roach MA, Vogelmann BS. Evidence-based medicine training during residency: A randomized controlled trial of efficacy. *BMC Med Educ.* 2010;10:59.
14. Liabsuetrakul T, Suntharasaj T, Tangtrakulwanich B, Uakritdathikarn T, Pornsawat P. Longitudinal analysis of integrating evidence-based medicine into a medical student curriculum. *Fam Med.* 2009;41:585–8.
15. Srinivasan M, Winer M, Breitfeld PP, Brahma F, Dickerson KL, Winer G. Early introduction of an evidence-based medicine course to preclinical medical students. *J Gen Intern Med.* 2002;17:58–65.
16. Cobos AH, Pérez CP, Vázquez GA, Cobos HCD, Tapia OJA. Construction and validation of an instrument to evaluate critical reading in research papers. *Proceedings of ICERI.* 2011: 5286–94.
17. Pérez-Padilla JR, Viniegra VL. Método para calcular la distribución de las calificaciones esperadas por azar en un examen del tipo falso, verdadero, no sé. *Rev Invest Clin.* 1989;41: 375–9.
18. Crowe M, Sheppard L. A review of critical appraisal tools show they lack rigor: Alternative tool structure is proposed. *J Clin Epidemiol.* 2011;64:79–89.
19. Cobos AH, Pérez CP, Bracho VLA, Garza GMA, Dávila RG, López JDO, et al. Habilidades docentes en alumnos tutores en lectura crítica de investigación médica durante el internado de pregrado. *Inv Ed Med.* 2014;3:92–9.
20. Cobos AH, Insfrán SMD, Pérez CP, Elizaldi LNE, Hernández DE, Barrera MJ. Lectura crítica de investigación en el internado de pregrado en hospitales generales. *Rev Med IMSS.* 2005;43:117–24.
21. Cobos AH, Pérez CP, Ramírez MM, Tapia OJ. Lectura crítica de investigación en estudiantes de medicina. Efectos a un año. *Rev Med Inst Mex Seguro Soc.* 2006;44 Supl 3:85–91.
22. Abdullah M, Firmansyah MA. Critical appraisal on journal of clinical trials. *Acta Med Indones-Indones J Intern Med.* 2012;44:337–43.
23. Ghali WA, Saitz R, Eskew AH, Gupta M, Quan H, Hersman W. Successful teaching in evidence-based medicine. *Med Educ.* 2000;34:18–22.
24. Khan KS, Coomarasamy A. A hierarchy of effective teaching and learning to acquire competence in evidence-based medicine. *BMC Med Educ.* 2006;6:59–68.
25. Taylor R, Reeves B, Ewings P, Binns S, Keast J, Mears R. A systematic review of the effectiveness of critical appraisal skills training for clinicians. *Med Educ.* 2000;34:120–5.
26. Taylor R, Reeves B, Mears R, Keast J, Binns S, Ewings P, et al. Development and validation of a questionnaire to evaluate the effectiveness of evidence-based practice teaching. *Med Educ.* 2001;35:544–7.