

Validación de escala MinCir de estudios de pronóstico*

Drs. HÉCTOR LOSADA M.^{1,2}, CARLOS MANTEROLA D.^{1,2}, VIVIANA PINEDA N.^{1,2}, MANUEL VIAL G.^{1,2}, ANTONIO SANHUEZA C.¹. GRUPO MINCIR (Metodología e Investigación en Cirugía)

¹ Departamento de Cirugía y Traumatología. Universidad de La Frontera.

² Servicio de Cirugía. Hospital Hernán Henríquez Aravena. Temuco, Chile.

Abstract

Methodological quality in prognostic papers (scale MinCir Pr)

Introduction: Our team reported the development of a scale for assessing methodological quality in prognostic papers (scale MinCir Pr). The aim of this paper is to report the process of validation of scale MinCir Pr. **Material and Methods:** We reviewed 121 papers about prognosis. Score was performed according to the scale MinCir Pr and level of evidence. In addition we recorded the number of publications of the author. For the construct validity of the measurement was made through the technique known extreme groups, which claims that the evaluations differ regarding the critical attribute. To this end, we used the level of evidence items 2 and 4, waiting for a difference in scale scores of articles published at the ends of the score. **Results:** For validation of extreme groups we compare the performance of the scale in relation to the level of evidence 2 and 4, since in these groups had more papers (55 and 48 respectively). The average score of the scale in papers with level of evidence 2 was 82.21 ± 13.79 compared with the score in papers with level of evidence 4 which was 64.97 ± 17.10 ($p = 0.000$). **Conclusion:** We performed the validity of the scale builder MinCir Pr using the technique known extreme groups.

Key words: Prognosis (MeSH), evidence based medicine (MeSH), design of scale, methodological quality.

Resumen

Introducción: Nuestro equipo de trabajo reportó el desarrollo de una escala de evaluación de calidad metodológica de estudios de pronóstico (escala MinCir Pr). El objetivo del presente trabajo es reportar el proceso de validación de la escala MinCir Pr. **Material y Método:** Se revisaron 121 estudios de pronóstico. Se realizó la puntuación de los trabajos según la escala MinCir Pr y según el nivel de evidencia. Además se registró el número de publicaciones del autor. Para la validez de constructo se realizó la medición por medio de la técnica de grupos extremos conocidos, que presume que las evaluaciones diferirán respecto al atributo crítico. Para este fin, se utilizaron los artículos con nivel de evidencia 2 y 4, esperando una diferencia en la

*Recibido el 17 de abril de 2011 y aprobado para publicación el 15 de mayo de 2011.

Financiado por el proyecto DIDUFRO 120629.

Correspondencia: Dr. Héctor Losada M.
Casilla 54-D. Temuco, Chile.
hflosada@ufro.cl

puntuación de la escala de los artículos publicados en los extremos de la puntuación. **Resultados:** Para realizar la validación de grupos extremos comparamos el desempeño de la nueva escala en relación a los niveles de evidencia 2 y 4; ya que en estos grupos había mayor número de artículos (55 y 48 respectivamente). El promedio de puntuación de la nueva escala de los trabajos con nivel de evidencia 2 fue de $82,21 \pm 13,79$ en comparación con la puntuación de los trabajos con nivel de evidencia 4 que fue de $64,97 \pm 17,10$ ($p = 0,000$). **Conclusión:** Se realizó la validez de constructor de la escala MinCir Pr utilizando la técnica de grupos extremos conocidos.

Palabras clave: Pronóstico (MeSH), medicina basada en evidencia (MeSH), diseño de escala, calidad metodológica.

Introducción

Uno de los aspectos más importantes del trabajo del grupo MinCir en diseño de escalas de evaluación de calidad metodológica es el desarrollo del criterio de multi-dimensionalidad¹⁻³. Para nuestro equipo el concepto de calidad metodológica de un estudio tiene relación con múltiples aspectos: diseño del estudio, tamaño de la muestra, metodología usada en el estudio, calidad del reporte, entre otras.

El propósito de diseñar escalas de evaluación de calidad metodológica con puntuación numérica tiene como objetivo utilizar nuestra escala en revisiones sistemáticas. Nuestro equipo ha empleado en su metodología el cálculo de promedio ponderado para este fin.

Otro aspecto importante es la aplicabilidad de la escala a diferentes diseños de investigación; lo cual la hace aplicable tanto en revisiones sistemáticas donde existen estudios de buena calidad de evidencia, como en revisiones sistemáticas donde existen estudios de calidad de evidencia heterogénea.

Nuestro equipo de trabajo reportó el desarrollo de una escala de evaluación de calidad metodológica de estudios de pronóstico (escala MinCir Pr), cumpliendo todos los pasos metodológicos del diseño de escalas en investigación¹.

El objetivo del presente trabajo es reportar el proceso de validación de la escala MinCir Pr.

Material y Método

Diseño

Esta etapa consiste en la aplicación de la escala la determinación de la validez de constructo. Para la validez de constructo se realizó la medición por medio de la técnica de grupos extremos conocidos, que presume que las evaluaciones diferirán respecto al atributo crítico. Para este fin, se utilizaron los artículos con nivel de evidencia 2 y 4, esperando una diferencia en la puntuación de la escala de los artículos publicados en los extremos de la puntuación.

Población

La población fueron 122 estudios de pronóstico

que fueron revisados por 4 profesionales con experiencia en investigación, diferentes a los participantes en la etapa de diseño de la escala. Además se analizó el número de publicaciones del autor y se calificó el nivel de evidencia de cada artículo según el centro de evidencia de Oxford^{1,4}.

Tamaño de la muestra

El tamaño de la muestra para el proceso de generación de la escala fue 122 artículos. Conservamos este número para el proceso de validación¹.

Selección de artículos

Para el proceso de validación se incluyeron los artículos que fueron utilizados en el proceso de generación de la escala¹.

Selección de investigadores

Los criterios de inclusión para los investigadores en esta segunda etapa fueron:

1. Certificación de epidemiólogo clínico o bioestadístico.
2. Pertenecer a una unidad o centro de epidemiología clínica.
3. Aceptación de participación en el protocolo.

Plan de análisis

Se utilizó estadística descriptiva, con cálculo de medidas de tendencia central, dispersión y tendencia extrema; análisis divariado utilizando χ^2 de Pearson o test exacto de Fischer para variables categóricas y t-test, ANOVA y Kruskal Wallis para variables continuas. Los datos fueron analizados utilizando el programa Stata® 8.0.

Aspectos éticos y de financiamiento: Este protocolo fue financiado por el proyecto DIDUFRO N° 120629 de la Dirección de Investigación y Desarrollo de la Universidad de La Frontera.

Resultados

Recordemos que la escala MinCir Pr tiene un total de 25 ítems, 6 dominios y un coeficiente de alfa de Cronbach de 0,89. La puntuación es dicotómica, utilizando 1 para el ítem que no fue cumplido por

el artículo y 5 para el ítem que fue cumplido por el artículo. Teóricamente el valor más alto de la escala es de 125 y el valor más bajo es de 25.

La escala (MinCir Pr) se describe en la Tabla 1.

La puntuación promedio de la escala para el total de los estudios fue de $74,88 \pm 16,99$; con una mediana de 78,5 y valores extremos de 29 para el estudio peor puntuado y 109 para el estudio mejor puntuado.

La distribución de la puntuación general de la escala se describe en la Figura 1.

La puntuación de nuestro escore en relación a los diferentes diseños de investigación se describe en la Tabla 2.

El comportamiento de nuestro escore en relación al número de publicaciones del autor se describe en la Figura 2.

Tabla 1. Escala MInCir para estudios de pronóstico (MInCir Pr).

	1	5
Dominio 1. (Variable de desenlace)		
Se define la variable de desenlace		
El método para medir la variable de desenlace es válido y confiable		
Se define las unidades de medida de la variable de desenlace		
Una adecuada proporción de la muestra tiene datos completos de la variable de desenlace		
El método y las unidades de medición de la/las variables de desenlace son los mismos para todos los pacientes		
Dominio 2. (Variable de exposición)		
Se define la variable de exposición		
El método para medir la variable de exposición es válido y confiable		
Se define las unidades de medida de la variable de exposición		
Una adecuada proporción de la muestra tiene datos completos de la variable de exposición		
El método y las unidades de medición de la/las variables de exposición son los mismos para todos los pacientes		
Dominio 3. (Seguimiento)		
Se enuncia el porcentaje de pérdida de seguimiento		
Se describen las características de los sujetos perdidos		
Se hicieron esfuerzos por recolectar la información de los sujetos perdidos		
Las causas de pérdida de seguimiento son descritas		
Se analizaron las diferencias entre las características de los sujetos con pérdida de seguimiento y los que no		
Dominio 4. (Análisis y conclusiones)		
Con los datos presentados es posible determinar las medidas de riesgo y asociación		
Se realizan modelos de asociación o predictivos		
El modelo elegido es adecuado según el diseño del estudio		
Las conclusiones del estudio son coherentes con el objetivo y la metodología empleadas		
Dominio 5. (Diseño y variables de confusión)		
Se enuncia claramente el diseño del estudio		
Considera variables de confusión		
Se definen adecuadamente las variables de confusión		
Los métodos para medir las variables de confusión son válidos y confiables		
Dominio 6. (Tamaño de la muestra)		
Hay cálculo del tamaño de la muestra		
Se justifica el tamaño de la muestra		
Total		

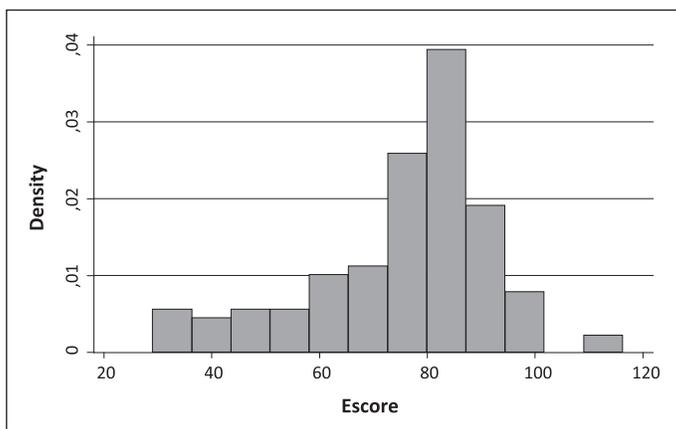


Figura 1. Distribución de la puntuación de la escala MInCir Pr. Escala MInCir Pr = escala de pronóstico MInCir. MInCir = metodología e investigación en cirugía. Escore = puntuación de la escala MInCir Pr.

Tabla 2. Puntuación escala MInCir Pr según diseño de investigación

Diseño de investigación	n	Escore Escala MInCir Pr (X ± DE)
Meta-análisis	1	73
Revisión sistemática de la literatura	1	49
Cohorte prospectiva	53	83,39 ± 11,16
Cohorte retrospectiva	28	81 ± 9,03
Corte transversal	1	77
Serie de casos	38	59,18 ± 17,44
Total	122	74,88 ± 16,99

X = promedio. DE = desviación estándar.

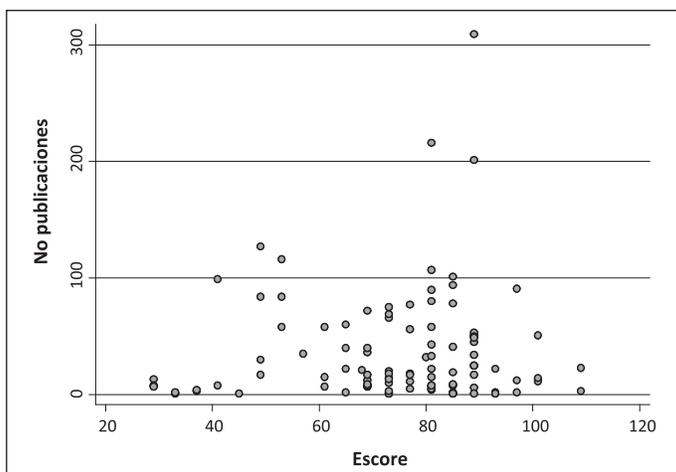


Figura 2. Puntuación escala MInCir Pr según número de publicaciones del autor. Escala MInCir Pr = escala de pronóstico MInCir. MInCir = metodología e investigación en cirugía. Escore = puntuación de la escala MInCir Pr. No publicaciones: número de publicaciones del autor.

El coeficiente de correlación entre nuestro escore y el número de publicaciones del autor fue de 0,0645.

El comportamiento de la nueva escala en relación a los niveles de evidencia se describe en la Figura 3.

Si analizamos el comportamiento de la nueva escala en relación a los niveles de evidencia observamos que existen diferencias estadísticamente significativas (ANOVA, $p = 0,0001$).

Al comparar el desempeño de la nueva escala en relación a los niveles de evidencia decidimos realizar un análisis de grupos extremos considerando los niveles de evidencia 1 y 4.

El promedio de puntuación de la nueva escala de los trabajos con nivel de evidencia 1 fue de $69,91 \pm 18,29$ en comparación con la puntuación de los trabajos con nivel de evidencia 4 que fue de $46,83 \pm 21,21$ ($p = 0,000$).

La mediana del escore en los artículos de nivel de evidencia 1 fue de 73 y en los artículos de nivel de evidencia 4 fue de 54.

Discusión

La escala MInCir Pr fue diseñada para la utilización en revisiones sistemáticas. El proceso de diseño y validación cumplió varios pasos metodológicos importantes que vale la pena recordar¹:

1. Extracción de posibles ítems con base en la experiencia de los autores y la literatura.
2. Revisión de los posibles ítems por un comité de expertos tratando de excluir aquellos que evaluaran los mismos aspectos y optimizando la traducción al castellano.
3. Aplicación del cuestionario inicial a 121 artículos de pronóstico.
4. Realización de análisis factorial para extraer ítems y dominios.
5. Cálculo de la consistencia interna.
6. Aplicación de la nueva escala y calificación con niveles de evidencia a los artículos descritos.

Los dominios de la escala son bastante lógicos en la evaluación de la calidad metodológica de estudios de pronóstico:

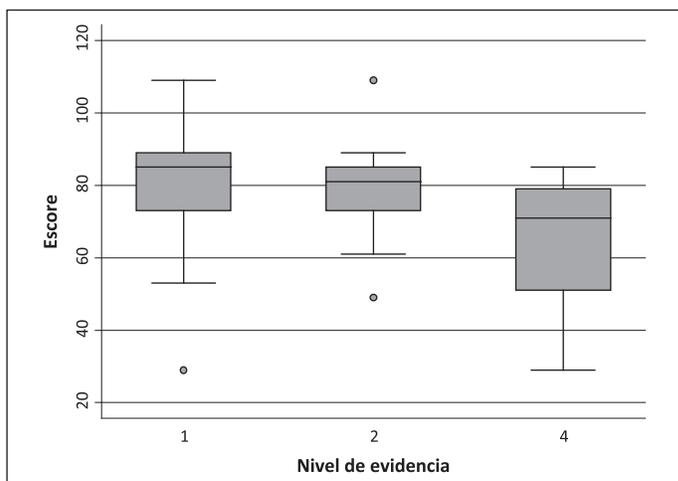


Figura 3. Puntuación escala MInCir Pr según niveles de evidencia. Escala MInCir Pr = escala de pronóstico MInCir. MInCir = metodología e investigación en cirugía. Escore = puntuación de la escala MInCir Pr.

1. Dominio 1: Variable de desenlace.
2. Dominio 2: Variable de exposición.
3. Dominio 3: Seguimiento.
4. Dominio 4: Análisis y conclusiones.
5. Dominio 5: Diseño y variables de confusión.
6. Dominio 6: Tamaño de la muestra.

Esto es importante ya que estos dominios que empíricamente parecían importantes en la evaluación de la calidad metodológica, fueron comprobados después de un proceso metodológico de diseño y validación; corrigiendo así una de las críticas realizadas a la escala inicial diseñada por el equipo de investigación para evaluar la calidad metodológica de estudios de terapia. Además, esta nueva escala cumple el criterio teórico de multidimensionalidad de la calidad metodológica.

Es importante destacar los coeficientes de confiabilidad interna (alfa de Cronbach) para los diferentes dominios que varían entre 0,95 y 0,75, mostrando una buena consistencia interna de los diferentes dominios.

En cuanto a la validación de la escala decidimos usar el nivel de evidencia para comparar nuestra escala debido a que es la herramienta más usada para evaluar calidad de evidencia. Sin embargo, somos conscientes que el nivel de evidencia es una herramienta básicamente monodimensional que tiene en el diseño de los estudios el enfoque de su evaluación^{5,6}.

Al revisar la puntuación de la escala según el diseño de los estudios; encontramos que sólo hay un artículo evaluado como meta-análisis; revisión sistemática y corte transversal.

La puntuación de la escala fue ligeramente mayor en las cohortes prospectivas en comparación con las

cohortes retrospectivas ($83,39 \pm 11,16$ y $81 \pm 9,03$ respectivamente); lo cual podría indicar que hay aspectos metodológicos diferentes al diseño que pueden ser evaluados por la escala y que explicarían esta poca diferencia en la puntuación.

Esto contrasta con la puntuación de las series de casos; que muestra una importante diferencia con la puntuación de la cohorte retrospectiva ($59,18 \pm 17,44$ y $81 \pm 9,03$ respectivamente).

La puntuación del escore según el nivel de evidencia se puede observar en la Tabla 3.

Para realizar la validación de grupos extremos comparamos el desempeño de la nueva escala en relación a los niveles de evidencia 2 y 4; ya que en estos grupos había mayor número de artículos (55 y 48 respectivamente). El promedio de puntuación de la nueva escala de los trabajos con nivel de evidencia 2 fue de $82,21 \pm 13,79$ en comparación con la puntuación de los trabajos con nivel de evidencia 4 que

Tabla 3. Puntuación escala MinCir Pr según nivel de evidencia

Nivel de evidencia	n	Escore Escala MInCir Pr (X ± DE)
1	1	73
2	55	82,21 ± 13,79
3	3	74,33 ± 22,03
4	48	64,97 ± 17,10
5	15	59,93 ± 11,65
Total	122	74,88 ± 16,99

X = promedio. DE = desviación estándar.

fue de $64,97 \pm 17,10$ ($p = 0,000$). Si comparamos la puntuación de la escala MinCir Pr en los trabajos de nivel de evidencia 1 y 2 con los trabajos de nivel de evidencia 4 y 5 las diferencias estadísticas se conservan ($p = 0,001$). Esto quiere decir que nuestra escala puede discriminar entre los artículos mejor y peor evaluados según los niveles de evidencia.

Esperamos que la escala MinCir Pr pueda ser útil para los profesionales que realicen revisiones sistemáticas; sobre todo en el mundo quirúrgico donde para diversas preguntas de investigación muchas veces no es posible obtener estudios con diseño que corresponda a un buen nivel de evidencia.

Conclusión

Se ha diseñado una escala válida y confiable para la evaluación de calidad metodológica de estudios de pronóstico (MinCir Pr) cumpliendo el criterio de multidimensionalidad de la calidad metodológica. Esta escala puede ser utilizada en la realización de revisiones sistemáticas.

Referencias

1. Losada H, Manterola C, Pineda V, Vial M, Sanhueza A, Grupo MinCir. Diseño de una escala para la evaluación de calidad metodológica de estudios de pronóstico. *Rev Chil Cir.* 2009;61:59-72.
2. Manterola C, Pineda V, Vial M, Losada H, Muñoz S. Revisión sistemática de la literatura. Propuesta metodológica para su realización. *Rev Chil Cir.* 2003;55:204-8.
3. Pineda V, Manterola C, Vial M, Losada H. ¿Cuál es la calidad metodológica de los artículos referentes a terapia publicados en la Revista Chilena de Cirugía? *Rev Chil Cir.* 2005;57:500-7.
4. www.cebm.net/levels_of_evidence.asp. Centre for Evidence-Based Medicine de Oxford. Levels of Evidence and Grades of Recommendation. 2007. Ref Type: Internet Communication.
5. Bornhöft G, Maxon-Bergemann S, Wolf U. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. *BMC Med Res Methodol.* 2006;11:56.
6. GRADE working group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.