

Validación cultural de un instrumento para medir el nivel de conocimiento de bioestadística

Cultural validation of an instrument to measure the level of knowledge in biostatistics

Eduardo Espinoza^{1,a}, Diego Garcés^a

RESUMEN

Objetivo: Realizar la validación cultural y analizar la fiabilidad de un cuestionario para medir el nivel de conocimientos en bioestadística. **Material y método:** El cuestionario fue inicialmente traducido al castellano por 3 traductores, y seguidamente retraducido al inglés. Las traducciones y retraducciones fueron analizadas midiendo la correlación y el grado de acuerdo entre ellas. Posteriormente un comité de expertos construyó un documento de consenso y luego se evaluó el grado de comprensibilidad en 30 sujetos. Finalmente, se realizó el test – retest para evaluar fiabilidad del documento validado. **Resultados:** La correlación entre las parejas de traducciones del inglés al castellano fue $r_s = 0,21$; $r_s = 0,25$ y $r_s = 0,90$ ($p < 0,05$). El grado de acuerdo entre las parejas de traducciones del inglés al castellano fue Kappa = $-0,08$ ($p=0,08$), Kappa = $0,67$ ($p < 0,0000$), y Kappa = $0,58$ ($p < 0,0000$). La correlación entre las parejas de retraducciones del castellano al inglés fue $r_s = 0,97$ ($p < 0,0000$), $r_s = -0,03$ ($p=0,70$) y $r_s = 0,18$ ($p=0,03$). El grado de acuerdo entre las parejas de retraducciones del castellano al inglés fue Kappa = 1 , Kappa = 0 , y Kappa = 0 ($p < 0,05$). El documento de consenso probó ser comprensible en un 82% (IC al 95% de 80,2 a 83,2%). El test-retest en dos periodos mostró un $r_p = 0,94$ ($p < 0,0000$). **Conclusiones:** La adaptación cultural del cuestionario de conocimientos en bioestadística dio origen a un documento final de alta confiabilidad y comprensibilidad.

PALABRAS CLAVE: Validación, bioestadística, estudios de validación, conocimiento. (**Fuente:** DeCS BIREME).

SUMMARY

Objective: To make the cultural validation and measure the reliability of a questionnaire which will assess level of knowledge of bio-statistics. **Methods:** The questionnaire was initially translated to Spanish by 3 translators. Then each translation was retranslated to English. The translations and re-translations were assessed by correlation and measure of agreement. A committee of experts built a consensus document and afterwards degree of comprehension was assessed on a 30 subject group. Finally a test/re-test was done in order to assay the temporal stability of the instrument. **Results:** The correlation between the couples of translations from English to Spanish was $r_s = 0.21$, $r_s = 0.25$ y $r_s = 0.90$ ($p < 0.05$). The degree of agreement between the couples of translations from English to Spanish was Kappa = -0.08 ($p=0.08$), Kappa = 0.67 ($p < 0.0000$), y Kappa = 0.58 ($p < 0.0000$). The correlations between couples of re-translations from Spanish to English was $r_s = 0.97$ ($p < 0.0000$), $r_s = -0.03$ ($p=0.70$) y $r_s = 0.18$ ($p=0.03$). The degree of agreement between the couples of re-translations from Spanish to English was Kappa = 1.0 ($p = 0.00$), Kappa = 0 ($p < 0.05$), y Kappa = 0 ($p < 0.05$). The consensus document showed to be comprehensible in 82% (95% CI 80.2-83.2) Correlation between the two times the questionnaire was taken was 0.94 ($p < 0.0000$). **Conclusions:** Cultural adaptation of the questionnaire about knowledge of bio-statistics originated a final instrument of high compressibility and reliability.

KEYWORDS: Validation studies, biostatistics, validation. (**Source:** MeSH NLM).

¹ Clínica Delgado, Lima-Perú.

^a Médico-cirujano

INTRODUCCIÓN

En Medicina resulta indispensable la participación de otras disciplinas y una de las más importantes es la bioestadística y cada vez es más necesario recurrir a literatura basada en demostraciones estadísticas para llevar a cabo una adecuada práctica clínica, pero muchas veces la comprensión de la Medicina basada en evidencia (MBE) representa un reto para la mayoría de los prestadores de salud (1,2).

La MBE incluye al uso de la mejor evidencia disponible sumada a la experiencia clínica en la toma de decisiones clínicas y realiza la verificación de la evidencia mediante la investigación (3,4,5).

En un estudio realizado entre médicos en actividad, se encontró que los participantes presentaron bajo grado de confianza respecto de sus conocimientos bioestadísticos (6). La actitud de los residentes frente a la MBE fue evaluada en un estudio realizado a nivel nacional en los Estados Unidos. A pesar de que los residentes de medicina interna eran evaluados frecuentemente sobre la aplicación de MBE, estos catalogaron como insuficientes sus conocimientos (1).

En dos hospitales del Perú, se observó que los médicos tenían una actitud favorable frente a la MBE, pero existen marcadas deficiencias en la incorporación de conocimientos provenientes de la misma (7).

En un estudio en Dinamarca se concluyó que el conocimiento de estadística de la mayoría de médicos impedía llegar a conclusiones adecuadas a partir de los análisis estadísticos (2).

Consideramos que una evaluación de los médicos y estudiantes de los últimos años, resulta relevante, no sólo para conocer el nivel en que se encuentran sino también para plantear mejoras debido a que son importantes para que el médico y el estudiante tengan una correcta lectura e interpretación de los artículos científicos.

En el Perú no se cuenta con un instrumento validado para poder realizar ésta medición, por lo que planteamos usar la traducción de un cuestionario ya validado usado en otro idioma para lo cual es necesario realizar una validación cultural de la misma (8)

La adaptación de un instrumento a otro idioma es un proceso complejo y debido a las diferencias culturales, no se puede realizar una simple traducción. En la

literatura se recomienda adaptar escalas y cuestionarios ya validados, argumentando que eso reduce los costos y facilita el intercambio de información entre la comunidad científica (8,9,10).

La adaptación cultural es un proceso importante cuando el instrumento es usado en un idioma diferente, un contexto y tiempo distinto al original, porque de esta manera se puede evitar el riesgo de introducir sesgos en el estudio (11,12).

En el proceso de adaptación a otra lengua se deben considerar de forma distinta la cultura y el país, teniendo en cuenta el idioma, el contexto cultural y el estilo de vida. Este punto debe ser considerado principalmente en las poblaciones hispánicas que viven en otros países, como los Estados Unidos, y que, aun teniendo el mismo idioma, poseen dialectos propios que necesitan ser tomados en consideración. Igualmente menciona que no hay una estandarización para la realización de la validación cultural (9).

El objetivo del trabajo fue realizar la validación cultural de un cuestionario para medir el nivel de conocimientos en bioestadística en externos, internos y residentes y evaluar la fiabilidad del instrumento validado.

MATERIAL Y MÉTODOS

El cuestionario fue elaborado por Donna M. Windish, MD, MPH del Departamento de Medicina Interna de la Escuela de Medicina de la Universidad de Yale, la misma que autorizó la traducción y su uso. Originalmente fue escrita en inglés y consta de preguntas sobre conocimiento acerca de métodos estadísticos, diseño de estudios e interpretación de resultados de estudios (13). Las aptitudes sobre estadística y las preguntas sobre el grado de confianza fueron adaptadas de la encuesta: "Evaluación de las herramientas para el mejoramiento del pensamiento estadístico" llamada ARTIST por sus siglas en inglés (14). Respecto a las preguntas sobre conocimiento de la bioestadística: dos fueron adaptadas de un estudio danés sobre conocimiento estadístico (2), siete fueron adaptadas del curso de estadística de la Escuela de Johns Hopkins Bloomberg School of Public Health (15) y las preguntas restantes fueron desarrolladas por la misma autora.

Si bien no existe un consenso estricto acerca de la metodología para realizar la validación cultural de un instrumento, se trataron de utilizar las recomendaciones

INVESTIGACIÓN ORIGINAL / ORIGINAL RESEARCH

propuestas en la literatura. En la figura 1 se puede apreciar un flujograma de la metodología usada.

Para analizar las coincidencias entre las parejas se contaron el número de palabras en las que coinciden cada par de traducciones y se obtuvo un porcentaje; luego se analizó el grado de acuerdo (concordancia) para lo cual se leyó cada pregunta y cada alternativa de la encuesta traducida y se comparó si se entiende lo mismo entre cada pareja de traducciones y se sacó un porcentaje de las preguntas en las que sí se entiende lo mismo y de las que no; posteriormente se hizo lo mismo con las retro-traducciones. Para correlacionar variables continuas (la comparación entre el conteo de palabras) se utilizó el coeficiente de correlación de Spearman por tratarse de variables sin distribución Normal; y para evaluar la concordancia entre variables categóricas (el acuerdo entre las traducciones, es decir, si se entiende lo mismo por cada pregunta traducida) se usó el estadístico Kappa. Se consideró como estadísticamente significativo un valor de $p < 0,05$. Los datos fueron registrados en el programa Microsoft Excel 2007. Para el análisis, los

datos se transfirieron a un software estadístico STATA versión 11.

Luego, se prosiguió a la evaluación de la comprensibilidad de cada pregunta del cuestionario traducido. Se evaluó que el lenguaje sea claro, que la redacción sea adecuada, y que haya comprensión de las preguntas. Posteriormente se calculó la media del grado de entendimiento de cada pregunta y la media del grado de entendimiento del cuestionario en general con un intervalo de confianza (IC) del 95%. Luego de este proceso se obtuvo el cuestionario validado culturalmente.

Finalmente se analizó la fiabilidad del cuestionario previamente validado, mediante el método test-retest. Para esto se aplicó el cuestionario a 30 personas, las cuales tuvieron que resolver las preguntas, marcando la opción que consideren correcta a cada una de las preguntas de opción múltiple; luego de un plazo entre 20 a 30 días se volvió a tomar el cuestionario a las mismas 30 personas y se evaluó las semejanzas en sus respuestas mediante el coeficiente de correlación de Pearson. La diferencia de

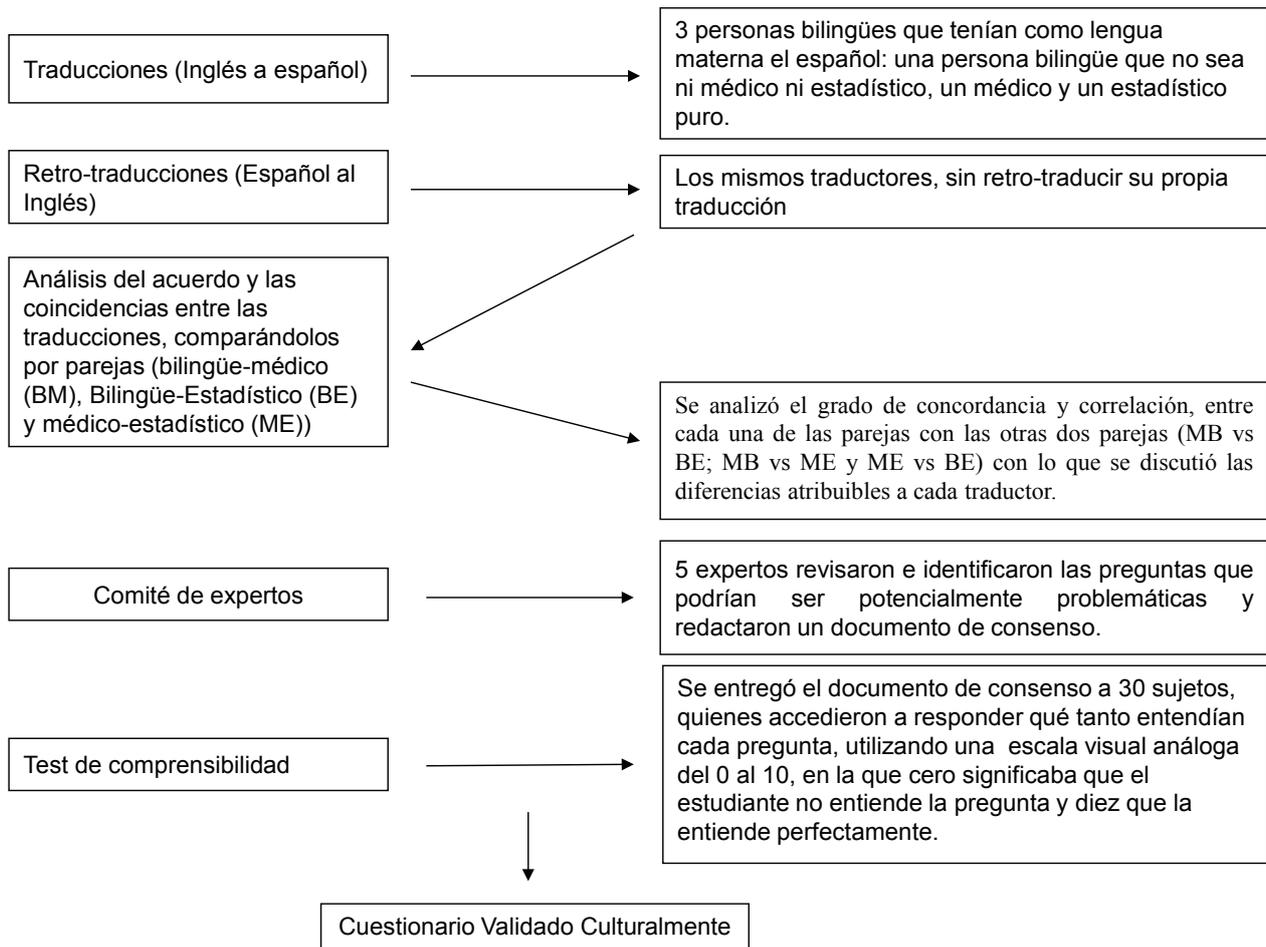


Figura 1. Flujograma de la metodología.

Tabla 1. Grado de acuerdo entre grupos de traductores.

Evaluadores	Acuerdo observado	Acuerdo Esperado	Kappa	p
TRADUCCIÓN DEL INGLÉS AL CASTELLANO				
Médico - Bilingüe vs Médico - Estadístico	84,06%	85,32%	-0,0847	0.8445
Médico - Bilingüe vs Bilingüe - Estadístico	92,75%	78,13%	0,6686	0,0000
Médico - Estadístico vs Bilingüe - Estadístico	92,30%	79,12%	0,5835	0,0000
TRADUCCIÓN DEL CASTELLANO AL INGLÉS				
Médico - Bilingüe vs Médico - Estadístico	100%	97,14%	1	0,0000
Médico - Bilingüe vs Bilingüe - Estadístico	98,55%	98,55%	0.00	>0,05
Médico - Estadístico vs Bilingüe - Estadístico	98,55%	98,55%	0,00	>0,05

Tabla 2. Correlación entre grupos de traducciones y re-traducciones.

	r	p
Traducciones		
Médico - Bilingüe vs Médico - Estadístico	0,21	0,0124
Médico - Bilingüe vs Bilingüe - Estadístico	0,25	0,0029
Médico - Estadístico vs Bilingüe - Estadístico	0,90	0,0000
Re-traducciones		
Médico - Bilingüe vs Médico - Estadístico	0,97	0,0000
Médico - Bilingüe vs Bilingüe - Estadístico	-0,03	0,7009
Médico - Estadístico vs Bilingüe - Estadístico	0,18	0,0338

días permite evitar que las personas recuerden posibles respuestas y que aprendan más acerca del tema. Este análisis nos permitió analizar la estabilidad temporal del instrumento cuando es medido en las mismas condiciones y bajo los mismos parámetros.

La muestra de sujetos, tanto para la evaluación de comprensibilidad como para el análisis de la fiabilidad, fue reclutada de alumnos matriculados de 6to, 7mo año y médicos residentes de la Universidad Peruana Cayetano Heredia (UPCH) que tuvieran como idioma materno el castellano.

El tamaño de la muestra para evaluar la comprensibilidad y el test – retest fue de 10 de 6to año, 10 del 7mo año y 10 médicos residentes de la UPCH, basándonos en la cantidad necesaria descrita en la literatura. Sus nombres permanecieron ocultos y se les entregó una hoja informativa.

RESULTADOS

La evaluación de la concordancia por grupos de parejas de traducciones mostró diferentes grados de acuerdo en los tres grupos. El *Kappa* fue -0,0858 para

Médico-Bilingüe (M-B) vs Médico-Estadístico (M-E) ($p=0,84$); $Kp=0,6686$ para el grupo de M-B vs Bilingüe-Estadístico (B-E) ($p < 0,0000$) y $Kp=0,4183$ para el grupo de M-E vs B-E. ($p < 0,0000$). La concordancia, entre las re-traducciones (castellano al inglés) del M-B vs M-E fue alta (*Kappa*: 1; $p = 0,0000$), en comparación al resto de grupos en los cuales no hubo acuerdo (Tabla 1).

La correlación por grupos de parejas de traducciones en todos los grupos de traducciones mostró una correlación significativa entre ellos. La más alta fue entre M-E vs B-E.

Al evaluar la correlación por grupos de parejas de las re-traducciones, al emparejar M-B vs M-E se encontró una alta correlación ($r_s: 0,9770$, $p: 0,0000$) Al emparejar M-E vs B-E la correlación fue baja y al emparejar M-B vs B-E no hubo correlación (Tabla 2).

El documento de consenso realizado por el comité de expertos puede ser visto en la Figura 2.

La evaluación del grado de comprensibilidad del documento de consenso encontró que la media de

Tabla 3. Grado de comprensibilidad del documento de consenso.

	Media	Desviación estándar	Error estándar	Intervalo de confianza al 95%
Pregunta 1	8,1	1,54	0,28	7,54 a 8,66
Pregunta 2	8,2	1,95	0,36	7,49 a 8,91
Pregunta 3	8,7	1,29	0,24	8,23 a 9,17
Pregunta 4	7,6	1,57	0,29	7,03 a 8,17
Pregunta 5	8,8	1,4	0,26	8,29 a 9,31
Pregunta 6	8,47	1,22	0,22	8,02 a 8,91
Pregunta 7	8,3	1,26	0,23	7,84 a 8,76
Pregunta 8	8,53	1,48	0,27	7,99 a 9,07
Pregunta 9	7,8	1,77	0,32	7,15 a 8,45
Pregunta 10	8,13	1,59	0,29	7,55 a 8,71
Pregunta 11	8,37	1,56	0,29	7,80 a 8,94
Pregunta 12	8,03	1,63	0,3	7,44 a 8,63
Pregunta 13	8,03	1,77	0,32	7,39 a 8,68
Pregunta 14	8,33	1,54	0,28	7,77 a 8,90
Pregunta 15	7,5	1,59	0,29	6,92 a 8,08
Pregunta 16	8,37	1,47	0,27	7,83 a 8,90
Documento de consenso	8,2	0,36	0,09	8,02 a 8,32

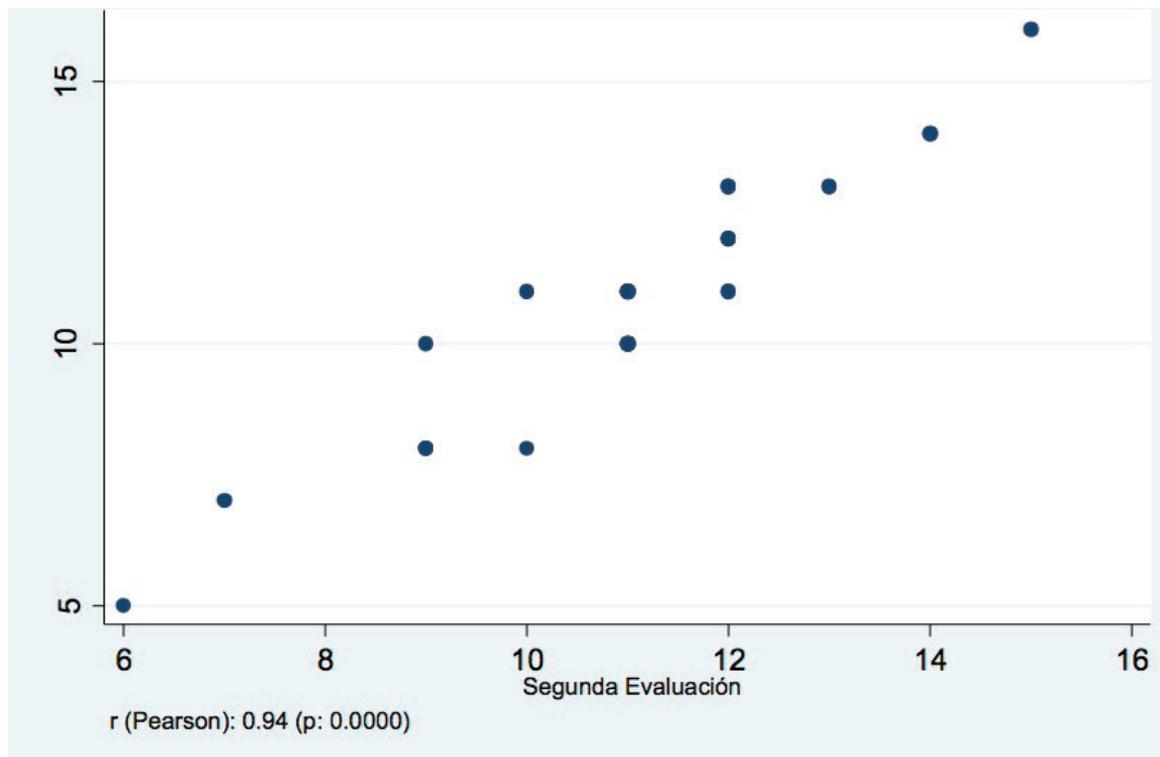


Gráfico 1. Correlación de los scores de conocimientos en bioestadística en dos períodos de evaluación.

cada pregunta varió entre 7,5 y 8,8, siendo la media del documento de consenso de 8,20 (IC al 95%: 8,02 a 8,32) (Tabla 3). Con este último análisis se obtuvo el cuestionario validado culturalmente.

Finalmente se analizó la fiabilidad del documento validado culturalmente para lo que se realizó la toma

del documento de consenso a 30 personas las cuales respondieron el cuestionario en 2 periodos, separados entre sí un mínimo de 20 días y un máximo de 30 días, luego del cual se analizaron las semejanzas en sus respuestas obteniéndose una correlación alta $r_p: 0,94$ ($p=0,000$) (Gráfico 1).

CUESTIONARIO

Por favor escoja la mejor respuesta para cada una de las siguientes preguntas:

1. Un estudio quiere evaluar las características de nacimiento en una población. ¿Cuál de las siguientes variables describe la escala o tipo de medida más adecuada?

(Complete cada espacio en blanco de la siguiente parte con su respuesta. Use cada letra cuantas veces crea apropiada)

- A. Discreta
- B. Continua
- C. Ordinal
- D. Nominal
- E. Dicótoma

- a. ____ Peso de nacimiento en gramos
- b. ____ Peso de nacimiento clasificado como bajo, medio, alto
- c. ____ Tipo de parto clasificado como cesárea, natural, inducido

2. Para determinar si el ayuno está asociado con la fiebre por dengue, se recolectaron datos de 40 pacientes con dengue. Estos pacientes fueron pareados por edad, sexo y raza con 40 pacientes sin dengue. Luego se revisaron las historias clínicas de estos pacientes para determinar si ellos también habían ayunado antes de su enfermedad. Este tipo de estudio es conocido como:

- a. Estudio transversal
- b. Estudio de cohorte prospectivo
- c. Estudio de casos – controles
- d. Estudio de cohorte retrospectivo
- e. Estudio clínico randomizado

3. El propósito de un estudio doble-ciego o doble enmascaramiento es:

- a. Permitir la comparación entre los sujetos tratados y no tratados.
- b. Reducir los efectos de la variación por muestreo.
- c. Evitar el sesgo del observador y del sujeto
- d. Evitar el sesgo del observador y de la variación por muestreo

4. Un estudio prospectivo estudió: obesidad, dieta y hábitos de ejercicio de un grupo de individuos. Correlacione el método analítico adecuado para cada una de las siguientes hipótesis. (Complete cada espacio en blanco de la siguiente parte con su respuesta. Use cada letra cuantas veces crea necesaria.)

- A. T-student para comparar la media de 2 poblaciones
- B. Análisis de Varianza (ANOVA)
- C. Análisis de correlación
- D. Test de homogeneidad de Chi-cuadrado
- E. Regresión logística

- a. ____ La edad media no varía a través de 4 grupos de ingesta de grasas.
- b. ____ El uso de multivitamínicos no varía entre 4 grupos de ingesta de grasas.
- c. ____ La media del IMC es la misma para el grupo de baja y alta ingesta de grasas.

5. Cualquier error en el diseño, realización o análisis que resulta en un falso estimado del efecto de la exposición en el riesgo de la enfermedad, es llamado:

- a. Confusor
- b. Sesgo
- c. Interacción
- d. Estratificación

6. En un estudio placebo-controlado del uso de aspirina y dipyridamol para prevenir la re-estenosis arterial después de una angioplastia coronaria, el 38% de los pacientes que recibieron tratamiento tuvieron re-estenosis, y el 39% de los pacientes que recibieron placebo tuvieron re-estenosis. Al reportar este resultado los autores afirmaron un $P>0.05$. Esto significa:

- a. Las probabilidades de encontrar una diferencia si el estudio se repitiera nuevamente es más que 1 en 20.
- b. La probabilidad de que una diferencia tan grande pueda ocurrir sólo por azar, es menor que 1 en 20.
- c. La probabilidad de que una diferencia tan grande pueda ocurrir sólo por azar, es mayor que 1 en 20.
- d. La probabilidad de que este estudio sea correcto, es de 95%.

7. En el mismo estudio de aspirina y dipyridamol, los investigadores querían evaluar si existían diferencias entre los grupos a través del tiempo con respecto al desarrollo de la re-estenosis, mientras se controlan otros factores de riesgo potenciales. ¿Qué método de análisis sería el más apropiado para esta pregunta de investigación?

- a. Análisis de Kaplan-Meier (curva de supervivencia)
- b. Regresión Logística
- c. Regresión Lineal
- d. Regresión proporcional de Cox
- e. Test de homogeneidad de chi cuadrado

8. En un estudio, la edad promedio de los participantes fue de 26 años \pm 5 años (media \pm Desviación estándar). ¿Cuál de las siguientes afirmaciones es la correcta?

- a. Es 95% correcto que la verdadera media se encuentra entre el intervalo de 16-36 años.
- b. La mayoría de los pacientes tenían 26 años; el resto de tenían entre 21 y 31 años.
- c. Aproximadamente el 95% de los pacientes tenían entre 16 y 36 años.
- d. Ningún paciente tenía menos de 16 o más de 36 años.

9. Unos investigadores miden los niveles de colesterol de una muestra de pacientes de Nueva Zelanda y Asia, y encuentran los siguientes resultados:

Región	Tamaño de la muestra	Media del nivel de colesterol (mmol/L)	Desviación Estándar (mmol/L)
Nueva Zelanda	100	5.4	1.2
Asia	150	4.9	1.3

Ellos calcularon la media y el 95% IC (intervalo de confianza) para la verdadera diferencia de la media de los niveles de colesterol entre las 2 poblaciones, y encontraron: Diferencia de la media del nivel de colesterol de 0.5mmol/L y 95% IC (0.18-0.82).

El 95% IC para la verdadera diferencia de la media de los niveles de colesterol entre las 2 poblaciones sugiere:

- a. No hay diferencia significativa en la media de los niveles de colesterol entre las 2 poblaciones.
- b. Hay una media de niveles de colesterol significativamente mayor en la población de Asia comparada con la población de Nueva Zelanda.
- c. Hay una media de niveles de colesterol significativamente mayor en la muestra de Nueva Zelanda comparada con la muestra de Asia.
- d. Hay una media de niveles de colesterol significativamente mayor en la población de Nueva Zelanda comparada con la población de Asia.

10. Unos investigadores diseñaron un estudio que evaluaba las muertes cardiovasculares comparando una nueva droga contra placebo. Determinaron que necesitarían 200 pacientes en cada grupo para detectar una diferencia del 15% en las muertes cardiovasculares con una potencia del 90% y una significancia del 0.01.

¿Cuál de los siguientes cambios requeriría que los investigadores aumenten el tamaño de la muestra?

- a. Querer detectar una diferencia de 20%
- b. Especificar una potencia de 80%
- c. Utilizar un nivel de significancia de 0.05
- d. Querer detectar una diferencia de 10%

Figura 2. Cuestionario final (Parte 1).

11. En un programa de detección de diabetes, el punto de corte de nivel de glicemia asignado para la prueba A está fijado en 130mg/100ml y para la prueba B está fijado en 160mg/100ml. Esto significa:
- La sensibilidad de la prueba B es mayor que la de la prueba A.
 - La especificidad de la prueba B es mayor que la de la prueba A.
 - La sensibilidad y la especificidad es la misma para las dos pruebas.
 - El número de falsos positivos es mayor con la prueba B que con la prueba A.
12. El tercer Estudio Nacional de Salud y Nutrición fue llevado a cabo en EEUU durante los años 90 y estudió la relación entre obesidad y depresión. Los autores investigaron la asociación entre depresión mayor e índice de masa corporal (IMC) en hombres y mujeres. (American Journal of Epidemiology. 2003; 158: 1139-1147)

Tabla. Odds Ratios no corregido de depresión mayor durante el último mes

Categoría de IMC, kg/m ²	Odds Ratio no corregidos	95% Intervalo de Confianza
Peso normal (IMC 18,5-24,9)	1.00	
Bajo peso (IMC<18,5)	1.17	0.49-2.80
Sobrepeso (IMC 25-29,9)	0.86	0.53-1.41
Obesidad (IMC>30)	1.88	1.02-3.46
Clase 1 (IMC 30-34,9)	1.28	0.64-2.56
Clase 2 (IMC 35-39,9)	1.76	0.78-3.95
Clase 3 (IMC>= 40)	4.98	2.07-11.99

De la tabla en la parte superior, ¿Cuál es la interpretación correcta del valor 0.86 en el grupo con sobrepeso?

- La posibilidad (odds) de una persona con sobrepeso de sufrir de depresión mayor es 14% menor que las de los individuos con peso normal.
 - La posibilidad (odds) de una persona con sobrepeso de sufrir de depresión mayor es 14% mayor que las de los individuos con peso normal.
 - Una persona con sobrepeso tiene una probabilidad de 0.86 de sufrir de depresión mayor.
 - Una persona con sobrepeso tiene una posibilidad (odds) de 0.86 de sufrir de depresión mayor.
13. Los autores del estudio previo también analizaron en la misma población el odds ratio corregido para depresión mayor durante el mes previo, y los resultados fueron:

Basado en la tabla, cuál de las siguientes afirmaciones es correcta:

- La posibilidad de depresión aumenta de forma estadísticamente significativa en personas con IMC mayor de 40 comparadas con personas con IMC de otras categorías, controladas para todas las otras co-variables.
- La posibilidad de depresión en personas con IMC normal es estadísticamente significativo menor que otras personas con IMC mayor a 40, controlada para todas las otras variables.
- La posibilidad de depresión disminuye significativamente al aumentar la edad, controlada para todas las otras co-variables.
- No hay diferencia estadísticamente significativa para la posibilidad de depresión entre los sexos, controlada para todas las otras co-variables.

Tabla. Odds Ratio ajustado para depresión mayor durante el mes pasado

Categoría de IMC, kg/m ²	Odds Ratio ajustado	95% Intervalo de Confianza
Peso normal (IMC 18,5-24,9)	1.00	
Bajo peso (IMC<18,5)	1.13	0.43-3.01
Sobrepeso (IMC 25-29,9)	0.96	0.57-1.64
Obesidad (IMC>30)	1.84	0.95-3.55
Clase 1 (IMC 30-34,9)	1.33	0.57-3.13
Clase 2 (IMC 35-39,9)	1.90	0.79-4.60
Clase 3 (IMC>= 40)	4.63	2.06-10.42
Sexo		
Masculino	1.00	
Femenino	2.62	1.76-3.92
Edad		
15-19	1.00	
20-24	0.80	0.36-1.76
25-29	0.61	0.22-1.69
30-34	0.64	0.30-1.39
Raza/Etnia		
Blanca	1.00	
Afroamericano	0.80	0.48-1.32
Hispana /otra	1.02	0.53-1.94

14. Como parte del Estudio sobre la Salud de la Mujer, los investigadores querían investigar el rol de la inflamación sistémica como predictor del riesgo cardiovascular en la mujer (N Engl J Med 2002; 347(20):1557-1565).

Los investigadores usaron un diseño de estudio prospectivo con un gran tamaño muestral. Los niveles de proteína C reactiva (PCR) fueron medidos al inicio y las mujeres fueron seguidas por un promedio de 8 años. La siguiente tabla muestra los riesgos relativos estimados para un evento cardiovascular (infarto de miocardio agudo (IMA) o Accidente cerebrovascular (ACV)) dentro de los siguientes 5 años para los quintiles de los niveles de PCR de las 30000 mujeres estudiadas. El primer quintil se usa como categoría de referencia.

	1	2	3	4	5
Quintil de PCR	0.49 mg/dl	>0.49-1.08 mg/dl	>1.08-2.09 mg/dl	>2.09-4.19 mg/dl	>4.19 mg/dl
Riesgo Relativo	1.0	1.8	2.3	3.2	4.5
Número de mujeres	6000	6000	6000	6000	6000

Basados en los datos de riesgo relativo en la parte superior, se puede concluir:

- No hay riesgo de IMA/ACV en mujeres con PCR dentro del primer quintil.
 - La disminución del nivel de PCR aparentemente aumenta el riesgo de IMA/ACV
 - El aumento del nivel de PCR aparentemente aumenta el riesgo de IMA/ACV.
 - No parece haber asociación entre los valores de PCR y el riesgo de IMA/ACV.
15. El Estudio Nacional de Evaluación del Riesgo de Osteoporosis estudió en EEUU a 200,160 mujeres postmenopáusicas a partir de los 50 años. Al inicio del estudio 14,412 de estas mujeres tenían osteoporosis definida con una densitometría ósea con score T ≤ 2,5 (JAMA. 2001;286 [22]:2815-2822)

La siguiente tabla muestra los resultados del modelo de regresión logística multivariada de correlación con osteoporosis.

Tabla3. Modelo de Regresión Logística Multivariada. Correlación estadísticamente significativa para score T ≤ -2.5

Figura 2. Cuestionario final (Parte 2).

Factores de Riesgo	OR (Intervalo de Confianza al 95%)
Grupo Edad, años	
50 a 54	1.00 (referencia)
55 a 59	1.79 (1.56-2.06)
60 a 64	3.84 (3.7-4.37)
65 a 69	5.94 (5.24-6.74)
70 a 74	9.54 (8.42-10.81)
74 a 79	14.34 (12.64-16.26)
>80	22.56 (19.82-25.67)
Años desde la menopausia	
<5	1.00 (referencia)
6 a 10	0.79 (0.70-0.89)
11 a 15	0.83 (0.76-0.91)
16 a 20	0.96 (0.89-1.03)
21 a 25	1.01 (0.95-1.08)
26 a 30	1.02 (0.95-1.09)
31 a 35	1.10 (1.03-1.19)
36 a 40	1.14 (1.05-1.24)
> 41	1.24 (1.14-1.35)
Educación universitaria pregrado o superior	0.91 (0.87-0.94)
Estado de salud auto-tasado	
excelente	1.00 (referencia)
muy bueno	1.04 (0.97-1.13)
bueno	1.23 (1.14-1.33)
justo-pobre	1.62 (1.50-1.76)
Historia de Fractura	
cañera	1.96 (1.75-2.20)
muñeca	1.90 (1.77-2.03)
esquina	1.34 (1.17-1.54)
costilla	1.43 (1.32-1.56)
Historia matema de osteoporosis	1.08 (1.01-1.17)
Historia matema de fractura	1.16 (1.11-1.22)
Raza/etnia	
blanco	1.00 (referencia)
afro-americano	0.55 (0.48-0.62)
nativo americano	0.97 (0.82-1.14)
hispano	1.31 (1.19-1.44)
asiático	1.56 (1.32-1.85)
Índice de masa corporal (IMC), kg/m²	
<23	1.00 (referencia)
23.01-25.99	0.46 (0.44-0.48)
26-29.99	0.27 (0.26-0.28)
>30	0.16 (0.15-0.17)
Uso de medicación actual	
cortisona	1.63 (1.47-1.81)
diuréticos	0.81 (0.76-0.85)
Uso de estrógenos	
anterior	0.77 (0.73-0.80)
actual	0.27 (0.25-0.28)
Fumar cigarrillos	
anterior	1.14 (1.10-1.19)
actual	1.58 (1.48-1.68)
Ejercicio regular	0.86 (0.82-0.89)
Uso de alcohol, bebidas/semana	
ninguna	1.00 (referencia)
1 a 6	0.85 (0.80-0.90)
7 a 13	0.76 (0.69-0.83)
>14	0.62 (0.54-0.71)

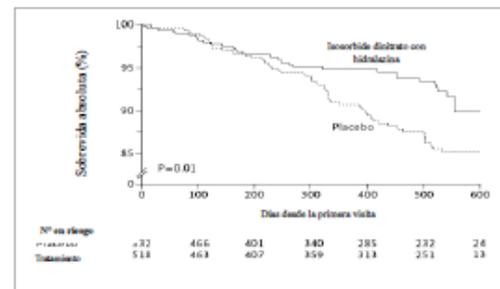
Tecnología	
Absorciometría de Rayos X de energía única simple de talón	1.00 (referencia)
Absorciometría de doble energía de rayos X periférica de antebrazo	2.86 (2.75-2.99)
Absorciometría de doble energía de rayos X periférica de dedo	4.86 (4.56-5.18)
Ecografía de talón	0.79 (0.70-0.90)

Además de la edad, la asociación más fuerte con osteoporosis es:

- fumar cigarrillos
- años transcurridos desde la menopausia
- IMC
- uso de estrógenos
- historia matema de fractura

16. Un estudio randomizado controlado fue conducido para determinar si la combinación de isorbide dinitrato con hidralazina en pacientes afro-americanos con insuficiencia cardiaca, era superior al placebo con respecto a la sobrevivida. (N Eng J Med. 2004;351[20]:2049-2057).

Un análisis de Kaplan Meier se realizó y es mostrado en la parte inferior.



Basados en esta información se puede concluir:

- El riesgo absoluto de muerte es significativamente menor en el grupo de tratamiento vs el de placebo comparando todos los puntos de tiempo entre el día 0 y el día 600.
- El riesgo absoluto de muerte es significativamente menor en el grupo de tratamiento vs el de placebo en el día 600.
- El riesgo absoluto de muerte es significativamente mayor en el grupo de tratamiento vs el de placebo en el día 600.
- El riesgo absoluto de muerte es aproximadamente el mismo en ambos grupos.

Figura 2. Cuestionario final (Parte 3).

DISCUSIÓN

El proceso de adaptación cultural está orientado a lograr la medición de un mismo fenómeno en diferentes culturas; esencialmente es la producción de un instrumento equivalente adaptado a otra cultura. En el Perú no existe un instrumento validado para medir el nivel de conocimiento en Bioestadística en estudiantes de medicina ni médicos residentes.

Si bien no existe un consenso estricto en la literatura de cómo debe realizarse la adaptación cultural de un instrumento, se propone que para realizarla es importante tener en cuenta dos aspectos: la validación cultural que realiza una evaluación lingüística que analiza aspectos de equivalencia entre el instrumento original y la versión traducida; y la evaluación de las propiedades psicométricas del instrumento validado (16,17).

Se recomienda que para las traducciones participen al menos 2 traductores cuya lengua materna sea la del idioma de destino (17). Nosotros hemos utilizado 3, los cuales por su ámbito profesional abarcan los tópicos que son abordados en el cuestionario. Al hacer las traducciones del inglés al castellano se encontró que la concordancia entre el M-B vs B-E fue sustancial, esto se podría explicar porque el estadístico al momento de hacer la traducción al español es el que ha usado el lenguaje más técnico entre los tres, debido a que es el que maneja el nivel más alto de conocimiento respecto al tema, mientras que el médico lo tradujo de manera más sencilla tratando de buscar el sentido de lo que traducía y el bilingüe lo hizo de manera más coloquial, lo cual lo hacía más entendible para la mayoría; es por esto que la versión hecha por el bilingüe tuvo mayor relevancia al comparar la concordancia. En cambio, la correlación fue significativa para todas las parejas aunque en diferente grado, siendo el más alto en la del M-E vs B-E, lo cual se puede entender debido a que tanto el médico como el bilingüe trataron de usar palabras que se adecuaban al lenguaje técnico del tema.

Para las retro-traducciones, la literatura recomienda que sean otros traductores, cuya lengua materna sea la del cuestionario original (17). En nuestro caso, por dificultad para conseguir profesionales con lengua materna inglesa, recurrimos a los mismos traductores, lo que explica los resultados encontrados en el análisis de las retro-traducciones. El grado de correlación y concordancia entre las re-traducciones fue significativo únicamente al comparar las parejas M-B vs M-E. Esto podría ser explicado porque es el médico el que maneja un mejor vocabulario para traducir al inglés términos

estadísticos y hacerlos entendibles, buscando el sentido de cada uno de los ítems evaluados; y no solamente una traducción literal (que es hecha por el bilingüe) ni una muy técnica (que es hecha por el estadístico). Además hay que tener en cuenta que los que realizaron las retro-traducciones ya tenían conocimiento del tema, debido a que habían realizado las traducciones al castellano previamente.

En base a las traducciones y retro-traducciones, el comité de expertos, que estuvo conformado por cinco médicos con estudios que respaldan sus conocimientos en bioestadística, (4 con idioma materno español y uno inglés), elaboró un documento único de consenso, que al ser evaluado mostró un alto grado de comprensibilidad en cada uno de las 16 preguntas por parte de los externos, internos y residentes de la Facultad de Medicina de la UPCH.

Después del análisis estadístico realizado a las traducciones, la evaluación por el comité de expertos y la prueba de comprensibilidad a la que se sometió el cuestionario, creemos que se ha preservado el significado original de cada uno de las preguntas, logrando la equivalencia de significados entre la cultura en que originalmente se desarrolló el cuestionario y la nueva población a la cual está dirigida. Luego de esto se puede decir que el cuestionario está validado culturalmente.

La literatura menciona que la correcta traducción y adaptación cultural de un cuestionario no siempre garantiza la preservación de sus propiedades psicométricas: fiabilidad, consistencia interna, validez de contenido, de criterio y sensibilidad al cambio (9,11,17).

En este trabajo únicamente se evaluó la fiabilidad intra-observador del instrumento mediante la técnica del test-retest; donde mostró una alta confiabilidad. La variabilidad inter – examinadores (inter-observadores) no fue medida debido a que el cuestionario únicamente debe ser entregado a la persona que lo va a responder y esto no variaría los resultados del cuestionario. De esta forma quedaría pendiente analizar las otras propiedades psicométricas del cuestionario.

La metodología utilizada para obtener un instrumento validado culturalmente, puede ser empleada en diversos ámbitos; de esta forma se puede recolectar información valiosa en un país tan heterogéneo como el nuestro.

Dentro de las limitaciones del estudio, debemos considerar que no hay un consenso estricto para realizar la validación cultural, así que nos hemos basado en uno

de los métodos más comunes, dejando de lado otras metodologías para realizar la validación cultural, los cuales difieren en análisis posteriores a la obtención del cuestionario traducido, que forman parte de la evaluación psicométrica del documento final.

En conclusión, la adaptación cultural del cuestionario de conocimientos en bioestadística dio origen a un documento final de alta comprensibilidad y confiabilidad que puede ser usada para analizar el nivel de conocimiento en bioestadística de externos, internos y residentes.

Declaración de financiamiento y de conflictos de intereses:

El estudio fue financiado por los autores y declaran que no hay conflicto de intereses entre los mismos.

Contribución de autoría:

Ambos autores han contribuido en igual proporción a la elaboración del artículo.

Correspondencia:

Eduardo Antonio Espinoza Olórtegui
Av. La Alborada 1684 Lima 1, Perú.
Correo electrónico: eduardo.espinoza@upch.pe
Teléfono: +51 993494856

REFERENCIAS BIBLIOGRÁFICAS

1. Dellavalle RP, Stegner DL, Deas AM, et al. Assessing evidence-based dermatology and evidence based internal medicine curricula in US residency training programs: a national encuesta. *Arch Dermatol.* 2003; 139(3):369-372.
2. Wulff HR, Andersen B, Brandenhoff P, Guttler F. What do doctors know about statistics? *Stat Med.* 1987; 6(1):3-10.
3. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA.* 1992; 268(17):2420-5.
4. Ashcroft R. Current epistemological problems in evidence based medicine. *J Med Ethics.* 2004; 30: 131-135.
5. Sánchez-Mendiola M. La medicina basada en evidencias y la escuela de Medicina. *Gac Méd Méx.* 2004; 140(3): 314-316.
6. Young JM, Ward JE. Evidence-based medicine in general practice: beliefs and barriers among Australian GPs. *J Eval Clin Pract.* 2001;7(2):201-210.
7. Canelo C. Conocimientos, actitudes y prácticas de la medicina basada en evidencias en médicos asistentes y residentes en dos hospitales de Lima-Perú. *Rev Med Hered.* 2007;18 (2):76-84.
8. Alexandre N, Guirardello E. Adaptación cultural de instrumentos utilizados en salud ocupacional. *Revista Panamericana de Salud Pública* 2002; 11(2):109-111.
9. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology.* 1993; 46: 1417-1432.
10. Hutchinson A, Bentzen N, Konig-Zanhn C. Cross cultural health outcome assessment: a user's guide. Groningen: European Research Group on Health Outcomes; 1997.
11. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol.* 2010; 10:13.
12. Herdman M, Fox-Rushby J, Badia X: A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res.* 1998; 7:323-335.
13. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA.* 2007; 298(9):1010-22.
14. Assessment Resource Tools for Improving Statistical Thinking (ARTIST). ARTIST Online Tests Minnesota: University of Minnesota. Citado el 9 de enero del 2007. Disponible en: <https://apps3.cehd.umn.edu/artist/>
15. Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health. Course materials from Statistical Methods in Public Health II and III, 2003-2004 academic year. Baltimore: Johns Hopkins Bloomberg School of Public Health (Citado el 2 de agosto del 2007. Disponible en: <http://www.biostat.jhsph.edu/courses/bio622/index.html>
16. Guyatt G, Patrick D. Evaluation of translation and cultural adaptation in global clinical trials included in Cochrane reviews. Sao Paulo, Brazil.:XV Cochrane Colloquium; Oct 23-27. 2007.
17. Ramada-Rodilla JM, Serra-Pujadas C, Delclós-Clanchet GL. Adaptación cultural y validación de cuestionarios de salud: revisión y recomendaciones metodológicas. *Salud Pública de México.* 2013; (55) 57-66.

Recibido: 16/03/2016
Aceptado: 03/06/2016