



# REVISTA IBEROAMERICANA DE PSICOLOGÍA Y SALUD



Revista Oficial de la Federación Iberoamericana de Asociaciones de Psicología (FIAP)  
[Official Journal of the Latin-American Federation of Psychological Associations]

AVANCE ONLINE

## Generalización de la consistencia interna del *Historical-Clinical-Risk Management-20* (HCR-20) en población penitenciaria: Un meta-análisis

Pablo Prieto-Macías, Alejandro Guillén-Riquelme\* y Raúl Quevedo-Blasco

Centro de Investigación, Mente, Cerebro y Comportamiento (CIMCYC). Universidad de Granada, España

• Recibido: 23 - 03 - 2020 • Aceptado: 27 - 04 - 2020 • Avance online: 10 - 06 - 2020

**RESUMEN.** La *Historical-Clinical-Risk Management-20* (HCR-20) es un instrumento ampliamente usado para la evaluación del riesgo de violencia en muestra general y penitenciaria. Algunos estudios han demostrado una buena consistencia interna en muestra penitenciaria, pero no existe un meta-análisis que proporcione una visión global de todos estos resultados. El presente meta-análisis tiene el objetivo examinar los alfas medios de este instrumento en instituciones penitenciarias y analizar el efecto de los moderadores en la fiabilidad. Los datos fueron recogidos a través de las bases datos de *Web of Science*, *Scopus*, *ProQuest*, *Psycodoc* y *Cochrane*. Se incluyeron 25 estudios que examinaban la consistencia interna en esta herramienta, reuniendo un total de 2.517 participantes de diferentes países con una media de edad de 29,7 años. Se demostró que el HCR-20 posee una óptima consistencia interna, influenciada por diversos moderadores. La principal conclusión de este meta-análisis es, por tanto, que el HCR-20 es un instrumento fiable para este tipo de muestra. Esta información es importante por la repercusión que tiene este instrumento en numerosas decisiones trascendentales.

**PALABRAS CLAVE:** HCR-20, Evaluación riesgo violencia, Prisión, Consistencia interna, Meta-análisis.

### Generalization of the internal consistency of the *Historical-Clinical-Risk Management-20* (HCR-20) in prison sample: A meta-analysis

**ABSTRACT.** The *Historical-Clinical-Risk Management-20* (HCR-20) is a widely used instrument for assessing risk of violence in civil and correctional sample. Previous studies have demonstrated good reliability in correctional sample, but it does not exist a meta-analysis that provides an overview vision for all these results. This meta-analysis has the aim to examine the average alphas of this instrument in prison and analyze the effect of moderators on reliability. Data were collected through formal and informal sources. Twenty five studies, that examined the internal consistency of this tool, were included, collecting a total of 2,517 participants from different countries with an average age of 29.7 years. It was demonstrated that HCR-20 has an optimal internal consistency, influenced by various moderators. The main conclusion of this meta-analysis is, therefore, that HCR-20 is a reliable instrument for this type of sample. This information is important because of the impact that this instrument has on numerous transcendental decisions.

**KEYWORDS:** HCR-20, Violence risk evaluation, Jail, internal consistency, Meta-analysis.

La violencia es un aspecto de especial importancia, analizada incluso desde las propuestas de la Clasificación Internacional de Enfermedades, versión 11 [CIE-11] (Heyman et al., 2018), aunque existe una falta de consenso

en la definición, que muchas veces dificulta la comparación de resultados y el establecimiento de políticas adecuadas (Loinaz, 2017). Son múltiples los estudios que analizan los ámbitos donde se producen comportamientos violentos/agresivos (Álvarez-García, Núñez, García, y Barreiro-Collazo, 2018; Gallego, Novo, Fariña, y Arce, 2019; Juarros-Basterretxea, Herrero, Fernández-Suárez, Pérez, y Rodríguez-Díaz, 2018), los posibles factores asociados (Férriz, Sobral, y Gómez-Fraguela, 2018; Fernández-

\*Correspondencia: Alejandro Guillén-Riquelme.  
Centro de Investigación, Mente, Cerebro y Comportamiento (CIMCYC). Universidad de Granada, España  
C.P: 18011, Granada, España.  
E-mail: agr@ugr.es

Suárez, Pérez, Herrero, Juarros-Basterretxea, y Rodríguez-Díaz, 2018; Juarros-Basterretxea et al., 2018; Juarros-Basterretxea, Overall, Herrero, y Rodríguez-Díaz, 2019; Martín y Moral, 2019), las consecuencias en distintos contextos (Morales-Toro, Guillén-Riquelme, y Quevedo-Blasco, 2019), e incluso, las posibles intervenciones e impacto de las mismas (Cantos, Kosson, Goldstein, y O'Leary, 2019; Yanez-Peñúñuri, Martínez-Gómez, y Rey-Anacona, 2019). En la actualidad, la validación de instrumentos es un objetivo prioritario para cualquier evaluación, y en concreto, para analizar los distintos tipos de violencia (v.gr., Contreras, Bustos-Navarrete, y Cano-Lozano, 2019; Santirso, Martín-Fernández, Lila, Gracia, y Terreros, 2018). De forma concreta, la evaluación del riesgo de violencia en el ámbito penitenciario resulta fundamental debido a su gran influencia en algunas resoluciones del sistema penal o la importancia que tiene en la intervención preventiva posterior (Geraghty, 2015). Esto es imprescindible ya que alrededor de un 1,4 millones de personas murieron en 2016 por causas violentas (World Health Organization, 2016), representando el 2,5% de las muertes mundiales (Mathers, Stevens, Hogan, Mahanani, y Ho, 2017).

La escala *Historical-Clinical-Risk Management-20* (HCR-20; Webster, Eaves, Douglas, y Wintrup, 1995) es uno de los instrumentos más usados para la evaluación del riesgo de violencia en las instituciones penitenciarias, además de una de las herramientas sistemáticas predictivas más exitosas para medir el alcance del peligro y enfocar con ello la gestión preventiva (Arai, Takano, Nagata, y Hirabayashi, 2017; Arbach y Andrés-Pueyo, 2007; Sada et al., 2016), incluso en muestra psiquiátrica (Strand y Selenius, 2019). El HCR-20 presenta cuatro versiones (la versión 1, 2, 3 y FAM). Para la versión 1 hay evidencias de una adecuada fiabilidad en Suecia (Belfrage, 1998), Canadá (Mills y Gray, 2013) o EE.UU. (Douglas et al., 1998). En la versión 2, estos buenos resultados se extienden a países como Alemania (Dahle, 2006), Bélgica (Claix y Pham, 2004), Brasil (Telles, Day, Folino, y Taborda, 2009), Canadá (Grimbos et al., 2016), EE.UU. (Neal, Miller, y Shealky, 2015), Francia (Pham,

Chevrier, Nioche, Ducro, y Réveillère, 2005), Japón (Arai et al., 2017), o Lituania (Sellbom, Laurinavičius, Ustinavičiūtė, y Laurinaitytė, 2018). Asimismo, los estudios que demuestran buenas propiedades psicométricas en la versión 3 del HCR-20 proceden de países como Reino Unido (Geraghty, 2015) o Noruega (Bjørkly, Eidhammer, y Selmer, 2014). La adaptación FAM (Female Additional Manual) de esta misma versión, es la que muestra una más baja fiabilidad (Geraghty, 2015). A pesar de los avances, se ha demostrado que la fiabilidad y validez cambian si se lleva a cabo, o no, en entornos controlados (Neal et al., 2015). Además, no existe un meta-análisis sobre la fiabilidad del HCR-20. Otros aspectos importantes son el bajo número de participantes que tienen los estudios dicha temática, y la gran variabilidad que existe entre los valores de alfa en el HCR-20, que varían desde 0,58 (Geraghty, 2015) hasta 0,95 (Belfrage, 1998).

Atendiendo a todo lo anterior, el objetivo del presente estudio es establecer la fiabilidad media del HCR-20 en diversos estudios con muestras penitenciarias, es decir, la generalización de la consistencia interna. Los objetivos específicos son: a) estimar la consistencia interna media del HCR-20 y sus subescalas, y b) realizar un análisis de los moderadores de la consistencia interna del HCR-20 y sus subescalas. De acuerdo con las investigaciones ya mencionadas y los objetivos planteados, se establecen tres hipótesis: (a) el tamaño del efecto de la consistencia interna del HCR-20 global tendrá una puntuación superior a 0,7, que es el punto de corte para el que se considere un nivel adecuado para el alfa (Muñiz, 2018; Oviedo y Campo-Arias, 2005); (b) los resultados del alfa medio de las tres subescalas también superarán el punto de corte de 0,7; y (c) la mayoría de los moderadores con más efecto serán aquellos que se refieran a características de la muestra.

## MÉTODO

### • PROCEDIMIENTO DE BÚSQUEDA DE ESTUDIOS

Se realizó la búsqueda mediante: *Web of Science* (Colección principal), *Scopus*, *Psycodoc*, *Cochrane* y *ProQuest* (incluyendo *PsycINFO*, *Sociology Collection*, *Medline*, *Applied Social*

Sciences Index & Abstracts, Criminology Collection, Psychology Database, Criminal Justice Database, ProQuest Dissertations & Theses Global, PsycARTICLES, International Bibliography of the Social Sciences, Health & Medical Collection, Nursing & Allied Health Database, PsycEXTRA, y PsycTESTS).

Se buscó en los campos de "título, palabras claves y resumen" la siguiente fórmula: (jail\* OR prison\* OR "penal institution\*" OR "prison house\*" OR correctional OR gaol\* OR prisoner\* OR inmate\* OR imprison\* OR sentence\* OR remand\* OR criminal\* OR \*offender\*) AND (HCR OR "HCR-20" OR "assessing risk for violence"

OR "historical-clinical-risk management" OR "historical-clinical-risk management-20" OR "historical clinical risk management\*"), adaptada a cada buscador/base de datos. El rango abarcó desde el año de publicación de la primera versión del HCR-20 (1995) hasta el 31 de diciembre de 2018 (año cerrado). Solo se incluyeron aquellos estudios cuyo idioma de publicación o traducción fuera inglés o español. Se reunió un total de 1.194 artículos. Después de aplicar todos los criterios de inclusión, 19 artículos fueron incluidos, con un total de 25 muestras. El diagrama de flujo PRISMA del proceso de cribado de la literatura (ver Urrutia y Bonfill, 2010) se presenta en la Figura 1.

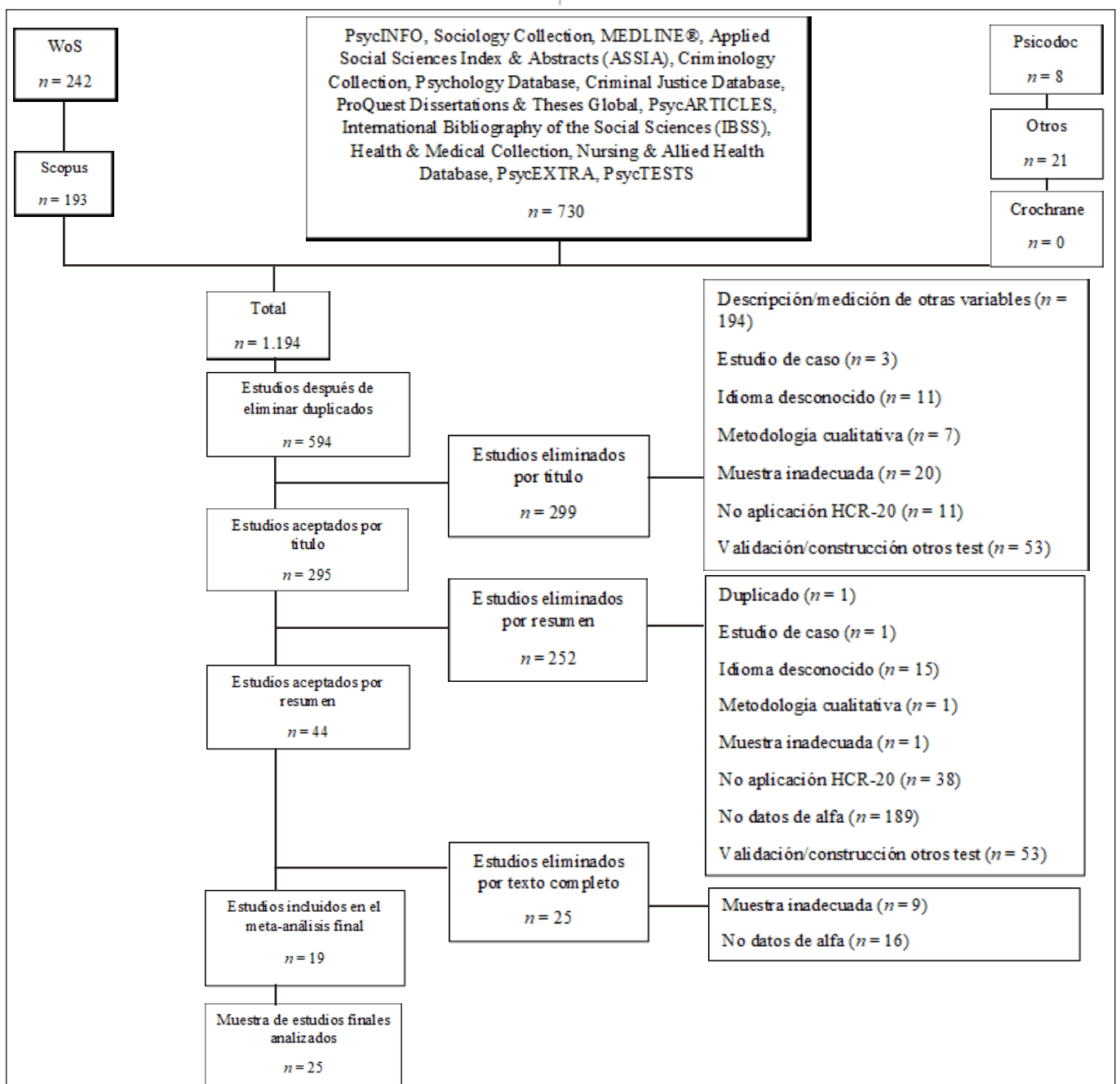


Figura 1: Diagrama de flujo del proceso de cribado

## • CRITERIOS DE SELECCIÓN DE ESTUDIOS

Los estudios tuvieron que tener los siguientes criterios para ser incluidos: (a) mayores de 18 años de ambos sexos internos en prisión, que lo hayan estado anteriormente, o que se encontraran reclusos en algún tipo de institución penitenciaria (como hospitales psiquiátricos penitenciarios u hospitales psiquiátricos forenses); (b) que la evaluación implicara la aplicación del HCR-20 en cualquiera de sus versiones (versión 1, 2, 3 o FAM); (c) estudios que proporcionen información del alfa de Cronbach y (d) trabajos experimentales o cuasiexperimentales siempre que incluyeran datos pretratamiento.

## • CODIFICACIÓN DE VARIABLES MODERADORAS

De cada estudio se codificaron los siguientes datos: (a) el año de publicación; (b) país de procedencia de la muestra; (c) versión del HCR-20; (d) número de participantes de la muestra; (e) media y desviación típica (DT) de la edad; (f) porcentaje de hombres; (g) institución donde se realiza el estudio; (h) si la muestra incluye pacientes con trastornos mentales; (k) porcentaje de cada trastorno; (l) alfa global del HCR-20 y de sus subescalas; (m) medias y DT, del global y de sus subescalas y (n) calidad metodológica del estudio.

## • FIABILIDAD DE LA CODIFICACIÓN

Todas las variables de todos los estudios se codificaron por dos jueces independientes. En caso de discrepancia se solucionó por consenso. Para las variables categóricas se empleó la Kappa de Cohen, donde se obtuvieron valores entre 0,71 y 1. Para las variables continuas se empleó la correlación intraclass entre 0,93 y 1. Además, para los diferentes alfas recogidos se calculó la Kappa, obteniendo un acuerdo perfecto para el alfa global, así como para las escalas H y R, mientras que en la escala C se obtuvo un valor de 0,93. Se puede afirmar que los valores de ambos índices se sitúan en niveles muy altos.

## • ANÁLISIS ESTADÍSTICOS

Se asumió un modelo de efectos aleatorios (Borenstein, Hedges, Higgins, y Rothstein, 2010). Los coeficientes de consistencia interna medios, se transformaron a puntuaciones T (Sánchez-Meca y López-Pina, 2008). A continuación, se estimó la fiabilidad media, ponderando estas puntuaciones T por la inversa de las varianzas muestrales. Por último, se retransformaron los resultados a valores alfa. El análisis de los moderadores se realizó a través de un meta-ANOVA para estudiar las fuentes de variabilidad en el caso de variables categóricas, y mediante la meta-regresión en el caso de variables continuas. La estimación de la homogeneidad se evaluó a través de los estadísticos  $Q$  e  $I^2$ . El posible sesgo de publicación, se analizó mediante el Test Egger y del Test de Rosenthal.

Para la interpretación de la magnitud del tamaño del efecto se tomó la Probabilidad de Superioridad del Tamaño del Efecto (PSTE) que informa de la superioridad del tamaño observado en relación a todos los posibles mayores de 0,50 (umbral mínimo de fiabilidad; < 0,50 fiabilidad inaceptable) y el Índice de Incremento del Efecto (IIE) que, adaptado a este caso, indica el incremento del efecto sobre el criterio de fiabilidad (valor de prueba), esto es, el valor para que un instrumento sea fiable como alfa 0,71 (50% de la varianza explicada) (Redondo et al., 2019).

Todos los análisis se realizaron con el programa estadístico *R*, mediante el paquete *metafor* (Viechtbauer, 2010).

## RESULTADOS

Se obtuvieron 19 estudios que incluían 25 muestras, con un total de 2.517 sujetos. La edad media fue de 29,7 años ( $DT = 9,02$ ). Un 80% de las muestras incluían presos con trastornos mentales. Los datos descriptivos de la muestra y de los estudios se resumen en la Tabla 1. En la Tabla 2 se presenta el tamaño de efecto medio para la puntuación global. La fiabilidad es alta y se observa una elevada heterogeneidad ( $I^2 = 84,98\%$ ).

Tabla 1  
Características descriptivas de la muestra y de los estudios

Edad media (DT)	29,7 (9,02)
Sexo (%)	
Hombres	1.841 (73,14%)
Mujeres	676 (26,86%)
País de la muestra (%) <sup>a</sup>	
Alemania	1 (4%)
Bélgica	1 (4%)
Brasil	1 (4%)
Canadá	7 (28%)
EE.UU	3 (12%)
Francia	1 (4%)
Japón	1 (4%)
Lituania	1 (4%)
Noruega	2 (8%)
Reino unido	4 (16%)
Suecia	2 (8%)
Sin especificar	1 (4%)
Trastornos mentales primarios <sup>b</sup>	
Orgánicos	1,89%
Debidos al consumo de sustancias psicótropas	21,32%
Esquizofrenia, trastornos esquizotípico y trastornos de ideas delirantes	56,13%
Del humor (afectivos)	7,29%
Neuróticos y somatomorfos	0,63%
Asociados a disfunciones fisiológicas y factores somáticos	0,85%
De la personalidad	27,55%
Retraso mental	4,69%
Del desarrollo	0,05%
Institución penitenciaria <sup>a</sup>	
Prisión	7 (28%)
Hospital psiquiátrico	17 (68%)
Ambos	1 (4%)
Versión del HCR-20 (%) <sup>a</sup>	
Versión 1	3 (12%)
Versión 2	17 (68%)
Versión 3	3 (12%)
Versión 2 FAM	1 (4%)
Versión 3 FAM	1 (4%)
<sup>a</sup> Datos y porcentajes procedentes del total de estudios ( $k = 25$ ), no de la muestra de participantes.	
<sup>b</sup> Porcentaje procedentes del total de participantes de estudios con muestra con trastornos mentales ( $n = 1.857$ ).	

*Tabla 2*  
Correlaciones medias para la escala del HCR-20 y subescalas analizadas

Escala	<i>k</i>	Correlaciones Medias	<i>EE</i>	IC inf.	IC sup.	<i>Z</i>	<i>Q</i>	<i>I</i> <sup>2</sup>
HCR global	20	0,81	0,02	0,76	0,85	32,3***	126,48***	84,98
Escala H	16	0,83	0,03	0,74	0,88	15,07***	429,04***	96,50
Escala C	16	0,75	0,03	0,66	0,82	21,41***	184,92***	91,89
Escala R	15	0,77	0,04	0,68	0,84	13,66***	414,97***	96,63

Nota. *k* = número de estudios; *EE* = error estándar; IC inf. = intervalo de confianza inferior; IC sup. = intervalo de confianza superior; los grados de libertad del estadístico *Q* son *k* - 1. \**p* < 0,05; \*\**p* < 0,01; \*\*\**p* < 0,001

*Tabla 3*  
Consistencia interna media del HCR-20 y de sus subescalas por país

País	Alfa medio	<i>EE</i>	IC inf.	IC sup.
HCR global				
Alemania	0,84***	0,25	0,79	0,88
Bélgica	0,74*	0,04	0,63	0,82
Brasil	0,79	0,06	0,64	0,88
Canadá	0,84	0,03	0,80	0,87
EE.UU	0,72**	0,03	0,63	0,78
Francia	0,84	0,04	0,76	0,90
Japón	0,82	0,04	0,75	0,87
Lituania	0,86	0,04	0,80	0,90
Noruega	0,84	0,06	0,70	0,92
Reino Unido	0,68***	0,03	0,62	0,74
Suecia	0,95***	0,04	0,91	0,97
Escala H				
Bélgica	0,61	0,16	-0,21	0,93
Brasil	0,63	0,24	-0,22	0,95
Canadá	0,82	0,24	0,56	0,94
EE.UU	0,87	0,19	0,67	0,97
Japón	0,76	0,19	0,13	0,97
Lituania	0,64	0,23	-0,14	0,94
Noruega	0,85	0,24	0,31	0,99
Suecia	0,96	0,24	0,69	0,99
Escala C				
Bélgica	0,47	0,13	-0,23	0,83
Brasil	0,51*	0,18	-0,25	0,87
Canadá	0,74	0,19	0,52	0,88
EE.UU	0,84	0,15	0,67	0,93
Japón	0,72	0,15	0,24	0,93
Lituania	0,60	0,18	0,01	0,89
Noruega	0,59	0,18	-0,15	0,91
Suecia	0,89	0,20	0,59	0,98



Tabla 3 (Continuación)  
Consistencia interna media del HCR-20 y de sus subescalas por país

País	Alfa medio	EE	IC inf.	IC sup.
Escala R				
Bélgica	0,54	0,21	-0,70	0,95
Brasil	0,69	0,30	-0,36	0,98
Canadá	0,64	0,30	-0,01	0,92
EE.UU	0,83	0,26	0,51	0,97
Japón	0,77	0,24	-0,09	0,99
Lituania	0,85	0,30	0,14	0,99
Noruega	0,81	0,30	-0,03	0,99
Suecia	0,85	0,30	0,13	0,99

Nota. EE= error estándar; IC inf.= intervalo de confianza inferior; IC sup.= intervalo de confianza superior; los grados de libertad del estadístico Q son  $k - 1$ . \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ .

Se obtuvo un tamaño del efecto medio del alfa de la escala global con un valor adecuado próximo a 0,8. En cuanto a las subescalas, todas obtienen valores entre 0,7 y 0,9, el intervalo recomendado para cualquier instrumento. Los valores del alfa medio del HCR-20 y de sus subescalas son significativos y medio-altos o altos. Los forest plot se pueden observar en las Figuras 2, 3, 4 y 5.

### •ANÁLISIS DE MODERADORES CATEGÓRICOS

Debido a la alta heterogeneidad, se procedió al análisis de moderadores. En primer lugar, el país resultó significativo ( $Q(10) = 73,43$ ;  $p < 0,001$ ). Se puede ver con más detalle en la Tabla 3. En cuanto a las subescalas del HCR-20, el único resultado no significativo se presentó en la escala C para la población brasileña, mucho menos fiable que el resto.

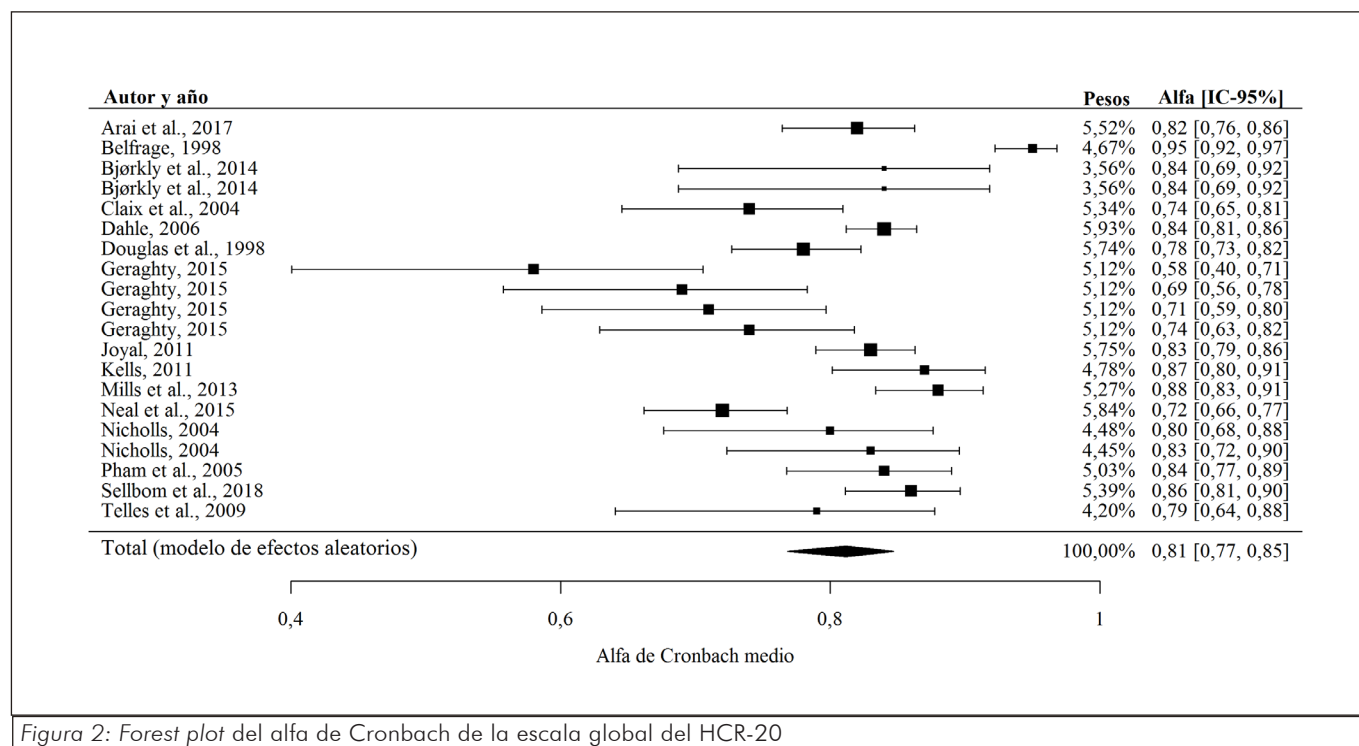


Figura 2: Forest plot del alfa de Cronbach de la escala global del HCR-20

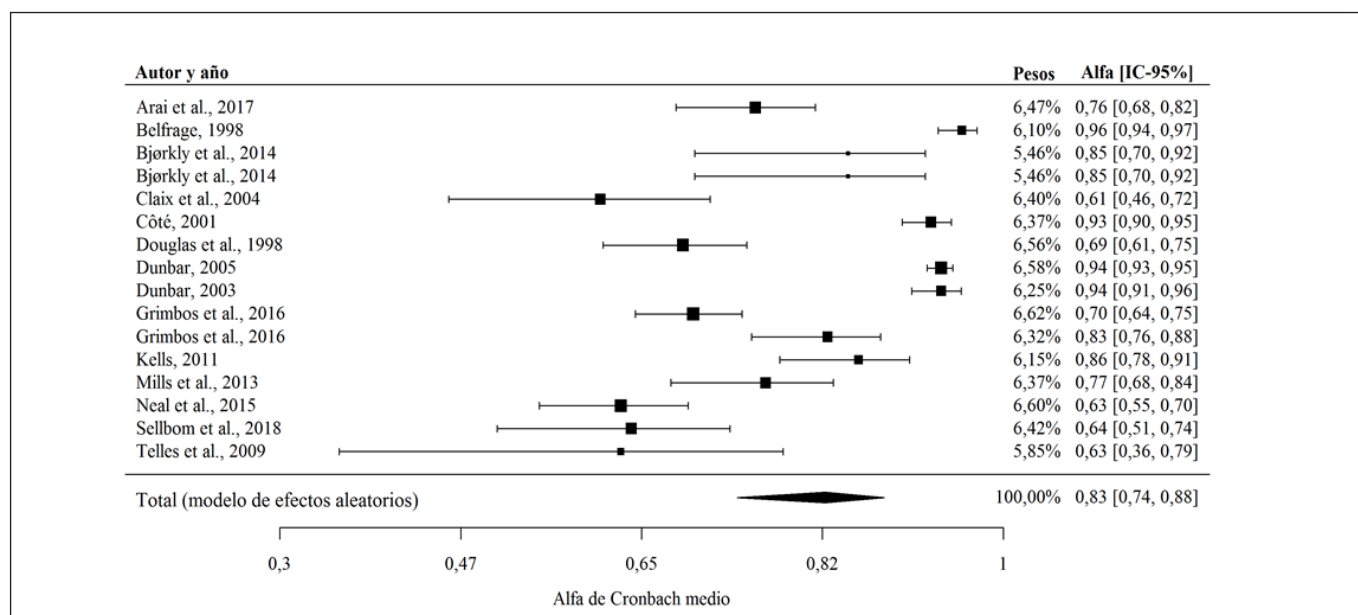


Figura 3: Forest plot del alfa de Cronbach de la escala H del HCR-20

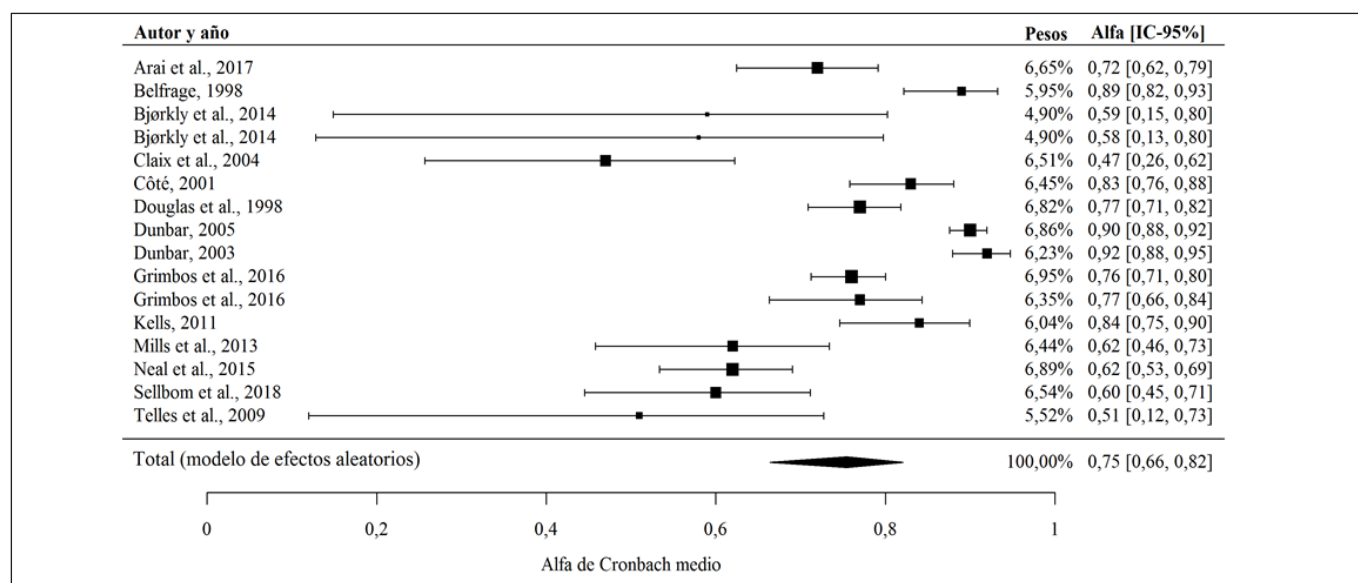


Figura 4: Forest plot del alfa de Cronbach de la escala C del HCR-20

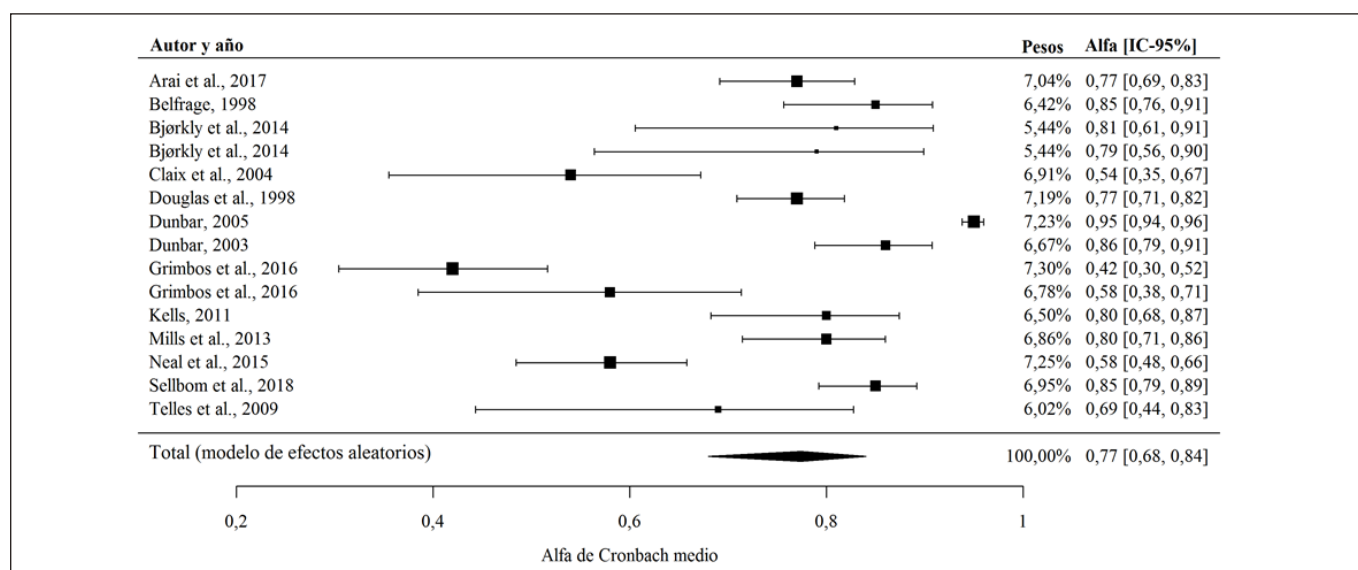


Figura 5: Forest plot del alfa de Cronbach de la escala R del HCR-20



Tabla 4  
Consistencia interna media del HCR-20 y de sus subescalas por tipo de institución

Institución	Alfa medio	EE	IC inf.	IC sup.
<b>HCR global</b>				
Ambas	0,83***	0,07	0,65	0,93
Hospital psi. penitenciario	0,78	0,07	0,73	0,82
Prisión	0,85	0,08	0,79	0,90
<b>Escala H</b>				
Hospital psi. penitenciario	0,79	0,06	0,69	0,87
Prisión	0,85***	0,07	0,73	0,93
<b>Escala C</b>				
Hospital psi. penitenciario	0,71	0,04	0,62	0,79
Prisión	0,81***	0,05	0,70	0,88
<b>Escala R</b>				
Hospital psi. penitenciario	0,69***	0,05	0,58	0,78
Prisión	0,86**	0,06	0,77	0,92

Nota. EE= error estándar; IC inf. = intervalo de confianza inferior; IC sup. = intervalo de confianza superior; los grados de libertad del estadístico Q son  $k - 1$ . \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ .

Tabla 5  
Consistencia interna media del HCR-20 y de sus subescalas por versión del HCR

Institución	Alfa medio	EE	IC inf.	IC sup.
<b>HCR global</b>				
Versión 1	0,88***	0,04	0,81	0,93
Versión 2	0,80*	0,04	0,75	0,84
Versión 2 FAM	0,71*	0,08	0,45	0,86
Versión 3	0,78	0,06	0,66	0,87
Versión 3 FAM	0,74	0,08	0,50	0,88
<b>Escala H</b>				
Versión 1	0,83***	0,09	0,62	0,95
Versión 2	0,80	0,10	0,69	0,88
Versión 3	0,85	0,14	0,55	0,97
<b>Escala C</b>				
Versión 1	0,77***	0,06	0,59	0,89
Versión 2	0,75	0,07	0,66	0,82
Versión 3	0,58	0,12	0,14	0,83
<b>Escala R</b>				
Versión 1	0,80***	0,10	0,51	0,95
Versión 2	0,73	0,12	0,57	0,85
Versión 3	0,80	0,17	0,37	0,96

Nota. EE = error estándar; IC inf. = intervalo de confianza inferior; IC sup. = intervalo de confianza superior; los grados de libertad del estadístico Q son  $k - 1$ . \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ .

En la Tabla 4 se recoge el uso del tipo de centro como moderador. Los resultados fueron significativos para la escala global cuando mezclaban ambas poblaciones, y para las subescalas la única institución significativa para las tres variables fue la prisión. Por último, la consistencia se calculó para cada versión, siendo la "1" la única significativa para todas las escalas (ver Tabla 5).

### •ANÁLISIS DE MODERADORES CONTINUOS

Para el moderador año de publicación, los resultados fueron significativos para la escala H y para la escala C, aunque todas las betas eran inferiores a 0,01. En cuanto a la influencia de la calidad metodológica, ninguna de las escalas resultó afectada por este moderador. Para la variable edad media de la muestra, únicamente resultó significativa la escala C, de nuevo con un beta de 0,01. En cuanto al porcentaje de hombres como moderador, ninguna subescala obtuvo la suficiente significación, excepto la escala global ( $\beta = 0,001$ ).

La influencia de la puntuación media del HCR-20 y sus subescalas en los valores de alfa no resultó significativa para ninguna de las escalas presentadas; además todas las puntuaciones se mostraron homogéneas. En cuanto a la desviación típica de las puntuaciones medias del HCR-20, utilizada para comprobar si la variabilidad de las puntuaciones promedio de los distintos estudios afectaba a los resultados del alfa, se demostró que únicamente la escala global del HCR-20 ( $\beta = -0,02$ ;  $p = 0,04$ ) y la subescala R ( $\beta = -0,20$ ;  $p = 0,003$ ) se mostraba significativa en este moderador. Como ocurría con la significación, solo la escala global HCR-20 ( $Q = 3,94$ ;  $p = 0,04$ ) y la subescala R ( $QR = 8,56$ ;  $p = 0,003$ ) eran homogéneas en sus puntuaciones.

Seguidamente, se calculó la magnitud del efecto observado. Para la subescala H fue superior al 94,9% de todos los efectos posibles aceptables e incrementa el 6,60% (IIE = 0,066) la estimación de la fiabilidad con respecto al umbral de fiabilidad. La magnitud del efecto observado para la subescala C fue superior al 89% de todos los efectos posibles aceptables e incrementa el 2,91% (IIE = 0,066) la fiabilidad con respecto al umbral de fiabilidad. Para

la subescala R es superior al 91,1% de todos los efectos posibles aceptables e incrementa el 3,95% (IIE = 0,0395).

### •ANÁLISIS ADICIONALES PARA EL HCR-20 Y SUS SUBESCALAS

Los resultados procedentes del *Test Egger* indican que la única escala que estaba afectada por el sesgo de publicación es la escala C ( $Z = 2,22$ ;  $p = 0,02$ ). Sin embargo, en el *Test Rosenthal* todas las escalas obtuvieron un valor significativo superior a 25.000 estudios ( $p < 0,001$ ). Por lo que se demuestra que, a nivel global, el presente meta-análisis está libre de sesgo de publicación.

## DISCUSIÓN

Las revisiones en el ámbito Jurídico-Forense son cada vez más necesarias (Quevedo-Blasco, Guillén-Riquelme, y Buela-Casal, 2019) e importantes para el abordaje de problemáticas con gran impacto en la actualidad (v.gr., Cano-Montalbán y Quevedo-Blasco, 2018; Selaya-Berodia, Quevedo-Blasco, y Neufeld, 2018). Los Meta-análisis sobre las propiedades psicométricas de instrumentos, con el fin de analizar la idoneidad de su aplicación (v.gr., Salvador, Arce, Rodríguez-Díaz, y Seijo, 2017), son fundamentales para el avance en la evaluación. El presente estudio es el primero en examinar la fiabilidad del HCR-20 a través de un meta-análisis. Se observan unos excelentes índices medios del alfa del HCR-20 y de sus subescalas. Los valores de alfa de la escala global y de sus subescalas varían en todo momento dentro de un intervalo adecuado, siendo la escala C la que presenta un valor más bajo (Cronbach y Meehl, 1955). Además, para todas las subescalas, la magnitud del efecto observado es alta y muy por encima de los valores aceptables. Por tanto, se podría concluir que el presente instrumento en general es fiable para la evaluación del riesgo de violencia en muestra penitenciaria, confirmándose la primera y segunda hipótesis del presente meta-análisis. Todos estos resultados son consistentes con algunos de los estudios más recientes (Arai et al., 2017; Grimbois et al., 2016; Neal et al., 2015; Sellbom et al., 2018).

Para el segundo objetivo propuesto, los hallazgos muestran que los moderadores categóricos que se revelan explicativos de la variabilidad son la versión del HCR-20 y el tipo institución donde se realizaba el estudio. A pesar de que la variable país no actúa como moderadora de la escala global, sí que explica la variabilidad en el caso de todas las subescalas. La explicación a este hecho no está clara, aunque una posible explicación podría ser que, al ser distinto el nivel de homogeneidad para las subescalas, la unión de estas en una puntuación total provocaría la heterogeneidad de la escala global, es decir, esta última representaría en realidad la propia variabilidad entre dichas subescalas al unirlas. Esto se observa en la diferencias de fiabilidad para versiones según el país (v.gr., Bjørkly, Eidhammer, y Selmer, 2014; Geraghty, 2015).

Para la variable tipo de institución, la escala global muestra su mayor índice de fiabilidad para la muestra procedente de la prisión en comparación con las otras poblaciones, al igual que ocurre en las subescalas. Consecuentemente, se puede concluir que, aunque este instrumento está diseñado con el propósito de aplicarse en una variedad de entornos forenses (Webster et al., 1995), es más fiable su aplicación en entornos penitenciarios puros. Coincidiendo con Neal et al. (2015), quienes afirmaban que la fiabilidad varía cuando los estudios se llevan a cabo en entornos muy controlados.

Para la variable versión del HCR-20, la escala global presenta su mayor coeficiente alfa en su versión 1, pudiendo indicar que la mejora de muchos aspectos en las sucesivas versiones ha supuesto una cierta pérdida de fiabilidad en la prueba. Por otro lado, la versión que se muestra más inconsistente es la versión 2. Estos resultados coinciden con los del estudio de Geraghty (2015), donde la versión 3 era más fiable que la versión 2 original y que la adaptación a mujeres (FAM).

En el caso de los moderadores continuos, los resultados muestran que la fiabilidad del HCR-20 global está significativamente mediada por las variables año de publicación, calidad metodológica, edad media y puntuación media del instrumento, para la escala global. La consistencia interna de este instrumento es mayor

cuanto menor es el año de publicación, la calidad metodológica, y la puntuación media de la escala, y cuanto más joven es la muestra. El hecho de que la consistencia interna sea menor conforme aumenta los años de publicación puede estar informando de tres hechos: El primero, podría ser otra señal (como pasaba con las versiones del HCR-20 anteriormente comentada) de que la mejora en las definiciones e información de la herramienta, la adecuación a los desarrollos en el área, y la mejora del carácter idiosincrático y de la toma de decisiones que se han ido desarrollando en el instrumento a lo largo de los años, han ido en detrimento de la fiabilidad de la prueba; segundo, que el constructo que pretende medir la prueba ha perdido validez; o tercero, que la conceptualización teórica en la que se basa dicho constructo no es tan correcta en la actualidad (Oviedo y Campo-Arias, 2005). Es importante remarcar que estos resultados son muestra de que las conclusiones deben tomarse con mayor precaución cuando la muestra obtenga puntuaciones muy altas en su nivel de riesgo y cuanto más edad tenga esta.

Los hallazgos presentados en este estudio son importantes para la práctica psicológica, principalmente porque el HCR-20 es uno de los instrumentos más utilizados para la evaluación del riesgo de violencia, y el hecho de demostrar su fiabilidad, supone una evidencia robusta para certificar que las medidas están libres de errores y que el constructo de la escala es válido. El uso de escalas fiables es fundamental cuando el instrumento se utiliza para tomar decisiones que pueden tener trascendentales consecuencias para el sujeto evaluado y para la población en general (Geraghty, 2015).

Los elementos que pueden ser fuente de sesgo potencial por su amenaza a la validez interna serían: la diferencia entre el número de mujeres y hombres (aunque esto podría ser un mero reflejo de una realidad social donde el porcentaje de hombres presos es mayor al de mujeres); la falta de datos en países no europeos; la diferencia entre la muestra procedente de hospitales psiquiátricos penitenciarios y la de prisiones (ya que la mayoría de los estudios incluidos en este trabajo utilizaron la versión 2 del HCR-20); y que la escala C puede verse afectada por el sesgo de publicación.

En cuanto a las limitaciones, la principal es el bajo número de estudios encontrados, debido a la poca investigación dirigida este coeficiente en instituciones penitenciarias. La segunda es que, al ser el primer meta-análisis dirigido a la consistencia del HCR-20, no existen otros estudios de este tipo con los que comparar los resultados. Y por último, la posible influencia de otros moderadores no tenidos en cuenta que pudieran explicar más detalladamente la variabilidad de los resultados.

En futuras investigaciones podría examinarse la fiabilidad de esta herramienta usando otros estimadores, profundizar en el estudio de la consistencia interna en muestras penitenciarias de otros países, como España; o el análisis comparativo de la fiabilidad del HCR-20 en hospitales psiquiátricos y prisiones.

• **Conflicto de intereses.**

Los autores declaran no tener ningún conflicto de intereses.

## REFERENCIAS

Las referencias marcadas con un asterisco (\*) son las incluidas en el meta-análisis

- Álvarez-García, D., Núñez, J.C., García, T., y Barreiro-Collazo, A. (2018). Individual, family, and community predictors of cyber-aggression among adolescents. *European Journal of Psychology Applied to Legal Context*, 10(2), 79-88. <http://dx.doi.org/10.5093/ejpalc2018a8>
- \*Arai, K., Takano, A., Nagata, T., e Hirabayashi, N. (2017). Predictive accuracy of the Historical-Clinical-Risk Management-20 for violence in forensic psychiatric wards in Japan. *Criminal Behaviour and Mental Health*, 27(5), 409-420. <https://doi.org/10.1002/cbm.2007>
- Arbach, K., y Andrés-Pueyo, A. (2007). Valoración del riesgo de violencia en enfermos mentales con el HCR-20. *Papeles del Psicólogo*, 28(3), 174-186. Recuperado de <http://www.papelesdelpsicologo.es/pdf/1501.pdf>
- \*Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: Integrating research and clinical practice. *Journal of Forensic Psychiatry*, 9(2), 328-338. <https://doi.org/10.1080/09585189808402200>
- \*Bjørkly, S., Eidhammer, G., y Selmer, L. E. (2014). Concurrent Validity and Clinical Utility of the HCR-20V3 Compared With the HCR-20 in Forensic Mental Health Nursing: Similar Tools but Improved Method. *Journal of Forensic Nursing*, 10(4), 234-242. <https://doi:10.1097/JFN.0000000000000047>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., y Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 1(2), 97-111. <https://doi.org/10.1002/jrsm.12>
- Cano-Montalbán, I., y Quevedo-Blasco, R. (2018). Sociodemographic Variables Most Associated with Suicidal Behaviour and Suicide Methods in Europe and America. A Systematic Review. *European Journal of Psychology Applied to Legal Context*, 10(1) 15-25. <http://dx.doi.org/10.5093/ejpalc2018a2>
- Cantos, A. L., Kosson, D. S., Goldstein, D. A., y O'Leary, K. D. (2019). Treatment impact on recidivism of family only vs. generally violent partner violence perpetrators. *International Journal of Clinical and Health Psychology*, 19(3), 171-180. <https://doi.org/10.1016/j.ijchp.2019.05.002>
- \*Claix, A., y Pham, T. H. (2004). Evaluation of the HCR-20 Violence Risk Assessment Scheme in a Belgian forensic population. *L'Encéphale*, 30(5), 447-453. [https://doi.org/10.1016/S0013-7006\(04\)95459-0](https://doi.org/10.1016/S0013-7006(04)95459-0)
- Contreras, L., Bustos-Navarrete, C., y Cano-Lozano, M. C. (2019). Child-to-parent Violence Questionnaire (CPV-Q): Validation among Spanish adolescents. *International Journal of Clinical and Health Psychology*, 19(1), 67-74. <https://doi.org/10.1016/j.ijchp.2018.09.001>
- \*Côté, G., Hodgins, S., y Daigle, M. (Abril, 2001). *Violent behaviour, PCL-R and HCR-20 among involuntary inpatients, forensic patients and severely mentally disordered inmates*. Trabajo presentado en la Primera Conferencia Anual de la International Association of Forensic Mental Health Services, Vancouver.

- Cronbach, L. J., y Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. <https://doi.org/10.1037/h0040957>
- \*Dahle, K. P. (2006). Strengths and limitations of actuarial prediction of criminal reoffence in a German prison sample: A comparative study of LSI-R, HCR-20 and PCL-R. *International Journal of Law and Psychiatry*, 29(5), 431-442. <https://doi.org/10.1016/j.ijlp.2006.03.001>
- \*Douglas, K. S., Klassen, C., Ross, D., Hart, S. D., Webster, C. D., y Eaves, D. (Agosto, 1998). *Psychometric properties of HCR-20 violence risk assessment scheme in insanity acquittees*. Trabajo presentado en la Conferencia Anual de la American Psychological Association, San Francisco.
- \*Dunbar, E. (2003). Symbolic, relational, and ideological signifiers of bias-motivated offenders: Toward a strategy of assessment. *American Journal of Orthopsychiatry*, 73(2), 203-211. <https://doi.org/10.1037/0002-9432.73.2.203>
- \*Dunbar, E., Quinones, J., y Crevecoeur, D. A. (2005). Assessment of hate crime offenders: The role of bias intent in examining violence risk. *Journal of Forensic Psychology Practice*, 5(1), 1-19. [https://doi.org/10.1300/J158v05n01\\_01](https://doi.org/10.1300/J158v05n01_01)
- Fernández-Suárez, A., Pérez, B., Herrero, J., Juarros-Basterretxea, J., y Rodríguez-Díaz, F. J. (2018). The role of psychopathic traits among intimate partner violent men: A systematic review. *Revista Iberoamericana de Psicología y Salud*, 9(2), 84-114. <https://doi.org/10.23923/j.rips.2018.02.017>
- Férriz, L., Sobral, J., y Gómez-Fraguela, J.A. (2018). Empatía y delincuencia juvenil: Un meta-análisis sobre la relación. *Revista Iberoamericana de Psicología y Salud*, 9(1), 1-16. <http://dx.doi.org/10.23923/j.rips.2018.01.011>
- Gallego, R., Novo, M., Fariña, F., y Arce, R. (2019). Child-to-parent Violence and Parent-to-child Violence: A Meta-analytic Review. *European Journal of Psychology Applied to Legal Context*, 11(2) 51-59. <https://doi.org/10.5093/ejpalc2019a4>
- \*Geraghty, K. A. (2015). *Assessing risk in female offenders* (Tesis doctoral). University of Birmingham, United Kingdom. Recuperado de <https://search.proquest.com/docview/1779966245?accountid=14542>
- \*Grimbos, T., Penney, S. R., Fernane, S., Prosser, A., Ray, I., y Simpson, A. I. (2016). Gender comparisons in a forensic sample: Patient profiles and HCR-20V2 reliability and item utility. *International Journal of Forensic Mental Health*, 15(2), 136-148. <https://doi.org/10.1080/14999013.2016.1152617>
- Heyman, R. E., Kogan, C. S., Foran, H. M., Burns, S. C., Smith Slep, A. M., Wojda, A. K., ..., y Reed, G. M. (2018). A case-controlled field study evaluating ICD-11 proposals for relational problems and intimate partner violence. *International Journal of Clinical and Health Psychology*, 18(2), 113-123. <https://doi.org/10.1016/j.ijchp.2018.03.001>
- \*Joyal, C. C., Côté, G., Meloche, J., y Hodgins, S. (2011). Severe mental illness and aggressive behavior: On the importance of considering subgroups. *The International Journal of Forensic Mental Health*, 10(2), 107-117. <https://doi.org/10.1080/14999013.2011.577136>
- Juarros-Basterretxea, J., Herrero, J., Fernández-Suárez, A., Pérez, B., y Rodríguez-Díaz, F. J. (2018). Are generalist batterers different from generally extra-family violent men? A study among imprisoned male violent offenders. *European Journal of Psychology Applied to Legal Context*, 10(1), 8-14. <http://dx.doi.org/10.5093/ejpalc2018a1>
- Juarros-Basterretxea, J., Overall, N., Herrero, J., y Rodríguez-Díaz, F. J. (2019). Considering the effect of sexism on psychological intimate partner violence: A study with imprisoned men. *European Journal of Psychology Applied to Legal Context*, 11(2), 61-69. <https://doi.org/10.5093/ejpalc2019a1>
- \*Kells, M. (2011). *The psychometric assessment of offenders with an intellectual disability* (Tesis doctoral), University of Birmingham, United Kingdom. Recuperado de <https://core.ac.uk/download/pdf/1631818.pdf>



- Loinaz, I. (2017). *Manual de evaluación del riesgo de violencia*. Metodología y ámbitos de aplicación. Madrid: Ediciones Pirámide (Grupo Anaya, S.A.).
- Martín, B., y Moral, M. V. (2019). Relación entre dependencia emocional y maltrato psicológico en forma de victimización y agresión en jóvenes. *Revista Iberoamericana de Psicología y Salud*, 10(2), 75-89. <https://doi.org/10.23923/j.rips.2019.02.027>
- Mathers, C., Stevens, G., Hogan, D., Retno Mahanani, W., y Ho, J. (2017). Global and Regional Causes of Death: Patterns and Trends, 2000-15. En D.T. Jamison, H. Gelband, S. Horton, P. Jha, R. Laxminarayan, C.N. Mock, y R. Nugent (Eds.), *Disease Control Priorities* (3a. ed., Vol. 9., pp. 65-104). Washington, D.C.: World Bank.
- \*Mills, J. F., y Gray, A. L. (2013). Two-tiered violence risk estimates: A validation study of an integrated-actuarial risk assessment instrument. *Psychological Services*, 10(4), 361-371. <https://doi.org/10.1037/a0032608>
- Morales-Toro, V., Guillén-Riquelme, A., y Quevedo-Blasco, R. (2019). Maltrato infantil y trastornos mentales en delinquentes juveniles: Una revisión sistemática. *Revista de Investigación en Educación*, 17(3), 218-238. Recuperado de <http://reined.webs.uvigo.es/index.php/reined/article/view/423/435>
- Muñiz, J. (2018). *Introducción a la Psicometría*. Teoría clásica y TRI. Madrid: Pirámide (Grupo Anaya, S. A.).
- \*Neal, T. M., Miller, S. L., y Shealy, R. C. (2015). A field study of a comprehensive Violence Risk Assessment Battery. *Criminal Justice and Behavior*, 42(9), 952-968. <https://doi.org/10.1177/0093854815572252>
- \*Nicholls, T. L. (2004). *Violence risk assessments with female NCRMD acquittees: Validity of the HCR-20 and PCL-SV* (Tesis doctoral). Simon Fraser University, Canadá.
- Oviedo, H. C., y Campo-Arias, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572-580. Recuperado de: <https://www.redalyc.org/pdf/806/80634409.pdf>
- \*Pham, T. H., Chevrier, I., Nioche, A., Ducro, C., y Réveillère, C. (2005). Psychopathy, risk assessment and rehabilitation. *Annales Medico-Psychologiques*, 163(10), 878-881. <https://doi.org/10.1016/j.amp.2005.10.016>
- Quevedo-Blasco, R., Guillén-Riquelme, A., y Buela-Casal, G. (2019). Análisis bibliométrico de las revistas de Psicología afines al ámbito Jurídico-Forense atendiendo a la WoS y el JCR (2018). *Revista de Investigación en Educación*, 17(3), 165-178. Recuperado de <http://reined.webs.uvigo.es/index.php/reined/article/view/419>
- Redondo, L., Fariña, F., Seijo, D., Novo, M., y Arce, R. (2019). A meta-analytical review of the responses in the MMPI-2/MMPI-2-RF clinical and restructured scales of parents in child custody dispute. *Anales de Psicología*, 35(1) 156-165. <http://dx.doi.org/10.6018/analesps.35.1.338381>
- Sada, A., Robles-García, R., Martínez-López, N., Hernández-Ramírez, R., Tovilla-Zarate, C. A., López-Munguía, F., ..., y Fresan, A. (2016). Assessing the reliability, predictive and construct validity of Historical, Clinical and Risk management-20 (HCR-20) in Mexican psychiatric inpatients. *Nordic Journal of Psychiatry*, 70(6), 456-461. <https://doi.org/10.3109/08039488.2016.1159330>
- Salvador, B., Arce, R., Rodríguez-Díaz, F. J., y Seijo, D. (2017). Evaluación psicométrica de la psicopatía: una revisión metaanalítica. *Revista Latinoamericana de Psicología*, 49(1), 36-47. <http://dx.doi.org/10.1016/j.rlp.2015.09.015>
- Sánchez-Meca, J., y López-Pina, J. A. (2008). El enfoque meta-analítico de generalización de la fiabilidad. *Acción Psicológica*, 5(2), 37-64. <https://doi.org/10.5944/ap.5.2.457>
- Santirso, F. A., Martín-Fernández, M., Lila, M. Gracia, E., y Terreros, E. (2018). Validation of the Working Alliance Inventory–Observer Short Version with male intimate partner violence offenders. *International Journal of Clinical and Health Psychology*, 18(2), 152-161. <https://doi.org/10.1016/j.ijchp.2018.02.003>



- Selaya-Berodia, A., Quevedo-Blasco, R., y Neufeld, C. B. (2018). Personalidad, estado emocional y ansiedad: Una revisión sistemática de su influencia sobre las falsas memorias. *Revista Mexicana de Psicología*, 35(1), 5-25. Recuperado de <http://comeppsi.com/images/rmp/RMP%20351%20005-025.pdf>
- \*Sellbom, M., Laurinavičius, A., Ustinavičiūtė, L., y Laurinaitytė, I. (2018). The Triarchic Psychopathy Measure: An examination in a Lithuanian inmate sample. *Psychological Assessment*, 30(7), e10-e20. <https://doi.org/10.1037/pas0000603>
- Strand, S. J. M., y Selenius, H. (2019) Assessing risk for inpatient physical violence in a female forensic psychiatric sample – comparing HCR-20v2 with the female additional manual to the HCR-20v2. *Nordic Journal of Psychiatry*, 73(4-5), 248-256. <https://doi.org/10.1080/08039488.2019.1613447>
- \*Telles, L. E., Day, V. P., Folino, J. O., y Taborda, J. G. (2009). Reliability of the Brazilian version of HCR-20 assessing risk for violence. *Brazilian Journal of Psychiatry*, 31(3), 253-256. <https://doi.org/10.1590/S1516-44462009005000001>
- Urrutia, S., y Bonfill, X. (2010). Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Medicina Clínica*, 135(11), 507-511. <http://doi.org/10.1016/j.medcli.2010.01.015>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48. <http://doi:10.18637/jss.v036.i03>
- Webster, C. D., Eaves, D., Douglas, K. S., y Wintrup, A. (1995). The HCR-20 scheme: *The assessment of dangerousness and risk*. Burnaby, Canadá: Mental Health, Law, and Policy Institute, and Forensic Psychiatric Services Commission of British Columbia.
- World Health Organization (2016). *WHO methods and data sources for country-level causes of death 2000-2015*. Ginebra, Suiza: World Health Organization. Recuperado de [https://www.who.int/healthinfo/global\\_burden\\_disease/GlobalCOD\\_method\\_2000\\_2015.pdf](https://www.who.int/healthinfo/global_burden_disease/GlobalCOD_method_2000_2015.pdf)
- Yanez-Peñúñuri, L. Y., Martínez-Gómez, J. A., y Rey-Anacona, C. A. (2019). Therapeutic intervention for victims and perpetrators of dating violence: a systematic review. *Revista Iberoamericana de Psicología y Salud*, 10(2), 107-121. <https://doi.org/10.23923/j.rips.2019.02.029>