

Biomédica 2005;25:366-76

ARTÍCULO ORIGINAL

Información de salud en la prensa diaria argentina: adaptación al español y validación del cuestionario *Index of Scientific Quality* para medir su calidad

Emiliano Biondo, Marina Claudia Khoury

Introducción. La prensa diaria constituye una fuente de acceso a la información de salud y puede influenciar sobre la utilización de servicios sanitarios. La calidad de las noticias publicadas en periódicos de países desarrollados es subóptima.

Objetivo. Adaptar al castellano y evaluar la confiabilidad del cuestionario *Index of Scientific Quality*; medir la calidad de la información de salud publicada en la prensa diaria argentina durante los años 2000 a 2002; determinar el nivel de lectura requerido (en grados de escolarización) para comprender los textos; analizar los factores asociados a la calidad.

Materiales y métodos. La adaptación consistió en la traducción del *Index of Scientific Quality*, retrotraducción y prueba piloto. Para la confiabilidad, se seleccionaron artículos con más de 300 palabras referidos a tratamiento, prevención, diagnóstico, efectos del estilo de vida y exposiciones peligrosas sobre la salud. Dos médicos calificaron independientemente 129 artículos. Se cuantificó la concordancia interobservador e intraobservador para cada ítem mediante el coeficiente de correlación intraclass (IC95%). La calidad se evaluó en 210 artículos seleccionados al azar. Se asignaron valores entre 0-100 puntos. El nivel de lectura se evaluó mediante el método gráfico de FRY. Se exploró la relación entre calidad y otras variables mediante un modelo de regresión lineal múltiple.

Resultados. La concordancia interobservador osciló entre 0,48 (0,34-0,61) y 0,67 (0,56-0,75); la concordancia intraobservador entre 0,51 (0,37-0,63) y 0,95 (0,93-0,96); la consistencia interna (alfa de Cronbach), 0,88; la calidad [mediana (rango intercuartílico)]=25 puntos (16,7 y 33,3), y el nivel de lectura [promedio (IC95%)]=10,4 grados (10,2 y 10,6).

Conclusiones. La calidad de la información fue deficitaria. No se encontraron factores que expliquen la baja calidad. El nivel de lectura resultó una barrera de acceso a la información.

Palabras claves: información de salud, prensa diaria, confiabilidad, cuestionario, calidad científica.

Validation of a questionnaire to assess the quality of health information in Argentinian newspapers

Introduction. The daily press is an important source of health information and may influence health care utilization. However, medical reports published in newspapers from developed countries have shown a poor quality.

Objectives. The reliability of the questionnaire *Index of Scientific Quality* was evaluated by using it to measure the scientific quality of health information published by Argentinian newspapers between 2000 and 2002. It assessed the readability of the texts in grade levels and explored the relationship between quality and other factors.

Materials and methods. The Spanish adaptation of the instrument consisted in translation, back-translation and pilot study. The reliability was assessed by applying the instrument to newspaper articles with more than 300 words that discussed therapy, diagnosis, prevention, lifestyle effects, and hazardous exposure. Two physicians independently graded 129 articles. Inter-observer and intra-observer concordance was quantified for each item with the intra-class correlation coefficient (CI95%). To measure scientific quality, a randomized sample of 210 articles was assessed. Each received a mark that ranged from 0 to 100. Readability was determined by the FRY graph method. The relationship between quality and other variables was explored with multiple linear regression analysis.

Results. The inter-rater concordance varied between 0.48 (0.34-0.61) and 0.67 (0.56-0.75). Intra-rater concordance varied from 0.51 (0.37-0.63) to 0.95 (0.93-0.96). The internal consistency (Cronbach's alpha) was 0.88. The quality-of-health-information was rated at 25 points (16.7, 33.3) [median (intercuartil range)]. The reading level was assessed to be at the 10.4 grade (10.2-10.6) [mean (CI 95%)].

Conclusions. Quality of the health information was greatly deficient; however, no specific factors were related with quality. Readability was probably a key barrier for access to the health information.

Key words: health information, daily press, reliability, questionnaire, scientific quality.

Para muchos sectores de la población, los medios masivos de comunicación constituyen la principal fuente de acceso a la información relacionada con la salud (1,2).

Entre los distintos medios, la prensa diaria ha incrementado el espacio para difundir información de salud. Asimismo, a partir de las ediciones publicadas en internet ha pasado a tener mayor circulación. Esta situación ha posibilitado que gran parte de la comunidad pueda acceder al conocimiento de la salud a través de este medio. Esto adquiere gran relevancia si se considera que las campañas realizadas a través de la prensa pueden influir de manera positiva sobre la utilización de los servicios de salud (3). Sin duda que la información, además de ser accesible, debe ser válida y precisa.

Varios estudios recientes que evaluaron las características de la información de salud en países desarrollados, describieron una baja calidad de la misma (4-7). A tal fin se usaron diferentes métodos de medición para determinados temas específicos relacionados con salud; sólo un estudio utilizó un instrumento genérico correctamente desarrollado y validado que permitió medir la calidad de varios tópicos (8). Por otro lado, se han identificado algunos factores que podrían actuar como barreras para mejorar la calidad (9).

En nuestro país se observa con preocupación el incremento de la información en salud y el tipo de fuentes empleadas en los principales medios

gráficos, así como la aparición de empresas que resumen y difunden las mismas noticias sin control. No se encontró una investigación que evalúe la calidad de la información en Argentina; tampoco se disponen de instrumentos válidos y confiables adaptados al español para medir la misma.

Los objetivos del estudio fueron: adaptar al castellano y evaluar la confiabilidad del cuestionario *Index of Scientific Quality* (ISQ); medir la calidad de la información de salud publicada en la prensa diaria argentina durante los años 2000 a 2002; determinar el nivel de lectura requerido (en grados de escolarización) para comprender los textos; analizar la relación entre calidad de información, medios de prensa, año de publicación, nivel de lectura, y extensión del artículo.

Materiales y métodos

Población de estudio

Se realizó un estudio de corte transversal. Se consideró como unidad de análisis aquellos artículos referidos a temas de salud que fueron publicados en diarios nacionales y regionales con circulación diaria promedio mayor a 20.000 ejemplares. Los datos de circulación se obtuvieron del Instituto de Verificación de Circulación de Medios (IVC) de la República Argentina. El período de estudio abarcó los años 2000 a 2002. Se incluyeron artículos que desarrollaron tópicos referidos a intervenciones terapéuticas, o procedimientos diagnósticos, o intervenciones preventivas, o efecto del estilo de vida sobre la salud, o efecto de exposiciones potencialmente peligrosas sobre el estado de salud. En los casos en que dos tópicos coexistieran en un mismo artículo, se optó por seleccionar aquel más desarrollado. A su vez, para poder ser elegible el

Correspondencia:

Emiliano Biondo, Pellegrini 241, Esquel, 9200, Chubut, Argentina.

Teléfono: (02945) 453 234 y (02945) 451 461
lapiruja@yahoo.es

Recibido: 20/12/04; aceptado: 09/06/05

reporte debía contener al menos 300 palabras, y no pertenecer a cartas de lectores.

Los artículos fueron extraídos de una base de datos electrónica que indexa todos los reportes referidos a salud publicados en los diarios argentinos (INFOSALUD®). El autor del proyecto seleccionó y compiló todos aquellos que cumplieron con los criterios de inclusión. Para cada artículo se recolectó la siguiente información: medio de prensa donde fue publicado, número total de palabras y fecha de publicación. En total se obtuvieron 767 artículos que cumplieron con los criterios de selección (261 del año 2000, 182 del año 2001, y 324 del año 2002).

Adaptación transcultural

El cuestionario ISQ se desarrolló y validó originalmente por investigadores de la Universidad de McMaster en Canadá, con la finalidad de medir la calidad de la información relacionada con salud en la prensa gráfica (8). Está constituido por 8 ítems con escala de Likert con rango de 1 a 5 puntos.

En primer lugar, se contactó con el autor del cuestionario quien autorizó su utilización.

Se realizaron dos traducciones iniciales desde el idioma original al español. Las mismas fueron hechas por dos traductores argentinos en forma independiente. Se evaluaron las diferencias entre ambas, y se logró una primera versión argentina.

Para probar la validez del proceso, una persona nativa de los EE.UU. llevó a cabo una retrotraducción (*back translation*) del cuestionario al idioma inglés. Se resolvieron las diferencias mediante el acuerdo de los participantes del proceso, y se llegó a una versión preliminar argentina del ISQ.

Por último, se llevó a cabo una prueba piloto de la versión preliminar. Participaron de la misma una médica pediatra y un médico de familia, ambos con experiencia en el área de investigaciones clínicas y cuestionarios. Cada médico evaluó en forma independiente cinco artículos que cumplieran con los criterios de inclusión. Los mismos fueron seleccionados por el autor del proyecto, procurando un espectro variado de calidad. Al concluir la evaluación, los dos médicos junto con

el autor analizaron las dificultades en la utilización del instrumento a través de la técnica de grupo de discusión. Finalmente, se obtuvo la versión definitiva argentina del ISQ.

Evaluación de la confiabilidad

Para alcanzar este objetivo, se utilizó una muestra de 129 artículos extraída a partir de la población de estudio. El autor del proyecto seleccionó aquellos textos que, según criterios surgidos de recomendaciones de expertos (10-13), presentaron un espectro variado de calidad.

Para calcular el tamaño muestral se asumió obtener un acuerdo entre ambos observadores de 0,7, medido a través del coeficiente de correlación intraclase (CCI). A su vez, se estimó un rango de variabilidad entre 0,6 y 0,8, tolerando un error alfa de 0,05. Previo al cálculo, se realizó la transformación z de Fisher del coeficiente de correlación intraclase para poder asumir su distribución normal.

La variabilidad interobservador se midió entre los dos médicos que participaron en la prueba piloto. Cada uno de ellos evaluó los 129 artículos en forma independiente.

No se dio a conocer la fecha ni el medio de publicación.

La variabilidad intraobservador fue medida en el médico de familia, quien reevaluó los mismos 129 artículos luego de 30 días de haber finalizado su primer calificación.

La consistencia interna se midió en forma individual para cada uno de los dos observadores.

Medición de la calidad y nivel de lectura

La médica pediatra que participó de las etapas anteriores fue la encargada de realizar ambas actividades. Se utilizó una muestra de 210 artículos. Los mismos se seleccionaron a partir de la población de estudio mediante un muestreo aleatorio simple. Se analizaron 70 artículos por cada año. No se dio a conocer la fecha ni el medio de publicación.

Para calcular el tamaño muestral se realizaron las siguientes asunciones: 1) obtener un poder para el estudio de 0,85; 2) aceptar un error alfa de

0,05; 3) poder diferenciar 6 puntos en el nivel de calidad, sobre un marcador total de 45 puntos, entre los diferentes años que se iban a comparar; 4) se consideró obtener una desviación estándar de la calidad de 10 puntos; 5) dado que se compararon tres grupos, se utilizó un factor de corrección de 0,82 en la fórmula empleada.

El instrumento utilizado para medir la calidad fue la versión definitiva del ISQ adaptado y validado al idioma español en Argentina (ver anexo). La escala numérica de medición, resultante de la sumatoria de todos los ítems (marcador total), se transformó en un marcador de calidad con rango de 0 a 100 puntos para facilitar su comparación.

El nivel de lectura requerido para la comprensión de textos se midió a través del método gráfico de FRY validado al idioma español (14). Utiliza una escala de medición numérica, en el que el valor resultante representa el equivalente en grados de escolarización. Dicho valor se obtiene al relacionar a través de un gráfico cartesiano, el número de sílabas cada 100 palabras con el número de oraciones cada 100 palabras.

Análisis estadístico

Para medir la concordancia interobservador e intraobservador se utilizó el coeficiente de correlación intraclase (CCI). La estimación central se calculó con base en el análisis de varianza (ANOVA) para medidas repetidas. Los IC95% se calcularon mediante la transformación z de Fisher.

La consistencia interna se midió a través de la correlación entre cada ítem con el marcador total, y mediante el coeficiente alfa de Cronbach. Ambas mediciones se corrigieron mediante la eliminación del ítem analizado en cada cálculo de correlación.

Para explorar la relación entre calidad de información, medios de prensa, año de publicación, extensión del artículo y nivel de lectura, se construyó un modelo de regresión lineal múltiple. La variable medios de prensa se incluyó al modelo en forma dicotómica. En la categoría 1 se incluyeron las notas publicadas en *Clarín* o *La Nación*, ya que son los medios más reconocidos y de mayor circulación a nivel nacional. En la categoría 0 se agruparon las notas publicadas en el resto de los medios.

La variable año de publicación se incluyó como variable indicadora, tomando como categoría de referencia el año 2001. El resto de las variables se incluyeron de forma numérica.

Para seleccionar las variables asociadas a calidad, se comenzó incorporándolas al modelo todas juntas. El efecto confundidor se evaluó al retirar las variables de a una cada vez, tomando como parámetro la modificación del coeficiente β en más del 15%. Los supuestos de homocedasticidad, linealidad, normalidad e independencia se cumplieron en forma satisfactoria. No se observó multicolinealidad entre las variables regresoras.

Se tomó como valor de significancia estadística un valor alfa de 0,05.

Se utilizó el programa estadístico Stata 7.0 para el análisis de los datos.

Resultados

Adaptación transcultural

Durante el proceso de adaptación transcultural, previo a la prueba piloto, se decidió anexas un ítem local (ítem 4) al cuestionario original, quedando un instrumento definitivo con 9 ítems (ver anexo). Las opciones de respuesta para el nuevo ítem fueron escogidas por los autores de la investigación, y su validez fue probada en la fase piloto por los evaluadores.

Evaluación de la confiabilidad

En el cuadro 1 se resume la calificación asignada por cada observador. Puede notarse un efecto piso marcado en todos los ítems excepto en el ítem 1, en el cual contrariamente se observó un efecto techo.

La concordancia de las calificaciones entre ambos médicos (acuerdo interobservador) mostró un rango entre 0,48 (0,34-0,61) para el ítem 1, y 0,67 (0,56-0,75) para el ítem 9. El acuerdo intraobservador osciló entre 0,51 (0,37-0,63) para el ítem 8, y 0,95 (0,93-0,96) para el ítem 6. En el cuadro 2 se muestra el acuerdo inter e intraobservador para cada uno de los ítems.

La correlación promedio entre todos los pares de ítems (alfa de Cronbach global) fue de 0,87 y 0,89 para cada uno de los médicos. La correlación entre

Cuadro 1. Variabilidad inter-observador. Estadística descriptiva (n =129)
Calificación asignada por cada observador para evaluar la confiabilidad.

Ítems	Rango posible	Media-Mediana		Desvio St.		Skewness		Kurtosis		% al techo§		% al piso©	
		F*	P†	F	P	F	P	F	P	F	P	F	P
Item 1	1-5	4,26 - 5	3,71 - 4	0,93	0,94	-0,95	-0,49	2,73	3,14	54,26	20,93	0	2,3
Item 2	1-5	2,13 - 2	1,95 - 2	1,1	0,81	0,37	0,36	1,93	2,25	0,78	0	39,53	33,33
Item 3	1-5	1,64 - 1	1,65 - 1	0,95	0,86	1,33	1,1	3,8	3,2	0,78	0	62,79	56,59
Item 4	1-5	1,98 - 2	2 - 2	1,15	0,94	0,85	0,4	2,68	2	3,1	0	49,61	37,98
Item 5	1-5	2 - 2	1,47 - 1	1,1	0,76	0,8	1,53	2,8	4,6	3,1	0	44,96	66,67
Item 6	1-5	1,38 - 1	1,1 - 1	0,78	0,36	2,4	4,12	8,85	19,8	0,78	0	74,42	93,02
Item 7	1-5	1,67 - 1	1,16 - 1	0,88	0,46	1,1	3,01	3,2	11,25	0	0	55,04	88,37
Item 8	1-5	1,91 - 1	1,26 - 1	1,1	0,59	0,71	2,56	2,18	9,85	0,78	0	52,71	79,84
Item 9	1-5	1,76 - 2	1,73 - 1	0,89	0,92	0,95	1,23	3	3,98	0	0,78	48,84	51,16

* F = Médico Familia.

† P = Médico Pediatra.

§ % al techo P % de ítems con score = 5

© % al piso P % de ítems con score = 1

Cuadro 2. Acuerdo inter-observador e intra-observador.

Item	Concordancia inter-observador *CCI (IC 95%)	Concordancia intra-observador CCI (IC 95%)
Ítem 1	0,48 (0,34-0,61)	0,64 (0,53-0,73)
Ítem 2	0,58 (0,45-0,68)	0,76 (0,68-0,82)
Ítem 3	0,66 (0,55-0,75)	0,86 (0,81-0,90)
Ítem 4	0,55 (0,41-0,66)	0,89 (0,85-0,92)
Ítem 5	0,66 (0,56-0,75)	0,84 (0,78-0,88)
Ítem 6	0,61 (0,49-0,71)	0,95 (0,93-0,96)
Ítem 7	0,57 (0,44-0,67)	0,63 (0,51-0,72)
Ítem 8	0,53 (0,39-0,64)	0,51 (0,37-0,63)
Ítem 9	0,67 (0,56-0,75)	0,71 (0,61-0,79)

*Coeficiente de correlación intraclase e intervalo de confianza del 95%.

el ítem 1 y el marcador total de calidad (correlación ítem-marcador total corregido) presentó los valores más bajos. En el cuadro 3 se muestra la consistencia interna comparativa entre ambos médicos, y se analiza la variabilidad de la misma al eliminar el nuevo ítem incluido (ítem 4).

Medición de la calidad, nivel de lectura y relación entre variables

El tiempo promedio empleado para medir la calidad científica fue de 15 minutos por artículo. En el cuadro 4 se resume el marcador de calidad asignado y su distribución, según cada ítem y según la sumatoria de los 9 ítems. Pudo notarse un efecto piso significativo en los ítems 3, 5, 6,

7, 8 y 9; el más relevante se notó en el ítem 6, el cual se refiere a la precisión estadística de los resultados.

En el cuadro 5 se describen todas las variables analizadas. En los tres años, al menos, 50% de los artículos presentó un marcador de calidad de 25 puntos o menos, con un rango intercuartil entre 16,7 y 33,3. El marcador más alto correspondió a los artículos publicados en el 2002. Para comprender los artículos de los tres años se requirió un promedio de 10,4 grados de escolarización, con un IC95% entre 10,2 y 10,6.

A través del análisis bivariado se halló asociación estadísticamente significativa entre calidad y año de publicación, y entre calidad y número de palabras del artículo (cuadro 6). Los artículos que fueron publicados en el 2002 tuvieron, en promedio, un marcador de calidad de 9,17 puntos mayor que aquellos artículos publicados en el 2001. Por su parte, la magnitud de la asociación entre el número de palabras y la calidad carece de significado relevante, ya que es necesario incrementar 120 palabras para que la calidad mejore 1 punto sobre un máximo posible de 100.

Al incorporar las 4 variables al modelo múltiple, no pudo probarse la existencia de asociación entre el marcador de calidad y el grado de lectura, ni tampoco entre el marcador de calidad y el medio de prensa. Se mantuvo la asociación observada

Cuadro 3. Consistencia interna comparativa entre ambos médicos, con y sin la inclusión del nuevo ítem.

Ítem	Correlación ítem-score total (corregida)				Alpha de Cronbach (corregido)			
	con inclusión del ítem 4		sin inclusión del ítem 4		con inclusión del ítem 4		sin inclusión del ítem 4	
	F*	P†	F	P	F	P	F	P
Ítem 1	0,34	0,53	0,35	0,49	0,34	0,53	0,87	0,89
Ítem 2	0,74	0,68	0,74	0,69	0,74	0,68	0,83	0,86
Ítem 3	0,74	0,77	0,69	0,77	0,74	0,77	0,83	0,85
Ítem 4	0,57	0,62			0,57	0,62		
Ítem 5	0,71	0,72	0,70	0,73	0,71	0,72	0,83	0,85
Ítem 6	0,65	0,64	0,62	0,65	0,65	0,64	0,84	0,87
Ítem 7	0,50	0,54	0,51	0,54	0,50	0,54	0,85	0,87
Ítem 8	0,39	0,60	0,43	0,60	0,39	0,60	0,87	0,87
Ítem 9	0,90	0,87	0,88	0,85	0,90	0,87	0,81	0,84

Alpha de Cronbach global con inclusión ítem 4 F = 0,87; P = 0,89

Alpha de Cronbach global sin inclusión ítem 4 F = 0,86; P = 0,88

* F = Médico Familia.

† P = Médico Pediatra.

Cuadro 4. Evaluación de la calidad.

Score de calidad asignado sobre el total de artículos (n 210). Estadística descriptiva según cada ítem, y sobre la suma de los 9 ítems

Ítem	Media-Mediana	DS	Skewness	Kurtosis	% al techo	% al piso
Ítem 1	3,38-3	0,94	-0,34	2,9	10	3,33
Ítem 2	2,64-3	1,07	0,11	2,71	5,71	18,57
Ítem 3	1,65-1	0,87	0,95	2,59	0	58,57
Ítem 4	2,63-3	0,94	-0,24	2,78	1,9	16,2
Ítem 5	1,84-2	0,96	0,94	3,24	1,43	48,1
Ítem 6	1,18-1	0,54	3,02	11,22	0	88,57
Ítem 7	1,48-1	0,86	1,86	5,95	0,95	70,95
Ítem 8	1,98-2	0,98	0,84	3	0,95	37,62
Ítem 9	1,78-2	0,84	0,91	3,44	0,48	44,29
Total	18,56-18	5,25	1,13	5,24	0	0

en el análisis divariado. No se comprobó efecto confundidor entre las variables. El modelo definitivo se muestra en el cuadro 7. La capacidad pronóstica de este modelo (R^2 ajustado) fue de 0,0881. Esto implica que el año de publicación y el número de palabras permite explicar menos del 10% de la variabilidad del marcador de calidad.

Discusión

Adaptación transcultural

La decisión de anexar un ítem local (ítem 4) al cuestionario original se debió a que se consideró de fundamental importancia para la calidad de un reporte, describir con claridad la intervención. La

mayoría de las guías de evaluación de estudios científicos, basadas en consenso de expertos, incluyen como uno de los principales ítems la descripción detallada de la intervención, sea ésta una prueba diagnóstica, una exposición a una daño, o una intervención terapéutica o preventiva (10-13). Dichas guías fueron desarrolladas posteriormente al ISQ, por lo cual se consideró pertinente actualizar el instrumento agregando un ítem referido a la intervención.

Confiabilidad del instrumento

Para poder interpretar los valores del coeficiente de correlación intraclase (CCI), una de las clasificaciones más utilizadas es la de Fleiss (15). Valores mayores a 0,90 indican una concordancia muy buena, mientras que un coeficiente menor a 0,30 es malo o nulo; entre 0,31 y 0,50, el acuerdo es mediocre; entre 0,51 y 0,70, moderado, y entre 0,71 y 0,90, bueno.

Por otro lado, los valores de alfa de Cronbach entre 0,70-0,90 indican una consistencia interna óptima (16).

Con base en estas clasificaciones, se puede decir que la concordancia interobservador fue moderada para la mayoría de los ítems y, en algunos casos, mediocre. En el ítem 1 se observó el menor grado de acuerdo, lo cual resultó en una clasificación discordante en 87 pares de artículos.

Cuadro 5. Características los artículos, según año de publicación.

Variables	Score de Calidad (escala 0-100) [¶]	Número de palabras por artículo	Nivel de lectura (años de escolaridad)	% de artículos en Clarín y La Nación
Año 2000 (n=70)				
X (IC 95%)	25,1 (22 ; 28,2)	689 (626,9; 751,1)	10,6 (10,2; 11)	44.3 (32;56)
Pc50 (RI) [‡]	23, 6 (16,7;30,6)	605,5 (481 ; 862)	10 (9 ;12)	
Desvio St.	13	260,4	1,7	
Skewness	0,7	0,7	0,4	
Kurtosis	3	2,6	2,6	
Año 2001(n=70)				
X (IC 95%)	22,7(20,2 ; 25,2)	737,6 (665,3; 809,8)	10,6 (10,2 ; 11)	52,9 (41; 65)
Pc50 (RI)	20,8(16,7 ; 30,7)	717,5 (503 ; 906)	10 (9 ; 12)	
Desvio St.	10,4	303,1	1,7	
Skewness	0,9	1,7	0,7	
Kurtosis	4,9	9,9	3	
Año2002 (n=70)				
X (IC 95%)	31,9 (27,6; 36,1)	744,9 (656,3 ; 833,5)	10 (9,6 ; 10,5)	30 (19;41)
Pc50 (RI)	30,6 (16,7; 41,7)	668,5 (461 ; 942)	10 (9 ; 11)	
Desvio St.	17,9	371,6	1,8	
Skewness	0,8	1,9	0,1	
Kurtosis	4,3	8,6	4,4	

[‡]Mediana (Rango Intercuartilo)

$$\text{score de calidad} = \frac{\text{Score total observado (resultante de la suma de los 9 ítems) - Mínimo valor posible del score total (9 puntos)}}{\text{Máximo valor posible del score total (45 puntos) - Mínimo valor posible del score total (9 puntos)}} \times 100$$

Cuadro 6. Análisis divariado mediante regresión lineal simple (n=210)

Score_cal	Coef. β	P> t	[95% CI]	
Nº palabras	0,0083	0,009	0,0021	0,0146
nivel_lectura	0,4248	0,469	-1,5793	0,7298
año 2000 [¶]	2,3816	0,319	-2,3215	7,0846
año 2002[¶]	9,167	0,000	4,4640	13,8700
Clarín-La Nación	1,3220	0,517	-2,6960	5,3400

[¶] La variable año fue analizada como indicadora. Nivel de referencia: año 2001

Cuadro 7. Modelo de regresión lineal múltiple (n = 210).

Score_cal	Coef. B	P> t	[95% CI]	
año 2000 [¶]	2,7711	0,241	-1,8764	7,4187
año 2002[¶]	9,1077	0,000	4,4692	13,7462
nºpalabras	0,0080	0,010	0,0019	0,0141

[¶] La variable año fue analizada como indicadora. Nivel de referencia: año 2001.

Por su parte, la concordancia intraobservador fue buena y, en algunos casos, moderada. En el ítem 6 hubo una concordancia excelente. Una posible explicación sería que esta pregunta presentó las

calificaciones más bajas en ambas oportunidades, mostrando un efecto piso del 74,4% en la primera evaluación y 72% en la segunda. Al ser claramente mala la calidad de este punto, es de esperar que la correlación sea elevada ya que se facilita la calificación. Esto puede reafirmarse aún más, si tenemos en cuenta que el marcador de este ítem fue el más bajo en toda la población de estudio (cuadro 4).

Por último, la consistencia interna fue óptima y no se observaron diferencias al realizar el análisis con el ítem local y sin él (ítem 4). La correlación entre el ítem 1 y el marcador total de calidad fue la más baja de todos los pares (cuadro 3). Esto puede atribuirse a la calificación elevada del ítem 1 en contraposición con el resto. Dicha calificación podría explicarse con base en que en los artículos periodísticos seleccionados realmente se detalla con claridad la población; o bien, debido a que el ítem es confuso, lo cual puede inferirse a partir de su concordancia interobservador mediocre. El ítem 9 mostró la correlación más alta con el marcador total de calidad, lo cual reafirma su carácter transicional (cuadro 3).

Un aspecto importante a tener en cuenta al interpretar los resultados de la confiabilidad, es que los evaluadores fueron personas con experiencia en investigación y, además, participaron del proceso de adaptación transcultural. Esto podría sobreestimar los resultados y, además, limitar la utilización del instrumento a profesionales entrenados. Este sesgo se introdujo intencionalmente con la finalidad de obtener resultados más precisos en el proceso de evaluación de calidad.

Como conclusión podemos decir que, a pesar de contar con un instrumento con propiedades psicométricas aceptables, es necesario seguir mejorando las mismas.

Calidad de la información, nivel de lectura y relación entre variables

Con respecto a la calidad de la información puede concluirse, pese a no contar con parámetros previos de comparación, que la misma fue deficitaria. Dicho déficit fue, en promedio, del 73,5%.

Con base en el censo del 2001, aproximadamente, 33% de la población mayor de 15 años tenía un nivel de escolarización mayor al 10 grado. Por consiguiente, dado que el nivel de lectura requerido para comprender los textos fue, en promedio, de 10,4 grados de escolaridad, al menos, 66% de la población mayor de 15 años no pudo comprender adecuadamente la información en salud publicada en la prensa gráfica. Sería recomendable tratar de lograr una redacción acorde a un nivel de 7 grado ya que, de esta forma, 83% de los ciudadanos mayores de 15 años podría comprender la información.

Entre las variables analizadas no se encontraron asociaciones relevantes que ayuden a comprender la baja calidad.

Con base en los resultados obtenidos, surge la necesidad de identificar barreras y facilitadores para la mejora de la calidad de la información, proponer intervenciones y evaluar las mismas con un instrumento adecuado que permita el seguimiento de la información a lo largo del tiempo.

En síntesis, se ha obtenido una versión argentina del ISQ para evaluar la calidad de la información

de salud en la prensa gráfica con aceptable confiabilidad, y se observó una baja calidad de la información. Con base en estos resultados, surge la necesidad de proponer ciclos continuos de mejora de la calidad de la información de salud.

Agradecimientos

A Andrew Oxman por autorizar la utilización y adaptación del ISQ.

Conflicto de intereses

Los autores declaran no tener conflictos de intereses.

Financiación

Esta investigación ha sido subsidiada por la Comisión Nacional de Programas de Investigación en Salud (CONAPRIS) del Ministerio de Salud y Ambiente de la Nación Argentina, beca "Carrillo-Oñativia 2003" con apoyo del Centro de Investigaciones Epidemiológicas (CIE) de la Academia Nacional de Medicina.

Referencias

1. **Lupton D, Chapman S.** Two studies of public health news. In: Chapman S, Lupton D, eds. *The fight for public health: principles and practice of media advocacy*. London: BMJ Publishing Group;1994:58-95
2. **Wright WR.** Mass media as sources of medical information. *J Commun* 1975;25:171-3
3. **Grilli R, Freemantle N, Minozzi S, Domenighetti G, Finer D.** Mass media interventions: effects on health services utilisation. *Cochrane Database Syst Rev* 2000;2:CD000389.
4. **Mac Donald MM, Hoffman-Goetz L.** A retrospective study of the accuracy of cancer information in Ontario daily newspapers. *Can J Public Health* 2002;93:142-5.
5. **Johansen LW, Bjorndal A, Flottorp S, Grotting T, Oxman AD.** Evaluation of health information in newspapers and brochures. What can we believe? *Tidsskr Nor Laegeforen* 1996;116:260-4.
6. **Mercado-Martínez FJ, Robles-Silva L, Moreno-Leal N, Franco-Almazan CJ.** Inconsistent journalism: the coverage of chronic diseases in the Mexican press. *J Health Commun* 2001;6:235-47.
7. **Wells J, Marshall P, Crawley B, Dickersin K.** Newspaper reporting of screening mammography. *Ann Intern Med* 2001;135:1029-37.
8. **Oxman AD, Guyatt GH, Cook DJ, Jaeschke R, Heddle N, Keller J.** An index of scientific quality for health reports in the lay press. *J Clin Epidemiol* 1993; 46:987-1001.

9. **Larsson A, Oxman AD, Carling C, Herrin J.** Medical messages in the media-barriers and solutions to improving medical journalism. *Health Expect* 2003;6: 323-31.

10. **Jaeschke R, Guyatt G, Sackett D.** Users' guides to the medical literature, III: how to use an article about a diagnostic test, A: are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994; 271:389-91

11. **Guyatt G, Sackett D, Cook D J.** Users' guides to the medical literature, II: how to use an article about therapy or prevention, A: are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993; 270:2598-601.

12. **Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V.** Users' guides to the medical literature, IV: how to use an article about harm. Evidence-Based Medicine Working Group. *JAMA* 1994;271:1615-9.

13. **Greenhalgh T.** How to read a paper. Assesing the methodological quality of published papers. *BMJ* 1997; 315:305-8

14. **Gilliam B, Pena S, Moutain L.** The Fry graph applied to Spanish readability. *The Reading Teacher* January 1980:426-430

15. **Fleiss JL.** The measurement of interrater agreement. En: Fleiss JL, editor. *Statistical methods for rates and proportions*. 2ª ed. Toronto: Willey press, 1981; p 212-36.

16. **Streiner DL, Norman GR.** Health measurement scales. A practical guide to their development and use. 2ª ed. Oxford: Oxford University Press, 1995; p 65.

APENDICE

Versión definitiva del ISQ en español

Índice de la calidad de la información científica publicada en la prensa de difusión general

1. Aplicabilidad

¿Está claro a quién es aplicable la información contenida en el artículo (es decir, ¿en qué población se puede aplicar la evidencia)?

NO*	PARCIALMENTE			SI
Se induce a potenciales interpretaciones erróneas	Falta de claridad mínima			Ambigüedad mínima
1	2	3	4	5

*es decir que se podría inducir al lector a creer que la información le es aplicable cuando en realidad no lo es.

2. Opiniones versus hechos

¿Los hechos se distinguen claramente de las opiniones?

NO*	PARCIALMENTE			SI
Se induce a potenciales interpretaciones erróneas	Las afirmaciones se atribuyen a las fuentes pero la evidencia subyacente es ambigua			La evidencia subyacente a los puntos principales se cita con claridad
1	2	3	4	5

*es decir que existen afirmaciones concretas que sugieren la existencia de evidencia cuando no se la cita, y no es seguro que exista.

3. Validez

¿La evaluación de la credibilidad (metodología) de la evidencia es clara y tiene buenos fundamentos (no es confusa)? (Al realizar esta evaluación se debe considerar sólo el mensaje explícito, no la evaluación implícita.)

NO*	PARCIALMENTE			SI
No se evalúa o se induce a potenciales interpretaciones erróneas	Se reporta el diseño del estudio o el tipo de evidencia, pero no se evalúa adecuadamente			La solidez de los métodos de investigación está evaluada adecuadamente

1	2	3	4	5
---	---	---	---	---

*es decir que no se da ningún indicio de lo buena que es la evidencia subyacente, o se tergiversa la calidad de la evidencia; donde "calidad" se refiere a la solidez de los métodos de investigación, no a la significancia estadística o coherencia entre los estudios.

4. Intervención

¿Se describe en forma clara la intervención a la cual se le adjudican los resultados principales?

NO*	PARCIALMENTE	SI		
No se describe o se induce a potenciales interpretaciones erróneas	Se menciona, pero en forma poco específica	Se describe de manera detallada y comprensible		
1	2	3	4	5

*es decir que no es posible identificar la intervención a la cual se refieren los resultados.

5. Magnitud

¿Se reporta con claridad el valor o la magnitud de los resultados (efectos, riesgo o costo) que son el foco principal del artículo?

NO*	PARCIALMENTE	SI		
No se reporta o se induce a potenciales interpretaciones erróneas	La magnitud de los efectos o riesgos se reporta en forma incompleta o ambigua (incluyendo, si es relevante, relación dosis-respuesta)	Se reporta con claridad la magnitud de los efectos o riesgos principales el riesgo basal y la		
1	2	3	4	5

*es decir que no se reporta la magnitud (importancia práctica) de los efectos, riesgo o costos que son el foco principal del artículo, o se reporta de manera que probablemente resulte confusa.

6. Precisión

¿Hay una evaluación clara y bien fundada (no confusa) de la precisión de toda estimación que se presenta o de la probabilidad de que algunos de los resultados reportados podrían ser producto del azar?

NO*	PARCIALMENTE	SI		
No se evalúa o se induce a potenciales interpretaciones erróneas	En forma indirecta o incompleta; por ejemplo, se reporta el tamaño muestral pero no se evalúa adecuadamente	Se evalúa adecuadamente la significancia estadística o la precisión		
1	2	3	4	5

*es decir que se puede inducir al lector a tener una suposición errónea sobre la precisión de algunas estimaciones que se reportan, sobre la significancia estadística de los resultados, o sobre a la probabilidad de un error tipo II si se reportan resultados "negativos".

7. Coherencia

¿Se considera la coherencia de la evidencia (entre los estudios), y la evaluación está bien fundamentada (no confusa)? (Al realizar esta evaluación se debe considerar el mensaje explícito, no la evaluación implícita.)

NO*	PARCIALMENTE	SI
-----	--------------	----

No se considera o se induce a potenciales interpretaciones erróneas

Se analiza más de un estudio, pero hay cierta ambigüedad en cuanto a la cantidad de estudios existentes o a su coherencia

Se reportan con claridad la cantidad de estudios y la coherencia (con respecto a la dirección de sus resultados)

1	2	3	4	5
---	---	---	---	---

*es decir que no se cita evidencia, o sólo se cita un estudio sin referencia a otros estudios relevantes (o falta de otros estudios relevantes), o es probable que los estudios citados sean una muestra sesgada de los estudios relevantes disponibles.

8. Consecuencias

¿Se encuentran identificadas todas las consecuencias importantes (beneficios, riesgos y costo) de interés, relativas al tema central del artículo?

NO*	PARCIALMENTE			SI
1	2	3	4	5
Se induce a potenciales interpretaciones erróneas	No se consideran los beneficios, riesgos o costos potencialmente importantes			Se identifican con claridad los beneficios, riesgos y costos más importantes

*es decir que las consecuencias aparentemente importantes (beneficios, riesgos, o costos) relativas a las asociaciones (efectos o riesgo) sobre las cuales está escrito el artículo, no se han considerado

9. Evaluación total de la calidad

Según sus respuestas a las preguntas anteriores, ¿cómo calificaría la calidad científica general del artículo?

BAJA	INTERMEDIA			ALTA
1	2	3	4	5
Falencias críticas o considerables	Falencias potencialmente importantes pero no críticas			Falencias mínimas