

# MÁSTER EN PSICOLOGÍA LEGAL Y FORENSE



## TRABAJO FIN DE MÁSTER

**Revisión Sistemática:**

### **Instrumentos de Valoración de Riesgo de Violencia de Uso Forense en España**

**Alumno:** José M<sup>º</sup> Otín del Castillo

Dirigido por: Prof. Beatriz García Rodríguez

30 de octubre de 2022

## Índice

Índice.....	2
Resumen .....	3
1. Introducción y justificación .....	4
2. Metodología.....	6
2.1. Fuentes de Información.....	6
2.2. Procedimiento .....	6
2.3. Estrategia de Búsqueda .....	6
2.4. Criterios de Inclusión .....	7
2.5. Proceso de Extracción de Datos .....	7
3. Resultados.....	9
4. Discusión .....	40
4.1. Evidencias de validez predictiva.....	40
4.2. Metodología .....	42
4.3. Contextos de aplicación .....	42
4.4. Tamaño de las muestras .....	43
4.5. Discusión.....	43
5. Conclusiones .....	46
6. Referencias.....	48
7. Anexos.....	51

## Resumen

La valoración psicológica del riesgo de violencia es una de actuales competencias del psicólogo forense en el contexto de su actuación en el ámbito judicial, por lo que dicha actividad ha propiciado el desarrollo de numerosas herramientas destinadas a tal fin en distintos ámbitos y para diferentes poblaciones. Utilizando las principales bases de datos electrónicas y repositorios de investigación en el contexto psicológico forense, se llevó a cabo una revisión sistemática de los instrumentos de evaluación del riesgo de violencia que son utilizados actualmente en España. Después de la revisión, 11 estudios cumplieron con los criterios de inclusión. Se extrajeron, analizaron y sintetizaron sistemáticamente los datos sobre los instrumentos estudiados, los detalles de la muestra, el diseño del estudio, los métodos de análisis, la validez, y los contextos de aplicación. Los resultados indican que los investigadores de cinco países han probado once herramientas distintas de evaluación del riesgo de violencia, diseñadas para su uso en contextos diversos (policial, penitenciario, asistencial o de la administración de justicia). Tres estudios se centraron en la evaluación de riesgo de agresores de pareja, tres en violencia doméstica, cuatro en delincuentes juveniles y uno en agresores sexuales ingresados en centros psiquiátricos penitenciarios en Estados Unidos. La mayoría de los estudios fueron administrados o codificados por los investigadores en lugar de ser administrados en entornos reales.

La investigación actual respecto a los instrumentos de valoración de riesgo de violencia más utilizados en el contexto psicológico forense español se centran especialmente en los ámbitos de la predicción de riesgo de violencia juvenil y de pareja, donde el SAVRY y el sistema VioGén han mostrado buenas propiedades de validez en distintos contextos de aplicación. No obstante, se necesitan más investigaciones con muestras suficientes y una mayor homogeneidad en las fuentes primarias de obtención de datos que permitan una mayor fiabilidad en los resultados; así como respecto al diseño específico de instrumentos de predicción de riesgo de violencia en la pareja respecto a jóvenes adolescentes.

Palabras clave: riesgo de violencia, evaluación, forense, validez predictiva.

## 1. Introducción y justificación

La valoración psicológica del riesgo de violencia es una de actuales competencias del psicólogo forense en el contexto de su actuación pericial en el ámbito judicial, que se justifica en su utilidad por su capacidad para aportar datos de carácter científico al órgano judicial competente para la adopción de medidas de gestión del riesgo, en especial la supervisión y control del presunto agresor y la protección de la potencial víctima si fuera conocida (Esbec, 2003; Muñoz y López-Ossorio, 2016; Amor, Echeburúa y Bohórquez, 2020).

Las graves consecuencias que pueden derivarse de errores en la valoración predictiva, que pueden materializarse tanto en restricción de derechos y libertades y estigmatización para los evaluados a causa de los falsos positivos, como en riesgo de revictimización para las posibles víctimas por los falsos negativos, suponen una importante responsabilidad para el psicólogo forense, que debe extremar el rigor técnico y científico; y ha supuesto que exista cierta controversia en este tipo de periciales desde el ámbito jurídico, en el que algunas corrientes doctrinales estiman que atentan contra la presunción de inocencia y el principio del libre albedrío (Martínez-Garay, L., y Montes-Suay, 2018; Muñoz y López-Ossorio, 2016),

En este contexto, especialmente en las últimas dos décadas han surgido nuevos instrumentos para predecir la conducta violenta gracias al aumento de conocimiento de la naturaleza y procesos relacionados con el fenómeno y el cambio de paradigma que supuso la sustitución del concepto de *peligrosidad criminal* como atributo disposicional de carácter estático y dicotómico por el más acertado de *riesgo de violencia*, pasando así de la *estimación de peligrosidad* del sujeto a la *valoración del riesgo de reincidencia delictiva* (Andrés-Pueyo, 2013; Horcajo-Gil et al., 2019).

Esta valoración, de carácter probabilístico, encuentra su fundamento en la actual consideración del origen multicausal de la delincuencia, que apunta a la interacción y concurrencia de factores de protección y de riesgo, que permiten abordar el delito en términos de probabilidad, aunque no de determinación (Mampaso et al., 2014; en Horcajo-Gil et al., 2019).

Este nuevo enfoque de la cuestión ha permitido también la evolución del objetivo principal de estos instrumentos (la predicción del riesgo de violencia) hacia uno más amplio orientado también a la gestión de ese riesgo, con el beneficio que supone para una mejora de la prevención de las conductas violentas (Loínaz, 2017).

El objetivo del trabajo es efectuar una revisión sistemática de los principales instrumentos de valoración de riesgo de violencia de uso forense en España, planteando como preguntas de investigación: ¿cómo se ha probado la validez predictiva de los instrumentos seleccionados? ¿en qué entornos, poblaciones y ambientes se han evaluado? ¿qué resultados han obtenido? ¿qué interés tienen en el contexto psicológico forense español?

Se espera que los resultados de la presente revisión sean útiles para identificar los puntos fuertes y las lagunas metodológicas en la literatura especializada actual.

## 2. Metodología

La metodología descrita se ha seguido atendiendo a las recomendaciones de Sánchez-Meca (2010), Cajal et al. (2020) y PRISMA [*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*] (2020) para la elaboración de revisiones sistemáticas.

### 2.1. Fuentes de Información

Se han utilizado distintas fuentes de información basadas tanto en criterios formales (bases de datos indexadas, referencias extraídas de artículos primarios) e informales.

### 2.2. Procedimiento

En primer lugar, se ha realizado una revisión bibliográfica para contrastar los diferentes instrumentos de evaluación de riesgo de violencia en el ámbito forense que están siendo utilizados a nivel internacional con el fin de actualizar el estado de la cuestión. En una segunda fase se realizó una búsqueda utilizando los acrónimos de aquellos instrumentos más detectados en la búsqueda inicial. Finalmente se realizó una selección de trabajos sobre aquellos instrumentos más utilizados en el ámbito pericial forense español.

### 2.3. Estrategia de Búsqueda

Entre el 30 de junio y el 15 de julio de 2021 se realizó una búsqueda sistemática avanzada en las bases de datos especializadas de EBSCOhost (PsycInfo, PsycArticles y PSICODOC)=179 y SciELO=15, resultados totales obtenidos N=194, con los siguientes parámetros:

- Limitada al período 2010-2021
- Revisados por pares
- Uso de descriptores del tesoro APA “risk assessment” y “violence” conectados por marcador booleano AND, e “instruments”, “tools”, “measures”, “questionnaires”, “scales” conectados entre sí por el operador OR. Como términos en español se utilizaron “evaluación”, “riesgo” y “violencia” conectados por el operador Y.

Esta búsqueda se completó con referencias de artículos primarios (4) y búsqueda informal (10). Así, el número de artículos recuperados en bruto fue de 208. Una vez eliminados duplicados quedaron 185.

Dado el excesivo número de artículos se realizó un primer filtrado, dejando únicamente aquellos publicados entre 2018 y 2022, reduciendo la muestra a 60 trabajos.

Se realizó un segundo filtrado, seleccionando aquellas publicaciones con factor de impacto JCR situado en los Q1 y Q2 (2021), reduciendo la muestra a 29 trabajos. Adicionalmente, se realizó una

búsqueda entre las referencias de los trabajos seleccionados, añadiendo cuatro resultados más, quedando un total de 33 trabajos.

Finalmente, se seleccionaron 11 trabajos que cumplían todos los criterios de inclusión.

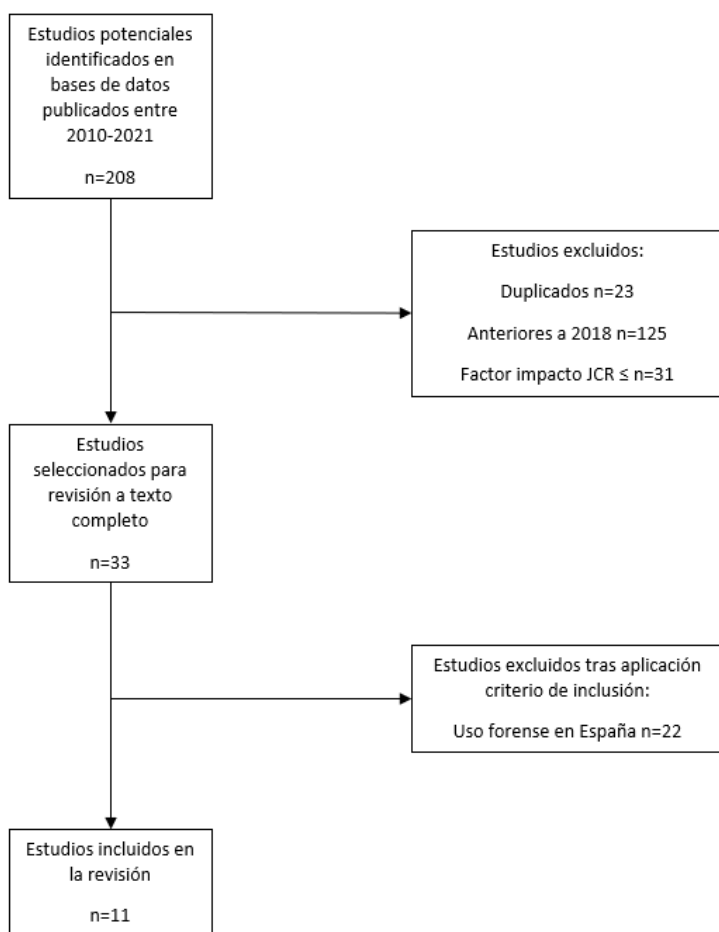
#### **2.4. Criterios de Inclusión**

- Fecha de publicación: 2018-2022
- Publicados en revistas con factor de impacto Q1-Q2 (2021)
- Estudios empíricos
- Uso forense en España de los instrumentos revisados (Andrés-Pueyo y Echeburúa, 2010): VRAG, PCL-R, HCR-20, SVR-20, SARA, EPV y SAVRY. Se incluye también VioGén por ser el instrumento oficial para las Fuerzas y Cuerpos de Seguridad del Estado.

#### **2.5. Proceso de Extracción de Datos**

Se efectuó un primer filtro atendiendo a la valoración cualitativa y dicotómica (sí/no) basado en la presencia o ausencia de los criterios establecidos eliminando las duplicidades detectadas; excluyendo aquellos trabajos que no cumplían todos los criterios de inclusión, tal como se muestra en el diagrama de búsqueda y selección de estudios para revisión:

Fig. 1. Diagrama de búsqueda y selección de estudios para revisión



Fuente: Elaboración propia



### 3. Resultados

Los trabajos estudiados evalúan once instrumentos de valoración de riesgo de violencia para diferentes contextos y poblaciones, si bien únicamente ocho son de uso en España. Tres estudios se centraron en la evaluación de riesgo de agresores de pareja, tres en violencia doméstica, cuatro en jóvenes delincuentes y uno en agresores sexuales ingresados en centros psiquiátricos penitenciarios en Estados Unidos. En la Tabla 1 se especifican la denominación y abreviaturas de los instrumentos analizados.

Se realiza a continuación un resumen de cada estudio siguiendo básicamente el esquema de redacción de trabajos científicos (justificación, metodología, discusión y resultados, conclusiones).

#### Estudio 1: *Intimate Partner Violence Risk Assessment in Community Health Facilities: A Multisite Longitudinal Study* (Arbach, K., y Bobbio, A, 2018)

---

Estudio prospectivo multicéntrico con un seguimiento a mínimo de un mes cuyo objetivo principal es analizar la validez predictiva del instrumento RVD-Bcn para la violencia de género en una muestra de mujeres usuarias de centros comunitarios de Córdoba (Argentina).

Se elige el RVD-Bcn por estar basado en el modelo de juicio profesional estructurado para evaluar el riesgo de violencia física grave por parte de la pareja y diseñado para ser completado cuando la principal fuente de información es la denuncia de la víctima y no se dispone de una entrevista con el agresor; contando con buenos niveles de consistencia interna ( $\alpha = .71$ ) y validez predictiva para las formas más graves de violencia física durante un seguimiento de seis meses (con un área bajo las curvas ROC en torno a  $a = .72$ )

La muestra estaba compuesta por 321 mujeres que acudían a cinco centros de salud públicos de la comunidad, con una edad media en el momento de la evaluación era de 37,14 años (SD = 13,54, rango 18-81). Los criterios de selección de fueron: tener más de 18 años y haber tenido al menos una pareja masculina en su vida.

Las evaluaciones se realizaron entre el 1 de octubre de 2014 y el 1 de agosto 2016, en el contexto de la práctica profesional habitual en los cinco centros de salud pública de la ciudad, previa formación específica de 40 evaluadores (un trabajador social, una enfermera y varios psicólogos, médicos y estudiantes avanzados de Psicología) para la aplicación del instrumento mediante entrevista con las participantes durante su práctica profesional rutinaria. Se les pidió que informaran de una calificación de riesgo resumida para poder comparar mediante puntuaciones numéricas.

En una segunda fase entre el 22 de enero de 2015 y el 19 de diciembre de 2016 se realizó un seguimiento (para la muestra osciló entre 47 y 311 días) mediante entrevistas personales o telefónicas por el mismo evaluador inicial en el caso de las entrevistas personas, y por estudiantes de psicología entrenados para ello en las telefónicas.

Para el análisis estadístico se utilizaron medidas de control y el análisis de la curva de características operativas del receptor (ROC) para examinar la validez predictiva de la puntuación total del instrumento y de las puntuaciones de riesgo resumidas para predecir la violencia de género. El nivel de significación estadística utilizado fue  $p < 0,05$ .

Como resultados se obtuvieron una fiabilidad de moderada a alta ( $\alpha = 0,85$ ) de la escala utilizada, una asociación significativa entre la calificación de riesgo resumida estimada en la evaluación inicial y la ocurrencia de violencia durante seguimiento,  $\chi^2(2, N = 225) = 68,99, p < .001$ ; y un AUC de la puntuación total fue de 0,89 (IC del 95% = 0,79 a 0,98,  $p < 0,001$ ), y de 0,88 (IC del 95% = 0,80 a 0,97,  $p < 0,001$ ) para la puntuación de riesgo resumida.

En la discusión se indica que el objetivo principal de la investigación era comprobar si el RVD, un instrumento para estimar el riesgo de violencia de pareja en los centros de salud de la comunidad es aplicable a las poblaciones de Argentina. Se informa que los resultados muestran que la precisión de la predicción de futuros eventos de violencia de género fue buena para una versión adaptada del instrumento de 14 ítems y se asemejó mucho a la eficacia de predicción de la versión original de 16 ítems del RVD y otras medidas de evaluación del riesgo en el mismo contexto; y también que los juicios de riesgo resumidos aumentan el poder predictivo de la puntuación numérica del RVD en concordancia con resultados similares informados para instrumentos similares.

Los autores señalan que este estudio proporciona una base empírica sólida para el uso de instrumentos estructurados de evaluación del riesgo de violencia de la violencia en los centros comunitarios a los que acuden las mujeres regularmente, del que pueden beneficiarse tanto las víctimas como los profesionales implicados en su trato, así como para la planificación de la seguridad, la defensa y el asesoramiento. No obstante, señalan con prudencia que ello no significa necesariamente que estos instrumentos sean también eficaces en cualquier población femenina o que otros instrumentos también serán eficaces en este grupo.

Indican también que sus resultados apoyan la evidencia convergente que muestra que la combinación de varios factores de riesgo posibles puede utilizarse para producir un instrumento de predicción eficaz y libre de influencia de factores culturales, por lo que este instrumento podría resultar eficaz en otros países latinoamericanos.

Se apunta en el trabajo que, en atención a la escasez de estudios prospectivos y longitudinales en la literatura especializada, el diseño del estudio es un punto fuerte, ya que la ocurrencia de la violencia se registró prospectivamente, lo que minimizó los sesgos confirmatorios en el proceso de evaluación del riesgo.

Como ejemplo de la utilidad del instrumento exponen que los profesionales pudieron utilizar los resultados de las calificaciones de la RVD para las decisiones relativas a las solicitudes de protección o para animar a las mujeres a presentar una denuncia. Las participantes de alto riesgo fueron derivadas a servicios de atención especializada donde recibirían medidas de protección. De este modo, apuntan que el uso de una calificación de riesgo RVD podría provocar una mayor vigilancia y una gestión y tratamiento del riesgo más eficaces, así como a un menor número de episodios de violencia que observar durante el seguimiento.

Los autores concluyen que el RVD ha demostrado ser una herramienta útil para la evaluación y gestión del riesgo a corto plazo (el periodo crítico de mayor reincidencia de los agresores). Como limitaciones para la interpretación de sus resultados señalan la especificidad geográfica y demográfica de la muestra, la fiabilidad entre calificadores del instrumento y que los datos relativos a los resultados violentos pueden no ser siempre fiables por distintos factores asociados a su obtención.

Apuntan como futuras líneas de investigación la comprobación de la utilidad del instrumento más allá del corto plazo, y examinar si la consideración de una lista predeterminada de lista de factores de riesgo, y el uso de RVD en particular, mejora los esfuerzos de gestión del riesgo y, en última instancia, reduce la prevalencia y la gravedad de los resultados violentos.

Finalizan haciendo hincapié en la especificidad del uso del instrumento, recomendándolo exclusivamente para evaluar el riesgo de violencia física; y en la necesidad de que una evaluación exhaustiva de la violencia debe incluir la consideración de otros factores contextuales o institucionales que han sido identificados como variables importantes que aumentan o disminuyen la probabilidad de la violencia de género.

---

*Estudio 2: Predictive Validity of HCR-20, START, and Static-99R Assessments in Predicting Institutional Aggression Among Sexual Offenders (Cartwright, Desmarais, Hazel, Griffith y Azizian, 2018)*

---

La justificación de este estudio se basa en la escasez de estudios que hayan examinado la validez de las evaluaciones de riesgo para predecir la violencia general no sexual en agresores sexuales institucionalizados, a pesar de existir evidencia indicativa del mayor riesgo de sufrir violencia no sexual que sexual de esta población.

Los objetivos específicos del estudio fueron: examinar la distribución de las puntuaciones de las evaluaciones HCR-20, START y Static-99R y las estimaciones de riesgo; evaluar la concordancia entre las evaluaciones HCR-20, START y Static-99R; y probar la validez predictiva de las evaluaciones HCR-20, START y Static-99R en la predicción de la agresión institucional durante 90 y 180 días.

La muestra se compuso de 152 varones detenidos o internados civilmente en un hospital forense de seguridad de Estados Unidos, con edades entre 22 y 79 años ( $M = 51,23$ ;  $SD = 10,04$ ), con distintos diagnósticos psiquiátricos según DSM-IV-TR: pedofilia (51,1%,  $n = 72$ ), trastorno antisocial de la personalidad (40,4%,  $n = 57$ ), parafilia no especificada (41,8%  $n = 59$ ), trastorno del estado de ánimo (20,7%  $n = 29$ ) y trastorno psicótico (15,7%  $n = 22$ ).

Los datos de seguimiento fueron recogidos durante períodos de 90 y 180 días después de la fecha de finalización de su evaluación final de riesgo de violencia (que incluye el uso del HCR-20, el START y el Static-99R en el plazo de una semana desde el ingreso), entre 2012 y 2014. Las evaluaciones se completaron utilizando toda la información disponible, incluida la revisión de los registros penales y hospitalarios, la observación del comportamiento del paciente en el hospital, las entrevistas con los pacientes en el contexto de la planificación del tratamiento y la revisión de casos, y la consulta con los miembros del equipo de tratamiento. Se aplicaron procedimientos internos de calidad para garantizar que los evaluadores cumplieran criterios de calificación, complementados con formaciones de refuerzo impartidas por los autores de los instrumentos de evaluación de riesgos.

Respecto a los instrumentos utilizados, el estudio examinó las puntuaciones de los ítems del HCR-20, las puntuaciones de tres subescalas, las puntuaciones totales y las estimaciones de riesgo de violencia; las calificaciones de los ítems del START, las puntuaciones totales de Fuerza y Vulnerabilidad y las estimaciones de riesgo de violencia; y las puntuaciones totales del Static-99R, excluyendo las categorías de riesgo. Incluyó el análisis de datos sobre incidentes violentos (la agresión física y verbal a otros internos y al personal, así como de daños materiales). Se calcularon estadísticas descriptivas para examinar la prevalencia de la agresión verbal, los daños materiales y la agresión física hacia otros durante el seguimiento de 90 y 180 días y para investigar las propiedades psicométricas del HCR-20, el START y el Static-99R.

Los resultados encontraron altas tasas de concordancia entre los resultados de las evaluaciones del HCR-20 y del START, pero bajas tasas de concordancia entre las puntuaciones totales del HCR-20 y del START con las puntuaciones totales del Static-99R y altas tasas de discordancia entre las puntuaciones totales del HCR-20 y del START con el Static-99R, si bien los autores atribuyen las discrepancias al distinto marco temporal de predicción entre los distintos instrumentos. Señalan que, no obstante, estos resultados indican que el HCR-20 y el START miden constructos y riesgos distintos de los que mide el Static-99R; y apoyan su uso en la práctica clínica con delincuentes sexuales. Inciden en que los resultados de los

análisis de validez predictiva proporcionaron un mayor apoyo al uso del HCR-20 y del START, en comparación con el Static-99R, para evaluar el riesgo de diferentes formas de agresión institucional entre los delincuentes sexuales.

Respecto a los distintos instrumentos evaluados, indican que en general, el rendimiento de las evaluaciones del HCR-20 fue mayor para la predicción de la agresión en el período de seguimiento de 180 días que en el de 90 días, demostrando una validez predictiva de buena a excelente, por lo que estos resultados se suman a la evidencia empírica que apoya el uso del HCR-20 para identificar el riesgo de violencia a medio plazo (es decir, seis meses) entre los delincuentes sexuales. Del mismo modo, apuntan como prometedores los resultados del START y apoyan su uso en la evaluación y gestión del riesgo de agresión institucional a corto plazo entre los delincuentes sexuales, atendiendo a que las puntuaciones totales de Vulnerabilidad y Fortaleza, así como las estimaciones de riesgo de violencia, mostraron una validez entre buena y excelente para predecir cualquier agresión y la agresión verbal durante los períodos de seguimiento de 90 y 180 días. Ponen de relieve que, de todas las evaluaciones, la puntuación total de Vulnerabilidad de START superó a cualquier otra subescala o puntuación total de HCR-20, START o Static-99R, demostrando los valores AUC más altos observados.

Por último, respecto al Static-99R indican que sus evaluaciones mostraron una validez de regular a buena para predecir cualquier agresión y la agresión verbal, así como los daños materiales, pero no la agresión física hacia otros (inciden en que éste es el más importante de sus resultados). Atendiendo al inferior rendimiento de este instrumento frente a los otros, sugieren que el Static-99R es el más adecuado para estimar el riesgo de reincidencia sexual y que los instrumentos de evaluación del riesgo de violencia general, como el HCR-20 o el START, deberían utilizarse para evaluar la agresión general en los delincuentes sexuales.

Como limitaciones señalan la limitación de la generalización de los resultados del estudio por el tamaño de la muestra y su especificidad (delincuentes sexuales masculinos de una sola institución). Añaden además que los datos sobre las conductas agresivas se recopilaron a partir de informes de incidentes críticos, lo que puede subestimar la prevalencia de los actos agresivos, y que no se dispone de datos sobre la fiabilidad entre los evaluadores de las variables del estudio. Por último, señalan que se utilizó la versión 2 del HCR-20 y que sería necesario por tanto comprobar la posibilidad de reproducir los resultados actuales con el HCR-20 V3.

Sugieren que las investigaciones futuras deberían examinar las evaluaciones del HCR-20 y del START realizadas en delincuentes sexuales en múltiples entornos y contextos, incluir a las mujeres delincuentes sexuales, complementar los registros oficiales con otras fuentes, comprobar la fiabilidad entre los evaluadores y examinar la validez predictiva en función de los diferentes tipos de agresión (por

ejemplo, psicótica, impulsiva, depredadora). También la posibilidad de comprobar si la misma muestra podría producir resultados únicos a otras poblaciones de abusadores sexuales.

Entre las conclusiones del estudio destacan que el estudio apoya la validez de las evaluaciones HCR-20, START y, en menor medida, la Static-99R para predecir la agresión institucional entre los pacientes detenidos o internados civilmente en virtud de la ley de depredadores sexuales; que los resultados indican que los instrumentos diseñados para evaluar el riesgo de reincidencia sexual, y el Static-99R en particular, son limitados en su capacidad para evaluar el riesgo de violencia general (no sexual); que los hallazgos sugieren que el uso del HCR-20 o del START para identificar el riesgo general de violencia entre los delincuentes sexuales beneficiaría la gestión de los casos y el tratamiento, así como ayudaría en las decisiones relativas a la supervisión y la liberación; que las evaluaciones del HCR-20 y del START pueden ser útiles para distinguir entre los pacientes más o menos propensos a tener conductas agresivas, incluso dentro de una población algo homogénea y de alto riesgo, así como para informar sobre las decisiones de supervisión y las estrategias de gestión de riesgos (por ejemplo, identificar qué pacientes requieren niveles de seguridad más altos).

Estudio 3: *Predictors of Intimate Partner Homicide in a Sample of Portuguese Male Domestic Offenders* (Cunha y Goncalves, 2019)

---

El propósito de este trabajo fue examinar las diferencias entre la violencia de pareja y el homicidio por parte de la pareja íntima e identificar las variables que predicen éste, atendiendo a los datos que sugieren una dinámica diferente entre ambas conductas, aunque también muchos puntos en común. Los autores justifican la necesidad del trabajo en la escasez de estudios sobre el homicidio por parte de la pareja íntima en Portugal, tomando como hipótesis que dicha conducta es el último eslabón de la cadena de conductas violentas, por lo que es importante determinar los posibles predictores de la misma. La predicción del riesgo de violencia en la pareja y de homicidio en este contexto facilita la toma de conciencia del problema y la búsqueda de las mejores soluciones.

Comienza el trabajo realizando una revisión de estudios sobre factores de riesgo y perfiles de agresores y víctimas en este ámbito, planteando que, a pesar de la amplia literatura existente sobre los factores de riesgo de ambas conductas (violencia y homicidio), una cuestión central es si existen diferencias en los factores de riesgo relevantes para ambas. Señalan en este sentido que la literatura proporciona apoyo a dos perspectivas diferentes. Por un parte, el asesinato en la intimidad como acontecimiento repentino o inesperado sin una historia previa de comportamiento violento. El otro punto de vista afirma que el asesinato es la culminación de una historia de violencia y no es en sí mismo un fenómeno discreto.

Así, a partir de estos hallazgos contradictorios los autores estiman necesario comprender y aclarar si existen diferencias entre los factores de riesgo de ambos fenómenos, para lo que realizaron un estudio cuantitativo utilizando un instrumento específico para analizar los factores de riesgo, la Guía de Evaluación del Riesgo de Agresión Conyugal (Spousal Assault Risk Assessment Guide, SARA) entre hombres portugueses que asesinaron a sus parejas o exparejas y hombres que maltrataron. Los objetivos de este estudio fueron identificar los factores de riesgo de violencia en la pareja y de homicidio por parte de la pareja íntima, examinar si existen diferencias en términos de factores de riesgo e identificar las variables que mejor predicen el homicidio.

La muestra del estudio se compuso de 172 hombres adultos, distribuidos en dos grupos (maltratadores,  $n = 137$ , y homicidas ( $n = 35$ ), reclutados de la prisión ( $n = 96$ ; 55,8%) y de la comunidad ( $n = 76$ ; 44,2%). Los agresores de la comunidad asistían a programas de intervención contra la violencia doméstica (21,1%) o estaban bajo la supervisión de los servicios de libertad condicional (52,6%) o de los servicios de protección de menores (26,3%). Los criterios de inclusión fueron: ser varón; mayor de edad (más de 18 años); tener una relación íntima actual o anterior con una víctima femenina (es decir, cónyuge o excónyuge, pareja o expareja, novia o exnovia); haber ejercido violencia física, psicológica y/o sexual contra la víctima; y tener las habilidades cognitivas para responder a la entrevista.

La recogida de datos se efectuó mediante entrevistas con los participantes entre 2012 y 2013, utilizando el instrumento de evaluación de riesgo actuarial SARA, por sus apropiadas cualidades psicométricas, su capacidad de discriminación entre los agresores con y sin antecedentes de violencia conyugal y entre los agresores conyugales reincidentes y no reincidentes; la posibilidad de diferenciar a los agresores en términos de factores de riesgo individuales y de riesgo global percibido; la inclusión de factores de riesgo identificados sobre la base de una revisión de la literatura científica relevante, así como una consideración de las cuestiones clínicas y legales pertinentes; y que el grado global de riesgo planteado por el individuo tiene en cuenta la naturaleza, la gravedad, la probabilidad, la frecuencia y la inminencia de cualquier violencia futura.

El instrumento fue puntuado de forma independiente por dos codificadores con formación en psicología utilizando el coeficiente Kappa de Cohen para medir el grado de fiabilidad entre ambos, que osciló entre 0,72 y 0,96. Las calificaciones del SARA se basaron en una entrevista con el delincuente y en una revisión de toda la información pertinente del expediente, tal como se recomienda en el manual. Los datos sociodemográficos fueron codificados conforme a un cuestionario sociodemográfico y se analizaron los expedientes judiciales y organizativos individuales de los autores para obtener información sobre los antecedentes penales (condenas anteriores por violencia doméstica y otros delitos) y los delitos perpetrados (tipo de delito, contexto). Los ítems y variables se codificaron de forma dicotómica para el análisis estadístico, en el que se utilizaron estadísticas descriptivas, pruebas  $t$  y chi-cuadrado para

examinar las diferencias entre el tipo de delito y todas las variables. Aquellas más significativas se incluyeron en el posterior análisis de regresión logística multivariante. Por último, el poder predictivo del modelo generado por este procedimiento se estimó mediante el análisis ROC.

Como resultados más relevantes cabe destacar que los homicidas tienen 1 vez más probabilidades de ser mayores, 13 veces más probabilidades de tener ideación/intento suicida y/o homicida en el año anterior al delito, y casi 18 veces más probabilidades de utilizar armas o amenazas de muerte creíbles. A la inversa, la probabilidad de perpetrar violencia en la pareja disminuye en un 95% si el hombre no es soltero, disminuye en un 96% si el hombre tiene un historial de violación de la libertad condicional o de la supervisión comunitaria y disminuye en un 77% si tiene un historial de victimización y/o exposición a la violencia familiar en la infancia.

Los resultados del estudio sugieren que, a pesar de que ambos grupos comparten algunas características, otros factores y variables permiten distinguir entre homicidas y maltratadores, desafiando la suposición de que el asesinato es un acontecimiento inexplicable y repentino y revelando que estos delitos son explicables. Señalan que la mayor edad de los homicidas y ser solteros son dos factores que parecen ser predictores del homicidio, de acuerdo con la literatura anterior.

Resaltan que, aunque se han identificado factores familiares, individuales y culturales como precursores o correlatos del maltrato grave por parte de la pareja, los homicidas tenían menos probabilidades de experimentar y/o ser testigos de violencia en la infancia, de abusar de sustancias, de presentar un historial de violación de órdenes de libertad condicional o de restricción, y de minimizar y negar un historial de agresión conyugal que los maltratadores.

Sin embargo, en este estudio, sólo los antecedentes de victimización y/o exposición a la violencia familiar y la violación de la libertad condicional aparecieron como predictores, y ambos disminuyeron sustancialmente (en más del 90%) la probabilidad de que un hombre cometiera homicidio, lo que podría sugerir que haber sido víctima o testigo de la violencia en la infancia puede hacer que los individuos sean más conscientes de las consecuencias de las formas letales de la violencia en la pareja y que la aplicación de órdenes de protección civiles o penales es probable que prevenga el homicidio. Al mismo tiempo, estos resultados, especialmente los relacionados con la violación de las órdenes de libertad condicional o de restricción, sugieren que los autores de delitos graves contra las parejas íntimas parecen tener un comportamiento social adecuado y pueden ser menos antisociales que los autores de violencia menos grave.

Identifican también como importantes predictores del homicidio el mayor historial de ideación o intención suicida/homicida y un fuerte historial de uso de armas y amenazas de muerte creíbles contra la víctima, aumentando notablemente (más de 10 veces) la probabilidad de que un hombre cometa un



asesinato contra su pareja; en concordancia con la literatura existente, indicando que este hallazgo respalda los resultados anteriores sobre el poder predictivo de la violencia doméstica previa en el homicidio posterior y refuerzan la opinión de que el homicidio en el contexto de la pareja no es un acontecimiento repentino o inesperado ni un fenómeno imprevisto perpetrado por un hombre relativamente convencional.

Los autores reflexionan sobre la posibilidad de que, en algunos casos, las evaluaciones de riesgo formales e intervenciones previas podrían haber subestimado el potencial predictor de riesgo de violencia letal de la existencia de amenazas homicidas o suicidas y el uso de armas y/o amenazas de muerte creíbles durante el año anterior al delito y que al menos algunos homicidios o intentos de homicidio podrían ser delitos potencialmente prevenibles.

Se señalan también las similitudes compartidas entre ambos grupos como el historial de agresiones a miembros de la familia y de agresiones a extraños, un historial continuo de problemas en las relaciones íntimas, un historial de agresiones físicas, agresiones sexuales o celos contra la mujer, una escalada continua de agresiones en el año anterior al incidente, violaciones de órdenes de contacto y actitudes legitimadoras de la violencia; llamando la atención sobre la importancia de las cogniciones y orientaciones de los hombres con respecto a las relaciones íntimas.

Apuntan que los resultados del estudio llaman la atención sobre la necesidad de desarrollar campañas públicas educativas, promover la formación profesional en el sistema de aplicación de la ley y para los profesionales de la salud, desarrollar y aplicar instrumentos de evaluación del riesgo para ambas conductas, desarrollar políticas gubernamentales e institucionales comunitarias para promover la igualdad de relaciones de género, y desarrollar medidas legislativas y judiciales más eficaces para eliminar la violencia contra las mujeres en Portugal.

Como limitaciones, indican que se utilizó un enfoque retrospectivo basado en el discurso de los agresores, pudiendo verse afectados los informes por distintos sesgos; que sólo incluyó factores relacionados con el agresor y los delitos y omitió las variantes relacionadas con la víctima y el relato de los hechos por parte de ésta; la escasa representatividad de la muestra; el uso de relaciones predictivas lineales estáticas sin incluir efectos de interacción en la predicción de la gravedad de la violencia; la ausencia de análisis de otros factores relevantes que se han considerado importantes en relación con el homicidio en la relación de pareja; y el hecho de que el estudio se realizó con una muestra de agresores identificados por el sistema judicial, limitando así las conclusiones.

Finalizan reflejando que las investigaciones posteriores deberían incluir estas variables para determinar si estas cuestiones se aplican al contexto portugués; así como el interés de la inclusión en los

análisis de factores contextuales como el patriarcado, el sexismo y los roles de género por su posible relación con estas manifestaciones de violencia.

Estudio 4: *Predictive Validity of the Structured Assessment of Violence Risk in Youth (SAVRY) Among a Sample of Asian Canadian Youth on Probation* (Li, Viljoen, Christiansen y Muir, 2020)

---

Se trata de un estudio basado en un diseño prospectivo de campo realizado en Canadá sobre una muestra de jóvenes asiáticos canadienses frente a jóvenes blancos canadienses, con el objetivo de analizar la validez predictiva del instrumento de evaluación de riesgo de violencia en la juventud SAVRY (Borum et al., 2006), que utiliza el modelo de juicio profesional estructurado y que, en poblaciones de raza blanca, ha mostrado buenas propiedades psicométricas: de acuerdo con estudios anteriores, tiene una fiabilidad Inter observadores de buena a excelente (ICC = 0,72 a 0,97), una alta consistencia interna (alfa de Cronbach = 0,82 a 0,90), y una buena validez predictiva (AUC = 0,74 a 0,80). Sin embargo, se ha cuestionado su validez en otras poblaciones como las minorías raciales y étnicas.

Los autores justifican la necesidad de su estudio poniendo de manifiesto la falta de investigación sobre personas de ascendencia asiática en este campo por diversos motivos como los estereotipos o la escasez de muestras y de información sobre la raza; y en particular de las poblaciones asiáticas en países occidentales. Sin embargo, indican que las herramientas de evaluación del riesgo tienden a reflejar las normas occidentales y, por lo tanto, pueden no incluir factores de riesgo que son relevantes para las minorías raciales y étnicas (aculturación, diferencias intergeneracionales, discriminación...).

En el estudio se plantean tres preguntas de investigación: ¿Es más probable que los agentes de libertad condicional califiquen a los jóvenes canadienses del este/sudeste de Asia y del sur de Asia como de menor riesgo en el SAVRY en comparación con los jóvenes canadienses blancos? ¿Cuál es la validez predictiva de las puntuaciones totales de riesgo del SAVRY y las calificaciones de riesgo resumidas para los jóvenes canadienses del este/sudeste de Asia en comparación con los jóvenes canadienses blancos para la reincidencia violenta y no violenta? y ¿Modera la raza la relación entre las puntuaciones SAVRY y el tiempo que transcurre hasta la primera reincidencia violenta y no violenta de los jóvenes? A su vez, se plantean como hipótesis que, a pesar de que las puntuaciones totales de riesgo SAVRY de los jóvenes y las calificaciones de riesgo sumarias predecirían la reincidencia violenta y no violenta para ambos grupos asiático-canadienses, la validez predictiva para estos fuera menor que para los jóvenes blanco-canadienses, porque los asiático-canadienses pueden tener factores de riesgo que no son captados en el SAVRY (por ejemplo, la aculturación), y ciertos ítems del SAVRY (por ejemplo, la supervisión de los padres) pueden no ser como predictivo para los canadienses de origen asiático, que a su vez serían considerados de menor riesgo que los jóvenes canadienses de raza blanca debido al mito de la “minoría modelo”

(estereotipo contemporáneo de que las personas con ascendencia asiática son "minorías modelo" que tienen éxito social y económico, son trabajadores y respetan la ley).

La muestra se compuso de 573 jóvenes (445 canadienses blancos, 56 canadienses del este/sudeste de Asia y 72 canadienses del sur de Asia) en libertad condicional en una provincia canadiense. Se realizó un seguimiento prospectivo de los jóvenes durante una media de 1,97 años ( $SD = 0,56$  años) para determinar si posteriormente eran acusados de delitos violentos o no violentos; mediante la implantación del SAVRY en las evaluaciones de los agentes de libertad condicional previamente instruidos para ello.

Para evaluar la reincidencia se utilizaron los registros judiciales. La reincidencia violenta y la no violenta se examinaron por separado. Los datos de reincidencia recogen tanto los delitos cometidos en la juventud como en la edad adulta (es decir, a partir de los 18 años). También se calculó el tiempo hasta la primera reincidencia violenta y no violenta después de la evaluación SAVRY.

La validez predictiva se midió mediante análisis ROC, calculándose también la sensibilidad (proporción de reincidentes con calificaciones de alto riesgo) y la especificidad (proporción de no reincidentes con calificaciones de bajo riesgo) para examinar las diferencias de grupo en las proporciones de reincidentes y no reincidentes correctamente identificados. También se calcularon los valores predictivos positivos y negativos, los falsos positivos y los falsos negativos.

Los resultados indicaron que, en comparación con los canadienses blancos, ambos subgrupos de asiáticos canadienses tenían puntuaciones de riesgo SAVRY totales, históricas e individuales/clínicas significativamente más bajas. Además, los canadienses del sur de Asia eran significativamente menos propensos que los canadienses blancos a ser calificados como de alto riesgo de reincidencia no violenta. No surgieron diferencias significativas entre los subgrupos de canadienses asiáticos.

En general, las puntuaciones totales y de dominio del riesgo SAVRY predijeron la reincidencia violenta y no violenta con tamaños de efecto grandes para ambos subgrupos asiático-canadienses. Los autores destacan al respecto que, a pesar de que los agentes de libertad condicional calificaron a los jóvenes asiáticos canadienses como de menor riesgo, no tenían una probabilidad significativamente menor que los jóvenes blancos canadienses de reincidir, lo que podría deberse a las percepciones y mitos contemporáneos sobre los individuos de ascendencia asiática. Las puntuaciones de riesgo más bajas de los jóvenes asiáticos canadienses podrían deberse, en parte, a un énfasis cultural en la privacidad familiar (por ejemplo, las familias de Asia oriental, sudoriental y meridional podrían ser menos propensas a revelar información negativa, como un historial de maltrato infantil).

No obstante, destacan que el SAVRY mostró una fuerte validez predictiva para los subgrupos asiáticos, en contraposición a su hipótesis inicial, dándose la circunstancia de que los agentes de libertad

condicional identificaron mejor la reincidencia de los jóvenes asiático-canadienses que de los jóvenes blanco-canadienses. Así, aunque la hipótesis inicial era que tendrían más probabilidades de calificar incorrectamente a los jóvenes canadienses de origen asiático como de bajo riesgo, lo que indicaría una infravaloración del riesgo, esta hipótesis no se confirmó. Por ejemplo, casi todos los asiáticos canadienses que los evaluadores calificaron como de bajo riesgo de reincidencia violenta no cometieron ningún delito violento. Sugieren por tanto que el uso del SAVRY podría ayudar a contrarrestar los estereotipos del mito de la minoría modelo, al proporcionar una estructura estandarizada para evaluar el riesgo.

En general, la validez predictiva del SAVRY fue similar para los canadienses del este/sudeste y del sur de Asia; aunque apuntan que la influencia de las tasas base de reincidencia podría tener incidencia en los resultados obtenidos, por lo que sería importante comprobar si los resultados varían con las muestras de canadienses asiáticos que tienen tasas de base más altas.

Como limitaciones del estudio señalan el pequeño tamaño de las muestras, la imposibilidad de examinar las diferencias entre los canadienses de Asia Oriental y del Sudeste, y la posible influencia de los prejuicios raciales en los registros oficiales utilizados para el estudio.

Concluyen indicando que sus hallazgos sugieren la utilidad del SAVRY para la predicción de la reincidencia entre los jóvenes canadienses de Asia oriental, sudoriental y meridional, y que a pesar de ello los profesionales que apliquen el instrumento con estas poblaciones deben tratar de aumentar la formación intercultural para reducir los sesgos y estereotipos en la evaluación.

En cuanto a futuras líneas de investigación, sugieren el interés en examinar la validez predictiva del SAVRY para los jóvenes de ascendencia asiática en diferentes países y contextos, incluyendo la predicción de reincidencia en función del género. Asociado a esto, indican la necesidad de investigar la posible influencia de los sesgos y antecedentes raciales, étnicos y culturales de los propios evaluadores.

*Estudio 5: Validation and Calibration of the Spanish Police Intimate Partner Violence Risk Assessment System (VioGén), (López-Ossorio, González, Muñoz, Urruela y Andrés-Pueyo, 2019)*

---

Este estudio describe los fundamentos, el desarrollo y la validación de los formularios de evaluación policial del riesgo de violencia de pareja del Sistema VioGén del Ministerio del Interior español (VPR4.0 y VPER4.0), utilizados en el contexto policial, principal vía de acceso de la víctima al sistema judicial y por lo tanto idóneo para adoptar las primeras medidas de protección.

La muestra fueron 6613 nuevos casos de violencia de pareja incluidos en el Sistema VioGén durante un periodo de dos meses seguidos longitudinalmente durante seis meses. Las víctimas femeninas tenían una edad media de 34,5 años ( $DE = 13,03$ ) y un rango de edad de 13 a 68 años. En cuanto a los

agresores denunciados, la edad media fue de 36,51 años ( $DE = 14,57$ ) y un rango de edad de 14 a 68 años. El 65,7% de esta muestra y el 67,9% de los agresores eran de nacionalidad española, distribuyéndose el resto porcentualmente entre 89 países diferentes de forma similar en ambos sexos.

Los instrumentos utilizados fueron los nuevos protocolos de evaluación policial del riesgo *Evaluación Policial del Riesgo/Valoración Policial del Riesgo (VPR4.0)* y *Evaluación Policial de la Evolución del Riesgo/Valoración Policial del Riesgo (VPER4.0)*. Ambos instrumentos se construyeron tras revisión bibliográfica de indicadores de riesgo de violencia de pareja en bases de datos y fuentes documentales por expertos universitarios y profesionales de la policía; abarcando exhaustiva revisión de estudios (validez racional), consultas a usuarios profesionales (validez de la respuesta) y opinión de expertos (validación por parte de los jueces) para una adecuada validez de contenido. Estos indicadores recopilados se filtraron para cada uno de los formularios, con el fin de generar cinco niveles de riesgo (no apreciado, bajo, medio, alto y extremo) y cinco dimensiones temáticas (gravedad del episodio reportado, factores relacionados con el agresor, indicadores de vulnerabilidad de la víctima, circunstancias agravantes e indicadores para seguimiento de riesgo y actualización de medidas de protección aplicadas).

Se calculó la fiabilidad intraobservador y la validez convergente mediante el análisis de los datos por un grupo de psicólogos entrenados para confirmar la presencia de los indicadores de las nuevas formas y calcular la correlación con los indicados por los agentes de la ley (ciegos a esta información) y el Protocolo de Evaluación del Riesgo de Violencia contra las Mujeres por parte de sus parejas o exparejas (RVD-Bcn).

Para los análisis estadísticos las variables independientes correspondieron a los formularios de evaluación del riesgo VPR4.0 y VPER4.0. Las variables dependientes fueron la reincidencia en los casos de VPI en un periodo de tiempo de tres a seis meses para la muestra de validación del VPR4.0 ( $n = 6613$ ; reincidencia de 490 casos) y la gravedad de la violencia, que se considera la variable más importante del análisis por su relevancia ( $n = 3725$ ; violencia grave en 487 casos). Para la VPER4.0, la variable dependiente de la reincidencia procede de su propia muestra ( $n = 3973$ ; reincidencia de 246 casos).

Los resultados mostraron que los instrumentos utilizados son comparables con los reportados en la literatura nacional e internacional para herramientas similares en cuanto a validez predictiva. La distribución de la escala de los puntos de corte de la VPR4.0 en función de la probabilidad de reincidencia, la gravedad y la capacidad de respuesta de las fuerzas de seguridad mostró buenos resultados para una herramienta de cribado utilizada por los agentes durante el primer acercamiento al caso a través de una información limitada; y sugieren que puede detectar eficientemente a individuos con bajo riesgo de reincidencia, por lo que las herramientas actuariales parecen encajar mejor en los planes de gestión de riesgos (especialmente en el contexto policial) en grupos de riesgo seleccionados, comparando a los individuos con el grupo de referencia. En cuanto a los resultados predictivos sobre el riesgo de

reincidencia de violencia grave resultaron moderados, sugiriéndose que pudiera deberse a las dificultades que tienen los policías para recopilar información en las primeras fases, limitación inherente al contexto de exploración inicial.

Para lograr la implementación de un nuevo formulario VPER con indicadores de riesgo de reincidencia y factores de protección relacionados con aquella, los autores desarrollan un nuevo formulario de carácter semiactuarial de funcionamiento autónomo mediante su propio algoritmo capaz de realizar predicciones a través de un conjunto mixto (de protección del riesgo) de indicadores que proporciona la posibilidad de evaluar la probabilidad de reincidencia y, al mismo tiempo, monitorizar los cambios que se producen en el tiempo, mejorando los parámetros de actuación de VPR4.0 y superando sus limitaciones de acceso y calidad de la información disponible durante la primera evaluación, posicionándose como el elemento clave del protocolo general. Así, cuando los factores de protección incluidos en el VPER4.0 están presentes, proporcionan una buena predicción de los casos, cuya evolución tanto positiva como negativa puede ser abordada mediante el uso de las dos versiones del formulario.

Señalan que la versión del VPER-S4.0 es capaz de monitorizar el progreso de los eventos y reducir el nivel de riesgo, a la vez que es sensible a las incidencias y características del agresor que pueden aumentar las probabilidades de un nuevo evento violento. El VPER-C4.0 no sólo incluye los indicadores de la versión corta, sino que también integra un conjunto de indicadores centrados en la investigación del suceso violento o las características de un posible incumplimiento de las medidas judiciales con el objetivo adicional de reducir la multirreincidencia.

Concluyen que los resultados (teniendo en cuenta que los valores obtenidos están dentro de los márgenes reportados por diferentes metaanálisis respecto a este tipo de instrumentos) muestran la idoneidad de utilizar procedimientos que, de forma coordinada, incorporen dos instrumentos de evaluación del riesgo, uno para una primera evaluación de cribado y un segundo para reevaluar periódicamente las situaciones de peligro de violencia de pareja; lo que apoya su uso para la práctica profesional. Apuestan por las herramientas de carácter actuarial como las utilizadas ante las metodologías basadas en el juicio profesional estructurado, por su transparencia y fiabilidad para realizar predicciones.

Como limitaciones señalan las inherentes a los estudios de investigación en la vida real y entornos aplicados; el hecho de que en este tipo de predicciones de probabilidad de violencia futura no es posible saber con certeza si la intervención policial tuvo algún efecto en el resultado final de la reincidencia; el dilema que supone la perspectiva ideográfica frente la nomotética en estos diseños, donde las variables de riesgo de un caso concreto pueden no estar bien representadas en las evaluaciones actuariales; y que el tipo de diseño utilizado no ofrece las mismas garantías de los estudios longitudinales prospectivos.

Estudio 6: *Intimate Partner Homicide Risk Assessment by Police in Spain: The Dual Protocol VPR5.0-H* (López-Ossorio, González-Álvarez, Loinaz, Martínez y Pineda, 2020)

---

Este trabajo, que puede decirse es una continuación del anterior, describe la evolución de la herramienta VPR5.0 (Valoración Policial del Riesgo del Sistema VioGén) a un algoritmo diferenciado cuyo objetivo es predecir mejor los homicidios de mujeres por su pareja, cuestión que los autores estiman de importancia no tanto por su incidencia estadística, con una baja prevalencia en España (estimada en 0.30 por cada 100.000 mujeres), como por su impacto personal y social.

Destacan los autores que los índices de predicción de las evaluaciones policiales del riesgo resultan bajos para la violencia letal, a causa de la insuficiente identificación de indicadores de riesgo adecuados. Sugieren que las razones para ello pueden ser la ya mencionada baja prevalencia de estos casos; la variedad de perfiles de este tipo de agresores; y la inexistencia en muchos casos de antecedentes de violencia.

Formulan como objetivos del estudio determinar los factores de riesgo específicos del homicidio contra la pareja, en contraste con los de la violencia de pareja general en España y, en segundo lugar, mejorar la predicción de los homicidios de mujeres por su pareja mediante el desarrollo y la validación de una nueva escala para proporcionar una estimación del riesgo de homicidio complementaria al algoritmo principal de la escala VPR5.0-H.

La muestra se compuso de 2.159 registros, 159 de los cuales correspondían a homicidios de mujeres por su pareja ocurridos entre febrero de 2006 y septiembre de 2018; dividiendo la muestra en dos grupos comparables con casos de homicidios de mujeres por su pareja y controles (sin homicidios de mujeres por su pareja) correspondientes a denuncias aleatorias registradas en el Sistema VioGén en el último trimestre de 2016 para validar los análisis. En el grupo de homicidas la edad media de las mujeres fallecidas fue de 41,49 años ( $DT = 14,66$ , rango = 14-77) y el 68,0% eran españolas; el 27,6% de la muestra había denunciado previamente al agresor por al menos un episodio de violencia (porcentaje similar al 25,9% de denuncias presentes en el total de casos de HIP del citado periodo). En cuanto a los homicidas, la edad media fue de 46,38 años ( $DE = 14,56$ , rango = 19-86) y el 70,7% eran españoles; el 48,7% de los casos de HIP ocurrieron en zonas urbanas y el 51,3% en zonas rurales. El porcentaje de suicidios consumados tras el homicidio de la muestra analizada fue del 22,0%, y en todos los homicidios del periodo fue del 20,2%. En el grupo control las víctimas tenían una edad media de 34,5 años ( $DE = 13,03$ , rango = 13-68). La edad media de los agresores denunciados fue de 36,51 años ( $DT = 14,57$ , rango = 14-68); el 65,7% de las víctimas y el 67,9% de los denunciados eran de nacionalidad española; el 65,0% de las denuncias se presentaron en zonas urbanas y el 35,0% en zonas rurales.

El instrumento utilizado fue el ya mencionado formulario VPR5.0, compuesto por 35 factores de riesgo codificados de forma dicotómica (presente/ausente), proporciona cinco niveles de riesgo: no apreciado, bajo, medio, alto y extremo; agrupándose los factores en cinco dimensiones: antecedentes de violencia de pareja; factores relacionados con el agresor; indicadores relacionados con la vulnerabilidad de las víctimas; circunstancias relacionadas con los menores; y circunstancias agravantes y aspectos subjetivos del propio riesgo de la víctima.

Se utilizó un diseño epidemiológico, analítico observacional de casos y controles que permite la construcción de un modelo de nivel predictivo basado en la estadística bivariante para conocer las odds ratio (*OR*) de los factores predictores de homicidios de mujeres por su pareja, y procedimientos multivariantes para analizar la confusión de medidas. Se establecieron diversas matrices de datos para obtener los factores de riesgo y la validación de la escala H, adoptando medidas de control para evitar duplicidades. Los factores de riesgo obtenidos debían reunir como requisitos de validez demostrar su asociación con el desenlace mortal, obteniendo su propio peso matemático (*OR*), sin importar que fuera diferente al utilizado para la recaída de menor intensidad; ser compartidos con el instrumento base ya utilizado (VPR5.0) para evitar que los agentes tuvieran que realizar dos evaluaciones; y el modelo predictivo definitivo debería mejorar el actual VPR5.0.

Los datos se analizaron tomando como variables independientes los factores de riesgo de VPR5.0, y como variable dependiente el resultado fatal; utilizando tablas de contingencias y estadísticos específicos para determinar la significación estadística; destacando el uso de las curvas ROC para la validez predictiva.

Los resultados indicaron que 13 de los 35 factores de riesgo presentes en el instrumento se asociaron significativamente con el homicidio de mujeres por su pareja, con un tamaño del efecto diferente al de no homicidio de mujeres por su pareja (con valores *OR* entre 1.507 y 8.087). La regresión logística binaria mostró seis factores significativos que clasificaban correctamente el 86,3% de los homicidios de mujeres por su pareja. Destaca especialmente el hecho de que cuatro factores de la esfera de la víctima fueron relevantes, lo que no se había observado en estudios anteriores respecto a la predicción del riesgo de reincidencia no letal, y que confirmaría las diferencias señaladas entre la violencia y el homicidio en el ámbito de la pareja, asumiendo no obstante la existencia de factores de riesgo comunes en ambas formas de violencia. Destaca también la aparición de nuevas circunstancias relevantes para la predicción del homicidio de mujeres por su pareja no detectadas en estudios anteriores: las amenazas de suicidio del agresor; la violación de órdenes de alejamiento; la ideación e intentos de suicidio; el hecho de que la víctima padezca algún tipo de enfermedad mental y la experiencia de homicidio de mujeres por su pareja en el entorno familiar de la víctima. Así, los autores indican que las



circunstancias clínicas del agresor (especialmente en la esfera parasuicida) y de la víctima parecen ser importantes en este tipo de delitos.

Los hallazgos sugieren que otros factores como el historial de quejas sobre episodios previos de violencia, la presencia de violencia física o sexual y el uso de armas no resultaron ser un factor diferenciador entre un episodio repetido no letal y un homicidio de mujeres por su pareja, atribuyendo como posible explicación que la muestra de homicidio de mujeres por su pareja utilizada en el estudio procede no solo de casos con denuncias previas por violencia (27,6%), sino también de otros que presentan características diferentes, incluyendo un 22% de casos de homicidio seguidos de suicidio.

Así, la Escala H compuesta por los 13 factores significativamente asociados al homicidio de mujeres por su pareja, presenta una sensibilidad del 84% y una especificidad del 60%, que los autores estiman como valores adecuados para este contexto de aplicación; significando que a diferencia de estudios similares como los desarrollados para VPR mediante la estimación de la reincidencia, en su estudio la mayoría de los casos no se encontraban bajo protección policial u otras circunstancias que pudieran modificar los resultados de validez predictiva del instrumento.

Señalan interesantes implicaciones prácticas de las conclusiones del estudio. Así, los resultados indican la conveniencia de un doble mecanismo de medición para mejorar las estimaciones combinadas de reincidencia y homicidio, lo que se consigue con el VPR5.0 para la reincidencia y la Escala H para el homicidio, aportando un mayor grado de protección para los casos identificados como más graves, estableciendo la denominada VPR5.0-H que proporciona dicha información de manera independiente e integrada en un único informe de riesgo. Por otra parte, la nueva herramienta es más precisa respecto a la predicción de casos potenciales de violencia letal considerando la posibilidad de que no haya habido denuncia previa.

Proponen una valoración complementaria del riesgo por parte de los profesionales del equipo forense mientras se gestiona el riesgo por parte de los agentes de policía. La nueva herramienta está actualmente calibrada para su uso en el ámbito policial, donde prima la evaluación de la gestión del riesgo. Sin embargo, en otros ámbitos profesionales, como el de la Justicia, su uso podría beneficiarse de una recalibración.

Como limitaciones señalan que los casos de homicidio utilizados pueden no haber incluido toda la casuística de un fenómeno tan complejo; la existencia de otro tipo de indicadores importantes asociados al fenómeno que no han sido tenidos en cuenta porque no son accesibles para los agentes de policía; y el hecho de que se trata de una herramienta informática utilizada por fuerzas policiales de España en el marco del sistema VioGén, lo que limita la replicabilidad de sus resultados.

Finalizan indicando que las investigaciones futuras deberán estudiar estos aspectos para determinar si existen otro tipo de factores de riesgo que permitan detectar y prevenir estos hechos, especialmente en contextos sanitarios o sociales, que faciliten la protección de las mujeres que no denuncian a la policía.

*Estudio 7: An Evaluation of the Predictive Validity of the SAVRY and YLS/CMI in Justice-Involved Youth With Fetal Alcohol Spectrum Disorder (McLachlan, Gray, Roesch, Douglas y Viljoen, 2018)*

---

Este estudio investiga el uso de herramientas de riesgo de evaluación de la violencia en población penal juvenil afectada por trastornos del espectro del denominado síndrome alcohólico fetal (SAF), concretamente la validez predictiva del instrumento SAVRY (Evaluación Estructurada del Riesgo de Violencia en la Juventud) y el Nivel de Servicio de la Juventud-Inventario de Gestión de Casos (YLS/CMI).

Los autores revisan en primer lugar los aspectos que caracterizan el trastorno y su relación con la conducta y el entorno social de los afectados, quienes presentan unas mayores tasas de factores de riesgo criminógenos y de probabilidad de mostrar conductas delictivas y terminar en contacto con el sistema penal. Aducen que la falta de formación para reconocer o diagnosticar esta discapacidad lleva a una cierta invisibilidad del trastorno.

Justifican su trabajo indicando que a pesar del acuerdo científico respecto a la utilidad y validez de los instrumentos mencionados para la evaluación del riesgo de violencia en población juvenil y su amplia difusión en este ámbito, no existen estudios sobre su uso en jóvenes con SAF. Revisan brevemente algunos trabajos sobre la validez del SAVRY y el YLS/CMI con poblaciones juveniles afectadas de distintas discapacidades y psicopatologías (discapacidad intelectual, TDAH, trastornos de conducta y mentales), señalando que no existe una relación clara entre la posible generalización de estos resultados a los jóvenes delincuentes afectados por SAF. Añaden además la necesidad de examinar la validez de las interpretaciones de las puntuaciones de riesgo mediante el uso de estas herramientas en esta población para cumplir los criterios de admisibilidad de pruebas en los tribunales norteamericanos; y señalan las consecuencias que sobre los propios afectados o la sociedad pueden tener los juicios inexactos sobre el riesgo.

La pregunta de investigación planteada en el estudio se centró en evaluar si las puntuaciones totales de SAVRY y YLS/CMI, las calificaciones de riesgo y los juicios de riesgo profesionales estructurados podrían predecir la aparición de nuevos cargos durante un período de seguimiento de un año; estableciendo la hipótesis de que las calificaciones de riesgo y las puntuaciones serían más altas en los jóvenes involucrados en la justicia con SAF en comparación con aquellos que no lo padecieran.

La muestra consistió en 100 jóvenes involucrados en el sistema judicial divididos en dos grupos de 50 (afectados con SAF y no afectados), con edades comprendidas entre 12 y 23 años, reclutados en dos áreas urbanas canadienses. Para la selección de afectados con SAF se revisaron los informes clínicos para confirmar el diagnóstico de acuerdo a los criterios aceptados, previa remisión de expedientes de delinquentes juveniles procedentes de distintas agencias y actores del sistema judicial canadiense; en tanto que los participantes del grupo de comparación fueron reclutados en las oficinas de probación y en los centros penitenciarios ubicados en las mismas regiones geográficas y seleccionados mediante autoinforme y revisión de expedientes.

Las medidas utilizadas incluyeron el historial delictivo y de reincidencia, la Escala Abreviada de Inteligencia de Wechsler (WASI, 1999), la Evaluación Estructurada del Riesgo de Violencia en los jóvenes (SAVRY) y el Nivel de Servicio de los Jóvenes/Inventario de Gestión de Casos YLS/CMI. Esta es una herramienta actuarial de evaluación de riesgo de reincidencia y ayuda en la planificación de gestión de los casos, que clasifica el riesgo en cuatro niveles (bajo, moderado, alto y muy alto) en función del análisis de 42 factores de riesgo presente/ausente en ocho dominios; mostrando una buena validez predictiva.

Para el análisis de datos se utilizaron diversos estadísticos, destacando el cálculo ROC para la evaluación de la validez predictiva.

Destacan en los resultados obtenidos que los jóvenes del grupo SAF registraron su primera acusación oficial aproximadamente un año antes que el grupo de comparación, mostraban un coeficiente intelectual significativamente más bajo, un mayor número de contactos con el sistema penal, una mayor probabilidad de reincidencia violenta durante el período de seguimiento y con menor tiempo entre delitos. En lo referente a las puntuaciones obtenidas con los instrumentos del estudio, resultó que en la mayoría de los dominios y en las puntuaciones totales de ambos, el grupo de afectados por SAF obtuvo puntuaciones significativamente más altas; excepto en las referentes a los factores de protección, en las que fueron significativamente más bajas. Los resultados sugieren que ni el SAVRY ni el YLS/CMI predijeron la reincidencia de manera diferente para los grupos con SAF y los grupos de comparación, ya sea a nivel de calificación de riesgo o de puntuación total de riesgo, pero un hallazgo interesante fue la relación observada entre los factores de protección del SAVRY y la mayor asociación con el desistimiento de la reincidencia violenta y general en los jóvenes con SAF, incluso después de controlar el tiempo pasado en custodia, siendo el predictor más consistente de la ausencia de reincidencia a través de los análisis. Así, a pesar de que los jóvenes con SAF tienen menos factores de protección en relación con los jóvenes del grupo de comparación la asociación mencionada subraya la necesidad de identificar y reforzar los factores de protección para esta población, en línea con lo propuesto en otras investigaciones centradas en el papel clave de los factores de protección, la resiliencia y las fortalezas para las personas con SAF en el

sistema de justicia penal así como la investigación que vincula los factores de protección, como la evaluación temprana, el diagnóstico y la intervención, con resultados más positivos.

Los autores concluyen manifestando que este estudio resulta pionero en la detección del fenómeno de la "puerta giratoria" respecto a los jóvenes delincuentes con SAF, como muestra su hallazgo de su mayor y más rápida reincidencia durante el período de seguimiento, poniendo de relieve la necesidad de mayor investigación para mejorar la evaluación del riesgo en esta población.

Señalan como puntos fuertes del estudio el diseño prospectivo y la incorporación de información de entrevistas y revisión de expedientes, más próximos al uso clínico de los instrumentos utilizados; así como el hecho de contrastar a jóvenes con SAF respecto a otros que han experimentado similares factores de riesgo de desarrollo y ambientales en ausencia del trastorno. Como limitaciones indican el modesto tamaño de la muestra y la imposibilidad de que los calificadores fueran ciegos a la detección del trastorno, con el sesgo que ello pudiera implicar.

Los resultados de este estudio sugieren que las herramientas de evaluación del riesgo, como el SAVRY y el YLS/CMI, pueden proporcionar información importante sobre el riesgo de reincidencia de un joven y, de manera crítica, información potencialmente importante necesaria para desarrollar planes de gestión comprensivos (aunque esto último queda por probar empíricamente).

Finalizan sugiriendo como orientaciones futuras la cuestión de si los jóvenes involucrados en la justicia con SAF requieren servicios especializados y protocolos de evaluación, o si los modelos actuales de evaluación e intervención pueden ser utilizados para hacer frente a sus necesidades (por ejemplo, utilizando el SAVRY y YLS / CMI), y la necesidad de mayor investigación para entender si se requiere una formación específica sobre el trastorno del espectro de síndrome fetal alcohólico para asegurar que las herramientas de riesgo como el SAVRY y el YLS/CMI resulten en una evaluación de riesgo eficaz y en la planificación de la gestión del caso, incidiendo especialmente en la cuestión referente a si los jóvenes con SAF implicados en la justicia responden igualmente a los enfoques de gestión e intervención diseñados para reducir el riesgo y tratar sus necesidades clínicas. Las evaluaciones de riesgo que se llevan a cabo con el único propósito de hacer predicciones sobre el riesgo, sin elaborar planes de gestión e intervención informados sobre este, pueden dar lugar a que muchos jóvenes con SAF implicados en la justicia sean considerados de alto riesgo de reincidencia. Esta circunstancia, unida a la falta de implementación de planes y apoyos informados, puede conducir a una respuesta pobre a la "gestión habitual" y a tasas potencialmente más altas de cargos de la administración de justicia, reincidencia y encarcelamiento, y en última instancia contribuir a el problema de la "puerta giratoria" en los jóvenes con SAF.

Estudio 8: *Predictive Validity of the Structured Assessment of Violence Risk in Youth (SAVRY) With Indigenous and Caucasian Female and Male Adolescents on Probation* (Muir, Viljoen, Jonnson, Cochrane y Rogers, 2020)

---

Este estudio de campo prospectivo examina la validez predictiva de las calificaciones del instrumento SAVRY en una muestra compuesta de jóvenes canadienses caucásicos e indígenas americanos en libertad condicional, con el fin de comprobar si son adecuadas para este último grupo poblacional, que pudiera estar sobrerrepresentado en el sistema de justicia canadiense en atención a los datos estadísticos al respecto, que indican que en 2107-2018 su tasa de encarcelamiento es seis veces mayor que la de la población juvenil general, en tanto que representan únicamente el 8% de la población del país. Para el desarrollo del estudio se tuvieron en cuenta las consultas y aportaciones de expertos de los pueblos indígenas

Inician el trabajo analizando las posibles causas de esta sobrerrepresentación, en especial el colonialismo y sus consecuencias, como las experiencias negativas de los internados en los que muchos niños indígenas fueron ingresados a la fuerza y su relación con la discriminación sistémica, la marginación, el bajo nivel educativo, entre otros factores.

Justifican la pertinencia de su estudio examinando la validez predictiva de las puntuaciones del instrumento en relación a otras poblaciones, y destacando que únicamente se han realizado dos estudios que han examinado el SAVRY en poblaciones indígenas, de los que destacan importantes limitaciones como el pequeño tamaño de las muestras y su carácter de estudios de revisión de archivos y no de estudios de campo; sugiriendo que por tanto existen lagunas en la investigación en este campo, como la ausencia de análisis de la naturaleza y los tipos de errores cometidos en las predicciones de riesgo para los jóvenes indígenas. Indican que es posible que los tipos de error difieran para los jóvenes indígenas y no indígenas, incluso si las puntuaciones del SAVRY tienen una validez predictiva adecuada para los jóvenes indígenas en general; pudiendo estar afectados por el sesgo de estereotipos incorrectos de los agentes de libertad condicional, produciendo un aumento de tasas de falsos positivos. Añaden la escasez de investigación sobre la validez predictiva del SAVRY respecto a las mujeres en general (con resultados contradictorios) y mujeres indígenas en especial, con tasas de sobrerrepresentación en el sistema penal superiores incluso a las de los hombres indígenas.

A lo anterior añaden otras circunstancias, como el tratamiento de la cultura en las investigaciones como una variable dicotómica a pesar de la enorme variabilidad entre los jóvenes indígenas en este aspecto; y que pocos estudios han comprobado si los jóvenes indígenas y no indígenas tienen puntuaciones y perfiles diferentes en el SAVRY.

Para lograr el objetivo del estudio (comparar la validez predictiva de las puntuaciones SAVRY para los jóvenes indígenas y caucásicos en libertad condicional), mediante un diseño de estudio de campo prospectivo los autores atendieron a cuestiones tales como probar la validez predictiva por separado para mujeres y hombres; si la validez predictiva para los jóvenes indígenas variaba dependiendo de si los jóvenes vivían en la reserva o fuera de ella; examinaron no sólo los índices de discriminación sino también el poder predictivo positivo y negativo; y probaron las diferencias medias en las puntuaciones totales y de dominio del SAVRY para los jóvenes indígenas y caucásicos. Además, utilizamos un, en el que los OPJ evaluaron a los jóvenes con el SAVRY.

La hipótesis era que las puntuaciones del SAVRY mostrarían una menor validez predictiva y mayores tasas de falsos positivos para los jóvenes indígenas en comparación con los jóvenes caucásicos. Se esperaba también que el SAVRY tuviera una menor validez predictiva para las mujeres en comparación con los hombres, y para los jóvenes indígenas que viven en la reserva en comparación con los que viven fuera de ella. Predijeron que los jóvenes indígenas tendrían una puntuación más alta en los factores de riesgo y más baja en los factores de protección que los jóvenes caucásicos.

La muestra consistió en 744 jóvenes en libertad condicional en una provincia canadiense, incluyendo 299 jóvenes indígenas (80 mujeres, 219 hombres) y 445 jóvenes caucásicos (88 mujeres, 357 hombres), con una edad media de los jóvenes en el momento de la evaluación de 17,04 ( $SD = 1,33$ , rango = 12 a 20). Casi el 60% de los jóvenes tenían cargos previos y el 31,3% habían sido encarcelados anteriormente. El grupo de indígenas pertenecía a tres pueblos distintos (Primeras Naciones, Inuit y Métis), en tanto que el grupo de comparación de caucásicos lo componían jóvenes cuyos padres eran blancos.

Para la extracción y codificación de datos se recurrió a la base de datos provincial de justicia juvenil por un grupo de asistentes de investigación previamente formados, con medidas de control de la fiabilidad de los evaluadores. Se utilizaron calificaciones de evaluación con SAVRY, información demográfica y de reincidencia tanto violenta como no violenta.

Se comprobaron las diferencias de grupo en las puntuaciones SAVRY y la reincidencia utilizando pruebas  $t$  y análisis chi-cuadrado, los índices de discriminación mediante pruebas ROC, los valores predictivos positivos y negativos y el análisis de supervivencia para la reincidencia.

Los resultados indicaron que tanto en varones como en mujeres indígenas se obtuvieron puntuaciones significativamente más altas tanto totales de riesgo como en los dominios de riesgo que el grupo control, pero a la vez significativamente más bajas en los factores de protección; así como una mayor propensión a ser calificados como de alto riesgo y acusados de cualquier tipo de reincidencia. En la predicción de la reincidencia violenta, se identificaron correctamente más mujeres indígenas como de

alto riesgo (88,9%) en comparación con las mujeres caucásicas (62,5%). Sin embargo, menos mujeres indígenas fueron identificadas correctamente como bajo riesgo (46,4%) en comparación con las mujeres caucásicas (72,3%). Los patrones de sensibilidad y especificidad para los hombres fueron similares, aunque no tan exagerados. Datos similares se obtuvieron en ambos sexos respecto a la predicción de reincidencia general. No se encontraron diferencias significativas entre los jóvenes indígenas que vivían en reservas y fuera de ellas.

Se encontró que una mayor proporción de jóvenes indígenas de alto riesgo volvió a delinquir violentamente en comparación con los jóvenes caucásicos de alto riesgo. Por el contrario, los resultados fueron ligeramente inferiores para las mujeres y los hombres indígenas respecto a las mujeres y los hombres caucásicos. Un mayor número de jóvenes indígenas de bajo riesgo reincidió en la violencia respecto a los jóvenes caucásicos de bajo riesgo. Los patrones fueron similares para las predicciones de cualquier reincidencia. Respecto a los falsos positivos y falsos negativos cabe señalar que se halló una asociación significativa entre el origen étnico y la reincidencia violenta para los jóvenes indígenas de alto riesgo (tenían más del doble de probabilidades de reincidir violentamente que los jóvenes caucásicos). Contrariamente, en el caso de los jóvenes de bajo riesgo, la asociación entre el origen étnico y la reincidencia violenta no fue significativa, pero sí lo fue en la reincidencia general (tenían más del doble de probabilidades de reincidir en cualquier caso en comparación con los jóvenes caucásicos).

Del análisis de los resultados los autores concluyen que el SAVRY predijo significativamente la reincidencia violenta y de cualquier tipo entre las mujeres y los hombres indígenas y caucásicos. Contrariamente a las hipótesis de partida, surgieron relativamente pocas diferencias significativas en la validez predictiva de las puntuaciones SAVRY entre los jóvenes indígenas y caucásicos; destacando el dominio histórico como predictor más fuerte de cualquier reincidencia para los varones caucásicos que para los varones indígenas, sugiriendo que podría deberse a que el SAVRY no capta los factores de riesgo históricos relacionados con el colonialismo (es decir, el trauma histórico). Alternativamente, los factores históricos pueden no ser tan eficaces para diferenciar a los varones indígenas que reinciden y a los que no reinciden. Los factores de protección también predijeron cualquier reincidencia significativamente mejor para los hombres caucásicos que para los indígenas. Hallaron también que el poder predictivo positivo fue mayor en los jóvenes indígenas que en los caucásicos, y que la validez predictiva era comparable para los jóvenes indígenas que vivían dentro y fuera de reservas.

Como limitaciones más importantes al estudio señalan los posibles sesgos derivado del uso de registros oficiales; el escaso tamaño de la muestra de jóvenes que vivían en reserva; y la imposibilidad de examinar la validez predictiva de otros grupos minoritarios.

Como implicaciones para la práctica del estudio señalan la importancia de la consideración de los factores culturales específicos de esta población por los evaluadores, incluyendo la alta variabilidad de

diferencias culturales, tanto para los juicios de riesgo como para la formulación de casos específicos y las recomendaciones de tratamiento. En cuanto a las implicaciones para futuras investigaciones, señalan también el importante papel de los predictores culturales destacados y la necesidad de mayor investigación respecto a cómo pueden afectar a la reincidencia en esta población. Sugieren asimismo la necesidad de evaluar la utilidad de varias herramientas de evaluación del riesgo para conceptualizar y manejar el riesgo en los jóvenes indígenas; e investigar la influencia de los sesgos psicométrico y del evaluador en los procesos de evaluación atendiendo a las diferencias culturales.

Finalizan remarcando la necesidad de un mayor esfuerzo para mejorar la comprensión y evaluación de los factores que contribuyen a la delincuencia entre los jóvenes indígenas mediante la búsqueda comunitaria y la consulta con los pueblos indígenas.

*Estudio 9: A twist in Intimate Partner Violence Risk Assessment Tools: Gauging the contribution of exogenous and historical variables (Quijano-Sánchez, Liberatore, Rodríguez-Lorenzo, Lillo y González-Álvarez, 2021)*

---

Trabajo centrado en el análisis del uso de un modelo basado en aprendizaje automático y ciencia de datos para mejorar la capacidad predictiva del riesgo de violencia del sistema VioGén utilizado por la Secretaría de Estado de Seguridad del Ministerio del Interior estudiando el valor predictivo de variables exógenas e históricas.

Se justifica por la necesidad de minimizar la posibilidad de reincidencia en los casos de violencia de pareja a través de la determinación del nivel de protección más adecuado para las víctimas, en tanto que dicho nivel de protección está asociado al riesgo de reincidencia. Pretende aprovechar las capacidades que ofrecen las técnicas de ciencia de datos y aprendizaje automático que se vienen utilizando con éxito en otros ámbitos para evaluar el actual sistema VioGén.

Los autores parten de la premisa de que, con el objetivo de garantizar la seguridad de la víctima y la no reincidencia del agresor, el plan de protección debe ajustarse al máximo para optimizar los recursos disponibles maximizando la precisión del modelo al tiempo que minimizando la subestimación de los planes de riesgo.

Los datos fueron extraídos de la base de datos del sistema VioGén con la evolución de 46.047 nuevos casos registrados entre octubre de 2016 y diciembre de 2017.

En la primera sección del trabajo se revisan los fundamentos metodológicos de los instrumentos de evaluación de riesgo de violencia señalando que están enfocados desde el campo de las ciencias sociales, y señalan la naturaleza dinámica del riesgo y procesual de su evaluación, para lo cual resulta



idóneo el enfoque propuesto; y revisa en detalle el funcionamiento del sistema VioGén sobre el que se propone el uso de las nuevas herramientas. Concluyen este apartado señalando las limitaciones encontradas (la naturaleza estática de los enfoques desarrollados hasta ahora;) la falta de homogeneidad en las definiciones de reincidencia y los planes de protección más adecuados asociados; y la falta de diversidad en los modelos de predicción estudiados, de carácter actuarial.

La metodología utilizada pretende abordar dichas limitaciones estudiando el impacto de la evaluación del historial del caso así como de las variables exógenas; definiendo un nuevo paradigma para el cálculo del plan de protección más adecuado, asociándolo a las ventanas temporales de reincidencia; y abordando la identificación del plan de protección más adecuado mediante métodos de aprendizaje automático; todo ello centrándose únicamente en el modelo predictivo del Formulario de Evolución de la Evaluación de Riesgo Policial (VPER).

Se definen las variables de entrada del sistema, añadiendo a las utilizadas por el VPER (información sobre el formulario), información exógena que incorpora información relativa al caso que no forma parte de dicho formulario, e información histórica que incluye variables que representan la evolución del caso. A continuación, se define el denominado “nivel óptimo de protección”, cuyo objetivo es calcular directamente el plan de protección más adecuado para cada caso, minimizando la probabilidad de reincidencia y optimizando los recursos policiales destinados a ello, evitando los riesgos de reincidencia entre períodos de revisión de los planes. Abordan este doble problema mediante la aplicación de modas de aprendizaje automático con algoritmos matemáticos para ajustar la variable de respuesta capaces de estimar correctamente el riesgo de infravaloración.

En el diseño de la investigación se realizó en primer lugar una depuración de los datos (eliminación de duplicados y de casos incompletos, comprobación y corrección de problemas de coherencia en las respuestas de los formularios), un análisis de los distintos subtipos de reincidencia en función de la ventana temporal, el nivel de protección de VioGén (destacando que el porcentaje de acierto y subestimación es de 80,57% y 15,54%, respectivamente) y el nivel de protección aplicado.

Las preguntas de investigación se centraron en la posibilidad de mejora en el rendimiento del modelo propuesto respecto al actual mediante la inclusión de variables exógenas e históricas, y la existencia de diferencias significativas entre el uso de un modelo multiclase y otro ordinal en el problema considerado.

En el conjunto de datos seleccionados se aplicaron diferentes modelos matemáticos de aprendizaje automático, siendo el modelo XGBoost el que proporcionó el mejor rendimiento. Los resultados permitieron a los investigadores determinar que la inclusión de las variables históricas da lugar a una mejora significativa del rendimiento del modelo, mientras que añadir sólo la información exógena

no produce un modelo que sea significativamente mejor que VioGén, aunque sí mejora significativamente el rendimiento de un modelo.

Así, dan respuesta a sus preguntas de investigación concluyendo que: (a) sí existen diferencias significativas entre utilizar un modelo multiclase y uno ordinal en el problema considerado, (b) la inclusión de variables exógenas no supone una mejora en el rendimiento del modelo, aunque sí una mejora cuando se combina con datos históricos, y (c) la inclusión de variables históricas supone una mejora en el rendimiento del modelo.

Como aportaciones más relevantes de su investigación destacan implicaciones inmediatas para los sistemas policiales predictivos, con la introducción de un nuevo paradigma para el diseño de instrumentos de evaluación de riesgos de violencia en la pareja mediante el cálculo directo del nivel óptimo de protección en vez de asignar un plan de protección basado en la reincidencia; y la mejora que supone la introducción de técnicas de aprendizaje automático en el modelo, que habría corregido más del 25% de los casos originalmente infravalorados por el sistema VioGén, al tiempo que mejora la precisión global del mismo, con la importancia que ello supone atendiendo a que el plan de protección está directamente correlacionado con el riesgo que pueda sufrir la víctima y la urgencia de su protección.

Como líneas futuras de investigación más destacables proponen la importancia de seguir buscando variables exógenas que representen el entorno en el que se produce el caso (como la tasa de paro, la tasa de criminalidad de la localidad en la que se produce el incidente, los informes de las prisiones o la información de los casos que se archivan judicialmente), generar un conocimiento más detallado de la evolución de los hechos, definir con mayor precisión los rangos temporales de revisión de los planes de protección.

---

Estudio 10: *Predictive Validity of the SAVRY, YLS/CMI, and PCL:YV Is Poor for Intimate Partner Violence Perpetration Among Adolescent Offenders* (Shaffer, Viljoen y Douglas, 2022)

Este estudio longitudinal evalúa la validez predictiva de tres instrumentos de predicción de riesgo de violencia en población juvenil (SAVRY, YLS/CMI y PCL:YV) en el ámbito de la evaluación y gestión del riesgo de violencia de pareja entre los adolescentes, mediante el análisis de factores tales como la edad, el género, la raza/etnia y los antecedentes de violencia de pareja como variables de control y el examen de las puntuaciones totales y de subescala de SAVRY y YLS/CMI y las puntuaciones de resumen de riesgo y las puntuaciones totales y de subescala de PCL:YV para comprobar si eran predictivas de la perpetración de violencia de pareja durante un seguimiento de 2 años.

Tras una breve introducción en la que se discuten las consecuencias físicas y psicológicas para las víctimas de la violencia de pareja entre los adolescentes y el interés social y científico en lograr su reducción mediante diversos enfoques y estrategias, destacan el uso de instrumentos de predicción y evaluación de riesgo de violencia como una de las mejores prácticas para determinar el potencial de un individuo para participar en la violencia y luego implementar estrategias para controlar o mitigar este potencial, destacando el hecho de que a pesar del elevado número de instrumentos desarrollados (más de 400), en la actualidad no existen directrices estructuradas para evaluar el riesgo de los adolescentes de sufrir violencia de pareja, lo que tiene como consecuencia que los profesionales encargados de identificar a los jóvenes en riesgo de perpetrar violencia de pareja tienen poca orientación en esta área para garantizar que están evaluando la información pertinente y aplicando eficazmente las intervenciones y los tratamientos adecuados.

Se describen brevemente los fundamentos de los instrumentos seleccionados, apuntando que la evidencia sugiere que pueden ser relevantes para el ámbito de la violencia de pareja entre adolescentes dada la coincidencia entre factores de riesgo compartidos para la perpetración de violencia de pareja, violencia general y delincuencia que son medidos en dichos instrumentos; si bien la mayoría de las investigaciones se han centrado en su precisión predictiva con respecto a cualquier tipo de violencia o cualquier delito, pero no respecto a la comisión de violencia de pareja específicamente.

Los autores señalan no disponer de hipótesis específicas sobre la validez incremental y predictiva de estas herramientas para la perpetración de la violencia de pareja, dada la escasez de investigaciones al respecto.

La muestra estaba compuesta por 163 adolescentes con edad media de los participantes de 16,41 años ( $SD = 1,14$ ), y criterios de inclusión: tener entre 12 y 18 años; estar incluidos en el sistema de justicia penal juvenil; estar bajo supervisión comunitaria activa; residentes en el Distrito Regional de Vancouver de la Columbia Británica, Canadá; y ser capaces de hablar un inglés fluido. La composición de género y racial/étnica fue similar a la reportada en las estadísticas nacionales de delinquentes adolescentes canadiense, por lo que se sugiere que fue suficientemente representativa en estos aspectos.

La selección de los participantes se realizó por asistentes de investigación previamente entrenados mediante entrevistas semiestructuradas, cuestionarios de autoinforme, información de archivos y compleción de los instrumentos de evaluación con control de la fiabilidad entre evaluadores.

Las medidas utilizadas fueron las características demográficas (género, edad y raza/etnia); el historial de violencia de pareja a lo largo de la vida; los instrumentos SAVRY, YLS/CMI y PCL:YV; y los resultados de reincidencia obtenidos de los registros oficiales durante 2 años en tres modalidades (violencia contra la pareja, delitos violentos y delitos en general).

El análisis de datos de la validez predictiva se realizó mediante la fuerza de las asociaciones entre el SAVRY, el YLS/CMI y el PCL:YV y la perpetración de violencia de pareja y cualquier delito violento y cualquier delito a lo largo de los 2 años de seguimiento utilizando análisis ROC. Para evaluar la validez predictiva incremental se emplearon modelos de regresión logística jerárquica, y se estimó la potencia mediante diversos procedimientos estadísticos para comprobar si el tamaño de la muestra actual era adecuado para los análisis bivariados y de regresión incremental planificados.

Los resultados mostraron que, para las puntuaciones totales, en todos los instrumentos fueron significativamente predictivas de reincidencia violenta y general, pero no significativa para la violencia de pareja. Lo mismo ocurrió con las subescalas o las calificaciones de riesgo resumidas. Ninguna de las tres herramientas proporcionó una validez incremental sobre las variables de control demográficas y la historia auto informada de perpetración de violencia de pareja o victimización en la predicción de la futura perpetración de violencia de pareja, lo que sugiere que el triaje para los creadores de riesgo estático (por ejemplo, preguntar a los jóvenes sobre la violencia de pareja anterior) podría ser un mejor punto de partida para los profesionales para evaluar el riesgo de violencia de pareja.

Puntualizan no obstante los autores que las medidas utilizadas por los instrumentos pueden no captar factores de riesgo únicos, ya que no fueron específicamente diseñadas para ello, por lo que las dinámicas de relación negativas entre los agresores y las víctimas potenciales (por ejemplo, el conflicto verbal, los celos, el engaño) y las características que interfieren con la voluntad, la capacidad o la oportunidad de las víctimas de adoptar un comportamiento de autoprotección (por ejemplo, un apoyo inadecuado) pueden tener que ser consideradas además de factores de riesgo del agresor para proporcionar evaluaciones más precisas del riesgo de violencia de pareja. Sugieren por tanto que la evaluación del riesgo de violencia de pareja adolescente mejoraría si se evaluaran separadamente los factores de riesgo de violencia general y de cualquier delito y los factores de riesgo específicos de la violencia de pareja.

Añaden que, aunque los predictores de las diferentes formas de violencia pueden ser similares, las definiciones operativas pueden variar en función del tipo de violencia en cuestión, lo que indicaría la necesidad de un conocimiento especializado sobre la violencia de pareja para refinar las definiciones de los ítems y las instrucciones de codificación incluidas en los instrumentos para producir una evaluación precisa del riesgo de violencia de pareja.

Señalan que sus hallazgos son coherentes con los resultados comunicados por los estudios que examinan la utilidad de las herramientas generales de violencia y delincuencia entre los adultos que ejercen la violencia de pareja. Sugieren que dada la evidencia de un solapamiento de los factores de riesgo de la violencia de pareja en diferentes grupos de edad y que la manifestación de la violencia de pareja en las relaciones íntimas de los adolescentes es en cierto modo análoga a la manifestación de la violencia de

pareja entre los adultos, las herramientas de violencia de pareja de los adultos pueden tener cierta utilidad con aquellos también.

Sin embargo, puntualizan que las diferencias de desarrollo entre adultos y adolescentes pueden hacer que el uso de las herramientas de evaluación de los adultos con los jóvenes sea problemático, ya que pueden juzgar erróneamente comportamientos que son comunes en la adolescencia (como la propensión a participar en conductas de riesgo y ser menos capaces de utilizar habilidades de relación saludables para manejar conflictos). A ello añaden las importantes diferencias en la presencia y relevancia de factores específicos en las distintas etapas de desarrollo, que implican una mayor influencia de los factores ambientales en los adolescentes, por lo que puede ser necesario enfatizar estos factores en las evaluaciones. Por otra parte, aducen que la inclusión de los factores de protección en las evaluaciones de los adolescentes podría proporcionar una imagen más equilibrada de ellos, evitar los efectos de etiquetado y permitir el desarrollo de intervenciones más eficaces.

Señalan como limitaciones más importantes: la posibilidad de infradenuncia de casos de violencia de pareja en los adolescentes, que tendría su reflejo en los registros oficiales que se utilizaron como fuente de datos, por lo que sería conveniente utilizar un método multifuente para la obtención de datos; el pequeño tamaño de la muestra y la baja tasa base de violencia de pareja en ella (11,5%), que podría limitar la potencia de los análisis; la imposibilidad de comprobar las diferencias entre subgrupos en la precisión predictiva de los tres instrumentos para la perpetración de la violencia de pareja; las distintas definiciones empleadas para operacionalizar el historial de perpetración de violencia de pareja a lo largo de la vida y durante el período de seguimiento, lo que puede producir una estimación conservadora de la tasa de prevalencia de violencia de pareja anterior; y la antigüedad de los datos, recogidos entre 2008 y 2012.

Concluyen manifestando que sus hallazgos sugieren que el uso de las herramientas existentes para adolescentes para evaluar y gestionar el riesgo de violencia de pareja podría ser problemático. Así, el SAVRY y el YLS/CMI proporcionan una imagen amplia de las características y circunstancias de un delincuente adolescente, lo que limita su capacidad para pronosticar con precisión el riesgo de perpetración de violencia de pareja. Además, indican que los rasgos psicopáticos son sólo un potencial predictor de la perpetración de violencia de pareja, y que los adolescentes perpetradores de violencia de pareja no tienen por qué obtener puntuaciones altas en la PCL:YV; por lo que sería necesario el desarrollo de instrumentos específicos para los adolescentes en este ámbito.

Estudio 11: *The Validity of Risk Assessment Instruments for Transition-Age Youth* (Vincent, Drawbridge y Davis, 2019)

---

Trabajo dedicado a investigar la validez predictiva diferencial de instrumentos de evaluación de riesgo de adultos HCR-20 y VRAG, y de jóvenes YLS/CMI y SAVRY, en población juvenil en edad de transición (definida por los autores entre la adolescencia tardía y los 25 años) mediante comparación de tres grupos de edad: 16 a 24 años, 12 a 15 años y 25 a 40 años.

Se justifica la elección de esta población diana porque en el período de transición los jóvenes son más parecidos a los adultos en algunos aspectos como la toma de decisiones y más parecidos a los adolescentes en otros, como la impulsividad; circunstancia que motiva que la línea para tratarlos en el sistema de justicia penal sea difusa y heterogénea, por lo que resulta pertinente examinar la validez del uso de los enfoques actuales de evaluación de riesgos con jóvenes en ese rango de edad.

Tras realizar un repaso a los principales enfoques seguidos en los instrumentos estandarizados de evaluación de riesgo (actuarial y de juicio clínico estructurado), los autores plantean algunas cuestiones de interés al respecto, como la variación de la importancia relativa de los factores de riesgo y las curvas de actividad delictiva en función de la edad; y hacen notar que en general los estudios al respecto no han examinado específicamente el período de transición.

Para el estudio se combinaron conjuntos de datos secundarios obtenidos de otros investigadores para las dos evaluaciones de riesgo para adolescentes más ampliamente validadas (SAVRY e YLS/CMI) y dos de las evaluaciones de riesgo para adultos mejor validadas (HCR-20 y VRAG). Se definió el grupo objetivo (jóvenes en edad de transición) incluyendo las edades de 16 a 24 años, un segundo grupo de adolescentes (entre 12 y 15 años) y otro de adultos entre 25 y 41 años, excluyéndose los mayores de esta edad por sus bajas tasas de reincidencia y puntuaciones en las herramientas de evaluación de riesgos en relación con otros adultos, lo que podría sesgar los resultados de las comparaciones relacionadas con la edad.

En relación con la posible existencia de un sesgo predictivo relacionado con la edad se evaluó previamente si el grupo de edad moderaba la validez predictiva examinando las diferencias de grupo en este aspecto. Así, para las evaluaciones de adolescentes, se comparó el grupo de 16 a 24 años con el de adolescentes más jóvenes (de 12 a 15 años), y para las evaluaciones de adultos se comparó al grupo objetivo con los adultos mayores (de 25 a 40 años). También se examinaron las diferencias medias en las puntuaciones totales y las categorías de nivel de riesgo para cada instrumento de riesgo para determinar si su uso podría conducir a un impacto diferencial en los sistemas de justicia.

Para la selección de los instrumentos se realizó un estudio previo que abarcó un amplio período de tiempo y diferentes países, a través del que seleccionaron nueve conjuntos de datos (no se especifican).

Las medidas utilizadas fueron las proporcionadas por los instrumentos seleccionados, operativizando la reincidencia como un evento de ocurrencia (sí/no) de dos tipos: *violenta* (delitos relacionados con el daño real o la amenaza de daño a las personas) o *cualquiera* (cualquier reincidencia, incluidos los delitos violentos, exceptuando las violaciones de la libertad condicional). No fue posible examinar *cualquier reincidencia* para el VRAG, porque sólo se disponía de la reincidencia violenta.

Para el análisis de datos se utilizaron diversos estadísticos como Chi-cuadrado para identificar diferencias significativas entre los grupos de edad en cuanto a género, raza (minorías frente a blancos), país y entorno (psiquiátrico forense, psiquiátrico civil, preadjudicación, delincuente adjudicado o encarcelado); egresiones jerárquicas moderadas de Cox para comprobar el sesgo predictivo; análisis de curvas ROC para las puntuaciones totales y los niveles de riesgo de cada instrumento de riesgo; o comparaciones de grupos medios para comprobar el impacto dispar.

Los resultados sugirieron que, en general, tanto las evaluaciones de riesgo de los adolescentes como las de los adultos son predictores válidos y significativos de la reincidencia de los jóvenes en edad de transición y no funcionan de forma diferente en función del grupo de edad; que tanto los instrumentos de evaluación del riesgo de los adolescentes como los de los adultos pueden utilizarse para estos jóvenes sin dar lugar a un tratamiento diferenciado; y que las personas encargadas de realizar las evaluaciones de riesgo de los jóvenes en edad de transición implicados en la justicia pueden confiar en la capacidad de los instrumentos para menores y adultos para predecir si estos individuos tienen probabilidades de reincidir, siempre y cuando el instrumento haya sido bien validado, aunque su utilidad para la gestión de riesgos puede ser más limitada en función de si el instrumento utilizado contempla factores de riesgo variables o necesidades criminógenas específicas para esta población.

Señalan los autores como limitaciones del estudio que se basó en datos secundarios que no se recopilaron originalmente para examinar las diferencias relacionadas con la edad en la validez predictiva, provocando el sesgo hacia determinadas edades producto de los diferentes tamaños de las muestras. Otra limitación derivada del uso de conjuntos de datos secundarios es que algunos no incluían un registro adecuado de la raza y el origen étnico, lo que no pudo controlarse en los análisis de YLS/CMI o VRAG. También indican como limitación que las calificaciones fueron puntuadas por expertos de distintos campos (la mayoría de los datos del YLS/CMI, por ejemplo, fueron puntuados por personal de justicia en el terreno, mientras que gran parte de los datos del VRAG y del HCR-20 fueron puntuados por asistentes de investigación o médicos). Finalmente, señalan que algunos de los instrumentos han sido objeto de revisiones recientes, por lo que sería necesario realizar nuevos estudios con estas versiones.

## 4. Discusión

Los estudios seleccionados resultaron ser muy heterogéneos en cuanto a muestras, ámbitos y contextos de aplicación de los instrumentos. En la Tabla 2 se resumen sus características más relevantes.

### 4.1. Evidencias de validez predictiva

En todos los trabajos excepto uno se ha tomado como referencia principal para evaluar la validez predictiva el análisis de las curvas ROC [*Receiver Operating Characteristics*] y la denominada área bajo la curva [*área under the curve*, AUC], medida muy utilizada particularmente en los instrumentos de evaluación de riesgos. La característica operativa del receptor (ROC) representa la sensibilidad frente a la especificidad en forma de curva, y el área bajo la curva (AUC) indica la probabilidad de predicción. El AUC oscila entre 0 y 1, donde 0.5 refleja una incapacidad de predicción y las puntuaciones más altas reflejan una mejor capacidad de predicción positiva, estimándose que los valores comprendidos entre 0.75 y 0.90 indican una buena capacidad predictiva del instrumento (Horcajo-Gil et al., 2019; Loinaz, 2017), siendo considerada una medida idónea para proporcionar el mejor conocimiento sobre la calidad técnica de los instrumentos de evaluación de riesgo de violencia (Muñoz-Vicente y López-Ossorio, 2016).

Puede definirse de una manera sencilla como la probabilidad de que el instrumento clasifique correctamente a un par de sujetos con relación al criterio de referencia (por ejemplo, la reincidencia).

Para el presente estudio se han tenido en cuenta únicamente los valores totales obtenidos en el estadístico de referencia, toda vez que la gran disparidad de parámetros evaluados dificulta en gran medida la comparación entre todos ellos. No se han representado los valores correspondientes a instrumentos que no son de uso habitual en el contexto forense español (YLS/CMI, Static-99R y START).

El RVD-Bcn del estudio 1 (Arbach, K., y Bobbio, A., 2019) obtuvo unos resultados de AUC=0.88 en las medidas iniciales y AUC=0.89 tras un seguimiento medio de 133 días.

Los datos obtenidos en el estudio 2 de Cartwright, Desmarais, Hazel, Griffith y Azizian (2018) en las medidas iniciales a 90 días y tras seguimiento a 180 días fueron de AUC=0.69 y 0.75 para el HCR-20. Para este mismo instrumento, en el estudio 11 (Vincent, Drawbridge y Davis, 2019) se obtuvieron valores de AUC=0.62.

El SARA (estudio 3, Cunha y Goncalves, 2019) obtuvo un AUC=0.95 en las medidas obtenidas para el modelo predictivo.

Respecto al SAVRY, en el estudio 4 (Li, Viljoen, Christiansen y Muir, 2020) se obtuvieron resultados de AUCs situados entre 0.63 y 0.89, dependiendo del grupo estudiado y el tipo de reincidencia (violenta o no violenta).



En el estudio 7 (McLachlan, Gray, Roesch, Douglas y Viljoen, 2018), los valores obtenidos oscilaron entre AUC=0.54 y 0.73 dependiendo del grupo estudiado y el tipo de reincidencia (violenta o no violenta).

En el estudio 8 (Muir, Viljoen, Jonnson, Cochrane y Rogers, 2020), los valores se encuentran entre AUC=0.65 y 0.79, en función del sexo y el origen étnico de los grupos estudiados.

En el estudio 10 (Shaffer, Viljoen y Douglas, 2022), se obtuvieron valores entre AUC=0.58 y 0.75 dependiendo del tipo de reincidencia (de violencia de pareja, general o violenta).

Finalmente, en el estudio 11 (Vincent, Drawbridge y Davis, 2019), los valores obtenidos respecto al grupo de análisis (jóvenes delincuentes en edad de transición) fueron de AUC=0.74 para la reincidencia violenta.

Este mismo estudio obtuvo valores AUC=0.63 para el grupo de análisis con el instrumento VRAG.

Los estudios 5 (López-Ossorio et al., 2019), 6 (López-Ossorio et al., 2020) y 9 (Quijano-Sánchez, et al., 2021), analizan el instrumento VioGén utilizado por el Ministerio del Interior. El estudio 5 obtiene para distintas versiones, periodos de seguimiento y tipo de reincidencia valores que oscilan entre AUC=0.60 y 0.82; en tanto que el estudio 6 obtiene un valor AUC=0.80 para la versión VPR5.0-H. El estudio 9 analiza el Formulario de Evolución de la Evaluación de Riesgo Policial (VPER), para el que no ofrece valores AUC al utilizar otros parámetros estadísticos.

El instrumento PCL:YV es evaluado únicamente en el estudio 10, obteniendo valores entre AUC=0.59 y 0.76 dependiendo del tipo de reincidencia (de violencia de pareja, general o violenta). Respecto a este instrumento se han señalado como limitaciones a tener en cuenta que la presencia de psicopatía en las escalas únicamente actúa como un predictor más, junto con el resto de los factores presentes, y que algunos autores dudan que pueda ser acertado establecer un punto de corte hasta haber obtenido evidencia empírica consistente respecto a la estabilidad de los rasgos de psicopatía a lo largo de la adolescencia y durante la transición a la edad adulta (Fernández-Sastrón y Esbec, 2003; Horcajo-Gil et al., 2019).

Loinaz (2017) señala como valores medios de AUC descritos en revisiones o metaanálisis 0.70 para el HCR-20, 0.74 para el VRAG, 0.63-0.70 para el SARA, y 0.71 para el SAVRY.

Los parámetros estudiados en esta revisión son muy heterogéneos, con una amplia variabilidad de muestras y resultados, si bien todos los instrumentos analizados muestran una validez predictiva entre moderada y alta y se encuentran próximos a los valores medios señalados por Loinaz, dependiendo de las medidas analizadas; destacando especialmente los resultados obtenidos por el RVD-Bcn y el SARA, ambos orientados a la predicción de riesgo de conducta violenta en el ámbito de la pareja.

## 4.2. Metodología

Los instrumentos utilizados se basan en los dos enfoques metodológicos más aceptados y utilizados en la actualidad, el juicio profesional estructurado y el método actuarial. No existe un acuerdo unánime en la comunidad científica acerca de la superioridad de una u otra metodología, pues ambas cuentan con ventajas e inconvenientes y están influidas por circunstancias tales como el contexto de aplicación o la subjetividad del evaluador, si bien cuentan con el reconocimiento de su validez (Fernández-Sastrón y Esbec, 2003; Loinaz, 2017).

Así, el VRAG, el PCL:YV, el VioGén y el VPR5.0-H utilizan un enfoque actuarial; en tanto que el HCR-20, el SARA, el RVD-Bcn y el SAVRY utilizan el juicio profesional estructurado.

## 4.3. Contextos de aplicación

Un único estudio (el número 3) se llevó a cabo en un contexto penitenciario sobre una muestra de hombres condenados por violencia contra la pareja encarcelados y en programas de intervención en Portugal, con la finalidad de identificar variables predictoras de homicidio en agresores de violencia de pareja.

Tres estudios (los números 5, 6, y 9) han utilizado muestras de población en España, y han sido desarrollados en el contexto de aplicación policial, utilizando la amplísima base de datos gestionada por el sistema VioGén, que incluye información muy completa sobre los casos de violencia de pareja denunciados; y están orientados a la mejora de la evaluación policial del riesgo de violencia de pareja y la adopción de medidas de protección, así como a la mejora de la predicción de riesgo de homicidio en ese ámbito.

Dos estudios (el 1 y el 2) se llevaron a cabo en un contexto asistencial, con una muestra argentina de mujeres víctimas de violencia de pareja usuarias de centros comunitarios con la finalidad de utilizar los resultados de las calificaciones de la herramienta para las decisiones relativas a las solicitudes de protección o para animar a las mujeres a presentar una denuncia; y población de internos varones en un hospital forense norteamericano con el fin de mejorar la predicción de riesgo de agresión institucional. El estudio 7 se llevó a cabo en un doble contexto asistencial y del sistema de justicia con una muestra de delincuentes juveniles afectados de síndrome alcohólico fetal y un grupo de control sin esta patología, teniendo como objetivo la mejora tanto de la predicción del riesgo de violencia en jóvenes afectados por síndrome alcohólico fetal como para la detección y tratamiento de sus necesidades clínicas.

El resto de los estudios (los números 4, 7, 8, 10 y 11) fueron efectuados en el marco del sistema de justicia penal de Canadá y Estados Unidos, sobre muestras de delincuentes juveniles y adultos en situación de libertad condicional; con la finalidad de comprobar la validez de los instrumentos comparados

en distintas poblaciones específicas atendiendo a criterios de raza, etnia, violencia específica en el ámbito de relaciones de pareja o edades de transición entre juventud y edad adulta.

#### **4.4. Tamaño de las muestras**

En general, el tamaño de las muestras es reducido, existiendo una amplísima diferencia entre estudios, oscilando entre N=100 (mínimo) y N=46047 (máximo). Cabe señalar al respecto que las muestras más amplias y representativas se encuentran en los estudios españoles, atendiendo al potencial ofrecido por el sistema estadístico VioGén, que incluye datos muy completos (respecto a víctima, agresor, hechos, circunstancias concurrentes...) de todos los casos de violencia de pareja denunciados en España.

#### **4.5. Discusión**

Los instrumentos más estudiados son el SAVRY, con cinco estudios en cuatro de los cuales se comparó con otros instrumentos, y el VioGén, con tres estudios; seguidos por el HCR-20 con dos (en ambos casos junto con otros instrumentos). El RVD-Bcn, SARA y VRAG fueron empleados en un único estudio cada uno de ellos (en combinación con otros en el caso del VRAG).

Estos resultados, en combinación con las poblaciones diana estudiadas sugieren el interés de los investigadores norteamericanos (Estados Unidos y Canadá) por validar instrumentos para la predicción del riesgo de violencia de delincuentes juveniles como es el SAVRY en poblaciones de diferente origen étnico y racial e incluso rango de edad, especialmente en el contexto de la aplicación de medidas tanto de prevención como de protección en el sistema de justicia; en tanto que los investigadores europeos se centran más en la predicción del riesgo de violencia grave en la pareja en el contexto policial, en especial en la mejora de predicción del riesgo de homicidio sobre la mujer.

Los estudios dirigidos a comprobar la validez del SAVRY en distintas poblaciones son coincidentes en mostrar valores situados en el rango medio-alto, lo que lleva a concluir a los investigadores que en general el instrumento resulta aceptable para ser utilizado en esas poblaciones especiales por razón de etnia, raza, o padecimiento de patología psiquiátrica, si bien aportan sugerencias de mejora en determinados aspectos como la formación multicultural de los evaluadores o la necesaria atención a los distintos contextos culturales. Los resultados de los estudios comparativos entre esta herramienta y otras del mismo ámbito (VRAG, HCR-20, YLS/CMI o PCL:YV) señalan a un ligeramente mayor poder predictivo del SAVRY, con la única excepción a la validez de este instrumento señalada en el estudio 10, cuyos hallazgos sugieren que el uso de las herramientas existentes para adolescentes para evaluar y gestionar el riesgo de violencia de pareja podría ser problemático puesto que proporcionan una visión general de las características y las características y circunstancias de un delincuente adolescente que pueden ser

propias de ese período evolutivo, lo que limitaría su capacidad para pronosticar con precisión el riesgo de perpetración de violencia de pareja.

Los resultados sobre validez predictiva con relación a los instrumentos más analizados (SAVRY y VioGén) sugieren cierta superioridad en esta medida para las herramientas actuariales del sistema VioGén, encontrando valores más altos en términos absolutos. Estos resultados cobran más valor atendiendo tanto al tamaño muestral de los diferentes estudios como a la heterogeneidad de las poblaciones utilizadas por VioGén, que incluyen un amplio rango de edades, ambos sexos, agresores y víctimas, lo que permite otorgarles un mayor poder de generalización.

Especialmente precisos parecen los instrumentos utilizados para la predicción de riesgo de violencia mortal (homicidio), que alcanzan valores AUC de 0.80 para el VPR5.0-H y 0.95 para el SARA en la muestra penitenciaria portuguesa, si bien en este habría que precisar aquí la importante limitación que supone su reducido tamaño (N=172). Estos resultados son superiores a los encontrados en un reciente metaanálisis de la validez predictiva de instrumentos de evaluación de riesgo de violencia doméstica (Van der Put, Gubbels y Assink, 2019), que encontró una validez predictiva de este tipo de herramientas considerada baja, con AUC=0.643. El SARA está ampliamente representado en dicho metaanálisis, en tanto que no figura en el mismo ningún estudio respecto a VioGén.

A pesar de los buenos resultados obtenidos por el SARA en el estudio 3 y por el VPR5.0-H del estudio 6 en cuanto a la predicción del homicidio en el ámbito de la pareja, muestran algunos resultados divergentes. Así, en el estudio 3 se destaca que los homicidas de pareja tienen 18 veces más probabilidades de utilizar armas, en tanto que el VPR5.0-H no identifica el uso de armas como un factor diferenciador entre un episodio repetido no letal y un homicidio de mujeres por su pareja, atribuyendo como posible explicación que la muestra de homicidio de mujeres por su pareja utilizada en el estudio procede no solo de casos con denuncias previas por violencia (27,6%), sino también de otros que presentan características diferentes. Del mismo modo, el estudio 3 señala como un importante factor predictor del menor riesgo de violencia letal la existencia de un historial de violación de libertad condicional o supervisión comunitaria, en tanto que para el estudio 6 esta misma circunstancia es considerada relevante para la predicción de dicho riesgo.

La revisión realizada cuenta con algunas limitaciones significativas. Así, respecto a la evidencia incluida no se han identificado los posibles sesgos que la literatura científica destaca y que han podido afectar a los resultados de los estudios revisados, como el riesgo del evaluador, el de publicación y el editorial. En ello ha influido la ya mencionada heterogeneidad de los estudios seleccionados, que dificulta la realización de los necesarios análisis estadísticos para la detección de estos sesgos; y ha determinado la necesidad de realizar una revisión de carácter cualitativo. Otras limitaciones importantes respecto al proceso de selección de estudios para esta revisión han sido la ausencia de revisión por pares y la reducida

elección de instrumentos incluidos en los criterios de inclusión, que ha limitado enormemente los resultados obtenidos, dejando una muestra final bastante reducida.

## 5. Conclusiones

Amor, Echeburúa y Bohórquez (2020) reseñan como avances más significativos en el ámbito de la evaluación de riesgo de violencia la sustitución del término *valoración de la peligrosidad* por el de *riesgo de violencia*, el abandono de los métodos de evaluación basados en juicio clínico no estructurado a favor de aquellos actuariales o basados en el juicio clínico estructurado, y el acervo de instrumentos adaptados al español o creados en España con capacidad predictiva en contextos específicos. Los resultados obtenidos en esta revisión son concordantes con estas conclusiones, pues todos los estudios manejan el concepto de riesgo de violencia, han utilizado alguno de los métodos de evaluación señalados y efectivamente se han encontrado instrumentos creados en España (VioGén) que ha demostrado una buena capacidad predictiva en el ámbito de violencia de pareja.

El uso final de los instrumentos para la valoración de riesgo de violencia difiere en función de diferentes factores, en especial el tipo de profesional que lo aplica y el contexto en el que se desenvuelve (Graham, Sahay, Rizo, Messing y Macy, 2021). Así, los servicios policiales los emplean para reducir el riesgo de nuevas agresiones a las víctimas y mejorar su protección, los profesionales de la salud para evaluar la pertinencia de recibir un tratamiento para agresores, y los operadores jurídicos para gestionar el riesgo y como ayuda en la decisión de adopción de medidas legales. En el contexto pericial psicológico forense en España, cualquier informe solicitado respecto a evaluación de riesgo de violencia puede servir a esos objetivos.

No obstante, cabe precisar que el contexto donde quizá adquieran mayor relevancia estas periciales es en la administración de justicia, especialmente en el ámbito de protección y reforma de menores, donde la vigente Ley Orgánica 5/2000 de 12 de enero reguladora de la responsabilidad penal de los menores contempla el establecimiento de planes de intervención flexibles e individualizados para la reeducación y reinserción social del menor; y en la valoración de riesgo de violencia de género, que constituye uno de los servicios incluidos en la Carta de Servicios 2021-2024 de las Unidades de Valoración Forense Integral (UVFI) de los Institutos de Medicina Legal y Ciencias Forenses del Ministerio de Justicia.

Los estudios analizados en esta revisión se centran especialmente en los ámbitos de la predicción de riesgo de violencia juvenil y de pareja, destacando los instrumentos SAVRY y VioGén en los respectivos contextos. Ambos son los más empleados en España, el SAVRY, que cuenta con adaptación española de Hilterman y Andrés-Pueyo (2008) en el contexto de la evaluación e intervención con menores infractores (Luengo, Cutrín y Maneiro, 2015); en tanto que para la violencia de pareja está estandarizado para su uso por las Fuerzas y Cuerpos de Seguridad por el Ministerio del Interior el sistema VioGén.

Los hallazgos de la revisión indican que todos los instrumentos han sido correctamente validados mediante parámetros estadísticos científicamente reconocidos, y han sido aplicados en poblaciones muy diversas. Específicamente el SAVRY ha mostrado una capacidad general aceptable para ser empleado con jóvenes de muy diversos entornos y procedencias, aunque una interesante implicación práctica de los resultados al respecto sugiere que los evaluadores deben recibir una mejor y mayor formación en aspectos relacionados con la multiculturalidad (cada vez más presente en todas las capas sociales) como el fenómeno de la aculturación y los estereotipos y prejuicios, necesitándose más investigación sobre la posible influencia de los sesgos y antecedentes raciales, étnicos y culturales de los propios evaluadores.

Como lagunas metodológicas cabría destacar el escaso tamaño de las muestras en todos los estudios y la heterogeneidad tanto en las fuentes de obtención de datos como en la administración de los instrumentos, limitaciones en las que prácticamente coinciden todos los trabajos a excepción de los efectuados con VioGén, y que afectan a la calidad de las conclusiones. Una mejora sustancial en estos aspectos se considera esencial para futuras investigaciones y el avance del conocimiento científico en esta área. Sería necesaria igualmente mayor investigación respecto al diseño específico de instrumentos de predicción de riesgo de violencia en la pareja respecto a jóvenes adolescentes, atendiendo a los resultados de los estudios analizados.

## 6. Referencias

- Alicia, M. J., & Miguel Ángel, C. O. (2011). Riesgo de Violencia y Trastorno Mental. *Portada de Anuario de Psicología Jurídica*, 21, 69-81. <https://doi.org/10.5093/jr2011v21a7>
- Amor, P., Echeburúa, E., y Bohórquez, Isaura A. (2020). Psicología forense y derecho penal (II): tipos de delinquentes violentos y evaluación del riesgo de violencia. Miguel A. Carrasco y Fernando Rubio (Eds.), *Psicología Jurídica y Forense, Vol. II: Ámbitos de aplicación* (págs. 163-220)
- Andrés-Pueyo, A., y Echeburúa Odriozola, E. (2010). Valoración del riesgo de violencia. Instrumentos disponibles e indicaciones de aplicación. *Psicothema*, Vol. 22, nº 3, 403-409.
- Andrés-Pueyo, A. (2013). Peligrosidad criminal: análisis crítico de un concepto polisémico. Maroto Calatayud, M., y Demetrio Crespo, E. (Coords). *Neurociencias y derecho penal: nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, 483-504.
- Arbach, K., & Bobbio, A. (2018). Intimate Partner Violence Risk Assessment in Community Health Facilities: A Multisite Longitudinal Study. *Psychosocial Intervention*, 27(2), 105-112. <https://doi.org/10.5093/pi2018a13>
- Barra, S., Bessler, C., Landolt, M. A., & Aebi, M. (2018). Testing the validity of criminal risk assessment tools in sexually abusive youth. *Psychological Assessment*, 30(11), 1430-1443. <https://doi.org/10.1037/pas0000590>
- Cajal, B., Jiménez, R., Gervilla, E., & Montaña, J. J. (2020). Doing a Systematic Review in Health Sciences. *Clínica y Salud*, 31(2), 77-83. <https://doi.org/10.5093/clysa2020a15>
- Cartwright, J. K., Desmarais, S. L., Hazel, J., Griffith, T., & Azizian, A. (2018). Predictive validity of HCR-20, START, and static-99R assessments in predicting institutional aggression among sexual offenders. *Law and Human Behavior*, 42(1), 13-25. <https://doi.org/10.1037/lhb0000263>
- Consejo Médico Forense. Protocolo de valoración forense urgente del riesgo de violencia de género. Ministerio de Justicia, Secretaría General Técnica. Madrid, 2020.
- Cunha, O. S., & Gonçalves, R. A. (2019). Predictors of Intimate Partner Homicide in a Sample of Portuguese Male Domestic Offenders. *Journal of Interpersonal Violence*, 34(12), 2573-2598. <https://doi.org/10.1177/0886260516662304>
- Esbec Rodríguez, E. (2003). Valoración de la peligrosidad criminal (Riesgo-Violencia) en Psicología forense. Aproximación conceptual e histórica. *Psicopatología Clínica Legal y Forense*, Vol. 3, Nº 2, 45-64.



- Fernández-Sastrón, O., Esbec Rodríguez, E. (2003). Valoración de la peligrosidad criminal (Riesgo-Violencia) en Psicología forense. Instrumentos de evaluación y perspectivas. *Psicopatología Clínica Legal y Forense*, Vol. 3, Nº 2, 65-90.
- García-López, E. (Ed.). (2019). *Psicopatología de la violencia: Aspectos jurídicos y evaluación criminológica*. Editorial El Manual Moderno.
- Graham, L. M., Sahay, K. M., Rizo, C. F., Messing, J. T., & Macy, R. J. (2021). The Validity and Reliability of Available Intimate Partner Homicide and Reassault Risk Assessment Tools: A Systematic Review. *Trauma, Violence, & Abuse*, 22(1), 18-40. <https://doi.org/10.1177/1524838018821952>
- Horcajo-Gil, P. J., Dujo-López, V., Andreu-Rodríguez, J. M., & Marín-Rullán, M. (2019). Valoración y Gestión del Riesgo de Reincidencia Delictiva en Menores Infractores: Una Revisión de Instrumentos. *Anuario de Psicología Jurídica*, 29(1), 41-53. <https://doi.org/10.5093/apj2018a15>
- Jiménez Gómez, F. (2009). *Evaluación psicológica forense*. Solo Soluciones.
- Li, S. M. Y., Viljoen, J. L., Christiansen, A. K., & Muir, N. M. (2020). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) among a sample of Asian Canadian youth on probation. *Law and Human Behavior*, 44(6), 485-501. <https://doi.org/10.1037/lhb0000431>
- Loinaz, I. (2017). *Manual de evaluación del riesgo de violencia. Metodología y ámbitos de aplicación*. Editorial Pirámide.
- López-Ossorio, J. J., González-Álvarez, J. L., Loinaz, I., Martínez-Martínez, A., & Pineda, D. (2020). Intimate partner homicide risk assessment by police in Spain: The dual protocol VPR5.0-H. *Psychosocial Intervention*, 30(1), 47-55. <https://doi.org/10.5093/pi2020a16>
- López-Ossorio, J. J., González-Álvarez, J. L., Muñoz Vicente, J. M., Urruela Cortés, C., & Andrés-Pueyo, A. (2019). Validation and Calibration of the Spanish Police Intimate Partner Violence Risk Assessment System (VioGén). *Journal of Police and Criminal Psychology*, 34(4), 439-449. <https://doi.org/10.1007/s11896-019-09322-9>
- Luengo, M. Á., Cutrín, O., y Maneiro, L. (2015). Protocolo de Valoración del Riesgo en Adolescentes Infractores: Una herramienta informatizada para la gestión del riesgo. *Infancia, Juventud y Ley*, 51-58.
- Martínez Garay, L., y Montes Suay, F. (2018). El uso de valoraciones del riesgo de violencia en Derecho Penal: algunas cautelas necesarias. *Indret, Revista para el análisis del Derecho*, 2-47.
- McLachlan, K., Gray, A. L., Roesch, R., Douglas, K. S., & Viljoen, J. L. (2018). An evaluation of the predictive validity of the SAVRY and YLS/CMI in justice-involved youth with fetal alcohol spectrum disorder. *Psychological Assessment*, 30(12), 1640-1651. <https://doi.org/10.1037/pas0000612>

- Ministerio de Justicia (2021). Unidades de Valoración Forense Integral (UVFI). Institutos de Medicina Legal y Ciencias Forenses del Ministerio de Justicia. Carta de Servicios 2021-2024.  
<https://www.mjusticia.gob.es/es/ministerio/organismos-entidades/medicina-legal-ciencias/unidades-de-valoracion-forense-integral>
- Muir, N. M., Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., & Rogers, B. J. (2020). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) with Indigenous and Caucasian female and male adolescents on probation. *Psychological Assessment, 32*(6), 594-607.  
<https://doi.org/10.1037/pas0000816>
- Muñoz Vicente, J. M., & López-Ossorio, J. J. (2016). Valoración psicológica del riesgo de violencia: Alcance y limitaciones para su uso en el contexto forense. *Anuario de Psicología Jurídica, 26*(1), 130-140.  
<https://doi.org/10.1016/j.apj.2016.04.005>
- Quijano-Sánchez, L., Liberatore, F., Rodríguez-Lorenzo, G., Lillo, R. E., & González-Álvarez, J. L. (2021). A twist in Intimate Partner Violence Risk Assessment Tools: Gauging the contribution of exogenous and historical variables. *Knowledge-Based Systems, 234*, 107586.  
<https://doi.org/10.1016/j.knosys.2021.107586>
- Sánchez-Meca, J. (2010). Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta, Vol. 38*, Nº 2, 53-64.
- Sánchez-Meca, J., y Botella Ausina, J. (2010). Revisiones sistemáticas y metaanálisis: herramientas para la práctica profesional. *Papeles del Psicólogo, Vol. 31*, Nº 1, 7-17.
- Shaffer, C. S., Viljoen, J. L., & Douglas, K. S. (2022). Predictive validity of the SAVRY, YLS/CMI, and PCL:YV is poor for intimate partner violence perpetration among adolescent offenders. *Law and Human Behavior, 46*(3), 189-200. <https://doi.org/10.1037/lhb0000483>
- Van der Put, C. E., Gubbels, J., & Assink, M. (2019). Predicting domestic violence: A meta-analysis on the predictive validity of risk assessment tools. *Aggression and Violent Behavior, 47*, 100-116.  
<https://doi.org/10.1016/j.avb.2019.03.008>
- Vincent, G. M., Drawbridge, D., & Davis, M. (2019). The validity of risk assessment instruments for transition-age youth. *Journal of Consulting and Clinical Psychology, 87*(2), 171-183.  
<https://doi.org/10.1037/ccp0000366>

## 7. Anexos

**Tabla 1***Denominación y Abreviaturas de los Instrumentos Analizados en esta Revisión*

Denominación del Instrumento	Abreviatura
Structured Assessment of Violence Risk in Youth	SAVRY
Historical Clinical Risk Management-20	HCR-20
Spousal Assault Risk Assessment Guide	SARA
Protocolo de Valoración de Riesgo de Violencia Contra la Mujer por Parte de su Pareja o Expareja	RVD-Bcn
Youth Level of Service/Case Management Inventory <sup>a</sup>	YLS/CMI
Short-Term Assessment of Risk and Treatability <sup>a</sup>	START
Violence Risk Appraisal Guide	VRAG
Psychopathy Checklist-Youth Version	PCL:YV
Static-99R <sup>a</sup>	Static-99R
Spanish Police Intimate Partner Violence Risk Assessment System	VioGén
Valoración Policial del Riesgo (Homicidio)	VPR.0-H

<sup>a</sup> Para esta revisión no se consideró este instrumento por no ser de uso en el ámbito forense español

**Tabla 2***Características más Relevantes de los Estudios Seleccionados*

Autor y año	Título	Instrumentos	Muestra	País	Contexto de Aplicación
Arbach, K., y Bobbio, A. (2019)	Intimate Partner Violence Risk Assessment in Community Health Facilities: A Multisite Longitudinal Study	RVD-Bcn	Mujeres víctimas de violencia doméstica usuarias de centros comunitarios (N=321)	Argentina	Asistencial
Cartwright, Desmarais, Hazel, Griffith y Azizian. (2018)	Predictive Validity of HCR-20, START, and Static-99R Assessments in Predicting Institutional Aggression Among Sexual Offenders	HCR-20 START Static-99R	Varones detenidos/internados civilmente en hospital forense de seguridad (N=152)	Estados Unidos	Asistencial
Cunha y Goncalves (2019)	Predictors of Intimate Partner Homicide in a Sample of Portuguese Male Domestic Offenders	SARA	Varones encarcelados/en programas de intervención (N=172)	Portugal	Penitenciario
Li, Viljoen, Christiansen y Muir (2020)	Predictive Validity of the Structured Assessment of Violence Risk in Youth (SAVRY) Among a Sample of Asian Canadian Youth on Probation	SAVRY	Delincuentes juveniles asiáticos y blancos canadienses (N=573)	Canadá	Sistema de Justicia (Libertad Condicional)
López-Ossorio, González, Muñoz, Urruela y Andrés-Pueyo (2019)	Validation and Calibration of the Spanish Police Intimate Partner Violence Risk Assessment	VioGén	Víctimas y agresores en casos de violencia de pareja (N=6613)	España	Policial

	System (VioGén)				
López-Ossorio, González-Álvarez, Loinaz, Martínez y Pineda (2020)	Intimate Partner Homicide Risk Assessment by Police in Spain: The Dual Protocol VPR5.0-H	VPR5.0-H	Víctimas y agresores en casos de violencia de pareja con/sin homicidio (N=2159)	España	Policial
McLachlan, Gray, Roesch, Douglas y Viljoen (2018)	An Evaluation of the Predictive Validity of the SAVRY and YLS/CMI in Justice-Involved Youth With Fetal Alcohol Spectrum Disorder	SAVRY YLS/CMI	Delincuentes juveniles con/sin síndrome alcohólico fetal (N=100)	Estados Unidos	Asistencial Sistema de Justicia (Libertad Condicional)
Muir, Viljoen, Jonnson, Cochrane y Rogers (2020)	Predictive Validity of the Structured Assessment of Violence Risk in Youth (SAVRY) With Indigenous and Caucasian Female and Male Adolescents on Probation	SAVRY	Delincuentes juveniles indígenas y caucásicos (N=744)	Canadá	Sistema de Justicia (Libertad Condicional)
Quijano-Sánchez, Liberatore, Rodríguez-Lorenzo, Lillo y González-Álvarez (2021)	A twist in Intimate Partner Violence Risk Assessment Tools: Gauging the contribution of exogenous and historical variables	VioGén	Víctimas y agresores en violencia de pareja (N=46047)	España	Policial
Shaffer, Viljoen y Douglas (2022)	Predictive Validity of the SAVRY, YLS/CMI, and PCL:YV Is Poor for Intimate Partner Violence	SAVRY YLS/CMI PCL:YV	Delincuentes juveniles (N=163)	Canadá	Sistema de Justicia (Libertad Condicional)

	Perpetration Among Adolescent Offenders				
Vincent, Drawbridge y Davis (2019)	The Validity of Risk Assessment Instruments for Transition-Age Youth	HCR-20 VRAG YLS/CMI SAVRY	Datos sobre delincuentes juveniles (N=1365) y adultos (N=1988)	Estados Unidos	Sistema de Justicia (Libertad Condiciona)

---