

Tres sistemas de policía predictiva en España: VioGén, RisCanvi y VeriPol

Evaluación desde una perspectiva de derechos humanos

Lucía Martínez Garay (coord.)



Tres sistemas de policía predictiva en España: VioGén, RisCanvi y VeriPol

Evaluación desde una perspectiva
de derechos humanos

Tres sistemas de policía predictiva en España: VioGén, RisCanvi y VeriPol

Evaluación desde una perspectiva
de derechos humanos

Lucía Martínez Garay, coord.

Colección: Desarrollo Territorial
Papers, 10

Dirección de la colección: Maria Dolores Pitarch

Consejo de dirección: Josep Vicent Boira, Sacramento Pinazo, Joan Romero i Ana Sales



Esta obra está bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.

Esta publicación no puede ser reproducida, ni total ni parcialmente, ni registrada en, o transmitida por, un sistema de recuperación de información, de ninguna forma ni por ningún medio, sea fotomecánico, fotoquímico, electrónico, por fotocopia o por cualquier otro, sin el permiso de la editorial. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

Título original:

Three predictive policing approaches in Spain: VioGén, RisCanvi and VeriPol. Assessment from a human rights perspective

© Universitat de València, 2024

© Del texto: los autores y las autoras, 2025

© De esta edición: Universitat de València, 2025

Publicacions de la Universitat de València

puv.uv.es

publicacions@uv.es

Coordinación editorial: Amparo Jesús-Maria Romero

Corrección y maquetación: Letras y Píxeles, S. L.

Traducción: traductor automático, supervisado por los autores y las autoras

Diseño de la cubierta: Celso Hernández de la Figuera

ISBN papel: 978-84-1118-659-9

ISBN OA: 978-84-1118-660-5

DOI: <https://doi.org/10.7203/PUV-OA-9788411186605>

Edición digital

ÍNDICE

CONTRIBUCIONES	11
AGRADECIMIENTOS	12
LISTADO DE ABREVIATURAS	13
RESUMEN EJECUTIVO	15
1. Génesis del informe y elección de las herramientas estudiadas	16
2. Condicionantes y límites de la realización del informe	16
3. Conclusiones generales aplicables a las tres herramientas analizadas	18
4. Principales conclusiones respecto de VioGén	21
5. Principales conclusiones respecto de RisCanvi.....	22
6. Principales conclusiones respecto de VeriPol	24
VIOGÉN.....	25
1. Descripción general	25
2. Datos	27
3. Eficacia del enfoque policial predictivo para prevenir la delincuencia	51
4. Evaluación de impacto sobre los derechos humanos.....	66
5. Nuevos desarrollos sobre VeriPol entre enero y abril de 2023.....	73
6. Fuentes	76
RISCANVI.....	79
1. Descripción general	79
2. Datos	85
3. Eficacia del enfoque policial predictivo para prevenir la delincuencia	113
4. Evaluación del impacto sobre los derechos humanos	118
5. Fuentes	133

VERIPOL	135
1. Descripción general	135
2. Datos	137
3. Eficacia del enfoque policial predictivo para prevenir la delincuencia	145
4. Evaluación del impacto sobre los derechos humanos	161
5. Fuentes.....	165
 ANEXOS SOBRE VIOGÉN.....	167
Anexo I: dificultades para obtener información y datos sobre VeriPol.....	167
Anexo II: Los factores de riesgo de VioGén	172
Anexo III: Medidas de protección policial para cada nivel de riesgo según la Instrucción 4/2019 de la Secretaría de Estado de Seguridad.....	179
 ANEXOS SOBRE RISCANVI	187
1. Riscanvi completo.....	187
2. <i>Riscanvi Screening</i>	209
 ADENDA: PREGUNTAS Y RESPUESTAS	213
 ADENDA AL INFORME «TRES ENFOQUES POLICIALES PREDICTIVOS EN ESPAÑA»:.....	213
1. Preguntas específicas sobre VioGén	213
2. Preguntas específicas sobre RisCanvi.....	215
3. Preguntas específicas sobre VeriPol	228

Informe realizado por:

Lucía Martínez Garay. Investigadora responsable. Departament de Dret Penal e Institut Universitari d'Investigació en Criminologia i Ciències Penals, Universitat de València.

Andrés Boix Palop. Departament de Dret Administratiu i Processal, Universitat de València.

Álvaro Briz Redón. Departament d'Estadística i Investigació Operativa, Universitat de València.

Fernando Flores Giménez. Departament de Dret Constitucional, Ciència Política i de l'Administració e Institut de Drets Humans, Universitat de València.

Andrea García Ortiz. Institut Universitari d'Investigació en Criminologia i Ciències Penals, Universitat de València.

Mireia Molina Sánchez. Departament de Dret Administratiu i Processal, Universitat de València.

Francisco Montes Suay. Departament d'Estadística i Investigació Operativa, Universitat de València.

Adrián Palma Ortigosa. Departament de Dret Administratiu i Processal, Universitat de València.

Alfred Peris Manguillot. Institut de Matemàtica Pura i Aplicada, Universitat Politècnica de València.

Alba Soriano Arnanz. Departament de Dret Administratiu i Processal, Universitat de València.

Financiación y conflictos de intereses

El origen de este trabajo se encuentra en un encargo de una organización no gubernamental de derechos humanos para estudiar las herramientas policiales predictivas existentes en España, organización que también proporcionó apoyo financiero. Este informe es fruto de la actividad investigadora desarrollada en el marco del proyecto de investigación «Algorithmic Law» (CIPROM/2024/16) de la Universitat de València, del que forman parte los siguientes autores: Andrés Boix Palop, Lucía Martínez Garay, Alba Soriano Arnanz, Adrián Palma Ortigosa y Mireia Molina Sánchez; en el proyecto de investigación PID2021-123441NB-I00 (financiado por MCIN/AEI/ 10.13039/501100011033 y por «FEDER A way of making Europe», del que son miembros los siguientes autores: Lucía Martínez Garay y Andrea García Ortiz; y por el proyecto de investigación PID2023-152781NB-I00 «Administrative Planning» (PlanAdm), del que forman parte los siguientes autores: Andrés Boix Palop, Alba Soriano Arnanz y Adrián Palma Ortigosa.

Las opiniones expresadas en este informe son exclusivamente las de los autores y no representan las de las instituciones a las que pertenecen, ni las de las instituciones que encargaron o financiaron la investigación.



Limitaciones

Este informe se cerró en noviembre de 2022, pero diversas razones técnicas y personales han retrasado su publicación.

Con posterioridad a esa fecha solo se han incluido de forma muy fragmentaria algunas informaciones y noticias aparecidas sobre VioGén y RisCanvi hasta abril de 2023, así como algunas referencias muy breves al informe de auditoría publicado sobre RisCanvi en 2024.

La compatibilidad de las tres herramientas de policía predictiva con los derechos humanos se ha evaluado en relación con la normativa vigente en España y Europa a la fecha de cierre del informe. Por ello, en este informe no se evalúa la conformidad de las herramientas con las obligaciones establecidas en el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial en la Unión Europea.

Contribuciones

Lucía Martínez Garay ha sido la investigadora responsable y ha dirigido y coordinado el informe. Participó en el análisis de datos cuantitativos de VioGén y RisCanvi, entrevistó a agentes de policía, abogados, médicos forenses y a un magistrado, fue responsable de la redacción preliminar del capítulo de VioGén, participó en la redacción preliminar de los capítulos de RisCanvi y VeriPol, y ha supervisado toda la redacción final del informe completo.

Francisco Montes Suay y Álvaro Briz Redón realizaron el análisis cuantitativo de los datos sobre RisCanvi y colaboraron en el análisis de algunas de las muestras de VioGén.

Alfred Peris Manguillot realizó el análisis cuantitativo de los datos públicamente disponibles sobre VioGén y ayudó a revisar el análisis de las muestras de VioGén.

Alba Soriano Arnanz y Adrián Palma Ortigosa analizaron las sentencias de las audiencias provinciales sobre RisCanvi. Se encargaron de la redacción preliminar de los capítulos de RisCanvi y VeriPol, y entrevistaron a los abogados. Adrián Palma Ortigosa entrevistó a algunos agentes de policía sobre VeriPol y Alba Soriano Arnanz ha traducido la mayor parte del informe al inglés y ha revisado la traducción realizada por los demás miembros del equipo.

Andrea García Ortiz y Mireia Molina Sánchez participaron en el análisis de datos cuantitativos sobre VioGén y entrevistaron a los jueces. Mireia Molina Sánchez analizó las sentencias de las audiencias provinciales sobre VioGén, y Andrea García Ortiz entrevistó a una trabajadora de la Oficina de Asistencia a la Víctima del Delito.

Andrés Boix Palop coordinó las entrevistas a los jueces, realizó algunas de ellas, supervisó las partes relativas a las obligaciones de rendición de cuentas de cada herramienta y revisó la redacción final de todo el informe.

Fernando Flores Jiménez facilitó el contacto con la Secretaría de Estado de Seguridad y revisó la redacción final de todo el informe.

Todas las conclusiones han sido debatidas y asumidas por todos los autores. Las opiniones expresadas en este informe son las de sus autores y no representan las de las instituciones a las que pertenecen, ni tampoco las de las instituciones que encargaron o financiaron la investigación.

Agradecimientos

Queremos agradecer a las siguientes personas su ayuda o contribución a este informe:

Carlos Castillo y Marzieh Karimi Haghighi, de la Universitat Pompeu Fabra, por los cálculos que hicieron para nosotros con sus muestras de datos de RisCanvi.

Virginia Álvarez, por su apoyo durante la elaboración del informe.

Verónica Gisbert Gracia, de la Universitat de València, por sus orientaciones y consejos sobre metodologías cualitativas y en la preparación de las entrevistas.

Antonio Andrés Pueyo, de la Universitat de Barcelona, por facilitar información sobre la descripción detallada de algunos de los factores de riesgo de RisCanvi.

El Tribunal Superior de Justicia de la Comunidad Valenciana, la Asociación de Psicología Forense de la Administración de Justicia, la Secretaría de Estado de Seguridad del Ministerio del Interior, la Unidad de Planificación y Coordinación Estratégica de la Policía Nacional, las jefas de las Unidades de Valoración Forense de los Institutos de Medicina Legal y Ciencias Forenses de Valencia y Alicante, la persona del Departament de Justícia de la Generalitat de Catalunya que contestó a nuestros correos electrónicos, los jueces, policías, abogados, personal de la Oficina de Asistencia a las Víctimas del Delito de Valencia, y todas las personas que estuvieron dispuestas a darnos información sobre alguna de las herramientas de evaluación del riesgo o nos facilitaron el contacto con personas que podían hacerlo.

Listado de abreviaturas

AEPD	Agencia Española de Protección de Datos
CGPJ	Consejo General del Poder Judicial
RGPD	Reglamento General de Protección de Datos
HCR-20	<i>Historical Clinical Risk Management–20</i>
IMLCF	Instituto de Medicina Legal y Ciencias Forenses
JVM	Jueces de Violencia sobre la Mujer
LSI-R™	<i>Level of Service Inventory–Revised™</i>
TM	Trastorno mental
PCL-R	Lista de Comprobación de Psicopatía - Revisada
JVP	Juez de vigilancia penitenciaria
REVI	Escala de riesgo de reincidencia violenta de RisCanvi
RisCanvi-C	RisCanvi completo
RisCanvi-S	RisCanvi <i>Screening</i>
SARA	<i>Spousal Assault Risk Assessment Guide</i>
SIDENPOL	Sistema de Denuncias Policiales de la Policía Nacional
SIPC	Sistema de Información Penitenciaria de Cataluña
SOS-RisCanvi	Equipo de apoyo, orientación y seguimiento en la aplicación del RisCanvi
SES	Secretaría de Estado de Seguridad
SVR-20	<i>Sexual Violence Risk–20</i>
VFR	Valoración forense del riesgo
VPER	Valoración policial de la evolución del riesgo
VPR	Valoración policial del riesgo

RESUMEN EJECUTIVO

El informe que se enlaza a continuación, *Three predictive policing approaches in Spain: Viogén, RisCanvi and VeriPol (Assessment from a human rights perspective)*, ha sido elaborado por un grupo de profesores de la Universitat de València y la Universitat Politècnica de València, especialistas en distintas ramas del derecho (penal, administrativo y constitucional), la estadística y la matemática, bajo la coordinación de la Dra. Lucía Martínez Garay, profesora de Derecho Penal de la Universitat de València. El trabajo evalúa tres herramientas algorítmicas (con uso de inteligencia artificial, IA, en una de ellas) empleadas por la Administración española para realizar análisis predictivos en los ámbitos policiales y penitenciarios, desde la perspectiva de su respeto a los derechos fundamentales de los ciudadanos que son objeto de las evaluaciones para las que estas herramientas son empleadas. Los objetivos de este informe han sido: 1) analizar la capacidad predictiva de los tres sistemas y su eficacia para alcanzar el objetivo que pretenden; 2) analizar su grado de transparencia, tanto para la ciudadanía en general como para las personas directamente afectadas por su utilización; 3) identificar posibles sesgos estadísticos; 4) analizar si respetan el vigente marco jurídico europeo y español, tanto en materia de protección de datos como respecto del empleo de este tipo de modelos; y, finalmente, 5) valorar si los costes para los derechos humanos exceden lo que sería legítimo, necesario o proporcionado en relación con el beneficio actual o probable en términos de prevención del delito. Aunque estos han sido los objetivos comunes para las tres herramientas, el análisis se ha tenido que adaptar a las características de cada una de ellas y, sobre todo, al tipo y a la cantidad de datos e información de los que hemos podido disponer, que ha sido dispar en cada una de las tres. Asimismo, en el análisis se han combinado métodos cuantitativos y cualitativos, de nuevo de forma distinta para cada herramienta según la información a la que nos ha sido posible acceder.

1. GÉNESIS DEL INFORME Y ELECCIÓN DE LAS HERRAMIENTAS ESTUDIADAS

El origen de este informe es un encargo de una ONG de ámbito europeo e internacional que pretende realizar un estudio comparativo respecto de los riesgos que las herramientas de policía predictiva utilizadas en distintos estados europeos pueden plantear para los derechos fundamentales de los ciudadanos. El presente trabajo se refiere a la parte del análisis sobre las herramientas de uso predictivo empleadas en España, y ha sido realizado, en el marco de este encargo, a partir de un contrato de transferencia del conocimiento con la Universitat de València, bajo la dirección y coordinación de Lucía Martínez Garay. Sus resultados son, además, y en parte, fruto de la tarea de investigación realizada por la mencionada profesora y el resto del equipo en varios proyectos de investigación («Algorithmical Law», PROMETEU/2021/009; PID2021-123441NB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y «ERDF A way of making Europe»; y «Digital economy regulation: equality guarantees provided by public powers and algorithmical tools», PID2019-108745GB-I00). En todo caso, los análisis, las conclusiones y las opiniones expresadas en este informe son únicamente los de sus autores y no representan los de las instituciones a las que pertenecen ni los de las que han financiado estas investigaciones, así como tampoco a la ONG que lo ha encargado.

Este informe no analiza todas las herramientas de esta índole que en estos momentos emplean las Administraciones españolas, sino únicamente tres de ellas, que seleccionó la ONG que encargó el informe a partir de un listado más amplio proporcionado por el equipo investigador. Fue también la ONG quien definió los aspectos que debían ser analizados en cada una. En concreto, se han estudiado dos herramientas algorítmicas utilizadas por la policía (VeriPol, empleada para la detección de posibles denuncias falsas de robo; y VioGén, utilizada para estimar los niveles de riesgo para las víctimas de violencia de género, a fin de decidir mejor las medidas de protección policial necesarias en cada caso) y una tercera que utiliza la Administración autonómica catalana para la gestión de sus prisiones (RisCanvi, que estima tanto el riesgo de reincidencia como el de otras conductas peligrosas por parte de la población reclusa).

2. CONDICIONANTES Y LÍMITES DE LA REALIZACIÓN DEL INFORME

Este informe fue finalizado y entregado en noviembre de 2022, aunque su publicación por parte de la ONG que lo encargó, junto al resto de informes de otros países europeos, aún no ha tenido lugar. No obstante, en esta versión se han añadido, aunque de manera sucinta y fragmentaria, ciertos datos e información sobre VioGén y RisCanvi obtenidos hasta abril de 2023, así como algunas referencias a la auditoría que las propias autoridades autonómicas catalanas encargaron respecto de RisCanvi y que ha sido publicada en el primer trimestre de 2024. Estas fechas han condicionado también el marco normativo de referencia respecto al cual se ha evaluado la compatibilidad de las herramientas con los derechos fundamentales: las regulaciones española y europea vigentes en materia de empleo de herramientas

automatizadas por los poderes públicos, legislación sobre modelos predictivos y normativa sobre protección de datos de carácter personal, vigente en 2022. Por esta razón, el informe no evalúa el grado de cumplimiento de las herramientas con las obligaciones que el Reglamento Europeo en materia de Inteligencia Artificial (AI Act), aprobado por el Parlamento Europeo el 13 de marzo de 2024 y publicado por el *Diario Oficial de la Unión Europea* en junio de 2024, establece respecto del uso de sistemas algorítmicos por parte de los poderes públicos. Téngase en cuenta, a estos efectos, que la limitada definición de «Inteligencia Artificial» que el artículo 3 de este reglamento contiene en su versión final, según la cual esta requiere de un sistema diseñado para funcionar no solo con diversos niveles de autonomía, sino que, además, sea capaz de mostrar capacidad de adaptación tras su despliegue, es decir, una capacidad de evolución/aprendizaje mínimamente autónomo, hace que de las tres herramientas analizadas solo VeriPol pueda entenderse que entra en la actualidad en sentido estricto en esta definición y debería, por tanto, cumplir con todas las previsiones del reglamento cuando entre en vigor para los usos de IA por parte de los poderes públicos (lo que no está previsto hasta dentro de unos años por el propio reglamento, en todo caso, lo que hace que este análisis sea más importante en cuanto a pautas y directrices que aplicar que de cumplimiento normativo estricto). Por lo demás, esto podría cambiar en un futuro, dado que hay, en el caso de VioGén, evoluciones en proceso dirigidas a incorporar moderadamente sistemas de IA para la mejora evolutiva de la herramienta a fin de incrementar su capacidad predictiva, y de que respecto de RisCanvi se han propuesto también modificaciones en este sentido (como resultado de la auditoría externa encargada por la Generalitat). En todo caso, la mayor parte de los juicios de valor sobre adaptación de las herramientas analizadas al marco jurídico y al cumplimiento de ciertas exigencias de transparencia y de auditabilidad se corresponden con el grueso de las exigencias que esta norma jurídica de derecho europeo ha incorporado para los supuestos de utilización de IA por parte de los poderes públicos. Pues entendemos que, si bien no como marco normativo cuyo cumplimiento sea ya directamente exigible a los poderes públicos, sí constituyen guías y directrices que, como mínimo, han de ser tenidas en consideración cuando emplean estas herramientas.

Por último, queremos señalar que la información sobre la que se ha elaborado el informe es en parte de acceso público, y en parte nos ha sido proporcionada o bien por las instituciones que gestionan cada una de las herramientas analizadas, o bien por personas que las utilizan o conocen por su actividad profesional. Respecto de esta segunda clase de información, solo hemos reproducido aquí los datos o extremos que no estaban expresamente cubiertos por un acuerdo de confidencialidad. También hemos mantenido en todo caso el anonimato de todas las fuentes empleadas cuando así lo han requerido y cuando su identificación personal no era relevante a efectos del informe realizado.

3. CONCLUSIONES GENERALES APLICABLES A LAS TRES HERRAMIENTAS ANALIZADAS

Las principales conclusiones que se extraen del informe, y que son predicables de las tres herramientas analizadas, son las siguientes.

- En general, todas ellas cumplen con el marco normativo vigente en España en materia de empleo de herramientas algorítmicas y sistemas automatizados para el apoyo a la acción administrativa, lo que en gran parte tiene que ver no con un diseño de estas particularmente riguroso y exigente en cuanto a sus garantías, sino con la muy limitada ambición del marco jurídico español vigente en esta materia. Igualmente, no se han detectado incumplimientos relevantes respecto de las obligaciones de derecho europeo y nacional en materia de protección de datos de carácter personal, lo que también es consecuencia de que resulta relativamente fácil cumplirlas en estos casos porque el marco no es particularmente exigente con los poderes públicos (que han de cumplir con las exigencias técnicas, pero ni siquiera siempre con la necesidad de consentimiento, que puede ser sustituido por una habilitación legal). En todo caso, en el informe se proponen mejoras en todos los casos en esta materia, que tienen que ver con una mejor auditoría y evaluación de los riesgos, que no parece haber sido realizada con la intensidad y profundidad exigida por el marco jurídico vigente, tanto en uno como en otro ámbito, cuando estamos ante sistemas que pueden tener una afección notable a los derechos de los ciudadanos. Asimismo, entendemos que hay que garantizar mejor el derecho de acceso de los ciudadanos a sus datos personales empleados por el sistema.
- El principal punto negro que comparten las tres herramientas, si bien en diversos grados, es la falta de transparencia, información y suministro de datos completos sobre su funcionamiento, que facilite una auditoría y control externos capaces de detectar eventuales sesgos discriminatorios o riesgos para los derechos de los ciudadanos, por una parte; y que permita, por otra, evaluar mejor la capacidad predictiva y su eficacia en la prevención del delito, sin tener que confiar para ello únicamente en los informes y evaluaciones de las propias administraciones públicas que las emplean o han diseñado. En relación con este punto, es importante, a nuestro juicio, diferenciar dos cuestiones. Una cosa es que el diseño, la implementación o incluso el control del funcionamiento de estos sistemas hayan sido rigurosos, realizados por especialistas, muchas veces en saludable colaboración entre las administraciones y las universidades –lo que es sin duda una primera garantía de que las herramientas están construidas con rigor y calidad desde el punto de vista técnico–, y otra bien distinta que a pesar de ello sigan siendo necesarios tanto la transparencia de cara a la ciudadanía como el control externo por parte de terceros no implicados en el diseño ni el funcionamiento diario de los sistemas. Lo primero, porque tanto el diseño de estas herramientas como el establecimiento de las pautas sobre cómo usarlas implican decisiones tanto técnicas como sobre qué fines deben ser priorizados, y qué costes se considera aceptable asumir para alcanzar aquellos fines. Decisiones que son por lo tanto, en última instancia,

de naturaleza política, y sobre las cuales los ciudadanos tienen derecho a estar informados, por razones de confianza de la ciudadanía en las herramientas y en los poderes públicos que las emplean, y para poder controlar la acción de estos. Especialmente en el caso de herramientas que, como las tres analizadas, usan por los poderes públicos como apoyo para tomar decisiones que pueden redundar en restricciones de derechos fundamentales. Y lo segundo, porque también un principio básico del método científico es la publicidad de los métodos, de los datos y de los resultados, para que puedan ser confirmados y refutados por otros investigadores, lo que redundaría en la detección y corrección de errores, la discusión y el debate bien informados y, en último término, la formulación de críticas y propuestas de mejora y el avance del conocimiento. Si una de las razones que se alegan para utilizar algoritmos en apoyo de la toma de decisiones es que la cuantificación y matematización de la información produce decisiones más objetivas y eficaces, y mejores porque están basadas en la evidencia científica, debe ser posible que dicha mayor objetividad y eficacia sean contrastables científicamente, lo que requiere compartir los datos, los códigos y los resultados.

- A este respecto, y de acuerdo con la información disponible, las tres herramientas parecen *a priori* construidas y diseñadas de manera seria y rigurosa. Sin embargo, todas presentan problemas en relación con la transparencia hacia los ciudadanos y hacia la comunidad de expertos, porque faltan datos esenciales para poder realizar una auditoría externa completa totalmente fiable: no se encuentran disponibles todos los *dataset*, no se conoce el diseño del algoritmo ni el peso relativo de cada uno de los factores predictivos y, en el caso de VeriPol, es incluso complicado saber si se usa de manera generalizada y consistente por parte de las autoridades. Dicho esto, es necesario no obstante singularizar el caso de RisCanvi, que merece un juicio aparte y mucho más favorable, en la medida en que la Generalitat de Catalunya está haciendo esfuerzos muy apreciables en los últimos tiempos para suministrar en abierto mucha más información sobre la herramienta y su funcionamiento, como destacamos en el apartado correspondiente.
- Además, la lógica del Reglamento Europeo de Inteligencia Artificial, al que ya hemos hecho referencia, obligaría a una mayor transparencia para el caso de que estas herramientas incorporaran en el futuro IA con capacidad de aprendizaje y adaptativa, y ello en dos sentidos. Por un lado, en cuanto a las obligaciones de una mejor y más completa explicación de la lógica de funcionamiento de las herramientas en abierto con carácter general. Y, por otro lado, y con mucho más detalle, respecto de la información que será obligatorio suministrar a los controladores públicos y autoridades de verificación privadas, para que las administraciones públicas en el ámbito de la Unión Europea puedan emplear sistemas de esta naturaleza.
- En cuanto a la capacidad predictiva y a la eficacia para alcanzar el objetivo propuesto de reducir la comisión de ciertos delitos o la prevalencia de determinadas conductas, es difícil pronunciarse de manera general, porque la cantidad y la calidad de la información disponible sobre cada herramienta analizada son muy dispares, por lo que remitimos al análisis efectuado en el

informe respecto de cada una en concreto. Lo único que destacaríamos aquí como consideración general es que las alegaciones de los responsables del despliegue de estos sistemas, en el sentido de que han mejorado los resultados previos a su introducción, parecen consistentes (con la excepción de VeriPol, donde es imposible cualquier verificación de estos extremos por la falta de información disponible), aunque tienen limitaciones importantes: en algunos casos, la información y los datos sobre los que se sustentan estas afirmaciones son muy escasos, y en general es difícil hacer una comparación con lo que ocurriría en ausencia de estos sistemas. Pero, en todo caso, echamos en falta una reflexión más profunda que compare lo que se gana en términos de prevención del delito con el coste en cuanto a restricción de derechos fundamentales, para poder realizar un juicio sobre la proporcionalidad del empleo de estos sistemas. Juicio que además, en nuestra opinión, debería ser diferenciado según la decisión para la que se utilice el algoritmo: un determinado nivel de acierto en las estimaciones de riesgo puede ser suficiente para tomar decisiones beneficiosas para el condenado o acusado, pero quizá no para tomar decisiones que restrinjan sus derechos; o un determinado nivel de acierto puede bastar para adoptar medidas a escala policial en favor de la víctima, y sin embargo sería discutible que bastara para imponer medidas cautelares al acusado. Consideramos que en este aspecto queda mucho por hacer, y que para que este necesario debate sea siquiera posible es imprescindible un mejor y más completo acceso a la información sobre los algoritmos, su funcionamiento y sus resultados, como hemos señalado *supra*.

- En cuanto a posibles sesgos discriminatorios o impacto dispar de las estimaciones de riesgo sobre grupos diferentes de población, la información publicada por los responsables de estas herramientas es mucho más escasa aún que la disponible sobre capacidad predictiva, y también lo son los datos a partir de los cuales poder hacer una evaluación externa (si bien de nuevo aquí hay importantes diferencias entre las tres herramientas analizadas, siendo RisCanvi –otra vez– la que ofrece más y mejor información; y VeriPol, la más opaca). Nosotros hemos detectado algunos problemas en este sentido, tanto en VioGén como en RisCanvi (y los intuimos en VeriPol, pero aquí la ausencia de datos impide formular nada más que meras intuiciones), si bien las conclusiones que hemos alcanzado han de ser tomadas con mucha cautela, porque están elaboradas sobre una base muy endeble, como hemos tratado de subrayar a lo largo del informe. Ahora bien, precisamente esto es lo que a nuestro juicio resulta más preocupante: que hoy en día (y con la importante excepción parcial de RisCanvi, como destacamos en el apartado correspondiente) resulte imposible estudiar algo tan relevante como los eventuales sesgos o efectos discriminatorios que estas herramientas pueden estar produciendo –o no–, habida cuenta de la ausencia de datos de acceso público con los que trabajar. Que una cuestión de esta relevancia no pueda ser investigada por personas externas a quienes son responsables de la gestión de estas herramientas (y que, por otra parte, ni siquiera ha sido casi estudiada hasta ahora por estos responsables) nos parece un problema grave.

4. PRINCIPALES CONCLUSIONES RESPECTO DE VIOGÉN

Por lo que se refiere a VioGén, las principales conclusiones del estudio son la ya referida necesidad de una mayor información sobre su funcionamiento. Si bien hay varios trabajos sobre distintos aspectos del funcionamiento de la herramienta publicados en revistas académicas de prestigio, eso no permite el tipo de control externo necesario al que nos hemos referido *supra*. En particular, con VioGén podría seguirse un modelo de facilitación de datos anonimizados en abierto similar al que se ofrece, sin problema alguno, con RisCanvi (véase *infra*). O podrían incluirse en el portal estadístico de la web de la Delegación del Gobierno contra la Violencia de Género muchos más datos de los que ahora tiene, que son escasos y aportan poca información. Ello impide una evaluación rigurosa y fiable del funcionamiento de la herramienta desde fuera, tanto en cuanto a su eficacia y fiabilidad como respecto de la identificación de posibles problemas y riesgos.

Con todo, de la información existente se desprende que las afirmaciones de sus responsables sobre el buen funcionamiento de la herramienta en cuanto a su capacidad para reducir la revictimización de las mujeres que denuncian haber sufrido violencia de género parecen consistentes con los datos disponibles en abierto, y con la ampliación de estos que nos ha sido proporcionada. En particular, creemos necesario ser prudentes en relación con las críticas que ha recibido VioGén por no ser capaz de detectar mejor los casos en que la mujer está en riesgo de homicidio a manos de su pareja o expareja. En nuestra opinión, el Ministerio ha realizado un indiscutible esfuerzo por mejorar la predicción de esos casos, que por otra parte son muy difíciles de detectar dada su bajísima prevalencia, e implican siempre una tasa muy elevada de falsos positivos. Sin perjuicio de las mejoras que puedan seguir haciéndose en este ámbito (y que no necesariamente tienen que pasar siempre por la actuación policial o judicial), nos parece necesario ser conscientes de los límites a que la predicción de este tipo de eventos está sometida, y reconocer que ningún sistema de estimación de riesgos puede aspirar razonablemente a reducir estas tragedias a cero.

La capacidad predictiva de VioGén es buena por lo que hace a la detección del riesgo bajo (en el sentido de que en la inmensa mayoría de los casos de riesgo bajo no hay agresiones posteriores), pero más débil por lo que se refiere a los casos de riesgo alto, pues muchos casos en los que se afirma riesgo alto o extremo no van seguidos tampoco de agresiones ulteriores. Sin perjuicio de reconocer que la propia implementación de medidas de protección contribuye seguramente en parte a este efecto, hay que ser conscientes de que el sistema, como reconocen sus propios responsables, asume un nivel elevado de falsos positivos para poder maximizar la sensibilidad (esto es, no dejar sin protección a ninguna mujer que realmente pueda estar en riesgo). Y esta opción, que policialmente no es problemática desde la perspectiva de los derechos fundamentales, porque las medidas de protección de las mujeres inciden muy poco en los derechos de los varones denunciados, resultaría sin embargo mucho más cuestionable si se utilizaran los niveles de riesgo de manera automática en el ámbito judicial para fundamentar la imposición de medidas cautelares, especialmente las más restrictivas de derechos. Esta utilización generalizada no está ocurriendo hoy en día, al menos hasta donde hemos podido

averiguar, pero es algo frente a lo que conviene estar alerta porque, en primer lugar, ya en la actualidad es posible tener en cuenta el resultado de VioGén para tomar estas decisiones, aunque no se perciba como determinante; y, en segundo lugar, porque es probable que este uso pase a ser mayor en el futuro. En tal caso, el juicio emitido en este sentido debiera ser, como es obvio, notablemente revisado.

Por lo que hace a posibles sesgos o efectos discriminatorios de VioGén, ya hemos dicho que la muy limitada información disponible dificulta mucho poder analizar esta cuestión con un mínimo de rigor. Nosotros hemos detectado indicios que apuntan a una protección menor, o menos eficiente, de las mujeres nacidas fuera de España, tanto frente a nuevas agresiones de género como frente a homicidios. Señalamos en el estudio la metodología empleada para analizar los datos disponibles y las razones por las que llegamos a esta conclusión, que ciertamente se basa en información incompleta, pues incompleta es la información disponible al respecto, pero que consideramos que, en ausencia de más datos, y más completos, debería llevar al menos a una revisión del funcionamiento de la herramienta, tanto para la estimación del riesgo de violencia en general como para la estimación del riesgo de homicidio con la nueva escala H.

En cuanto al cumplimiento normativo, se echan en falta evaluaciones más completas de riesgos, tanto en cuanto a los datos de carácter personal empleados y los riesgos que de su uso puedan derivarse como, sobre todo, en materia de las evaluaciones sobre riesgos para derechos fundamentales y de prevención de sesgos que un sistema de este tipo debería incorporar. De hecho, en el modelo de cumplimiento del Reglamento Europeo en materia de inteligencia artificial estas evaluaciones pasarán a ser obligatorias para sistemas de esta índole desde el momento en que incorporen IA con capacidad de evolución o aprendizaje (lo que la versión actual de VioGén parece que no hace, pero que en cuanto pase a ser empleada una que sí lo haga deberá adaptarse a esta exigencia). Por último, también son susceptibles de una mejora importante los mecanismos que garanticen el acceso efectivo de los ciudadanos a sus evaluaciones y a todos sus datos de carácter personal almacenados por el sistema, algo que en la actualidad no se garantiza suficientemente.

5. PRINCIPALES CONCLUSIONES RESPECTO DE RISCANVI

RisCanvi es la herramienta más transparente de las analizadas. Si bien el diseño de los algoritmos se mantiene secreto (con la excepción del que estima el riesgo de reincidencia violenta, que hemos podido conocer gracias a la colaboración de colegas de la Universitat Pompeu Fabra), la Generalitat de Cataluña ha ido poniendo en práctica, de manera creciente con los años, una política de *open data* con publicación de datos anonimizados en formato Excel en su página web, que es reseñable en un sentido muy positivo, que creemos que debería servir de modelo a otras administraciones, y que permite analizar el funcionamiento de los algoritmos a cualquier investigador interesado sin comprometer en modo alguno la privacidad de las personas afectadas por las estimaciones de riesgo. Además, la herramienta ha sido auditada en cuanto a su funcionamiento recientemente, y los resultados de esta evaluación han sido también publicados. Desde este punto de vista, si en un futuro la

herramienta incorporara para la mejora de sus estimaciones de riesgo sistemas de IA en el sentido definido por el Reglamento europeo, probablemente podrá cumplir las exigencias de transparencia y auditabilidad del sistema con bastante facilidad.

Queremos subrayar que de los cinco tipos de riesgo que estima RisCanvi, en este informe solo hemos analizado el de reincidencia violenta, por lo que todas nuestras consideraciones se limitan exclusivamente a este aspecto.

La capacidad predictiva de RisCanvi es bastante buena en cuanto a la detección de las personas con riesgo bajo, pero mucho menor en cuanto a la detección de las personas con riesgo alto. Las estimaciones de riesgo incluyen un número elevado de falsos positivos y en esto coinciden tanto las investigaciones llevadas a cabo desde organismos dependientes de la Generalitat, como los cálculos que nosotros hemos podido realizar a partir de los datos disponibles, así como los resultados de la auditoría externa. Creemos que los responsables de la herramienta son conscientes de esta realidad y están extrayendo de ella la conclusión, razonable desde nuestro punto de vista, de que las evaluaciones de riesgo deberían conducir a una suavización de los regímenes de vida y a mayores posibilidades de acceso a la libertad para las personas de riesgo bajo, pero no llevar automáticamente a un endurecimiento del régimen ni de las condiciones de cumplimiento para aquellos clasificados como de riesgo alto. Sería deseable, no obstante, que existieran pautas claras de actuación a este respecto, y también que este conocimiento se extendiera al ámbito judicial, donde se revisan las decisiones de clasificación de las juntas de tratamiento, y donde la Fiscalía tiende a dar una importancia muy elevada a las clasificaciones de riesgo alto o medio, sin tener en cuenta las importantes limitaciones que esas estimaciones tienen.

Algunos de los factores que RisCanvi tiene en cuenta para evaluar el riesgo afectan muy directamente a la vida privada e intimidad de las personas y a su libre desarrollo de la personalidad o a su entorno de socialización. Puesto que no son públicos los algoritmos (con la excepción antes mencionada) no se puede saber cuál es la capacidad predictiva de esta información tan sensible. Y sin embargo puede resultar cuestionable que aspectos sobre los que la persona no tiene control (antecedentes delictivos de sus familiares, por ejemplo) o forman parte de su derecho al libre desarrollo de la personalidad (número y tipo de actividades sexuales preferidas, por ejemplo) puedan constituir una base legítima para fundamentar medidas restrictivas de derechos (por ejemplo, la denegación del acceso a un tercer grado o a la libertad condicional). Sería necesario comparar el rendimiento de estos ítems (es decir, cuánto se pierde en cuanto a capacidad predictiva, si se eliminan del algoritmo) con su coste en términos de afectación a los derechos (a la intimidad, a la igualdad, al respecto a la vida privada y familiar, etc.), para tomar una decisión sobre si resulta proporcionado emplearlos en las predicciones. Además, en algunos de los ítems puede incluso cuestionarse que sea pertinente tenerlos en cuenta incluso en el supuesto de que tengan valor predictivo (por ejemplo, la presencia de discapacidad). Creemos que sería necesaria una auditoría profunda respecto del riesgo que comportan estos elementos para los derechos fundamentales de los reclusos, que no nos consta que se haya llevado a cabo, así como una transparencia mucho mayor sobre las evidencias empíricas que, en su caso, podrían avalar que se entiendan estos factores como estadísticamente significativos.

Por último, nuestro análisis estadístico ha revelado cierto trato extraordinariamente severo para algunos colectivos de reclusos, que no parece justificado y que podría retroalimentar sesgos en la herramienta en perjuicio de estos colectivos, esencialmente las personas con problemas de salud mental o de adicciones, o personas con una situación socioeconómica muy desfavorable. No es el caso de Ris-Canvi, y creemos que esto es una virtud propia que debe ser destacada, en cuanto a un trato más severo con los extranjeros, que no se da según los datos disponibles. Entendemos que el funcionamiento de la herramienta debería al menos revisarse a partir de los datos estadísticos que aportamos, y valorarse alguna vía para solucionar estos problemas.

6. PRINCIPALES CONCLUSIONES RESPECTO DE VERIPOL

VeriPol es la herramienta de la que menos información se dispone, lo que agrava el problema de falta de transparencia y las dificultades, cuando no imposibilidad, para realizar la tarea de control externo que hemos pretendido con este informe. Agradecemos a la Policía Nacional la información y los datos que nos proporcionó para al menos podernos hacer alguna idea sobre esta herramienta y el tipo de uso que se le da en la práctica, pero aun así se trata de una situación muy deficiente en cuanto al cumplimiento de los estándares mínimos de transparencia que consideramos requeridos para el empleo de herramientas de este tipo, y aunque en el informe hemos analizado toda la información de la que hemos podido disponer, muchas conclusiones no pasan de ser conjeturas, a falta de datos contrastables con los que poder trabajar.

En cuanto a los problemas que pueden derivarse de su afección a derechos fundamentales o en cuanto a la posible generación de sesgos, esta misma falta de información impide una valoración externa sobre el funcionamiento práctico del sistema. El informe ha tratado, al menos, de establecer un marco teórico respecto de las cautelas mínimas exigibles para evitar que modelos como este, basados en IA y en aprendizaje a partir de sistemas de procesamiento de lenguaje natural, incurran en sesgos para las funciones que se le presuponen (detección de denuncias falsas por parte de los ciudadanos y ciudadanas), que, por ejemplo, puedan afectar más a la parte de la población que maneja peor el idioma por razones que pueden ser de muy diverso tipo (nivel socioeconómico, origen, nivel de estudios e incluso la mera diferencia diatópica dialectal en aquellos casos en que pueda emplearse una variedad diferente a la que ha sido mayoritariamente empleada para entrenar el modelo). Asimismo, y como para el resto de las herramientas, trata de sugerir mejoras para una mejor evaluación de los posibles riesgos para los derechos fundamentales que puedan derivarse de su utilización, riesgos tanto mayores cuanto más extendido sea su uso.

VIOGÉN

1 DESCRIPCIÓN GENERAL

Antecedentes y ámbito de aplicación

En 2007, la Secretaría de Estado de Seguridad (en adelante, SES) del Ministerio del Interior desarrolló e implantó el «Sistema de Seguimiento Integral en casos de violencia de género» (Sistema VioGén), que integra las valoraciones policiales del riesgo de revictimización y las medidas para la protección de las mujeres víctimas en función del nivel de riesgo evaluado. En el diseño inicial del sistema, y en las sucesivas actualizaciones y estudios realizados sobre el funcionamiento de VioGén, han participado tanto expertos de los cuerpos policiales como académicos.

Las evaluaciones de riesgo se realizan a través de dos formularios: el VPR (valoración policial de riesgo) y el VPER (valoración policial de la evolución del riesgo), ambos integrados en una aplicación web. Solo los agentes autorizados de las Fuerzas y Cuerpos de Seguridad (Policía Nacional y Guardia Civil,¹ y las policías autonómicas y locales que voluntariamente se adhieran) pueden realizar las evaluaciones de riesgo, pero a los resultados y a la información contenida en VioGén pueden acceder otros muchos órganos administrativos implicados, bien en la protección y atención a las víctimas, bien en la supervisión y control de los agresores (instituciones penitenciarias, juzgados, institutos de Medicina Legal y Ciencias Forenses, oficinas de asistencia a la víctima, fiscalías, delegaciones y subdelegaciones del Gobierno y, por último, servicios sociales y organismos de igualdad de las distintas comunidades autónomas). Desde 2020 el personal de los institutos de Medicina

¹ La Guardia Civil es, junto con la Policía Nacional española, uno de los dos cuerpos de seguridad de ámbito nacional que existen en España. La Policía Nacional es responsable de las zonas urbanas, mientras que la Guardia Civil es responsable de las zonas rurales.

Legal y Ciencias Forenses (en adelante, IMLCF) también puede utilizar el formulario VPR para realizar valoraciones forenses de riesgo cuando un juez así se lo requiera.

Ni el VPR ni el VPER se utilizan en las comunidades autónomas de Cataluña y el País Vasco, que aplican sus propias herramientas de evaluación (RVD-BCN en el primer caso, y EPV-R en el segundo).

La variable estimada por VioGén es un nuevo acto de violencia de género, entendida como cualquier tipo de violencia: física, sexual, psicológica o amenazas, incluido el incumplimiento de resoluciones judiciales. La violencia de género es entendida en VioGén como la violencia ejercida sobre la mujer por su cónyuge actual, excónyuge o por quienes mantengan o hayan mantenido con ella una relación de afectividad análoga, aun sin convivencia, de acuerdo con lo previsto en el artículo 1 de la Ley Orgánica 1/2004, de Medidas de Protección Integral contra la Violencia de Género. Para VioGén, por tanto, violencia de género es sinónimo de violencia de pareja.

Desde la primera versión en 2007, el VPR y el VPER han sido revisados y actualizados en varias ocasiones, hasta la versión 5.0 del VPR y la versión 4.1 del VPER, actualmente en vigor (Instrucción 4/2019, de la Secretaría de Estado de Seguridad, por la que se establece un nuevo protocolo para la valoración policial del nivel de riesgo de violencia de género). La última actualización del formulario VPR (5.0) incluye además una escala específica para estimar la probabilidad de violencia muy grave o letal, así como alertas específicas para cuando se detecten menores en situación de vulnerabilidad o menores en riesgo, a las que nos referiremos más adelante. Según una nota de prensa del Ministerio del Interior, realizada con motivo del 15 aniversario de VioGén en julio de 2022 (<https://links.uv.es/EM8m6QQ>), el sistema había analizado hasta esa fecha 701.563 casos y realizado más de 5,4 millones de evaluaciones policiales de riesgo desde su creación. Del total de casos registrados en esos 15 años, el 10,4 % permanecían activos (73.072) y recibían protección policial. El sistema cuenta con 2.839 agentes que ofrecen protección a más de 73.000 mujeres y sus hijos. Además, más de 39.000 usuarios están autorizados a interactuar con VioGén.

Fuentes de información para este informe

VioGén se utiliza principalmente para las valoraciones de riesgo que realizan los agentes policiales cuando la mujer víctima de violencia de género presenta una denuncia. Pero también puede ser utilizado (desde 2020) por el personal de los institutos de Medicina Legal y Ciencias Forenses (IMLCF) para realizar valoraciones forenses (que son más exhaustivas que las realizadas por la policía) cuando así lo solicite un juez.

Sobre la base del nivel de riesgo indicado por los formularios VPR y VPER, la policía adopta medidas para proteger a la mujer frente a una nueva victimización. Pero VioGén también puede utilizarse para adoptar medidas de protección a escala judicial: el juez puede tener en cuenta ambas evaluaciones de riesgo (policial y forense) a la hora de ordenar medidas cautelares como órdenes de alejamiento o incluso prisión provisional.

Dado que hay diferentes agentes implicados en las decisiones que pueden tomarse a partir de las valoraciones de riesgo efectuadas con VioGén, hemos buscado información para este informe en distintas fuentes:

- Nos pusimos en contacto con la Secretaría de Estado de Seguridad y le pedimos muestras de evaluaciones policiales de riesgo y también de evaluaciones forenses de riesgo.
- Hemos revisado los datos públicos disponibles sobre VioGén que se pueden encontrar en las páginas web del Ministerio del Interior y del Ministerio de Igualdad.
- Solicitamos al Consejo Médico Forense datos e información sobre las evaluaciones forenses de riesgo.
- Nos pusimos en contacto con la Asociación de Psicología Forense de la Administración de Justicia para pedir información sobre las evaluaciones forenses de riesgo programadas.
- Buscamos en bases de datos las resoluciones judiciales que mencionaran VioGén y, como había muy pocas, pedimos permiso al Tribunal Superior de Justicia de la Comunitat Valenciana para acceder a las resoluciones dictadas por los Juzgados de Violencia sobre la Mujer en la Comunitat Valenciana.

Dado que en cada uno de los casos nos hemos encontrado con diversas dificultades, que sería demasiado largo explicar aquí en detalle, las hemos resumido en el anexo *Dificultades para obtener información y datos sobre VioGén*, y en los siguientes apartados nos referimos a los concretos datos y fuentes de información utilizados a la hora de analizar cada uno de los aspectos de VioGén. En cualquier caso, queremos señalar, por un lado, que en varias ocasiones no hemos podido obtener datos cuantitativos –o estos han sido escasos e incompletos–, por lo que hemos tenido que trabajar solo o principalmente con información cualitativa (entrevistas a policías, personal de los institutos de Medicina Legal, Jueces, etc.). Por otro lado –y lo mismo cabe decir de las otras dos herramientas examinadas en este informe–, las dificultades en la obtención de información y datos condicionan la calidad y solidez de nuestras conclusiones, algo que también hemos tratado de poner de manifiesto a lo largo de las páginas que siguen en el análisis de las distintas cuestiones abordadas.

2. DATOS

Datos de entrada

Datos que se introducen en el sistema

La versión actual del VPR 5.0 incluye 35 indicadores de riesgo agrupados en 5 dimensiones (historia de violencia, características del agresor, vulnerabilidad de la víctima, circunstancias relacionadas con menores y circunstancias agravantes); el formulario VPER 4.1 incluye 37 indicadores agrupados en 6 dimensiones algo diferentes.

La descripción detallada de todos los indicadores de riesgo puede encontrarse en la Guía de Procedimiento VPR 5.0 y VPER 4.1 (véase Fuentes); en los anexos sobre VioGén del presente informe ofrecemos una descripción resumida de estos.

Recogida de datos y exactitud

La valoración policial del riesgo con el formulario VPR se lleva a cabo cuando la policía recibe una denuncia por un acto de violencia de género. La principal fuente de información es, por tanto, la víctima, aunque se anima a los agentes de policía a contrastar esta información con otras fuentes, como los agentes de policía que hayan intervenido en el caso, el agresor, vecinos o testigos, familiares y, si están disponibles, informes técnicos como los médicos (las fuentes utilizadas deben anotarse en el formulario). Y el agente que rellena el formulario VPR siempre tiene que indicar a qué fuentes de información ha podido acceder. Sin embargo, la principal fuente de información siempre ha sido y continúa siendo la víctima, lo que significa que cuando la mujer no quiere denunciar la protección que recibe es menor, o nula. Para hacer frente a esta situación (alrededor del 70 % de las mujeres que son asesinadas por sus parejas no habían presentado denuncia previa) la Instrucción 4/2019 SES ya establecía que la policía debía investigar todos los casos de violencia de género de los que tuviera conocimiento, así como realizar la valoración del riesgo, aunque no hubiera denuncia. Más recientemente la Instrucción 5/2021 SES ha establecido lo que se denomina «Protocolo 0», con instrucciones detalladas para los policías sobre cómo proceder en estos casos.

Al introducir los datos en el sistema, los agentes de policía tienen instrucciones de no formular preguntas directas a la víctima, sino dejar que se explique y extraer toda la información que necesiten de lo que ella está narrando. La Guía de Procedimiento para el uso del VPR y el VPER publicada por el Ministerio del Interior dice que

es muy importante dejar que la mujer empiece relatando de forma libre y espontánea, evitando así la introducción de sesgos. Si es necesario obtener aclaraciones o más información, el agente podrá preguntar a la víctima, de forma flexible y en el momento oportuno. Las preguntas no deben hacerse como si se tratase de la administración de un cuestionario, sino con fórmulas de preguntas narrativas, abiertas, empezando con una repetición de las mismas palabras de la víctima, lo que ayuda a obtener un relato ajustado a los hechos y proporciona seguridad a la víctima como muestra de escucha activa (Guía de Procedimiento VPR 5.0 y VPER 4.1, p. 5).

Entrevistamos para este informe a dos policías (una de ellas trabaja en una oficina de denuncias y la otra tiene experiencia como agente de protección de la mujer), a las jefas de las unidades de valoración forense integral de los IMLCF de Valencia y Alicante, y a una persona que trabaja en una Oficina de Atención a las Víctimas del Delito, y todas ellas subrayaron que la mujer se encuentra en una situación de mucho estrés y angustia cuando presenta la denuncia ante la policía. Puede tener miedo (de que el varón se vuelva más agresivo por haber presentado ella la denuncia, o de perderlo por ello...), vergüenza (de lo que pensarán de ella su familia y sus hijos, de no haber sido capaz de resolver el problema por sí misma, de lo que lleva años soportando, etc.), estar en estado de *shock* si se acaba de producir una agresión violenta, no estar segura aún de lo que quiere (dejar a su pareja, intentar darle

otra oportunidad...), etc. Por eso, la información que da en ese momento puede ser incompleta, o puede no ser capaz de expresarse con claridad. Esta es una de las razones por las que, si días después se hace una valoración forense del riesgo, el resultado puede ser diferente al que obtuvo la policía, y eso no significa necesariamente que la policía no hiciera bien su trabajo.

¿Es probable que los datos contengan sesgos?

La Guía de Procedimiento publicada por el Ministerio del Interior contiene una descripción detallada de cada factor de riesgo (ítem), explicando qué información se debe obtener con cada uno de ellos, e incluso añadiendo ejemplos para orientar a los policías. Sin embargo, es interesante lo que nos comentaron las dos agentes de policía a las que pudimos entrevistar (y en este punto también coincidió la jefa de la Unidad de Valoración Forense Integral de Valencia): en su opinión, por muy bien descritos que estén los ítems y por mucho que se intente objetivar la recogida de información, prácticamente todos ellos requieren una valoración previa por parte del agente actuante, en la que inevitablemente influyen sus sesgos y prejuicios.

Por ejemplo, todos los ítems que requieren evaluar la gravedad de la violencia mostrada, o los que requieren valorar si el hombre es exageradamente celoso, muestra comportamientos controladores, o si se han producido relaciones sexuales forzadas, etc. En algunos de estos casos, qué comportamiento debe considerarse «voluntario» o «forzado», «normal» o «exagerado» depende de las creencias de cada persona sobre los papeles de hombres y mujeres, y sobre la dinámica de las relaciones de pareja. Por ejemplo, un encuentro sexual al que la mujer accede sin protestar y sin que el hombre utilice ninguna amenaza o violencia física, pero porque su negativa en ocasiones anteriores ha provocado gritos e insultos feroces que teme volver a soportar, podría ser valorado como «forzado» por unos y no por otros.

Marcar sí o no en alguno de estos ítems es también una imagen muy «fría» de la compleja realidad que puede estar viviendo la mujer, o de la situación psicológica exacta del hombre, cuyo conocimiento más detallado podría explicar mejor muchas cosas. La jefa de la Unidad de Valoración Forense Integral del IMLCF de Valencia comentó que, cuando se hace un informe forense para el juez, sí se incluye toda esta información más detallada, pero no se introduce en VioGén. Cuando se le preguntó si creía que debería estar en el sistema, expresó sus dudas de que las mujeres estuvieran satisfechas, ya que toda su vida privada (relaciones sexuales, parejas anteriores, problemas psicológicos y sociales, etc.) estaría en una base de datos accesible a miles de funcionarios del Estado, aunque sea para su protección.

Ambas policías (las dos eran mujeres) hicieron mucho hincapié en lo necesario que es haber participado en cursos de formación para entender cuál es la perspectiva de género, la situación de las mujeres en comisaría, cómo valorar correctamente sus declaraciones, etc., y subrayaron que la formación que imparte el Ministerio del Interior es de gran calidad. Pero también señalaron que estos cursos no pueden ser realizados por todo el mundo (en primer lugar, porque no hay suficientes para todos los policías, pero también por el hecho de que tienen unos requisitos previos que hay que cumplir).

Análisis y datos de salida

Procesamiento de los datos

Tanto el VPR como el VPER son herramientas actuariales que realizan una estimación automática del riesgo mediante un algoritmo. Existe ponderación, aunque el algoritmo se mantiene en secreto y no tenemos información sobre la fórmula actual (se puede encontrar alguna información sobre versiones anteriores en López-Ossorio, 2016). Parece que no todos los ítems de los formularios contribuyen al nivel de riesgo, algunos están ahí solo para dar más información sobre las situaciones de la víctima, por ejemplo, su país de origen (López Ossorio, 2016: 230).

El nivel de riesgo puede ser modificado al alza por los agentes de policía si, basándose en circunstancias no reflejadas en los factores de riesgo del formulario, lo consideran necesario para una mejor protección de la víctima.

El sistema fue diseñado, y desde entonces ha sido revisado varias veces, por expertos de los cuerpos policiales y del mundo académico.

¿Es un sistema de aprendizaje automático?

Ni el VPR ni el VPER son sistemas de aprendizaje automático. Sin embargo, varios académicos de universidades madrileñas y un miembro de la SES han publicado recientemente un estudio que compara la eficiencia predictiva del actual VioGén con la que resultaría de aplicar un sistema de aprendizaje automático (NC, *nearest centroid*). El estudio concluye que la eficiencia predictiva aumenta con el aprendizaje automático, e incluso diseña un modelo híbrido que permitiría una transición progresiva del modelo actual basado en estadística convencional al nuevo modelo de aprendizaje automático (González Prieto et al., 2021). Al preguntar sobre ello al personal de la SES, nos dijeron que se trata de un mero experimento y que no hay ningún plan para aplicarlo en la práctica.

Por otro lado, parece que al menos algunos de los procedimientos internos de VioGén se realizan con *machine learning*, o al menos se está implementando. Según una nota de prensa hecha pública el 15 de diciembre de 2020 por el Gobierno de España,

La Secretaría de Estado de Seguridad del Ministerio del Interior ha apostado por incorporar la analítica avanzada y la inteligencia artificial para reforzar la eficacia del proceso policial de valoración del riesgo en los casos de violencia de género. [...] El Área de Violencia de Género, Estudios y Formación de la Secretaría de Estado de Seguridad ha incorporado la plataforma analítica de la empresa de *software* SAS Iberia, que facilitará actualizaciones mucho más rápidas y eficientes del Protocolo de Valoración Policial del Riesgo del Sistema VioGén, ponderando mejor los actuales indicadores de riesgo de reincidencia e identificando nuevas variables que ayuden a afinar aún más esta valoración del riesgo mediante el análisis automatizado y en tiempo real de grandes cantidades de datos. [...] Hasta el momento, se han actualizado los algoritmos matemáticos de los dos formularios de valoración policial del riesgo de reincidencia del Sistema VioGén (VPR y VPER) mediante la selección de muestras de casos y el análisis manual de algunos indicadores de riesgo, tras su seguimiento durante varios años. El *software* de SAS Iberia

incorporará tecnología de analítica avanzada e inteligencia artificial que automatizará el análisis de una mayor cantidad de datos de criminalidad, incluso combinados con datos de fuentes abiertas, lo que ayudará a ponderar mejor los algoritmos, identificando nuevos indicadores de riesgo, y en periodos de tiempo mucho más cortos. Además, son algoritmos más sensibles a la evolución de la criminalidad y mejoran así la predicción de aquellos casos en los que es previsible la reincidencia («Interior recurre a la tecnología de inteligencia artificial para mejorar la valoración policial de riesgo en casos de violencia de género» (<<https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/interior/Paginas/2020/151220-inteligencia.aspx>>; consulta: 08/10/2022).

¿Cómo se presenta el resultado?

La estimación del riesgo la da automáticamente el sistema tras haber introducido la información los formularios en una escala ordinal de 5 niveles de riesgo: «no apreciado», «bajo», «medio», «alto» y «extremo».

En 2019 se introdujo una nueva escala en VioGén. En España, algunos casos de violencia de género que acabaron en homicidio, pero en los que la valoración policial había asignado un riesgo bajo o nulo, han generado una considerable alarma social. Para mejorar las predicciones en estos casos, y asumiendo la hipótesis de que los factores que predicen la violencia letal pueden no ser necesariamente los mismos que los que predicen cualquier tipo de reincidencia en violencia de género, se realizó un estudio que encontró que solo 13 de los 35 indicadores de riesgo del VPR correlacionaban de forma estadísticamente significativa con una mayor probabilidad de homicidio. Por ello, se desarrolló una escala H dentro del formulario VPR 5.0, que evalúa este riesgo específico de violencia muy grave o letal en paralelo a la estimación general del riesgo de reincidencia (López-Ossorio et al, 2021).

Cuando la escala H detecta un riesgo alto de homicidio, emite una advertencia automática de «caso de especial relevancia» y, además, si el nivel de riesgo de reincidencia general en el VPR había sido «no apreciado» o «bajo», el formulario lo eleva automáticamente a «medio» o «alto». La versión 5.0 del VPR incorpora además advertencias específicas para «menores en situación de vulnerabilidad» (cuando la mujer víctima manifiesta que cree que se encuentran en una situación objetiva de amenaza o que pueden estar en riesgo) y «menores en riesgo» (cuando se da la combinación de un caso de especial relevancia y el de menores en situación de vulnerabilidad, y además factores clínicos de riesgo en el agresor; López-Ossorio, Muñoz Vicente et al., 2020: 6). En todos estos casos el formulario recomienda que se solicite al fiscal y al juez una valoración forense del riesgo.

Capacidad predictiva, falsos negativos y falsos positivos

La capacidad predictiva ha sido testada en varios estudios, publicados por expertos que trabajan para el Ministerio y algunos académicos de distintas universidades, en revistas científicas con *peer review*.

Un trabajo publicado en 2016 (López-Ossorio et al., 2016) estudió una muestra de 407 casos, de los que 211 (51,8 %) presentaban riesgo «no apreciado» en la VPR, 157 eran casos de riesgo «bajo» (38,6 %), 38 casos de riesgo «medio» (9,3 %) y un

caso de riesgo «alto» (0,2 %). Tras un seguimiento de tres y seis meses de las nuevas denuncias presentadas contra el mismo delincuente, la tasa de reincidencia (medida como nuevas denuncias) fue del 4,91 % (20 casos) a los tres meses, y del 12,04 % (49 casos) a los seis meses. Los autores del estudio informan que el 81,6 % de los nuevos delitos denunciados fueron de violencia leve, y el 18,4 % de violencia grave (entendiendo como violencia tanto la física y la psicológica, así como la violencia sexual y el acoso) (López-Ossorio et al., 2016: 5).

Para el análisis de la eficacia predictiva del VPR, los autores combinaron los cinco niveles de riesgo en dos: riesgo NO (no apreciado) y riesgo SÍ (que incluía los niveles de riesgo bajo, medio y alto), que dio los siguientes resultados tras un seguimiento de tres meses:²

Tabla 1
Número de casos, nivel de riesgo en VPR y reincidencia registrada al cabo de 3 meses

<i>Riesgo estimado</i>	<i>Reincidencia SÍ</i>	<i>Reincidencia NO</i>	<i>Total</i>	<i>Valor predictivo</i>
SÍ	17	179	196	Positivo: 8,6 %
NO	3	208	211	Negativo: 98,5 %
Total	20	387	407	
	Sensibilidad 85 %	Especificidad: 53,7 %		

Fuente: elaboración propia a partir de los datos de la tabla 2 de López-Ossorio et al. (2016: 5).

¿Qué significan estos datos? La sensibilidad del 85 % significa que, de cada 100 mujeres que fueron agredidas de nuevo por su pareja o expareja, VioGén clasificó 85 como de riesgo (es decir, riesgo bajo, medio o alto). La especificidad significa que de cada 100 mujeres que no fueron agredidas de nuevo VioGén clasificó a 53,7 como de no riesgo. Los valores predictivos, por su parte, arrojan la siguiente información: de cada 100 mujeres que fueron consideradas por VioGén como de riesgo, el 8,6 % fueron nuevamente agredidas por su pareja o expareja, y de cada 100 mujeres que fueron consideradas por VioGén como casos de no riesgo, el 98,5 % no denunciaron una nueva agresión.

Un estudio más reciente (López-Ossorio et al., 2019) puso a prueba el funcionamiento de la versión 4.0 de VPR analizando 7.147 nuevas denuncias por violencia

² Los autores también proporcionan el valor del área bajo la curva ROC (0,714), así como un análisis mediante el modelo de regresión logística, y la ODDS ratio (6,585). A los seis meses, los valores de sensibilidad, ODDS ratio y área bajo la curva fueron inferiores: 61,2, 0,58 y 1,826 %, respectivamente (el estudio no proporciona valores predictivos para la estimación del riesgo a los seis meses).

de género, con un seguimiento de 10 meses (320 días). En este estudio, la variable «riesgo» se dicotomizó de forma diferente: riesgo no apreciado y riesgo bajo, por un lado, frente a riesgo medio, alto y extremo. La reincidencia policial fue del 8 % a los 3 meses y del 14,3 % a los 10 meses (56,5 % en los 3 primeros meses). Los resultados de este estudio fueron: sensibilidad, 79 %; especificidad, 35 %; VPP (valor predictivo positivo) 17 %; VPN (valor predictivo negativo) 91% (y AUC 0,62).

¿Qué conclusiones pueden extraerse de estos datos? VioGén parece tener pocos falsos negativos. Cuando el riesgo estimado es no apreciado o bajo, en más del 90 % de los casos no hay nuevas agresiones a la mujer. Este alto valor predictivo negativo apoya el uso de esta herramienta como criterio de exclusión, es decir, se puede confiar en que un resultado de ausencia de riesgo irá acompañado muy probablemente de una ausencia real de nuevas agresiones, y puede considerarse un buen argumento para no adoptar medidas limitativas de los derechos del acusado, al menos a corto plazo. VioGén también parece detectar bastante bien a las mujeres que están en situación de riesgo, ya que la sensibilidad ronda el 80 % en ambos estudios.

Los documentos oficiales y los estudios publicados sobre VioGén admiten expresamente que la herramienta está diseñada para maximizar la sensibilidad, es decir, para detectar como de alto riesgo el mayor número posible de casos en los que efectivamente la víctima vuelve a ser agredida, y así no dejar desprotegida a ninguna víctima que necesite protección policial (González Álvarez et al., 2018:50; López-Ossorio et al., 2019:56; López-Ossorio et al., 2020: 179, 180).

Por el contrario, y debido en parte a esto, VioGén tiene muchos falsos positivos. Cuando el riesgo estimado es medio, alto o extremo, menos del 20 % de las mujeres sufrirán una nueva agresión. ¿Significa esto que VioGén funciona mal en este sentido? La respuesta debe ser prudente. VioGén podría estar sobrestimando mucho el riesgo. O también podría ocurrir que las medidas de protección que se ponen en práctica para garantizar la seguridad de las víctimas (anexo *Medidas de protección policial para cada nivel de riesgo...*) consigan precisamente evitar la reincidencia que de otro modo se habría producido.

¿Es posible saber cuál es el caso? El mismo estudio de López-Ossorio et al. (2019) contiene información interesante sobre este punto. Consiguieron construir un grupo de control formado por 194 casos en los que, a pesar de haberse apreciado algún riesgo en el VPR, los casos fueron inactivados y no obtuvieron protección policial (sobre las causas de inactivación de casos, ver *infra*, apartado siguiente), pero fueron reactivados más de tres meses después debido a una nueva denuncia. La tasa de reincidencia en este grupo fue del 38,7 %, frente a una tasa del 13,6 % en los casos que sí habían tenido protección policial (López-Ossorio et al., 2019: 56).

Aunque este grupo de control es bastante reducido (n = 194), y no nos consta que se haya realizado un experimento similar en más ocasiones, lo que obliga a ser cautos a la hora de valorar la relevancia de este hallazgo, creemos que debe concluirse que VioGén es eficaz, y que la aplicación de las medidas que se activan en cada nivel de riesgo sí reduce la reincidencia. Sin embargo, es probable que ello coexista con una importante sobreestimación del riesgo, pues aun aceptando que las medidas adoptadas para proteger a las mujeres reduzcan la reincidencia en esos 20 puntos porcentuales, el volumen de falsos positivos alcanza al menos el 80 % de los casos con algún nivel de riesgo estimado, según los mismos estudios. De hecho,

tanto los documentos oficiales como los estudios publicados sobre VioGén admiten expresamente que la herramienta está diseñada para maximizar la sensibilidad, es decir, detectar como en riesgo el mayor número posible de casos en los que la víctima será realmente agredida de nuevo, y así no dejar desprotegida a ninguna víctima necesitada de protección policial (González Álvarez et al., 2018:50; López-Ossorio et al., 2019: 56; López-Ossorio et al., 2020: 179, 180), asumiendo así que el número de falsos positivos tendrá que seguir siendo alto. Reflexionaremos sobre si esto es un problema desde la perspectiva de los derechos humanos implicados *infra*, en el apartado sobre evaluación de impacto en derechos humanos.

Una cosa más que destacar sobre la precisión de VioGén es que parece diferenciar bien entre los distintos niveles de riesgo, ya que cuanto mayor es el nivel de riesgo, mayor es también la tasa de reincidencia. Se encontró mayor reincidencia, más grave y multirreincidencia en los niveles de riesgo más altos (López-Ossorio et al., 2019: 55).

También se dispone de algunos datos sobre la capacidad predictiva de la escala H: según López-Ossorio et al. (2021), la sensibilidad oscila entre el 81 y 84 %, la especificidad entre el 61 y 60 %, el valor predictivo positivo es del 19 % y el valor predictivo negativo del 97 %. Sin embargo, estos parámetros no corresponden a la aplicación real de la escala a casos nuevos, sino a la muestra de casos antiguos de homicidio con la que se desarrolló y validó la escala. Todavía es pronto para intentar analizar su eficacia en la prevención de nuevos casos de homicidio; no obstante, añadiremos algunas consideraciones sobre ello en la siguiente sección.

No hemos podido comprobar los datos sobre la capacidad predictiva por nosotros mismos, porque las muestras de datos que obtuvimos de la SES no permitían ningún análisis significativo. Como explicamos en el Anexo (apartado *Intentos de obtener muestras de datos e información sobre el algoritmo*), la primera muestra que nos facilitaron no contenía ningún dato sobre reincidencia. En el segundo envío aparece la variable «Incidencia VPER-C». Sin embargo, como la propia SES nos advertía en el correo electrónico, una VPER-C no solo se cumplimenta cuando hay reincidencia (violenta o no), sino también, en otros casos, por lo que la existencia de estas VPER-C es solo una aproximación a una medida de la reincidencia, pero no indica cuál es la tasa de reincidencia real ni sus características. Según lo que se nos dijo en ese correo electrónico, el número de VPER-C más o menos duplica la prevalencia de la reincidencia violenta. Con esa información, era imposible hacer ningún análisis de la exactitud de las estimaciones de riesgo.

Transparencia, evaluación y revisión

El sistema no es transparente. Existe información pública sobre VioGén, principalmente artículos académicos publicados en revistas científicas, guías de procedimiento, y algunos informes elaborados por la SES. Pero nunca se han hecho públicos los detalles de los algoritmos que hay detrás de los formularios VPR y VPER: se sabe qué factores se tienen en cuenta, pero no cómo se combinan ni qué peso relativo tiene cada uno de ellos.

El Ministerio del Interior no tiene una política de datos abiertos. Que sepamos, ninguna de las muestras de los estudios publicados en revistas académicas se ha

hecho pública, por lo que es imposible para los investigadores que no estén trabajando ya con el Ministerio contrastar los resultados, o intentar indagar en cuestiones que aún no se hayan abordado en estos estudios. Bajo la etiqueta «Estadísticas del sistema VioGén», la página web del Ministerio del Interior solo contiene informes mensuales en formato PDF, que incluyen información sobre las valoraciones de riesgo realizadas por provincia y comunidad autónoma, por nivel de riesgo (incluyendo los nuevos «casos de especial relevancia» y los casos con menores en situación de vulnerabilidad o de riesgo), y por cuerpo policial (<<https://www.interior.gob.es/opencms/es/servicios-al-ciudadano/violencia-contra-la-mujer/estadisticas-sistema-viogen/>>).

Existen algunos datos estadísticos sobre VioGén en la página web del Ministerio de Igualdad (Delegación del Gobierno contra la Violencia de Género), pero son pobres (<<https://violenciagenero.igualdad.gob.es/violenciaEnCifras/>>). Esta web ofrece, por un lado, un portal estadístico y, por otro, informes periódicos («Boletines estadísticos») y fichas informativas (ambos en formato PDF). Los datos del portal estadístico sobre VioGén solo hacen referencia a: el número de casos por año y mes, el número de casos en cada nivel de riesgo y el número de casos por provincia y comunidad autónoma. Los informes periódicos incluyen algo más de información: el número de casos por edad de la víctima y el número de casos en función de si la mujer ha nacido en España o en otro país, pero cualquier análisis que se quiera hacer con ellos tiene que hacerse manualmente, buscando los datos en los diferentes informes (formato PDF) publicados por mes o año. Además, los formularios VPR y VPER trabajan con muchos otros factores de riesgo sobre el varón y la mujer de los que no es posible encontrar información en ningún sitio. Tampoco hay información disponible en la página web de ninguno de los dos ministerios (ni en ningún otro lugar) sobre la tasa de reincidencia o el tipo de nuevos delitos cometidos en los casos bajo vigilancia policial.

La información del portal estadístico (web del Ministerio de Igualdad, Delegación del Gobierno contra la Violencia de Género) es un poco más completa para los casos en los que la mujer fue asesinada por su pareja o expareja:³ además de los datos que ya hemos mencionado, hay información sobre el tipo de relación y convivencia con el agresor, edad de la víctima y del agresor, país de origen de ambos, si se había presentado denuncia previa, y si el agresor ha intentado o conseguido suicidarse después del crimen. La página web también incluye fichas informativas en formato PDF sobre los casos de homicidio, que contienen toda esta información y también si se habían puesto en práctica medidas de protección, y si estas medidas se habían incumplido. Pero no hay información, por ejemplo, sobre cuál fue el resultado de las evaluaciones de riesgo realizadas con VioGén en los casos que acabaron en homicidio cuando había habido denuncias previas.

Los responsables del diseño e implementación de VioGén en la SES nos han dicho en varias ocasiones que son transparentes porque han publicado los resultados

³ Un tipo de información similar se puede encontrar en el portal estadístico también sobre los casos en los que los menores han sido asesinados por sus padres o por la pareja o expareja de la madre.

de varios estudios realizados sobre VioGén en revistas científicas, que tienen un estricto proceso de revisión por pares, por lo que la calidad de la información difundida está garantizada. En nuestra opinión, sin embargo, confunden dar información sobre los resultados, que es una cosa, con permitir el acceso de terceros a los datos, para que puedan ser analizados de forma independiente por personas no implicadas, que es otra cosa diferente.⁴

Nuestras peticiones de muestras para poder hacer nuestros propios análisis externos para este informe han tenido poco éxito, ya que se han encontrado con los problemas que hemos explicado detalladamente en el anexo *Dificultades para obtener información y datos sobre VioGén*.

¿Cuál es el riesgo de que el resultado tenga sesgos?

Debido a la escasez de datos disponibles públicamente y a las dificultades para obtener información que explicamos en Anexo *Dificultades para obtener información y datos sobre VioGén*, no es mucho lo que podemos decir sobre esta cuestión. A continuación, explicaremos, en primer lugar, lo que se puede concluir a partir de los datos de que la Secretaría de Estado de Seguridad nos facilitó en abril y mayo de 2022; y, en segundo lugar, nuestras conclusiones tras haber analizado los datos disponibles públicamente.

Análisis de los datos de valoraciones policiales facilitados por la Secretaría de Estado de Seguridad (ses)

Ninguna de las muestras que se nos facilitaron contiene información sobre prácticamente ninguno de los factores de riesgo utilizados por VioGén. Por ejemplo, aunque una de las hojas Excel tenía una columna para la variable «nacionalidad del agresor» y otra para la variable «nacionalidad de la víctima», resulta que en más del 70 % de los casos pone «país no codificado», lo que nos impide analizar en esta muestra si existe alguna asociación significativa, y en qué sentido, entre la nacionalidad de la víctima o del agresor y el nivel de riesgo asignado por VioGén (análisis que, por las razones que explicaremos más adelante, hubiera sido muy interesante).

Hicimos algunos cálculos para intentar detectar asociaciones entre las escasas variables contenidas en estas muestras, pero o bien son poco relevantes porque se trata de información que ya se ofrece de VioGén en abierto (por ejemplo, provincia y comunidad autónoma donde ocurrieron los hechos), o bien darían información sobre el fenómeno de la violencia de género en general, pero no sobre el funcionamiento o características de la herramienta VioGén, que es el objetivo de este

⁴ Lo mismo se subraya en el informe *Éticas sobre VioGén* (2022, p. 5): «la mayoría de los estudios sobre VioGén han sido realizados por los mismos investigadores que contribuyeron a su desarrollo (López-Ossorio et al., 2019; López-Ossorio, González-Álvarez, et al., 2020), e individuos que trabajan o tienen intereses creados en el Ministerio y las fuerzas policiales. Esto refuerza nuestro argumento a favor de la necesidad de una supervisión independiente del sistema».

informe. Por ejemplo, los datos muestran que las víctimas denuncian menos cuando el agresor es el novio o pareja, y más cuando es el cónyuge o expareja. También existe una fuerte asociación en la muestra entre todos los niveles de riesgo, de bajo a extremo, y el hecho de que la víctima haya presentado denuncia, mientras que cuando la víctima no ha presentado denuncia (y es la policía quien hace el informe de oficio, o a partir de denuncias de terceros) hay una mayor probabilidad de que el resultado de VioGén sea «riesgo no apreciado». Esto quizás podría interpretarse en el sentido de que las víctimas denuncian siendo conscientes del riesgo que corren. Sin embargo, también podría tener que ver con el hecho de que cuando la víctima no denuncia, y por tanto no se dispone de la declaración de la mujer, es probable que la información disponible para completar los datos solicitados por VioGén sea mucho menos completa, y por tanto los niveles de riesgo sean menores.

En definitiva, la información proporcionada por el Ministerio en estas dos muestras no permitía realizar ningún análisis relevante sobre las características o el funcionamiento de VioGén, y mucho menos una auditoría externa del sistema.

Análisis de un posible sesgo contra mujeres nacidas fuera de España en las valoraciones policiales del riesgo, basado en datos públicos.

Como hemos explicado anteriormente, el país de origen de la víctima es uno de los pocos ítems que aparecen en los datos disponibles públicamente sobre valoraciones de riesgo realizadas con VioGén. Además, algunos informes (Retamozo y Monteros, 2020; GREVIO, 2020) habían llamado la atención sobre el hecho de que las mujeres nacidas en otros países sufren más violencia por parte de sus parejas o exparejas, y la denuncian con mayor frecuencia, por lo que centramos nuestra atención en esta cuestión, con los siguientes resultados.

Según los datos disponibles, en España las mujeres nacidas en el extranjero sufren más violencia a manos de su pareja que las mujeres nacidas en España. Este fenómeno, que ya se detectó en la Macroencuesta sobre violencia contra las mujeres realizada en 2015 por el Ministerio de Sanidad, Servicios Sociales e Igualdad (pp. 349-355), se mantiene en la Macroencuesta sobre violencia contra las mujeres realizada en 2019 por la Delegación del Gobierno contra la Violencia de Género (Ministerio de Igualdad). Los datos que se presentan a continuación proceden de la Macroencuesta de 2019, por ser los más recientes.

Si tenemos en cuenta la prevalencia a lo largo de la vida, el 12,7 % de las mujeres nacidas en España ha sufrido violencia física o sexual por parte de su pareja (2,6 % de la pareja actual y 19,8 % de parejas anteriores) y el 29,9 % ha sufrido algún tipo de violencia psicológica (13,0 % de la pareja actual y 40,9 % de parejas anteriores). Estos porcentajes son del 24,7 % (5,3 % de la pareja actual y 31,4 % de las parejas anteriores) y 45,6 % (24,3 % de la pareja actual y 52,7 % de las parejas anteriores), respectivamente, entre las mujeres nacidas en el extranjero (p. 268).

Esta desproporción se mantiene si tenemos en cuenta no la prevalencia a lo largo de la vida, sino la violencia sufrida en los últimos cuatro años y en el último año. En cuanto a la prevalencia de violencia en los 12 meses anteriores a las entrevistas, el 1,5 % de las mujeres nacidas en España ha sufrido violencia física o sexual por parte de su pareja (1,3 % de pareja actual y 1,2 % de pareja anterior) y el 9,5 % algún tipo

de violencia psicológica (8,8 % de pareja actual y 6,9 % de pareja anterior). Estos porcentajes son del 3,9 % (3,0 % de pareja actual y 2,7 % de parejas anteriores) y 17,8 % (18,2 % de pareja actual y 9,2 % de parejas anteriores), respectivamente, entre las mujeres nacidas en el extranjero (p. 270).

Además, y según la misma fuente (p. 275), las mujeres nacidas en el extranjero denuncian haber sufrido violencia física, sexual, emocional o haber sentido miedo por parte de su pareja en mayor medida (28,6 %) que las mujeres nacidas en España (20,0 %).

Como se señala en ambas macroencuestas (2015: 353; 2019: 267), en estos resultados hay que tener en cuenta el efecto de la edad, ya que la estructura de edad de la población nacida en el extranjero es más joven que la de la población nativa, es decir, las mujeres nacidas en el extranjero están sobrerrepresentadas entre las mujeres más jóvenes e infrarrepresentadas en los grupos de mayor edad. Esto significa que a veces la razón de un resultado estadísticamente significativo podría estar más relacionada con la edad de la mujer que con el país de origen. En la encuesta de 2015, sin embargo, los modelos logísticos aplicados confirmaron las relaciones observadas por lugar de nacimiento (cf. 2015: 353); en la encuesta de 2019 no se han realizado tales análisis de regresión logística (2019: 267).

Dicho esto, si las mujeres nacidas en el extranjero son víctimas de violencia de género en mayor medida que las españolas, y además denuncian más violencia, sería lógico que las mujeres nacidas en el extranjero estuvieran sobrerrepresentadas en el sistema VioGén, y efectivamente así es. Sin embargo, esta sobrerrepresentación no es homogénea en los distintos niveles de riesgo, sino que es más acentuada en el grupo de casos inactivos y en los niveles de riesgo bajo y no apreciado.

Casos inactivos:

Los casos inactivos en VioGén son aquellos que, por determinadas circunstancias, temporalmente no son objeto de seguimiento policial. Un caso inactivo puede reactivarse en cualquier momento si un cambio en las circunstancias así lo aconseja. Según la Instrucción 4/2019 de la Secretaría de Estado de Seguridad, por la que se establece el actual protocolo para la valoración policial del riesgo con VioGén, los supuestos en los que se inactiva un caso son los siguientes:

- Con carácter general, cuando el caso haya sido valorado como de riesgo no apreciado y, excepcionalmente, cuando el caso haya sido valorado como de riesgo bajo, siempre que para ambos casos se cumplan las siguientes premisas: que no existan medidas judiciales en vigor, que las circunstancias concretas del caso no aconsejen lo contrario a juicio del agente encargado de la valoración y que se haya contactado al menos una vez con la víctima y se haya realizado una nueva valoración del riesgo sin que este se haya elevado.
- También pueden inactivarse los casos cuando concurren circunstancias que hagan imposible o innecesaria la protección de la víctima o la prevención efectiva de una nueva agresión: la no localización de la víctima durante al menos un mes a pesar de los continuos esfuerzos por dar con su paradero, el fallecimiento de la víctima o del agresor, la salida al extranjero o el ingreso en

un centro de custodia/establecimiento penitenciario de cualquiera de ellos, u otra circunstancia similar.

- Por último, también podrá inactivarse un expediente cuando la víctima renuncie voluntaria, manifiesta y reiteradamente a la protección policial, siempre que no existan medidas judiciales en vigor.

Según los informes periódicos disponibles en la web de la Delegación del Gobierno contra la Violencia de Género, a 31 de diciembre de 2021 en los casos inactivos (acumulados desde el inicio del funcionamiento de VioGén) había 377.463 mujeres nacidas en España y 212.999 mujeres nacidas en otros países. Pero en España viven muchas más mujeres nacidas aquí que mujeres nacidas en el extranjero. Si normalizamos según el número de mujeres de cada grupo que residen en nuestro país en 2021 (los datos se pueden encontrar en la web del Instituto Nacional de Estadística, <https://www.ine.es/>), obtenemos la siguiente tabla:

Tabla 2
Ratio de casos inactivos en VioGén a 31 de diciembre de 2021 para mujeres nacidas en España y en el extranjero por 10.000 mujeres

<i>Año</i>	<i>Casos inactivos mujeres nacidas en España por cada 10.000</i>	<i>Casos inactivos mujeres nacidas en otros países por cada 10.000</i>
2021	217,72	669,12

Fuente: elaboración propia a partir de los datos de los informes disponibles en la página web de la Delegación del Gobierno contra la Violencia de Género y en la página web del Instituto Nacional de Estadística: <<https://www.ine.es/>>.

Es decir, los casos inactivos de mujeres nacidas en otros países son más de tres veces superiores a los de mujeres nacidas en España, estandarizados por población. Si tenemos en cuenta que la prevalencia de la violencia de género en todas sus formas es más frecuente entre las mujeres nacidas fuera de España, llama la atención que en este colectivo los casos se inactiven con mucha más frecuencia que los de las mujeres nacidas en España.

A priori, estos datos podrían apuntar a la existencia de un sesgo en VioGén, en el sentido de que tiende a asignar a las mujeres nacidas fuera de España un riesgo menor del que correspondería según la tasa de violencia de género que sufren (casos de inactivación correspondientes al supuesto a)). Pero podría haber otras explicaciones para este fenómeno. Por ejemplo, teniendo en cuenta que España es un país con un alto nivel de turismo, podría ser que algunas denuncias fueran presentadas por mujeres extranjeras que han pasado aquí unos días o semanas de vacaciones, pero que luego han regresado a su país, y en consecuencia el caso ha tenido que ser inactivado ante la imposibilidad de hacer un seguimiento y dar protección desde España (supuesto de inactivación b)). Algo similar podría decirse respecto al colectivo de mujeres inmigrantes, que pueden pasar unos meses o años en España, denunciar, y posteriormente regresar a sus países, lo que haría necesaria la inactivación por las mismas razones. Sin embargo, y aceptando que probablemente estas circunstancias expliquen al menos en parte la sobrerrepresentación de mujeres nacidas fuera de

España en los casos inactivos, ocurre que la distribución de mujeres nacidas dentro y fuera de España en los distintos niveles de riesgo de los casos activos también muestra una cierta desproporción, como ahora explicaremos.

Casos activos:

Los casos activos en VioGén son aquellos que son objeto de seguimiento y protección por la policía. Estos casos pueden estar en uno de los siguientes niveles de riesgo: no apreciado, bajo, medio, alto o extremo. Los datos son los siguientes:

Tabla 3
Casos activos en VioGén para mujeres de 15 años o más, nacidas en España

Año	N.º de víctimas mortales	Riesgo no apreciado	Riesgo bajo	Alto riesgo	High risk	Riesgo extremo	Población	Riesgo por 10.000	Asesinatos por 1 millón
2015	38	25.035	10.350	2.265	80	6	17.354.619	1,35	2,19
2016	27	21.010	14.169	2.969	133	8	17.343.375	1,79	1,56
2017	32	19.152	16.913	3.522	150	13	17.332.935	2,13	1,85
2018	34	17.626	19.697	4.431	180	17	17.328.504	2,67	1,96
2019	32	20.758	17.547	4.565	299	10	17.339.761	2,81	1,85
2020	30	20.796	18.467	4.649	321	7	17.328.451	2,87	1,73
2021	25	20.846	20.660	6.017	427	8	17.333.723	3,72	1,44

Fuente: elaboración propia a partir de los datos disponibles en el portal estadístico del Ministerio de Igualdad y los boletines estadísticos del Ministerio de Igualdad. El número de víctimas mortales comprende aquí tanto los casos con denuncia previa como sin ella. Las columnas relativas al nivel de riesgo corresponden a los casos activos cada año. El «riesgo» de la penúltima columna incluye los niveles medio, alto y extremo. La columna de población se ha extraído de la página web del Instituto Nacional de Estadística <<https://www.ine.es/>>.

Tabla 4
Casos activos para mujeres de 15 años o más, nacidas en otros países

2015	22	10.357	3.226	665	20	1	2.847.655	2,41	7,73
2016	22	8.658	4.835	814	36	3	2.905.781	2,94	7,57
2017	18	8.034	5.868	1.069	66	6	2.987.176	3,82	6,03
2018	19	7.809	7.180	1.470	82	6	3.124.048	4,99	6,08

Tabla 4 (cont.)
Casos activos para mujeres de 15 años o más, nacidas en otros países

2019	23	9.702	6.685	1.684	102	3	3.332.289	5,37	6,90
2020	18	10.255	7.387	1.672	102	0	3.513.068	5,05	5,12
2021	22	10.626	8.395	2.336	151	3	3.573.020	6,97	6,16

Fuente: elaboración propia a partir de los datos disponibles en el portal estadístico del Ministerio de Igualdad y los boletines estadísticos del Ministerio de Igualdad. El número de víctimas mortales comprende aquí tanto los casos con denuncia previa como sin ella. Las columnas relativas al nivel de riesgo corresponden a los casos activos cada año. El «riesgo» de la penúltima columna incluye los niveles medio, alto y extremo. La columna de población se ha extraído de la página web del Instituto Nacional de Estadística: <<https://www.ine.es/>>.

Una vez normalizado el número de mujeres nacidas fuera y dentro de España en cada nivel de riesgo en función de sus respectivas poblaciones, se observa que las mujeres nacidas fuera de España están sobrerrepresentadas respecto a las nacidas en España, como cabría esperar dado que sufren en mayor proporción la violencia de género. Sin embargo, esta sobrerrepresentación es más acusada en los niveles más bajos de riesgo, y no tanto en los más altos. Es decir (y siempre normalizando el número de mujeres de cada grupo en función de su población), nos encontramos con que, de media para los años 2015 a 2021, en el nivel de riesgo no apreciado hay 2,45 mujeres nacidas en el extranjero por cada mujer nacida en España; en el nivel de riesgo bajo la ratio es de 2,01; y en los niveles más altos, tomando conjuntamente los niveles medio, alto y extremo, la ratio baja a 1,86.

Tabla 5
Ratios de mujeres nacidas en España y en el extranjero por niveles de riesgo en VioGén

<i>Periodo</i>	<i>Ratio de riesgo no apreciado otros países/España normalizado por población</i>	<i>Ratio de bajo riesgo otros países/España normalizado por población</i>	<i>Ratio de riesgo (medio, alto y extremo) otros países/España normalizado por población</i>
2015-2021	2,45	2,01	1,86

Fuente: elaboración propia a partir de los datos de las tablas 3 y 4.

En estos casos, que son los activos, no resultan aplicables las posibles explicaciones que hemos barajado para la gran diferencia observada entre los dos grupos de mujeres en los casos inactivos. VioGén parece asignar niveles de riesgo más bajos a las mujeres nacidas en el extranjero que a las nacidas en España, normalizando por población, lo que podría considerarse una discriminación indirecta.

Tuvimos ocasión de comentar este resultado con personal de la Secretaría de Estado de Seguridad, y en su opinión estas estadísticas ofrecen una imagen parcial de la situación. Nos explicaron que, según sus datos, las mujeres extranjeras efectivamente denuncian más agresiones inicialmente, pero que luego sufren menos reincidencia que las españolas. La variable «país de origen» según sus datos no tiene por tanto peso estadístico en el riesgo de volver a sufrir violencia de género, y por eso, de hecho, aunque en el formulario VPR aparece como un ítem (ser extranjera o no), no tiene peso en el resultado del algoritmo.

Para demostrar esta afirmación, nos remitieron una muestra de datos con datos de reincidencia y la variable «víctima extranjera»⁵ sobre 32.466 casos, de los cuales 23.610 son nacidos en España y 8.857 nacidos en otro país (anexo *Dificultades para obtener información y datos sobre VioGén*, para una descripción más detallada). La muestra se compone, según su información, de todas las valoraciones VPR realizadas entre el 13 de marzo de 19 y diciembre de 2019, y tiene un periodo de seguimiento de 1 año. La reincidencia (ya sea con una sola nueva agresión o con más de una en ese periodo) se registró en 4.267 mujeres nacidas en España, y en 812 mujeres nacidas en el extranjero. Estas cifras dan una ratio de 5,25 mujeres nacidas en España revictimizadas por cada mujer nacida en el extranjero, pero en esta muestra hay más mujeres nacidas en España que en el extranjero, por lo que entendemos que debemos normalizar los datos según el número de mujeres nacidas en España y en el extranjero en la hoja Excel (en este caso, no normalizamos sobre las respectivas poblaciones residentes en España, sino sobre los datos de la muestra). Después de hacerlo, obtenemos estos resultados

Tabla 6
Ratio de mujeres revictimizadas

<i>Nacidas en España:</i>	<i>Nacidas en el extranjero:</i>
4.267 / 23.610 = 0,18	812 / 8.857 = 0,09

Fuente: elaboración propia a partir de una muestra de datos enviada por la SES (anexo *Dificultades para obtener información y datos sobre VioGén*).

⁵ Aquí es necesario hacer una aclaración. En los datos disponibles públicamente sobre VioGén (portal estadístico, boletines estadísticos, etc.) se distingue entre mujeres nacidas en España y mujeres nacidas en el extranjero, que no es exactamente lo mismo que mujeres españolas y extranjeras, aunque ambas se solapan en gran medida. En nuestros cálculos siempre hemos utilizado los datos del Instituto Nacional de Estadística sobre mujeres nacidas dentro y fuera de España para normalizar por población (no los datos por nacionalidad). En cuanto a la hoja Excel que nos envió la ses el 3 de octubre, no estamos seguros de si la variable «víctima extranjera» se refiere a una u otra. Tras consultar con el personal de la ses, se nos indicó que se codifica como extranjera según la registren los funcionarios implicados, normalmente en función del origen.

De modo que, tal como nos había dicho el personal de la SES, efectivamente encontramos que las mujeres nacidas en España tienen en esta muestra el doble de tasa de reincidencia que las nacidas en el extranjero.

Esta muestra enviada por la SES es la única información que tenemos sobre tasas de reincidencia según país de origen, por lo que a falta de otros datos con los que contrastarla, supondremos que efectivamente refleja una diferencia en la tasa de revictimización –según los datos de que dispone la policía– entre las mujeres nacidas en España y las nacidas en el extranjero. Si las mujeres nacidas en el extranjero tienen menor riesgo de revictimización, podría ser coherente que VioGén tendiera a clasificarlas como de menor riesgo.

Sin embargo, estos datos nos llevan a preguntarnos sobre las razones por las que esto puede ser así. Llama la atención que, si según todos los estudios disponibles las mujeres extranjeras parecen sufrir más de violencia de género a manos de sus parejas o exparejas, una violencia más grave, y además la denuncian con mayor frecuencia (según los datos de la macroencuesta a la que nos hemos referido anteriormente), ¿cómo es posible que una vez interpuesta la primera denuncia, su probabilidad de volver a ser agredidas sea menor? Hicimos esta pregunta al personal de la SES y nos dijeron que probablemente se debía a muchos factores, sin concretar más. En uno de los documentos que nos habían enviado junto a las primeras muestras de datos en abril, se menciona un posible factor explicativo: que la reincidencia de los hombres extranjeros es menor (11,7 %) que la de los españoles (15,2 %). Esta explicación merece un análisis.

Por un lado, parece presumir que las mujeres nacidas fuera de España solo son agredidas por hombres nacidos en el extranjero, y las mujeres nacidas aquí solo por hombres nacidos aquí. Pero no siempre es así: según los datos que hemos podido encontrar en la web del Instituto Nacional de Estadística (que no se refieren a denuncias policiales, es importante señalarlo, sino a medidas de protección dictadas por los juzgados),⁶ en el año 2021 hay un 10,6 % de mujeres nacidas en España cuyo agresor ha nacido en el extranjero, y otro 10,6 % en el que la situación es a la inversa.

Por otro lado, no existen datos disponibles públicamente sobre el país de origen del agresor: el informe confidencial de la SES al que hemos aludido solo ofrece los porcentajes que acabamos de mencionar,⁷ y ni el portal estadístico de la Delegación del Gobierno contra la Violencia de Género ni los boletines estadísticos que se publican en su página web contienen la variable del país de origen del agresor. Esta variable solo aparece en los casos de mujeres asesinadas por sus parejas o exparejas, por lo que incluiremos su análisis en el siguiente apartado.

⁶ <<https://www.ine.es/jaxiT3/Tabla.htm?t=28232&L=0>>.

⁷ Si no nos equivocamos, estas cifras corresponden al estudio de reincidencia realizado con una muestra de $n = 7.147$ nuevas denuncias y un seguimiento de 320 días, cuyos resultados se publican en López Ossorio et al. (2019). La reincidencia policial fue del 14,3 % en esta muestra (1.020 nuevas denuncias), pero el trabajo solo ofrece información sobre el seguimiento general de VioGén, sin analizar si los resultados difieren entre grupos. No conocemos la distribución de los hombres por país de origen en esta muestra, ni en el conjunto de todos los casos activos de VioGén.

Pero, antes de abordar esta cuestión, es importante comentar la situación de vulnerabilidad de las mujeres nacidas en el extranjero, especialmente cuando se trata de mujeres inmigrantes en situación irregular. Como ya han advertido otros estudios⁸ y nos explicó la persona de la Oficina de Atención a las Víctimas del Delito con la que tuvimos la oportunidad de hablar, estas mujeres se encuentran en una situación de especial vulnerabilidad. Tienen miedo a denunciar porque temen que, si la policía descubre su situación irregular, sean deportadas a sus países o incluso detenidas en un centro de internamiento de extranjeros (CIE). La legislación española prevé la concesión de permisos de residencia temporal para las mujeres víctimas de violencia de género,⁹ pero según nos explicaron desde la Oficina de Atención a las Víctimas del Delito, la mayoría de ellas desconocen esta posibilidad. Es más, los agresores utilizan su situación irregular contra ellas, amenazándolas con que si cuentan lo que está pasando o denuncian serán deportadas. Las policías con las que hablamos para este informe también señalaron que las mujeres extranjeras irregulares son especialmente vulnerables: suelen estar más solas, no tienen apoyo familiar, el agresor se aprovecha de esta situación y la explota, y las amenaza con que se quedarán solas, en la calle, o que serán deportadas. Y también comentaron estas policías que a veces existe un prejuicio en las personas que están en contacto con estas víctimas (tanto policías como jueces, fiscales e incluso personal de oficinas de asistencia u otros organismos) cuando la mujer finalmente se decide a denunciar, en el sentido de que asumen que si la víctima es inmigrante indocumentada, cuando denuncia está mintiendo y lo único que quiere es obtener un permiso de residencia. Hay que tener en cuenta, en cualquier caso, que estos permisos no son automáticos¹⁰ y además solo son temporales: si el procedimiento penal termina con una sentencia no condenatoria o con una resolución de la que no se deduce la situación de violencia de género, a la mujer extranjera se le denegará (o perderá), la autorización de residencia temporal y de trabajo por circunstancias excepcionales y, en su caso, la solicitada a favor de sus hijos. Y se iniciará o continuará el procedimiento administrativo sancionador por estancia irregular en territorio español (incluida la expulsión).

Y en relación con la posibilidad de que las mujeres inmigrantes, una vez interpuesta la primera denuncia, no denuncien nuevas agresiones (y por tanto su tasa de revictimización «parezca» menor que la de las mujeres nacidas en España), algunos estudios han señalado que las mujeres inmigrantes «tienen más riesgo de abandonar los servicios para continuar o reanudar la convivencia con su agresor». Al igual que las mujeres nacionales, «se sienten muy culpables», pero «esta culpabilidad se magnifica en las mujeres inmigrantes. ¿Por qué? Porque si yo denuncio a

⁸ Informe de GREVIO sobre España (2020, n.º 20); Informe de Eticas sobre VioGén (2022: 22 y s.); Martínez-Román et al. (2017); entre otros.

⁹ Artículos 131-134 del Real Decreto 557/2011, de 20 de abril, por el que se aprueba el Reglamento de la Ley Orgánica 4/2000, sobre derechos y libertades de los extranjeros en España y su integración social.

¹⁰ Para conceder el permiso es necesario o bien que el juez haya dictado una orden de protección a favor de la mujer extranjera, o bien que el Ministerio Fiscal emita un informe que constata la existencia de indicios de violencia de género.

este hombre, y tiene antecedentes penales o policiales, no se le renovará el permiso de residencia o de trabajo, entonces es mi responsabilidad que a este hombre se le renueve o no el permiso, que se le permita volver a España, ‘¿qué van a pensar mis hijos?’» (Martínez-Román et al., 2017: 208). Es, a nuestro juicio, especialmente cruel para las mujeres en situación irregular que no solo se encuentren con obstáculos mucho mayores que las españolas o las residentes legales para pedir ayuda (por su dependencia económica del agresor, por el miedo a ser deportadas, por la falta de apoyo familiar y social, porque a veces por su entorno cultural no son conscientes del maltrato al que están siendo sometidas, etc.), sino también que, además, la actual normativa de extranjería en España añade en su caso la carga de saber que pueden ser responsables de que sus parejas o exparejas se enfrenten también a consecuencias más duras que los españoles como consecuencia de la denuncia.

A la vista de todo ello, nos preguntamos si acaso el hecho de que las mujeres extranjeras, una vez interpuesta la primera denuncia, acudan menos que las españolas a denunciar nuevas agresiones o incidentes, no podría deberse a estas circunstancias, de modo que no es que sufran menos revictimizaciones una vez interpuesta la primera denuncia, sino que no vuelven a la policía tanto como las mujeres nacidas en España a denunciar nuevas agresiones o incidentes, y por tanto las estadísticas oficiales están subestimando su tasa real de revictimización.

Y creemos que sería muy necesario analizar este problema, porque los datos de mujeres asesinadas por sus parejas o exparejas, a los que ahora nos referimos, también parecen apuntar a una desprotección de este colectivo.

Mujeres asesinadas:

Las diferencias que existen entre las mujeres nacidas en España y las nacidas en el extranjero se agudizan si nos fijamos en las mujeres que mueren a manos de su pareja o expareja.

Del total de mujeres asesinadas por violencia de género entre 2015 y 2021, 218 eran nacidas en España (60,22 % de las víctimas), y 144 eran nacidas en el extranjero (39,77 %). Sin embargo, hay que tener en cuenta que en España hay menos mujeres nacidas en el extranjero que nacidas aquí, por lo que si normalizamos las cifras según las respectivas poblaciones obtenemos que (de media en el periodo de 7 años estudiado) los asesinatos afectaron a 1,79 mujeres nacidas en España por cada millón de mujeres nacidas en España, pero afectaron a una media de 6,46 mujeres nacidas en el extranjero por cada millón de mujeres nacidas en el extranjero.

Si calculamos la proporción de las poblaciones respectivas para los años considerados, obtenemos las siguientes cifras:

Tabla 7
Ratio de mujeres asesinadas nacidas en otros países/en España,
normalizada por población

<i>Año</i>	<i>Ratio mujeres asesinadas otros países/España normalizada por población</i>
2015	3,53
2016	4,86
2017	3,26
2018	3,10
2019	3,74
2020	2,96
2021	4,27

Fuente: elaboración propia a partir del portal estadístico sobre violencia de género y datos del INE sobre población.

Además, entre las mujeres nacidas fuera de España que mueren a manos de sus parejas, existía una denuncia previa con más frecuencia que entre las mujeres nacidas en España.

El número total de víctimas mortales nacidas en España en el periodo 2015-21 es de 218, de las cuales en 43 casos (19,7 %) había una denuncia previa por violencia de género. Esto da una cifra de 2,48 víctimas mortales nacidas en España con denuncia previa en el periodo 2015-21 por cada millón de mujeres nacidas en España (normalizado a la media poblacional de este periodo).

El número total de víctimas mortales nacidas en otros países en el periodo 2015-21 es de 144, de las cuales en 41 casos (28,5 %) había una denuncia previa por violencia de género, un porcentaje superior, como puede observarse, al de víctimas nacidas en España (19,7 %). A su vez, esto arroja un número de 12,88 víctimas mortales nacidas en otros países con denuncia previa, en el periodo 2015-21, por cada millón de mujeres nacidas en otros países (normalizado a la media poblacional de este periodo). Es decir, más de cinco veces la ratio correspondiente a las víctimas nacidas en España, normalizada a la población en este periodo.

Si comparamos la ratio referida anteriormente, entre mujeres nacidas en el extranjero y nacidas en España clasificadas en niveles de riesgo medio o alto (normalizada según población y teniendo en cuenta la media de los años analizados, de 2015 a 2021), y la comparamos con la ratio de mujeres asesinadas nacidas en el extranjero y nacidas en España con denuncia previa, nos encontramos con lo siguiente:

Tabla 8
Comparación de las ratios de mujeres en riesgo nacidas en otros países/en España, y mujeres asesinadas nacidas en otros países/en España que habían presentado denuncia previamente, normalizadas por población

<i>Periodo</i>	<i>Ratio de riesgo otros países/España normalizado por población</i>	<i>Ratio para otros países asesinados/España, con denuncia, normalizado por población</i>
2015-2021	1,86	5,19

Fuente: elaboración propia a partir del portal estadístico sobre violencia de género y datos del INE sobre población.

En nuestra opinión, esta proporción es preocupante. Incluso teniendo en cuenta que en torno al 70 % de las mujeres que mueren cada año no habían presentado denuncia previa, si nos fijamos en el 30 % que sí lo había hecho, nos encontramos con una sobrerrepresentación de las mujeres nacidas fuera de España, es decir, que suelen pedir ayuda policial con más frecuencia, pero acaban asesinadas cinco veces más (en proporción a su población respectiva) que las nacidas en España. En otras palabras: al menos en este nivel de los asesinatos no se cumple la afirmación de que la reincidencia sea menor para las mujeres nacidas en el extranjero.

Creemos que algo debe de estar fallando en la protección que el sistema penal-policial proporciona a estas mujeres. Cuidado: ese algo no tiene por qué ser necesariamente VioGén, ya que la competencia para dictar las medidas de protección más intensas (órdenes de alejamiento, prisión provisional) no corresponde a la policía sino a los jueces y, como veremos, los jueces dan una importancia muy relativa al nivel de riesgo estimado por VioGén. Sin embargo, hay un punto que puede afectar a VioGén sobre el que nos gustaría llamar la atención.

Como hemos explicado anteriormente, en 2019 se introdujo en VioGén una nueva escala H para estimar específicamente el riesgo de violencia letal. Se trata de un segundo algoritmo, que discurre en paralelo al que estima el riesgo de cualquier revictimización, y que emite una alerta automática si detecta una determinada combinación de algunos factores de riesgo presentes en el VPR que correlacionan con una mayor probabilidad de que las mujeres sean asesinadas. A la luz de lo que venimos explicando, nos resulta llamativo que entre los factores de riesgo que conforman la escala H no figure el país de origen de la víctima, dada la notable desproporción entre el número de mujeres asesinadas cuando su país de origen es el extranjero y cuando es España. Esta desproporción es muy notable (5,19 veces más mujeres asesinadas, entre las que habían presentado denuncia previa, y normalizando la tasa por población), y se mantiene estable desde 2015. En el artículo científico que explica cómo se eligieron los factores de riesgo específicos de la escala H, los autores (algunos de ellos miembros de la Secretaría de Estado de Seguridad) explican que seleccionaron aquellos que demostraron tener una relación estadísticamente significativa en términos de ODDS ratio con el resultado de homicidio, y el factor «país de origen de la víctima» no la tenía (López-Ossorio, González-Álvarez,

Loinaz, Martínez-Martínez y Pineda, 2021, tabla de la p. 50). Nos preguntamos si esto puede deberse a que, en la muestra de homicidios con la que trabajaron, el porcentaje de mujeres asesinadas nacidas en otros países era del 32 %, muy inferior al de las españolas (68 % de la muestra; cf. *Ibíd.*, p. 49). Estos porcentajes representan efectivamente la distribución de las mujeres víctimas mortales de violencia de género entre las nacidas dentro y fuera de España, pero, como ya hemos señalado, el hecho de que entre las víctimas haya aparentemente «menos» mujeres nacidas en el extranjero oculta que, si se normaliza su número en relación con su población, son muchas más que las nacidas en España, en términos relativos.

Por último, en relación con las mujeres asesinadas podemos incluir un análisis de algunas variables relativas al agresor, que no hemos podido realizar respecto a las víctimas de violencia de género en general porque el portal estadístico de la página web de la Delegación del Gobierno contra la Violencia de Género solo incluye estos datos para los casos de homicidio. Las variables estudiadas son: el país de nacimiento, tanto de las víctimas como de los agresores, la existencia de denuncia previa y el suicidio del agresor (con tres posibilidades: se suicidó, no hubo intento o el intento fracasó). Nos centramos en el periodo 2006-2022, para el que se dispone de datos sobre las variables anteriores, por lo que estudiamos un total de 968 casos de víctimas mortales de violencia de género.

Tabla 9
Origen de las víctimas y de los agresores en los homicidios de mujeres cometidos por su parejas o expareja

<i>Agresores \ víctimas</i>	<i>España</i>	<i>Otros países</i>
España	542	94
Otros países	67	265

Fuente: elaboración propia a partir de la información disponible en el portal estadístico de la página web de la Delegación del Gobierno contra la Violencia de Género.

En esta tabla reflejamos los cruces de país de nacimiento de agresores y víctimas, y observamos que en el 14,8 % de los asesinatos cometidos por agresores nacidos en España, las mujeres (víctimas) habían nacido en otros países, mientras que en el 20,2 % de los asesinatos cometidos por agresores nacidos en otros países las mujeres (víctimas) eran de origen español.

Existe un contraste interesante al comparar la tasa de suicidio de los agresores, como puede verse en la siguiente tabla:

Table 10
Suicidio entre varones que matan a su pareja por país de origen

<i>Agresores \ suicidio</i>	<i>No intentado</i>	<i>Consumado</i>	<i>Intento fallido</i>
España	374	168	94
Otros países	248	43	41

Fuente: elaboración propia a partir de la información disponible en el portal estadístico de la página web de la Delegación del Gobierno contra la Violencia de Género.

Mientras que el 58,8 % de los agresores nacidos en España no intentaron suicidarse, el 74,7 % de los agresores nacidos en otros países no lo intentaron. Además, si comparamos los casos en que el suicidio ni siquiera se intenta con aquellos en que se consuma, para los agresores nacidos en España hay 2,2 veces más casos de no tentativa que de consumación del suicidio, mientras que para los agresores nacidos en otros países la proporción es de 5,8. Pensamos que se trata de una diferencia notable de comportamiento, pero no podemos hacer un análisis significativo de esta porque probablemente esté relacionada con rasgos personales y psicológicos y situaciones de cada caso de las que no tenemos información.

En cuanto a la existencia de una denuncia previa, ya hemos explicado que hay una diferencia significativa entre las víctimas nacidas en España o nacidas en otros países. La diferencia es aún mayor si nos centramos en los agresores, como puede verse en esta tabla:

Tabla 11
Autores de homicidios en la pareja por país de origen y existencia de denuncias previas

<i>Agresores</i>	<i>Sí</i>	<i>No</i>
España	142	494
Otros países	107	225

Fuente: elaboración propia a partir de la información disponible en el portal estadístico de la página web de la Delegación del Gobierno contra la Violencia de Género.

Es decir, solo el 22,3 % de los asesinatos cometidos por agresores nacidos en España tenían una denuncia previa, en contraste con un 32,2 % en el caso de los agresores nacidos en otros países. Además, la ODDS ratio de denuncia previa (nacidos en España / otros países) es de 0,6, lo que también aporta en este sentido. Y, por último, al igual que hicimos en el caso de las víctimas, una normalización de algunos indicadores de los agresores respecto a las poblaciones medias arroja más luz sobre el estudio. En el periodo 2006-2022, la población media de hombres

nacidos en España, de 20 años o más,¹¹ fue de 15.459.941 y la población media de hombres nacidos en otros países fue de 2.702.893.

Si normalizamos el número de agresores, obtenemos 4,1 cada 100.000 en el caso de los agresores nacidos en España, y 12,3 (el triple) cada 100.000 para los agresores nacidos en otros países.

Si nos restringimos a los agresores contra los que se había presentado una denuncia antes de que cometieran el homicidio, entonces tenemos 0,92 cada 100.000 casos nacidos en España, y 3,96 (más de cuatro veces) cada 100.000 agresores nacidos en otros países.

Estas dos normalizaciones muestran que los hombres nacidos en otros países están sobrerrepresentados en los casos de homicidio en la pareja, si tenemos en cuenta la normalización con la población.

Creemos que se trata de un resultado que hay que manejar con cautela. Podría llevar fácilmente a la conclusión de que los hombres nacidos fuera de España son más «agresivos» o «peligrosos» con las mujeres, pero en nuestra opinión sería una conclusión demasiado precipitada. Podría influir la mayor vulnerabilidad de la víctima, ya que en la mayoría de los casos en los que el asesino no ha nacido en España, la mujer también es extranjera y hemos visto que este grupo de mujeres parece ser más vulnerable. En otras palabras, si las mujeres nacidas fuera de España son más vulnerables a esta violencia, y la mayoría de ellas tienen también como pareja o expareja a hombres nacidos en el extranjero, los hombres no nacidos en España aparecerán con mayor frecuencia como autores de homicidios en la pareja, pero sin necesidad de que sus características personales y su situación sean la causa principal de esta sobrerrepresentación.¹²

También es importante tener en cuenta que no sabemos si la sobrerrepresentación de hombres nacidos en el extranjero existe también respecto al conjunto de casos de violencia de género, ya que solo hemos podido comprobarlo para los casos de homicidio, pues el portal estadístico solo ofrece esa variable para estos últimos.

Por otro lado, lo que vemos en estas cifras es que los hombres nacidos en el extranjero están sobrerrepresentados en los casos de homicidio en la pareja, no que VioGén esté asignando a estos hombres un mayor riesgo. Al contrario, como explicaremos más adelante (apartado *Prevención de homicidios*), la mayoría de los casos de homicidio en los que se había realizado una evaluación policial del riesgo habían clasificado a la víctima como de bajo riesgo o incluso sin riesgo de revictimización.

¹¹ Cuando hemos analizado en páginas anteriores la distribución de las mujeres extranjeras por niveles de riesgo, hemos tenido en cuenta a las mujeres de 15 años o más, mientras que aquí se trata de hombres de 20 años o más. La razón de la diferencia es que la distribución de la población en el portal estadístico de VioGén es diferente para hombres y mujeres (cuando hay víctimas mortales), y la normalización se ha hecho respecto al rango de edad que mejor se ajusta.

¹² Sin embargo, hay que tener en cuenta que la mayoría de los factores de riesgo de VioGén están relacionados con el comportamiento o las características del agresor (anexo *Los factores de riesgo de VioGén*: todos los incluidos en los grupos I y II, es decir, 23 de un total de 35 factores).

De modo que, si miramos el problema desde este punto de vista, quizá VioGén podría estar subestimando el riesgo de homicidio en la mayoría de los casos, independientemente del país de origen del delincuente.

Y, siguiendo esta misma línea de pensamiento, aunque no sabemos cómo se comportan las valoraciones de riesgo de VioGén en relación con el país de origen del agresor a la hora de valorar el riesgo de violencia de género, porque no se dispone de datos, sí sabemos dos cosas: 1) las mujeres nacidas en el extranjero sufren violencia de género en mayor proporción que las mujeres nacidas en España, pero están infrarrepresentadas en los niveles más altos de riesgo de VioGén; y 2) las mujeres nacidas en el extranjero tienen como pareja más frecuentemente a hombres nacidos en el extranjero. En conclusión, podría ser (esto es una mera hipótesis) que las evaluaciones de riesgo de VioGén también estén subestimando el riesgo para los hombres nacidos en el extranjero.

En cualquier caso, el problema importante aquí es, a nuestro juicio, que las mujeres nacidas en el extranjero son –según los datos existentes y según nuestro propio análisis– víctimas de violencia de género y asesinatos por razón de género en mayor medida que las nacidas en España, pero están infrarrepresentadas en los niveles de mayor riesgo de VioGén. No existen datos públicos disponibles sobre reincidencia ni revictimización por país de origen, y los datos que obtuvimos de la SES sobre esta cuestión, que reflejan una menor tasa de revictimización para las mujeres nacidas en el extranjero, parecen no ser coherentes con la mayor vulnerabilidad de este colectivo y, por tanto, podrían estar infraestimando su verdadera tasa de revictimización. La falta de información hace muy difícil valorar si VioGén podría estar discriminando (o no) a ciertos grupos de víctimas según su país de origen (así como en función de otros criterios, como discapacidades, nivel socioeconómico, edad, etc.).

En cuanto a los hombres acusados de violencia de género, también es necesario disponer de datos sobre sus características. Aunque VioGén se concibe como una valoración del riesgo de la mujer a sufrir una nueva agresión, es al mismo tiempo una valoración del riesgo que tiene el hombre de reincidir, y debería poder estudiarse si VioGén puede tener algún efecto discriminatorio entre grupos de hombres, ya sea por país de origen o en función de otros rasgos o circunstancias.

3. EFICACIA DEL ENFOQUE POLICIAL PREDICTIVO PARA PREVENIR LA DELINCUENCIA

Respuesta de las fuerzas del orden en función del resultado

Para cada uno de los 5 niveles de riesgo que pueden derivarse del VPR, la Instrucción número 4/2019 de la SES establece diferentes medidas de protección para la víctima que debe aplicar la policía, que van desde dar a la mujer información sobre los recursos a su disposición (oficinas de atención a la víctima, servicios sociales) o recomendaciones de autoprotección, hasta facilitarle números de teléfono de emergencia, o teleasistencia móvil, pasando por mantener contacto telefónico o personal

con ella, realizar vigilancia de su domicilio y trabajo, acompañamiento a procedimientos judiciales, etc., y, en casos de riesgo extremo, protección permanente de la víctima y vigilancia de los colegios de sus hijos. El contenido de las medidas que corresponden a cada nivel de riesgo puede consultarse en la Instrucción 4/2019, de la Secretaría de Estado de Seguridad, que lo hemos reproducido en el anexo.

Es importante subrayar que, policialmente, la puntuación actuarial dada por el VPR (y VPER) determina automáticamente el grado de protección que asignar a la víctima, aunque el funcionario actuante puede modificarla al alza si considera que las circunstancias lo justifican. Pero las medidas que deben adoptarse son las recogidas en la Instrucción 4/2019 del SES, y el funcionario actuante no tiene margen de maniobra para decidir aplicar otras distintas.

La individualización de la protección se produce a través del Plan Personalizado de Seguridad, que el funcionario encargado del caso debe elaborar también para cada víctima, teniendo en cuenta sus circunstancias especiales (véanse las directrices para el plan en el anexo *Medidas de protección policial para cada nivel de riesgo...*). Pero este plan coexiste con las medidas automáticas de protección que dependen del nivel de riesgo, y no puede desactivar ninguna de ellas.

Las situaciones son muy diferentes desde los puntos de vistas forense y judicial. Desde el primero, la puntuación actuarial del VPR es solo uno de los elementos con los que el perito forense emitirá un juicio clínico estructurado sobre el nivel de riesgo. El juez, a su vez, puede tener en cuenta el nivel de riesgo estimado por el VPR, así como la evaluación más exhaustiva del riesgo realizada por el médico forense, a la hora de tomar una decisión sobre medidas de protección para la víctima que limiten los derechos del acusado, pero también tendrá en cuenta los argumentos presentados por los abogados de las partes y por las Fiscalía. Estas cuestiones se explicarán con más detalle a continuación.

Evaluaciones forenses del riesgo

La valoración forense del riesgo en casos de violencia de género es la valoración que realiza el personal de los institutos de Medicina Legal y Ciencias Forenses (IMLCF) a petición del juez, sobre el nivel de probabilidad de que la mujer vuelva a ser agredida por su pareja o expareja, para que el juez adopte las medidas cautelares más adecuadas en función de esta información.

Para realizar las valoraciones de riesgo en los casos de violencia de género, el Ministerio de Justicia elaboró en 2011 un Protocolo médico forense para la valoración urgente del riesgo de violencia de género (Ministerio de Justicia, 2011), en el que se recomendaba seguir una metodología que combinara el análisis clínico de cada caso con la ayuda de una herramienta de valoración actuarial. La herramienta actuarial recomendada en dicho protocolo era la escala EPV-R (Echeburúa et al., 2010). Sin embargo, en 2020 se aprobó un nuevo protocolo de valoración forense urgente del riesgo de violencia de género que mantiene el enfoque de juicio profesional estructurado, pero la herramienta actuarial que ahora se recomienda utilizar es el formulario VPR del sistema VioGén. El resultado del VPR se integrará junto a información adicional en un juicio clínico estructurado, que se entregará al juez para que este lo tenga en cuenta al adoptar medidas cautelares contra el agresor.

Estas medidas, según la legislación española, pueden ser una orden de protección más o menos intensa –que incluye tanto medidas penales (prohibición de aproximación a la víctima, prohibición de residencia, prohibición de comunicación con la víctima, retirada de armas) como civiles (atribución del uso y disfrute del domicilio familiar; determinación del régimen de custodia, visitas, comunicación y estancia con los hijos; régimen de alimentos; y cualquier medida para proteger al menor de peligros o daños)– o la prisión provisional. De ahí el interés en analizar las valoraciones forenses de riesgo que se han realizado desde 2020, para comprobar, por un lado, cuántas son, y, por otro, si el resultado de estas valoraciones, que se realizan con el mismo algoritmo que las policiales, difiere o no de estas últimas y, en caso afirmativo, por qué.

Cabe destacar también que, como decíamos antes, la alarma social generada por algunos casos que acabaron en homicidio, pero en los que la valoración policial había estimado un riesgo bajo o inexistente impulsó la creación de la nueva escala H dentro del formulario VPR 5.0, que valora este riesgo específico de violencia muy grave o letal en paralelo a la estimación general del riesgo de reincidencia. Cuando la escala H detecta un alto riesgo de homicidio, emite un aviso automático de «caso de especial relevancia». La versión 5.0 del VPR incorpora además advertencias específicas para «menores en situación de vulnerabilidad» (cuando la mujer víctima manifiesta que cree que se encuentra en una situación objetiva de amenaza o que puede estar en riesgo) y «menores en riesgo» (cuando se da la combinación de un caso de especial relevancia y el de menores en situación de vulnerabilidad, y además factores clínicos de riesgo en el agresor; López-Ossorio, Muñoz Vicente et al., 2020:6). En todos estos casos, el formulario recomienda que se solicite al fiscal y al juez una valoración forense del riesgo.

Existen dos tipos de evaluaciones forenses de riesgo: las urgentes, que deben realizarse en un plazo de 72 horas para que el juez pueda tomar decisiones inmediatas estando de guardia, y en las que solo interviene el médico forense (ya que solo él tiene también la obligación de estar de guardia, no el psicólogo forense ni los trabajadores sociales), y las programadas, en las que participan tanto el médico forense como el psicólogo forense y el trabajador social, cuando se solicitan con un plazo más dilatado y se dispone de más tiempo para buscar y analizar la información relevante.

Por lo que sabemos, no se están realizando evaluaciones programadas. Como hemos explicado en anexo *Dificultades para obtener información y datos sobre VioGén*, los médicos forenses que hemos entrevistado no tienen conocimiento de que se esté realizando ninguna, y de la encuesta que enviamos a los psicólogos forenses no se recibió respuesta. Así que nos centraremos en las evaluaciones urgentes.

Desde la aprobación de la Ley Orgánica 1/2004, se ha dicho con frecuencia que las valoraciones forenses de riesgo de violencia de género son muy escasas. En la Memoria de la Fiscalía General del Estado correspondiente al año 2016 (Fiscalía General del Estado, 2017), se afirma que

Pese a la gran utilidad del PMFVUR [protocolo médico forense para la valoración urgente del riesgo de violencia de género] como herramienta para efectuar la valoración de riesgo objetivo su uso sigue siendo muy escaso, habiéndose emitido exclusivamente 284 informes de conformidad con el protocolo, número muy inferior

al del año anterior (421) y a todas luces insuficiente e insignificante si se tiene en cuenta el número de solicitudes de OP [órdenes de protección] en las que su utilización podría haber sido de gran utilidad (0,74 % del total de aquellas) (p. 464, datos referidos solo a las comunidades autónomas dependientes del Ministerio de Justicia en esta materia).

No parece que en años posteriores su número haya aumentado: una noticia publicada en el diario digital *elDiario.es*, citando datos facilitados por el Ministerio de Justicia, ofrece las siguientes cifras: «en 2019, último año disponible para todo el territorio nacional, los Institutos de Medicina Legal y Ciencias Forenses hicieron 2.152 estudios de este tipo. El número de sentencias dictadas entonces, aunque no todas emanan de procesos iniciados ese mismo año, permite hacer una aproximación: la cifra supone el 4,15 % de los 51.790 casos juzgados». «[...] [D]atos de Justicia más actualizados solo para las comunidades que no tienen las competencias transferidas revelan que [...] en 2020, un año alterado por la COVID, se hicieron 149 valoraciones y hasta junio de 2021, solo 109» (<https://www.eldiario.es/sociedad/juzgados-infrautilizan-valoraciones-forenses-riesgo-casos-violencia-machista_1_8159760.html>).

Gracias a la muestra de valoraciones forenses que nos ha facilitado la Secretaría de Estado de Seguridad, sabemos que desde el 25 de marzo de 2021 hasta el 30 de marzo de 2022 se han realizado 117 valoraciones forenses, utilizando el VFR (es decir, el formulario VioGén VPR para uso forense). Los datos facilitados por la SES no informaban de la provincia o comunidad autónoma a la que corresponden, pero en la medida en que se supone que todos los IMLCF acceden a VioGén (incluidos los de las comunidades autónomas que tienen transferidas las competencias de Justicia), estas 117 corresponden probablemente a las realizadas en toda España –excepto seguramente Cataluña y País Vasco, por la distribución competencial española– en este periodo de 12 meses.

La escasez de valoraciones forenses se atribuye en la Memoria de la Fiscalía General del Estado para 2020 a la escasez de personal y precariedad de medios de las unidades de valoración forense integral en los IMLCF (Fiscalía General del Estado, 2021, pp. 728, 734-736), habiendo algunas provincias que aún carecen de ellas (por ejemplo, León, Madrid, Baleares o Tarragona). Se afirma que, a pesar de los esfuerzos que se han venido realizando en los últimos años y especialmente desde la firma del Pacto de Estado contra la Violencia de Género en diciembre de 2017, existen grandes contrastes en el número y composición de las oficinas de atención a las víctimas del delito y de las unidades de valoración forense integral, y el informe menciona quejas de fiscales de distintas provincias por el retraso en la realización de las valoraciones de riesgo, siendo muy difícil obtenerlas durante el turno de guardia y especialmente antes de la celebración de la vista en la que se decide la orden de protección.

Sin embargo, nos gustaría resaltar que probablemente esta no sea la única razón por la que las valoraciones forenses de riesgo son tan escasas. En nuestras entrevistas con las responsables de las unidades de valoración forense integral de Valencia y Alicante, nos manifestaron que, aunque la implantación de estas unidades es aún muy precaria en otros territorios, y existen grandes diferencias entre unos lugares y otros, en la Comunitat Valenciana hay personal y medios suficientes en todas las provincias, y pueden realizar valoraciones urgentes en el tiempo de guardia. Según

lo que nos comentaron, el principal motivo por el que no se realizan es esta comunidad autónoma sería que los jueces no las solicitan.

Hemos conseguido entrevistar a 9 de los 16 jueces de violencia sobre la mujer que existen en la Comunitat Valenciana, y les hemos preguntado, entre otras cosas, si solicitan valoraciones forenses y las razones por las que lo hacen o no. Los jueces afirman que no suelen solicitar estas valoraciones, y que solo lo hacen en los casos en los que tienen serias dudas sobre el patrón de violencia, en los casos en que el nivel de riesgo en VioGén es extremo o alto, y en los casos de violencia habitual. Teóricamente, las partes también pueden pedir al juez que solicite una valoración forense del riesgo, pero esto tampoco ocurre; solo un juzgado nos dijo que fue solicitada en una ocasión por el abogado de la víctima.

El principal motivo para no solicitarlos es, según los jueces, la falta de medios personales y materiales. Solo hay un médico forense por juzgado, que no está exclusivamente al servicio del Juez de Violencia sobre la Mujer, sino que atiende al resto de juzgados de la demarcación (penales, civiles, etc.), por lo que la solicitud de este tipo de informes, que deberían realizarse lo antes posible al encontrarnos en juzgados de guardia, es muy escasa. Los juzgados que no se encuentran en la capital de provincia no disponen de forenses en sus propios juzgados que puedan realizar este tipo de valoración forense del riesgo, sino que deben remitirlo a la Unidad de Valoración Forense del Riesgo que se encuentra en las capitales de provincia. Es importante destacar que en el caso de Valencia la situación es un poco diferente: los jueces de la ciudad de Valencia sí manifestaron que solicitan evaluaciones forenses de riesgo con cierta regularidad. Más o menos, una cada dos o tres meses.

El hecho de que rara vez se soliciten estas evaluaciones sugiere que los jueces no las consideran de mucha relevancia. No obstante, los pocos casos en los que se solicitan parecen ser casos en que concurren particularidades poco frecuentes en los que se requiere un dictamen pericial. Así que en estos casos difíciles sí se tienen en cuenta para aclarar el caso.

Destacamos esta cuestión del bajo número de valoraciones forenses porque, como hemos explicado en el apartado *Capacidad predictiva, falsos negativos y falsos positivos*, los algoritmos de VioGén están diseñados para minimizar el riesgo de falsos negativos a escala policial (es decir, evitar que mujeres que realmente pueden estar en riesgo no reciban protección policial), y a cambio asumen un alto porcentaje de falsos positivos, es decir, casos en los que se asigna un riesgo alto, aunque la mujer no vuelva a ser agredida. En la Guía para la aplicación del formulario VFR 5.0 en la valoración forense del riesgo, publicada por el Ministerio del Interior (López Ossorio et al., 2020), se afirma que una de las funciones de la valoración forense sería reducir la tasa de falsos positivos (p. 16), aplicando los conocimientos especializados del médico forense y una información más amplia que la que habrá servido de base a la valoración policial. Sin embargo, parece claro que este filtrado no se está produciendo, a la vista del escaso número de evaluaciones forenses de las que se tiene constancia (y del número aún menor de evaluaciones programadas, si es que las hay).

Por otra parte, se plantea la cuestión de qué resultados se obtienen en los –pocos– casos en que sí se llevan a cabo evaluaciones forenses del riesgo. Para ello

contamos con la información cuantitativa derivada de la muestra de evaluaciones forenses facilitada por la SES. En esta muestra de 117 valoraciones, realizadas entre marzo de 2021 y marzo de 2022, lo más frecuente es que la valoración forense mantenga el mismo nivel de riesgo que la valoración policial (en el 45,3 % de los casos). En el 34,19 % de los casos la valoración forense corrigió al alza la valoración policial, y en el 20,5 % restante la valoración forense dio un nivel de riesgo inferior al de la valoración policial. En la gran mayoría de los casos se han impuesto medidas judiciales (solo en 6 casos no hay medidas), incluyendo la mayor parte de los casos en los que el resultado de la evaluación forense es de bajo riesgo o incluso de ningún riesgo. Sin embargo, la información facilitada por la SES no explica el contenido de estas medidas, ni si ya se habían decretado únicamente sobre la base de la evaluación policial, o si se impusieron o modificaron como consecuencia de la evaluación forense.

En cuanto a las razones por las que una valoración forense puede cambiar el resultado obtenido en la valoración policial, las responsables de las unidades integrales de valoración forense de Valencia y Alicante han insistido en que no es porque la valoración policial se haya hecho «mal», sino fundamentalmente porque:

- los médicos forenses son quienes mejor pueden valorar la presencia de trastornos mentales en el acusado, así como en la víctima, factores cuya ausencia o presencia parecen hacer variar enormemente el resultado de la valoración;
- una evaluación forense requiere una entrevista con el acusado, mientras que en la mayoría de las evaluaciones policiales no se dispone de esta fuente de información, ya que los acusados no suelen declarar ante la policía;
- cuando la víctima declara ante la policía, a menudo se encuentra en una situación de estrés, miedo, angustia, vergüenza, quizás todavía en estado de *shock* por la agresión sufrida, si es muy reciente, y su relato puede omitir cosas que, en una nueva situación más tranquila, sí aparecen cuando habla con el médico forense.¹³

Y, en cualquier caso, no hay que olvidar que la valoración forense, a diferencia de la policial, utiliza el formulario VioGén, pero no acaba ahí; es un juicio clínico estructurado en el que el perito tiene en cuenta y valora cualquier otra información que considera relevante. Por otra parte, en los datos sobre valoraciones forenses que obtuvimos de la SES en una hoja Excel, parece que lo que se indica en la columna «Riesgo VFR» no es el resultado final de la valoración forense, sino el resultado que el médico forense ha obtenido rellenando el formulario VioGén VFR. Sin embargo, como este formulario es solo un elemento de juicio que el médico combina con otra información, no podemos estar seguros de que el resultado final de la evaluación forense global coincida siempre con el nivel de riesgo que el médico ha obtenido rellenando el formulario VFR.

¹³ Así lo pone de manifiesto también el informe de Éticas sobre VioGén: tras entrevistar a 31 mujeres que habían sufrido violencia de género y habían pasado por el sistema VioGén, el 80 % de ellas señalaron diferentes problemas a la hora de responder a las preguntas (2022: 24-25)

En definitiva, no parece que hasta el momento las valoraciones forenses estén actuando como filtro para reducir el nivel de falsos positivos que genera VioGén en los niveles altos de riesgo porque, en primer lugar, se realizan muy pocas y, en segundo lugar, de los datos de que disponemos no se puede concluir que la valoración forense tienda a corregir a la baja la valoración policial, sino que parece que en la mayoría de los casos mantiene el mismo nivel de riesgo o lo eleva. Desde el punto de vista de los derechos fundamentales del imputado, esta situación debería ser motivo de preocupación, si los jueces dieran una importancia decisiva al nivel de riesgo indicado por las valoraciones (forenses o policiales), especialmente si las utilizaran como base principal para adoptar las medidas cautelares más graves, especialmente la prisión provisional. Sin embargo, este no parece ser el caso.

Relevancia de la evaluación policial o forense del riesgo judicialmente

El nivel de riesgo de una mujer de volver a ser agredida puede llevar al Juzgado de Instrucción a dictar medidas cautelares para protegerla. Sin embargo, por las razones expuestas en el anexo *Dificultades para obtener información y datos sobre VioGén*, no hemos podido analizar directamente las resoluciones de los jueces de violencia sobre la mujer (JVM) o de los juzgados de instrucción que conceden o deniegan estas medidas. Por ello, hemos analizado las –escasas– sentencias de las audiencias provinciales que se dictan cuando se interpone un recurso contra la resolución del JVM, y hemos entrevistado a varios JVM de la Comunitat Valenciana. Los resultados de ambas investigaciones se presentan por separado.

Resoluciones judiciales dictadas por audiencias provinciales sobre VioGén

Si se introduce la palabra VioGén en la base de datos Aranzadi, solo aparecen 209 sentencias para toda España hasta el momento (excluyendo Cataluña y el País Vasco, donde no se utiliza esta aplicación). Además, entre estas 200, hay muchas que no son relevantes para nuestro análisis, ya que no tienen que ver con el uso de VioGén en las evaluaciones de riesgo, sino que la palabra se menciona en el texto, pero en relación con otras cuestiones. Si reducimos las sentencias a aquellas que sí abordan de alguna manera el nivel de riesgo indicado por VioGén para tomar decisiones sobre medidas de protección para la víctima (y correlativa restricción de derechos del imputado), el número se reduce a poco más de 100 en toda la base de datos (concretamente, 104, según la última consulta realizada en julio de 2022. Se trata de un número llamativamente bajo, teniendo en cuenta que cada año los JVM dictan miles de resoluciones sobre medidas de protección (según datos del CGPJ del año 2021, solo en ese año se solicitaron 37.270, de las que 133 fueron inadmitidas, 26.254 concedidas y 10.883 denegadas; *cfr.* Informe Anual sobre Violencia de Género, Año 2021 disponible en: <<https://www.poderjudicial.es/cgpj/es/Temas/Estadistica-Judicial/Estadistica-por-temas/Datos-penales--civiles-y-laborales/Violencia-domestica-y-Violencia-de-genero/Datos-sobre-Violencia-sobre-la-mujer-en-la-estadistica-del-CGPJ/>>). Por ello, aunque incluimos a continuación un breve análisis de su contenido, queremos subrayar que, a nuestro juicio, y a diferencia de lo que ocurre con las relativas a RisCanvi, en este caso son demasiado escasas

como para considerarlas representativas de lo que está ocurriendo a escala nacional en los juzgados de violencia sobre la mujer.

Lo primero que hay que destacar es que, aunque el número de resoluciones es muy bajo, su número parece haber aumentado en los años más recientes, ya que no hemos incluido ningún filtro temporal a la hora de buscar en la base de datos, por lo que podría haber sentencias desde 2007, año en el que se puso en marcha VioGén, pero la gran mayoría de ellas corresponden a 2019 y años posteriores, siendo muy escasas las sentencias anteriores a 2019.

Por otro lado, la distribución territorial de las 104 resoluciones que consideramos relevantes es muy dispar: la mayoría han sido dictadas por las audiencias provinciales de la Comunidad Autónoma de Andalucía (por ejemplo, 36 solo en la Audiencia Provincial de Cádiz) y hay muchas menos en otras: 23 en la Audiencia Provincial de Madrid (aunque la población de esta provincia es 6 veces mayor que la de Cádiz), 4 en Murcia, ninguna en la Comunitat Valenciana, etc. No sabemos qué puede explicar esta situación: podría estar relacionada con el hecho de que en Andalucía los JVM utilizan en mayor medida VioGén para justificar sus decisiones, y por esta razón la palabra aparece con más frecuencia en las decisiones de las audiencias provinciales que, como tribunales superiores, revisan las decisiones de los JVM. Sin embargo, dado el escaso número de resoluciones encontradas, cualquier conclusión sería aventurada. También podría deberse simplemente al hecho de que los abogados de Cádiz, Málaga o Granada son más combativos que sus colegas de otras provincias y recurren con más frecuencia los autos que imponen medidas cautelares, y por esta razón los casos llegan con más frecuencia ante las audiencias provinciales de esas provincias.

Los criterios de los jueces sobre la utilidad de VioGén son también muy dispares, por lo que no es fácil identificar tendencias predominantes. Si acaso, y con cautela, podría decirse que en la mayoría de los casos el nivel de riesgo indicado por VioGén no es el único criterio para decidir sobre las medidas cautelares, sino que, para decretarlas, los jueces exigen que existan otras evidencias que también avalen su necesidad (y que deben extraerse de las declaraciones de víctima y agresor, de testigos o informes médicos de lesiones, en su caso, o de un informe pericial). Estas evidencias, además, deben acreditar no solo la existencia de riesgo para la mujer, sino también que el hecho denunciado se ha cometido probablemente, ya que el artículo 544 ter de la Ley de Enjuiciamiento Criminal exige como requisito previo para dictar tanto la orden de protección como otras medidas cautelares que

existiendo indicios fundados de la comisión de un delito o falta contra la vida, integridad física o moral, libertad sexual, libertad o seguridad de alguna de las personas mencionadas en el artículo 173.2 del Código Penal, resulte una situación objetiva de riesgo para la víctima que requiera la adopción de alguna de las medidas de protección reguladas en este artículo.

Este es el criterio mayoritario, por ejemplo, en las sentencias de la Audiencia Provincial de Madrid, en las que incluso hemos encontrado una en la que, a pesar de que VioGén indicaba un riesgo extremo para la mujer, no se adoptaron medidas de protección, porque era la única prueba que apunta a la existencia de riesgo y existían declaraciones contradictorias de la mujer y del hombre sobre los hechos controvertidos, sin ningún otro dato o prueba.

Sin embargo, en general, cuando el nivel de riesgo indicado por VioGén es extremo, se observa que los JVM sí adoptan la orden de protección (aunque para ello las resoluciones también tienen en cuenta la presencia de otros indicios), que suele consistir en la prohibición de comunicación con la víctima y la obligación de alejamiento. Solo a veces se incluye la obligación de llevar dispositivos electrónicos para vigilar el cumplimiento de la orden de alejamiento. Y es excepcional que se decrete la prisión provisional, que en las resoluciones que hemos revisado nunca se basa exclusivamente en la valoración del riesgo de VioGén, sino también en la presencia de otros elementos de juicio (como que el agresor haya vuelto a agredir a la víctima tras la interposición de la denuncia, por ejemplo).

Un criterio distinto es el seguido por la Audiencia Provincial de Cádiz, que da más importancia a la valoración del riesgo realizada por VioGén, imponiendo órdenes de alejamiento del agresor, aunque la única prueba que conste en la causa «hasta el momento» sea la valoración policial del riesgo, y aunque el nivel de riesgo sea bajo. Sin embargo, este tribunal sí coincide con los demás en lo relativo a la prisión provisional, respecto a la cual sostiene que «una medida tan gravosa como la prisión provisional no siempre será necesaria cuando el riesgo apreciado sea elevado», si es posible controlarlo a través de medidas menos restrictivas de derechos, como las órdenes de alejamiento controladas por dispositivos telemáticos.

También cabe destacar que existen dos sentencias de la Audiencia Provincial de Málaga que cuestionan el sistema VioGén, afirmando que no sirve para estimar el riesgo y que carece de soporte científico.

Por último, hemos encontrado tres sentencias del Tribunal Superior de Justicia de Andalucía (Granada) que hacen referencia al nivel de riesgo no para adoptar medidas cautelares, sino para justificar un cambio del juzgado competente para conocer del caso. Según la legislación vigente (art. 15 bis de la Ley de Enjuiciamiento Criminal), el Juzgado competente es el del domicilio de la víctima. Existe una sentencia de la Sala de lo Penal del Tribunal Supremo de 31 de enero de 2006 que establece que por domicilio de la víctima debe entenderse el que tenía cuando se produjeron los hechos punibles. Pero en ocasiones es precisamente a raíz del inicio del proceso judicial cuando la mujer abandona el domicilio que hasta entonces compartía con su pareja y se traslada a otro que considera más seguro, por ejemplo, el de sus padres, precisamente por miedo al agresor. En estos casos (en los que VioGén había dado riesgo medio o alto), el Tribunal Superior de Justicia de Andalucía considera que el JVM competente para conocer del caso es el del lugar al que la víctima se ha trasladado con intención de quedarse y por una causa justificada.

En cualquier caso, queremos insistir una vez más en lo poco representativas que nos parecen estas sentencias, dado su escasísimo número en comparación con los miles de autos que dictan cada año los JVM sobre medidas de protección para las mujeres que denuncian ser víctimas de violencia de género.

Resultados de las entrevistas con los jueces y juezas de violencia sobre la mujer de la Comunitat Valenciana

La opinión entre los jueces es bastante similar en cuanto al valor que dan a VioGén a la hora de adoptar medidas cautelares penales: sí se fijan en el nivel de riesgo, pero en ningún

caso lo consideran un elemento determinante o vinculante. Lo utilizan como una herramienta complementaria que les ayuda a decidir sobre las medidas cautelares a adoptar.

Junto con VioGén, los otros elementos que tienen en cuenta los jueces son las declaraciones de la víctima y del acusado, los antecedentes penales, la gravedad de los hechos, el tipo de delito cometido, la reincidencia, el apoyo social de la víctima y del presunto agresor, los informes médicos de lesiones, el testimonio de testigos, el consumo de drogas y alcohol y la presencia de enfermedades mentales. Si se solicita una evaluación forense del riesgo o una evaluación médica forense de la víctima, también se tendrá en cuenta junto a los demás elementos.

En su mayoría, los jueces dicen que no consultan directamente la aplicación VioGén, sino que son los funcionarios judiciales quienes realizan esta tarea. Solo dos jueces nos han dicho que han consultado la aplicación en algún momento, y uno dice que consulta VioGén en todos los casos sin excepción.

La mayor ventaja que ven en VioGén es que les permite obtener una visión general del caso. Los que más utilizan la herramienta afirman que es muy útil porque en una sola herramienta pueden tener toda la información del caso concreto. Los jueces que le dan menos relevancia consideran que es útil para los jueces que están empezando su práctica judicial en este ámbito y tienen todavía poca experiencia.

Por otro lado, no mencionan ningún inconveniente relevante. Solo subrayan que hay que ser cauteloso con la información que proporciona VioGén porque solo ofrece información relacionada con la declaración de la víctima.

VioGén tampoco resulta determinante a la hora de decidir si se impone al acusado un dispositivo de vigilancia telemática. Para ello, los jueces tienen en cuenta no solo el nivel de riesgo, sino también el resto de la información disponible sobre el caso. Afirman que hay circunstancias en las que no es aconsejable colocar el dispositivo si no están relacionadas con el nivel de riesgo, por ejemplo, cuando la víctima no quiere llevarlo o cuando se encuentra en situación de desamparo.

Hay una cuestión interesante que merece la pena mencionar en este contexto. Las medidas cautelares dictadas por los JVM consisten mayoritariamente en las denominadas «órdenes de protección», reguladas en el artículo 544 ter de la Ley de Enjuiciamiento Criminal. Estas órdenes suelen contener medidas de alejamiento y prohibición de comunicación entre el acusado y la víctima. Cuando existan menores que convivan con la víctima y dependan de ella, el juez deberá pronunciarse en todo caso, incluso de oficio, sobre la procedencia de adoptar también medidas civiles, que podrán consistir en la forma en que se ejercerá la patria potestad, la atribución del uso del domicilio familiar, la determinación del régimen de guarda y custodia de los hijos, la suspensión o mantenimiento del régimen de visitas, comunicación y estancia con los menores, el régimen de alimentos, así como cualquier disposición que se considere oportuna para apartarlos del peligro o evitar que sufran daños.

En los últimos años ha crecido en España la preocupación por la situación de los hijos de las mujeres que sufren violencia de género. A veces son testigos de la violencia ejercida contra sus madres, hay maltratadores que directamente también les agreden y, en ocasiones, también les matan. Desde 2013 hasta mediados de 2021, un total de 39 niños han sido asesinados por sus padres o las parejas de sus madres. Para mejorar la protección de estos menores se aprobó la Ley Orgánica 8/2021, de 4 de junio, de protección integral a la infancia y a la adolescencia frente a la

violencia. Esta ley reformó el artículo 94 del Código Civil, para establecer que «no procederá el establecimiento de un régimen de visita o estancia» respecto del progenitor «que esté incurso en un proceso penal iniciado por atentar contra la vida, la integridad física, la libertad, la integridad moral o la libertad e indemnidad sexual del otro cónyuge o sus hijos», ni tampoco si el juez advierte «indicios fundados de violencia doméstica o de género». Esto se aplica tanto a padres como a madres, aunque las estadísticas muestran que hay muchas más agresiones de hombres a mujeres que a la inversa. El precepto especifica, no obstante, que el juez podrá mantener las visitas si dicta una resolución «motivada en el interés superior del menor». Es decir, la norma general será prohibirlas, pero podrá no hacerlo si está justificado. Esta reforma legal responde, según el Ministerio de Igualdad, que la ha promovido, a la idea de que «un maltratador no puede ser un buen padre», y a anteponer el «interés superior del menor» a cualquier otra consideración.

La Ley 8/2021 también reformó el artículo 544 ter de la Ley de Enjuiciamiento Criminal, que, aunque no alcanza el grado de semiautomatismo introducido en el Código Civil, establece ahora que

Cuando se dicte una orden de protección con medidas de contenido penal y existieran indicios fundados de que los hijos e hijas menores de edad hubieran presenciado, sufrido o convivido con la violencia a la que se refiere el apartado 1 de este artículo, la autoridad judicial, de oficio o a instancia de parte, suspenderá el régimen de visitas, estancia, relación o comunicación del inculcado respecto de los menores que dependan de él. No obstante, a instancia de parte, la autoridad judicial podrá no acordar la suspensión mediante resolución motivada en el interés superior del menor y previa evaluación de la situación de la relación paterno filial.

Antes de la reforma, el juez ya podía decidir si mantenía o no las visitas a los hijos del acusado. Pero, en la práctica, la suspensión se producía en muy pocos casos. Las JVM que entrevistamos afirman que la suspensión del régimen de visitas no era la regla general en los casos de violencia de género. Y que daban muy poca relevancia a VioGén a la hora de decidir sobre esta cuestión, porque consideraban que debía tenerse en cuenta la relación entre el presunto agresor y sus hijos, así como el hecho de que los niños hubieran presenciado los episodios violentos y que hubieran sufrido algún tipo de agresión. Solo una de las juezas nos dijo que, en su opinión, el régimen de visitas debería suspenderse siempre en situaciones de violencia de género, porque estos agresores no pueden desarrollar una buena relación paterno filial.

Tras la reforma, los jueces nos dicen que la suspensión se ha convertido en la norma.

El partido político Vox (extrema derecha) presentó un recurso de inconstitucionalidad contra esta reforma (diversos jueces y juristas también habían expresado sus críticas a esta), y el 13 de septiembre de 2022 se conoció la posición del Tribunal Constitucional español al respecto: mantiene la validez de la norma, que considera ajustada a la Constitución, pero entendiendo que no priva automáticamente del régimen de visitas o estancias, ni siquiera en los casos en los que el progenitor esté incurso en un proceso penal. Interpreta que el juez debe valorar siempre la situación en su conjunto y que si decide suspender el régimen de visitas debe hacerlo mediante una resolución motivada, en la que valore los indicios razonables de los hechos delictivos que han dado lugar al proceso penal, así como la necesidad, idoneidad y proporcionalidad de las medidas adoptadas.

Desde 2019, VioGén ha incorporado un sistema de alerta automática para casos especialmente relevantes de peligro de violencia letal, y también cuando considera que hay menores en situación de vulnerabilidad o en riesgo. Estos casos no son inusuales: a 31 de agosto de 2022, VioGén tiene activos 8.019 casos de menores en situación de vulnerabilidad, y 913 de menores en riesgo. No sabemos si VioGén, tras la sentencia del Tribunal Constitucional que impide la aplicación automática de la suspensión del régimen de visitas, tendrá más relevancia en estos casos de la que los jueces venían reconociendo hasta ahora.

¿Es probable que la respuesta de las fuerzas del orden prevenga la delincuencia?

Prevención de cualquier tipo de revictimización

Ya hemos dado una respuesta parcial a esta pregunta anteriormente (apartado *Capacidad predictiva, falsos negativos y falsos positivos*), al analizar la tasa de falsos positivos y falsos negativos en las evaluaciones con el VPR, y hemos concluido que, en nuestra opinión, VioGén sí reduce en cierta medida las revictimizaciones de mujeres. Aquí nos gustaría completar esta afirmación previa, derivada únicamente de los datos del estudio de López-Ossorio et al. (2019), con alguna información cualitativa.

Pudimos entrevistar a dos agentes de policía para este informe sobre su experiencia con la herramienta y su opinión al respecto. Su opinión sobre VioGén es muy buena, la consideran un avance muy importante en la mejora de la protección y seguridad de las mujeres. Cuando se les pregunta en qué consiste esta mejora, no mencionan en primer lugar la reducción del número de nuevas agresiones o de reincidencias graves, sino que destacan sobre todo el hecho de que toda la información esté centralizada en una misma aplicación, lo que facilita enormemente la coordinación entre distintos cuerpos policiales, entre agentes de un mismo cuerpo y distintas localidades, y entre la policía y otros organismos con competencias en la protección de la mujer.

La persona de la Oficina de Atención a la Víctima del Delito que entrevistamos es de la misma opinión. Dice que VioGén tiene la enorme ventaja de tener toda la información de cada víctima centralizada y actualizada, porque tanto la policía como los demás organismos que pueden acceder a VioGén (incluidas las oficinas de atención a la víctima) llevan un registro de novedades, nuevas denuncias, cambios de domicilio, etc. VioGén también permite que cada persona que entra en contacto con la víctima no tenga que volver a preguntarle por muchas cosas de las que a la mujer le puede resultar difícil hablar o doloroso recordar, es decir, evita la revictimización de la mujer.

También comentó que mejora la protección de las mujeres porque muchas veces tienden a minimizar el riesgo y la gravedad de la situación, no siempre cuentan si han vuelto a ser agredidas, pero si otra oficina o comisaría recibe la denuncia, la información entra en el sistema y cualquiera que tenga contacto con la mujer, aunque ella no lo diga, conoce los datos actualizados.

Prevención de homicidios

En España, y desde que tenemos datos sobre este fenómeno (desde 2003), en la mayoría de los casos de mujeres que mueren a manos de su pareja o expareja no había denuncia previa (en torno al 70 %), por lo que no estaban «en el radar» de VioGén, y por tanto la policía no podía darles protección. Es decir, que, aunque VioGén funcionara perfectamente y diera riesgo alto o extremo a todas las mujeres que realmente están en riesgo de ser asesinadas, y las medidas de protección tomadas por la policía o acordadas por los jueces fueran eficaces, esto solo reduciría la tasa de asesinatos en un 30 %. Pero, en cualquier caso, dejando esto a un lado, para aquellas mujeres que sí denunciaron, a veces se ha acusado a VioGén de no protegerlas suficientemente, porque muchas de las que murieron a manos de sus parejas habían recibido una calificación de riesgo bajo en VioGén.

Tabla 12
Resultados de la evaluación de riesgo con VioGén en mujeres víctimas de homicidio en la pareja

Año	Mujeres asesinadas	Sin denuncia previa	Con denuncia previa	Riesgo no apreciado	Riesgo medio	Riesgo extremo	Riesgo alto	Riesgo extremo
2007-2016 ¹⁴	602	508	94	40	24	13	4	0
2019 ¹⁵	56	44	12	4	2	4	0	0
2020 ¹⁶	46	38	8	2	1	2	1	0

Fuente: elaboración propia a partir de los datos contenidos en Santos Hermoso y González Álvarez (2017: 5) y en los informes del Consejo General del Poder Judicial sobre víctimas mortales de violencia de género y doméstica en el ámbito de la violencia de pareja o expareja.

¹⁴ Los datos de estos años se han extraído de Santos Hermoso y González Álvarez (2017: 5).

¹⁵ En este año ya estaba operativa la nueva escala H de VioGén, específica para la estimación del riesgo de homicidio. En uno de los procedimientos, el instructor no estuvo de acuerdo con la valoración realizada automáticamente por el sistema y reclasificó el riesgo de no apreciado a bajo. En ninguno de los casos el algoritmo estimó que se tratara de un caso de especial relevancia (es decir, con riesgo específico de muerte) (Consejo General del Poder Judicial: *Informe sobre víctimas mortales de la violencia de género y doméstica en el ámbito de la pareja o expareja año 2019*, p. 43).

¹⁶ Este año, la nueva escala H también estaba operativa en VioGén. En uno de los procedimientos, el instructor discrepó de la valoración automatizada realizada por el sistema y reclasificó el riesgo de bajo a medio. En otro caso, aunque el resultado fue de riesgo medio, el algoritmo del sistema determinó que se trataba de un caso de especial relevancia (es decir, con riesgo específico de muerte) (Consejo General del Poder Judicial: *Informe sobre víctimas mortales de la violencia de género y doméstica en el ámbito de la pareja o expareja año 2020*, p. 50). Este informe, así como el citado en la nota anterior, pueden descargarse en: <<https://www.poderjudicial.es/cgpj/es/Temas/Violencia-domestica-y-de-genero/Actividad-del-Observatorio/Informes-de-violencia-domestica-y-de-genero/?filtroAnio=>>>.

Creemos que, a partir de estos datos, la valoración de la eficacia de VioGén en la protección de las mujeres frente al asesinato a manos de sus parejas o exparejas debe ser cautelosa. En primer lugar, porque el número de mujeres asesinadas al año es bajo en relación con la población española (en torno a 20 millones de mujeres mayores de 14 años), y si lo reducimos a aquellas sobre las que la policía podía ejercer alguna protección porque existía denuncia previa, es aún menor. Es decir, un fenómeno tan infrecuente es intrínsecamente muy difícil de predecir. Además, según Santos Hermoso y González Álvarez (2019: 5), aunque la mayoría de los casos que acabaron en homicidio habían sido clasificados en los niveles de riesgo más bajos, hay que tener en cuenta que la mayoría de los casos en VioGén pertenecen a los niveles de riesgo más bajos. Entre 2013 y 2015, el promedio de casos activos en el sistema VioGén fue de 60.004. En este mismo periodo, el promedio de casos en riesgo «extremo» fue de 10 (0,02 %), en «alto» de 124 (0,21 %), en «medio» de 2.861 (4,77 %), en «bajo» de 13.200 (22 %) y en «no apreciado» de 43.809 (73,01 %). Teniendo esto en cuenta, aunque la mayoría de los casos en los que la mujer fue asesinada fueron clasificados en riesgo «no apreciado», su porcentaje es sólo del 0,01 % respecto al total de casos en ese nivel de riesgo; los de riesgo bajo supusieron el 0,02 % del total de casos en ese nivel; los de riesgo medio supusieron el 0,05 % del total en ese nivel; y los clasificados en riesgo alto supusieron el 0,32 % de la media de casos en riesgo alto. Estos autores concluyen que

Pese a que los formularios de valoración policial del riesgo no están diseñados para predecir el feminicidio [antes del desarrollo de la escala H], sino para estimar el riesgo de que se produzca un nuevo episodio violento, los resultados encontrados sugieren que hay cierta relación entre el nivel de riesgo, el intervalo temporal y la muerte. En proporción a todos los casos de VCP registrados en el Sistema, han fallecido más mujeres en los niveles de riesgo elevados («alto», 0,32 %; y «medio», 0,05 %) que en los inferiores («bajo», 0,02 %; y «no apreciado», 0,01 %) (2017, p. 7).

En segundo lugar, porque las medidas de protección más contundentes no son las que la policía puede decidir por su cuenta en función del nivel de riesgo dado por VioGén, sino que dependen de que el juez ordene medidas cautelares de protección para la mujer durante la instrucción del procedimiento. Y a la hora de dictar estas órdenes, como hemos explicado en el apartado anterior, los jueces no solo se guían por el nivel de riesgo de VioGén, que en ningún caso es vinculante para ellos, sino que tienen en cuenta otros muchos factores.

En tercer lugar, porque VioGén se construyó para evaluar el riesgo de que una mujer vuelva a sufrir violencia de género de cualquier tipo y gravedad, no específicamente el riesgo de homicidio. Si este último suceso, especialmente grave, estuviera asociado a factores de riesgo específicos que no fueran los mismos que los de otros tipos de violencia de pareja contra la mujer, no se podría culpar a VioGén por infraestimar un tipo de suceso para cuya predicción no fue diseñado. Esto es precisamente lo que sostienen quienes estuvieron al frente del diseño de VioGén (López-Ossorio, González-Álvarez, Loinaz, Martínez-Martínez y Pineda, 2021; véase también Santos Hermoso y González Álvarez, 2017), por lo que en 2019 se desarrolló e implementó la nueva escala H. Su algoritmo solo tiene en cuenta 13 de los 35 indicadores del formulario VPR 5.0, y activa automáticamente una alerta

cuando detecta riesgo de homicidio, sea cual sea el riesgo inicial que VioGén haya asignado al caso.

También por esta razón (que VioGén no estaba diseñado, al menos hasta 2019, para estimar el riesgo específico de homicidio) no parece acertado concluir que es la mayor protección policial lo que está detrás de la correlación entre el descenso del número de víctimas mortales y el aumento del número de casos calificados como de riesgo medio o alto, que se observa tanto entre las mujeres nacidas en España (más claramente) como entre las nacidas en el extranjero. En otras palabras, si las valoraciones de bajo o nulo riesgo en los casos que acabaron en homicidio hasta 2019 no deben considerarse (por todo lo expuesto anteriormente) un «fracaso» de VioGén, tampoco debe atribuirse a VioGén como un «éxito» de la herramienta el hecho de que la tasa de homicidios descienda al tiempo que aumenta el número de casos clasificados en los niveles de mayor riesgo.

Por otro lado, no parece que el descenso de la tasa de mujeres asesinadas nacidas en España observado en 2019, 2020 y 2021 pueda atribuirse a los efectos de la escala H, ya que este descenso se viene produciendo al menos desde 2015, es decir, desde cuatro años antes de la implantación de esta escala. Lo mismo ocurre en el caso de las mujeres asesinadas nacidas en el extranjero, donde el descenso es más atenuado.

En cualquier caso, en nuestra opinión aún es pronto para valorar la eficacia de esta escala H, que apenas lleva tres años en funcionamiento, y en uno de ellos (2020) los datos están muy condicionados por el confinamiento domiciliario y las restricciones a la movilidad que afectaron durante meses a la población española con motivo de la COVID-19.

Y, por último, a la hora de valorar la eficacia de esta escala para proteger a las mujeres frente los homicidios, no podemos fijarnos solo en los casos en los que VioGén valoró un riesgo bajo y la mujer fue asesinada. Hay que tener en cuenta también lo que ya hemos advertido a lo largo de este informe (y resaltaremos también en relación con RisCanvi): en ocasiones puede existir un riesgo realmente alto de homicidio, que VioGén detecta de manera correcta, y precisamente por la eficacia de las medidas de protección puestas en práctica, la agresión finalmente no se produce. Lo que puede parecer un «error» por sobreestimación del riesgo puede ser en algunos casos una prueba de la eficacia de las medidas policiales. Este efecto es más difícil de medir, pero hay que tener en cuenta que lo más probable es que también exista. La siguiente tabla muestra en cuántos casos la escala H emite una advertencia (solo hemos considerado los meses de enero y junio, añadiendo octubre de 2022 por ser la última información disponible antes de cerrar este informe):

Tabla 13
Casos activos de especial relevancia (con alerta de riesgo de homicidio emitida por la escala H) por niveles de riesgo

<i>Fecha</i>	<i>Medio</i>	<i>Alto</i>	<i>Extremo</i>	<i>Total</i>
31.01.2020	4.965	4.380	999	10.344

30.06.2020	3.113	414	17	3.544
31.01.2021	3.400	513	12	3.925
30.06.2021	3.853	672	17	4.542
31.01.2022	4.971	704	17	5.692
30.06.2022	9.226	957	20	10.203
30.10.2022	6.758	651	10	7.419

Fuente: elaboración propia a partir de las Estadísticas del sistema VioGén publicadas por el Ministerio del Interior en su página web (<<https://www.interior.gob.es/opencms/ca/servicios-al-ciudadano/violencia-contra-la-mujer/estadisticas-sistema-VioGén/VioGén/>>).

4. EVALUACIÓN DE IMPACTO SOBRE LOS DERECHOS HUMANOS

Ponderación entre la eficacia y la protección de los derechos fundamentales del acusado y de la víctima

VioGén está sometido a mucha presión pública para evitar sobre todo los falsos negativos. Se ha sugerido que no protege suficientemente a las mujeres (por ejemplo, Grevio, 2020: 66, 67; Eticas, 2022: 30, 32), porque la gran mayoría de los casos en el sistema se califican como de riesgo bajo o inexistente, o porque las mujeres que han sido asesinadas habían sido calificadas en ocasiones como de bajo riesgo. Sin embargo, no estamos seguros de que esta acusación sea la más pertinente.

Hasta ahora, los datos indican una capacidad bastante buena de la herramienta para discriminar entre los grupos de mayor y menor riesgo. Los estudios que se han publicado sobre la capacidad predictiva de VioGén (eso sí, siempre realizados por personas con intereses creados en la herramienta; véase el siguiente apartado para una valoración crítica de esta cuestión) muestran que ha sido deliberadamente diseñada para maximizar la sensibilidad, es decir, para detectar como de riesgo el mayor número posible de casos en los que la víctima es efectivamente agredida de nuevo, y tratando así de no dejar desprotegida a ninguna mujer necesitada de protección policial. Según López-Ossorio et al. (2016), de cada 100 mujeres que volvieron a ser agredidas por su pareja o expareja, VioGén clasificó a 85 como de riesgo, y de cada 100 mujeres que fueron consideradas por VioGén como casos de no riesgo, el 98,5 % no volvieron a ser agredidas. Y esto se está consiguiendo a costa de una alta tasa de falsos positivos, lo que significa que se están desplegando recursos policiales para proteger a mujeres que sin protección policial no habrían sido revictimizadas, pero asumiendo esto como un coste necesario para garantizar la protección al máximo número posible de mujeres que sí se enfrentan a un riesgo. En cuanto a los casos de homicidio, la prevalencia de este fenómeno (y también de los homicidios de niños, que también son motivo de gran preocupación pública) es muy

baja, lo que hace muy difícil elevar los niveles de precisión (aunque la SES ha hecho esfuerzos considerables para mejorarlos, con el diseño de la escala H). Cualquier intento de aumentar la proporción de mujeres clasificadas en alto riesgo, ya sea de cualquier tipo de violencia de género o de homicidio, probablemente elevaría aún más las tasas de falsos positivos. Y este es un aspecto que también hay que tener en cuenta, no solo por la propia sostenibilidad del sistema (diseñado para facilitar y optimizar la asignación de unos recursos que son limitados), sino también teniendo en cuenta que en algunos casos las medidas adoptadas para garantizar la protección de la mujer pueden suponer restricciones de los derechos del varón acusado.

Pensamos que el alto índice de falsos positivos puede asumirse policialmente. En este sentido, el impacto de VioGén sobre los derechos fundamentales puede considerarse muy limitado, o incluso merecedor de un juicio favorable, en la medida en que no restringe los derechos de los potenciales agresores, sino que ofrece asistencia y protección a las víctimas, y los datos e informaciones disponibles parecen mostrar que sí reduce en cierta medida la tasa de revictimización de las mujeres. Es cierto que las medidas policiales de protección incluyen también algunas que afectan al agresor (anexo *Medidas de protección policial para cada nivel de riesgo...*), pero ninguna de ellas limita directa y gravemente sus derechos (por ejemplo, informarle de que el caso está sometido a control policial; requerirle para que entregue voluntariamente las armas si tiene licencia –y solicitar la retirada judicial de estas–; acompañarle al domicilio común para retirar sus pertenencias, si el juez acuerda que debe abandonarlo; revisar sus permisos o salidas penitenciarias; velar por el cumplimiento de las medidas judiciales si se ordenan; instar a la Fiscalía a que le vigile con dispositivos telemáticos; control esporádico o intensivo de sus movimientos y contacto con personas de su entorno).

Dicho esto, es importante destacar que estas altas tasas de falsos positivos empezarían a ser un problema grave si las evaluaciones de riesgo con VioGén se utilizaran de forma rutinaria a nivel judicial, ya que los jueces pueden adoptar medidas que limitan de forma mucho más drástica los derechos de los acusados. Por la información disponible no parece que VioGén esté siendo determinante en la adopción de tales medidas hasta el momento; al contrario, su relevancia desde un punto de vista judicial parece ser, al menos por ahora, bastante baja. Pero si las estimaciones de riesgo de VioGén se convirtieran en la base sobre la que decretar, por ejemplo, prisión provisional, entonces la elevada tasa de falsos positivos y la sobreestimación del riesgo que sin duda presenta VioGén conducirían fácilmente a una limitación generalizada y –a nuestro juicio– injustificada de los derechos de los varones imputados.

La violencia de género es un problema social estructural con profundas raíces culturales y educativas, que se manifiesta no solo en las relaciones de pareja, sino en cualquier contexto (laboral, escolar, en los medios, en la publicidad, etc.). Es obvio que la violencia de género y, sobre todo, los asesinatos por razón de género tienen que ser perseguidos y castigados, y, cuando sea posible, prevenidos con medidas de protección policiales o judiciales. Pero también debería ser obvio que el problema de las mujeres que sufren violencia a manos de sus parejas o de otros hombres no se resolverá ni erradicará solo con más vigilancia policial o con sanciones penales más duras.

En España hace años que se optó políticamente por priorizar la vía penal y judicial en la lucha contra la violencia de género. Lo afirma abiertamente la Memoria de

la Fiscalía General del Estado para 2020 (Fiscalía General del Estado, 2021: 729): «El sistema normativo español ha optado por la punición como uno de los pilares en la lucha contra esta lacra [= VG]». Si la intervención policial-judicial es la principal respuesta, pero el fenómeno persiste, entonces la explicación se busca en supuestos fallos de esta respuesta penal-policial, lo que conduce a una reacción punitiva cada vez más intensificada. Y así VioGén se sitúa en el centro de la polémica, pues se espera que consiga lo imposible (detectar a todas las mujeres en riesgo, especialmente a las de alto riesgo de homicidio), lo que paradójicamente puede llevar a que el sistema quede desacreditado por no lograr lo que es imposible de alcanzar. Y mientras tanto otros aspectos del problema pueden no estar recibiendo la misma atención (posible discriminación por nacionalidad o país de origen, cuáles son las razones por las que muchas mujeres siguen sin denunciar, etc.).

Es cierto que con los años esto ha cambiado un poco; por ejemplo, ya no es necesario, como lo era antes, que la mujer haya presentado una denuncia ante la policía para que pueda acceder a centros de acogida para mujeres maltratadas, o a otro tipo de prestaciones sociales. Pero, en nuestra opinión, la opinión pública y las autoridades siguen contemplando este fenómeno principalmente desde el punto de vista de que se trata de violencia, por lo tanto, de un delito, y por ello debe ser tratado como los delitos: mediante el sistema de justicia penal. Por supuesto, esto es cierto, y ha supuesto una mejora importante en comparación con lo que ocurría en España hace cuarenta años, cuando la violencia de género ni siquiera se percibía como un problema social relevante y el hecho de que los hombres a veces pegaran (o controlaran, maltrataran, insultaran) a sus mujeres se consideraba casi normal. Pero considerar que el sistema penal es la única salida a la violencia de género puede no estar ayudando a todas las mujeres, y puede no ser eficaz para cambiar las profundas raíces socioculturales de este tipo de violencia.

Transparencia y sesgos

VioGén no es transparente. No se puede compartir la opinión expresada en repetidas ocasiones por los responsables de su diseño y supervisión, en el sentido de que hay suficiente información disponible sobre la herramienta en los estudios y artículos de revista publicados sobre esta. El hecho de que se hayan publicado varios artículos sobre VioGén en revistas científicas revisadas por pares merece, desde luego, un juicio favorable, en la medida en que ciertamente aporta información sobre algunos aspectos de la herramienta y garantiza la calidad de las investigaciones realizadas sobre ella. Pero en ningún caso puede considerarse equivalente a una evaluación externa. Una cosa es la calidad de un artículo científico, y otra la transparencia sobre las herramientas utilizadas para la adopción de decisiones por parte de la Administración pública (policía y juzgados en este caso). Que sepamos, nunca ha existido un estudio externo independiente sobre VioGén, realizado por autores ajenos a su diseño o implementación, y es imposible hacerlo porque los escasísimos datos que están públicamente disponibles sobre VioGén presentan todas las carencias a las que nos hemos referido en el apartado *¿Cuál es el riesgo de que el resultado tenga sesgos?*.

Las razones esgrimidas por los desarrolladores de la herramienta para mantener en secreto el diseño del algoritmo y no compartir muestras anonimizadas para que investigadores externos puedan trabajar con ellas (anexo *Anexos sobre VioGén*) no son convincentes. En el contexto español, la Generalitat de Cataluña dispone de una herramienta, a la que nos referiremos en el siguiente apartado, que trabaja con datos tan sensibles como VioGén o más (relacionados con las variables utilizadas para evaluar el riesgo de reincidencia en la población penitenciaria) y ha publicado hojas Excel con miles de datos anonimizados en abierto sin ningún problema.

El hecho de que los datos disponibles en abierto sobre VioGén sean tan escasos y pobres impide contrastar los resultados obtenidos en los estudios existentes, así como analizar aspectos que hasta ahora no han sido estudiados por los responsables de la herramienta (o sobre los que no han publicado artículos o informes científicos) como posibles sesgos. En nuestro informe hemos detectado que puede haber una insuficiente protección de las mujeres nacidas fuera de España, cuestión que ya había sido señalada por otros antes que nosotros, pero es imposible saberlo con seguridad o analizar el alcance de este posible sesgo debido a la falta de datos. Tampoco es posible analizar si las estimaciones de riesgo de VioGén afectan de forma dispar a otros colectivos vulnerables (por ejemplo, por edad, nivel socioeconómico, posibles discapacidades, etc.) ni proponer mejoras que puedan corregirlo.

Obligaciones activas de rendición de cuentas

Con respecto a esta cuestión, en nuestra opinión, las autoridades públicas responsables de VioGén o con derecho a utilizarla también deberían adoptar algunas medidas, a fin de mejorar la rendición de cuentas, la transparencia y el respeto de las directrices del RGPD sobre el uso de datos personales respecto del empleo del sistema. En todo caso, la información sobre el tratamiento de datos efectuado por VioGén sí se encuentra publicada en el registro del Ministerio del Interior. Esta información puede consultarse en el siguiente enlace (p. 129): <<https://transparencia.gob.es/transparencia/dam/jcr:b3eb17bc-b1a7-4092-9c99-f11f62d868ae/20220922%20Registro%20de%20Actividades%20de%20Tratamiento%20del%20Ministerio%20del%20Interior.pdf>>.

No obstante, consideramos que, respecto a estos tratamientos, deberían introducirse algunas mejoras:

1. En primer lugar, debería llevarse a cabo una evaluación de impacto de la protección de datos completa y exhaustiva, que por el momento parece inexistente o, al menos, no se ha publicado. Así lo prevé el Reglamento para aquellos tratamientos de datos que puedan suponer un alto riesgo para los derechos y libertades de las personas físicas. Además, ha de ser tenido en cuenta que se presume en todo caso que existe un alto riesgo cuando estos tratamientos se realizan con nuevas tecnologías (art. 35 RGPD). Como se ha indicado, no sabemos si se ha llevado a cabo efectivamente la evaluación de impacto relativa a la protección de datos, dado que no aparece publicada. A estos efectos, el Ministerio del Interior es la autoridad responsable de llevar a cabo la evaluación

de impacto sobre la protección de datos y quien debería proporcionar la información requerida en relación con este punto.

2. Deben establecerse mayores garantías para las operaciones de tratamiento que impliquen datos de categoría especial, siguiendo las directrices e instrucciones establecidas por el RGPD europeo y de acuerdo con los requisitos detallados para esta categoría de datos declinados por las autoridades de protección de datos (en España, la Agencia Española de Protección de Datos – AEPD– ha realizado un exhaustivo y completo trabajo en la materia).
3. Por último, consideramos que debería elaborarse un registro completo de los tratamientos de datos que sean realizados empleando la herramienta. Este registro ha de ser público, de acuerdo con el artículo 32.4 de la Ley Orgánica 7/2021, de 26 de mayo, de protección de datos de carácter personal tratados con fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales. El Ministerio del Interior es también la autoridad responsable de la publicación de esta información.

Acceso a los datos y posibilidad de impugnar los resultados de VioGén

El artículo 53.1 a) de la Ley 39/2015, de Procedimiento Administrativo Común (LPAC), establece un derecho absoluto a favor de las personas físicas de acceso a toda la información contenida en los expedientes administrativos relativos a procedimientos que puedan afectarles. Asimismo, el artículo 70 de esta misma norma enumera los materiales que deben formar parte de los expedientes administrativos, con algunas posibles exclusiones enumeradas en el artículo 70.4 LPAC, pero esta norma indica explícitamente que en ningún caso podría excluirse del expediente administrativo una evaluación o informe directamente relacionado con la decisión administrativa finalmente adoptada. Por lo tanto, si bien no se recoge ninguna excepción para VioGén ni en la legislación procesal penal ni en la normativa policial, el acceso a los datos de las valoraciones de VioGén ha de ser siempre completamente posible de acuerdo con la actual legislación española, cuando sea realizado por las mujeres que facilitaron los datos con el fin de obtener protección frente a posibles resultados de violencia.

- Además, los titulares de los datos personales tenidos en cuenta o introducidos en los distintos sistemas algorítmicos tienen una serie de derechos según la Ley Orgánica 7/2021, de protección de datos de carácter personal tratados con fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales.
- Derecho a obtener determinada información sobre el tratamiento de datos efectuado por estos sistemas algorítmicos (art. 21, LO 7/2021).
- El derecho de acceso a los datos personales utilizados en el tratamiento de sus datos personales y a los resultados generados por sistemas algorítmicos respecto de los tratamientos que les afecten. Este último no está expresamente reconocido por la norma, pero creemos que también es posible exigirlo (ex artículo 22, LO 7/2021).

- Derecho de rectificación y supresión de datos personales (art. 23 LO 7/2021). En este sentido, el Ministerio del Interior ha habilitado un canal específico para solicitar la supresión de los datos personales que obren en VioGén (<<https://www.interior.gob.es/opencms/es/servicios-al-ciudadano/violencia-contra-la-mujer/supresion-de-datos-sistema-viogen/>>). Este derecho puede ser ejercido tanto por la víctima como por el presunto infractor. Se aceptará la supresión de los datos personales presentes en el sistema VioGén a petición de los interesados en los casos en que exista una resolución judicial firme de sobreseimiento definitivo o sentencia absolutoria firme. No obstante, dichas solicitudes de supresión de datos podrán ser denegadas cuando sigan siendo necesarios para la consecución de los fines para los que fueron recabados (protección de las víctimas y prevención de infracciones penales relacionadas con la violencia de género), o cuando se haya producido alguna reiteración, reincidencia o quebrantamiento de medidas judiciales o penales.
- Se prevé la prohibición del tratamiento de datos personales que dé lugar a la adopción de decisiones totalmente automatizadas o algorítmicas, a menos que una ley prevea dicho tratamiento y establezca al menos el derecho de los interesados en obtener la intervención humana en el proceso de revisión de la decisión adoptada (art. 14, LO 7/2021).
- En todo caso, y respecto del caso de VioGén, hay que recordar que en determinadas acciones el resultado del algoritmo está totalmente automatizado, por lo que se le aplicarían también a estos supuestos las normas referidas a tratamientos automatizados del artículo 41 LPAC.

En cuanto a las víctimas, son informadas por la policía del nivel de riesgo que VioGén predice para ellas y además se les facilita una hoja informativa genérica en la que se explican las distintas variables que se tienen en cuenta, así como los posibles niveles de riesgo que puede predecir el sistema y las medidas de protección asociadas a cada uno de ellos. Es importante destacar que esta información no es personalizada, sino que consiste simplemente en una explicación general del funcionamiento del sistema.

En todo aquello referido a los derechos que asisten a los potenciales agresores cuando son procesados por la policía y por los tribunales, hay que tener en cuenta que uno de los problemas de VioGén es que se concibe como una herramienta utilizada para proteger a las víctimas y, en consecuencia, no se han tenido excesivamente en cuenta los efectos que tiene sobre los potenciales agresores, dificultando así los derechos procesales que puedan tener para impugnar sus resultados. En cualquier caso, consideramos relevante identificar las posibilidades que pueden tener los potenciales agresores de impugnar la valoración: 1) cuando están siendo procesados por la policía; y 2) cuando están siendo procesados por los tribunales.

En cuanto a las posibilidades de obtener información al ser procesado por la policía, no existe un mecanismo formal que garantice a los presuntos infractores la posibilidad de acceder a la lógica subyacente tras la predicción de VioGén. Por lo tanto, los infractores generalmente no obtienen información sobre sus puntuaciones para cada una de las variables ni sobre el peso atribuido a cada una de ellas. No obstante, en las entrevistas mantenidas con los agentes de policía hemos podido

comprobar que, en casos muy excepcionales, los presuntos infractores reciben esta información de manera informal, si la policía considera que comunicársela contribuirá a mejorar su conducta.

En cualquier caso, pensamos que de acuerdo con la regulación contenida en el artículo 22 de la Ley Orgánica 7/2021, la persona denunciada como posible agresor debería tener al menos acceso a la información y a la codificación de los factores contenidos en VioGén que se relacionan con él (sus delitos pasados, sus conductas hacia la víctima, sus celos, sus problemas psiquiátricos, intentos de suicidio, etc., es decir, los factores 1 y 2 del VPR; anexo *Los factores de riesgo de VioGén*). En la medida en que VioGén solo haya dado lugar a la adopción de medidas de protección para la víctima, pero ninguna medida restrictiva de derechos para el acusado, entendemos que este no podría reclamar un derecho a conocer la información contenida en VioGén relativa a la mujer (especialmente el factor 3 del VPR; anexo *Los factores de riesgo de VioGén*), porque en estos casos no está claro que el hombre sea «el interesado» al que se refiere el artículo 22 de la Ley Orgánica 7/2021, ya que las medidas no producen efectos sobre él sino sobre ella. Téngase en cuenta, además, que el algoritmo contiene datos personales de la mujer relacionados con su esfera íntima (problemas mentales, situación familiar, adicciones, etc.). Por el contrario, si como consecuencia de esta evaluación fueran adoptadas medidas restrictivas sobre la persona denunciada como posible agresor, bien por la policía o por algún órgano administrativo, entonces pensamos que debería tener acceso a toda la información contenida en VioGén, porque en tal caso sí sería indudablemente «interesado», y el artículo 53.1.a) de la Ley 39/2015 (LPAC) le daría en todo caso acceso a todo el expediente, así como también prevé el artículo 22 de la Ley Orgánica 7/2021.

Por el momento, la normativa vigente en España no prevé la adopción de medidas administrativas negativas para el imputado que dependan de su nivel de riesgo en VioGén, y las medidas que la policía puede adoptar sobre él se limitan a controlar sus movimientos o entrevistarse con él para informarle de que está sometido a control o para exigirle la entrega voluntaria de sus licencias de armas, etc. (anexo *Medidas de protección policial para cada nivel de riesgo...*). Cosa distinta podría ocurrir si se utilizara VioGén, por ejemplo, para denegar el acceso a los permisos durante el cumplimiento de la condena: sabemos que la administración penitenciaria tiene acceso a VioGén para realizar consultas y consideramos que, en este caso, si se tiene en cuenta el nivel de riesgo para tomar la decisión, el interno debería poder acceder a la información contenida en el algoritmo (en este caso, además, también sería de aplicación la normativa penitenciaria que regula el derecho de los presos a la información contenida en su expediente penitenciario; sobre esta cuestión, ver más abajo lo que explicamos en relación con RisCanvi).

Por otro lado, la evaluación de riesgos generada por VioGén podría ser cuestionada o revisada en los tribunales. Sin embargo, no hemos encontrado sentencias que analicen esta cuestión. Solo hemos encontrado un caso en el que un presunto agresor solicitó la emisión de un informe por parte de la policía sobre la valoración de alto riesgo que VioGén había indicado sobre él. En este caso, sin embargo, el juez denegó dicho informe al considerar que no presentaba especial relevancia probatoria y la Audiencia Provincial confirmó la decisión en apelación (Auto n.º 654/2021,

Audiencia Provincial de Málaga, ECLI:ES:APMA:2021:1560A), con el siguiente razonamiento:

En efecto, en modo alguno puede considerarse útil para el esclarecimiento de los hechos y, menos aún, decisiva en términos de defensa, la práctica de las diligencias de investigación solicitadas, pues es sabido que las valoraciones de riesgo que se realizan en los atestados policiales se basan en la aplicación del sistema VioGén, sistema informático del que se obtiene una calificación de riesgo siempre aproximada y con las naturales reservas, tras procesar mediante algoritmos un cuestionario que es contestado por la persona a proteger (la presunta víctima), por lo que ni en fase de instrucción, y menos aún en el plenario, puede otorgarse a estos informes especial relevancia probatoria.

Creemos que tiene razón la Audiencia Provincial de Málaga cuando afirma que la valoración del riesgo de VioGén no tiene relevancia alguna como medio de prueba de si el acusado cometió o no los hechos que se le imputan. En cambio, la argumentación es problemática si el acusado solicitó el informe para impugnar las medidas cautelares que se le impusieron durante la pendencia de la causa, a partir de la afirmación de que VioGén le había asignado un riesgo alto. Como no hemos tenido acceso a la resolución del juez no sabemos si fue así, pero, si lo fue, denegar el acceso a esta solicitud sería, a nuestro juicio, contrario al derecho a un proceso con todas las garantías reconocido en el artículo 24.2 de la Constitución Española, el artículo 6.1 del CEDH y el artículo 118 de la Ley Española de Enjuiciamiento Criminal, que otorga a los imputados el derecho a ser informados de los hechos que se le atribuyen, así como de cualquier cambio relevante en el objeto de la investigación y de los hechos imputados, y también el derecho a la revisión de las actuaciones. También los artículos 234 y 454 de la Ley Orgánica 6/1985 reguladora del Poder Judicial, que obligan a las oficinas judiciales a facilitar a los interesados cuanta información soliciten sobre el estado de las actuaciones judiciales, que podrán examinar y conocer, salvo que sean o hayan sido declaradas secretas conforme a la ley.

5. NUEVOS DESARROLLOS SOBRE VIOGÉN ENTRE ENERO Y ABRIL DE 2023

Durante el mes de diciembre de 2022 y los primeros días de enero de 2023, varias mujeres fueron asesinadas por sus parejas o exparejas.

Aunque finalmente el número total de asesinatos en 2022 se mantuvo muy similar al de años anteriores (50; <<https://links.uv.es/o8hEOja>>), la concentración de muchos casos en pocas semanas creó una importante alarma y los ministerios de Interior e Igualdad se reunieron para, entre otras cosas, estudiar si era necesario revisar VioGén para mejorar la protección de las mujeres.

El resultado de todo ello fue una nueva Instrucción de la Secretaría de Estado de Seguridad, la Instrucción 1/2023, de 9 de febrero, que ha incluido, entre otras, estas dos novedades:

1. Los agentes encargados del plan de seguridad personalizado comunicarán a la mujer si su agresor ha sido denunciado por otra u otras víctimas anteriores de violencia de género, cuando consideren que esta circunstancia constituye un

especial factor de riesgo para ella y las personas a su cargo. La comunicación no será generalizada: los agentes valorarán en cada caso concreto el grado de riesgo que los antecedentes del agresor suponen para la actual pareja y la necesidad de que ella y su entorno lo conozcan para mejorar su percepción del peligro que corren. La comunicación requerirá que el caso sea considerado de riesgo alto o extremo, o de riesgo medio con especial relevancia (especial relevancia significa que concurren circunstancias que aumenten significativamente la probabilidad de que el agresor ejerza violencia muy grave o letal contra la mujer), y que concurren además algunas otras circunstancias adicionales enumeradas en la instrucción.

Por otro lado, la comunicación a la víctima se limitará a hacer constar que el agresor tiene antecedentes por violencia en el sistema VioGén, sin aportar más detalles sobre los hechos anteriores que se le atribuyen. Solo se tendrán en cuenta las conductas violentas previas o que hayan supuesto un riesgo cierto para la seguridad de la mujer agredida o de su entorno denunciadas por víctimas anteriores durante los cinco años previos. Se excluirán los hechos prescritos o respecto de los cuales el sistema VioGén registre la existencia de sentencia absolutoria o sobreseimiento libre.

2. La Instrucción 1/2023 reitera también las previsiones que una instrucción anterior, la 8/2021, ya había establecido respecto a los delincuentes reincidentes: cuando el sistema VioGén detecte la existencia de un delincuente reincidente, e independientemente del nivel de riesgo asignado al caso concreto, generará una alerta que será notificada automáticamente a las autoridades judiciales y fiscales competentes, para que dispongan de esta información a la hora de adoptar las decisiones que estimen oportunas. Esta consideración de «caso persistente» se reflejará también, de forma automatizada, en toda la documentación que manejen los funcionarios policiales encargados de la protección de las víctimas actuales y antiguas.

La verdadera novedad, por tanto, de la Instrucción 1/2023 es la posibilidad, en determinados casos, de que la policía advierta a la mujer de que su agresor tiene antecedentes de violencia doméstica ejercida sobre otras mujeres. La cuestión más problemática en relación con esta medida es la protección de la intimidad y los datos personales del agresor (en España, los antecedentes penales no son públicos, y VioGén incluye no solo los antecedentes penales en el sentido de debidos a una sentencia firme, sino también información sobre denuncias presentadas, pero aún pendientes, retiradas, etc.). La Instrucción 1/2023 subraya que no se vulnera la normativa de protección de datos aplicable (constituida por la Ley Orgánica 7/2021), porque se respeta el principio de minimización de daños, haciendo un uso no excesivo de los datos para una finalidad muy concreta. Tanto la Fiscalía General del Estado como la Abogacía del Estado emitieron informes a petición del Ministerio del Interior sobre esta cuestión, y no se opusieron.

Conseguimos acceder a estos informes a través de una solicitud de derecho de acceso a información pública. Lo que se había preguntado a ambos organismos era si es conforme a la normativa de protección de datos revelar a una mujer que

presenta una denuncia por violencia de género que su agresor tiene antecedentes en VioGén respecto a otras mujeres. La respuesta de ambos organismos es afirmativa, siempre y cuando no se haga de forma generalizada, sino valorando en cada caso concreto si es una medida necesaria para lograr una mejor protección de la mujer. En principio, y teniendo en cuenta el marco legal aplicable al respecto en España, nos parece razonable la argumentación y la respuesta.

Por otro lado, la acumulación de varios casos de homicidio en unas semanas en torno a la Navidad merece un breve comentario sobre si existen determinadas épocas del año o contextos específicos en los que tienden a acumularse los homicidios de mujeres a manos de sus parejas o exparejas, y cuáles serían las medidas más adecuadas que adoptar en estos casos.

La Delegación del Gobierno contra la Violencia de Género publicó en 2021 un estudio sobre «Análisis temporal de los asesinatos de mujeres por violencia de género en España a lo largo de 15 años (2003-2017)» (Sanz Barbero et al., 2021), que incluía, entre otras cuestiones, un análisis de la ocurrencia de este suceso en meses o días concretos de la semana. Los resultados fueron los siguientes:

En la distribución mensual de los asesinatos, se observa que los meses de verano (julio y agosto) son los que más asesinatos acumulan a lo largo de todo el periodo (86 y 81 víctimas, respectivamente), además de enero (83 víctimas). En el extremo opuesto, abril es el mes con menos asesinatos (64 víctimas).

En la distribución de los asesinatos por día de la semana, destacan el domingo y el lunes como días con mayor número de asesinatos (146 y 145, respectivamente). Pasados estos días, la frecuencia decrece hasta alcanzar el mínimo el jueves (110 víctimas), para posteriormente incrementarse de nuevo.

Otro estudio, sin embargo, utilizando como conjunto de datos todos los feminicidios ocurridos en España en 2007-2017, había concluido que «no se han identificado diferencias significativas entre trimestres, meses o días de la semana» (Torrecilla, Quijano-Sánchez, Liberatore, López-Ossorio y González-Álvarez, 2019: 12).

Ninguno de los dos estudios señala ningún efecto de las fiestas navideñas, o de las reuniones familiares y el estrés que conllevan, sobre los asesinatos por razón de sexo. Ambos, en cualquier caso, destacan que a lo largo de la serie temporal hay periodos concretos de pocos días en los que el número de asesinatos es inusualmente alto. El primer estudio concluye que

Si bien se han observado episodios llamativos, con un número de asesinatos inusual para lo breve de su duración, se constata que son simplemente un reflejo del efecto derivado de la mayor frecuencia de asesinatos en determinados años, meses y, muy especialmente, en determinados días de la semana. No se ha encontrado evidencia sobre la existencia de ninguna agrupación de asesinatos.

El trabajo de Torrecilla et al. (2019) subraya que no se encuentra efecto imitación (*copycat effect*) y concluye que

Los resultados del análisis muestran que pueden producirse clústeres de sucesos, aunque no existe correlación temporal entre ellos. Esta investigación informa a los responsables políticos y activistas de que las actividades de prevención contra los feminicidios, aunque necesarias, deben planificarse cuidadosamente y no ser el resultado de una reacción a un aumento aparente de la ratio de homicidios.

Por otro lado, hay un interesante artículo de Ansgar Seyfferth que se publicó en prensa en 2019, cuando también se había creado alarma social porque se habían acumulado muchos homicidios en la pareja, en esta ocasión en julio. En él se explica de forma sencilla que este tipo de acumulaciones pueden ser simplemente fruto de la mera casualidad; el texto se puede encontrar en: <https://www.huffingtonpost.es/entry/la-violencia-de-genero-y-el-azar_es_5d7bc154e4b03b5fc882a875.html>.

6. FUENTES

- ASOCIACIÓN ETICAS INVESTIGACIÓN E INNOVACIÓN (2022). En línea: <<https://eticasfoundation.org/gender/the-external-audit-of-the-VioGén-system/>>
- GREVIO(Group of Experts on Action against Violence against Women and Domestic Violence) (2020): *Baseline Evaluation Report Spain*, Council of Europe, en línea: <<https://rm.coe.int/grevio-s-report-on-spain/1680a08a9f>>.
- ECHEBURÚA, Enrique; Pedro Javier AMOR; Ismael LOINAZ y Paz DE CORRAL (2010). «Escala de Predicción del Riesgo de Violencia Grave contra la pareja – Revisada– (EPV-R)». *Psicothema* 4(22), pp. 1054-1060.
- GONZÁLEZ ÁLVAREZ, José Luis; Juan José LÓPEZ OSSORIO y Marina MUÑOZ RIVAS (2018): «La valoración policial del riesgo de violencia contra la mujer pareja en España. Sistema de seguimiento integral en los casos de violencia de género», Ministerio del Interior – Secretaría General Técnica, en línea: <<https://cpage.mpr.gob.es/producto/la-valoracion-policial-del-riesgo-de-violencia-contra-la-mujer-pareja-en-espana-2/>> (consulta: 27/12/2021).
- GONZÁLEZ-PRieto, Ángel; BRÚ Antonio, Juan Carlos NUÑO y José Luis GONZÁLEZ-ÁLVAREZ (2021): «Machine learning for risk assessment in gender-based crime», en línea: <<https://arxiv.org/abs/2106.11847>>.
- Guía de Procedimiento VPR5.0 y *uper*4.1. Protocolo de valoración policial del riesgo y gestión de la seguridad de las víctimas de violencia de género (2019): Área de Violencia de Género, Estudios y Formación, Ministerio del Interior, en línea: <<https://violenciadegenerotic.wordpress.com/wp-content/uploads/2019/05/instruccion-4-2019.pdf>>.
- Instrucción 1/2023, de la Secretaría de Estado de Seguridad, por la que se desarrolla la obligación de informar a la víctima de los antecedentes del agresor en los denominados casos de «agresores persistentes» y otras medidas de protección y seguridad en casos de violencia de género.
- Instrucción 4/2019, de la Secretaría de Estado de Seguridad, por la que se establece un nuevo protocolo para la valoración policial del nivel de riesgo de violencia de género (Ley Orgánica 1/2004), la gestión de la seguridad de las víctimas y el seguimiento de los casos a través del sistema de seguimiento integral de los casos de violencia de género. (Sistema VioGén). (2019). En línea: <<https://www.poderjudicial.es/cgpj/es/Temas/Violencia-domestica-y-de-genero/Guias-y-Protocolos-de-actuacion/Protocolos/>> (consulta: 27/12/2021).
- Entrevistas con José Luis González Álvarez y Juan José López Ossorio.
- LÓPEZ OSSORIO, Juan José (2016): *Construcción y validación de los formularios de valoración policial del riesgo de reincidencia y violencia grave contra la*

- pareja (VPR4.0 – VPER4.0) del Ministerio del Interior de España*. Tesis doctoral por compendio de publicaciones, en línea: <<https://repositorio.uam.es/handle/10486/678510>>
- LÓPEZ OSSORIO, Juan José; José Luis GONZÁLEZ-ÁLVAREZ y Antonio ANDRÉS-PUEYO (2016): «Eficacia predictiva de la valoración policial del riesgo de la violencia de género», *Intervención Psicosocial* 25, pp. 1-7
- LÓPEZ OSSORIO, Juan José; José Luis GONZÁLEZ-ÁLVAREZ; Ismael LOINAZ; Ana MARTÍNEZ-MARTÍNEZ y David PINEDA (2021). «La valoración policial del riesgo del homicidio de pareja en España: el protocolo dual VPR 5.0-H». *Intervención Psicosocial* 30(1), pp. 47-55, en línea: <<https://doi.org/10.5093/pi2020a16>>.
- LÓPEZ OSSORIO, Juan José; Ismael LOINAZ y José Luis GONZÁLEZ-ÁLVAREZ (2019): «Protocolo para la valoración policial del riesgo de violencia de género (VPR4.0): revisión de su funcionamiento», *Revista Española de Medicina Legal* 45(2), pp. 52-58.
- LÓPEZ OSSORIO, Juan José; José Manuel MUÑOZ VICENTE; Antonio ANDRÉS-PUEYO y Mar PASTOR BRAVO (2020): Guía de aplicación del formulario VPR 5.0-h en la valoración forense del riesgo, Ministerio del Interior – Área de Violencia de Género, Estudios y Formación / Gabinete de Coordinación y Estudios, en línea: <<https://www.otrosi.net/wp-content/uploads/2021/02/GUIA-VALORACION-DEL-RIESGO-FORENSE.pdf>>.
- MARTÍNEZ-ROMAN, María Asunción; Carmen VIVES-CASES y Carmen PÉREZ-BELDA (2017): «Immigrant Women Suffering From IPV in Spain: The Perspectives of Experienced Social Workers», *Affilia: Feminist Inquiry in Social Work* 2(32), pp. 202-216. DOI: 10.1177/0886109916667158.
- Memoria elevada al Gobierno de S. M. Presentada al inicio del año judicial por el Fiscal General del Estado Excmo. Sr. Don José Manuel Maza Martín. Sr. Don José Manuel Maza Martín (2017): Fiscalía General del Estado, vol. I, Madrid, en línea: <https://www.fiscal.es/memorias/memoria2017/FISCALIA_SITE/recursos/pdf/MEMFIS17.pdf>.
- Memoria elevada al Gobierno de S. M. Presentada al inicio del año judicial por la Fiscal General del Estado Excmo. Sra. Doña Dolores Delgado García. Sra. Doña Dolores Delgado García (2021): Fiscalía General del Estado, Madrid, en línea: <https://www.fiscal.es/memorias/memoria2021/FISCALIA_SITE/recursos/pdf/MEMFIS21.pdf>.
- Protocolo de valoración forense urgente del riesgo de violencia de género (2020): Consejo Médico Forense, Ministerio de Justicia, Secretaría General Técnica, Madrid, en línea: <<https://www.poderjudicial.es/cgpj/es/Temas/Violencia-domestica-y-de-genero/Guias-y-Protocolos-de-actuacion/Protocolos/>> (consulta: 27/12/21).
- Protocolo médico-forense de valoración urgente del riesgo de violencia de género (2011): Ministerio de Justicia, Madrid, en línea: <<https://violenciagenero.igualdad.gob.es/profesionalesInvestigacion/juridico/protocolos/docs/protocoloMedicoForense2011.pdf>>.
- RETAMOZO QUINTANA, Tatiana (coord.) y Silvina MONTEROS OBELAR (2020): *Mujeres migrantes víctimas de violencia de género en España. Segundo informe. Documento de análisis cuantitativo*, AIETI, en línea: <<https://>>

tirardelhilo.info/wp-content/uploads/2024/02/Documento_2DO-INFORME-VG-Y-MUJER-MIGRANTE_VF.pdf>.

SANTOS HERMOSO, Jorge y José Luis GONZÁLEZ-ÁLVAREZ (2017): «Homicidio de pareja (2007-2016): tiempo entre denuncia, valoración policial del riesgo y muerte», *Revista Comportamiento y Derecho* 3(1), pp. 1-10.

SANZ BARBERO, Belén; Fernando VALLEJO RUIZ DE LEÓN; Rocío CARMONA y Carmen VIVES-CASES (2021): *Análisis temporal de los asesinatos de mujeres por violencia de género en España a lo largo de 15 años (2003-2017)*, Delegación del Gobierno contra la Violencia de Género, Madrid, en línea: <https://violenciagenero.igualdad.gob.es/violenciaEnCifras/estudios/investigaciones/2021/estudios/analisis_asesinatos.htm>.

TORRECILLA, José Luis; Lara QUIJANO-SÁNCHEZ; Federico LIBERATORE; Juan José LÓPEZ OSSORIO y José Luis GONZÁLEZ-ÁLVAREZ (2019): «Evolution and study of a copycat effect in intimate partner homicides: A lesson from Spanish femicides», *PLoS ONE* 14(6): e0217914, en línea: <<https://doi.org/10.1371/journal.pone.0217914>>.

RISCANVI

1. DESCRIPCIÓN GENERAL

Antecedentes y ámbito de aplicación

RisCanvi es una herramienta de evaluación de riesgo que se utiliza en el sistema penitenciario catalán desde 2009. Su origen se encuentra en las conclusiones de la Comisión para el estudio de las medidas de prevención de la reincidencia en delitos graves, creada por el Departamento de Justicia de la Generalitat de Cataluña a raíz de la alarma social surgida en 2007 por la puesta en libertad de algunos presos condenados por delitos graves, que según los responsables penitenciarios no estaban rehabilitados. El RisCanvi surgió como respuesta a la preocupación de las autoridades del sistema penitenciario catalán por la reincidencia violenta de los exreclusos (especialmente los delincuentes sexuales) tras el cumplimiento de sus condenas (Andrés-Pueyo et al., 2018: 255). Como explica la propia Generalitat de Catalunya (2019: 3).

RisCanvi se diseñó para alcanzar dos objetivos principales en la gestión penitenciaria. El primer objetivo era mejorar las predicciones individualizadas del riesgo de violencia futura en forma de nuevos delitos violentos, quebrantamiento de permisos, y conductas violentas auto y hetero-dirigidas de las personas internadas en prisiones ordinarias y entre los liberados condicionales. Más concretamente, RisCanvi debía ser una herramienta profesional que consiguiera protocolizar la toma de decisiones con contenido y trascendencia pronóstica, reducir al máximo las decisiones erróneas y asegurar la transparencia de estas. Así, como segundo objetivo, se pretendía generalizar la aplicación de herramientas de valoración de riesgo como procedimiento de trabajo habitual para los profesionales penitenciarios y así poder introducir «buenas prácticas» en la gestión de la información dirigida a la toma de decisiones, compartir la información entre los diferentes técnicos, asegurando su constante actualización, y consolidar el uso de instrumentos de validez y utilidad contrastadas empíricamente.

RisCanvi fue diseñado y desarrollado en un proyecto conjunto de los servicios penitenciarios y el Grupo de Estudios Avanzados sobre Violencia de la Universidad de Barcelona entre 2007 y 2009. Inicialmente estimaba la probabilidad de cuatro conductas: a) violencia autodirigida (autolesiones, suicidio);

b) violencia intrainstitucional (dirigida a otros internos o al personal penitenciario); c) comisión de nuevos delitos violentos (reincidencia violenta); y d) riesgo de quebrantamiento de permisos penitenciarios o salidas, con escalas específicas para cada uno de estos criterios. Según Andrés-Pueyo et al. (2018: 261), en 2018 estaban trabajando en el desarrollo de una nueva escala de evaluación del riesgo de reincidencia general; la Generalitat de Catalunya afirma en 2019 que ya está desarrollada (2019: 3); y según un reciente artículo de prensa (Saura y Aragón, 2021) ya se está utilizando para predecir la reincidencia general.

En el caso de los presos preventivos, solo se aplicarán los resultados de los criterios de violencia autodirigida y violencia intrainstitucional (Generalitat de Catalunya, 2019: 8).

Según información de prensa (Saura y Aragón, 2021), desde su implantación, RisCanvi se ha utilizado más de 89.000 veces y se ha aplicado a más de 37.000 internos (el 96,6 % de la población penitenciaria catalana real ha sido evaluada al menos una vez). Cada interno es evaluado con el RisCanvi-S al entrar en el sistema y reevaluado cada seis meses. La versión RisCanvi-C¹ se aplica si hay cambios relevantes en la situación del interno, o si los niveles de riesgo son altos; también se aplica para ciertos grupos de casos como internos encarcelados por delitos violentos y graves.

Enfoque criminológico

El documento *El model de rehabilitació en les presons catalanes*, publicado por la Generalitat de Catalunya en 2011, explica que el modelo de rehabilitación en la ejecución de las sanciones penales se inspira en el modelo Riesgo-Necesidad-Responsividad desarrollado por Andrews y Bonta, un modelo muy influyente para la evaluación y tratamiento de los infractores (Andrews y Bonta, 2006). Sus principios básicos son:

- Riesgo: a quién dirigirse. El riesgo es la probabilidad de reincidencia del individuo. El riesgo puede evaluarse con las modernas herramientas de evaluación de riesgos y las intervenciones deben ajustarse al nivel de riesgo (deben ofrecerse intervenciones rehabilitadoras a los casos de riesgo moderado y alto, y los casos de bajo riesgo deben recibir una intervención mínima). Andrews y Bonta son los creadores de una herramienta de evaluación del riesgo ampliamente conocida y utilizada: el *Level of Service Inventory-Revised*TM (LSI-RTM) (<<https://storefront.mhs.com/collections/lis-r>>), y el *Level of Service / Case Management Inventory* (LS/CMI) (<<https://storefront.mhs.com/collections/lis-cmi>>).

¹ RisCanvi existe en dos versiones: la RisCanvi *Screening* (RisCanvi-S), que incluye 10 ítems de riesgo, y la versión completa, la RisCanvi *Complete* (RisCanvi-C), que incluye los 43 ítems de riesgo. Ambas estiman el riesgo para los cinco criterios diferentes (violencia autodirigida, violencia intrainstitucional, reincidencia violenta, riesgo de quebrantamiento de permisos o salidas de prisión y reincidencia general). Explicamos ambas escalas con más detalle a continuación.

- Necesidades: qué debe hacerse. Las intervenciones específicas deben centrarse en las necesidades criminógenas de cada individuo.
- Responsividad: cómo debe realizarse el trabajo. Este principio tiene dos partes: responsividad general y específica. La responsividad general promueve el uso de métodos cognitivos de aprendizaje social para influir en el comportamiento, porque se considera que estas estrategias son las más eficaces independientemente del tipo de delincuente. La responsividad específica establece que las intervenciones deben adaptarse, entre otras cosas, a los puntos fuertes, el estilo de aprendizaje, la personalidad, la motivación y las características biosociales (por ejemplo, sexo, raza) del individuo.

Según la Generalitat de Catalunya.

En Cataluña, el modelo de rehabilitación en el ámbito de la ejecución penal incorpora estas innovaciones teóricas y metodológicas y en 2007 el Departamento de Justicia Juvenil puso en marcha el Proyecto de Gestión del Riesgo con jóvenes infractores, con la herramienta de evaluación estructurada del riesgo SAVRY (*Structured Assessment of Violence Risk in Youth*) [...]. Posteriormente, en 2009, se inició la implantación del modelo de gestión de riesgos en el ámbito penitenciario con la herramienta principal de evaluación RisCanvi (2011: 15).

Fuentes de información para este informe

En febrero de 2022 presentamos una solicitud a la Generalitat de Catalunya para acceder a una muestra de datos de internos evaluados con RisCanvi, así como autorización para entrevistar a funcionarios de prisiones. Intercambiamos varios correos electrónicos para hacerles algunas preguntas sobre la herramienta, y en agosto el Departament de Justícia de la Generalitat de Catalunya nos comunicó que habían contratado a una consultora para realizar una auditoría externa del sistema RisCanvi, cuyos resultados se harían públicos a finales de 2022. También nos indicaron que, en función de los resultados de esta auditoría, estudiarían la posibilidad de dar acceso público a toda la base de datos del RisCanvi. Por esta razón, nos explicaron que habían dejado en suspenso la respuesta a nuestras peticiones hasta que finalizara esta auditoría externa, porque en el caso de que como resultado de esta decidieran hacer pública toda la base de datos, ya no sería necesario que nos dieran una muestra específica. En caso de que no hicieran pública la totalidad de la base de datos, y dado que nuestra solicitud seguiría pendiente, nos darían una respuesta más adelante, concediendo o denegando nuestra solicitud.

En febrero de 2023 recibimos respuesta de la Generalitat de Catalunya a nuestra petición de octubre de 2022 de que nos facilitaran la descripción detallada de los ítems del RisCanvi y de las instrucciones dadas a los trabajadores penitenciarios para codificarlos: nos enviaron la descripción de todos los factores de riesgo, tanto del RisCanvi-completo como del RisCanvi-*screening*. En 2024, la Generalitat ha colgado en su página web esta misma descripción de los ítems utilizados en ambas versiones de RisCanvi (*screening* y completo), junto con los resultados de la citada auditoría externa, a los que nos referiremos más adelante. Dado que nuestro informe ya estaba cerrado en esa fecha, hemos incluido la descripción de las variables de

ambos formularios, traducida al castellano (ya que estaba en catalán), como anejos de RisCanvi en este informe. Y hemos tenido en cuenta esta información para modificar parcialmente las conclusiones a las que habíamos llegado sobre RisCanvi en el apartado *Posibles sesgos y problemas de subjetividad*. En cualquier caso, es importante señalar que la descripción completa de los ítems que hemos recibido es un documento fechado en 2009. Sabemos que RisCanvi ha sido revisado desde entonces, por ejemplo, para añadir la nueva escala de riesgo de reincidencia general. También parece que la codificación de algunos ítems ha cambiado, ya que en el documento que nos han enviado todos los ítems de RisCanvi-C están codificados con tres respuestas posibles, mientras que la plantilla de codificación de variables de uno de los últimos estudios publicados por el CEFJE tiene algunas variables con cuatro respuestas posibles. Por lo tanto, podría ocurrir que la descripción actual de los ítems no sea exactamente la misma que se nos envió.

Aprovechamos en febrero de 2023 para preguntar a la Generalitat por la auditoría externa de RisCanvi que, según nos habían explicado en agosto de 2022, habían encargado y debía estar lista a finales de año. Nos contestaron que ya la habían recibido, que estaban analizando los resultados, y que a partir de marzo de 2023 decidirían qué tipo de difusión harían de la misma. El informe final de esta auditoría externa (así como un resumen ejecutivo) fue publicado por la Generalitat en acceso abierto en su página web en 2024 (<<https://links.uv.es/6BWTxrs>>). Dado que para entonces este informe ya estaba cerrado, solo hemos añadido algunas referencias muy breves a los resultados de esta auditoría relacionados con los aspectos que ya habíamos analizado y que la auditoría también aborda, en particular la capacidad predictiva de la escala de reincidencia violenta (apartado *Eficacia del enfoque policial predictivo para prevenir la delincuencia*), y los posibles sesgos hacia los extranjeros (en el apartado *Análisis cuantitativo de posibles sesgos*), pero no hemos incluido en este informe una valoración crítica de la auditoría en su conjunto.

Además de solicitar datos a la Generalitat, buscamos información sobre RisCanvi por otras vías.

La Generalitat de Catalunya ha publicado algunos estudios que contienen información sobre RisCanvi, en particular el informe de Capdevila et al. (2022) titulado «La reincidència en les excarceracions d'alt risc (2014-2016)», y ha colgado la base de datos en abierto en la web (<<https://cejfe.gencat.cat/ca/recerca/catalog/crono/2022/excarceracions-alt-risc/>>). El informe completo está disponible en esta web en catalán y en castellano; también hay disponible un resumen ejecutivo en inglés. Los resultados del análisis que hemos realizado con este conjunto de datos se resumen en Briz-Redón y Montes (2022), especialmente en las tablas 53-56.

La Generalitat de Catalunya había publicado previamente, en 2015, un estudio titulado *Taxa de reincidència penitenciària 2014* (Capdevila et al., 2015), sobre individuos excarcelados en 2010 y a los que se realizó un seguimiento hasta el 31 de diciembre de 2013. Tanto el documento como la base de datos asociada pueden encontrarse también en: <<https://cejfe.gencat.cat/ca/recerca/catalog/crono/2015/reincidencia-2014/>>. Sin embargo, no hemos realizado un análisis sobre los datos de este estudio porque de los 3.414 sujetos que componen la muestra solo 410 tenían al menos una evaluación RisCanvi en su versión completa, lo que conllevaba la existencia de una gran cantidad de datos perdidos, los correspondientes a las variables que forman parte del protocolo

y los relacionados con ellas. La baja proporción de evaluados con RisCanvi en este estudio previo (12 % del total de la muestra) se explica porque el protocolo se implantó en el sistema penitenciario catalán en 2009 y, por tanto, 2010 es el primer año en el que se dispone de datos para poder realizar una evaluación.

A finales de marzo de 2023, la Generalitat de Catalunya publicó un nuevo estudio sobre las tasas de reincidencia penitenciaria en Cataluña, CEJFE - Área de Investigación y Formación en Ejecución Penal: *Tasa de reincidencia penitenciaria 2020, Investigación propia*. Continúa la serie de estudios publicados sobre las tasas de reincidencia penitenciaria en Cataluña en los últimos años, el último de los cuales se publicó en 2015 (Capdevila et al., 2015; ver más arriba), y cuenta también con información sobre RisCanvi. La información sobre esta investigación se puede encontrar en línea: <<https://cejfe.gencat.cat/ca/recerca/catalog/crono/2023/taxa-reincidencia-penitenciaria/index.html>>. En esta misma página se puede descargar un resumen ejecutivo. Es un informe extenso, de más de 300 páginas, con mucha información interesante: ha seguido a personas condenadas y excarceladas en 2015 en Cataluña (N = 3.814 sujetos) hasta el 31 de diciembre de 2019, es decir, 4,5 años de media, con un mínimo de 4 y un máximo de 5 años. Además, siguiendo la política –parcial– de datos abiertos que la Generalitat ya había iniciado en estudios anteriores, también es posible descargar la hoja Excel con la codificación de todas las variables y los datos de reincidencia de cada sujeto desde la misma página web. Dado que este informe y los datos que lo acompañan se hicieron públicos cuando este informe ya había sido cerrado, no hemos realizado un análisis de estos, sino que únicamente hemos incluido algunas referencias muy puntuales y parciales a este estudio en nuestras conclusiones (apartados *Posibles sesgos y problemas de subjetividad y Transparencia*).

También nos pusimos en contacto con Carlos Castillo, profesor e investigador del Departament de Tecnologies de la Informació i les Comunicacions. Grupo de investigación en Ciencia Web y Computación Social, Universitat Pompeu Fabra (Cataluña), autor de algunos trabajos sobre RisCanvi. Él y su colega Marzieh Karimi-Haghighi disponían de una muestra de datos que habían obtenido del Departament de Justícia de la Generalitat de Catalunya, que no pudieron compartir con nosotros por estar sujetos a un acuerdo de confidencialidad, pero accedieron a realizar algunos análisis cuantitativos con su conjunto de datos para ayudar a nuestra investigación para el informe. Además, nos dieron información sobre los factores de riesgo que forman parte de cada una de las escalas de riesgo de RisCanvi (riesgo de reincidencia violenta, riesgo de quebrantamiento de permisos, riesgo de violencia intrainstitucional, riesgo de violencia autodirigida), aunque estas escalas pueden estar desactualizadas, porque la versión de la que ellos disponen es de 2019 y RisCanvi ha sido revisado desde entonces, y entre otras cosas se ha incluido la nueva escala de riesgo de reincidencia general.

Tanto los resultados del análisis que han realizado para nosotros como la descripción de la escala RisCanvi de reincidencia violenta pueden consultarse en el informe técnico que han publicado Karimi-Haghighi y Castillo (2022); y los cálculos que nosotros hemos realizado a partir de los datos contenidos en dicho informe pueden encontrarse en Briz-Redón y Montes (2022).

Es muy importante señalar que, aunque RisCanvi estima cinco tipos diferentes de riesgo (riesgo de violencia autodirigida, violencia institucional,

incumplimiento de permisos, reincidencia violenta y reincidencia general), solo hemos podido obtener datos relativos a las estimaciones realizadas con la escala que mide el riesgo de reincidencia violenta (escala REVI), porque es la única utilizada en el informe de Karimi-Haghighi y Castillo y en el estudio de Capdevila et al. (2022). Por lo tanto, nuestro análisis está muy limitado a un aspecto específico de RisCanvi y no puede considerarse una auditoría completa de esta herramienta en su conjunto.

Además, los datos de reincidencia de la muestra de Karimi-Haghighi y Castillo no se refieren a la comisión de un nuevo delito violento después de la fecha de puesta en libertad, sino a la comisión de un nuevo delito violento que haya conducido al encarcelamiento. La lista de delitos que se consideran «violentos» a efectos de su análisis puede consultarse en el apéndice de su informe, y son principalmente robos con violencia o intimidación.

El informe de Capdevila et al. (2022) sí incluye no solo la reincidencia violenta que conduce al encarcelamiento después de la fecha de puesta en libertad, sino también un análisis separado de la reincidencia violenta con otros tipos de penas y de la reincidencia general. Pero a efectos de nuestro informe nos centraremos (salvo que indiquemos lo contrario) en los resultados de la reincidencia violenta que conduce al encarcelamiento, porque así podemos comparar los resultados que obtenemos de ambas muestras, la de Capdevila (2022) y la de Karimi-Haghighi y Castillo.

La muestra de Castillo es de 2.514 personas, que son todas las excarceladas en Cataluña en 2016 en libertad definitiva o libertad condicional, con un periodo de seguimiento de cinco años. De ellas, según la tabla de la primera página, 1.707 han sido evaluadas con RisCanvi completo y 807 con *RisCanvi Screening*. Este último solo tiene dos niveles de riesgo: alto y bajo. Cuando el RisCanvi-S da riesgo bajo, esa es la clasificación que se introduce en el sistema; cuando da riesgo alto, lo que se hace es pasar el RisCanvi-C, que ya tiene los tres niveles de riesgo. Por lo tanto, para los 807 internos que solo fueron procesados por el RisCanvi-S, el sistema debió estimar un riesgo bajo de reincidencia. Solo el 10,5 % de los internos de esta muestra tenían un riesgo alto de reincidencia violenta según RisCanvi, el 16,7 % tenían un riesgo medio y el 72,8 % tenían un riesgo bajo.

La muestra de Capdevila et al. es diferente. Está formada por un grupo de estudio que comprende a todas las personas que salieron en libertad definitiva desde 1.^{er} o 2.^o grado penitenciario, o de medidas de seguridad, entre los años 2014-2016 y que, en el momento de la excarcelación, presentaban una evolución negativa en el centro penitenciario y un pronóstico alto de reincidencia violenta según los criterios RisCanvi (N = 352). También comprende, como grupos de control, a las personas que salieron en libertad definitiva desde 1.^{er} o 2.^o grado entre los años 2014-2016, y que, en el momento de la excarcelación, presentaban un pronóstico de riesgo medio o bajo de reincidencia violenta según los criterios RisCanvi (N = 355 para riesgo medio y N = 365 para riesgo bajo). El total de la muestra es, por tanto, de 1.072 sujetos.

Esta muestra tiene una composición muy peculiar porque contiene una proporción muy alta de individuos de alto riesgo (un tercio), que no se corresponde con el porcentaje de individuos de alto riesgo en el colectivo de personas excarceladas cada año en Cataluña (por ejemplo, si la comparamos con la muestra de Karimi-Haghighi y Castillo, este grupo representa solo el 10,5 % de los internos excarcelados).

El seguimiento en el estudio de Capdevila et al. duró desde el momento de su puesta en libertad definitiva en los años 2014, 2015 y 2016 hasta el 30 de septiembre de 2019, por lo que el periodo de seguimiento va desde un mínimo de 2,8 años hasta un máximo de 5,7 años.

2. DATOS

Datos de entrada

Datos de entrenamiento

Según Andrés-Pueyo et al. (2018: 258-9):

El punto de partida para el desarrollo del sistema RisCanvi fue un estudio retrospectivo de casos y controles con una muestra de 643 casos y 428 controles. El estudio analizó la prevalencia de cuatro criterios de riesgo relacionados con la violencia en las prisiones e identificó los factores relevantes para cada uno de ellos [...]. Se seleccionaron y analizaron cientos de registros penitenciarios que constituyeron la base empírica para la elaboración del instrumento. Su elaboración incluyó un análisis retrospectivo de 106 factores de riesgo delictivo y de violencia en casos y controles obtenidos de cuatro fuentes: otros protocolos de evaluación de riesgo penitenciario preexistentes (LSI-R, COMPAS², FOTRES, OASYS, HCR-20, etc.); pautas de evaluación de riesgo de violencia más específicas (SVR-20, SARA, etc.); tests clínicos y de personalidad (GHQ y NEO-PI-R); y otros factores de riesgo (basados en la experiencia profesional del personal). A partir de este amplio borrador inicial y mediante un estudio de casos y controles, pudimos identificar 43 factores de riesgo que nos permitieron determinar la probabilidad de cometer un nuevo acto violento o de quebrantar la libertad condicional. Una vez identificados estos factores, utilizamos la regresión logística binaria para seleccionar los factores de riesgo que eran específicos para predecir cada criterio.

Según información de prensa, el algoritmo ya ha sido recalibrado cuatro veces para mejorar su rendimiento (Saura y Aragó, 2021). En cuanto a los motivos de las sucesivas versiones, el *Manual d'aplicació del protocol de valoració RisCanvi* publicado por la Generalitat de Catalunya en 2019 (2019: 3) indica que «siguen apostando por la revisión y mejora de este protocolo de evaluación de riesgos». En esta tercera versión (RisCanvi v3), además del reajuste del algoritmo, un nuevo tipo de delito base (contra la seguridad del tráfico) y la ampliación del tipo de víctima, el cambio más relevante es la incorporación de un quinto criterio de predicción relativo a la reincidencia delictiva general. No hemos encontrado información adicional de acceso público sobre las razones más detalladas de las actualizaciones y revisiones de la herramienta.

² Quizá merezca la pena señalar que un informe publicado por Algorithmwatch (Bellio, 2021) incluye la opinión del profesor Andrés-Pueyo, indicando que RisCanvi «no se basa en Compas», sino que «gran parte del diseño de la herramienta se basa en el protocolo canadiense *Level of Service Inventory-Revised*».

Recogida de datos y codificación

RisCanvi-C utiliza 43 factores de riesgo relacionados con el historial delictivo/penitenciario del delincuente (como la duración de las condenas penales, el tiempo cumplido en prisión, el historial de violencia, el inicio de la actividad delictiva o violenta, etc.); información biográfica (incluidos elementos como la mala adaptación en la infancia, la distancia entre el lugar de residencia y el centro penitenciario, el nivel educativo, los problemas relacionados con el empleo, entre otros); factores relacionados con el entorno familiar/social (incluidos los antecedentes penales de la familia o los padres, las dificultades en la socialización o el desarrollo en la familia de origen, la falta de apoyo familiar o social, etc.); factores clínicos (como el abuso o la dependencia de las drogas o el alcohol, los trastornos mentales graves, la promiscuidad sexual o parafilia, entre otros); y rasgos de personalidad o actitudinales (como actitudes prodelictivas o antisociales, temeridad, hostilidad o irresponsabilidad). Factores de riesgo de las versiones screening y completa de RisCanvi (para una descripción más detallada de cada factor de riesgo, véanse los anexos *RisCanvi completo* y *RisCanvi Screening* del presente informe).

<i>RisCanvi Screening</i>	<i>RisCanvi Completo</i>	
<i>Factores de riesgo</i>	<i>Factores de riesgo</i>	
1. Inicio de la actividad delictiva o violenta.	Grupo 1 Criminal/ penitenciario	1. Delito base violento.
2. Historial de violencia.		2. Edad en el momento del delito base.
3. Problemas de conducta penitenciaria.		3. Intoxicación durante la comisión del delito base.
4. Evasiones. quebrantamientos o incumplimientos.		4. Víctimas con lesiones.
5. Problemas de consumo de drogas o alcohol.		5. Duración de la pena.
6. Respuesta limitada al tratamiento psicológico o psiquiátrico.		6. Tiempo ininterrumpido en prisión.
7. Intentos o conductas de autolesión.		7. Historial de violencia.
8. Falta de recursos económicos.		8. Inicio de la actividad delictiva o violenta.
9. Falta de apoyo familiar y social.		9. Aumento de la frecuencia, gravedad y diversidad de los delitos.
10. Actitudes hostiles o valores procriminales.		10. Conflictos con otros internos
	Grupo 2 Biografía	11. Incumplimiento de medidas penales.
		12. Expedientes disciplinarios.
		13. Evasiones o fugas.
		14. Regresión de grado.
		15. Quebrantamiento de permisos
		16. Desajuste infantil.
		17. Distancia entre la residencia habitual y el centro penitenciario.
		18. Nivel educativo.
		19. Problemas relacionados con el empleo.
		20. Falta de recursos económicos.
		21. Ausencia de planes viables de futuro.

<i>RisCanvi Screening</i>	<i>RisCanvi Completo</i>	
<i>Factores de riesgo</i>	<i>Factores de riesgo</i>	
	Grupo 3 Familia/ social	22. Antecedentes delictivos en la familia de origen.
		23. Socialización o crianza problemática en la familia de origen.
		24. Falta de apoyo familiar o social.
		25. Pertenencia a bandas delictivas o amigos delincuentes.
		26. Miembro de grupos sociales de riesgo delictivo.
		27. Rol delictivo destacado.
		28. Víctima de violencia de género (solo mujeres).
		29. Cargas familiares actuales.
	Grupo 4 Clínica	30. Abuso o dependencia de drogas.
		31. Abuso o dependencia del alcohol.
		32. Trastorno mental grave.
		33. Promiscuidad sexual y de riesgo. o parafilia.
		34. Respuesta limitada al tratamiento psicológico o psiquiátricos.
		35. Trastorno de la personalidad relacionado con la ira. la impulsividad o la violencia.
		36. Pobre afrontamiento del estrés.
		37. Intentos o conductas de autolesión.
	Grupo 5 Actitudes/ personalidad	38. Actitudes procriminales o valores antisociales.
		39. Baja capacidad mental e inteligencia.
		40. Temeridad.
		41. Impulsividad e inestabilidad emocional.
		42. Hostilidad.
		43. Irresponsabilidad.

Además de los factores de riesgo, todas las escalas incorporan cuatro variables que moderan el efecto de los factores incluidos en el algoritmo: la edad (superior o inferior a 30 años), el sexo (hombre o mujer), el país de origen (español o extranjero) y el estado o situación penitenciaria (preventivo o penado). Estas cuatro variables incluidas en todas las escalas de RisCanvi tienen diferente peso en función de los criterios predictivos de violencia autodirigida, violencia intrainstitucional, reincidencia delictiva general, reincidencia delictiva violenta o quebrantamiento de permisos (Generalitat de Catalunya, 2019: 10).

La información necesaria para codificar los factores de riesgo de RisCanvi es proporcionada por los diferentes profesionales que trabajan con el interno en el centro penitenciario. En Cataluña, cada interno tiene asignado un equipo multidisciplinar (arts. 37 y 38 del Decreto 329/2006, de 5 de septiembre, por el que se aprueba el Reglamento de organización y funcionamiento de los servicios de ejecución penal en Cataluña).³ Estos equipos están formados por (al menos): un psicólogo, un jurista, un educador, un docente, un médico, un trabajador social, un monitor ocupacional o insertor laboral, y un monitor deportivo. Estos profesionales trabajan regularmente con los reclusos. En cada centro penitenciario hay uno o varios equipos multidisciplinarios, dependiendo del número de reclusos que haya, de la plantilla del centro, del tipo de programas especializados que tenga, etc. Los miembros del equipo se reúnen todas las semanas (en las reuniones estarán presentes aquellos profesionales que tengan relación con los temas a tratar en cada ocasión) y entre sus funciones están elaborar el programa de tratamiento o modelo de intervención individual para cada interno, ponerlos en práctica, evaluar los objetivos alcanzados, realizar propuestas a la junta de tratamiento, recoger y atender las peticiones y quejas de los internos, etc. Los miembros de estos equipos multidisciplinarios son los encargados de recabar la información necesaria para codificar RisCanvi.⁴

A continuación, explicamos la recogida y codificación de los factores de riesgo de RisCanvi según la información contenida en el *Manual de aplicación del protocolo de valoración RisCanvi* publicado por la Generalitat de Catalunya (2019). La recogida de información se ha de realizar de forma continua y especialmente en los momentos en los que se han de actualizar las evaluaciones de riesgo. Es una tarea que forma parte de las funciones habituales de los profesionales de la rehabilitación, que tienen que adaptar sus métodos de recogida de información al contenido de los factores de riesgo: «Así, el contenido del RisCanvi es uno de los elementos (no el único) a tener en cuenta en la elaboración del guion de entrevista y otros procedimientos de evaluación por parte de los distintos profesionales en los diferentes contextos y momentos críticos (ingresos, clasificación, seguimiento, etc.)» (2019: 15).

Una vez recogida la información, los miembros de los equipos multidisciplinarios deben registrar en el Sistema de Información Penitenciaria de Cataluña (SIPC) las «evidencias» que han observado:⁵ las evidencias son la información objetiva que se ha obtenido respecto a cada factor de riesgo. Las evidencias deben ser descrip-

³ La composición y funciones de estos equipos son muy similares a las reguladas para el resto de España por los artículos 274 y 275 del Real Decreto 190/1996, de 9 de febrero, por el que se aprueba el Reglamento Penitenciario.

⁴ «Jordi Camps, jefe del servicio de rehabilitación del departamento catalán de atención a las víctimas, ha declarado a AlgorithmWatch que 750 personas trabajan a diario con la herramienta, entre psicólogos, criminólogos juristas, trabajadores sociales y educadores sociales. Entre ellos, cien personas actúan como “validadores”. Ellos deciden si los datos introducidos en RisCanvi por los demás profesionales coinciden con el nivel de riesgo que arroja el algoritmo» (Bellio López Molina, 2021)

⁵ El RisCanvi está integrado en el SIPC.

tivas, relevantes y breves. En el Manual de Aplicación de RisCanvi se dan algunos ejemplos de evidencias que podrían corresponder al factor de riesgo «Problemas con el consumo de drogas o alcohol»:

En los últimos 3 meses, de 12 análisis, 9 han dado positivo en cocaína u opiáceo. «Desde el 25 de marzo ha dejado el trabajo y ha comenzado el programa de centro de día por indicación del equipo». «Desde mayo, las comunicaciones con su pareja se han interrumpido porque ella intentó llevarle heroína en el último vis a vis». «El maestro informa de que en las últimas semanas ha bajado su rendimiento escolar: se distrae, ha faltado a varias clases y ha tenido discusiones con otros internos. Según él, parece que cuando ocurren estas cosas el interno está bajo los efectos de alguna sustancia», «en marzo del año pasado se encontraron ‘porros’ en la celda» (Generalitat de Catalunya, 2019: 16).

Cuando se introducen evidencias en el SIPC, el sistema registra qué profesional lo ha hecho. Además de la información, el profesional también debe registrar la fuente de la que la ha extraído. En cuanto a las fuentes y la objetividad de todas estas observaciones, el manual de aplicación de RisCanvi incluye algunas indicaciones que citamos textualmente:

Es muy importante que determinada información, como la relativa a actitudes, aspectos biográficos, planes de futuro, apoyo exterior o conducta habitual, se obtenga a través de distintas fuentes, además de la entrevista con el interno. En este sentido, la validez de la información mejora sensiblemente si se utilizan distintas fuentes (consulta de documentación, entrevistas a personas de su entorno, coordinación con otros servicios...), y si se recoge en distintos momentos. Por este motivo, los equipos han de procurar que la información sea contrastada mediante la utilización de distintas fuentes, obteniéndola en diferentes momentos e incluso (cuando pueda estar afectada por la subjetividad) por distintos profesionales. [...] [P]ara la mayoría de los casos será suficiente con los procedimientos básicos (entrevistas, consulta de documentación...). Sólo cuando estos procedimientos básicos indiquen que puede estar presente algún factor de riesgo que requiera un juicio profesional más específico (trastorno mental, rasgo de personalidad, discapacidad, apoyo social criminológico...) será necesario utilizar procedimientos de evaluación más complejos o más extensos, como autoinformes, pruebas psicológicas, observación de conducta, nuevas entrevistas en mayor profundidad, etc. (Generalitat de Catalunya, 2019: 15).

Cualquier profesional puede introducir evidencias sobre cualquier factor. No obstante, RisCanvi cuenta con una distribución de factores en función de las áreas de especialización de los diferentes miembros de los equipos multidisciplinares. Así, el profesional responsable de la codificación de cada factor es también responsable de asegurar que existe información contrastada y relevante para valorar los factores que son de su competencia y, como mínimo, introducirá evidencias en ellos, sin perjuicio de que también pueda registrar evidencias en todos los demás. La distribución es la siguiente:⁶

⁶ Para aquellos equipos en los que participen activamente pedagogos u otros profesionales en la evaluación RisCanvi, o en los que la organización del equipo tenga características espe-

Tabla 14
Personal penitenciario responsable de la recogida de pruebas sobre los factores de riesgo de RisCanvi

<i>Profesionales responsables</i>	<i>Items</i>
Jurista	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 14 y 15
Educador/a	10, 12, 16, 18, 25, 26, 27 y 43
Trabajador/a social	17, 19, 20, 21, 22, 23, 24 y 29
Psicólogo/a	28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 y 42

Fuente: Generalitat de Catalunya, 2019, p. 16

Una vez registradas las evidencias en el SIPC, se realiza la codificación de los factores de riesgo. La responsabilidad de la correcta codificación recae en el denominado «Validador». Según la información que se puede encontrar en el Manual de Aplicación de RisCanvi,

los validadores son responsables de los equipos multidisciplinares o, en algunos casos, técnicos especialmente acreditados para esta función. Los validadores de cada centro, ya sean responsables o técnicos, han recibido formación específica para esta función [...]. Tienen la responsabilidad de dirigir el proceso de recogida de información, de ponerla en común en las reuniones de equipo y de la valoración final del riesgo. Cuando ejecutan la acción de cerrar la evaluación, garantizan que ésta se ha realizado de acuerdo con el procedimiento establecido [...] (2019: 13-14).

Según la información que se puede encontrar en el Resumen RisCanvi publicado por la Generalitat de Catalunya.

La valoración del factor de riesgo (sí/no por la escala R-s y sí/?/no por la escala R-c) es tarea del equipo multidisciplinar encargado de realizar la evaluación, la cual es ejecutada por la figura del validador [...]. Esta valoración debe surgir necesariamente de la lectura de la evidencia que es la información que permite valorar la presencia o ausencia del factor. También se puede hacer una determinación de que no existen datos disponibles para poder valorar el factor en cuestión (s. f.: 6).

En cuanto a las instrucciones o pautas para codificar los factores de riesgo, la definición de cada factor RisCanvi puede encontrarse en numerosos documentos, por ejemplo, los anexos técnicos que se incluyen en los resultados de las investigaciones publicadas con datos abiertos del Departament de Justícia de la Generalitat de Catalunya (y disponible en inglés en el documento «Descripción de variables y codificación» que acompaña al estudio «Tasa de reincidencia penitenciaria 2014»). Pero se trata de descripciones muy breves que no permiten entender fácilmente qué

ciales, se podrán realizar modificaciones en la distribución de los ítems incluidos en esta tabla (Generalitat de Catalunya, 2019: 16).

situaciones o conductas podrían incluirse. Por ejemplo, algunas de las definiciones del documento que acabamos de citar son:

Tabla 15
Descripción de algunos factores de riesgo de RisCanvi

<i>Variable</i>	<i>Descripción</i>	<i>Codificación</i>
19. Problemas relacionados con el empleo	Desempleo crónico, inestabilidad laboral o muchos trabajos previos que duran poco tiempo. Dificultades de inserción laboral. Se niega a buscar trabajo, le despiden a menudo o abandona el trabajo con facilidad. <i>Considerar toda la vida del recluso</i>	1 = Sí 2 = No
20. Falta de recursos económicos	Nivel económico suficiente en el último año o antes de su ingreso en el caso de que haya estado más de 12 meses en el centro penitenciario.	1 = Sí 2 = No
25. Pertenencia a bandas	Es miembro de una banda organizada relacionada con actividades delictivas o se vincula con redes criminales. <i>Considerar el momento actual</i>	1 = Sí 2 = No
26. Pertenencia a grupo social de riesgo	Pertenece a grupos sociales que pueden estar en riesgo de cometer actos delictivos. <i>Considerar el momento actual</i>	1 = Sí 2 = No
36. Pobre afrontamiento del estrés	Tendencia a experimentar un alto nivel de estrés. Falta de recursos cognitivos y conductuales para afrontar (reducir, minimizar, dominar o tolerar) las demandas internas y externas de una determinada situación estresante. <i>Considerar el momento actual</i>	1 = Sí 2 = No

Fuente: documento «Descripción de variables y codificación» que acompaña al estudio «Tasa de reincidencia penitenciaria 2014», en línea: <https://repositori.justicia.gencat.cat/bitstream/handle/20.500.14226/270/variables_reincidencia_penitenciaria_2014.pdf?sequence=12&isAllowed=y>.⁷

El SIPC contiene una definición operativa más detallada de cada factor, pero estas definiciones no están disponibles públicamente. Preguntamos al Departament de Justícia de la Generalitat si podían facilitarnos las definiciones, pero no contestaron a este correo electrónico antes de la fecha de cierre de este informe a finales de 2022. Sin embargo, la Generalitat sí nos envió una descripción completa de los factores RisCanvi en febrero de 2023 (aunque desconocemos si se trata de una versión actualizada), que hemos incluido como anexo (ver anexos *RisCanvi completo* y *RisCanvi Screening* de este informe). Para la elaboración de este informe hemos trabajado, por tanto, con la descripción completa de solo algunos de los ítems, que el Dr. Antonio Andrés Pueyo, catedrático de Psicología de la Universitat de Barcelona y uno de los creadores de RisCanvi, compartió con nosotros:

⁷ Este documento contiene la descripción de muchas variables; las que son factores de riesgo de RisCanvi se identifican con una «@».

19. Problemas relacionados con el empleo: desempleo crónico, inestabilidad laboral o muchos empleos anteriores de corta duración. Dificultades para encontrar trabajo. Se niega a buscar trabajo, es despedido con frecuencia o lo abandona con facilidad. Considerar también el trabajo realizado dentro de la prisión. Los problemas laborales no están causados por una discapacidad física o una incapacidad legal. Considerar toda la vida del individuo.

Este ítem se refiere a las posibles dificultades del sujeto en relación con el entorno laboral y las relaciones prosociales que de él se derivan. Los individuos puntuarán en este factor de riesgo si se caracterizan por: negarse a buscar un trabajo legítimo (social y legalmente aceptable), tener un historial previo de muchos trabajos en cortos periodos de tiempo, haber sido despedido o abandonar frecuentemente el trabajo. El ítem evalúa posibles problemas actitudinales, de responsabilidad y relacionales del sujeto con el trabajo, no su capacidad/discapacidad física o mental para encontrar un empleo.

Los cambios recientes en la situación laboral oficial deben examinarse detenidamente para ver si pueden indicar una mejora. Sin embargo, no todos los cambios formales o legales en la situación laboral deben interpretarse como positivos a efectos de este punto. Por ejemplo, una persona con un historial previo de 15 años de problemas con el trabajo pero que ha obtenido recientemente una pensión por incapacidad laboral puede seguir puntuando «sí» (es decir, negativo) en este factor de riesgo.

En el caso de los inmigrantes, considerar si han mantenido una trayectoria laboral asalariado o constante, independientemente de su situación legal (incluido si tienen o no permiso de trabajo).

Tener en cuenta también la participación del sujeto en trabajos dentro del centro penitenciario o en programas de formación. Tener en cuenta que este factor de riesgo valora la disponibilidad mostrada por el sujeto para participar en actividades laborales tanto dentro como fuera del centro penitenciario.

NO: ha sido una persona activa y trabajadora, sin que consten problemas sistémicos relacionados con el trabajo.

?: presencia posible/moderada de problemas relacionados con el trabajo.

SÍ: ha tenido problemas laborales graves o muy graves.

20. Falta de recursos económicos: Nivel económico estimado en el último año o antes del ingreso en el caso de llevar más de 12 meses en el centro penitenciario. Se refiere al patrimonio y recursos económicos (ingresos, dinero) que tiene el interno para hacer frente, con suficiente independencia, a las exigencias económicas de la vida diaria en la calle. Debe considerarse conjuntamente la situación económica actual (o la previa al ingreso en prisión) y la que sería probable si el interno disfrutara de un permiso o de régimen abierto. Los recursos económicos pueden proceder del propio individuo, de su núcleo familiar o de los servicios sociales. Factores que pueden informar sobre el apoyo económico por parte de su familia son: hábitos de trabajo y nivel de estabilidad laboral de los miembros de la familia, bienes o ahorros de que dispone el sujeto, nivel de endeudamiento, equilibrio entre ingresos y gastos, y entre ingresos y *peculio* del interno,⁸ y nivel de cobertura de las necesidades mínimas.

⁸ A grandes rasgos, «peculio» es el nombre que la legislación penitenciaria española da a los depósitos que los familiares o allegados pagan a los presos para que puedan realizar compras dentro de los centros penitenciarios, en el economato de cada módulo, o los encargos que realizan al responsable, como café, refrescos, tabaco, mantas, televisores, artículos de aseo,

NO: Dispone de recursos económicos suficientes (incluso medios o elevados) para cubrir sus necesidades básicas (por ejemplo, alimentación o vivienda). Los recursos económicos no son un aspecto problemático para este sujeto.

?: Dispone de algunos recursos económicos, pero éstos son insuficientes o inestables para cubrir sus necesidades básicas, lo que hace que la vida cotidiana actual del interno sea crítica o problemática. Los recursos económicos son un aspecto problemático para este sujeto.

SÍ: el interno carece de apoyo económico para cubrir sus necesidades básicas. Su subsistencia o la de la familia a su cargo depende de la asistencia social.”

25. Pertenencia a bandas delictivas o amigos delincuentes: Es miembro de una banda organizada relacionada con actividades delictivas o se vincula a redes criminales. Considerar el momento actual.

El individuo tiene amigos (con los que pasa el tiempo libre, cuyas opiniones valora, que podrían ayudarlo en las dificultades...) con historial delictivo e involucrados en actividades delictivas. Las bandas organizadas o redes criminales se refieren a grupos de tres o más miembros que delinquen conjuntamente. Los miembros suelen tener una relación estrecha y una ideología común. Las redes criminales suelen tener motivaciones económicas o políticas.

NO: No hay evidencia de que el sujeto pertenezca a ninguna banda delictiva o tenga amigos delincuentes.

?: Hay evidencia parcial o dudosa de que el sujeto pertenece a una banda delictiva o tiene amigos delincuentes.

SÍ: Hay evidencia clara de que el sujeto pertenece a una banda criminal o tiene amigos implicados en actividades delictivas.

36. Pobre afrontamiento del estrés. Tendencia a experimentar un alto nivel de estrés. Falta de recursos cognitivos y conductuales para afrontar (reducir, minimizar, dominar o tolerar) las demandas internas y externas de una determinada situación estresante. Considerar el momento actual.

Este factor de riesgo implica anticipar qué fuentes de estrés puede encontrar el individuo y cómo reacciona y las afronta. Es necesario tener en cuenta la vulnerabilidad particular del individuo y la capacidad y los recursos de que dispone para superar y hacer frente a las dificultades. Puede ser útil disponer de algún indicador de su nivel de estrés, evaluado en función de los problemas a los que ha estado expuesto en su vida cotidiana. Deben tenerse en cuenta las habilidades de afrontamiento, ya que un alto nivel de estrés en combinación con buenas habilidades de afrontamiento puede dar lugar a una puntuación baja en este factor de riesgo.

La soledad y la alienación parecen favorecer la vulnerabilidad ante acontecimientos vitales negativos, que pueden conducir al hundimiento emocional y, en determinadas condiciones, a una acción suicida. Por lo tanto, estas situaciones deben tenerse especialmente en cuenta en la evaluación de este factor de riesgo.

tarjetas telefónicas, etc. Si el preso tiene un trabajo remunerado (por ejemplo, trabaja en la cocina de la prisión), su sueldo se ingresará en esta cuenta de «peculio» y, previa autorización, los fondos podrán transferirse fuera de la prisión. Por último, estas cuentas también pueden recibir el dinero que la persona lleve en efectivo el día de su ingreso en prisión. Como el uso de efectivo no está permitido en las prisiones, los reclusos reciben una tarjeta «peculio», como si fuera una tarjeta de crédito. Existe un límite de saldo y de gasto semanal, liquidándose el importe sobrante en el momento de la excarcelación. Regulación: artículos 319 a 324 del Reglamento Penitenciario de 1996.

NO: No experimenta estrés, el nivel de estrés experimentado es bajo o sus recursos de afrontamiento son adecuados.

?: Está expuesto a un nivel de estrés entre moderado y alto, pero lo ha afrontado adecuadamente en el pasado.

SÍ: Está experimentando un nivel estrés entre moderado y alto y no ha sido capaz de afrontarlo satisfactoriamente.

El profesor Andrés Pueyo destacó en su comunicación que todos los ítems tienen su definición operativa inicial y también recomendaciones para su aplicación que, además, cuentan con una «biblioteca» de ejemplos y casos prácticos que se estudian para acercarse a una mejor práctica profesional, y dijo que estos ítems están archivados.

Existe, además, un «Equipo de apoyo, orientación y seguimiento en la aplicación del RsCanvi» (equipo SOS-RisCanvi). El objetivo de este equipo es realizar un seguimiento minucioso de la aplicación de RisCanvi. Esto incluye acompañar a los equipos multidisciplinares en la implementación y aplicación de las escalas, ofreciendo apoyo y resolviendo las dudas y problemas que puedan surgir en su aplicación. El procedimiento consiste en que los validadores recogen las dudas y demandas planteadas por los diferentes profesionales y se ponen en contacto con el equipo SOS-RisCanvi, a través de una dirección de correo electrónico creada al efecto, o bien las consultas pueden realizarse telefónicamente. El equipo SOS también ha desarrollado un sistema de avisos recordatorios, sobre temas relacionados con la escala, su uso, la implementación, etc. Los subdirectores de Tratamiento de los centros penitenciarios y los responsables de los servicios sociales, así como los profesionales validadores, reciben periódicamente estos avisos en sus buzones, para su conocimiento y para que los difundan al resto de profesionales de rehabilitación (Generalitat de Catalunya, 2019: 24).

Nos parece que las autoridades y los profesionales responsables de RisCanvi han hecho un serio esfuerzo para diseñar la herramienta de forma que la información recogida sea lo más veraz, fiable y objetiva posible. El profesor Andrés-Pueyo también destacó en su comunicación por correo electrónico que todos los profesionales que trabajan con RisCanvi recibieron/reciben formación en la definición de los factores de riesgo y que se trata de una formación exhaustiva. Y añadió que uno de los puntos fuertes de la herramienta, en su opinión, es que el proceso de codificación está diseñado para garantizar la máxima validez y fiabilidad de las valoraciones.

Sin embargo, en nuestra opinión, la formación y las definiciones no pueden garantizar que la información sea siempre exacta y que no contenga prejuicios ni sesgos. Por un lado, porque las distintas fuentes de las que puede proceder la información son a su vez incompletas, inexactas o no del todo fiables. Por ejemplo, no podemos estar seguros de que los presos digan siempre la verdad cuando hablan con el personal penitenciario sobre su vida, su familia, sus antiguos delitos y sus planes para el futuro. El personal penitenciario está ahí para ayudarles, pero también para controlarlos, y los presos, especialmente cuando han sido asesorados por un abogado, saben que quizá les convenga no revelar algunos hechos (delitos pasados aún no descubiertos, u otros problemas). Por otro lado, los registros oficiales no siempre coinciden con los delitos realmente cometidos (delitos no descubiertos, condenas erróneas, conformidades que llevan a aceptar

una pena de prisión corta por temor a que el juicio pueda acabar con una pena más severa, etc.).

Además, la verificación de muchos de los factores con los que trabaja RisCanvi no es un proceso mecánico, sino que implica necesariamente la realización de valoraciones por parte de la/s persona/s que codifican la información. Por ejemplo, para codificar el factor «problemas relacionados con el empleo», la propia definición del factor exige comprobar no tanto si la persona tiene o ha tenido empleos estables, sino si tiene o no una actitud favorable hacia el trabajo legal. La concatenación de muchos trabajos en poco tiempo, que se pone como ejemplo, puede deberse a que las personas los abandonan porque se cansan o se aburren, pero también puede ser que las condiciones precarias del mercado laboral les obliguen a ello porque, aunque quisieran conservar su empleo, solo consiguen trabajos legales temporales de corta duración. RisCanvi pide a los calificadores que se posicionen al respecto y esta posición implica una valoración. Y aunque participen varias personas del equipo técnico, eso no significa que desaparezca la dimensión valorativa: puede ser una valoración compartida por varias personas, pero sigue siendo una valoración que depende de lo que piensen.

En resumen, coincidimos en considerar que cada ítem presente o ausente en estas herramientas depende mucho de cómo el evaluador humano interpreta la realidad, y esto también fue destacado por los abogados con los que hablamos en relación con RisCanvi, y también por varias personas entrevistadas en relación con VioGén.

Datos de salida

RisCanvi existe en dos versiones: *RisCanvi Screening* (RisCanvi-S), que contiene 10 factores de riesgo, y la versión completa, *RisCanvi Complete* (RisCanvi-C), que contiene los 43 ítems de riesgo. Ambas estiman el riesgo para cinco criterios diferentes (violencia autodirigida, violencia intrainstitucional, reincidencia violenta, riesgo de quebrantamiento de permisos o salidas de prisión, y reincidencia general). RisCanvi-S tiene dos niveles de riesgo –alto y bajo– que se obtienen mediante un procedimiento totalmente actuarial. Los factores de riesgo de la escala RisCanvi-S tienen dos categorías de respuesta (sí/no). «Sí» indica que el factor está claramente presente o hay indicios de que lo está, y «No» indica cuando el factor claramente no está presente. También se incluye una categoría para el caso en que no haya información sobre este factor (–no hay datos–), en cuyo caso se pide a los usuarios de la escala que tomen las medidas necesarias para encontrar la información que falta.

Si el RisCanvi-S da un riesgo bajo, en principio se cierra la evaluación y el sujeto será reevaluado cada seis meses. No obstante, se puede hacer una reevaluación antes si se produce un «evento crítico». El Manual de Aplicación de RisCanvi, publicado por la Generalitat de Catalunya (2019: 21), recoge como eventos especiales los siguientes:

Autolesión o intento de suicidio. Primera aparición de una conducta agresiva. Huelga de hambre. Recaída en el consumo de sustancias. Incumplimiento de una medida de supervisión comunitaria. Acontecimiento relevante en su entorno (separaciones, fallecimiento de familiares cercanos...). Cambio significativo de

actitud o comportamiento. Cambio de actitud en relación al tratamiento/intervención/equipo. Aparición de nueva información relevante.

Si el RisCanvi-S da un riesgo alto, la persona será entonces evaluada con el RisCanvi-C. La escala RisCanvi-S también puede recomendar una evaluación con la escala C cuando, a pesar de un riesgo bajo, existe lo que se denomina un «caso especial». Estos, según el propio Manual, son los siguientes:

- 1.- VIOFAM. Agresores domésticos violentos. Aquellos que han cometido delitos violentos físicos o sexuales, especialmente graves y continuados, contra miembros de la familia (consanguínea o adquirida), excluyendo a los agresores de pareja (o expareja).
- 2.- VIOPAR. Agresores de pareja. Los condenados por violencia física o sexual contra su pareja o expareja, convivieran o no en el momento de la agresión. También se incluirán los que realicen actos de violencia psicológica grave y continuada, especialmente amenazas y acoso (no sexual).
- 3.- VIOSEX. Agresores condenados por delitos sexuales violentos que preferentemente realizan su acción delictiva sobre personas que no son miembros de su familia, ni son su pareja o ex pareja sentimental, y por tanto que han actuado sobre conocidos o desconocidos y sin tener en cuenta el sexo o la edad de la víctima.
- 4.- VIOCRO. Agresores con un historial delictivo muy variado y crónico, con largos historiales y carreras delictivas, reincidencia y estancias en prisión.
- 5.- VIOTRAS. Agresores que han ejercido violencia física o sexual sobre cualquier tipo de víctima y que en el momento de cometer el delito o con anterioridad han padecido un trastorno mental grave o un trastorno de la personalidad (no psicopatía).
- 6.- PSICOPAT. Delincuentes de cualquier tipo, con o sin víctimas, en los que el inicio de su conducta delictiva es precoz y aumenta con la edad y que pueden ser susceptibles de tener una psicopatía (Generalitat de Catalunya, 2019: 12).

En el RisCanvi-C, existen tres niveles de riesgo: alto, medio y bajo. Los ítems o factores de riesgo de la escala RisCanvi-C tienen tres categorías de respuesta (sí/probable/no). «Sí» indica que el factor está claramente presente, «Probable» (o ?) se utiliza cuando el factor está probablemente presente, moderadamente presente o hay indicios de que está presente, «No» indica la ausencia del factor, es decir, cuando el factor no está presente. También se incluye una cuarta categoría para cuando no hay datos sobre este factor («sin datos» o S/D).

Una vez realizada la valoración con la escala RisCanvi-C, puede aparecer la recomendación de evaluar la reincidencia delictiva violenta a través de instrumentos más específicos de valoración del riesgo. Esta recomendación supone, como explican desde la Generalitat de Catalunya en el Manual que estamos citando, «que el sujeto por sus características personales y criminológicas pertenece a un grupo especial y, por tanto, el riesgo de cometer un determinado delito violento (sexual, contra la pareja, relacionado con trastorno mental, etc.) será valorado con mayor precisión mediante un instrumento específico para la evaluación de este tipo de conductas», (2019: 12). Y estos «grupos especiales», que pueden requerir la aplicación de estos instrumentos más específicos, son los mismos que acabamos de enumerar en el párrafo anterior. En función de estas tipologías delictivas, y de «otras razones criminológicas del interno» (2019: 12), las recomendaciones específicas de valoración que RisCanvi aconseja llevar a cabo para anticipar la probabilidad de un nuevo delito violento concreto, se realizan a través de las siguientes herramientas:

- HCR-20 *Historical Clinical Risk* – 20 (conjunto de directrices profesionales para la evaluación y gestión del riesgo de violencia, en línea: <<http://hcr-20.com/about/>>).
- SVR-20 *Sexual Violence Risk* – 20 (conjunto de directrices estructuradas de juicio profesional para realizar evaluaciones del riesgo de violencia sexual en contextos forenses penales y civiles, en línea: <<https://www.parrinc.com/Products/Pkey/4534>>).
- SARA Spousal Assault Risk Assessment Guide (Guía estructurada de 24 ítems para la evaluación del riesgo de violencia de pareja en personas sospechosas de maltrato de pareja o que reciben tratamiento por ello; en línea: <<https://www.rma.scot/wp-content/uploads/2023/01/Spousal-Assault-Risk-Assessment-Guide-Version-3-SARAv3.pdf>>).
- PCL-R Psychopathy Check List – Revised (Hare Psychopathy Checklist-Revised, o PCL-R, una herramienta de evaluación utilizada para evaluar los rasgos psicopáticos o antisociales de un individuo; en línea: <https://link.springer.com/referenceworkentry/10.1007/978-0-387-79948-3_837>).
- Cuando deban utilizarse estas escalas especializadas, la valoración será realizada por aquellos profesionales que estén acreditados para esta tarea en cada centro penitenciario por las autoridades competentes del Departament de Justícia, en función de su formación, experiencia y especialización profesional. No obstante, contarán con la colaboración del resto de los miembros del equipo para la recogida de información. Y el acto final de valoración del riesgo con estos instrumentos se realizará en las reuniones del equipo multidisciplinar (Generalitat de Catalunya, 2019: 17). En cualquier caso, parece que no siempre se siguen estas recomendaciones: al menos hasta 2016, y como se reconoce en el informe «Reincidencia en excarcelaciones de alto riesgo 2014-2016» (Capdevila, 2022). «En el momento en que se hicieron los protocolos de excarcelación de este estudio, el uso de estos instrumentos específicos era residual» (p. 67, también p. 129). Este informe incluye como propuesta «Incidir en la necesidad de utilizar de forma más frecuente y normalizada, para el colectivo de alto riesgo de reincidencia violenta, la evaluación mediante instrumentos específicos para la tipología violenta por la que estén condenados» (pp. 129 y 133).

Por último, hay que señalar que «excepcionalmente, en los casos de valoraciones de riesgo de internos con delitos violentos (contra las personas y contra la libertad sexual), el procedimiento permite la aplicación directa de la escala RisCanvi-C, es decir, no se aplicará previamente la escala RisCanvi-S». En estos casos, las reevaluaciones posteriores se harán también con la escala RisCanvi-C (Generalitat de Catalunya, 2019: 22).

En RisCavi-C los niveles de riesgo se calculan inicialmente de forma actuarial, pero luego, tras utilizar el método de juicio clínico estructurado, el evaluador decidirá qué nivel es el adecuado y puede desviarse del indicado por el algoritmo. Sin embargo, aunque RisCanvi-C admite una modificación en el nivel de riesgo si el validador no está de acuerdo con la puntuación actuarial, solo el 3,2 % de las evaluaciones han sido modificadas manualmente (Saura y Aragón, 2021). Según Bellio

(2021), «en el 95 % de los casos, si no más, el resultado del algoritmo coincide con el decidido por los validadores».

RisCanvi no se ha creado con aprendizaje automático ni lo utiliza.⁹ El sistema no tiene capacidad de aprendizaje autónomo, Sus diferentes escalas son una suma ponderada de los ítems puntuados. La fórmula para la reincidencia violenta puede encontrarse en el informe de Karimi-Haghighi y Castillo, aunque no estamos seguros de si ha sido revisada desde entonces.

Análisis cuantitativo de posibles sesgos

Para tratar de estudiar si pueden existir sesgos en RisCanvi, hemos intentado analizar si la herramienta arroja resultados diferentes para distintos colectivos, es decir, si tiende a asignar un mayor o menor nivel de riesgo a determinados colectivos y, en caso afirmativo, si estas diferencias se corresponden o no con una mayor reincidencia de dichos grupos.

Resultados por trastorno mental y consumo de sustancias

Tabla 16
Porcentaje de reclusos clasificados como de alto riesgo y porcentaje de reincidentes según trastorno mental y abuso de sustancias en la muestra de Karimi-Haghighi y Castillo

	<i>1. Ni TM ni abuso de sustancias</i>	<i>2. Solo TM</i>	<i>3. Solo abuso de sustancias</i>	<i>4. Tanto TM como abuso de sustancias</i>
% de reincidentes	3,4 %	7,3 %	8,7 %	8,0 %
% que RisCanvi clasifica como de alto riesgo	2,7 %	25,3 %	22,7 %	55,2 %
N	1,757	150	481	125

Fuente: tabla 3.6 en Karimi-Haghighi and Castillo, 2022.

⁹ Según Bellio (2021), «“la administración penitenciaria catalana pensó en actualizar RisCanvi con Inteligencia Artificial, pero decidió no hacerlo”, “Hace cinco años hicimos una petición para observar el resultado de introducir técnicas de aprendizaje automático y si esto mejoraría sustancialmente los resultados (...), El análisis de coste-beneficio no mostró una mejora significativa”, dijo el Sr, Camps».

Tabla 17
Porcentaje de internos clasificados como de alto riesgo y porcentaje de reincidentes por trastorno mental y abuso de sustancias en la muestra de Capdevila (2022)

	<i>1. No MHD or substance abuse</i>	<i>2. Only MHD</i>	<i>3. Only substance abuse</i>	<i>4. Both MHD and substance abuse</i>
% de reincidentes	7,16 %	13,13 %	11,69 %	15,27 %
% que RisCanvi clasifica como de alto riesgo	7,8 %	41,4 %	26,2 %	70,1 %
N	307	99	248	144

Fuente: tabla 55 en Briz-Redón y Montes (2022) a partir de los datos de la muestra de Capdevila.

En las tablas 16 y 17, el grupo «trastorno mental» (en adelante, TM) comprende a los sujetos que tienen un «sí» en uno o más de los siguientes ítems de RisCanvi: 32 (trastorno mental grave), 35 (trastorno de la personalidad relacionado con la ira, la impulsividad o la violencia), 39 (baja capacidad mental e inteligencia). El grupo «abuso de sustancias» comprende a las personas que tienen un «sí» en uno, o ambos, de los siguientes ítems: 30 (abuso o dependencia de drogas), 31 (abuso o dependencia del alcohol).

Vemos en las tablas anteriores que las personas con un trastorno mental grave, un trastorno de la personalidad o una discapacidad intelectual (TM) tienen un mayor riesgo de reincidencia según RisCanvi, al igual que las personas que abusan del alcohol o de otras drogas. En principio, esto es coherente con la constatación de que quienes se encuentran en alguna de estas situaciones reinciden más que quienes no lo están, y lo mismo ocurre con el abuso de sustancias. Sin embargo, se produce el siguiente efecto: los individuos con trastornos mentales (TM) reinciden casi el doble (7,3 %) que los que no padecen ni TM ni abuso de sustancias (3,45 %), y la reincidencia es ligeramente superior cuando los trastornos mentales se combinan con el abuso de sustancias (8 %). Pero el porcentaje de personas para las que RisCanvi predice un riesgo alto no se multiplica en la misma proporción, sino que es mucho mayor: es 8 veces mayor en el segundo grupo (25,3 %) que en el primero (2,7 %), y 18 veces mayor en el cuarto (55,2 %).

Esto tiene consecuencias en cuanto al tipo y número de errores generados por las predicciones. Como puede verse en la tabla 3.6 de Karimi-Haghighi y Castillo (2022), la sensibilidad (denominada en esta tabla «REVI alto %» entre los reincidentes) mejora a medida que aumenta el número de personas consideradas de alto riesgo de reincidencia en los distintos grupos; es decir, en el grupo sin TM ni abuso de sustancias, de los que efectivamente reincidieron, RisCanvi solo había clasificado correctamente como de alto riesgo al 13,6 %; en el segundo grupo (TM pero sin abuso de sustancias), al 36,4 %; en el tercer grupo (abuso de sustancias pero sin

TM), al 31,4 %; y en el último grupo (TM y abuso de sustancias), al 60 %. Sin embargo, el valor predictivo positivo empeora: en el primer grupo, de los que RisCanvi había clasificado como de alto riesgo, reincidió realmente el 16,7 %; en el segundo grupo, el 10,5 %; en el tercero, el 11,9 %; y en el último, solo el 8,7 %.

Si nos fijamos en los datos de la muestra de Capdevila (2022), los resultados son similares, e incluso más pronunciados: el porcentaje de personas que RisCanvi considera de riesgo alto se multiplica por diez en los casos en los que el sujeto padece un TM y además abusa de sustancias, mientras que la tasa de reincidencia solo se duplica en comparación con los que no presentan ninguna de estas dos circunstancias. En cuanto a la sensibilidad y al valor predictivo positivo, la sensibilidad aumenta del 18 % en el primer grupo al 77 % en el cuarto, pero el valor predictivo positivo se mantiene entre el 16 % y el 22 % en los cuatro grupos.

Los datos sugieren que RisCanvi tiende a sobrestimar el riesgo de reincidencia, en relación con el nivel real de reincidencia observado, entre las personas diagnosticadas con problemas de salud mental (tengan o no un historial de abuso de sustancias).

Podría concluirse que RisCanvi es extremadamente severo (incluso: ¿demasiado severo?) en la evaluación del riesgo de individuos con TM, y de aquellos con TM y abuso de sustancias. Cabe señalar que los trastornos mentales graves y los trastornos de la personalidad no se incluyen como ítems en la escala RisCanvi-C REVI (excepto la discapacidad intelectual, que sí se incluye) y tampoco se incluyen en RisCanvi-S, y sin embargo el porcentaje de personas que RisCanvi considera de alto riesgo para este grupo es mucho mayor. El abuso de sustancias sí puntúa tanto en la escala RisCanvi-C REVI como en el *Screening*.

En nuestra opinión se trata de un efecto preocupante, ya que son colectivos vulnerables (especialmente, los grupos 2 y 4), para los que la estancia en prisión puede resultar más perjudicial que para personas que no padecen trastornos. Desconocemos si el personal que trabaja con los internos es consciente del gran efecto que estas variables tienen en el nivel de riesgo global que RisCanvi asigna al sujeto, ya que –salvo error u omisión por nuestra parte– no hemos visto que se mencione esta cuestión en los trabajos e informes publicados hasta ahora sobre RisCanvi. Pero sería muy deseable que se conociera para evitar que se adopten decisiones restrictivas de derechos basadas en un elevado nivel de riesgo que es fundamentalmente consecuencia de estos factores, y que podrían acabar dando lugar a situaciones de discriminación de quienes requieren un mayor nivel de protección por parte de la Administración.

Resultados por nacionalidad

Tabla 18
Porcentaje de internos clasificados como de alto riesgo y porcentaje de reincidentes por origen nacional

	<i>Muestra de Karimi-Haghighi y Castillo</i>		<i>Muestra de Capdevila</i>	
	<i>Espanoles</i>	<i>Extranjeros</i>	<i>Espanoles</i>	<i>Extranjeros</i>
% de reincidentes	5,2 %	4,3 %	13,23 %	7,12 %
% que RisCanvi clasifica como de alto riesgo	13,5 %	5,8 %	38,2 %	24,1 %

Fuente: tabla 3.2 en Karimi-Haghighi y Castillo (2022) y tabla 55 en Briz-Redón y Montes (2022).

La escala RisCanvi REVI clasifica como de riesgo alto a un mayor porcentaje de españoles (13,5 %) que de extranjeros (5,8 %), y la tasa de reincidencia violenta es mayor para los españoles (5,2 %) que para los extranjeros (4,3 %). De hecho, ser extranjero es un factor que computa -2 en la escala de riesgo de reincidencia violenta. Estos datos serían coherentes con los publicados por Castillo / Karimi Haghighi en un trabajo anterior (2021), realizado con una muestra diferente, y también con los del estudio «Tasa de reincidencia penitenciaria 2014», en el que si bien el porcentaje de reincidencia total era mayor para los extranjeros (27,1 %) que para los españoles (21,1 %), los extranjeros reinciden más en delitos contra la propiedad y relacionados con las drogas, pero menos cuando se trata de delitos contra las personas físicas (Capdevila et al., 2015: 197 y 199). Aunque los «delitos violentos» no son exactamente lo mismo que los «delitos contra las personas físicas», ambas categorías se solapan en gran medida. Por otro lado, nuestros resultados también coincidirían con los obtenidos en este punto en la auditoría externa encargada por la Generalitat sobre RisCanvi, que también analiza posibles sesgos de la herramienta respecto a determinados colectivos protegidos, entre ellos los extranjeros. El informe de la auditoría utiliza métricas diferentes a las que nosotros hemos aplicado para analizar este aspecto (*error-rate balance*), pero llega a conclusiones similares, pues encuentran que la escala de reincidencia violenta etiqueta a los extranjeros con menor frecuencia que a los españoles como falsos positivos, y con mayor frecuencia que a los españoles como falsos negativos (Herron Mulet et al., 2024: 37, 73 y 75).

Cabe decir, por tanto, que en este punto en principio RisCanvi acierta, porque asigna un menor nivel de riesgo al grupo que realmente muestra una menor tasa de reincidencia violenta penitenciaria. Además, no se trata de un efecto casual o indirecto, sino que la herramienta está deliberadamente diseñada para conseguirlo, ya que ser extranjero computa -2 en la escala REVI; suponemos que esto se ha hecho a propósito, para tener en cuenta que los extranjeros tienen una menor tasa

de reincidencia violenta.¹⁰ En otras palabras, esta característica de RisCanvi debe valorarse positivamente, ya que la escala REVI de RisCanvi no parece discriminar ni directa ni indirectamente a los extranjeros, y tal vez en este sentido funcione mejor que los humanos evaluando el riesgo, donde la probabilidad de un sesgo xenófobo o racista nunca puede excluirse por completo (aunque carecemos de datos para confirmar esta hipótesis, ya que no disponemos de una muestra de sujetos cuyo riesgo se hubiera evaluado sin utilizar RisCanvi).

Sin embargo, a pesar de lo dicho anteriormente, cabe también señalar que la diferencia en las tasas de reincidencia en los datos de Karimi-Haghighi y Castillo es pequeña (5,2 % nacionales, 4,3 % extranjeros), pero la diferencia en el porcentaje de cada uno de los grupos que es clasificado por RisCanvi como de riesgo alto es mucho mayor (13,5 % nacionales, 5,8 % extranjeros). Teniendo esto en cuenta, quizás podría considerarse que RisCanvi es excesivamente «duro» con los nacionales, siguiendo el mismo razonamiento que hemos hecho respecto al consumo de drogas y los trastornos mentales.

Esta crítica, sin embargo, sería quizás prematura. Por un lado, porque en la otra muestra de datos de la que disponemos, el estudio de Capdevila (2022), el desequilibrio entre ambos grupos desaparece: la diferencia en las tasas de reincidencia es mucho mayor (13,23 % nacionales, 7,12 % extranjeros), y entre los nacionales, el porcentaje clasificado como de alto riesgo es del 38,2 %, y entre los extranjeros, del 24,1 %. Es decir, aquí la tasa de reincidencia de los nacionales es casi el doble que la de los extranjeros, pero el porcentaje que RisCanvi clasifica como de alto riesgo entre los nacionales es menos del doble que entre los extranjeros. Pero la muestra del estudio de Capdevila es menos representativa de la realidad, porque está construida *ad hoc* con tres grupos más o menos igual de grandes de personas de riesgo alto, medio y bajo, mientras que en las prisiones la mayoría de los internos son clasificados como de riesgo bajo (véase la tabla de la p. 3 de Karimi-Haghighi y Castillo (2022), cuya muestra es la totalidad de internos que salieron en libertad definitiva o provisional en el año analizado).

Es importante destacar que todo lo que hemos expuesto se refiere únicamente a la reincidencia penitenciaria violenta, es decir, a la comisión de un nuevo delito violento que ha dado lugar a un nuevo encarcelamiento. Desconocemos si RisCanvi discrimina o no a los extranjeros cuando evalúa el riesgo de reincidencia general, o el riesgo de ser condenado a una sanción penal que no conlleve pena de prisión, porque no hemos tenido acceso a muestras de datos en las que se hayan estimado estos tipos de riesgo y se hayan comparado con las tasas reales de reincidencia.¹¹

¹⁰ Desconocemos cómo se tiene en cuenta la nacionalidad en la escala de reincidencia general, ya que no hemos tenido acceso al diseño de esta escala.

¹¹ Hemos tenido acceso a este tipo de información (datos de reincidencia general, y datos de reincidencia con sanciones no privativas de libertad) en la muestra de Capdevila, pero –en nuestra opinión– no tiene sentido extraer conclusiones sobre la precisión o los sesgos para estos tipos de reincidencia a partir de los datos de esta muestra, porque: 1) esta muestra no es representativa de la población liberada cada año, sino que tiene una prevalencia muy alta de delincuentes de alto riesgo; y 2) esta muestra fue evaluada con la escala REVI, no con la nueva escala de reincidencia general, por lo que no consideramos adecuado juzgar el rendimiento

En definitiva, pensamos que RisCanvi no discrimina ni directa ni indirectamente a los extranjeros a la hora de evaluar el riesgo de reincidencia violenta penitenciaria, lo que probablemente sea una virtud de esta herramienta en comparación con las evaluaciones de riesgo que podrían hacer los individuos. Existe una tendencia a sobreestimar el riesgo de los españoles, aunque sólo la hemos observado en una de las muestras.

Resultados por estatus socioeconómico

RisCanvi contiene varios ítems que están o pueden estar vinculados a un estatus socioeconómico bajo o incluso a la marginalidad social. Hemos intentado analizar si podría haber algún efecto discriminatorio relacionado con este aspecto, es decir, si estos factores están asociados a aumentos del nivel de riesgo que no se corresponden con aumentos correlativos del nivel de reincidencia observada.

Tabla 19
Porcentaje de internos clasificados como de alto riesgo y porcentaje de reincidentes por estatus socioeconómico

	<i>Muestra de Karimi-Haghighi y Castillo</i>					<i>Muestra de Capdevila</i>				
Grupo	1	2	3	4	5	1	2	3	4	5
	Ni 19 ni 20	Solo 19 o 20	19 y 20	19 y 20 + otro de los otros 4	19 y 20 + todos los demás	Ni 19 ni 20	Solo 19 o 20	19 y 20	19 y 20 + otro de los otros 4	19 y 20 + todos los demás
% de reincidencia	3,7 %	6,6 %	9,34 %	9,25 %	10,5 %	7,2 %	12,6 %	12,7 %	12,9 %	7,7 %
% clasificado de alto riesgo	4,3 %	16,6 %	36,6 %	36,5 %	73,7 %	14,5 %	36,4 %	55,5 %	54,4 %	61,5 %
N	1.575	590	257	238	19	373	302	236	217	13

Fuente: tabla 3.7 en Karimi-Haghighi y Castillo (2022) y tabla 55 en Briz-Redón y Montes (2022).

El análisis no fue fácil, sobre todo porque no es fácil asociar con certeza determinados factores de riesgo con el estatus socioeconómico.

En un primer paso, nos fijamos en las variables 19 («problemas relacionados con el empleo») y 20 («falta de recursos económicos») del RisCanvi-C, y dividimos la muestra en tres grupos: sujetos con ninguno de los dos factores (grupo 1), sujetos con solo uno de ellos (grupo 2) y sujetos con ambos (grupo 3). Una vez hecho esto,

de RisCanvi en la predicción de la reincidencia general con niveles de riesgo derivados de una escala que fue diseñada para estimar el riesgo de reincidencia violenta.

intentamos ver qué ocurría si, además de las variables más directamente relacionadas con cuestiones económicas, añadíamos otras que pensábamos que podían indicar problemas sociales o de marginación.

En los grupos de la muestra de Karimi-Haghighi y Castillo, los resultados son los siguientes: la tasa de reincidencia de los que no tienen ninguna de las dos variables relacionadas con problemas económicos (ni la 19 ni la 20) es del 3,7 %; la de los que solo tienen una de las dos, del 6,6 %; y la de los que tienen las dos, del 9,34 %. Y el porcentaje de personas que RisCanvi considera de riesgo alto en cada uno de estos grupos es del 4,3 % en el primero, del 16,6 % en el segundo y del 36,6 % en el tercero.

En los datos de Capdevila, los resultados son muy similares: el porcentaje de reincidentes aumenta a medida que empeora la situación económica y laboral de las personas: 7,2 %, 12,6 % y 12,7 % para cada uno de los 3 grupos, respectivamente. Pero el porcentaje de personas dentro de cada grupo que RisCanvi clasifica como de riesgo alto aumenta en mucha mayor medida: 14,5, 36,4 y 55,5 %, respectivamente.

En otras palabras, en ambos grupos los datos se comportan de forma similar: la tasa de personas que RisCanvi considera de riesgo alto aumenta de forma desproporcionada con respecto al aumento de la reincidencia violenta real a medida que empeora la situación económica y laboral.

En consecuencia, podríamos hacer aquí una crítica similar a la que hemos hecho con respecto a las variables de trastorno mental y abuso de sustancias: RisCanvi perjudica a quienes tienen una situación económica precaria, ya que les asigna una puntuación de riesgo mucho más alta que a quienes no la tienen, mientras que las diferencias en las tasas de reincidencia, aunque existen, son mucho menores (algo que es especialmente evidente en el tercer grupo –personas con problemas de empleo y falta de recursos económicos–, en comparación con el primero).

Una vez visto cómo influyen los problemas económicos en el nivel de riesgo, hemos intentado ver qué ocurre si a los problemas económicos se añaden los sociales. La tarea no es fácil, porque es difícil saber qué combinación de los factores contenidos en RisCanvi podría reflejar mejor esta realidad (por ejemplo, hay variables como la inadaptación infantil, las amistades delictivas o la pertenencia a grupos de riesgo que quizá también podrían considerarse aquí), y porque no hemos tenido acceso directo a una muestra de todos los excarcelados en un periodo de tiempo determinado, y al depender de los cálculos que otro equipo de investigación nos ha hecho el favor de realizar para nosotros, el número de combinaciones posibles que nos ha parecido razonable pedir era limitado (la muestra de Capdevila (2022), sobre la que sí podíamos realizar análisis directamente, tiene algunas peculiaridades que, como ya hemos dicho, no la hacen adecuada). Sea como fuere, hemos construido un grupo (el número 4) formado por aquellas personas en las que concurren las variables 19 y 20, así como una o más de las cuatro siguientes: nivel educativo no superior a primaria (ítem 18 en RisCanvi); antecedentes delictivos en la familia de origen (ítem 22); socialización o crianza problemática en la familia de origen (ítem 23); y falta de apoyo familiar y social (ítem 24). Y lo que hemos observado es que la adición de estas variables, al menos en esta configuración (una o más), no produce cambios ni en el nivel de reincidencia ni en el porcentaje de personas a las que se les asigna alto riesgo, porque los resultados son muy similares a los del grupo inmediatamente anterior: la tasa de reincidencia es muy parecida, incluso algo menor (9,25 %), y el

porcentaje que RisCanvi clasifica como de alto riesgo aumenta un poco, pero no mucho (36,5 %). El mismo patrón se observa en los datos de la muestra de Capdevila: la tasa de reincidencia es del 12,9 %, y el porcentaje de personas de este grupo clasificadas como de alto riesgo es el 54,4 %.

Donde sí aparece una diferencia notable es cuando se acumulan todas estas variables, es decir, en el grupo en el que concurren acumulativamente todas estas situaciones, tanto problemas económicos como sociales (Grupo 5 en la tabla anterior; Grupo Socioecon3bis en la tabla 3.7 de Karimi-Haghighi y Castillo (2022)): en este caso, el 73,7 % del grupo es considerado de alto riesgo, a pesar de que la reincidencia es únicamente del 10,5 %, es decir, solo un punto porcentual más que en el grupo anterior (Grupo 4: 9,25 %), lo que significaría que RisCanvi asigna a estas personas un riesgo totalmente desproporcionado con respecto a su tasa de reincidencia real.

También en este caso los datos de Capdevila arrojan resultados similares a los de Karimi-Haghighi y Castillo: el Grupo 5 de la tabla anterior (Socioecon3 en la tabla 51 en Briz-Redón y Montes (2022)) son solo 13 personas, y llama la atención que la tasa de reincidencia (7,7 %; $n = 1$) sea inferior a la de los que solo tienen una de las variables relacionadas con problemas económicos, y también inferior a la de los que tienen ambas. Sin embargo, dentro de estos 13 individuos RisCanvi considera de alto riesgo al 61,5 % ($n = 8$). Además, la única persona registrada como reincidente violento en prisión no pertenecía al grupo de alto riesgo, por lo que la sensibilidad aquí es 0.

Ahora bien, es necesario ser muy cauteloso con estos datos, porque, aunque ambos estudios sugieren que RisCanvi puede sobrestimar enormemente el riesgo de las personas con dificultades socioeconómicas especialmente intensas, creemos que los grupos son demasiado pequeños ($n = 19$; $n = 13$) para poder extraer conclusiones sólidas.

Un último punto que probablemente merezca la pena mencionar en relación con estas variables es que, aunque todos estos ítems parecen influir en el nivel de riesgo asignado por RisCanvi, no todos ellos son factores de riesgo que puntúen en la escala REVI:

Tabla 20
Algunos de los factores de riesgo de la escala de riesgo de
reincidencia violenta en RisCanvi

18 (nivel educativo no superior a la enseñanza primaria)	-3
19 (problemas relacionados con el empleo)	No considerado
20 (falta de recursos económicos)	3
22 (antecedentes penales en la familia de origen)	2
23 (socialización o educación problemática en la familia de origen)	No considerado
24 (falta de apoyo familiar y social)	No considerado

Fuente: Karimi-Haghighi y Castillo (2022).

Comparación general entre los datos de ambas muestras

Aunque en términos generales los resultados del análisis que hemos realizado tienden a ser similares en ambas muestras, existen diferencias entre ellas: por ejemplo, la sensibilidad y el valor predictivo positivo tienden a ser –en general– mayores en los distintos grupos analizados en el estudio de Capdevila, 2022 (véanse las tablas 52 y 55 en Briz-Redón y Montes (2022)). Creemos que esto puede atribuirse, al menos en parte, a la diferente composición de las muestras. En la muestra de Karimi-Haghighi y Castillo, que comprende a todas las personas excarceladas en un año (2016) a libertad definitiva o condicional, la tasa de reincidencia penitenciaria violenta es solo del 4,9 %, y el grupo de alto riesgo es solo el 10,5 % de la muestra. En Capdevila (2022), por el contrario, la tasa de reincidencia es del 10,91 % (más del doble), probablemente porque el grupo de alto riesgo es un tercio de la muestra y, por tanto, está sobrerrepresentado en comparación con lo que debe ser la realidad de las personas excarceladas.

Relevancia de la evaluación de riesgos para la toma de decisiones (análisis cualitativo)

Análisis de las resoluciones judiciales que revisan las decisiones tomadas por el personal penitenciario utilizando las valoraciones de riesgo de RisCanvi

Para valorar hasta qué punto el nivel de riesgo que estima la herramienta es determinante en la toma de decisiones, hubiera sido ideal poder entrevistar a personal penitenciario. Sin embargo, como hemos explicado anteriormente, la petición que hicimos al Departamento de Justicia en este sentido no recibió respuesta.

Sin embargo, muchas de las decisiones adoptadas por el personal penitenciario en relación con los internos pueden ser recurridas por estos ante los tribunales, concretamente ante el juez de vigilancia penitenciaria (JVP). Examinar las resoluciones de estos jueces nos permitiría dos cosas: por un lado, vislumbrar, aunque fuera de forma indirecta, el peso que el personal penitenciario otorga al nivel de riesgo de RisCanvi, si el JVP se refiere o incluso reproduce en parte la documentación o los argumentos aportados por el personal penitenciario para justificar su propia decisión. Y, en segundo lugar, el examen de estas sentencias también nos permitiría ver qué relevancia dan los jueces a los niveles de riesgo de RisCanvi.

Pero, lamentablemente, no hemos podido examinar directamente estas sentencias porque en España solo se publican en bases de datos de acceso abierto y anonimizadas las sentencias judiciales dictadas por órganos colegiados, no las dictadas por órganos unipersonales como el JVP. Ahora bien, dado que las sentencias del JVP son a su vez recurribles ante las audiencias provinciales, que son órganos colegiados y cuyas resoluciones por tanto sí son accesibles, hemos examinado estas últimas.

Se han analizado un total de 540 resoluciones correspondientes desde 2015. Aunque RisCanvi comenzó a funcionar en 2009, hasta 2015 no aparecen resoluciones judiciales referidas a RisCanvi en la base de datos que hemos utilizado.

Utilizamos la base de datos de jurisprudencia de Aranzadi e introducimos la palabra clave *RisCanvi*. Centramos la búsqueda en las resoluciones dictadas por las cuatro audiencias provinciales de Cataluña porque son las que reciben a diario las previsiones de *RisCanvi* a la hora de tomar decisiones sobre salidas temporales de internos o sobre otro tipo de medidas.

Tabla 21
Número de resoluciones sobre *RisCanvi* dictadas por las cuatro audiencias provinciales de Cataluña encontradas en la base de datos Aranzadi para los años 2010-2022

	<i>Número de resoluciones judiciales que contienen la palabra «RisCanvi»</i>	<i>Número de prisiones</i>
Barcelona	2.657	8
Girona	309	2
Tarragona	141	2
Lleida	12	2

Fuente: elaboración propia a partir de la información contenida en la base de datos Aranzadi.

Llama la atención el número de resoluciones dictadas por las audiencias provinciales de Girona y Lleida. El número en Girona es especialmente alto, teniendo en cuenta que solo hay dos prisiones en esta provincia. Esto puede deberse a lo crítico que es este tribunal en particular con *RisCanvi*.

No hemos podido analizar las 3.119 resoluciones, sino solo una muestra:

- 200 resoluciones emitidas entre los años 2015 y 2019.
- 200 resoluciones emitidas en 2020.
- 70 resoluciones emitidas en 2021.
- 70 resoluciones emitidas en 2022.

Las resoluciones analizadas para el periodo comprendido entre 2010 y 2017 suponen el 100 % de las resoluciones emitidas en dicho lapso temporal (1 en 2010; 13 en 2011; 15 en 2012; 1 en 2015; 13 en 2016; y 32 en 2017). Las resoluciones analizadas para 2018 fueron un total de 63, que suponen un 19 % del total de resoluciones emitidas ese año. Las resoluciones analizadas para 2019 fueron un total de 62, que suponen un 9 % del total de resoluciones emitidas ese año. Las 200 resoluciones que se analizaron para 2020 suponen el 28 % de las dictadas ese año. Las 70 resoluciones analizadas para 2021 suponen el 5 % de las dictadas ese año. Las 70 resoluciones analizadas para 2022 suponen el 9 % de las dictadas ese año. Para los años 2018-2022, la selección se realizó utilizando un criterio temporal, por lo que el análisis se centró en las 63 primeras resoluciones de 2018; las 62 primeras dictadas en 2019; las 200 primeras de 2020; y las 70 primeras de 2021 y 2022, respectivamente, siendo 2022 el año en el que concluyó nuestro análisis.

Las conclusiones obtenidas tras el estudio de estas resoluciones son las siguientes.

A) Conclusiones generales

- RisCanvi no es, en términos generales, el elemento determinante utilizado por los jueces a la hora de conceder o denegar excarcelaciones temporales u otras medidas. En la mayoría de las resoluciones, RisCanvi no es más que uno de los diferentes factores que se tienen en cuenta al conceder o denegar permisos u otras medidas que benefician a los presos. En algunos casos, a pesar de la escasa relevancia otorgada a RisCanvi, se intenta justificar por qué no se han tenido en cuenta las indicaciones del algoritmo.
- Del mismo modo, las juntas de tratamiento de los centros penitenciarios no parecen tomar RisCanvi como un factor determinante a la hora de formular recomendaciones. Sí mencionan el nivel de riesgo pronosticado por la herramienta, pero lo consideran junto a otros muchos elementos que concurren en el caso de cada recluso.
- Son muy pocas las resoluciones en las que RisCanvi es el principal factor tenido en cuenta por los tribunales para denegar permisos u otras medidas cuando RisCanvi indica un riesgo medio/alto en los diferentes ítems que toma en consideración.
- Pero la Fiscalía utiliza sistemáticamente el riesgo indicado por RisCanvi como argumento para recurrir la concesión de libertades provisionales u otras medidas cuando es medio o alto.

B) Decisiones que tienden a seguir las predicciones de RisCanvi

- Hay un número relevante de resoluciones que tienen en cuenta RisCanvi cuando predice un riesgo bajo. En estos casos se suele conceder la libertad provisional u otra medida. Estas resoluciones han sido dictadas principalmente por la Audiencia Provincial de Barcelona. De hecho, hasta el año 2018, exactamente el mismo argumento por el que una predicción de riesgo bajo de reincidencia debería conllevar automáticamente la concesión de la libertad provisional, fue utilizado en 68 resoluciones diferentes dictadas por esta Audiencia.

C) Resoluciones críticas con RisCanvi (un número relevante). A continuación, resumimos las principales razones por las que los tribunales son críticos con RisCanvi:

- RisCanvi tiene en cuenta muchos elementos del pasado delictivo del recluso que no están actualizados. Por eso hay que tener en cuenta la situación actual del interno.
- RisCanvi considera factores estáticos del pasado del interno que este no puede modificar (gravedad del delito cometido, otros delitos por los que fue condenado, circunstancias relativas a su infancia, familia, etc.). Estos factores estáticos siguen afectando negativamente a los internos, ya que aumentan el riesgo indicado por el algoritmo, aunque la persona haya mejorado desde entonces su comportamiento o sus circunstancias actuales. Esto debe tenerse en cuenta a la hora de conceder o denegar las excarcelaciones temporales.

- RisCanvi es una herramienta automática de procesamiento de datos, mientras que la Junta de Tratamiento del Centro Penitenciario está formada por un grupo de personas que trabajan diariamente con el interno, por lo que las predicciones de RisCanvi no pueden prevalecer sobre la opinión profesional de la Junta. En otras palabras, cuando la Junta de Tratamiento propone un permiso y RisCanvi muestra un riesgo medio/alto, se suele tener en cuenta y seguir la opinión de la Junta.
- La Audiencia Provincial de Girona es, sin duda, la más crítica con la herramienta. En general, este tribunal es muy pro-reo. Es decir, en la mayoría de las ocasiones concede la libertad temporal de los internos, incluso en algunos casos en los que la Junta de Tratamiento no lo recomienda.
- Algunas de las resoluciones más recientes dictadas por estos tribunales han destacado que los internos deben ser evaluados periódicamente con RisCanvi. Aunque estas evaluaciones deberían tener lugar cada seis meses, no siempre es así. En particular, nos parece bastante problemático que la Audiencia Provincial de Tarragona haya denegado, al menos en dos casos, la concesión de excarcelaciones temporales de internos por no haberse realizado una nueva valoración del RisCanvi después de que el interno hubiera sido sometido a un programa de reinserción, por lo que el Tribunal consideró que no disponía de información suficiente para conceder dicha excarcelación ya que la última valoración con RisCanvi había estimado que el interno presentaba un alto riesgo de reincidencia. El hecho de que el personal penitenciario no haya evaluado al interno en el plazo establecido no debería servir, a nuestro juicio, como argumento para restringir su libertad, pues es una circunstancia de la que éste no es responsable.

Entrevista con un magistrado de la Audiencia Provincial de Girona

La Audiencia Provincial de Girona es uno de los cuatro tribunales de Cataluña competentes para conocer de los recursos presentados contra las resoluciones de los juzgados de vigilancia penitenciaria, por lo que es competente para conocer de los casos en los que se tenga en cuenta RisCanvi a la hora de conceder permisos penitenciarios o el tercer grado a favor de los internos.

De la entrevista con el magistrado de la Audiencia Provincial de Girona se pueden extraer las siguientes conclusiones:

En primer lugar, RisCanvi no suele ser el elemento determinante a la hora de conceder o denegar una excarcelación o un tercer grado. Teniendo en cuenta que los tribunales disponen de una importante cantidad de información sobre el interno, el resultado de RisCanvi es solo un factor más en la decisión judicial. Muchas veces, a pesar de que el sistema RisCanvi indique un riesgo alto en alguna de las escalas, el juzgado acaba concediendo el permiso si existen otros elementos o factores que contradicen las indicaciones del algoritmo. Los informes de la Junta de Tratamiento, que suelen ser bastante extensos e incluyen informes escritos del psicólogo, el trabajador social, el jurista y el educador, suelen tenerse muy en cuenta.

En segundo lugar, algunos jueces desconocen cómo funciona RisCanvi y las razones por las que el sistema da uno u otro nivel de riesgo. En los documentos e

información que los jueces reciben sobre RisCanvi suele haber poca información sobre las razones a las que obedecen sus resultados. Esto significa que las estimaciones que realiza RisCanvi a veces no se tienen en cuenta por falta de conocimiento. Este juez mencionó, y lo hemos visto en muchas sentencias, que los informes de la Junta de Tratamiento suelen incluir los cuatro niveles de riesgo de RisCanvi (por violencia autodirigida, por violencia intrainstitucional, por reincidencia violenta y por quebrantamiento de permisos o salidas), aunque la decisión que se ha de revisar en el caso concreto sea quizás sólo la concesión de una libertad temporal. Los jueces no acaban de entender cómo es posible que para un mismo preso uno de estos riesgos pueda ser bajo, otro medio y quizás otro alto, por lo que tienden a dar menos importancia a lo que no acaban de entender.

En tercer lugar, como regla general, el Ministerio Fiscal recurre las sentencias dictadas por los juzgados de vigilancia penitenciaria que conceden la libertad provisional o el tercer grado cuando RisCanvi indica un riesgo alto en alguna de las cuatro escalas.

Queremos señalar que esta opinión es la de un juez que forma parte de un tribunal especialmente crítico con las previsiones de RisCanvi. Por lo tanto, es posible que otros jueces no la compartan. Sin embargo, las opiniones expresadas por este juez coinciden en gran medida con el estudio de las sentencias de las demás audiencias provinciales que hemos analizado anteriormente.

Entrevista con abogados de Cataluña

Hemos entrevistado a tres abogados de Barcelona. Cada uno de ellos trabaja en despachos diferentes, y uno de ellos ha sido el presidente de la sección de Derecho Penitenciario del Colegio de Abogados de Barcelona. Obviamente, la opinión y la experiencia personales de tres individuos no pueden considerarse representativas de la opinión de todos los abogados que trabajan en Cataluña. Pero al menos los tres con los que hemos hablado tienen amplia experiencia en derecho penitenciario y con RisCanvi.

Todos los abogados coinciden en que tanto los jueces de vigilancia penitenciaria como el Ministerio Fiscal dan mucha importancia al nivel de riesgo que indica RisCanvi. En su opinión, todos estos agentes judiciales y trabajadores públicos suelen tener más en cuenta RisCanvi cuando este indica niveles medios o altos en las diferentes escalas que cuando predice niveles bajos de riesgo. Conviene señalar que esta afirmación difiere significativamente de las conclusiones extraídas de la entrevista con el juez de la Audiencia Provincial de Girona y de las resoluciones judiciales que hemos analizado. Sin embargo, creemos que esta diferencia puede explicarse probablemente por el hecho de que los abogados tratan principalmente casos en Juzgados de Vigilancia Penitenciaria, a cuyas resoluciones no hemos podido acceder, como se ha explicado anteriormente.

En lo que sí coinciden los abogados con el magistrado es en que la Fiscalía recurre sistemáticamente todas las salidas provisionales y los terceros grados cuando el sistema RisCanvi arroja una puntuación media o alta en alguna de las escalas.

En segundo lugar, los abogados destacan la dificultad de impugnar ante los tribunales los resultados indicados por RisCanvi. En muchos casos no es fácil

conocer las razones por las que el sistema ha pronosticado un concreto nivel de riesgo en una escala determinada, por lo que consideran que sus clientes se encuentran en una situación de indefensión considerable. Comentaron que si el abogado es persistente e insiste en hablar con el personal penitenciario (normalmente, el director de la prisión), a veces consigue una explicación sobre qué factores están impidiendo que el nivel de riesgo disminuya. También comentaron que saben que, en algunos casos, cuando el interno es educado, cortés, conoce las normas y la ley, se comporta correctamente en prisión y pide con educación explicaciones sobre lo que podría hacer para mejorar los resultados de la evaluación, el personal penitenciario sí le da explicaciones sobre la puntuación de los ítems en su RisCanvi. Pero los abogados no tienen constancia de que se informe de forma rutinaria a los presos sobre sus puntuaciones en RisCanvi ni sobre los motivos de estas.

En nuestra opinión, esto es preocupante, si tenemos en cuenta que el «interno educado» no es el perfil del interno «tipo». Una de las ventajas que supuestamente tiene RisCanvi es que trata a todos los presos por igual, estandarizando las evaluaciones y las decisiones, y evita que las preferencias personales de los trabajadores penitenciarios interfieran en el trato y la respuesta que reciben los presos. Pero, por lo que cuentan los abogados, parece que no siempre es así. Los presos menos informados, los que no pueden permitirse pagar a un abogado que insista en hablar con el personal penitenciario, los que no tienen el nivel educativo y cultural necesario para estar informados sobre sus derechos y saber cómo comunicarse eficazmente con el personal penitenciario, no se beneficiarán de estas ocasionales explicaciones sobre sus niveles de riesgo ni del asesoramiento de los trabajadores penitenciarios.

Además, los abogados creen que en muchos casos RisCanvi es utilizado por la Junta de Tratamiento para ratificar o confirmar las opiniones que ya se ha formado sobre un preso concreto. Cuando la Junta considera que la persona está preparada para ser clasificada en un régimen penitenciario menos restrictivo, o para que se le conceda un permiso penitenciario, revisa y actualiza sus valoraciones de RisCanvi, de forma que no es un cambio en el nivel de riesgo el responsable de un cambio en las condiciones o el régimen aplicable a los internos, sino que primero la Junta considera que es el momento de cambiar las condiciones, y esto es lo que lleva a una revisión de RisCanvi y a un cambio en los niveles de riesgo.

Los abogados también denuncian que, dado que en RisCanvi hay muchas variables que tienen en cuenta factores estáticos del pasado de los internos, en muchos casos es muy difícil rebajar el riesgo indicado por el sistema.

Por último, todos ellos coinciden en que la implantación de RisCanvi en todo el proceso penitenciario y judicial ha sido perjudicial. Según sus palabras, mientras que antes los abogados tenían cierto margen de maniobra para impugnar los dictámenes de la junta de tratamiento o de los jueces, ahora es muy difícil cuestionar el nivel de riesgo indicado por RisCanvi.

3. EFICACIA DEL ENFOQUE POLICIAL PREDICTIVO PARA PREVENIR LA DELINCUENCIA

Como también se ha sugerido en estudios anteriores (Capdevila, 2015). RisCanvi generalmente diferencia bien el riesgo relativo de los distintos grupos: según Karimi-Haghighi y Castillo (2022), en su muestra completa la tasa de reincidencia es mayor en los grupos a los que RisCanvi asigna una mayor probabilidad de riesgo, teniendo el grupo categorizado como de riesgo alto una tasa de reincidencia cuatro veces mayor que el grupo categorizado como de bajo riesgo, y esta es una propiedad deseable. Sin embargo, esto no siempre se cumple cuando se analiza cómo se comporta el riesgo en algunos subgrupos: por ejemplo, entre los extranjeros, el grupo de alto riesgo tiene una tasa de reincidencia menor que el grupo de riesgo medio (tabla 3.2). Y lo mismo ocurre con algunos otros de los subgrupos de las tablas 3.4., 3.5, 3.6 y 3.7 (Karimi Haghighi y Castillo (2022)).

Dicho esto, RisCanvi no es un instrumento bien calibrado: según los resultados de Karimi-Haghighi y Castillo (2022), el porcentaje de sujetos clasificados como de alto riesgo en la muestra total (10,5 %) duplica el de los que realmente cometieron un nuevo delito violento en el periodo de seguimiento (4,9 %). Esta proporción se mantiene similar para muchos de los subgrupos analizados. Esto significa que RisCanvi asigna sistemáticamente más riesgo del que realmente existe en los diferentes subgrupos y en la muestra total.

El primer informe público que evaluó la eficacia para medir el riesgo de reincidencia violenta de RisCanvi fue realizado por el Centro de Estudios Jurídicos de la Generalitat en 2015 (Capdevila Capdevila, 2015: 151), con una muestra relativamente pequeña ($N = 648$) debido a que RisCanvi aún no se había implantado totalmente en el sistema penitenciario catalán y todavía había pocos internos con evaluaciones RisCanvi. Los resultados fueron los siguientes: entre los que reincidieron tras salir de prisión, RisCanvi había conseguido identificar al 77,15 % de ellos como de riesgo alto o medio (sensibilidad 77,15 %); entre los que no reincidieron tras salir de prisión, RisCanvi había conseguido identificar al 57,26 % de ellos como de riesgo bajo (especificidad 57,26 %); entre aquellos a los que RisCanvi había asignado un riesgo alto o medio de reincidencia violenta, el porcentaje que efectivamente reincidió con un delito violento fue (valor predictivo positivo) del 17,94 %; y entre los que RisCanvi había etiquetado como de riesgo bajo, el porcentaje que no reincidió (valor predictivo negativo) fue del 95,4 % (el valor predictivo positivo no está mencionado expresamente en Capdevila Capdevila 2015, pero se puede calcular a partir de los datos allí aportados; véase Martínez Garay, 2016). Es decir, cuando RisCanvi dice que la persona es de riesgo bajo, hay un 95,4 % de probabilidad de que no cometa un nuevo delito violento, pero cuando RisCanvi dice que la persona es de riesgo alto o medio, la probabilidad de que cometa un nuevo delito violento es del 17,94 %.

El estudio de Capdevila et al. (2022) también analiza el nivel de acierto en las valoraciones de riesgo realizadas con RisCanvi, y concluye que no es alto, conclusión con la que estamos de acuerdo. Este informe analiza la reincidencia de un grupo de presos de alto riesgo ($N = 352$) que salieron de prisión entre 2014-2016 y fueron seguidos hasta el 30 de septiembre de 2019 (mínimo 2,7 años, máximo 5,7). El estudio informa sobre la capacidad predictiva de RisCanvi para 4 tipos de reincidencia:

reincidencia penitenciaria violenta (nuevo ingreso en el sistema penitenciario, ya sea en prisión preventiva o tras ser condenado, por un delito violento cometido con posterioridad a la fecha de puesta en libertad definitiva); reincidencia violenta en general (nuevo ingreso en el sistema de ejecución penal catalán por la comisión de un nuevo delito violento, cuando a los casos castigados con pena de prisión se suman aquellos en que el delincuente ha sido condenado con otro tipo de pena distinta de la privativa de libertad); cualquier reincidencia penitenciaria (nuevo ingreso en el sistema penitenciario por cualquier tipo de delito); y reincidencia general con cualquier tipo de sanción penal. Además, los resultados de este grupo se comparan con los de un grupo de control, formado por sujetos para los que RisCanvi había estimado un riesgo medio y bajo de reincidencia violenta.

Tras comparar el riesgo que RisCanvi había estimado antes de salir de prisión con los datos de reincidencia real tras el periodo de seguimiento, los autores observan (pp. 62 y ss.) que, entre quienes reincidieron tras salir de prisión, RisCanvi solo había conseguido identificar como de riesgo alto a cerca de la mitad de ellos (valores de sensibilidad que oscilan entre el 46 % y el 55,6 %) en los cuatro tipos de reincidencia, RisCanvi sí consigue mejores resultados en la detección de personas que no reinciden: entre este grupo, RisCanvi acierta –asignándoles un riesgo medio o bajo– en cerca del 70 % de los casos, respecto a los 4 tipos de reincidencia estudiados.

Como la muestra para este estudio está disponible como datos abiertos, también hemos podido calcular los valores predictivos para los 4 tipos de reincidencia. Los valores predictivos positivos son del 32,7 % para cualquier reincidencia penitenciaria, del 40,6 % para cualquier reincidencia, del 18,5 % para la reincidencia penitenciaria violenta y del 24,7 % para cualquier reincidencia violenta (Briz-Redón y Montes (2022); tablas 53-56). Es decir, entre el grupo considerado de alto riesgo, el porcentaje de reincidentes reales nunca alcanza 50 %, oscilando entre un mínimo del 18,5 % (para la estimación de la reincidencia penitenciaria violenta) y un máximo del 40,6 % (para cualquier reincidencia general).

Los valores predictivos negativos son del 85,4 % para cualquier reincidencia penitenciaria, del 75,4 % para cualquier reincidencia, del 92,8 % para la reincidencia penitenciaria violenta y del 85,8 % para cualquier reincidencia violenta. La capacidad predictiva de RisCanvi es mucho mejor para el grupo de riesgo medio y bajo: entre los sujetos que RisCanvi clasifica en esta categoría, el porcentaje de los que no reinciden durante el periodo de seguimiento es siempre superior al 75 %, hasta alcanzar el 92 % en lo que se refiere a la reincidencia violenta que conlleva un reingreso en prisión.

Tabla 22
Resumen de la capacidad predictiva de RisCanvi para diferentes tipos de reincidencia

<i>Estudios</i>	<i>Tipo de reincidencia</i>	<i>Sensibilidad</i>	<i>Especificidad</i>	<i>Valor predictivo positivo</i>	<i>Valor predictivo negativo</i>
Capdevila 2015	Reincidencia violenta	77,15 %	57,26 %	17,94 %	95,4 %
Capdevila 2022	Cualquier reincidencia penitenciaria	52,3 %	72,2 %	32,7 %	85,4 %
	Cualquier reincidencia	49,0 %	72,2 %	40,6 %	75,4 %
	Reincidencia penitenciaria violenta	55,6 %	70,0 %	18,5 %	92,8 %
	Cualquier reincidencia violenta	46,0 %	70,0 %	24,7 %	85,8 %

Fuente: datos extraídos de Capdevila (2015: 151), Capdevila (2022: 64 y ss.) y Briz-Redón y Montes (2022), tablas 53-56.

Estamos de acuerdo con Capdevila et al. (2022) cuando concluyen, para los datos de su estudio, que en lo que se refiere a la «predicción de la sensibilidad, es decir, la reincidencia que es objeto de este estudio, el grado de acierto es nulo, y la herramienta no da ninguna previsión fiable en los cuatro tipos de medida» (p. 67). Solo añadiríamos que lo mismo ocurre cuando se mide la capacidad predictiva en términos de valores predictivos positivos. Según estos autores, hay dos posibles explicaciones para este fenómeno, o ambas: «o existe una sobreestimación del riesgo en estos casos [...], o la misma detección y actuación posterior ha sido eficaz para rebajar el riesgo de reincidencia, aunque en ninguno de los dos sentidos puede hacerse una afirmación tazativa» (p. 122). A la vista de estos datos, parecen sensatas las recomendaciones que los autores hacen a las autoridades penitenciarias:

se podría concluir que con una persona con calificación de *medio/bajo* riesgo el equipo de tratamiento podría hacer una apuesta decidida por la progresión de grado se podría concluir que con una persona con calificación de *medio/bajo* riesgo el equipo de tratamiento podría hacer una apuesta decidida por la progresión de grado¹² y propuesta de salida a medio abierto, con probabilidades de error bajas

¹² Los presos en España se clasifican en uno de tres grados, cada uno de los cuales corresponde a un régimen penitenciario diferente: primer grado o régimen cerrado (con una vigilancia muy intensa y actividades muy limitadas con otros internos), segundo grado o régimen ordinario (vivir dentro de la prisión compartiendo actividades con otros internos y salir ocasionalmente de permiso), y tercer grado o régimen abierto (poder salir todos los días a trabajar y solo pasar las noches y fines de semana en prisión, o incluso permanecer toda la

respecto a la reincidencia. En el caso de la calificación de alto riesgo habría que mejorar mucho más su valoración antes de incluirla [a la persona] en el protocolo [el protocolo especial para la excarcelación de internos con alto riesgo de reincidencia violenta], como recomienda la misma herramienta RisCanvi, pasándole herramientas específicas en los casos de violencia general (HCR-20), violencia sexual (SVR-20) o violencia de pareja (SARA) (p. 123).

Los autores añaden que la aplicación de estas evaluaciones complementarias con dichas herramientas ha sido escasa en los casos que analizan en este informe (presos de alto riesgo excarcelados entre 2014 y 2016). No tenemos datos ni información sobre si, en los casos en los que RisCanvi da un riesgo alto, se aplican a los internos estas otras herramientas estructuradas de evaluación del riesgo a las que se refiere el párrafo transcrito, ni con qué resultados.

Como hemos comentado *supra*, no estamos seguros de que este enfoque cauteloso de los niveles de riesgo del RisCanvi, que sugiere el informe de Capdevila et al. (2022), se esté siguiendo realmente en la práctica. Sí lo parece en el ámbito de las audiencias provinciales, donde es muy poco habitual que un nivel de riesgo alto en RisCanvi determine la decisión que tomar sobre un interno, y normalmente se tienen en cuenta también otros factores. Pero, por lo que dicen los abogados, parece que los juzgados de vigilancia penitenciaria sí dan mucha relevancia a una predicción de riesgo alto en RisCanvi, y también sabemos que la Fiscalía siempre recurre las decisiones que conceden permisos o libertad condicional cuando el nivel de riesgo evaluado con RisCanvi es alto o medio. En cuanto a las decisiones de las juntas de Tratamiento, no tenemos información directa sobre este contexto y no nos atrevemos a hacer suposiciones.

En cualquier caso, hay una información interesante en el informe de Karimi-Haghighi y Castillo sobre hasta qué punto RisCanvi influye en las decisiones sobre los reclusos: en su muestra, de las personas liberadas a las que RisCanvi asigna un riesgo bajo de reincidencia violenta, entre un tercio y aproximadamente la mitad reciben la libertad condicional, mientras que de las personas liberadas a las que RisCanvi asigna un riesgo alto de reincidencia violenta, solo entre el 1 y el 3 % reciben la libertad condicional. La interpretación que hacen estos autores es que RisCanvi probablemente influye en el hecho de que una persona reciba o no la libertad condicional, pero no determina este resultado. En nuestra opinión, RisCanvi parece ser más determinante cuando el riesgo es alto (ya que solo entre el 1 y el 3 % de las personas con esta puntuación de riesgo obtienen la libertad condicional), pero no tanto cuando el riesgo asignado es bajo, ya que a más de la mitad de las personas con esta puntuación de riesgo no se les concede.

Por último, nos gustaría añadir una breve referencia a los resultados de la auditoría externa de RisCanvi encargada por la Generalitat de Catalunya y publicada en

semana en casa con algún tipo de vigilancia electrónica, etc.). Para ser clasificado en un grado menos restrictivo el interno tiene que cumplir unos requisitos, como obedecer las normas penitenciarias y mostrar una conducta aceptable, y ser considerado no peligroso. La expresión en español para la clasificación en mejor grado es «progresar de grado»; la clasificación en peor grado se dice «regresión de grado».

2024. Aunque esta auditoría analiza la capacidad predictiva de RisCanvi para los cinco tipos de reincidencia que predice la herramienta (violencia intrainstitucional, violencia autodirigida, quebrantamiento de permisos, reincidencia general y reincidencia violenta), nos referiremos aquí únicamente a los resultados relativos a la evaluación del riesgo de reincidencia violenta, por ser el único tipo de reincidencia que hemos estudiado en nuestro informe.

Según la auditoría, en relación con la reincidencia violenta, RisCanvi (versión completa) alcanza los siguientes valores de sensibilidad y especificidad:

Tabla 23
Sensibilidad y especificidad en las evaluaciones del riesgo de reincidencia violenta con RisCanvi versión completa según los resultados de la auditoría externa

	<i>Sensibilidad (tasa de verdaderos positivos)</i>	<i>Especificidad (tasa de verdaderos negativos)</i>
Agrupando riesgo medio con riesgo alto	27,71 %	90,07 %
Agrupando riesgo medio con riesgo bajo	55,21 %	69,31 %

Fuente: Herron Mulet et al. (2024: 34 y 72).

Por otra parte, aunque la auditoría no proporciona los valores predictivos, los hemos calculado utilizando los parámetros indicados en el informe, y hemos obtenido los siguientes resultados:

Tabla 24
Valores predictivos positivos y negativos en las evaluaciones del riesgo de reincidencia violenta con RisCanvi versión completa según los resultados de la auditoría externa

	<i>Valor predictivo positivo</i>	<i>Valor predictivo negativo</i>
Agrupando riesgo medio con riesgo alto	16,8 %	94,5 %
Agrupando riesgo medio con riesgo bajo	11,54 %	95,54 %

Fuente: elaboración propia a partir de los datos facilitados en Herron Mulet et al. (2024: 14, 21, 34 y 72) (N = 2,251 evaluaciones RisCanvi-completas; tasa de reincidencia violenta del 6,76 %; y datos de sensibilidad y especificidad recogidos en la tabla anterior).

Consideramos que estos resultados están en consonancia con los ya dados a conocer en estudios anteriores, en el sentido de que los valores predictivos positivos para las evaluaciones del riesgo de reincidencia violenta son muy bajos, los valores predictivos negativos, en cambio, son muy altos, y la especificidad es elevada. Llama la atención

el bajísimo porcentaje de sensibilidad encontrado en el informe de auditoría para el caso de la reincidencia violenta agrupando los niveles de riesgo medio y alto (27,71 %), que difiere llamativamente de los valores encontrados en estudios anteriores (nunca inferiores al 40 %), pero no sabemos a qué se debe, ya que la auditoría no lo indica.

En resumen, la auditoría sugiere, como también lo hicieron los anteriores estudios publicados, que RisCanvi tiene una buena capacidad predictiva para detectar y predecir a los delincuentes de bajo riesgo, pero una capacidad mucho más débil para los delincuentes de alto riesgo. Esto hace que los resultados de RisCanvi sean, en nuestra opinión, un buen argumento para conceder libertades y beneficios a los reclusos de bajo riesgo. Sin embargo, los resultados de RisCanvi no son un buen argumento para restringir la libertad o los derechos de los reclusos de alto riesgo, porque tal restricción de derechos tiene una base empírica mucho más débil. Las puntuaciones de alto riesgo solo deberían traducirse en medidas de reducción de riesgos que no impliquen la restricción de derechos (ampliación de la asistencia: programas de tratamiento, acompañamiento, asesoramiento, etc.).

4. EVALUACIÓN DEL IMPACTO SOBRE LOS DERECHOS HUMANOS

Queremos subrayar una vez más que hemos realizado un análisis limitado de RisCanvi. Los datos con los que hemos trabajado se refieren únicamente a la reincidencia violenta y a la escala REVI, mientras que RisCanvi evalúa otros cuatro tipos diferentes de riesgo. No hemos podido hablar con el personal penitenciario, ni analizar las resoluciones de los juzgados de vigilancia penitenciaria. Por lo tanto, las conclusiones que podemos extraer también tienen que ser prudentes y limitadas en su alcance a las cuestiones que hemos analizado.

Posibles sesgos y problemas de subjetividad

Queremos subrayar que RisCanvi se utiliza ahora para hacer predicciones de riesgo que eran hechas por humanos antes de que existiera la herramienta, porque hay muchas decisiones en el régimen penitenciario que siempre han dependido (total o parcialmente) de la estimación del riesgo de reincidencia. Podría ocurrir que los sesgos o efectos discriminatorios de los que ahora se podría acusar a RisCanvi también estuvieran presentes antes, cuando las estimaciones de riesgo no estaban automatizadas. Es difícil evaluar hasta qué punto RisCanvi funciona mejor o peor que antes en este sentido, en parte porque no hemos podido hablar con el personal penitenciario y pedirle su opinión sobre estos problemas. Pero si se pusieran a disposición datos abiertos sobre RisCanvi, y mejor aún si se pusieran a disposición las fórmulas actualizadas de los algoritmos de las diferentes escalas y los umbrales de discriminación, su análisis permitiría detectar estos posibles sesgos (independientemente de si los genera RisCanvi o de si RisCanvi solo arroja luz sobre los que siempre han estado ahí), y debatir sobre ellos y sobre las razones por las que se producen, así como sobre las mejores maneras de minimizarlos.

Consideramos que el hecho de que RisCanvi asigne un mayor riesgo de reincidencia a los españoles que a los extranjeros debe valorarse positivamente, ya que

RisCanvi no parece discriminar ni directa ni indirectamente a los extranjeros, y quizás en este sentido sus valoraciones de riesgo funcionen mejor que las valoraciones realizadas por una persona física, donde la probabilidad de un sesgo xenófobo o racista nunca puede excluirse por completo (aunque carecemos de datos para confirmar esta hipótesis, ya que no disponemos de una muestra de sujetos cuyo riesgo hubiera sido valorado sin utilizar RisCanvi).

RisCanvi asigna un mayor riesgo de reincidencia a las personas con un trastorno mental grave, un trastorno de la personalidad o una discapacidad intelectual (en adelante, TM) que a quienes no se encuentran en estas situaciones, y la proporción en que aumenta el nivel de riesgo en estos colectivos es muy superior al aumento de su tasa de reincidencia (y lo mismo ocurre con las personas que abusan del alcohol u otras drogas). Esto nos lleva a pensar que RisCanvi es extremadamente severo en su valoración del riesgo de las personas con TM, y también de las que tienen TM y abusan de sustancias, lo que nos parece preocupante. En España, como en muchos otros países, existe una sobrerrepresentación de personas con trastornos mentales en prisión; las cifras son difíciles de estimar y varían de un estudio a otro, pero se calcula que hay entre un 4 y un 5 % de personas con trastornos mentales graves en prisión, una tasa cuatro veces superior a la de la población general. Hay muchas causas detrás de este complejo fenómeno, pero en cualquier caso se trata de grupos vulnerables, para los que la estancia en prisión puede ser más perjudicial que para las personas que no sufren trastornos. Si el personal que trabaja con los internos no es consciente de la gran incidencia que tienen los trastornos mentales en el nivel de riesgo global que RisCanvi asigna al sujeto, esto podría favorecer que se adopten decisiones restrictivas de derechos sobre la base de un alto nivel de riesgo que se debe fundamentalmente a estos factores, y que podría acabar derivando en una discriminación de aquellos que requieren un mayor nivel de protección por parte de la Administración.

Por último, tanto los estudios de Capdevila como los de Karimi-Haghighi y Castillo sugieren que RisCanvi sobrestima el riesgo de las personas con problemas económicos (y puede que sobrestime de manera muy relevante el riesgo de las personas con dificultades socioeconómicas especialmente intensas, aunque es necesario ser cautos con respecto a este resultado porque los grupos en las muestras eran demasiado pequeños como para extraer conclusiones sólidas). No parece justo que se someta a unas personas a un régimen penitenciario más duro que a otras por el hecho de ser pobres, sobre todo cuando los datos parecen mostrar que el nivel de riesgo debido a estos factores aumenta a un ritmo mucho mayor que la tasa real de reincidencia de quienes se encuentran en estas circunstancias. En nuestra opinión, también sería necesario que esto se conociera (el personal penitenciario, los jueces y los fiscales), a fin de decidir cuándo puede ser injustificable tomar determinadas decisiones sobre las personas en función de su nivel de riesgo, si este viene determinado por este tipo de factores.

Llegados a este punto, consideramos oportuno hacer una reflexión. En el ámbito penitenciario, las leyes siempre han exigido que, para tomar determinadas decisiones, se tenga en cuenta el riesgo de reincidencia que presenta una persona (antes llamado «peligrosidad», terminología que todavía se utiliza en muchas leyes, incluido el Código Penal español). Existen otras herramientas policiales predictivas que

operan en ámbitos en los que antes no se realizaban estimaciones de riesgo, y en los que por tanto no es posible comparar si los algoritmos producen ahora mayores o menores problemas de sesgos y discriminaciones que los que antes producían los juicios individuales de peligrosidad.¹³ Pero respecto a RisCanvi sí es posible –e incluso necesario– preguntarse cuál de las dos alternativas plantea mayores problemas: si confiar las valoraciones de riesgo o de peligrosidad a los individuos, al personal que trabaja en prisión con sus distintas especializaciones profesionales, o introducir algoritmos. Y, en nuestra opinión, un análisis riguroso de esta cuestión debe reconocer que, cuando las valoraciones son realizadas por individuos, es muy posible que contengan sesgos y den lugar a discriminaciones contra personas o grupos por diversas razones (ya sea por prejuicios racistas, xenófobos, socioeconómicos o de otro tipo, ya sea porque cada persona tenga preferencias y simpatías –o antipatías– personales por determinados internos, o incluso por la arbitrariedad o falta de profesionalidad de algunos individuos).

Teniendo esto en cuenta, la introducción de algoritmos puede tener, en nuestra opinión, dos efectos posibles. Por un lado, los algoritmos corren el riesgo de «blanquear» tecnológicamente (*tech-washing*) los sesgos existentes, aparentando que las valoraciones son más objetivas que las que haría una persona, aunque los factores con los que trabajan los algoritmos obligan a los evaluadores a emitir juicios subjetivos, cuya naturaleza, sin embargo, se oculta tras la aparente objetividad y neutralidad del algoritmo. De este modo, los sesgos permanecen, pero se hacen invisibles, lo que, por un lado, legitima el sistema (ocultando que al castigar más severamente a determinados colectivos los discrimina injustamente) y, por otro, puede reforzar los sesgos: si, por ejemplo, el algoritmo otorga mayor riesgo a los pobres, estos recibirán un trato más severo, lo que probablemente reforzará los factores que determinarán que, en la siguiente ocasión, su riesgo vuelva a ser alto o incluso superior. El problema se agrava aún más porque es más fácil cuestionar y recurrir las decisiones individuales que el resultado de un algoritmo, cuyos detalles se mantienen en secreto.

En RisCanvi hay factores que desde el principio hacen sospechar que tras la objetividad del algoritmo se esconde una buena dosis de valoración subjetiva, y una vez conocida la descripción de los ítems creemos que esta objeción ha quedado plenamente confirmada. A pesar de la detallada descripción de los factores de riesgo, que revelan un serio y loable intento de garantizar que cualquiera que aplique esta herramienta lo haga de la misma manera, muchas de las variables (por ejemplo, de la 38 a la 43) son evaluaciones de rasgos de personalidad que inevitablemente implican comparar al sujeto con algún estándar de «normalidad». Por ejemplo, evaluar si tiene o no «arrepentimiento sincero» (n. 40); si sus comportamientos son

¹³ Es el caso de VioGén, ya que su implantación ha sido casi simultánea a la puesta en marcha de un marco de protección policial para las mujeres que sufren violencia de género que antes sencillamente no existía. Y es también en gran medida el caso de VeriPol, si tenemos en cuenta que la detección de denuncias falsas, además de hacerse en muy pocos casos, no era antes una comprobación que la ley obligara a hacer a los policías, sino el resultado al que podía conducir una investigación rigurosa.

«inapropiados a la luz del contexto en el que se producen» o si muestra «respuestas irracionales» (n. 41); si «se enfada por nimiedades» (n. 42); si tiene un «estilo de vida parasitario» o «se aprovecha» de su familia y amigos. Nótese que no estamos diciendo que no sea posible hacer una evaluación rigurosa de estas variables. Lo que decimos es que tal valoración es inevitablemente subjetiva, y que en estos aspectos las valoraciones algorítmicas no difieren –no pueden diferir– mucho de los antiguos juicios clínicos de peligrosidad.

También nos parece que la descripción de la variable 26, que ya no describe rasgos de personalidad, es extremadamente imprecisa y puede dar lugar a todo tipo de sesgos subjetivos por parte de los evaluadores.

Creemos que la Generalitat de Catalunya es consciente, al menos en parte, de este problema. En el estudio *Tasa de reincidencia penitenciaria 2020*, publicado en marzo de 2023, se afirma que:

21. Algunas de las variables dinámicas del RisCanvi presentan importantes diferencias en función de la valoración que, de cada centro, según los resultados anejos no publicados que hemos hecho centro por centro. Esto significa que la interpretación de la evidencia inicial de cada variable no es homogénea ni suficientemente estandarizada y plantea dudas en el equipo profesional sobre la eficacia de la herramienta de diagnóstico, que se trasladará a la toma de decisiones posterior (p. 305).

Y en la p. 310 incluyen, como propuesta para abordar este problema, lo siguiente:

I. Los resultados obtenidos por centros penitenciarios y no publicados en este informe respecto a las valoraciones e interpretaciones de las variables del RisCanvi nos confirman que no son homogéneas. Se impone como estrategia una revisión global de los criterios que sostienen los criterios de evidencias de estas variables, si bien se recomienda hacerlo de forma quirúrgica centro por centro y en supervisión de casos. Se recomienda formar un equipo de personas expertas en evaluación del riesgo y reconocidas dentro del Sistema que además tengan capacidad didáctica y pedagógica para orientar al resto de compañeros de centros que tengan que llenar el RisCanvi inicial. Si el primer RisCanvi está bien orientado, el resto de las revisiones resulta más fácil.

J. Se hace imprescindible disponer de una formación especializada de evaluación del RisCanvi y de gestión del riesgo aplicada. También de una formación inicial y de acogida que integre a los nuevos y las nuevas profesionales en esta técnica de trabajo penitenciaria.

Este problema podría compensarse en cierta medida con la mayor transparencia que pueden aportar las herramientas estructuradas de evaluación del riesgo en comparación con las evaluaciones clínicas de la peligrosidad (y este es el segundo efecto al que nos referíamos antes). Si los algoritmos son transparentes, entendiendo por ello no solo que la fórmula del algoritmo y la descripción de los ítems sean públicas, sino también que estén disponibles muestras de datos (anonimizados, obviamente) que contengan la codificación de cada ítem, los niveles de riesgo y los datos sobre reincidencia, se hace posible analizar cómo funciona el algoritmo, cómo se recoge la información, si da diferentes niveles de riesgo para distintos grupos, y si al hacerlo está discriminando por motivos constitucionalmente prohibidos. Los algoritmos transparentes pueden arrojar luz sobre los sesgos y las discriminaciones directas

o indirectas que producen las evaluaciones de riesgo, no solo las generadas por la herramienta, sino también las que quizás han estado inspirando las evaluaciones realizadas previamente por humanos. Creemos que esto puede ser positivo porque:

- 1) los ciudadanos (y sus abogados) pueden impugnar los resultados y defender mejor sus intereses;
- 2) se puede debatir qué factores es legítimo tener en cuenta en los algoritmos de predicción y cuáles no, qué valor dar a cada uno, si hay que asumir costes en términos de capacidad predictiva para minimizar los efectos discriminatorios, etc.; y
- 3) también porque puede entablarse un debate sobre cómo afectan a la delincuencia las propias dinámicas sociales y desigualdades estructurales y, por tanto, qué políticas podrían reducirlas con mayor eficacia. Es decir, si los algoritmos nos están diciendo que factores como el nivel educativo, el nivel de ingresos, los problemas de salud mental, etc., son factores que aumentan el riesgo de delinquir, parece bastante miope tenerlos en cuenta solo para medir el riesgo de reincidencia de cada individuo: el algoritmo nos está indicando que las propias desigualdades estructurales del sistema social generan riesgos, y por tanto pueden ofrecer pistas sobre hacia dónde dirigir políticas públicas que corrijan los factores sociales que favorecen la criminalidad (ampliación de los servicios públicos de atención a la salud mental; políticas económicas, fiscales y laborales que reduzcan las desigualdades, políticas de reforma educativa, etc.).

Por supuesto, para que todo esto sea posible, se requiere un nivel de transparencia muy elevado, que RisCanvi obviamente no alcanza por el momento (como tampoco ninguna de las otras dos herramientas analizadas en este informe).

En RisCanvi hay muchos puntos que son cuestionables desde el punto de vista de los derechos fundamentales. Por ejemplo, el hecho de que tener una baja capacidad intelectual (n. 39) aumente el riesgo puede equivaler a una discriminación por motivos de discapacidad, lo que es contrario a la Convención de la ONU de 2006. También es muy llamativo que en el factor número 33, tener muchos encuentros sexuales y aventuras de una noche, una preferencia por una amplia gama de actividades sexuales o ser infiel, eleve el nivel de riesgo aparentemente tanto como haber coaccionado sexualmente a otras personas o haber sido condenado por agresión sexual. Aparte del hecho de que las preferencias sexuales que no perjudican a otros son parte de la intimidad personal y del libre desarrollo de la personalidad, la valoración de este ítem corre el riesgo de que una persona sea considerada más peligrosa solo por tener una moral sexual diferente a la dominante, o a la del evaluador concreto que codifica este factor RisCanvi, lo que en nuestra opinión supondría una discriminación injustificada. Incluso en el caso de que estos factores pudieran tener un efecto de correlación demostrable con un mayor riesgo de reincidencia, algo que debería analizarse con mucha más transparencia de como hasta ahora se ha hecho el desarrollo de la herramienta, es muy preocupante desde la perspectiva de los derechos fundamentales su inclusión como factor predictivo, preocupaciones que deberían abordarse y debatirse abiertamente.

También es cuestionable utilizar en contra del sujeto el hecho de que tenga una opinión desfavorable del sistema de justicia penal, de la ley o del Gobierno (n.º 38). Creemos que esto forma parte del derecho a la libertad de conciencia, y no solo es un rasgo que caracteriza a muchos miembros de grupos políticos, movimientos de protesta civil, activistas del cambio climático o de los derechos humanos, sino que parece especialmente comprensible que tales actitudes sean mantenidas por quienes están sufriendo personalmente la dureza de las penas de prisión.

Dado que este tipo de factores implican *a priori* una limitación de derechos fundamentales, para que sea admisible tomarlos en consideración sería imprescindible justificar detalladamente en qué medida incrementan el riesgo de reincidencia, si lo hacen en mayor o menor medida que otros factores menos cuestionables y, por tanto, si es absolutamente necesario tomarlos en consideración para reducir sensiblemente el riesgo de reincidencia grave y violenta. En otras palabras, sería necesario llevar a cabo un análisis de proporcionalidad: un examen detallado de hasta qué punto la injerencia en los derechos fundamentales que supone la valoración de estos factores es necesaria y proporcionada, en comparación con el beneficio que supone para la sociedad la reducción de la reincidencia. Pero es imposible hacerlo mientras no se conozcan los detalles de la fórmula del algoritmo, y los estudios y cálculos en los que se ha basado la asignación de pesos a cada factor. Nosotros, en particular, no hemos podido realizar esta valoración porque, además de desconocer los trabajos preparatorios de RisCanvi y la ponderación y codificación exacta de cada factor, el único conjunto de datos al que hemos tenido acceso directo era una muestra poco representativa de la población reclusa (la muestra de Capdevila, en la que están sobrerrepresentados los internos de alto riesgo).

Otro factor de riesgo que consideramos de dudosa legitimidad es el número 17. Porque la distancia del centro penitenciario a la residencia del preso no depende de la elección de este último, sino que es la Administración penitenciaria la que decide dónde tiene que cumplir cada recluso su condena. Así pues, la propia Administración penitenciaria puede aumentar el nivel de riesgo al negarse a permitir que el recluso cumpla su condena cerca de su lugar de residencia y, a continuación, alegar este elevado nivel de riesgo como motivo para no concederle permisos o un régimen de vida menos severo.

También queremos destacar que muchos ítems dependen de situaciones que afectan al pasado de la persona o que, además de afectar a su pasado, son situaciones de las que no es responsable. Es el caso, por ejemplo, de los factores 8, 16, 22 o 23. Situaciones relacionadas con su infancia o juventud, o incluso con el presente, pero que él no puede cambiar (los antecedentes penales de miembros de su familia), pueden seguir aumentando el riesgo a pesar de los esfuerzos que haga el interno por mejorar en otros aspectos (controlar su ira, seguir un tratamiento, desintoxicarse de las drogas, etc.). Si los factores estáticos pesan mucho en el nivel final de riesgo, los esfuerzos del interno por mejorar su situación pueden estar condenados al fracaso, ya que el pasado será una pesada mochila de la que no podrá librarse. No sabemos si RisCanvi está diseñado para que estos factores pierdan peso con el tiempo, o si se instruye a las personas que trabajan con la herramienta para que anulen manualmente el nivel de riesgo dado automáticamente por la herramienta, de forma que se puedan evitar o mitigar estas consecuencias.

Somos conscientes de que probablemente cualquier profesional que trabaje en el ámbito de la evaluación de riesgos encontraría estas consideraciones desacertadas, por la siguiente línea de razonamiento: si se quiere evaluar el riesgo de reincidencia, hay que trabajar con toda la información que sea relevante, y si se empieza a cuestionar la legitimidad de los factores de riesgo por sus fricciones con los derechos fundamentales (o por otras consideraciones de equidad similares), ya no se podría hacer una valoración del riesgo con sentido, porque al eliminar información necesaria la evaluación será de mala calidad y dejará de ser útil. Y probablemente dirían, además, que criticar los algoritmos por estos motivos es hipócrita, ya que la misma información (aunque sin llamarla «factores de riesgo») es la que siempre se ha tenido en cuenta para hacer evaluaciones de peligrosidad.

Y nos parece que tendrían razón; la verdadera cuestión que está en juego es, sin embargo, qué conclusión debe extraerse de ello, teniendo en cuenta que las valoraciones de riesgo en el sistema de justicia penal y en las prisiones no solo tienen por objeto ayudar a la persona (por ejemplo, para decidir qué tratamiento es el mejor para ella, para asignar plazas escasas en programas de intervención, o para proporcionar asistencia social, etc.), sino que el nivel de riesgo también determina si se aplica un control penal más severo, y por tanto una restricción de derechos más intensa. Si los algoritmos fueran transparentes, permitirían una discusión más clara e informada (mejor de lo que permiten las valoraciones clínicas de peligrosidad) sobre hasta qué punto es justo o no utilizar cierto tipo de información para este fin, tanto si las valoraciones se hacen con algoritmos como si se realizan sin ellos. Sin embargo, si no son transparentes, esconden tras una pretendida cortina de cientificidad toda la carga subjetiva y normativa que conlleva cualquier valoración de riesgo (hecha con o sin algoritmos) y, al hacerlo, presentan las valoraciones algorítmicas como más legítimas que las anteriores, cuando es muy dudoso que lo sean, con lo que entorpecen el debate y contribuyen a la legitimación de un sistema de control penal basado en el riesgo.

Transparencia

Los algoritmos de RisCanvi son secretos.¹⁴ Las descripciones detalladas de los factores de riesgo y las instrucciones sobre cómo codificarlos no fueron públicas hasta 2024, aunque obtuvimos en febrero de 2023 de la Generalitat una descripción, fechada en 2009, y la hemos reproducido en el anexo de ResCanvi). A partir de 2024, la Generalitat de Catalunya ha publicado esta misma descripción de los factores de riesgo (de las dos versiones de RisCanvi, *screening* y completa) en su página web, junto al Informe final de auditoría y su resumen ejecutivo (<https://justicia.gencat.cat/ca/ambits/reinsercio_i_serveis_penitenciaris/serveis_penitenciaris/RisCanvi/RisCanvi/>).

¹⁴ El documento de Karimi-Haghighi y Castillo (2022) reproduce la fórmula para la escala REVI, pero no para las demás escalas, y no sabemos si esta fórmula sigue siendo la actual o ha sido revisada. Además, esta fórmula se ha hecho pública gracias a este informe; no habríamos podido encontrarla en ninguna parte de no haber sido por la ayuda del profesor Castillo.

No se nos ha concedido permiso para entrevistar al personal penitenciario, ni hemos obtenido respuesta a nuestra petición de una muestra de datos.

Dicho esto, hay que señalar que la Generalitat de Catalunya tiene una práctica de datos abiertos que nos ha permitido –y permite a cualquier investigador interesado– hacer algunos análisis: para algunos de los estudios que han realizado con RisCanvi, han publicado la base de datos en hojas Excel en internet, con la codificación de todos los factores de riesgo e información sobre reincidencia. Lo más interesante son los datos relacionados con el informe *Tasa de reincidencia penitenciaria 2020* (que se puede encontrar aquí: <[https://cejfe.gencat.cat/ca/recerca/opendata/pre-sions/taxa-reincidencia-2020/index.html#googtrans\(ca|en\)](https://cejfe.gencat.cat/ca/recerca/opendata/pre-sions/taxa-reincidencia-2020/index.html#googtrans(ca|en))>), que ya hemos mencionado en el apartado *Enfoque metodológico*, y que no hemos podido analizar para este informe porque se publicaron cuando ya estaba cerrado.

Es cierto que esta política de datos abiertos es limitada, y no incluye una explicación del diseño del algoritmo ni de los umbrales que separan los distintos niveles de riesgo. Pero, en nuestra opinión, esta práctica de datos abiertos debe valorarse muy positivamente, sobre todo teniendo en cuenta el secretismo con el que se han comportado y se siguen comportando todas las administraciones públicas que aplican sistemas automatizados de toma de decisiones en España, incluso cuando se trata de cuestiones mucho menos sensibles que el derecho penal y la reincidencia.¹⁵ Además, la Generalitat ha cedido previamente muestras a otros investigadores (equipo del profesor Castillo), y ha encargado una auditoría externa de RisCanvi, cuyos resultados se han hecho públicos en 2024. Por todo ello, y aunque la herramienta no es transparente, hay mucha más transparencia sobre RisCanvi que sobre las otras herramientas que hemos estudiado en este informe.

Por último, también nos gustaría destacar como un hecho muy positivo que la Generalitat de Catalunya haya encargado una auditoría externa del algoritmo y haya hecho públicos sus resultados. Aunque no comentamos aquí el alcance, la metodología o los resultados de esta auditoría, el mero hecho de que se haya llevado a cabo es una diferencia importante y, en nuestra opinión, muy bienvenida en comparación con las otras dos herramientas de policía predictiva analizadas en este informe.

Capacidad predictiva

RisCanvi no es un instrumento bien calibrado para evaluar la reincidencia violenta: el porcentaje de sujetos clasificados como de riesgo alto en las muestras totales y en muchos de los subgrupos analizados es superior al porcentaje de los que realmente cometieron un nuevo delito violento en el periodo de seguimiento. Esto significa que RisCanvi asigna sistemáticamente más riesgo del que realmente hay en los distintos grupos considerados y en la muestra total.

¹⁵ Los tribunales españoles han denegado (hasta ahora) el acceso al código que ejecuta un algoritmo que decide si los ciudadanos tienen o no derecho a una subvención en su factura de la luz; aquí puede encontrarse una explicación: <<https://civio.es/novedades/2024/05/08/la-audiencia-nacional-vuelve-a-rechazar-abrir-el-codigo-fuente-que-decide-quien-recibe-el-bono-social/>>.

Como hemos explicado detalladamente en el apartado *Eficacia del enfoque policial predictivo para prevenir la delincuencia de RisCanvi*, RisCanvi tiene un bajo nivel de precisión cuando el riesgo estimado es alto. Por el contrario, el nivel de precisión es mucho mejor cuando el nivel de riesgo estimado es bajo.

Dicho esto, queremos señalar que el nivel de riesgo asignado por RisCanvi conlleva la puesta en marcha de unas u otras medidas preventivas para evitar que se produzca el evento predicho. Por tanto, la actuación de la Administración, si es eficaz y consigue evitar que se produzca el nuevo delito que se considera altamente probable, puede paradójicamente hacer aparecer la predicción como un error. Estos efectos de profecía autocumplida también pueden producirse a la inversa, de forma que, por ejemplo, la concesión de un permiso o el acceso a un régimen penitenciario más benigno en un caso de riesgo medio o alto puede animar al preso, que ve que los profesionales confían en sus posibilidades, a realizar un mayor esfuerzo para evitar la reincidencia. No descartamos que estos fenómenos estén condicionando en alguna medida la capacidad predictiva que RisCanvi ha mostrado en los distintos estudios realizados; solo queremos señalar que con la información de que disponemos no podemos cuantificar estos efectos.

Cómo influye el resultado del algoritmo en la toma de decisiones

Como se ha indicado anteriormente, el Ministerio Fiscal concede gran importancia al resultado indicado por RisCanvi cuando el algoritmo asigna un riesgo medio o alto en las diferentes escalas. En algunas audiencias provinciales, el recurso interpuesto por el Ministerio Fiscal suspende el procedimiento de concesión del permiso de salida hasta que la Audiencia Provincial resuelva el asunto. Pueden pasar varios meses hasta que la Audiencia Provincial resuelva definitivamente el recurso. Durante este tiempo, el recluso debe permanecer en prisión y puede ocurrir que, cuando el tribunal se pronuncie a favor de la concesión del permiso, el preso ya haya sido puesto en libertad por haber cumplido su condena. En otras palabras, el efecto suspensivo y el hecho de que el fiscal recurra todas las decisiones por las que se conceden permisos temporales cuando RisCanvi indica un riesgo alto tienen consecuencias importantes en términos de derechos de los reclusos.

Por otro lado, aunque no hemos podido valorar directamente las decisiones de los jueces de vigilancia penitenciaria, ya que no hemos tenido acceso a ellas, según nos han contado los abogados con los que hemos hablado, parece que estos jueces dan mucha importancia al resultado de la evaluación de RisCanvi cuando es de riesgo alto, pero menos cuando indica bajo riesgo, lo que de ser cierto significaría que RisCanvi acaba teniendo un mayor efecto en las decisiones perjudiciales para el preso. Un dato que también parece apuntar en esta dirección es el análisis de los presos a los que se concede la libertad condicional en Karimi-Haghighi y Castillo (2022): de los excarcelados a los que RisCanvi asigna un riesgo REVI bajo, de un tercio a aproximadamente la mitad reciben la libertad condicional, mientras que de los excarcelados a los que RisCanvi asigna un riesgo REVI alto, solo entre el 1 y el 3 % reciben la libertad condicional.

Como no hemos podido hablar con miembros de las juntas de tratamiento, no podemos evaluar el grado de importancia que este órgano concede al riesgo que

RisCanvi estima, Solo contamos con la opinión de los abogados, según los cuales la Junta de Tratamiento concede demasiada importancia a RisCanvi. Sin embargo, es interesante en este contexto mencionar una de las conclusiones del informe *Tasa de reincidencia penitenciaria 2020*, publicado en marzo de 2023, que se refiere indirectamente a esta cuestión:

22. El equipo profesional de los centros penitenciarios todavía no aplican realmente el modelo RNR (riesgo, necesidad, responsividad) como herramienta de gestión del riesgo. Tampoco se acaban de creer la valoración de RisCanvi y confían aún más en sus propias percepciones que en la evidencia. Una tercera parte de los condenados que tienen un diagnóstico de riesgo bajo no tendrán ni un solo permiso ordinario durante toda la condena: estos triplican la tasa de reincidencia (31,2 %) respecto a quienes sí que han tenido permisos (10,3 %). Dos de cada cinco de riesgo bajo saldrán en libertad definitiva desde segundo grado y a una cuarta parte se les propondrá realizar el programa de tratamiento especializado (cuando en realidad no era necesario) y con los resultados poco eficientes.

En el texto citado, los autores del informe parecen centrarse en el problema de que el personal penitenciario no se cree del todo el RisCanvi cuando da un riesgo bajo, y no aplica a los internos las medidas más favorables que corresponderían a ese nivel de riesgo. Queda por ver qué ocurre en los casos en los que RisCanvi asigna riesgo medio o alto: si en estos casos este nivel sí condiciona más directamente el trato a los presos, o si el personal penitenciario no se cree la herramienta y aplica medidas más laxas que las que corresponderían a estos niveles de riesgo (lo que nos parece poco probable).

En cualquier caso, cabe destacar que las conclusiones extraídas en el informe *Tasa de reincidencia penitenciaria 2020* sobre qué hacer con las valoraciones de riesgo de RisCanvi¹⁶ no abogan por endurecer las condiciones penitenciarias de los internos de riesgo alto o medio, sino por flexibilizar el régimen de vida de los de riesgo bajo, adelantando la libertad condicional o la concesión de permisos y salidas penitenciarias:

los análisis estadísticos nos confirman que la salida desde medio abierto¹⁷ y lo más escalonada posible es la mejor solución para reducir la reincidencia, la apuesta del Departamento de Justicia debería ser fomentarlo de forma decidida y explicada desde el liderazgo. Invertir los esfuerzos en dotar al medio abierto de mayores y mejores recursos con el fin de asumir con garantías el control y seguimiento de estas personas en la comunidad y en vías de alcanzar el desistimiento delictivo y la reinserción social. B. Si el 70,8 % de los condenados valorados con RisCanvi dan riesgo bajo de reincidencia, la propuesta de clasificación inicial en tercer grado debería ser la medida automática para todos estos casos y solo corregida con argumentos contrastados y evidencias de otros factores negativos por parte de las personas profesionales de tratamiento (p. 308).

¹⁶ Los autores del informe trabajan principalmente para la Administración penitenciaria o para el Departament de Justícia de la Generalitat de Catalunya, y han contado también con la colaboración externa de diferentes investigadores y profesores de universidades catalanas.

¹⁷ «Medio abierto» o «régimen abierto» es el término español que se refiere a las penas de prisión que no se cumplen dentro de un centro penitenciario, sino en la comunidad, bien porque se ha concedido la libertad condicional, bien porque el recluso ha alcanzado el «tercer grado» (véase una explicación al respecto en la nota 30 de este informe).

Obligaciones activas de rendición de cuentas

Hay una serie de obligaciones activas de rendición de cuentas que las autoridades que utilizan RisCanvi también han de cumplir.

1. Al igual que en el caso de otras herramientas descritas en este informe, consideramos esencial, también en relación con RisCanvi, llevar a cabo una evaluación de impacto de protección de datos completa y exhaustiva de la herramienta y de su impacto y efectos (tanto en materia de protección de datos como más ampliamente, para el futuro y según normativa europea, sobre su impacto en materia de derechos fundamentales), que por el momento parece inexistente o, al menos, no se ha publicado. Así lo prevé el reglamento para aquellos tratamientos de datos que puedan suponer un alto riesgo para los derechos y libertades de las personas físicas. Se presume un alto riesgo cuando estos tratamientos se realicen con nuevas tecnologías (art. 35 RGPD). Desconocemos si esta evaluación de impacto sobre la protección de datos se ha llevado a cabo de forma efectiva, pero según toda la información que hemos recibido o que es de acceso público podemos deducir que en la fecha de realización de este informe seguía faltando, aunque con posterioridad a la misma se han publicado una auditoría externa y el compromiso por parte de las autoridades catalanas de llevarlas a cabo. En este caso, es el Departamento de Justicia de la Generalitat de Cataluña la autoridad responsable de llevar a cabo la evaluación de impacto sobre la protección de datos y quien debería proporcionar la información completa requerida en relación con este punto, lo que por el momento no ha sido el caso.
2. RisCanvi, además, y como es sabido, trata datos especialmente protegidos, por lo que deben adoptarse mayores garantías, siguiendo las directrices e instrucciones establecidas por el RGPD europeo y de acuerdo con los requisitos detallados para esta categoría de datos declinados por las autoridades de protección de datos (en España, la Agencia Española de Protección de Datos –AEPD– ha realizado un extenso trabajo en la materia, pero para el uso de datos realizado por las administraciones públicas en Cataluña, la Agència Catalana de Protecció de Dades –ACPD– tiene competencia plena en aquellas cuestiones relativas al control sobre el cumplimiento y posibles incumplimientos).
3. Se ha de elaborar un registro de las operaciones de tratamiento de datos que se lleven a cabo en el empleo de la herramienta. Este registro deberá ser público, según el artículo 32.4 de la Ley Orgánica 7/2021, de 26 de mayo, de protección de datos de carácter personal tratados para la prevención, detección, investigación y enjuiciamiento de infracciones penales y la ejecución de sanciones penales. El Departamento de Justicia de la Generalitat de Cataluña es la autoridad responsable de publicar esta información.

Los tratamientos de datos realizados por RisCanvi están publicados en el registro del Departamento de Justicia. Esta información puede consultarse en el siguiente enlace (p. 22): <<https://justicia.gencat.cat/ca/departament/protecciodades/registre-activitats-tractament/>>.

Acceso a los datos y posibilidad de cuestionarlos en Resultados de RisCanvi

A partir de la indagación que hemos realizado, y según nos consta, no existe una regulación específica sobre cómo y en qué condiciones los penados pueden acceder a la información que RisCanvi contiene sobre ellos o a las valoraciones de riesgo que arroja. Siendo así, la regulación aplicable, a nuestro entender, sería la general existente en España sobre el acceso de los internos a la información contenida en sus expedientes penitenciarios, con independencia del formato en que ésta se contenga.

Esta regulación ha venido constituida tradicionalmente por el artículo 15.2 de la Ley Orgánica General Penitenciaria, que establece que «(a) cada interno se le abrirá un expediente personal sobre su situación procesal y penitenciaria del que tendrá derecho a ser informado, y para cada penado se formará un protocolo de personalidad», y los artículos correspondientes del Reglamento Penitenciario: artículo 4.2. k) (los internos tienen «derecho a recibir información personal y actualizada sobre su situación procesal y penitenciaria») y artículo 18.1 («Admitido en el establecimiento un recluso, se procederá a verificar su identidad personal, efectuando la reseña alfabética, dactilar y fotográfica, así como a la inscripción en el libro de ingresos y la apertura de un expediente personal relativo a su situación procesal y penitenciaria de la que tendrá derecho a ser informado»). Estas disposiciones también son aplicables en Cataluña, a pesar de que esta comunidad autónoma tiene las competencias de gestión de los centros penitenciarios.

Este derecho a ser informado, sin embargo, tiene algunas excepciones, que han sido elaboradas a lo largo del tiempo por los fiscales y jueces de vigilancia penitenciaria y las audiencias provinciales, y que la Secretaría General de Instituciones Penitenciarias sistematizó en una Instrucción interna de 2019¹⁸ (<https://www.institucionpenitenciaria.es/documents/20126/78885/I-13-2019-ACCESO_EXPEDIENTES.pdf>), según la cual las pautas generales son las siguientes:

- El principio general es permitir el acceso al expediente penitenciario (ya sea mediante fotocopia si el número de hojas no excede de 15, o poniéndolo a disposición del abogado o familiar que se haya designado expresamente).
- El acceso solo puede ser restringido:
 - si existe riesgo para la seguridad de los técnicos que han emitido los informes cuyo acceso se solicita, en la medida en que el interno, si no le gusta lo que se dice de él en esos informes, podría albergar deseos de venganza contra los profesionales;

¹⁸ Las instrucciones internas de la Secretaría General de Instituciones Penitenciarias, organismo dependiente del Ministerio del Interior, no son vinculantes para la Generalitat de Catalunya, ya que esta tiene cedidas las competencias en materia de gestión penitenciaria. Sin embargo, que sepamos, la Generalitat no ha desarrollado una normativa propia sobre el acceso a los registros penitenciarios y, además, lo que refleja esta circular son los criterios que venían aplicando los jueces de vigilancia penitenciaria y las audiencias provinciales de toda España, incluida Cataluña.

- si se solicita el acceso de forma reiterada o abusiva o si se solicita información que ya debería estar a disposición del preso por habersele notificado previamente; y
- si el hecho de conocer el contenido de los informes técnicos puede poner en peligro la eficacia del tratamiento penitenciario, al romperse la relación de confianza entre los internos y los profesionales (el interno puede dejar de confiar en los profesionales que le tratan, si ve que en los informes dicen cosas de él que no esperaba, con las que no está de acuerdo, etc.).
- La denegación del acceso será siempre motivada y se notificará a los internos por escrito, mencionando expresamente la posibilidad que tienen de recurrir la decisión ante el Juzgado de Vigilancia Penitenciaria y la Agencia Española de Protección de Datos.

Estas posibilidades de denegar el acceso deben ser tratadas, en todo caso, como excepcionales, no solo porque así lo establece la normativa penitenciaria, que exige una completa justificación en cada caso, sino porque la legislación procesal general en España prevé el acceso a los expedientes personales en los expedientes administrativos del interesado, sin excepción posible, según el artículo 70.4 de la Ley 39/2015 (LPAC), cuando el expediente contenga valoraciones o evaluaciones directamente relacionadas con decisiones administrativas relativas al interesado .

Sin embargo, en la práctica, el acceso a los expedientes penitenciarios es más limitado de lo que la letra de la ley indica. Además de los tres motivos de denegación de acceso a los que nos acabamos de referir, jueces y tribunales han mencionado también el hecho de que los informes técnicos de los profesionales no conforman el «expediente penitenciario», sino el «protocolo de personalidad», que se menciona separadamente en el artículo 15.2 de la Ley Orgánica General Penitenciaria, y que no son informes cuyo destinatario sea el interno, sino que se dirigen a la Junta de Tratamiento para que adopte las decisiones correspondientes. Por otra parte, las decisiones relativas al acceso a los expedientes, tanto de los centros penitenciarios como del Juzgado de Vigilancia Penitenciaria, suelen ser estereotipadas, en el sentido de que para justificar la denegación de información se dice genéricamente que se pondría en peligro la seguridad del personal técnico, o que se podría poner en riesgo la eficacia del tratamiento y la relación de confianza entre el interno y los trabajadores penitenciarios, pero sin especificar por qué es así en el caso concreto de cada persona.

Es decir, a pesar de que aparentemente la normativa garantiza un derecho de acceso que solo puede ser exceptuado en casos excepcionales y de forma motivada, en la práctica no es nada fácil acceder a la información completa contenida en el expediente penitenciario y en el mencionado «protocolo de personalidad».

No hemos encontrado ninguna resolución que aborde el derecho de acceso a la información contenida en RisCanvi o al detalle de las valoraciones de riesgo, pero si algún interno lo reclamara, es muy probable que se encontrara con la misma respuesta que en la mayoría de los casos se está dando a quienes solicitan información

que no está contenida en los algoritmos.¹⁹ De hecho, como ya hemos señalado anteriormente, los abogados nos informan de que, en general, los internos no intentan impugnar los resultados de RisCanvi. Además, también nos dijeron que solo en algunos casos, si los reclusos se llevan bien con el personal penitenciario, pueden obtener información sobre su evaluación del RisCanvi a través de canales no oficiales. Por lo tanto, en la práctica, esta parece ser la única forma de obtener realmente una explicación. Así pues, probablemente la Administración penitenciaria catalana está siendo más restrictiva en el acceso a la información contenida en el RisCanvi que respecto del resto de la información de que dispone sobre los internos, pero hay que tener en cuenta que esto ocurre en un contexto en el que, a pesar de lo que dice la letra de la ley, el acceso a los datos e informes que la Administración tiene sobre los internos y sobre su evolución es bastante limitado en la práctica también respecto de las valoraciones no algorítmicas.

Sin embargo, recientemente se han producido algunos cambios relevantes en este ámbito. Por un lado, en 2016 el TEDH dictó la Sentencia Cano Moya c. España (11 de octubre de 2016), en la que se condenó al Estado español por haber vulnerado el artículo 34 del CEDH porque un juzgado de vigilancia penitenciaria no facilitó al interno copia del expediente completo que solicitó para interponer una demanda ante el TEDH.

En segundo lugar, en 2021 se introdujo un nuevo artículo 15 bis en la Ley Orgánica General Penitenciaria, con el siguiente tenor literal:

Artículo quince bis, Tratamiento de datos de carácter personal.

1. Admitido en el establecimiento un interno, se procederá a verificar su identidad personal, efectuando la reseña alfabética, dactilar y fotográfica, así como a la inscripción en el libro de ingresos y a la apertura de un expediente personal relativo a su situación procesal y penitenciaria, respecto del que se reconoce el derecho de acceso. Este derecho sólo se verá limitado de forma individualizada y fundamentada en concretas razones de seguridad o tratamiento.

2. El tratamiento de los datos personales de los internos se regirá por lo previsto en la Ley Orgánica de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales. Los datos personales de categorías especiales que no figuren en el apartado anterior se podrán tratar con el consentimiento del interesado. Sólo se prescindirá de dicho consentimiento cuando sea estrictamente necesario y se efectúe con las garantías adecuadas para proteger el derecho a la protección de datos de los interesados, atendiendo al tipo de datos que se traten y a las finalidades de los distintos tratamientos dirigidos a la ejecución de la pena, [...].

Y, en tercer lugar, el Tribunal Constitucional dictó la Sentencia 164/2021, de 24 de octubre (ECLI:ES:TS:2021:164). Esta sentencia, citando la STEDH Cano Moya c. España, anuló un auto del Juzgado de Vigilancia Penitenciaria que había

¹⁹ En las prisiones españolas gestionadas no por la Generalitat de Catalunya, sino por el Ministerio del Interior, también se utiliza un algoritmo predictivo, aunque solo para una cuestión concreta: evaluar el riesgo de incumplimiento de los permisos de salida. Se denomina tabla de variables de riesgo, o TVR, y ha sido ampliamente criticado por estar desactualizado y contener factores discriminatorios e injustos para los presos. La Instrucción 13/2019 no contiene ninguna previsión específica sobre cuál sería el acceso que tendrían los internos a los resultados de este algoritmo.

denegado al interno el derecho a obtener copia de los informes técnicos emitidos por el educador, el psicólogo y el trabajador social, en base a los cuales se le habían denegado los permisos que solicitaba. La Audiencia no cuestiona en esta sentencia los motivos por los que se puede denegar el acceso –recogidos en la Instrucción 13/2019 antes citada–, sino que señala que su aplicación a cada caso concreto debe basarse en las circunstancias específicas y particulares de este, y anula el auto del juez por ser una resolución estereotipada que no cumple este requisito.

Además, el Tribunal Constitucional también indica que un mecanismo alternativo que ayuda a los reclusos a acceder a sus expedientes es el derecho fundamental a la protección de datos. Los titulares de datos personales tenidos en cuenta o introducidos en los distintos sistemas algorítmicos tienen una serie de derechos. Estos derechos son esencialmente, como ya hemos referido anteriormente:

- Derecho a obtener determinada información sobre el tratamiento de datos realizado por estos sistemas algorítmicos (art. 21 de la Ley Orgánica 7/2021).
- El derecho de acceso a los datos personales utilizados por los sistemas de tratamiento y a los resultados generados por los sistemas algorítmicos. Este último no está expresamente reconocido por la norma, pero creemos que también es posible solicitarlo (art. 22 de la Ley Orgánica 7/2021).
- Derecho de rectificación y supresión de datos personales (art. 23 de la Ley Orgánica 7/2021).
- Ha de tenerse en cuenta que el ejercicio de estos derechos puede verse limitado por una serie de intereses generales, previa ponderación de estos.

Podría pensarse, por tanto, que se ha abierto una nueva etapa en relación con el derecho de los internos a acceder a toda la información que sobre ellos conste en el expediente penitenciario y en el protocolo de personalidad, y que a partir de ahora el acceso pleno debe ser la regla general y la denegación muy excepcional, y siempre debidamente motivada en relación con las circunstancias específicas que puedan concurrir en algunos casos concretos. Y, aunque no existe un mecanismo específico para acceder a la información de entrada y salida de RisCanvi, el derecho general que asiste a los internos para acceder a sus expedientes (bien a través del artículo 15,2 de la Ley Orgánica General Penitenciaria, bien a través de los derechos de Protección de Datos de Carácter Personal), debería aplicarse enteramente, en estos mismos términos, cuando se pretenda obtener información sobre las evaluaciones realizadas por RisCanvi respecto de los interesados.

Sin embargo, el efecto de estos cambios normativos y jurisprudenciales no está siendo inmediato. Aunque tras la sentencia del Tribunal Constitucional hay algunas resoluciones judiciales que son más estrictas que antes con la exigencia de motivación específica para denegar el acceso, los jueces de vigilancia penitenciaria en la última reunión que mantuvieron para unificar criterios, en mayo de 2022, ratificaron sin cambio alguno el criterio que venían siguiendo desde 2018 sobre el derecho de visita y, sin citar siquiera el nuevo artículo 15 bis de la Ley Orgánica General Penitenciaria, afirman respecto al artículo 15,2 que «los internos tienen derecho a ser informados de su situación penal y penitenciaria, pero no un derecho de acceso directo al contenido del expediente penitenciario, sin perjuicio del acceso a

los informes que obren en el procedimiento ante el Juez de Vigilancia Penitenciaria en los de la Ley Orgánica del Poder Judicial».

En nuestra opinión, sin embargo, probablemente con el paso del tiempo se vaya generalizando un planteamiento del problema en la línea de lo apuntado en la sentencia del Tribunal Constitucional 164/2021, y por tanto con mayores garantías de acceso a los datos que la Administración penitenciaria tiene sobre el interno, lo que debería conllevar también el derecho de acceso a los datos e información contenidos en algoritmos como RisCanvi. Por otra parte, creemos que la propia naturaleza de RisCanvi, como herramienta algorítmica que automatiza las valoraciones de las condiciones y situaciones de los internos, elimina o al menos minimiza algunos de los tradicionales riesgos de seguridad asociados a la emisión de valoraciones subjetivas realizadas por el personal penitenciario. La primera razón contenida en la citada Instrucción 13/2019 para denegar el acceso es que los internos podrían mostrarse agresivos con los funcionarios por lo manifestado en sus informes, pero dado que RisCanvi despersonaliza las valoraciones, que son realizadas por la «máquina» con la información recogida por todos los miembros del equipo multidisciplinar, es más difícil que el interno señale a una persona concreta contra la que dirigir su ira o deseo de venganza, por lo que resulta aún más difícil justificar este tipo de restricciones en el acceso de las valoraciones, que irían claramente contra lo dispuesto en el actual marco legal español.

5. FUENTES

- ANDRÉS-PUEYO, Antonio, Karin ARBACH-LUCIONI y Santiago REDONDO (2018): «The RisCanvi: A New Tool for Assessing Risk for Violence in Prison and Recidivism», en J. P. Singh et al. (Eds.): *Handbook of Recidivism Risk / Needs Assessment Tools*, Wiley-Blackwell, pp. 255-267.
- ANDREWS, D. y James BONTA (2006): *The Psychology of Criminal Conduct* (4.^a ed.), Cincinnati (EE. UU.): Anderson Publishing Co.
- BELLIO LÓPEZ-MOLINA, Naiara (2021): «In Catalonia, the RisCanvi algorithm helps decide whether inmates are paroled», *Algorithmwatch*, en línea: <<https://algorithmwatch.org/en/RisCanvi/>> (consulta: 25/05/21).
- BRIZ-REDÓN, Álvaro y Francisco MONTES (2022): «Análisis de los resultados del protocolo RisCanvi. Informe técnico». Departamento de Estadística e Investigación Operativa. Universitat de València, en línea: <<https://www.uv.es/montes/informe%20RisCanvi/informe.pdfRisCanvi/informe.pdf>>.
- CAPDEVILA CAPDEVILA, Manel (coord.) (2022): «La reincidencia en las excarcelaciones de alto riesgo (2014-2016)». Investigación propia, Centro de Estudios Jurídicos y Formación Especializada, Generalitat de Catalunya (el informe completo puede descargarse en castellano y catalán, así como un informe ejecutivo en inglés, en línea: <<https://cejfe.gencat.cat/ca/recerca/cataleg/crono/2022/excarceracions-alt-risc/index.html>>).
- CAPDEVILA CAPDEVILA, Manel (coord.); Marta BLANCH SERENTILL; Marta FERRER PUIG; Antonio ANDRÉS-PUEYO; Berta FRAMIS FERRER; Núria COMAS LÓPEZ; Albert GARRIGÓS BOU; Ariadna BOLDÚ PEDRO; Ares BATLLE MANONELLES y Judit MORA ENCINAS (2015): «Tasa de reincidencia penitenciaria 2014», Centro

de Estudios Jurídicos y Formación Especializada, Generalitat de Catalunya (el informe completo puede descargarse en castellano y catalán, así como un informe ejecutivo en inglés, en: <<http://cejfe.gencat.cat/ca/recerca/catalog/crono/2015/reincidencia-2014/>>).

El Model de rehabilitació a les presons catalanes (2011): Generalitat de Catalunya, Subdirecció General de Programas de Rehabilitación y Sanidad, en línea: <https://justicia.gencat.cat/ca/departament/publicacions/llobres_per_tematica/execucio_penal_i_justicia_juvenil/> (consulta: 19/11/2022).

HERRON MULET, Claudia; Xavier CLOTET FONS y Pol COLOMER DE SIMON (2024): *Informe Tiresias. Auditoria de l'algorisme RisCanvi*, Dribia Data Research / Generalitat de Catalunya, en línea: <<https://repositori.justicia.gencat.cat/bitstream/handle/20.500.14226/1321/auditoria-algorisme-riscanvi-informe-final.pdf>>.

KARIMI-HAGHIGHI, Marzieh y Carlos CASTILLO (2021): «Enhancing a Recidivism Prediction Tool With Machine Learning: Effectiveness and Algorithmic Fairness», ICAIL'21, 21-25 de junio, São Paulo, Brasil, en línea: <<https://doi.org/10.1145/3462757.3466150>>.

KARIMI-HAGHIGHI, Marzieh y Carlos CASTILLO (2022): «Quantitative analysis of disparate effects of RisCanvi for estimating the risk of violent recidivism», Technical Report, Web Science and Social Computing Research Group, Universitat Pompeu Fabra, en línea: <https://chato.cl/papers/karimi_haghighi_2022_quantitative_analysis_disparate_effects_RisCanvi_revi.pdf>.

Manual de aplicación del protocolo de valoración RisCanvi (2019): Generalitat de Catalunya, Departament de Justícia, Subdirecció General de Programas de Rehabilitació i Sanitat, en línea: <[https://justicia.gencat.cat/ca/ambits/reinsercio_i_serveis_penitenciaris/index.html#googtrans\(ca|es\)](https://justicia.gencat.cat/ca/ambits/reinsercio_i_serveis_penitenciaris/index.html#googtrans(ca|es))> (consulta: 19/11/2022).

MARTÍNEZ GARAY, Lucía (2016): «Errores conceptuales en la estimación de riesgo de reincidencia. La importancia de diferenciar sensibilidad y valor predictivo, y estimaciones de riesgo absolutas y relativas», *Revista Española de Investigación Criminológica* 14, en línea: <<https://reic.criminologia.net/index.php/journal/article/view/97>>.

Resum sobre RisCanvi, el protocol de valoració del risc, Generalitat de Catalunya, Departament de Justícia, en línea: <[https://justicia.gencat.cat/ca/ambits/reinsercio_i_serveis_penitenciaris/index.html#googtrans\(ca|es\)](https://justicia.gencat.cat/ca/ambits/reinsercio_i_serveis_penitenciaris/index.html#googtrans(ca|es))> (consulta: 19/11/2022).

SAURA, G. y L. ARAGÓ (2021): «El algoritmo de prisiones que no rinde cuentas a nadie», *La Vanguardia*, en línea: <<https://www.lavanguardia.com/vida/20211207/7911428/algoritmo-prisiones-rinde-cuentas-nadie.html>> (consulta: 27/12/21).

SAURA, G. y L. ARAGÓ (2021): «Un algoritmo impreciso condiciona la libertad de los presos», *La Vanguardia*, en línea: <<https://www.lavanguardia.com/vida/20211206/7888727/algoritmo-sirve-denegar-permisos-presos-pese-fallos.html>> (consulta: 27/12/21).

Tasa de reincidencia penitenciaria 2020. Investigación propia (2023): CEJFE – Área de Investigación y Formación en Ejecución Penal, en línea: <https://cejfe.gencat.cat/web/.content/home/recerca/catalog/crono/2023/taxa-reincidencia-penitenciaria/Tasa_reincidencia_penitenciaria_2020_CAST_acc_v2.pdf>.

VERIPOL

1. DESCRIPCIÓN GENERAL

Antecedentes y ámbito de aplicación

VeriPol es una herramienta informática, lanzada en 2018, que utiliza inteligencia artificial para estimar la probabilidad de que una denuncia presentada por una persona, por haber sido víctima de un robo con violencia, sea falsa. La herramienta solo se aplica en los casos de robo con violencia o intimidación (no a los robos con fuerza en las cosas), y a los hurtos cometidos por el llamado método del «tirón» (tirar repentinamente y con fuerza del bolso de la persona para arrebatarárselo).

Por el momento, el único cuerpo policial que lo utiliza es la Policía Nacional.

Ha recibido el Premio de Investigación de la Fundación Policía Española 2016-2017.¹

La hipótesis subyacente de la policía es que un elevado número de las denuncias presentadas por robos con violencia no corresponden a hechos reales, sino que son inventadas, a menudo con la intención de cobrar una indemnización de la compañía de seguros por los bienes robados.

Las razones por las que la policía asume la existencia de un elevado número de denuncias falsas en relación con los robos no están del todo claras. El artículo publicado por los diseñadores de VeriPol (Quijano-Sánchez 2018) afirma que este delito es «extremadamente común» y «generalmente llevado a cabo por ciudadanos que no tienen antecedentes penales» (p. 155). Afirman que según las estadísticas

¹ Según su propia página web (<<https://links.uv.es/vZl5PQd>>), esta fundación «es una entidad con personalidad jurídica independiente y patrimonio propio, gozando de plena capacidad jurídica y de obrar, sin ánimo de lucro, creada en el año 1998, y dedicada, casi exclusivamente, a cooperar con el Cuerpo Nacional de Policía en general y en beneficio de sus miembros en particular, proyectando su labor extraordinaria al resto de la sociedad».

oficiales «solo el 3,98 % de las denuncias por robos en España registradas en 2015 resultaron ser falsas» (p. 158), y explican que la cifra es muy diferente en las distintas zonas del país, oscilando entre el 0 y el 20 % (figura 1 en p. 156). Pero añaden que, aunque no se puede conocer la cifra real de denuncias falsas, debería ser mucho mayor, y sugieren que podría rondar el 57 %. Esta cifra resulta del siguiente razonamiento: el 81,25 % de los casos de robo registrados quedan sin resolver, y el departamento de policía con más éxito en la resolución de denuncias falsas de robo en 2015 (154 denuncias falsas probadas) experimentó una ratio de denuncias falsas de robo del 57,68 % en 2015. Si suponemos que la distribución de denuncias falsas es homogénea en todo el país, una estimación aproximada de la ratio real podría ser del 57 % (pp. 155 y 158).

Sin embargo, es cuestionable que el altísimo porcentaje de denuncias falsas que parecen suponer los creadores de VeriPol (más del 50 %) sea razonable. En nuestra opinión, el mero hecho de que más del 80 % de los robos queden sin resolver no es un argumento definitivo: podría indicar simplemente que la policía no es capaz de resolver los delitos, o que determinados tipos de robo, por sus circunstancias (comisión en un instante, asaltante desconocido, etc.) son en sí mismos difíciles de esclarecer.

En cualquier caso, según los desarrolladores de la herramienta y el comisario jefe al que entrevistamos, la introducción de VeriPol se justifica por tres razones:

- Las denuncias falsas suponen una importante merma de los recursos policiales. Estos recursos podrían utilizarse para esclarecer otras investigaciones por delitos reales.
- Cada denuncia falsa no detectada contamina las bases de datos policiales que se utilizan para realizar investigaciones, evaluar el riesgo de delincuencia en un territorio y tomar decisiones estratégicas.
- Las tasas estadísticas de denuncias falsas son relativamente elevadas.

Y VeriPol tiene dos objetivos:

- Disuadir a los ciudadanos de presentar denuncias falsas y, en consecuencia, prevenir la comisión de delitos, ya que fingir haber sido víctima de una infracción penal o denunciar una inexistente es un delito (simulación de delito) según el artículo 457 del Código Penal español.
- Evitar el uso innecesario de recursos policiales, optimizando así el esfuerzo y el tiempo en otras tareas.

Fuentes de información para este informe

Hemos obtenido menos información sobre esta herramienta algorítmica que sobre VioGén o RisCanvi. Por un lado, porque hay menos trabajos e informes publicados sobre ella (solo los dos de Quijano-Sánchez, Liberatore y Camacho-Collados, referidos en Fuentes), y también porque no hay datos públicos en ninguna web o lugar de acceso público a los que pudiéramos acceder para analizarlos por nosotros mismos. En la reunión a la que nos referimos a continuación solicitamos al comisario jefe

algunos datos, que nos enviaron una semana después; más abajo nos referiremos a ellos (véase el apartado *Eficacia del enfoque policial predictivo para prevenir la delincuencia de VeriPol*).

Según informes ocasionales que se pueden encontrar en los medios de comunicación, VeriPol sí se está utilizando en las oficinas de denuncias (por ejemplo, <<https://links.uv.es/uMM8uC3>>, fechado el 7 de agosto de 2022; o <<https://links.uv.es/Mfmj4o0>>, fechado el 30 de abril de 2022). Intentamos hablar con agentes de policía para saber más sobre el uso de VeriPol en la práctica diaria y sus resultados, pero fue muy difícil obtener información por esta vía. Solo conseguimos hablar con tres agentes (y con dos de ellos no directamente, sino a través de contactos), por lo que sus opiniones no pueden considerarse en modo alguno representativas. No obstante, y puesto que es una de las pocas fuentes de información de que hemos podido disponer, dejaremos constancia de los resultados de estas conversaciones, que cabe resumir como sigue: 1) No existe ningún protocolo o instrucción sobre cuándo o cómo utilizar VeriPol, cada agente es libre de utilizarlo o no. 2) Para uno de los agentes con los que hablamos, VeriPol es una buena herramienta, aunque son siempre los agentes quienes finalmente deciden si la denuncia es falsa o no en función de sus sospechas y de su experiencia. Los otros dos agentes con los que contactamos dijeron que VeriPol no es una herramienta fiable porque tiene un alto porcentaje de errores, y uno de ellos añadió que por este motivo no la utilizaban en su oficina. 3) A la hora de valorar si la denuncia es falsa, dan mucho más peso a las declaraciones del denunciante y a la propia experiencia de los policías que a VeriPol.

Finalmente, conseguimos tener una reunión con personas que trabajan en la Unidad de Planificación y Coordinación Estratégica de la Dirección General de la Policía Nacional. La reunión fue útil y gran parte de la información que hemos utilizado en este informe procede de lo que allí nos dijeron, así como de alguna información adicional y aclaraciones que nos facilitaron posteriormente por correo electrónico.

2. DATOS

Datos de entrada

Datos de entrenamiento

El sistema se entrenó con 1.122 denuncias de robo presentadas en España en 2015. El corpus incluye 534 denuncias verdaderas y 588 denuncias falsas. Tanto las denuncias verdaderas como las falsas pertenecen a casos policiales cerrados y verificados, en el sentido de que en las denuncias falsas había pruebas de que lo eran (en particular, porque el denunciante había reconocido los hechos), y en las consideradas verdaderas había pruebas de que el robo se había producido realmente.

El sistema analiza el texto del informe. Según los creadores de este sistema, hay determinadas palabras o expresiones presentes en un determinado informe que generan una mayor o menor probabilidad de que el informe sea falso o verdadero (Quijano-Sánchez et al., 2018: 155-168).

¿Es probable que estos datos contengan sesgos?

El sistema analiza el texto de la denuncia. El texto que entra en VeriPol no es una reproducción literal de la declaración del denunciante, con sus propias palabras y forma de expresarse, sino que es el texto que el policía redacta a partir de lo que el denunciante le cuenta que ha ocurrido. Creemos que esto es importante, porque VeriPol se basa en métodos de procesamiento del lenguaje natural para individualizar rasgos en las historias que las identifican como mentiras (el título del *paper* de Quijano-Sánchez et al. (2018) señala que su modelo ayuda a entender «cómo mentimos a la policía»), pero no analiza la historia que el eventual mentiroso está contando a la policía, sino que analiza la historia que el propio policía escribe. Y como VeriPol ha sido entrenado sobre textos escritos por los agentes, y no sobre la reproducción literal de los relatos de los denunciantes, podría ser que el sistema estuviera reproduciendo las opiniones de los agentes sobre la veracidad de las denuncias. Además, el hecho de haber sido diseñado desde el principio en función de denuncias que habían sido clasificadas como verdaderas o falsas por un agente experto podría tener como consecuencia que VeriPol no sea capaz de detectar como falsas el tipo de denuncias falsas que pasan desapercibidas para los agentes de policía experimentados y entrenados, ya que este tipo de denuncias falsas no forman parte de las denuncias falsas con las que se construyó el conjunto de datos de entrenamiento.

Comentamos a la policía que esta forma de actuar puede condicionar el funcionamiento del sistema. Por un lado, la propia impresión del policía sobre la credibilidad del atestado puede influir (de forma consciente o inconsciente) en su forma de redactar el texto que luego introducirá en la máquina, y esto introduciría un sesgo en el sistema. En otras palabras: si el informe redactado por el agente utiliza un lenguaje similar al de los informes falsos, el sistema lo detectaría como falso. Por otro lado, los agentes de policía tienen una forma bastante estereotipada de redactar las denuncias, lo que provoca una pérdida de riqueza en el texto que VeriPol analizará posteriormente. Respondieron que se trata de problemas de los que son conscientes, y que por estas razones (entre otras) el resultado de VeriPol nunca tiene un peso decisivo, sino que es siempre un agente de policía quien toma la decisión de si la denuncia es o no probablemente falsa, y qué medidas adoptar a partir de entonces.

Tememos que, aunque VeriPol no sea vinculante, saber que ha dado una cierta probabilidad de falsedad a la denuncia bien podría condicionar la decisión que el agente tome sobre cómo tratar el caso, por lo que a nuestro juicio hacer que no sea vinculante no es una solución real al problema. Una vez que la policía ha decidido implantar VeriPol, debería hacer un seguimiento exhaustivo y controlar cómo se está utilizando, así como la opinión que los agentes tienen de su utilidad, y tratar de averiguar, entre otras cosas, hasta qué punto la información que proporciona VeriPol, aunque no sea vinculante, sí condiciona la actuación de los agentes. Pero esta evaluación no se está haciendo por el momento (ver *infra*, datos de salida).

Además, el conjunto de datos no es muy grande: 1.122 denuncias, cuando cada año en España se presentan unas 30.000 denuncias por robos violentos ocurridos en la calle, y ya es un poco antigua (2015): en siete años puede haber cambiado mucho el lenguaje o las cosas que se roban con más frecuencia. Preguntamos en

la reunión de Madrid si habían comprobado que esas denuncias fueran representativas de diferentes edades, entornos socioeconómicos, diversidad geográfica y lingüística, etc., pero solo nos dijeron que debían ser representativas, sin más explicaciones. No hay ninguna garantía de que la herramienta no esté reforzando sesgos debidos a la composición del conjunto de datos inicial (véase la siguiente sección para más referencias a posibles sesgos).

Posibles problemas de representatividad de la muestra

Según Quijano-Sánchez et al. (2018: 157), todas las denuncias con las que se entrenó VeriPol habían sido anonimizadas mediante la eliminación de cualquier información personal, incluido el texto de la declaración del denunciante, y sin ninguna información adicional (por ejemplo, lugar, fecha, hora). No sabemos, sin embargo, si las nuevas denuncias introducidas en VeriPol se someten al mismo proceso: por lo que sabemos, VeriPol analiza directamente el texto del informe redactado por el agente, y este texto incluye, obviamente, todos los datos personales necesarios para identificar a la víctima, el lugar, la hora, etc.

La forma de expresarse de las personas varía mucho según la edad (la forma de hablar de los jóvenes, especialmente de los adolescentes, es muy diferente de la de las personas mayores), el nivel educativo y socioeconómico, el grado de fluidez en el idioma hablado (turistas o inmigrantes) e incluso la zona del país de la que procede la persona. Si la muestra con la que se construyó VeriPol no tuvo en cuenta estos factores, podría estar juzgando la veracidad de todas las nuevas denuncias sobre la base de un modelo que no es representativo de todas estas posibles variaciones y, por tanto, arrojar resultados inconsistentes en distintos grupos en función de la edad, entornos culturales y orígenes geográficos. En nuestra reunión con la policía señalamos este problema y nos dijeron que eran conscientes de ello. Preguntamos expresamente qué hacían con las denuncias presentadas en lenguas extranjeras o en alguna de las lenguas cooficiales que existen en algunas regiones de España: si no aplicaban VeriPol a estos casos, o si primero traducían la denuncia al castellano y luego aplicaban VeriPol a la traducción. En estos casos, el texto se traduce al español y luego se incorpora al algoritmo.² Y, en cualquier caso, recalcaron que VeriPol es solo una ayuda a la que el agente puede recurrir si lo desea, pero su resultado no es vinculante en ningún caso y la decisión siempre sigue siendo del agente.

² Esto es coherente con el hecho de que el texto analizado por VeriPol no es directamente el relato del denunciante, sino el informe redactado por el agente de policía al que se le cuenta lo sucedido. El agente redactará el informe en español, bien porque él mismo traduce lo que se le explica en otro idioma, bien porque, si él mismo no lo entiende, otra persona interpreta para el denunciante.

Datos de salida

Procesamiento de datos y resultado

El proceso funciona de la siguiente manera:

- La persona denuncia los hechos ante la policía.
- El agente de policía transcribe en un documento la denuncia presentada por el denunciante.
- Este documento se introduce en el sistema en formato PDF. VeriPol está actualmente integrado en el Sistema de Denuncias Policiales de la Policía Nacional (SIDENPOL), aplicación informática dedicada a la recepción y tramitación de denuncias y quejas.
- El funcionario que decide si se activa o no VeriPol no es el funcionario que recibió la denuncia, sino otro, perteneciente a las unidades de investigación. No es obligatorio utilizar VeriPol, depende del agente encargado de las investigaciones y, si quiere hacerlo, es tan sencillo como pulsar un botón en la pantalla del ordenador. A continuación, VeriPol procesa el texto descriptivo del informe y extrae la información necesaria para estimar la probabilidad de que el documento contenga declaraciones falsas. No existe ningún protocolo o instrucción que obligue a los agentes de policía a utilizar o no VeriPol (véase *infra*).
- Si se utiliza VeriPol, la probabilidad de que la denuncia sea falsa se obtiene inmediatamente y se expresa en porcentaje (%). Los porcentajes cercanos a 1 indican que la denuncia tiene una alta probabilidad de ser falsa, y viceversa. La denuncia se transmite siempre a la policía judicial para que continúe la investigación. Sin embargo, si VeriPol estima que existe una alta probabilidad de que la denuncia sea falsa, se tiene en cuenta, por ejemplo, para interrogar posteriormente al denunciante con más detalle para que la denuncia pueda considerarse falsa (especialmente cuando la policía tiene sospechas de que la denuncia es posiblemente falsa).

El sistema aplica primero un método de procesamiento del lenguaje natural, y luego una técnica de aprendizaje automático (regresión logística Ridge). Según los creadores de este sistema, existen ciertas palabras o expresiones que generan una mayor o menor probabilidad de que la denuncia sea falsa o verdadera (Quijano-Sánchez et al., 2018: 164-165). Los datos más relevantes para el sistema son:

- El modus operandi de la agresión: tirones o «empujones» por «detrás» de la víctima y, en general, cualquier agresión que venga por la «espalda», ataques de alguien que lleva un «casco» de moto, robos de «teléfonos móviles» caros, todo ello es característico de las denuncias falsas.
- Morfosintaxis de la denuncia: las denuncias veraces están más centradas en la dinámica del delito y en las interacciones tanto de la víctima como del agresor; tienen una forma más cercana y personal de narrar el incidente y tienden a ser más ricas en descripciones, mientras que las denuncias falsas presentan

frases que reflejan falta de información, altas ratios de negaciones (no puede recordar, no pudo ver, no pudo reconocer), parecen caracterizarse por frases más cortas y se centran más en los objetos que en la dinámica del delito.

- La cantidad de detalles descritos: las denuncias veraces aportan más detalles sobre el robo y el agresor, sobre otras personas que podrían aportar información extra, sobre el contacto con las fuerzas de seguridad justo después de la agresión.

Capacidad predictiva

La capacidad predictiva fue evaluada por los creadores del sistema en dos ocasiones: una al diseñar la herramienta y otra en un estudio piloto realizado en junio de 2017 en dos ciudades (Murcia y Málaga).

Según Quijano-Sánchez et al. (2018: 160 y ss.),

Para probar la eficacia de *VeriPol* frente a la evaluación manual de falsedad de cada denuncia registrada, dos funcionarios del Cuerpo Nacional de Policía español analizaron las denuncias del corpus. A estos dos expertos se les entregó una submuestra aleatoria del corpus compuesta por 659 denuncias [de las 1122 con las que se construyó *VeriPol*] y se les pidió que clasificaran cada denuncia como verdadera o falsa, y que evaluaran el grado de certeza de su respuesta utilizando una escala de 5 estrellas tipo Likert. Nótese que, para evitar sesgos en el experimento, los informes se barajaron aleatoriamente, se anonimizaron y se desmarcaron antes de entregarlos a los agentes.

Los resultados fueron los siguientes:

Tabla 25
Capacidad predictiva de VeriPol

	<i>Exactitud (IC 95 %)</i>	<i>Sensibilidad</i>	<i>Especificidad</i>	<i>Precisión (valor predictivo positivo)</i>
Evaluador 1	0,7596 (0,7248, 0,792)	0,8556	0,6291	0,7583
Evaluador 2	0,7255 (0,6895, 0,7594)	0,9683	0,3883	0,6873
VeriPol	0,9272 (0,9046, 0,9458)	0,9556	0,8877	0,9219

Fuente: datos extraídos de la tabla 12 de Quijano-Sánchez et al. (2018: 162).

Exactitud: número de predicciones correctas sobre el total de predicciones (TP + TN / TP + TN + FP + FN).

Sensibilidad: cuántos de los casos positivos (= documentos falsos) predijo correctamente el clasificador, sobre todos los casos positivos de los datos (TP / TP + FN).

Especificidad: cuántos de los casos negativos (= documentos verdaderos) predijo correctamente el clasificador, sobre todos los casos negativos de los datos (TN / FP + TN).

Precisión (o valor predictivo positivo): cuántas de las predicciones positivas realizadas (documentos etiquetados como falsos) son correctas (=son efectivamente falsos) (TP / TP + FP).

TP (*True Positive*) es el número de documentos falsos clasificados correctamente como falsos.

TN (*True Negative*) es el número de documentos verdaderos clasificados correctamente como verdaderos.

FN (falso negativo) es el número de documentos falsos clasificados erróneamente como verdaderos.

FP (falso positivo) es el número de documentos verdaderos clasificados erróneamente como falsos.

Los desarrolladores de la herramienta concluyen que

VeriPol supera significativamente a los expertos humanos (los IC del 95 % de VeriPol no se solapan con los de los expertos) obteniendo una mejora de más del 15 % frente a Evaluador 1 y de más del 20 % frente a Evaluador 2. El resultado es que VeriPol supera significativamente a los expertos humanos (los IC del 95 % de VeriPol no se solapan con los de los expertos). Curiosamente, Evaluador 2 consigue una Sensibilidad y *Recall* ligeramente mejores, sin embargo, esto se puede obtener clasificando como «falsos» la mayoría de los informes (Quijano-Sánchez, 2018: 161).

Y creemos que tienen razón en esta apreciación: debido a que este evaluador trabajaba de esta manera, su especificidad era mucho menor: 0,38.

El artículo de Quijano-Sánchez et al. también incluye una matriz de confusión para las clasificaciones de VeriPol en relación con toda la muestra de 1.122 informes verdaderos y falsos. Utilizando 0,5 como umbral (probabilidades superiores a 0,5 significan una denuncia falsa, probabilidades inferiores una verdadera), todos los indicadores (exactitud, especificidad, sensibilidad, valor predictivo positivo y valor predictivo negativo) oscilan entre el 90 y el 92 % (2018: 161, y tabla 14 en esta misma página). En un informe de prensa aparecido en abril de 2019, un comisario jefe a cargo de VeriPol afirmó que el nivel de acierto había mejorado desde entonces y ya se situaba en el 96 % (<<https://links.uv.es/RJYOKOL>>).

Los desarrolladores de VeriPol consideran que, de los dos posibles tipos de error, los falsos positivos (denuncias verdaderas que son erróneamente clasificadas como falsas) son los «más sensibles» (Quijano-Sánchez et al., 2018: 163 y ss.). Estamos de acuerdo, en la medida en que se trata de ciudadanos que realmente han sido víctimas de un robo con violencia, pero VeriPol marcará su denuncia como falsa. La tasa de falsos positivos es baja: de todas las denuncias verdaderas solo el 9 % se clasifican erróneamente como falsas, y como explican los autores, en estos casos de falsos positivos «la respuesta proporcionada por VeriPol es bastante incierta» porque la mayoría de estas denuncias mal clasificadas tenían una probabilidad predicha cercana a 0,5. Por eso recomiendan

que la probabilidad predicha sea interpretada por agentes que puedan ajustar el umbral de clasificación en función de su experiencia. Además, la probabilidad predicha podría utilizarse como factor de clasificación entre los informes para asignar prioridad a los casos, en lugar de para una clasificación adecuada, ya que existe un compromiso obvio entre la precisión y la tasa de falsos positivos (p. 164).

La policía ha seguido este consejo porque, como ya hemos explicado, VeriPol nunca es vinculante y el agente actuante siempre tiene la decisión final sobre si considerar falsa o no una denuncia.

Para comprobar la eficacia y efectividad de VeriPol, se realizó un estudio piloto en las zonas urbanas de Murcia y Málaga en junio de 2017. Según Quijano-Sánchez et al. (2018: 165), en cada ciudad «se enviaron dos agentes expertos en detección de denuncias falsas y en VeriPol para instalar el *software*, impartir un breve curso sobre su uso a los agentes e investigadores locales y supervisar toda la actividad. Después, todas las nuevas denuncias de robo y todos los casos de robo abiertos de 2017 fueron evaluados por VeriPol». En cuanto a los resultados obtenidos, los autores afirman:

por un lado, que

la implantación de VeriPol permitió un impresionante incremento de la productividad en cuanto a número de casos falsos de robo detectados y cerrados con éxito. De hecho, en sólo una semana se detectaron y cerraron 25 y 39 denuncias falsas de robo mientras que la media de casos falsos de robo detectados y cerrados en los meses de junio 2008-2016 es de 3,33 y 12,14, en Murcia y Málaga respectivamente.

Por otro lado, afirman que

Lo que es más importante, los ratios de éxito (es decir, el número de casos de robo falsos cerrados dividido por el número de denunciante interrogados) son del 81,58 % en Murcia y del 84,78 % en Málaga. Es importante tener en cuenta que un caso de robo se cierra con éxito como falso sólo si el denunciante confiesa su delito. Por lo tanto, estos ratios son necesariamente inferiores o iguales a la precisión real de VeriPol durante el estudio piloto, ya que potencialmente, no todos los denunciante culpables podrían haber confesado. En consecuencia, los resultados del estudio piloto permiten estimar un límite inferior de la precisión real. En concreto, VeriPol evaluó correctamente al menos el 83,54 % de las denuncias clasificadas como falsas (Quijano-Sánchez et al., 2018: 165 y ss.).

Tras considerar estos datos y las conclusiones que de ellos extraen los autores del estudio, hay que mencionar que no indican cuántos casos se evaluaron en cada ciudad, ni proporcionan una matriz de confusión con el número de falsos y verdaderos positivos/negativos para este estudio piloto. Sabemos que, de los casos que VeriPol marcó como falsos, en el 81,58 % de ellos el denunciante reconoció la falsedad en una entrevista posterior de la policía, pero no sabemos cuántos casos marcó VeriPol como falsos, ni qué porcentaje representan en relación con el número total de casos evaluados en este estudio piloto. Esta información habría sido muy interesante, ya que podría haber permitido una revisión crítica de la hipótesis asumida por los autores, según la cual el porcentaje de denuncias falsas por robo con violencia en España se sitúa en torno al 57 % de todas las denuncias por robo con violencia presentadas. Pero el documento no proporciona ninguna información al respecto.

En la entrevista que mantuvimos en Madrid preguntamos por este porcentaje, y qué razones tenía la Policía para suponer que es tan alto. Nos remitieron al artículo publicado (Quijano-Sánchez et al., 2018), pero la explicación que allí se ofrece es la que hemos relatado anteriormente (véase Antecedentes y alcance), y es, en nuestra opinión, poco satisfactoria.

Posibles sesgos

Creemos que existe el riesgo de que el resultado esté sesgado, porque VeriPol no trabaja directamente con las declaraciones hechas por el denunciante, sino con el texto escrito por la policía, y porque no se sabe si VeriPol ha tenido en cuenta las diferentes formas de hablar de los distintos grupos de personas: para una explicación detallada, véase *supra* las consideraciones que ya hicimos sobre los datos de entrada.

¿El sistema es de aprendizaje automático?

Podría serlo, ya que se basa en una tecnología de aprendizaje automático. Sin embargo, en la práctica no lo es, porque no se alimenta de las nuevas denuncias para que pueda aprender nuevos patrones o modificar las asociaciones. Para una explicación más detallada, véase la sección sobre evaluación y revisión.

Transparencia

El sistema no es transparente.

El artículo científico de Quijano-Sánchez et al. (2018) explica detalles técnicos del diseño de VeriPol y aporta información sobre resultados preliminares. Pero no hay datos públicos sobre su funcionamiento en la práctica: no se sabe cuántas veces se aplica, a qué porcentaje de denuncias, cuántas denuncias califica como falsas y cuántas como verdaderas, etc. Los datos de acceso público que existen sobre VioGén tienen muchas deficiencias, como hemos comentado anteriormente, pero, en comparación con VeriPol, VioGén parece ser un modelo de transparencia, porque no hay datos oficiales disponibles sobre VeriPol en absoluto. En algunos momentos de nuestra investigación incluso tuvimos serias dudas sobre si realmente se estaba aplicando en la práctica o no, hasta que conseguimos contactar con algunos policías, y sobre todo hasta que tuvimos la reunión con la Unidad de Planificación y Coordinación Estratégica de la Policía Nacional, que fue muy clarificadora.

La información disponible sobre VeriPol consiste principalmente en noticias de prensa. Creemos, y la reunión nos lo confirmó, que el Ministerio estaba interesado en publicitar la existencia de esta herramienta cuando se puso en marcha (2018) para que tuviera un efecto intimidatorio sobre quienes pudieran estar planteándose presentar una denuncia falsa. Además, los diseñadores de la herramienta también consideraron que podía tener este efecto: «La existencia de este sistema de apoyo a la toma de decisiones, si se publicita adecuadamente, puede ayudar a disuadir a los ciudadanos de presentar una denuncia falsa, evitando así un delito» (Quijano-Sánchez et al., 2018: 166).

Aunque no hemos hecho una búsqueda exhaustiva en internet, nos parece que la mayoría de las noticias de prensa en las que se menciona positivamente a VeriPol (por ejemplo, «La inteligencia artificial de la policía que desenmascara denuncias falsas», <<https://www.lavanguardia.com/tecnologia/20190414/461583468024/VeriPol-policia-nacional-inteligencia-artificial-algoritmo-denuncias-falsas.html>>; o «la policía no es tonta», <<https://www.elcorreoweb.es/sociedad/2020/11/24/veripol-policia-tonta-104578692.html>>)) son de los años 2018 y 2019, cuando creemos que el Ministerio hizo efectivamente ese esfuerzo de *marketing* para conseguir el efecto disuasorio deseado, mientras que a partir de 2020 y 2021, además de que las noticias son más escasas, ya hay algunas en las que aparecen opiniones críticas (por ejemplo: <<https://www.spmseguridad.com/VeriPol-poligrafo-de-la-policia-genera-dudas-por-expertos-en-etica/>> y <<https://elpais.com/tecnologia/2021-03-08/VeriPol-el-poligrafo-inteligente-de-la-policia-puesto-en-cuestion-por-expertos-en-etica-de-los-algoritmos.html>>).

Evaluación y revisión

Según nos explicaron en la reunión, los creadores de VeriPol cedieron gratuitamente el uso de la herramienta a la Policía Nacional. Sin embargo, para poder supervisar y actualizar la herramienta, piden una contraprestación. Para el pago de este canon es necesario redactar un convenio en el que se establezcan las obligaciones y compromisos de las partes y las cantidades que deben abonarse. Este convenio estaba tramitándose en el momento de cerrar este informe, pero no saben cuándo estará listo, y esta es la razón por la que VeriPol no se ha revisado ni actualizado con nuevas denuncias (salvo durante su primer año de funcionamiento, en el que se introdujeron algunas denuncias nuevas, unas mil), ni nadie ha realizado nuevos estudios para comprobar su funcionamiento. De modo que la policía no dispone de información sobre sus resultados, ni sobre si los niveles de precisión de la muestra original se mantienen cuando VeriPol se aplica diariamente en todas las comisarías.

3. EFICACIA DEL ENFOQUE POLICIAL PREDICTIVO PARA PREVENIR LA DELINCUENCIA

Respuesta de las fuerzas del orden en función del resultado

El sistema se utiliza para apoyar la decisión final. Un agente de policía que *a priori* ha recibido formación para comprender los resultados emitidos por VeriPol analiza el resultado (no obstante, véase infra más información sobre esta formación). El agente de policía es quien decide si la denuncia es falsa o no.

La decisión sobre qué hacer con la probabilidad de falsedad dada por VeriPol queda siempre en manos del agente. Si el agente cree (ya sea por la alta probabilidad dada por VeriPol, ya sea porque él/ella personalmente tiene esta impresión) que es probable que la denuncia sea falsa, el denunciante es citado de nuevo en comisaría para volver a ser interrogado sobre el caso. Según nos contaron en la reunión, es muy frecuente que en este nuevo interrogatorio el ciudadano acabe confesando la mentira.

En cualquier caso, no tenemos datos sobre cuántas veces los agentes se basan en la probabilidad que da VeriPol para decidir si vuelven a interrogar al denunciante (parece que estos datos no existen porque nadie ha realizado un seguimiento o revisión de cómo está funcionando VeriPol en la práctica). Ni siquiera sabemos con certeza con qué frecuencia se aplica VeriPol en la práctica (independientemente de lo que el agente decida hacer después); sobre esta cuestión, véase más adelante.

Probabilidad de bucles de retroalimentación, profecía autocumplida o sesgo de confirmación

En la práctica, el sistema no es de aprendizaje automático porque no se alimenta con nuevas denuncias. De modo que no disponemos de información para responder a esta pregunta. No obstante, tememos que pueda dar lugar a estos problemas por

las razones expuestas anteriormente: que el sistema no funciona con el texto de la denuncia tal y como lo cuenta/explica el denunciante con sus propias palabras, sino con un texto redactado por el agente de policía; que no sabemos si las muestras con las que se ha entrenado son representativas de los distintos grupos de población y sus diferentes formas de hablar; que VeriPol ha aprendido como denuncias «verdaderas» y «falsas» lo que los agentes de policía consideran una denuncia verdadera o falsa, etc.

Aplicación en la práctica

Ya hemos explicado anteriormente (en «Procesamiento de datos y resultado») cómo se utiliza VeriPol. Nos gustaría añadir ahora lo que se nos dijo en la reunión. La policía no ha querido hacer obligatorio el uso de VeriPol (cada agente decide si quiere ejecutarlo o no en cada caso), ni dar instrucciones obligatorias sobre cómo proceder una vez que ha dado el resultado (por ejemplo, que el agente actuante estuviera obligado a hacer más averiguaciones sobre si la probabilidad de falsedad dada por VeriPol superaba algún umbral predeterminado) por dos razones: no querían suscitar rechazos o actitudes negativas de los agentes contra VeriPol (que al parecer temen que pudieran surgir si su uso fuera obligatorio), y que son conscientes de que VeriPol necesita una evaluación y revisión para comprobar su funcionamiento, y aún no ha sido posible acometerla.

Formación de usuarios

En la reunión se nos dijo que se había impartido formación a los agentes que iban a utilizar VeriPol. Siempre los mismos expertos habían impartido formación en 21 Jefaturas de policía, por lo que se había formado a 670 policías (que después debían formar al resto de policías de sus unidades).

La formación no consistió en cómo utilizar la herramienta, ya que su funcionamiento es tan sencillo como pulsar el botón de la pantalla y el resultado aparece automáticamente. Tenía que ver con cómo realizar el interrogatorio posterior que los agentes deben hacer al denunciante si creen que hay una alta probabilidad de que la denuncia sea falsa, ya sea por el resultado de VeriPol o por su propio criterio. Nos dijeron que estas entrevistas tienen que realizarse con cuidado, los agentes tienen que ser educados y tener en cuenta que el denunciante podría ser inocente. Así que la formación impartida tenía que ver con técnicas sobre cómo detectar mentiras, pero al mismo tiempo sobre cómo realizar entrevistas teniendo en cuenta todas estas consideraciones. Nos dijeron expresamente que preferían asumir algunos falsos negativos (es decir, aceptar como verdaderas denuncias que son falsas) para evitar en la medida de lo posible tratar erróneamente como falsa una denuncia verdadera, porque eso dejaría a una víctima real sin protección de la policía.

Este enfoque cauteloso tiene sentido y es razonable mientras el conjunto de datos inicial sea pequeño, no se haya evaluado el rendimiento de la herramienta y aún no se hayan realizado comprobaciones básicas sobre posibles sesgos. Sin embargo, creemos que esta forma de hacer las cosas contrasta con la premisa básica a partir de la cual se creó VeriPol, que es que la policía parece desconfiar sistemáticamente

de muchas de las denuncias presentadas porque cree que más del 50 % de ellas son falsas. Uno de los principales objetivos de VeriPol, según sus desarrolladores, es detectar esta elevada cantidad de denuncias falsas ocultas, y si la policía cree, como ha dicho en repetidas ocasiones, que tiene una precisión muy alta, sería lógico que confiara en VeriPol cuando marca una denuncia como falsa, y tratará al denunciante en consecuencia, es decir, creyendo que miente.

Por otro lado, no estamos seguros de que esta formación se haya impartido realmente en todas las comisarías. Hemos encontrado algunas noticias de prensa según las cuales VeriPol ni siquiera se había instalado en algunas comisarías, y los agentes no habían recibido ninguna formación al respecto; véase, por ejemplo, la noticia aparecida el 30 de agosto de 2020: «La comisaría de Vigo lleva dos años sin usar la “máquina de la verdad”», según la cual «La versión que dan fuentes autorizadas de la comisaría de Vigo es que la aplicación no la usan todavía porque al no haber sido instalada no se ha dado la formación al personal que va a emplearla y nadie sabe cómo funciona» (<https://www.lavozdegalicia.es/noticia/vigo/vigo/2020/08/30/comisaria-vigo-lleva-dos-anos-usar-maquina-verdad/0003_202008V30C1991.htm>).

También queremos señalar una cuestión con la que nos hemos encontrado repetidamente al buscar información sobre VeriPol: la opinión de los policías sobre la herramienta. El trabajo de Quijano-Sánchez et al. (2018, pp. 166) informa de que cuando hicieron el estudio piloto para probar la herramienta en Murcia y Málaga, «se pidió a todos los agentes y funcionarios que participaron en el estudio piloto que respondieran a un cuestionario anónimo de forma voluntaria que abordaba su nivel de satisfacción con VeriPol y la utilidad percibida. En general, los 21 participantes coincidieron en que VeriPol es útil, fácil de usar y que debería ampliarse a otros tipos de delitos». Pero esta valoración positiva puede que no sea generalizada. La misma noticia de prensa que acabamos de citar añade que «El portavoz del sindicato UFP [Unión Federal de Policía, sindicato de policías] en Vigo, Agustín Vigo, explica que no solo no se usa en esta ciudad sino que «no se usa en muchos sitios porque al parecer [el programa] no anda muy bien. Algo pasó porque, que yo sepa, apenas se usa». Añade que «la formación es poca y, con el covid, menos». Y hemos oído opiniones similares, que expresan dudas sobre la fiabilidad de VeriPol, de algunos de los (pocos) agentes con los que hemos podido hablar para este informe o cuyas opiniones conocíamos indirectamente a través de contactos (aunque, como ya hemos subrayado, hemos hablado con muy pocos agentes y, por tanto, sus opiniones no pueden considerarse representativas).

En cualquier caso, preguntamos a las personas de la Unidad de Planificación y Coordinación Estratégica de la Policía Nacional si se había realizado alguna encuesta para conocer la opinión de los agentes sobre VeriPol desde aquella pequeñísima muestra del estudio piloto en Murcia y Málaga, y la respuesta fue que no. Añadieron que VeriPol puede ser útil sobre todo para ayudar a los agentes a evitar los sesgos que puedan tener, ya que si la herramienta da un resultado que no esperaban puede llevarlos a prestar atención a detalles o aspectos que quizás habían pasado por alto. Y también dijeron que VeriPol es más útil para los agentes que todavía tienen poca experiencia con denuncias falsas, y que los agentes con más experiencia no necesitan tanto la ayuda de VeriPol.

Sin embargo, nos tememos que esta línea de argumentación podría ser contradictoria: si VeriPol está para ayudar a superar sesgos que están instalados en la mente de los agentes, entonces probablemente los que más experiencia tienen serían los que más deberían utilizarlo, porque esa misma experiencia puede haber incorporado sesgos a lo largo de los años que podrían estar condicionando su trabajo ahora.

¿Es probable que la respuesta de las fuerzas del orden prevenga la delincuencia?

Para tratar de arrojar luz sobre esta cuestión analizaremos cuatro cuestiones distintas:

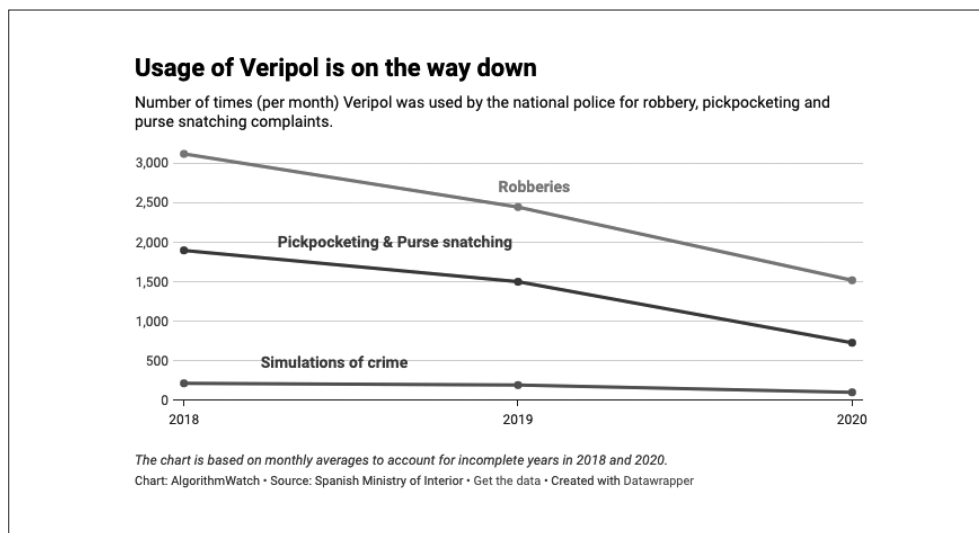
Frecuencia de uso de VeriPol y porcentaje de denuncias señaladas como falsas

Un informe de Naiara Bellio publicado en octubre de 2020 por Algorithmwatch («La policía española planea extender el uso de su detector de mentiras aunque su eficacia no está clara», en línea: <<https://algorithmwatch.org/en/spain-police-VeriPol/VeriPol/>>), que cita como fuente al Ministerio del Interior, decía que «VeriPol se ha utilizado en aproximadamente 84.000 denuncias desde su lanzamiento en octubre de 2018»; «En general, este departamento ministerial dijo que alrededor de un tercio de todas las consultas para estas categorías de delitos fue procesado por VeriPol»; «En 2019, único año completo del que se dispone de datos, VeriPol tramitó 49.702 denuncias, de las que 2.338 resultaron en un expediente por simulación de delito. La calificación de una denuncia como simulación no se basa únicamente en el resultado de VeriPol».

Según estas cifras, VeriPol señala aproximadamente una de cada veinte denuncias como simulación de delito, es decir, alrededor del 5 % (4,7 %). Si comparamos esto con el trabajo de Quijano-Sánchez et al., vemos que este porcentaje no está lejos del 3,98 % de denuncias que según las estadísticas oficiales se demostraron falsas en 2015 (2018: 158). Es decir, o bien VeriPol está refutando la suposición de los autores de que más del 50 % de las denuncias por robos violentos en España son falsas, o bien, si este alto porcentaje es correcto, VeriPol no está detectando la mayoría de las supuestas denuncias falsas.

Preguntamos a la policía si tenían datos más recientes. En la reunión dijeron que en 2022 se habían hecho 3.762 consultas en VeriPol, y que se habían detectado como falsas 511 denuncias por robo con violencia. Estas 511 son el total de denuncias falsas que se han detectado en 2022 por cualquier medio, no necesariamente VeriPol; explicaron que no pueden saber si en estos casos se ha utilizado VeriPol o no. Comparando las 3.762 denuncias tramitadas en 2022 (hasta octubre) con las 49.702 tramitadas en 2019, parece que VeriPol se está aplicando mucho menos que antes. El informe de Algorithmwatch, publicado en 2020, ya señalaba que el uso de VeriPol había ido disminuyendo desde 2018 e incluía el siguiente gráfico:

Gráfico 1
Número de veces (por mes) que VeriPol fue utilizado por la Policía Nacional para denuncias de robos, hurtos y robos de carteras



Fuente: Bellio (2020), en línea: <<https://algorithmwatch.org/en/spain-police-VeriPol/VeriPol/>> (consulta: 27/12/2021).

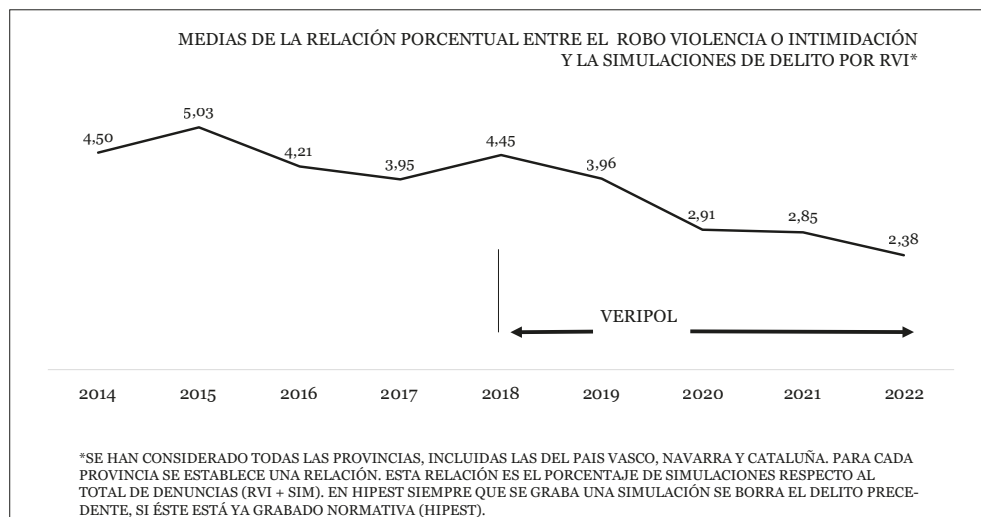
Preguntamos a la policía al respecto y nos contestaron que efectivamente su uso está disminuyendo, pero no especificaron en qué medida exactamente. Lo relacionaron con el problema de la falta de actualización de VeriPol: son conscientes de que la realidad de los hurtos y robos puede ir evolucionando con los años (mencionaron que por ejemplo los objetos más frecuentemente robados pueden ir cambiando), y que mientras que VeriPol no se actualice no quieren fomentar su uso, porque no está garantizado que su fiabilidad siga siendo la misma que al principio.

En cuanto al porcentaje de denuncias detectadas como falsas en 2022, 511 representan el 13,5 % de 3.762. Esta cifra es superior al 4,7 % detectado en 2019, según el informe de Algorithmwatch, pero aún está lejos del 57 % que los desarrolladores de VeriPol dan por falsas en el documento de 2018. Y, en cualquier caso, como no podemos saber en cuántos casos (de estos 511) se utilizó VeriPol, no se puede sacar ninguna conclusión sobre la eficacia o la precisión de la herramienta sobre la base de estas cifras.

Una semana después de la reunión, la policía nos envió algunos datos adicionales por correo electrónico, incluido el siguiente gráfico:

Gráfico 2

Medias de la relación porcentual entre robo con violencia o intimidación y delitos simulados de robo



Fuente: gráfico facilitado por la Policía Nacional. Se han considerado todas las provincias, incluidas las del País Vasco, Navarra y Cataluña. Para cada provincia se establece una relación. Esta relación es el porcentaje de simulaciones respecto al total de denuncias (robo con violencia o intimidación + simulaciones).

Hay dos cosas interesantes que señalar sobre estas cifras: por un lado, el porcentaje de denuncias falsas de robos violentos en relación con el número total de denuncias de robos violentos está disminuyendo, lo que efectivamente podría atestiguar el éxito de VeriPol como instrumento de intimidación de los posibles autores de este tipo de delitos. En segundo lugar, el porcentaje de denuncias por robo con violencia detectadas como falsas se ha mantenido entre el 2,38 y el 5 % desde 2014 (antes de la implantación de VeriPol) hasta 2022, es decir, muy lejos del 57 % supuesto en el artículo de Quijano-Sánchez et al.

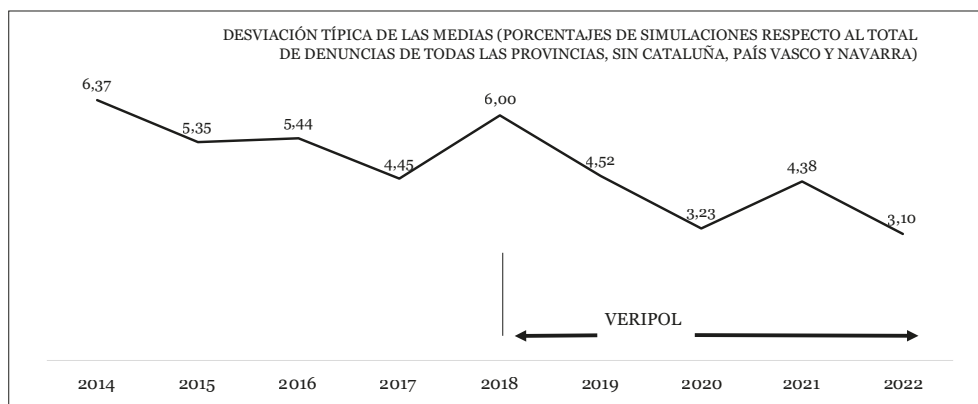
Pero es muy dudoso, en nuestra opinión, que estos datos atestigüen la eficacia de VeriPol en la prevención de robos violentos o simulaciones de delito. Por un lado, si VeriPol tiene una sensibilidad y precisión muy altas como aseguran sus desarrolladores, y detecta más del 90 % de las denuncias falsas, entonces el hecho de que tanto antes como sobre todo después de su implantación el porcentaje de denuncias detectadas como falsas sea muy bajo (entre el 2,5 y el 5 %) demostraría que ni antes ni ahora hay un problema importante de denuncias falsas en delitos de robo, y que ni antes ni ahora se malgastan recursos policiales en investigaciones que no llevarán a ninguna parte. Si, por el contrario, la policía tiene razón al creer que el porcentaje de denuncias falsas en robos con violencia es muy alto, entonces solo cabría concluir que VeriPol no está funcionando muy bien, y que su precisión no es ni mucho menos tan alta como parecía, ya que no está siendo capaz de detectar el gran número de denuncias falsas que se supone que existen.

En cuanto al efecto intimidatorio que haya podido tener VeriPol, no se puede descartar que el descenso a partir de 2018 se deba en parte a él. Dos puntos porcentuales menos pueden no parecer mucho, pero representan una disminución del 50 % en términos relativos. Sin embargo, en nuestra opinión, no está nada claro que deba atribuirse en su totalidad o en su mayor parte a un efecto intimidatorio de VeriPol, por dos razones: por un lado, porque no creemos que los ciudadanos sean tan conscientes de la existencia de esta herramienta. Es cierto que el Ministerio hizo un esfuerzo por publicitarla y darla a conocer en 2018 a través de los medios de comunicación, pero desde entonces ha pasado el tiempo y las noticias sobre su existencia son infrecuentes. Por otro lado, creemos que el hecho de que VeriPol se utilice cada vez menos con el paso de los años también implica que su existencia sea menos conocida. Y, en cualquier caso, estos datos se refieren a delitos detectados: también podría ser que los robos con violencia o la simulación de delito se sigan produciendo con una frecuencia similar, pero se detecten menos por otros motivos, o que estas cifras que nos han facilitado tengan otras inconsistencias (véase más adelante un análisis más detallado de este problema).

¿Está reduciendo VeriPol la disparidad en la detección de denuncias falsas entre las provincias españolas?

Quijano-Sánchez et al. señalaban en su trabajo de 2018 que la eficacia en la detección de denuncias falsas por robo era extremadamente heterogénea en España, oscilando entre el 0 y el 20 % en las distintas provincias. Preguntamos a la policía si sabían si la implantación de VeriPol había contribuido a nivelar estas diferencias. Dijeron que recopilarían los datos y nos enviaron posteriormente por correo electrónico el siguiente gráfico:

Gráfico 3
Desviación típica de las medias (porcentajes de simulación respecto al total de denuncias de todas las provincias, sin Cataluña, País Vasco y Navarra)



Fuente: gráfico facilitado por la Policía Nacional.

Es importante señalar que solo se nos ha proporcionado este gráfico, no las cifras brutas de las que se han derivado estos resultados. Según el gráfico, la diferencia entre provincias parece estar disminuyendo. Sin embargo, dudamos de hasta qué punto VeriPol ha contribuido a esta evolución, teniendo en cuenta que se está aplicando a un porcentaje cada vez menor de las denuncias presentadas a lo largo de los años, como se ha explicado anteriormente.

¿Está reduciendo VeriPol el número de denuncias falsas?

La eficacia de VeriPol podría demostrarse, paradójicamente, de dos maneras contradictorias. Por un lado, si aumenta el número de denuncias falsas detectadas, esto podría indicar que VeriPol funciona mejor que el criterio personal de los agentes y es capaz de detectar más casos de simulación de delito, que luego pueden ser perseguidos y sancionados. Así es como se midió la eficacia de VeriPol en el estudio piloto realizado en Murcia y Málaga en 2017, recogido en Quijano-Sánchez et al. (2018): un mayor número de denuncias falsas detectadas se consideró un éxito de VeriPol (véase *supra* 3.2.2.2 Capacidad predictiva).

Por otra parte, una reducción del número de denuncias falsas detectadas también podría interpretarse como un éxito, si suponemos que, como la gente conoce la existencia de la herramienta, se abstendrá de presentar denuncias falsas por miedo a ser descubierta y castigada. Esta es la forma en que el Ministerio ha venido presentando la eficacia de VeriPol. Por ejemplo, una noticia aparecida en abril de 2019 (<<https://links.uv.es/RJYOKOL>>) anunciaba que

Desde que llegó a todas las comisarías de la Policía Nacional el pasado mes de octubre –aún no está en Catalunya ni Euskadi, donde las policías propias tienen competencias en seguridad ciudadana– este tipo de denuncias fraudulentas ha descendido un 40 %. Aunque la Policía no quiere atribuir todo el mérito a la herramienta, que muchos denunciantes sean reiterativos o lo sea su entorno y ahora se lo piensen dos veces sí tiene mucho que ver con su implantación.

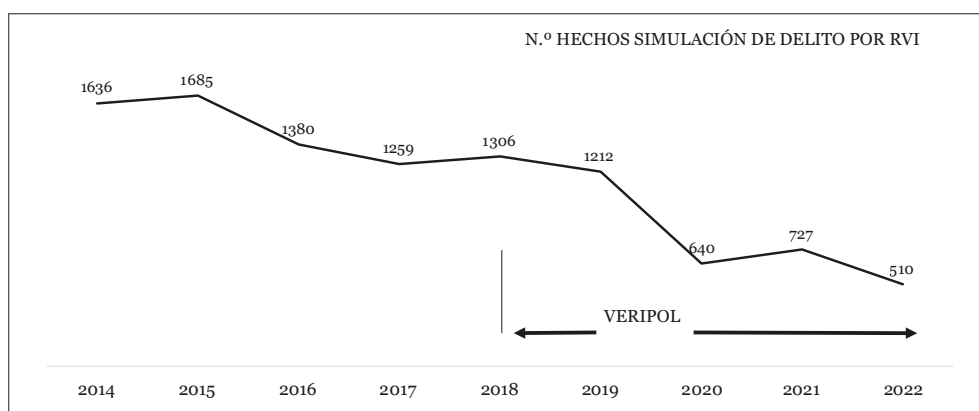
Un planteamiento muy similar encontramos en otra noticia aparecida en marzo de 2019: «El porcentaje de denuncias falsas detectadas por la Policía Nacional en España «ha disminuido un 40 % en los dos primeros meses de 2019 respecto a años anteriores», según el comisario Francisco Álvarez Fernández, jefe del Área de Seguimiento y Coordinación de la Unidad de Planificación Estratégica y Coordinación (UPEC). ¿Cuál es la razón? La presencia de Veripol. «Esto puede indicar que se ha producido un efecto disuasorio debido a la difusión que ha tenido la implantación de la herramienta. Para la Policía Nacional este efecto es el más deseado, ya que así se puede optimizar el uso de los recursos policiales y se hace una labor preventiva», añade (<<https://confilegal.com/20190326-VeriPol-la-discreta-herramienta-que-la-policia-nacional-utiliza-para-detectar-denuncias-falsas-empieza-a-dar-sus-frutos/>>).

Sin haber tenido la posibilidad de acceder directamente a datos sobre el número total de denuncias presentadas por robo con violencia en las calles cada año, el porcentaje al que se aplica VeriPol, el porcentaje de denuncias que VeriPol considera probablemente falsas, cuántas veces los agentes siguen las indicaciones de la

herramienta, el porcentaje de denunciantes que realmente confiesan, y sin haber tenido tampoco la oportunidad de entrevistar directamente a los escasos agentes que utilizan VeriPol a diario y que podrían comentar su opinión, no podemos saber cuál de estos dos posibles enfoques para evaluar la eficacia de VeriPol es el más acertado.

En la reunión con la policía preguntamos si tenían datos recientes sobre la evolución de las simulaciones de delito, y unos días después nos enviaron por correo electrónico los siguientes datos sobre esta evolución:

Gráfico 4
Número de hechos de simulación de delito por robo con violencia o intimidación



Fuente: gráfico facilitado por la Policía Nacional.

Cuando preguntamos por la fuente de los datos, se refirieron a la «Base de datos de hechos de interés policial y estadística de la Policía Nacional». Por lo que no incluye hechos conocidos por la Guardia Civil, ni por las policías locales o autonómicas.

El pronunciado descenso que se observa en 2020 hay que interpretarlo con cautela, ya que podría estar relacionado con el estricto confinamiento aplicado en España desde marzo hasta mayo debido a la pandemia de la COVID-19, que produjo una reducción de muchos tipos de delitos, en particular los que suelen ocurrir en la calle (como los robos con violencia, cuyas denuncias falsas VeriPol pretende detectar). Pero incluso teniendo esto en cuenta, según estos números la prevalencia de denuncias falsas por robos violentos habría ido disminuyendo desde 2015 al menos, y el descenso sería el más pronunciado desde 2018, año en el que se implantó VeriPol. En el correo electrónico que recibimos con estos datos la policía dice que «No podemos asegurar que la disminución de [...] delitos simulados se deba a un efecto disuasorio de VeriPol, ya que no disponemos de datos para distinguir las causas, aunque podría explicarlo al menos parcialmente».

Creemos que la policía tiene razón al no atribuir el descenso de las simulaciones de robo únicamente a VeriPol, dado que esta tendencia a la baja se viene apreciando al menos desde 2015, es decir, desde años antes de la implantación de VeriPol

en 2018, como puede verse en el gráfico. También hay que tener en cuenta que, en nuestra opinión, el supuesto efecto disuasorio de VeriPol es muy dudoso, como hemos explicado anteriormente.

Además, no disponemos de ninguna fuente con la que contrastar estos datos, ya que no existen datos públicos disponibles sobre la evolución de las simulaciones de robo conocidas por la policía. El Ministerio del Interior dispone de un portal estadístico (Portal Estadístico de Criminalidad, <<https://estadisticasdecriminalidad.ses.mir.es/publico/portalestadistico/>>) que, entre otras cosas, muestra, por tipo de delito, el número de hechos conocidos por la policía cada año. Sin embargo, es imposible obtener de este portal información sobre el número de delitos simulados que se producen cada año, porque el portal no desglosa todos los tipos de delito: el portal solo da el total de los «delitos contra la administración de justicia» conocidos por la policía cada año. Pero este es un apartado del Código Penal que incluye no solo la simulación de delito, sino también la prevaricación, el encubrimiento, el falso testimonio, el quebrantamiento de condena, etc.

Al no poder comprobar la información que nos han facilitado sobre las simulaciones de delito conocidas por la policía, hemos recurrido a otra fuente de información: el número de condenados por los tribunales por este tipo de delitos cada año.

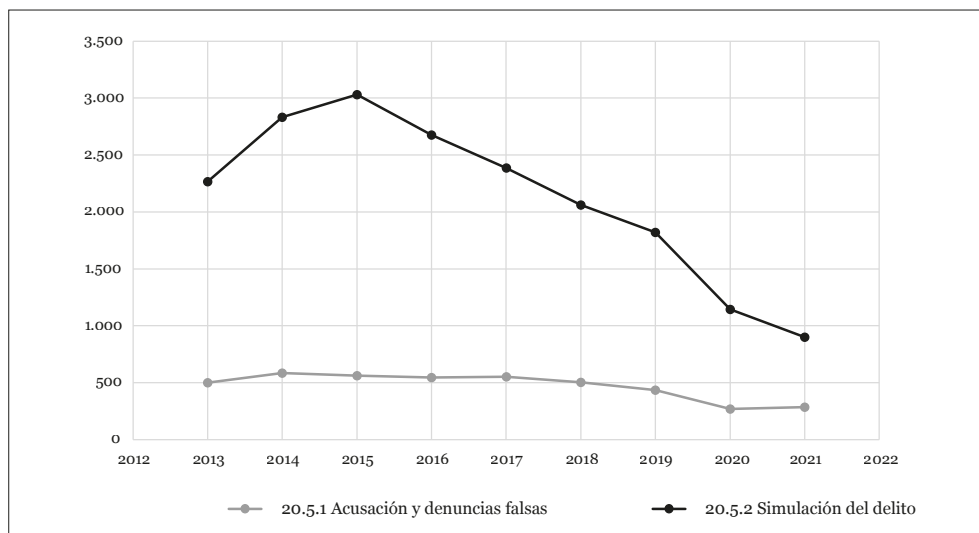
Tabla 26

Número de personas condenadas en España por acusación y denuncia falsas (art. 456 Código Penal) y simulación de delito (art. 457 Código Penal)

	2013	2014	2015	2016	2017	2018	2019	2020	2021
Acusación y denuncia falsas (art. 456 del Código Penal)	500	585	560	544	553	502	434	270	286
Simulación de delito (art. 457 del Código Penal)	2.267	2.833	3.031	2.677	2.388	2.060	1.821	1.143	899

Fuente: Instituto Nacional de Estadística (INE), en línea: <<https://www.ine.es/jaxiT3/Datos.htm?t=26017>>.

Figura 4. Número de personas condenadas en España por acusación y denuncia falsas (art. 456 Código Penal) y simulación de delito (art. 457 Código Penal)



Fuente: Instituto Nacional de Estadística (INE), en línea: <<https://www.ine.es/jaxiT3/Datos.htm?t=26017>>.

Como puede verse en la tabla y en el gráfico, las cifras siguen un patrón similar al que se muestra en el gráfico que nos envió la policía: se puede identificar una tendencia a la baja desde 2015 (tres años antes de la implantación de VeriPol), se produjo un descenso pronunciado en 2020 (que debe interpretarse en el contexto de los confinamientos por la COVID-19) y la tendencia a la baja continuó después (los datos del año 2022 aún no estaban disponibles en el momento de cerrar este informe).

Si la disminución del número de condenas por este delito apareció antes de la aplicación de VeriPol, es dudoso que la herramienta haya sido un factor significativo en esta evolución. Además, esta disminución del número de condenas está relacionada, al menos en cierta medida, con otro factor que no tiene nada que ver con VeriPol.

El Código Penal español tipifica dos delitos relativos a la acusación falsa: el artículo 456, que castiga con pena de prisión o multa la imputación a persona determinada de hechos falsos que fueran constitutivos de delito; y el artículo 457, que castiga con multa la simulación de haber sido responsable o víctima de un delito, o la denuncia de uno inexistente, «provocando actuaciones procesales».

Las denuncias falsas de robo que se realizan para obtener una indemnización de la compañía de seguros entrarían normalmente en el ámbito de aplicación del segundo artículo, ya que no identifican a ninguna persona concreta como presunto autor. Por lo tanto, si VeriPol considera que una denuncia es falsa, el denunciante podría ser acusado de cometer un delito tipificado en el artículo 457. Sin embargo, en 2015 se reformó el artículo 284 de la Ley de Enjuiciamiento Criminal, que ahora

establece que cuando en una denuncia no exista autor conocido la Policía Judicial conservará el atestado a disposición del Ministerio Fiscal y de la autoridad judicial, pero sin remitirlo a estos (existen algunas excepciones a esta regla que no son relevantes para este informe). Por tanto, cuando no hay autor conocido porque el denunciante no ha identificado a ninguna persona, que es lo que ocurre en los delitos simulados, esta denuncia falsa no «provocará actuaciones procesales», como exige el artículo 457 del Código Penal para que exista el delito, porque la denuncia no llegará al juez, y por tanto no se podrá perseguir a la persona por la comisión de este delito. Así lo ha interpretado el Tribunal Supremo en una sentencia de 2018 que sentó doctrina, y a la que han seguido varias más en el mismo sentido.³

Es decir, la implantación de VeriPol ha coincidido en el tiempo, suponemos que por casualidad, con esta reforma legal y con la posterior jurisprudencia del Tribunal Supremo, que restringe notablemente el ámbito de aplicación del delito de simulación de delito del artículo 457 del Código Penal. Podría ser que si los policías saben que no se va a perseguir el delito de simulación de robo, a pesar de que la denuncia sea falsa (por esta modificación legal y por la interpretación que los tribunales están haciendo de la misma), los agentes estén poniendo quizás menos esfuerzo en distinguir las denuncias reales de las falsas, y simplemente no tomen ninguna medida respecto a estas últimas, por lo que el número de denuncias falsas podría estar descendiendo en las estadísticas debido a este hecho: que los policías ya no las etiquetan como tales. No podemos asegurar que esto sea así, solo lo sugerimos como un factor que quizá podría explicar, al menos en parte, la tendencia a la baja no solo del número de condenados por simulación de delito, sino también del de hechos que la policía considera denuncias falsas.

Y, en cualquier caso, considerar la disminución del número de denuncias falsas como un éxito de VeriPol parece estar en contradicción con el hecho de que VeriPol solo se está utilizando en un pequeño porcentaje de las denuncias por robo con violencia, como hemos explicado anteriormente.

¿Está reduciendo VeriPol el número de denuncias por robo con violencia?

Si un porcentaje significativo de las denuncias de robos con violencia son denuncias falsas (como supone la policía), y VeriPol ha conseguido reducir la prevalencia de las denuncias falsas, es posible que el número total de denuncias por robos con

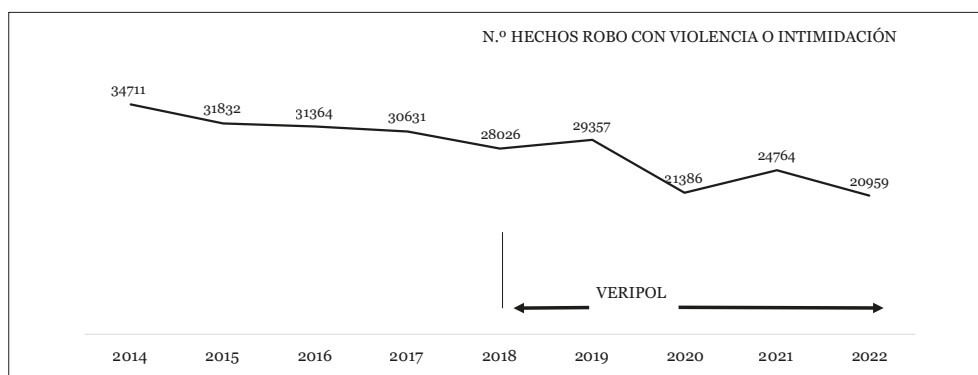
³ Esta jurisprudencia puede consultarse en:

- Tribunal Supremo. Sala de lo Penal. Fecha: 23/6/2020. Resolución: 347/2020 (ECLI:ES:TS:2020:2238) <<https://www.poderjudicial.es/search/AN/openDocument/9555aed50f121131/20200720>>.
- Tribunal Supremo. Sala de lo Penal. Fecha: 02/03/2022 Resolución: 195/2022 (ECLI:ES:TS:2022:823) <<https://www.poderjudicial.es/search/AN/openDocument/5de89019eee403d5/20220315>>.
- Tribunal Supremo. Sala de lo Penal. Fecha: 18/05/2022. Resolución: 485/2022 (ECLI:ES:TS:2022:1990) <<https://www.poderjudicial.es/search/AN/openDocument/c7c5ceb6152dc629/20220606>>.

violencia también disminuya, ya que solo quedarán las relacionadas con robos reales. Si esto es cierto, otra forma posible de comprobar la eficacia de VeriPol sería comprobar cómo ha evolucionado el índice de denuncias de robos con violencia.

En nuestra opinión, se trata de una hipótesis muy difícil de contrastar: la evolución de las denuncias por delitos de robo con violencia puede deberse a muchos otros factores distintos de la existencia de VeriPol o de su utilización, por lo que sin haber analizado previamente esas otras muchas posibles causas, establecer cualquier conexión entre la evolución de las denuncias por robo y la implantación o utilización de VeriPol nos parece una especulación sin base suficiente. No obstante, dado que la policía nos remitió también datos sobre la evolución de las denuncias por robo tras nuestra reunión en Madrid, comentaremos estos datos. El Power Point que enviaron contenía el siguiente gráfico:

Gráfico 5. Número de hechos por robo con violencia o intimidación



Fuente: gráfico facilitado por la Policía Nacional.

El gráfico 5 muestra un descenso de los robos con violencia e intimidación que se inició al menos en 2014, con una fuerte caída en 2020 que no puede atribuirse únicamente a VeriPol porque coincidió con el confinamiento por la COVID-19 como hemos explicado anteriormente, y luego una estabilización en 2021 y 2022 (y hay que tener en cuenta que los datos definitivos de 2022 serán más altos, ya que estas cifras no incluyen los robos cometidos en noviembre y diciembre).

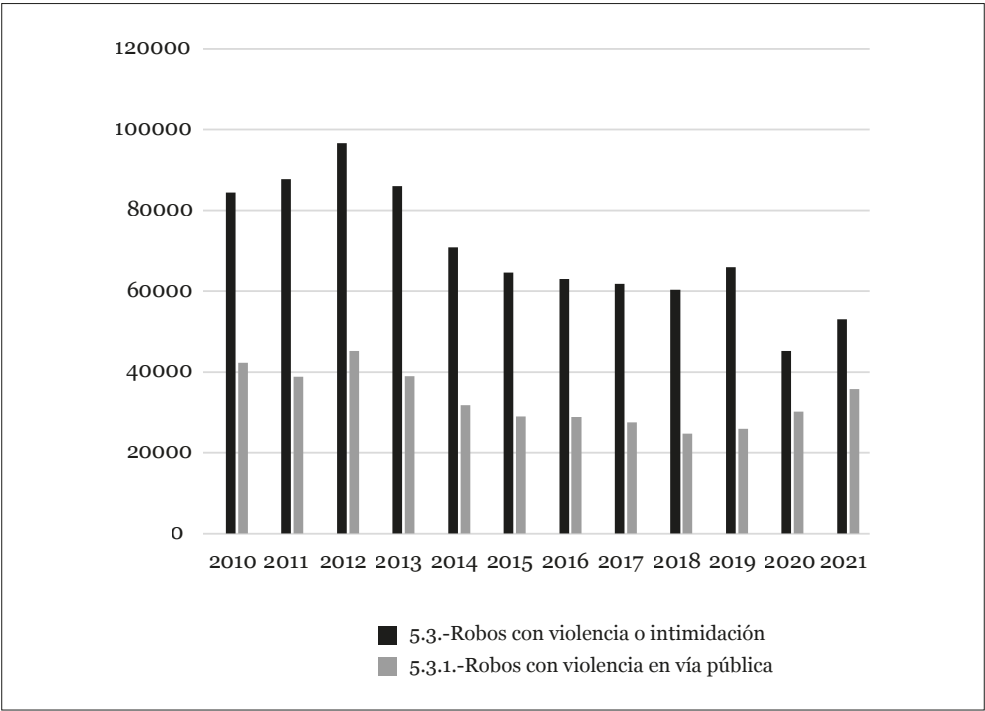
Esta misma evolución se puede encontrar en el portal estadístico del Ministerio del Interior (que en el momento de cerrar este informe disponía de datos sólo hasta 2021) para todos los robos con violencia e intimidación (las barras moradas del gráfico), pero no para los robos con violencia e intimidación en las calles, que parecen ser los únicos a los que se aplica VeriPol, que han ido subiendo desde 2018 (barras azules del gráfico):

Tabla 27
Robos con violencia o intimidación conocidos por la policía
por año en España

	2014	2015	2016	2017	2018	2019	2020	2021
5.3. Robo con violencia o intimidación	70.855	64.581	62.952	61.763	60.295	65.874	45.263	53.073
5.3.1. Robo con violencia en la vía pública	31.830	28.949	28.900	27.561	24.681	25.916	30.162	35.726

Fuente: Portal estadístico de criminalidad, Ministerio del Interior (<<https://links.uv.es/mUyT8yOmUyT8yO>>). Se computan datos provenientes de la Policía Nacional, Guardia Civil, policías autonómicas y policías locales que proporcionan datos al Sistema Estadístico de Criminalidad. No se incluyen datos de los Mossos d’Esquadra de robos con fuerza o con violencia/intimidación en establecimientos, ni de robos con violencia/intimidación en vía pública hasta 2019; a partir de 2020 se encuentran incluidos.

Figura 5
Robos con violencia o intimidación conocidos por la policía
por año en España



Fuente: Portal de estadísticas sobre delincuencia, Ministerio del Interior (<<https://links.uv.es/mUyT8yOmUyT8yO>>). Se computan datos provenientes de la Policía Nacional, Guardia Civil, policías autonómicas y policías locales que proporcionan datos al Sistema Estadístico de Criminalidad. No se incluyen datos de los Mossos d’Esquadra de robos con fuerza o con violencia/intimidación en establecimientos, ni de robos con violencia/intimidación en vía pública hasta 2019; a partir de 2020 se encuentran incluidos.

Pedimos a la policía alguna explicación sobre esta diferencia. Nos respondieron que los datos que habían enviado habían sido extraídos de la base de datos «Hechos de interés policial y estadístico de la Policía Nacional», que solo registra los delitos conocidos por este cuerpo de seguridad, mientras que la información del Portal de Estadísticas de Criminalidad del Ministerio del Interior contiene datos de la Policía Nacional, Guardia Civil, policías autonómicas y policías locales. También explicaron que «los datos de robos con violencia o intimidación en la vía pública por parte de los Mossos d’Escuadra [cuerpo policial en Cataluña] no se incluyen hasta 2020. De la Policía Nacional no se pueden aportar datos completos de las jefaturas del País Vasco, Navarra y Cataluña».

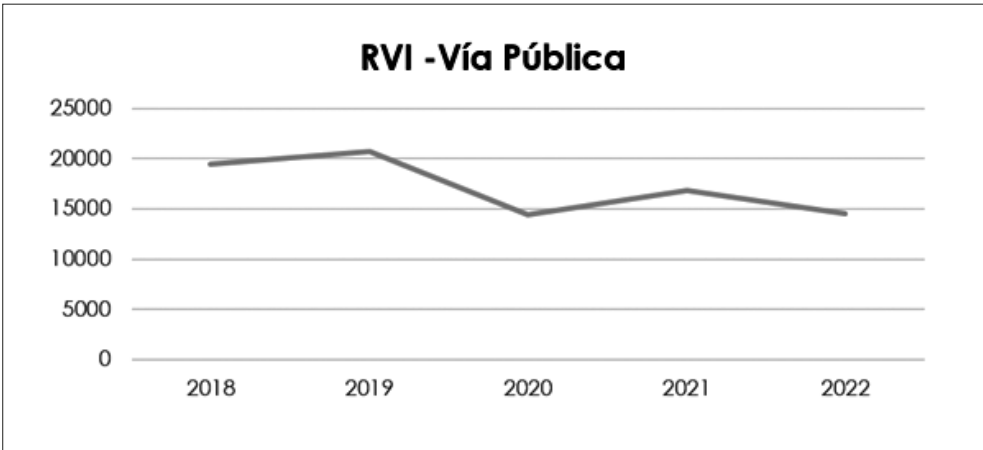
Añadieron que los datos que habían enviado una semana antes eran de todos los robos con violencia en todos los lugares, pero que filtrando por ubicación «vía pública» la tendencia sigue siendo a la baja, y enviaron la siguiente tabla y el siguiente gráfico:

Tabla 28
Número de hechos por robo con violencia o intimidación en la vía pública

	2018	2019	2020	2021	2022
RVI- Vía Pública	19,460	20,674	14,419	16,871	14,553

Fuente: cifras facilitadas por la Policía Nacional.

Gráfico 6
Número de hechos por robo con violencia o intimidación en la vía pública



Fuente: gráfico facilitado por la Policía Nacional.

No estamos del todo de acuerdo con esta conclusión: es cierto que la tendencia de estas cifras no va al alza, como ocurre con las cifras del portal estadístico (especialmente llamativa es la diferencia en 2020, año en el que estos delitos bajaron en la base de datos de la Policía Nacional, pero subieron en el portal estadístico). Pero más bien diríamos que (dejando a un lado el año 2020, que hay que interpretar en el contexto de la COVID-19) los robos con violencia en las calles conocidos por la Policía Nacional se han mantenido estables de 2020 a 2022. La cifra de 2022 ya es superior a la de 2020, y aún no incluye los robos cometidos en noviembre y diciembre.

En cualquier caso, es importante señalar que todas estas cifras tienen limitaciones muy relevantes: la propia policía, cuando envió los datos sobre los robos con violencia (gráfico 6), ya nos advirtió de que no disponía de los datos completos de algunas jefaturas en el País Vasco, Navarra y Cataluña, y de limitaciones en el portal Estadística, porque no incluye datos de robos con violencia o intimidación en la vía pública por parte de los Mossos d'Escuadra hasta 2020. Hay que tener en cuenta que Cataluña es la segunda comunidad autónoma más poblada de España, y junto con el País Vasco suman cerca del 20 % de la población española, por lo que no incluir datos o incluir datos incompletos de estas regiones conduce necesariamente a una imagen distorsionada de la criminalidad en España.

Llama la atención, por otro lado, que la policía nos advirtiera de estas limitaciones en cuanto a los datos de robos con violencia (gráfico 3), y sin embargo en otro de los gráficos que enviaron (gráfico 3) afirmen expresamente que «Se han considerado todas las provincias, incluidas las del País Vasco, Navarra y Cataluña». Nos preguntamos cómo puede ser que tengan datos completos de robos con violencia e intimidación para calcular el porcentaje de delitos simulados de robo, pero no tengan datos completos sobre la evolución de los robos con violencia. Este tipo de incoherencias en las estadísticas policiales vienen siendo puestas de manifiesto reiteradamente desde hace años por investigadores en España (véase por ejemplo el título del trabajo de Aebi y Linde (2010): «El misterioso caso de la desaparición de las estadísticas policiales españolas», *Revista electrónica de ciencia penal y criminología* 12, 2010, en línea: <<https://links.uv.es/PscvJ8t>>) y la situación no ha mejorado mucho desde entonces.

Conclusiones sobre la eficacia de la herramienta para prevenir la delincuencia

En conclusión: aunque los datos enviados por la policía apuntan a un descenso en la simulación de robos violentos desde 2018, año en el que se introdujo VeriPol, no creemos que sea seguro que VeriPol haya jugado un papel importante en ello: porque la tendencia a la baja era observable al menos cuatro años antes de la implantación de la herramienta; porque el descenso más significativo se produjo en 2020, año en el que los datos de criminalidad hay que interpretarlos con mucha cautela por el efecto del confinamiento de la población a causa de la COVID-19; porque en 2018 se produjo un cambio legal y jurisprudencial sobre la interpretación del delito de simulación de delito que también puede haber contribuido a su disminución; y, sobre todo, porque el uso de VeriPol es cada vez menor y solo se está aplicando a un pequeño porcentaje de las denuncias por robo con violencia.

En cuanto a los robos con violencia en las calles, no creemos que las cifras que obtuvimos de la policía, combinadas con lo que se puede encontrar en el portal estadístico del Ministerio del Interior, den fe de un claro descenso, ni de los hechos conocidos por el conjunto de las fuerzas policiales, ni de los conocidos solo por la Policía Nacional. Y aunque existiera, sería imposible saber hasta qué punto VeriPol ha sido un factor relevante.

4. EVALUACIÓN DEL IMPACTO SOBRE LOS DERECHOS HUMANOS

Posibilidad de sesgos y diferencias de resultados entre grupos

Como ya se ha explicado, es probable que VeriPol incluya sesgos y produzca resultados dispares para distintos grupos de población. En cuanto a los sesgos, se creó sobre la base de una clasificación de denuncias falsas y verdaderas realizada por un agente de policía experto, de modo que las denuncias falsas que este agente (o los agentes de policía en general) no ha sido capaz de detectar, VeriPol tampoco está en condiciones de reconocerlas como falsas. Además, el hecho de que no trabaje con el texto literal de la declaración del denunciante, sino con el texto que un agente de policía redacta a partir de lo que el denunciante le está contando, supone en nuestra opinión una contradicción con la premisa de partida de la herramienta: no es posible que detecte los patrones de cómo mienten los ciudadanos a la policía, mientras no trabaje con las narraciones que los ciudadanos hacen a la policía, sino con los textos que la policía elabora a partir de esas narraciones. Por último, las diferentes formas de expresarse de los distintos colectivos de la población (por nivel educativo, estatus socioeconómico, procedencia geográfica, edad o discapacidad, etc.) pueden condicionar en gran medida los resultados de una herramienta que trabaja procesando lenguaje natural, y a pesar de ello no hay constancia de que ni en la selección de las denuncias con las que se construyó el sistema, ni posteriormente en la supervisión de las denuncias sobre las que ha venido trabajando desde su implantación en las comisarías, se haya controlado si los resultados difieren en función de las características de estos distintos colectivos. En consecuencia, VeriPol podría estar discriminando sistemáticamente a algunos colectivos, sin que exista ningún control al respecto.

La importancia real de estos problemas puede ser limitada, en la medida en que según la información que hemos podido obtener parece que VeriPol se aplica cada vez menos, y que los agentes no confían mucho en sus resultados, por lo que es razonable pensar que la probabilidad de denuncia falsa que arroje VeriPol no condicionará mucho sus decisiones. Pero es imposible saber hasta qué punto esto es cierto o no, o si hay comisarías en las que la herramienta se aplica sistemáticamente a todos los casos, o en las que los agentes han decidido condicionar su actuación en función de estos niveles de probabilidad.

En nuestra opinión, la falta total de instrucciones sobre cuándo aplicar la herramienta, o cómo actuar en función de los resultados que arroja, está en contradicción con el hecho de que la policía haya hecho el esfuerzo de integrarla en el sistema

informatizado de tramitación de denuncias. Si se va a utilizar VeriPol, deberían existir directrices o protocolos que garanticen que se utiliza de forma coherente en todos los lugares, para evitar arbitrariedades o disparidades de trato. Por otro lado, el hecho de que los resultados de VeriPol no condicionen en modo alguno la actuación policial y el agente sea siempre libre de proceder como considere oportuno, aunque probablemente sea la mejor opción ante la falta de seguimiento y actualización de la herramienta, es indicativo en nuestra opinión de que la policía no confía en los resultados de VeriPol tanto como sugieren las declaraciones públicas de sus portavoces.

Evaluación de la proporcionalidad entre los beneficios preventivos y el impacto en los derechos humanos

En cuanto a la cuestión de la relación coste-beneficio, se nos dijo en la reunión que VeriPol merecía ciertamente la pena, en la medida en que no había costado dinero a la policía porque los creadores habían cedido gratuitamente su uso. En consecuencia, decían, si VeriPol detectaba una sola denuncia falsa que los agentes hubieran pasado por alto, estaría produciendo un beneficio superior a su coste, que era cero. Sin embargo, no estamos de acuerdo con esta valoración. Por un lado, aunque no haya habido que pagar ninguna licencia a los creadores para utilizarlo, implantarlo en el sistema de tramitación de denuncias de la policía sin duda habrá llevado tiempo y costado horas de trabajo a quienes han tenido que hacerlo, del mismo modo que ha sido necesario invertir tiempo y personal para impartir la formación sobre VeriPol que la policía dice haber realizado, lo que ha supuesto horas de trabajo de formadores y agentes formándose que no se han dedicado a otras tareas. Por otro lado, si finalmente se firma el convenio del que nos hablaron para que VeriPol pueda ser monitorizado y actualizado, esto supondrá un coste económico (que no cuantificaron), y que habrá que comparar con los beneficios que VeriPol podría generar.

Estos beneficios son, en nuestra opinión, muy modestos. Creemos que si para algo ha servido VeriPol ha sido precisamente para desmentir la suposición de sus creadores (y, al parecer, de la Dirección General de la Policía Nacional) de que más de la mitad de las denuncias presentadas por delitos de robo con violencia son falsas. Si realmente fuera así, VeriPol sería un fracaso, porque desde su implantación el porcentaje de denuncias detectadas como falsas se ha mantenido en los mismos niveles que antes, que ya eran muy bajos, e incluso ha disminuido. En otras palabras, VeriPol no ha conseguido sacar a la luz ese enorme número de denuncias supuestamente falsas. La policía quiere atribuir esto a un supuesto efecto disuasorio de VeriPol, que no negamos que haya podido existir quizás en las primeras fases, pero nos resulta difícil admitir que haya tenido un impacto significativo, cuando nosotros mismos, que hemos dedicado varios meses en esta investigación a buscar activamente información, hemos tenido en algunos momentos serias dudas de que la herramienta se estuviera utilizando realmente de forma habitual en las comisarías. En nuestra opinión, en cualquier caso, los datos disponibles (cuya fiabilidad es, a su vez, limitada) no avalan la eficacia de VeriPol en la prevención de delitos, ni de robos ni de simulación de delitos (estos últimos, por otra parte, no van a ser castigados por los tribunales en su inmensa mayoría, según la legislación vigente desde 2015 y la interpretación del Tribunal Supremo desde 2018).

En nuestra opinión, el éxito de VeriPol es que sus resultados cuestionan seriamente la hipótesis de que más del 50 % de las denuncias de robos con violencia son falsas, y esto debería llevar a la policía a reconsiderar si merece la pena el esfuerzo financiero de mejorar VeriPol cuando el rendimiento que puede proporcionar es muy limitado. En otras palabras, si realmente no parece haber un problema importante de denuncias falsas de robos con violencia, ¿para qué invertir dinero en una herramienta que aborde este problema? Una herramienta que además conlleva todos los peligros para los derechos humanos que ya hemos señalado: posible discriminación de colectivos cuya forma de expresarse difiere de la que ha sido mayoritaria en las denuncias utilizadas para diseñar el sistema, perpetuación de los sesgos de la propia policía, etc. Parece que la policía nunca ha considerado estas consecuencias como «costes».

El impacto sobre los derechos humanos de los falsos positivos podría parecer, a primera vista, limitado. En primer lugar, porque si los datos sobre la precisión de VeriPol dados por sus creadores en 2018 son correctos, se trata de un porcentaje pequeño (9 %); en segundo lugar, porque parece que la herramienta solo se está utilizando en un número reducido de casos; en tercer lugar, porque el hecho de que no se investigue la denuncia cuando se considera falsa no implica un resultado muy diferente respecto a lo que ocurre cuando la policía considera que la denuncia es cierta, ya que la policía no llega a esclarecer la inmensa mayoría de los robos con autor desconocido; y cuarto, porque la consecuencia más grave, que es convertir a la persona que dice haber sido víctima de un delito en sospechosa y acusada de un delito distinto (= simulación de delito) es improbable que se produzca, dada la nueva jurisprudencia del Tribunal Supremo explicada anteriormente.

Sin embargo, un examen más detallado de la cuestión sí genera motivos de preocupación. Si la herramienta se utilizara con más frecuencia (como podría ocurrir en caso de que prospere el acuerdo con sus creadores que se está negociando actualmente), ese 9 % de falsos positivos acabaría siendo un gran número de casos, teniendo en cuenta que el número de robos con violencia o intimidación en las calles ha rondado los 30.000 anuales en los últimos años. Además, como VeriPol no ha sido objeto de seguimiento desde su implantación, no podemos saber si las altas tasas de éxito comunicadas por sus creadores en 2018 se mantienen o no cuando se aplica en la práctica policial diaria, o si el número de falsos positivos sería ahora mayor. Además, si cambiara la regulación legal de la simulación de delitos o la interpretación del Tribunal Supremo respecto a este delito, las consecuencias para los falsos positivos serían mucho más graves. Y, muy importante, no hay forma de saber si los resultados difieren por sexo, raza, nacionalidad, territorio, nivel socioeconómico, etc., de los denunciantes. La base de datos, además de pequeña, fue seleccionada sin controlar estos factores, por lo que podría estar favoreciendo sesgos no detectados (o los favorecería si se utilizara la herramienta de forma más sistemática y con mayor intensidad), como ya hemos señalado anteriormente. Si así fuera, es decir, si los falsos positivos fueran más propensos a ser personas con escaso dominio del español (por bajo nivel educativo o pertenecientes a minorías que no hablan o hablan mal el español), ello podría dar lugar a que estos colectivos –que ya de por sí son bastante propensos a sufrir discriminación– perdieran aún más la confianza en la policía y en el «sistema», y tal vez no presentarían denuncias por los verdaderos robos de los que han sido víctimas.

Obligaciones activas de rendición de cuentas

En nuestra reunión con la policía se nos dijo que VeriPol no realiza tratamiento de datos personales en ningún momento. Nosotros trasladamos en esa misma reunión que, en nuestra opinión, la herramienta sí realiza tratamiento de datos personales: la denuncia que analiza VeriPol es información que puede vincularse a una persona concreta, por lo tanto, según el concepto de datos personales indicado en el RGPD, cuando VeriPol trata una denuncia está, efectivamente, tratando datos personales. Tras explicar las razones por las que creemos que VeriPol sí trata datos personales, la policía nos dijo que estudiarían esta posibilidad y consultarían con expertos dentro de la policía.

Dado que, como se ha señalado, consideramos que la normativa de protección de datos es de aplicación al tratamiento de datos que realiza VeriPol, entendemos que el Ministerio del Interior debería publicar el tratamiento de datos que realiza con la herramienta según el artículo 32.4 de la Ley Orgánica 7/2021, de protección de datos de carácter personal tratados con fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales.

En concreto, el Ministerio debería elaborar un registro de las operaciones de tratamiento de datos que realicen sobre VeriPol. La información que debería publicarse habría de incluir, entre otras cosas:

- la identificación del responsable del tratamiento y sus datos de contacto;
- los fines del tratamiento;
- la descripción de las categorías de interesados y de las categorías de datos personales; y
- el uso de perfiles.

El tratamiento de datos relacionados con el sistema VeriPol, en todo caso, y por el momento, no se publica.⁴ Como ya se ha dicho, en nuestra opinión, VeriPol lleva a cabo varias operaciones de tratamiento de datos, una de las cuales es la elaboración de perfiles, pero lo realiza sin que en ningún momento se mencione como posible situación de tratamiento de datos. El único tratamiento de datos en el que VeriPol podría considerarse integrado es el referido a SIDENPOL, pero tampoco de este modo se estaría cumpliendo con el marco legal vigente, pues según la Ley Orgánica 7/2021 es obligatorio mencionar el tratamiento específico realizado por VeriPol y, como consecuencia de ello, debería ser necesario mencionar expresamente la elaboración de perfiles, algo que no ocurre.

Por otro lado, al ser VeriPol una herramienta que utiliza nuevas tecnologías, entendemos que también le son de aplicación las previsiones del artículo 35 de la

⁴ Consulte el registro de actividades de tratamiento del Ministerio del Interior (p. 37) aquí: <<https://transparencia.gob.es/transparencia/dam/jcr:b3eb17bc-b1a7-4092-9c99-f11f62d868ae/20220922%20Registro%20de%20Actividades%20de%20Tratamiento%20del%20Ministerio%20del%20Interior.pdf>>. Registro actualizado el 22/09/2022 (consulta: 07/10/2022).

Ley Orgánica 7/2021, que establecen que es necesario realizar una evaluación de impacto sobre la protección de datos (EIPD) en estos casos. Como se ha señalado, la policía nos confirmó que no se ha realizado esta evaluación por el momento, dado que no han considerado que VeriPol trate datos personales

Acceso a los datos y posibilidad de impugnar los resultados de VeriPol

Los titulares de datos personales tenidos en cuenta o introducidos en los distintos sistemas algorítmicos, como ya se ha mencionado respecto de las otras herramientas analizadas en este informe, tienen una serie de derechos. Estos derechos son esencialmente:

- Derecho a obtener determinada información sobre el tratamiento de datos realizado por estos sistemas algorítmicos (art. 21 de la Ley Orgánica 7/2021).
- El derecho de acceso a los datos personales utilizados en el tratamiento de datos personales y a los resultados generados por sistemas algorítmicos. Este último no está expresamente reconocido por la norma, pero creemos que también es posible exigirlo (art. 22 de la Ley Orgánica 7/2021).
- Derecho de rectificación y supresión de datos personales (art. 23 de la Ley Orgánica 7/2021).
- En todo caso, y como ya se ha especificado también en el resto de los supuestos, recuérdese que el ejercicio de estos derechos puede verse limitado por una serie de intereses generales, pero no se están realizando tampoco estos juicios de necesidad y proporcionalidad al respecto.

Por otra parte, no sabemos si las personas acusadas de haber presentado una denuncia falsa pueden conocer e impugnar, en su caso, el resultado de VeriPol. Hasta la fecha, y que sepamos, ningún conflicto de esta naturaleza ha sido revisado por los tribunales.

5. FUENTES

BELLIO, Naiara (2020): «Spanish police plan to extend use of its lie-detector while efficacy is unclear», en línea: <<https://algorithmwatch.org/en/spain-police-VeriPol/>> (consulta: 27/12/21).

La Policía Nacional pone en funcionamiento la aplicación informática VeriPol para detectar denuncias falsas (2018): Policía Nacional, Ministerio del Interior del Gobierno de España, 27 de octubre, en línea: <http://www.interior.gob.es/ca/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864> (consulta: 22/12/21).

LIBERATORE, Federico; Lara QUIJANO-SÁNCHEZ y Miguel CAMACHO-COLLADOS (2018): «Applications of Data Science in Policing», *European Law Enforcement Research Bulletin* 4, SCE, pp. 89-96, en línea: <<https://orca.cardiff.ac.uk/id/eprint/135478/2/document.pdf>>.

QUIJANO-SÁNCHEZ, Lara; Federico LIBERATORE; José CAMACHO-COLLADOS y Miguel CAMACHO-COLLADOS (2018): «Applying automatic text-based detection

of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police», *Knowledge-Based Systems*, (149), pp. 155-168, en línea: <<https://doi.org/10.1016/j.knosys.2018.03.010>>.

ANEXOS SOBRE VIOGÉN

ANEXO I: DIFICULTADES PARA OBTENER INFORMACIÓN Y DATOS SOBRE VIOGÉN

En relación con esta herramienta, consideramos interesante explicar las formas en que hemos intentado obtener información sobre ella y los obstáculos que hemos encontrado.

Intentos de obtener muestras de datos e información sobre el algoritmo

Nos pusimos en contacto telefónico y por correo con la Dirección General de Coordinación y Estudios de la Secretaría de Estado de Seguridad, que finalmente accedió a una reunión que tuvo lugar en marzo de 2022. Inicialmente, la postura de nuestros interlocutores fue que ya colaboran con expertos de diferentes universidades en la revisión de VioGén, y que se han publicado varios artículos en revistas científicas revisadas por pares, lo que ya garantiza la calidad de la herramienta y su escrutinio externo. Nosotros señalamos que una cosa es la calidad de un artículo científico y otra la transparencia sobre las herramientas utilizadas por la policía y los juzgados de la jurisdicción penal para tomar decisiones, y que nunca ha habido un estudio externo independiente sobre VioGén, realizado por autores no implicados en su diseño o implementación, y se nos dijo que consultarían con sus superiores y nos responderían en unos días.

El 20 de abril de 2022 recibimos cuatro documentos de la Dirección General: dos hojas Excel con datos anonimizados (una relativa a las valoraciones policiales de riesgo y otra a las forenses), un informe confidencial sobre el diseño y sucesivas modificaciones del sistema VioGén y de los formularios de valoración policial del riesgo VPR y VPER, y un documento con un compromiso de confidencialidad. Tras la firma del compromiso y el análisis de la información que se nos remitió, comprobamos que el documento confidencial, aunque contenía abundante información sobre VioGén, básicamente explicaba cuestiones que ya son de dominio público por estar contenidas en artículos científicos, informes o tesis doctorales que ya han sido puestos a disposición del público con anterioridad. Este documento no detalla la fórmula del actual algoritmo utilizado por el VPR ni por el VPER, y las explicaciones que contiene sobre los umbrales de riesgo y los pesos de las variables corresponden

a una versión anterior del sistema, no a la vigente desde 2019. Por otro lado, la hoja Excel con datos sobre valoraciones forenses ofrecía información interesante que no habíamos podido obtener por otras vías sobre el número de estas valoraciones realizadas desde que los institutos de Medicina Legal tienen acceso a VioGén para utilizar el VFR. Sin embargo, no contenía información sobre lugares y fechas, ni sobre el tipo de medidas judiciales adoptadas tras la evaluación. Por último, la hoja Excel con datos anonimizados de las valoraciones policiales contenía solo 12 columnas con datos genéricos sobre víctima e imputado (edades, lugar, origen de la denuncia, convivencia o no, etc.), pero no contenía la codificación de ninguno de los factores de riesgo incluidos en el VPR, y la poca información que podía relacionarse con alguno de ellos (por ejemplo, si la víctima es nacional o extranjera) era inútil porque los datos relativos a esta variable no aparecían en la mayoría de las celdas. Además, aunque se mostraba el nivel de riesgo asignado por el VPR, no contenía ninguna información sobre si se habían producido o no nuevas agresiones después de la evaluación. En definitiva, no permitía ningún análisis del que pudieran extraerse conclusiones de interés sobre el funcionamiento de la herramienta.

En los meses siguientes intercambiamos algunos correos electrónicos y llamadas telefónicas. Les explicamos la imposibilidad de hacer una auditoría externa de VioGén con los datos que nos habían enviado, y el Ministerio nos explicó las dificultades que encontraba para acceder a nuestras peticiones. Por un lado, insistieron en que existe suficiente información sobre VioGén en los estudios y trabajos ya publicados. Por otra parte, mencionaron que, debido a la forma en que está construido el sistema, no es fácil extraer datos para la investigación, y también mencionaron la preocupación del Ministerio por la seguridad y la integridad del sistema VioGén frente a posibles ataques desde el exterior, y la absoluta necesidad de que el algoritmo siga siendo secreto tanto para evitar que sea hackeado como para garantizar que funcione correctamente y no pueda ser «trucado» cuando se aplique en la práctica. Discutimos algunas formas de resolver estas preocupaciones, y finalmente en junio de 2022 nos enviaron nuevos datos que consistían en lo siguiente:

- La misma hoja Excel de valoraciones forenses, pero esta vez incluyendo, como habíamos solicitado, la fecha. También habíamos pedido que se incluyera la comunidad autónoma, pero este dato no aparece.
- Otra hoja Excel con datos de evaluaciones policiales de riesgo (en el correo se decía que era una matriz «nueva», pero dado que contiene el mismo número de evaluaciones que la vez anterior y con las mismas fechas, entendemos que se trata de la misma muestra), en la que ha desaparecido la variable «localidad del suceso» y aparecen cuatro nuevas variables en cuatro columnas: número de agresores, número de víctimas, si ha habido o no una valoración de la evolución del riesgo en la que se haya anotado un incidente (VPERC), y las fechas de estas evaluaciones de seguimiento con incidente.

Como la muestra sobre evaluación del riesgo policial, por tanto, seguía sin incluir información útil, a partir de entonces centramos nuestros esfuerzos en otras fuentes de información, principalmente los pocos datos disponibles en las páginas web del Ministerio del Interior y, sobre todo, del Ministerio de Igualdad (nos referimos a

estas fuentes en detalle más arriba, apartado *Análisis de un posible sesgo contra mujeres nacidas fuera de España en las valoraciones policiales del riesgo, basado en datos públicos*).

En septiembre volvimos a ponernos en contacto con la SES, porque en el análisis que habíamos estado haciendo en julio y septiembre con los datos disponibles públicamente de las valoraciones de VioGén habíamos detectado diferencias notables entre las mujeres nacidas fuera de España y las nacidas en el país. Informamos muy brevemente de estos resultados y solicitamos si podían compartir con nosotros los datos de las valoraciones de riesgo de VioGén realizadas en los casos que habían acabado en homicidio, diferenciando por país de nacimiento. Tras tratar el tema por teléfono, recibimos una nueva muestra en octubre de 2022. Esta última muestra no contenía los datos que habíamos solicitado (las valoraciones de riesgo realizadas en los casos de víctimas mortales cuando se había presentado denuncia). En su lugar contenía, por primera vez desde que nos habíamos puesto en contacto con el Ministerio en marzo, datos sobre distintos tipos de reincidencia para más de 30.000 casos, y también la variable «víctima extranjera», pero no la codificación de ningún otro ítem del VPR o VPER, salvo la fecha de nacimiento de víctima y agresor, el nivel de riesgo y la fecha de las valoraciones. Nos referimos al contenido de esta muestra en el apartado *Análisis de un posible sesgo contra mujeres nacidas fuera de España en las valoraciones policiales del riesgo, basado en datos públicos*.

Intento de obtener información sobre evaluaciones forenses del riesgo

Nos pusimos en contacto con el Comité Técnico Científico del Consejo Médico Forense por correo electrónico en febrero de 2022. El Consejo Médico Forense es un órgano consultivo de ámbito nacional que tiene entre sus funciones la coordinación entre todos los institutos de Medicina Legal y Ciencias Forenses, y es el que ha elaborado el Protocolo de valoraciones forenses urgentes. Solicitamos los siguientes datos: cuántas valoraciones forenses de riesgo se realizan en los IMLCF, en qué casos se realizan, con qué resultado y si coinciden o no con el resultado previo de la VPR. Nos respondieron que, aunque la información que solicitábamos podría ser considerada más adelante, por el momento no lo veían factible porque el nuevo protocolo de uso de VioGén en las valoraciones forenses es de reciente implantación y no todos los IMLCF tienen acceso al sistema VioGén, por lo que aún se encuentra en una fase preliminar, a la espera de que se generalice su uso.

Les respondimos ofreciéndoles dos alternativas: que nos dieran solo información de las valoraciones realizadas por aquellos IMLCF que ya tuvieran acceso a VioGén y de los que se dispusiera de datos, o que nos dieran información de las valoraciones forenses realizadas durante la vigencia del anterior protocolo. Nos contestaron que volverían a plantear la cuestión en la próxima reunión del Comité, y en mayo de 2022 recibimos un mensaje indicando que «En respuesta a su solicitud de colaboración, indicarle que, tras revisarlo por parte del CCT, no podemos atender a la misma por no ser esta una función propia del Comité».

Por otro lado, paralelamente y con el fin de obtener información sobre las evaluaciones programadas, nos pusimos en contacto con la Asociación de Psicología Forenses de la Administración de Justicia (APF), entidad privada sin ánimo de lucro

a la que pueden pertenecer los psicólogos que trabajan en la Administración de Justicia de las distintas comunidades autónomas. La disposición de esta asociación a colaborar fue muy favorable desde el primer momento, aunque enseguida nos advirtieron de que, según su experiencia, se realizan muy pocas valoraciones de este tipo. Aceptaron enviar a sus asociados una encuesta, elaborada por nosotros, en la que se les preguntaba por el número de evaluaciones programadas que habían realizado, el resultado, la comparación con el resultado previo de VioGén y el tipo de medidas acordadas por el órgano judicial tras la valoración programada, si se conocían.

La encuesta se envió a los miembros en marzo, pero no se recibió ninguna respuesta, a pesar de haberseles enviado recordatorios en varias ocasiones. La dirección de la Asociación interpreta este hecho como un indicio de que no se están realizando evaluaciones de riesgo programadas, si no es en casos muy residuales. Y estamos de acuerdo con ellos: esta falta de respuesta no prueba que nunca se hayan realizado evaluaciones programadas, ya que no todos los psicólogos que trabajan en la Administración de Justicia pertenecen a la citada asociación (y además hay otros factores que afectan a las tasas de respuesta a las encuestas, como olvidos, falta de tiempo o motivación, etc.), pero compartimos la opinión de la dirección de la Asociación de que, si existen, probablemente sean muy pocas. En cualquier caso, y dado que no tenemos constancia de ninguna, no hemos podido hacer ningún análisis de estas para este informe.

Ante la falta de éxito que habíamos tenido en estas dos aproximaciones, en junio cambiamos de estrategia y solicitamos entrevistas con los responsables de las unidades de valoración forense integral de los institutos de Medicina Legal y Ciencias Forenses de Alicante y Valencia. En este caso sí lo conseguimos, y se realizaron ambas entrevistas, con los resultados que se han comentado en el informe.

Por otra parte, en lo que respecta a las evaluaciones forenses, disponemos también de los datos facilitados por la Secretaría de Estado de Seguridad (véase el apartado anterior).

Intento de examinar decisiones judiciales sobre medidas cautelares que utilicen la evaluación del riesgo de VioGén

Hemos intentado analizar las resoluciones judiciales que adoptan medidas cautelares en la instrucción de casos por violencia de género, para responder a la pregunta de qué relevancia dan los jueces al nivel de riesgo asignado por VioGén (ya sea el de la valoración policial, o bien el de la valoración forense, en caso de que esta exista). Sabíamos que aquí nos íbamos a encontrar con el mismo problema que ya hemos explicado en relación con RisCanvi: como estas decisiones las adopta el Juzgado de Violencia sobre la Mujer (en adelante, JVM), o, en las demarcaciones que carecen de este juzgado especial, el Juzgado de Instrucción, que son órganos unipersonales, no se incluyen en las bases de datos, y tendríamos que recurrir a las decisiones de las audiencias provinciales que resuelven los recursos interpuestos contra las decisiones del JVM. Sin embargo, mientras que las bases de datos de jurisprudencia contienen más de 3.000 sentencias sobre RisCanvi dictadas en una sola comunidad autónoma, si se introduce la palabra *VioGén* en la base de datos solo aparecen

209 sentencias para toda España (excluyendo Cataluña y País Vasco, donde no se utiliza esta aplicación). Además, entre estas 200 hay muchas que no son relevantes para nuestro análisis, ya que no tienen que ver con el uso de VioGén en las evaluaciones de riesgo, sino que la palabra se menciona en el texto, pero en relación con otras cuestiones. Por ejemplo, algunas resoluciones resuelven recursos contra el sobreseimiento provisional decretado por el JVM, donde la decisión debe adoptarse sobre la base de que existan indicios suficientes de que el hecho imputado se produjo para poder abrir juicio, pero, dado que aquí se valoran indicios de hechos pasados, el nivel de riesgo VioGén no es relevante, porque se trata de la probabilidad de que se cometan hechos futuros. En otras ocasiones, se trata de resoluciones que revisan la concesión o denegación de ayudas económicas a las víctimas o, curiosamente, de sanciones impuestas a agentes de la Guardia Civil por no cumplimentar correctamente el formulario VioGén o negarse a hacerlo.

Si reducimos las resoluciones a aquellas que sí atienden de algún modo al nivel de riesgo señalado por VioGén para tomar decisiones sobre medidas de protección para la víctima (y correlativa restricción de los derechos del imputado), el número se reduce a poco más de 100 en toda la base de datos (en concreto, 104, según la última consulta realizada en julio de 2022,). Se trata de un número llamativamente bajo, teniendo en cuenta que cada año se dictan miles de resoluciones de JVM sobre medidas de protección (según datos del CGPJ del año 2021, solo en ese año se solicitaron 37.270, de las que 133 fueron inadmitidas, 26.254 concedidas y 10.883 denegadas; cfr. Informe Anual sobre Violencia de Género, Año 2021, en línea: <<https://www.poderjudicial.es/cgpj/es/Temas/Estadistica-Judicial/Estadistica-por-temas/Datos-penales--civiles-y-laborales/Violencia-domestica-y-Violencia-de-genero/Datos-sobre-Violencia-sobre-la-mujer-en-la-estadistica-del-CGPJ/temas/Datos-penales-civiles-y-laborales/Violencia-domestica-y-Violencia-de-genero/Datos-sobre-Violencia-sobre-la-mujer-en-la-estadistica-del-CGPJ/>>). Por ello, aunque hemos incluido en el informe un breve análisis de su contenido, queremos subrayar que, en nuestra opinión, y a diferencia de lo que ocurre con las relativas a RisCanvi, en este caso son demasiado escasas como para considerarlas representativas de lo que está ocurriendo a escala nacional en los juzgados de violencia sobre la mujer.

Dado que, debido a la escasez de las sentencias de las audiencias provinciales, no es posible hacerse una idea de la relevancia dada por los tribunales a las evaluaciones de riesgo realizadas con VioGén, intentamos tener acceso directo a las resoluciones de los juzgados de Violencia sobre la Mujer, y en abril presentamos una solicitud al Tribunal Superior de Justicia de la Comunitat Valenciana para que nos autorizara a consultar una muestra de las dictadas en este territorio (nos limitamos a la Comunitat Valenciana porque el número de resoluciones dictadas a escala nacional habría excedido las posibilidades de análisis de nuestro equipo en el plazo de que disponíamos para este informe). La solicitud tropezó con un obstáculo que ya habíamos previsto: estas decisiones no han sido sometidas a un proceso de anonimización (ya que no se permite su consulta pública), por lo que darnos acceso a ellas plantearía problemas de protección de datos personales muy delicados. En nuestra opinión, la actual normativa española de protección de datos permitiría el acceso con fines estrictamente de investigación, pero, en cualquier caso, el Tribunal

Superior de Justicia remitió el asunto al Consejo General del Poder Judicial (CGPJ) para que emitiera su dictamen.

En vista de que en junio de 2022 seguíamos esperando una respuesta, y previendo que podría tardar algún tiempo y que, sobre todo, era probable que fuera negativa, enviamos otra solicitud al Tribunal Superior de Justicia de la Comunitat Valenciana para que nos autorizara a entrevistar a los jueces de violencia sobre la mujer, y así sustituir el análisis de las resoluciones por un análisis cualitativo de la opinión de quienes las dictan, basado en su experiencia profesional. La respuesta a esta petición no fue inmediata porque el Tribunal volvió a consultar con el CGPJ, pero en este caso el CGPJ contestó más rápidamente diciendo que no había ningún problema, y el 23 de julio de 2022 recibimos por fin la confirmación del Alto Tribunal para comenzar las entrevistas. Durante la última semana de julio comenzamos a contactar con los JVM de la Comunitat Valenciana (un total de 16) para agendar las entrevistas, pero agosto es el mes en el que la mayoría de los jueces disfrutan de sus vacaciones, por lo que la mayoría de ellos nos dijeron que volviéramos a llamar en septiembre. Así lo hicimos, pero septiembre tampoco es un momento muy adecuado para ellos, porque después de las vacaciones hay carga de trabajo acumulado, y algunos nos dijeron que no tenían tiempo para las entrevistas. Al final, pudimos hablar con una de las dos JVM de la provincia de Castellón, con cuatro de las siete de la provincia de Alicante y con cuatro de las siete de la provincia de Valencia, con los resultados que se han explicado en el informe.

ANEXO II: LOS FACTORES DE RIESGO DE VIOGÉN

Las evaluaciones de riesgo se realizan mediante dos formularios: el VPR (evaluación policial del riesgo) y el VPER (evaluación policial de la evolución del riesgo). El VPR se utiliza para la evaluación inicial de cada nuevo caso; el VPER se utiliza para realizar las evaluaciones posteriores. El VPER se cumplimenta cuando se produce una nueva agresión o incidente, pero también periódicamente aunque no haya habido nuevos incidentes, para supervisar la evolución del caso, con mayor o menor frecuencia en función del nivel de riesgo: en un plazo no superior a 72 horas en el nivel extremo, en un plazo de 7 días en el nivel alto, en un plazo de 1 mes en el nivel medio, en un plazo de 2 meses en el nivel bajo y en un plazo de 3 meses en el nivel no apreciado, siempre que los casos sigan activos.

Ítems VPR

Factor 1: Historia de violencia (no se limita al último episodio, sino que deben evaluarse todos los actos de violencia que hayan ocurrido y hayan sido denunciados).

Indicador 1. Vejaciones, insultos y humillaciones (violencia psicológica): manifestaciones verbales que minan la autoestima de la mujer, la desvalorizan, la humillan o

le provocan un estado de sufrimiento, tensión o miedo. El agente debe valorar si se considera leve, grave o muy grave.

Indicador 2. Violencia física: cualquier acto no accidental que cause daño físico o enfermedad a la víctima, o la ponga en grave riesgo de sufrir daño físico o enfermedad. Ejemplos: zarandeos, empujones, patadas, bofetadas, golpes, puñetazos, lanzamiento violento de objetos, mordiscos, estrangulamiento. El agente debe valorar si lo considera leve, grave o muy grave.

Indicador 3. Relaciones sexuales forzadas: toda conducta consumada o intentada del agresor en la que se presiona a la víctima para que realice prácticas sexuales no consentidas, mediante fuerza, coacción o amenazas, evitando así males mayores. El agente deberá valorar si se considera leve, grave o muy grave.

Indicador 4. Uso de armas u objetos contra la víctima.

Indicador 5. Existencia de amenazas o planes dirigidos a causar daño a la víctima. Este indicador valora la existencia de manifestaciones del agresor, a la víctima o a otras personas. El agente debe valorar si lo considera leve, grave o muy grave.

Indicador 6. En los últimos seis meses se ha producido un aumento de la escala de agresiones o amenazas.

Factor 2: Características del agresor

Indicador 7. En los últimos seis meses, el agresor muestra celos exagerados o sospechas de infidelidad.

Indicador 8. En los últimos seis meses, el agresor muestra comportamientos controladores.

El control se define como el hecho de que el agresor limite, inspeccione, vigile o registre diferentes ámbitos de la vida de la víctima.

Indicador 9. En los últimos seis meses, el agresor muestra conductas de acoso. El acoso se define como el seguimiento y persecución voluntarios, malintencionados y repetidos por parte del agresor, que amenazan la seguridad de la víctima.

Indicador 10. Existencia de problemas en la vida del agresor en los últimos seis meses, distinguiendo entre laborales/económicos (por ejemplo, despido, situación estresante en el trabajo, problemas financieros o deudas) y judiciales (no relacionados con la violencia de género).

Indicador 11. En el último año, el agresor causa daños materiales (por ejemplo, el agresor rompe objetos descargando su ira contra ellos).

Indicador 12. En el último año, se registran faltas de respeto hacia las autoridades o sus agentes (por ejemplo, el agresor hace comentarios como «la policía no sirve para nada», o en presencia de los agentes o cuando va a ser detenido, se enfrenta a ellos y los insulta).

Indicador 13. En el último año, ha agredido físicamente a otras personas o animales (incluidos otros miembros de la familia, especialmente niños).

Indicador 14. En el último año, se producen amenazas o desprecios hacia terceros (por ejemplo, el agresor insulta a un camarero por olvidarse de poner el azucarillo con el café; o el agresor desprecia a la cajera del supermercado llamándola lenta e inútil porque tarda más de lo debido en cobrarles).

Indicador 15. Existen antecedentes penales o policiales del agresor (por cualquier tipo de delito en el pasado).

Indicador 16. Existen quebrantamientos previos o actuales (cautelares o penales). Por cualquier tipo de delito (no solo por violencia de género).

Indicador 17. Existen antecedentes de agresiones físicas o sexuales. Este historial tan específico puede estar relacionado con cualquier persona del pasado.

Indicador 18. Existen antecedentes de violencia de género contra otra(s) pareja(s). Esto incluye también si el agresor tiene antecedentes en un país extranjero, independientemente de la fuente de esta información (fuente oficial o allegados del agresor, incluida la víctima).

Indicador 19. El agresor padece un trastorno mental o psiquiátrico. Se recomienda marcar este indicador, aunque el diagnóstico solo sea conocido por la víctima o por el propio delincuente, aunque el agente no haya tenido acceso al informe técnico correspondiente.

Indicador 20. El agresor presenta tentativas de suicidio o ideación suicida. Este indicador también puede ser reportado cuando el conocimiento proviene de expresiones que han sido transmitidas a terceros (o víctima).

Indicador 21. El agresor presenta algún tipo de adicción (alcohol, medicamentos o sustancias tóxicas legales o ilegales), de tal forma que esta conducta le genera problemas (incumplimiento de responsabilidades familiares o laborales, detenciones, agresiones físicas o peleas).

Indicador 22. El agresor tiene antecedentes familiares de violencia de género o doméstica. Este indicador incluye una historia o antecedentes de malos tratos (tanto denunciados como no denunciados) en la familia del agresor. Por ejemplo, existe la creencia de que el padre del agresor abusa o maltrató a la madre o al propio agresor.

Indicador 23. El delincuente tiene menos de 24 años.

Factor 3: Vulnerabilidad de la víctima

Indicador 24. Existencia de algún tipo de discapacidad, enfermedad física o psicológica grave.

Indicador 25. Víctima con ideación suicida o intentos de suicidio. Si es positivo, además de utilizar esta información para la evaluación del riesgo, debe recomendarse una derivación especializada.

Indicador 26. La víctima tiene algún tipo de adicción o conducta de abuso de sustancias (alcohol, drogas y fármacos).

Indicador 27. La víctima carece de apoyo familiar o social favorable. La víctima se encuentra sola en su situación de maltrato, no pudiendo acudir a familiares, amigos, asociaciones o grupos de apoyo.

Indicador 28. Víctima extranjera. Se valorará en función de la cultura de procedencia de la víctima y no tanto su nacionalidad actual.

Factor 4: Circunstancias relacionadas con los menores

Indicador 29. La víctima tiene menores a su cargo. En el hogar conviven hijos menores de edad (independientemente de que pertenezcan a ambos miembros de la pareja o sólo a uno de ellos). Indicar también cuándo no viven los hijos con el agresor (por separación u otras circunstancias).

Indicador 30. Existencia de violencia física o amenazas a la integridad física de los menores, que la víctima o terceras personas han percibido por parte del agresor y dirigidos a los niños, directa o indirectamente. Por ejemplo, el agresor ha dicho a la denunciante que «un día de estos acabaré con la vida de los niños y no los volverás a ver».

Indicador 31. La víctima teme por la integridad de los menores. La víctima está convencida de que el agresor puede causarles daño, y esto le produce miedo. Este es un indicador que tiene un componente más subjetivo que el anterior, mostrando este convencimiento por parte de la mujer denunciante. Por ejemplo, hay hijos menores conviviendo en el domicilio (independientemente de que sean hijos comunes o sólo hijos de la víctima de una relación anterior) y la víctima está preocupada porque puedan sufrir daños.

Factor 5. Circunstancias agravantes

Indicador 32. La víctima ha denunciado a otros autores en el pasado. Para notificar la presencia de este indicador, debe confirmarse a través de las bases de datos pertinentes marcando el indicador, independientemente de que la denuncia haya sido retirada posteriormente.

Indicador 33. Se han registrado episodios de violencia lateral recíproca.

Hay que tener especial cuidado con este indicador, dado que en ningún caso se puede preguntar literalmente a la víctima, sino que será tarea del agente inferir este tipo de conductas a partir de los datos previamente obtenidos y recogidos en el atestado policial sobre el comportamiento de la víctima antes, durante y después de los episodios de violencia del agresor. A la hora de valorar este indicador, será importante tener en cuenta aquellos casos en los que las víctimas, precisamente por su condición de víctimas, minimicen la violencia o incluso se culpabilicen a sí mismas para exculpar a sus agresores, en cuyo caso no se deberá marcar este indicador.

Si la fuente de información es el agresor, debe valorarse que su declaración sobre si la víctima agredió primero puede deberse a su intención de exonerarse y eludir responsabilidades. Este indicador no se refiere a lo que se denomina denuncias cruzadas.

Concretamente, el indicador se refiere a una situación en la que los miembros de la pareja se agreden física, sexual o psicológicamente, especialmente los primeros. Evalúa si la víctima tiende a mostrar un comportamiento violento hacia el agresor antes de que este muestre violencia hacia ella.

Por ejemplo, si una víctima insulta al agresor o incluso le agrede físicamente, y posteriormente se desarrolla un episodio de violencia de género, este indicador se evalúa como presente.

A las situaciones de este tipo de violencia se puede llegar a través de diferentes itinerarios, aunque para marcar este indicador lo realmente importante es que se haya llegado a estas situaciones conflictivas de violencia independientemente de cómo se haya llegado a la situación y sin culpabilizar a la víctima. También hay que prestar atención a aquellos casos en los que, siendo la violencia habitual (permanente y constante en el tiempo), el comportamiento de la víctima puede ser más bien una reacción defensiva en previsión de un probable nuevo acto de violencia por parte del agresor (sabe que la va a agredir y reacciona antes de que lo haga).

Indicador 34. La víctima ha manifestado al agresor su intención de romper la relación hace menos de seis meses.

Indicador 35. La víctima piensa que el agresor es capaz de agredirla muy violentamente o incluso de matarla.

Ítems VPER

Factor 1: Historia de violencia. Al igual que en el formulario VPR5.0, este conjunto de factores recoge los indicadores relacionados con los antecedentes de violencia de pareja, sin limitarse al último episodio.

Indicador 1. Vejaciones, insultos y humillaciones (violencia psicológica). El agente debe valorar si se considera leve, grave o muy grave.

Indicador 2. Violencia física. El agente deberá valorar si la considera leve, grave o muy grave.

Indicador 3. Sexo forzado. El agente deberá valorar si lo considera leve, grave o muy grave.

Indicador 4. Uso de armas u objetos contra la víctima.

Indicador 5. Existencia de amenazas o planes dirigidos a causar daño a la víctima. El agente deberá valorar si lo considera leve, grave o muy grave.

Factor 2: Incumplimiento de medidas cautelares judiciales o quebrantamiento de penas o medidas de seguridad penales, desde la última evaluación.

Indicador 6. El agresor se ha puesto en contacto telemático con la víctima. Este indicador evalúa la existencia de este tipo de contactos por parte del agresor cuando existe una medida judicial que lo impide.

Indicador 7. El agresor se ha acercado físicamente a la víctima. Este indicador valora la existencia de este tipo de contactos más directos por parte del agresor cuando existe una medida judicial que lo impide, aunque no se hayan producido episodios violentos en estas interacciones.

Indicador 8. Vulneración de dispositivos de control telemático de medidas y penas de alejamiento. Este indicador evalúa la violación «intencional» de los dispositivos de control telemático. Aunque es infrecuente, su ocurrencia es importante; por ello, si se da este indicador, debe señalarse también de forma complementaria el Indicador I-5 sobre la existencia de planes dirigidos a causar daño a la víctima, indicando la intensidad de tipo muy grave.

Factor 3: Características y comportamiento del agresor.

Indicador 9. En los últimos 6 meses, el agresor muestra celos exagerados o sospechas de infidelidad.

Indicador 10. En los últimos 6 meses, el agresor muestra conductas de control.

Indicador 11. En los últimos 6 meses, el agresor muestra conductas de acoso.

Indicador 12. El agresor está fugado (sobre él pende una orden de busca y captura) o en paradero desconocido.

Indicador 13. El agresor tiene antecedentes penales o policiales por cualquier tipo de delito en el pasado, también en el extranjero.

Indicador 14. Existen quebrantamientos de medidas cautelares o penales en relación con cualquier delito.

Indicador 15. El agresor presenta un trastorno mental o psiquiátrico. Se recomienda marcar este indicador, aunque solo se tenga constancia o sospecha de algún diagnóstico por parte de la víctima o del propio agresor, aunque el agente no haya tenido acceso aún a un informe técnico.

Indicador 16. El agresor presenta intentos de suicidio o ideación suicida.

Indicador 17. El agresor presenta algún tipo de adicción.

Indicador 18. Trámites judiciales de separación o divorcio no deseados por el agresor. Este indicador se señala cuando el agresor, directa o indirectamente, dificulta el proceso con acciones u omisiones que van más allá de un simple desacuerdo en algún aspecto legal de la separación.

Factor 4: Evidencias de comportamientos positivos del agresor

Indicador 19. El agresor se ha distanciado de la víctima. Señalar cuándo se tiene constancia de que el agresor no ha contactado con la víctima, salvo aquellas circunstancias legales derivadas de la custodia de los hijos o asuntos necesarios que lo requieran de forma natural.

Indicador 20. El agresor muestra una actitud pacífica, asume su situación con respecto a la víctima, sin deseos de venganza contra ella ni su entorno.

Indicador 21. El agresor muestra una actitud de respeto a la ley y de colaboración con los agentes. Este tipo de comportamiento se apreciará en todas las interacciones con la policía y el sistema judicial.

Indicador 22. El agresor muestra arrepentimiento por los hechos cometidos, en conversaciones con terceras personas o con los agentes; y si realiza conductas como intentar reparar el daño a la víctima.

Indicador 23. El agresor se acoge a programas de ayuda y muestra una integración adecuada en el programa de rehabilitación.

Indicador 24. El agresor cumple con el régimen de separación y cargas familiares. Por ejemplo: el delincuente abona la manutención de los hijos, cumple las normas de custodia de los hijos o de régimen de visitas.

Indicador 25. El agresor tiene una situación laboral y económica estable.

Indicador 26. El agresor cuenta con un apoyo social y familiar favorable.

Factor 5: Evidencias de comportamientos y circunstancias de la víctima.

Indicador 27. La víctima ha reanudado la convivencia con el agresor estando vigente una medida cautelar o penal de alejamiento.

Indicador 28. La víctima no desea declarar sobre episodios denunciados o, si lo ha hecho, manifiesta posteriormente su deseo de retirar la denuncia o rechazar la protección.

Indicador 29. La víctima realiza actividades que van en contra de su propia seguridad. Por ejemplo, encuentros o intentos de ponerse en contacto con el agresor; rechazar o abandonar el hogar de acogida; la víctima se quita o inhabilita el dispositivo telemático de seguridad.

Indicador 30. La víctima tiene algún tipo de discapacidad, enfermedad física o psiquiátrica grave. Esta es reconocida, ya sea física y englobando aquellas que requieren atención médica regular o medicación habitual como ceguera o sordera; o mental, como la discapacidad intelectual.

Por ejemplo, la víctima tiene un retraso mental y está incapacitada judicialmente.

Indicador 31. La víctima carece de apoyo familiar o social favorable. La víctima se encuentra sola en su situación de maltrato, sin poder recurrir a familiares, amigos, asociaciones o grupos de apoyo.

Indicador 32. La víctima muestra intentos de suicidio o ideación suicida. Hay conocimiento, ya sea por la víctima o sus allegados, de episodios de intentos de suicidio anteriores o actuales o de ideas recurrentes de acabar con su vida.

Indicador 33. La víctima tiene algún tipo de adicción o conductas de abuso al alcohol, las drogas o los fármacos.

Indicador 34. La víctima tiene a su cargo menores de edad. En el domicilio conviven hijos menores de edad (independientemente de que sean de ambos miembros de la pareja o sólo de uno de ellos). Indicar también cuándo los hijos no convivan con el agresor (por separación u otras circunstancias).

Indicador 35. Existen conflictos a causa de las responsabilidades de cuidado de los hijos. Este indicador debe tenerse especialmente en cuenta cuando existe un incumplimiento de la sentencia judicial en materia civil que regula estas responsabilidades.

Factor 6. Percepción y ajuste del riesgo.

Indicador 36. La víctima considera que su nivel actual de riesgo es: bajo, medio o alto.

Indicador 37. ¿Está usted de acuerdo con el riesgo percibido por la víctima? En el siguiente apartado de este indicador, el agente puede indicar si está de acuerdo con la percepción que la víctima tiene de su propio riesgo, si lo subestima o lo sobreestima.

ANEXO III: MEDIDAS DE PROTECCIÓN POLICIAL PARA CADA NIVEL DE RIESGO SEGÚN LA INSTRUCCIÓN 4/2019 DE LA SECRETARÍA DE ESTADO DE SEGURIDAD

Los niveles de riesgo «bajo», «medio», «alto» y «extremo» conllevarán, además de sus medidas de protección específicas, la aplicación de las previstas para los niveles anteriores que no se encuentren implícitas en ellas.

Además de las medidas obligatorias previstas en el presente anexo, para cada nivel de riesgo podrá aplicarse cualquier otra que a juicio de la unidad policial sea susceptible de aplicación.

<i>NIVEL DE RIESGO</i>	<i>MEDIDAS DE PROTECCIÓN OBLIGATORIAS</i>
NO APRECIADO	De gestión del caso en VioGén: La unidad policial especializada o la que gestione el caso se asegurará de que los datos contenidos en este son correctos y están debidamente actualizados, especialmente la identificación correcta, completa y actualizada del agresor y de la víctima, así como el contacto y domicilio de ambos y demás datos de interés. Igualmente deberá evitar y en su caso subsanar duplicidades de casos, así como de ID de víctimas y agresores principalmente, con el objeto de garantizar una adecuada protección de las primeras. De tipo informativo y asistencial: PARA LA VÍCTIMA • Las mismas, de tipo orientativo, informativo y asistencial que para cualquier otro ciudadano. Especialmente, información detallada de los recursos disponibles en su demarcación, tanto oficinas de atención a víctimas del delito (ámbito Justicia), como servicios sociales. • Facilitar recomendaciones en materia de autoprotección, pudiendo servir de orientación aquellas incluidas en el PSP y que se adapten a las circunstancias concretas de la víctima.

NIVEL DE RIESGO	MEDIDAS DE PROTECCIÓN OBLIGATORIAS
BAJO	De tipo informativo y asistencial: PARA LA VÍCTIMA: • Facilitar a la víctima números de teléfono de contacto permanente (24 horas) con las FFCCS de asignación del caso. • Facilitar información precisa sobre el servicio de teleasistencia móvil. • Derivación de la víctima a oficinas de atención a víctimas del delito (ámbito Justicia) y servicios sociales que correspondan a su lugar de residencia, donde deben informarle, sobre todo, sobre aquello que incida en su seguridad, en el caso de que proceda: puntos de encuentro, casas de acogida, etc. De tipo operativo: PARA LA VÍCTIMA: • Contactos telefónicos o personales con la víctima. En el caso de contactos personales, estos han de ser discretos, acordando con la víctima los mismos y la conveniencia de utilizar o no uniforme o vehículos con distintivos. • Indagar con la víctima para conocer momento de recepción y contenido de resoluciones judiciales que sobre el caso se hayan podido emitir en el ámbito civil (régimen de custodia y visitas, pago prestación por separación en favor de víctima/menores, adjudicación de la vivienda a la víctima, etc.) y que puedan repercutir en el grado de afectación/potencial peligrosidad del agresor, y por tanto en el contexto de la protección de la víctima. SOBRE EL AGRESOR: • Comunicación al agresor de que su caso de violencia de género está sometido a control/protección policial y, en su caso, contenido y alcance de las medidas judiciales/orden de protección acordadas. • Si el agresor tuviera licencia de armas, requerirle la entrega voluntaria de las armas (si tuviera) y permiso a la unidad policial actuante. Posteriormente, solicitar la retirada judicial de ambas. • En atención a las circunstancias del caso, acordar el acompañamiento del agresor al domicilio de la víctima, para retirar sus efectos, en el caso de que el juez acuerde su salida de este. • Revisar puntualmente la información penitenciaria en el sistema VioGén, con el objeto de conocer, si las tuviera, las salidas penitenciarias del agresor y contrastar dicha información con el centro penitenciario en aquellos supuestos en que la misma suscite dudas. • El sistema VioGén pone a disposición de los usuarios una ficha dossier, en formato imprimible, que, a juicio de la unidad policial gestora del caso, siempre para garantizar un mejor control de agresor y seguridad de víctima, puede ser difundida a unidades policiales de seguridad/patrullas de seguridad ciudadana.

<i>NIVEL DE RIESGO</i>	<i>MEDIDAS DE PROTECCIÓN OBLIGATORIAS</i>
MEDIO	De tipo operativo: PARA LA VÍCTIMA: • Entrevista personal con la víctima. • Si procede, en atención a las circunstancias del caso, efectuar traslado de la víctima para ingreso en centro de acogida. • Control ocasional de la víctima en domicilio, lugar de trabajo y centros escolares de los menores en horarios de entrada y salida. Para llevar a cabo este control es recomendable entrevista con oficinas de atención a víctima del delito (ámbito Justicia) o servicios sociales que atiendan a la víctima, al objeto de identificar modos efectivos de protección. • Acompañamiento a la víctima en actuaciones de carácter judicial, asistencial o administrativo, cuando se considere que puede existir algún tipo de riesgo para ella. SOBRE EL AGRESOR: • Instar a Fiscalía el seguimiento obligatorio del agresor a través de la asignación de un dispositivo telemático de control. • Comprobación periódica del cumplimiento por el agresor de las medidas judiciales de protección impuestas, lo que incluirá control ocasional del mismo en domicilio, lugar de trabajo u otros. • Comunicación/entrevista con el agresor en permisos penitenciarios o excarcelaciones para informarle de que su caso se encuentra sometido a control/protección policial, así como para contrastar información sobre sus datos de contacto/localización actuales.
ALTO	De tipo operativo: PARA LA VÍCTIMA: • En caso de no haberse podido localizar todavía al agresor, insistir a la víctima, para su más efectiva protección, en la posibilidad de traslado a centro de acogida, casa de un familiar o domicilio distinto. • Control frecuente y aleatorio en domicilio y lugar de trabajo de la víctima y, si procede en centros escolares de los hijos a la entrada y salida y contactos con personas de su entorno para mejor protección. SOBRE EL AGRESOR: • Control aleatorio de los movimientos del agresor y contactos esporádicos con personas que este frecuente o de su entorno.

NIVEL DE RIESGO	MEDIDAS DE PROTECCIÓN OBLIGATORIAS
EXTREMO	De tipo operativo: PARA LA VÍCTIMA: • Protección permanente de la víctima hasta que el propio agresor o sus circunstancias dejen de ser una amenaza inminente. • Si procede, vigilancia en centros escolares de los hijos de la víctima a la hora de entrada y salida. SOBRE EL AGRESOR: • Control intensivo de los movimientos del agresor hasta que este deje de ser una amenaza inminente para la seguridad de la víctima.

Plan de seguridad personalizado para cada víctima con medidas de autoprotección

Además de las medidas de protección que deben ponerse en marcha según el nivel de riesgo asignado por VioGén, la policía elabora un plan de seguridad personalizado para cada víctima de violencia de género. Este plan de seguridad personalizado se elaborará por primera vez con la primera VPER, independientemente del nivel de riesgo. Es en ese momento (tras la vista judicial, en su caso, la salida o no del agresor del domicilio, la adopción o no de medidas judiciales...) cuando los agentes actuantes conocen los datos necesarios para la correcta elaboración del PSP.

El plan de seguridad personalizado (PSP) se deberá confeccionar para cada víctima de forma individualizada, de tal manera que *TODAS y CADA UNA de las Medidas que contiene son susceptibles de DESELECCIONARSE* por los agentes intervinientes, bien el APARTADO entero, bien Medidas concretas dentro de cada APARTADO.

Los agentes actuantes deberán prestar especial atención a la hora de confeccionar el PSP *DESELECCIONANDO aquellas opciones que: NO sean posibles/factibles/asumibles, adecuadas o idóneas para la víctima en concreto, en atención a sus particulares circunstancias* (características físicas, posible discapacidad, edad, SI TRABAJA O NO, SI TIENE o NO menores a su cargo, edades de estos, tipo de vivienda que habita, habilidad y frecuencia en el uso de redes sociales, nivel formativo en general, arraigo familiar y/o social, disponibilidad económica, etc.).

El PSP *cobra fuerza y sentido como herramienta para autoprotección real de la víctima* cuando este ha sido confeccionado adecuada, rigurosa y proactivamente por el agente interviniente y las medidas de autoprotección que contiene guardan sentido y coherencia con el estilo de vida de la víctima en particular. En este sentido, el PSP se encuentra disponible en VioGén en diferentes idiomas.

Apartado 1:

PORTAR SIEMPRE UN TELÉFONO MÓVIL

- 1.1. *Cambiar el número de teléfono y solicitar a la compañía prestadora del servicio que sus datos no sean incluidos en ficheros ni guías de acceso público.*
- 1.2. *Bloquear las llamadas y resto de comunicaciones con el agresor a través de cualquier aplicación o cuenta (whatsApp, RR. SS., etc.)*
- 1.3. *No atender llamadas de números desconocidos.*
- 1.4. *Tener siempre activada la geolocalización solo en caso de que el agresor no conozca el nuevo número de móvil. Mantenga en el teléfono móvil o en papel los números de emergencia de la policía en un lugar preferente, así como el 112, 091, 062 y 092, vinculados a teclas de marcación rápida en caso de grabarlos en el teléfono móvil.*
- 1.5. *Guardar en el teléfono móvil o en papel los números de emergencia policiales en lugar preferente, así como los 112, 091, 062 y 092, vinculados a teclas de marcación rápida en el caso de grabación en el móvil.*
- 1.6. *Guardar en el teléfono móvil o en papel los teléfonos de asistencia en la demarcación: unidad de violencia, oficina de atención a víctimas, casa de acogida, selección de asociación/es de víctimas, servicios médicos, así como cualquier otro de confianza.*
- 1.7. *Instalar la app AlertCops (del Ministerio del Interior) en el teléfono móvil y explorar su manejo en la opción específica VdG. También, cualquier otra aplicación similar operativa en su demarcación.*

Apartado 2:

HACER UN USO SEGURO DE LAS NUEVAS TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN Y DE LAS REDES SOCIALES

- 2.1. *Restringir el acceso a su perfil/es.*
- 2.2. *Utilizar perfiles con «Nick» en lugar de mostrar el auténtico nombre.*
- 2.3. *No aceptar solicitudes de acceso o de amistad de desconocidos o del entorno del agresor.*
- 2.4. *Evitar difundir información personal en sus perfiles (fotografías, vídeos, localizaciones, amistades, etc.) e informar a sus contactos del riesgo de que ellos a su vez difundan dicha información.*

Apartado 3:

MEDIDAS DE AUTOPROTECCIÓN PERSONAL CON CARÁCTER GENERAL PARA TODAS LAS VÍCTIMAS

- 3.1. *Adoptar rutinas de seguridad en los desplazamientos y salidas para desarrollar cualquier actividad cotidiana.*

- 3.2. *Identificar a los vecinos/amigos de más confianza* que puedan ser contactados en caso de emergencia. Informarles de la situación y pedirles que llamen a Fuerzas y Cuerpos de Seguridad si advierten la presencia del agresor o si oyen gritos o sonidos propios de un ataque violento.
- 3.3. *No acudir a encuentros con el agresor.*
- 3.4. *Portar fotocopia de las disposiciones judiciales de protección/alejamiento.*
- 3.5. Interesarse por la realización de cualquier tipo de formación en: *defensa personal, gestión segura de redes sociales y aplicaciones móviles u otros que redunden en favor de su seguridad y autoconfianza.*
- 3.6. *En caso de cambio de domicilio, comunicarlo exclusivamente al cuerpo de seguridad encargado de su protección y, en su caso, a personas solo de su absoluta confianza.*

Apartado 4:

PLANIFICACIÓN DE RUTINA DE HUIDA EN CASO DE INTENTO DE NUEVA AGRESIÓN PARA TODAS LAS VÍCTIMAS

- 4.1. *Salir rápidamente del domicilio.*
- 4.2. Si no se puede salir, *confinarse en lugar seguro*, encerrarse en habitación segura portando teléfono móvil y llamar a las Fuerzas y Cuerpos de Seguridad.
- 4.3. *Pedir auxilio a través de ventanas* e incluso, de ser seguras, utilizarlas como vías de escape.
- 4.4. Mantenerse alejada de las habitaciones y dependencias del domicilio donde puedan guardarse *objetos contundentes o peligrosos* (cocina, garaje, gimnasio, etc.).
- 4.5. *Tener preparada una bolsa de emergencia*, por si es necesario abandonar el domicilio precipitadamente y esconderla en lugar seguro (la casa de un vecino o amigo), evitando el entorno de familia y amigos del agresor. La bolsa podría incluir, para la víctima y, en su caso, menores a su cargo: documentos esenciales (de identidad, relacionados con la vivienda, etc.) tarjetas bancarias o sanitarias, llaves (domicilio, vehículo, lugar de trabajo) dinero en efectivo, medicamentos o prescripciones, prendas de vestir, dispositivo móvil, etc.

Apartado 5:

MEDIDAS DE AUTOPROTECCIÓN ANTE UNA NUEVA AGRESIÓN SOLO CUANDO LA VÍCTIMA TIENE MENORES A SU CARGO

- 5.1. *Enseñarles a salir de la habitación* cuando se produzca un hecho violento.
- 5.2. Cuando los menores tengan edad suficiente, *enseñarles un sitio seguro donde dirigirse fuera de casa*, estableciendo una ruta segura donde poder encontrarse de nuevo tras el episodio violento.
- 5.3. *Disponer de una habitación segura en la casa*, preferiblemente con cerradura y teléfono y enseñarles a ir a ese lugar ante un episodio violento (es necesario subrayar la importancia de estar seguros).

- 5.4. *Avisar a los menores de qué es una situación violenta* (a menudo no la identifican fácilmente), acordando una palabra o gesto concreto clave que se pueda usar para que los niños comprendan que deben pedir ayuda.
- 5.5. *Enseñar a los menores a pedir ayuda a vecinos de confianza, familiares, etc.*, para que estos a su vez alerten a las Fuerzas y Cuerpos de Seguridad.
- 5.6. *Enseñar a los menores a ponerse en contacto con Fuerzas y Cuerpos de Seguridad*, tanto a través del teléfono fijo de casa como a través del teléfono móvil, cuando no les vea el agresor.
- 5.7. Es necesario que sepan que *deben decir su nombre completo y dirección a Fuerzas y Cuerpos de Seguridad, así como mensaje memorizado sencillo en caso de llamar. También que deben dejar el teléfono descolgado* para evitar una llamada de vuelta y por tanto una situación más peligrosa.
- 5.8. *Practicar con los menores estas estrategias* para que puedan ejecutarlas con destreza.
- 5.9. *Informar al colegio de la situación* y, en su caso, de las medidas judiciales en vigor, para que el centro adopte las medidas oportunas, entre otras, evitar convocar a reuniones de padres a ambos a la vez.

Apartado 6:

MEDIDAS DE AUTOPROTECCIÓN SOLO EN CASO DE QUE EL AGRESOR HAYA ABANDONADO EL DOMICILIO

- 6.1. *Cambio o refuerzo de cerradura en la entrada* o instalación en otras habitaciones o ventanas.
- 6.2. *Dejar las llaves puestas por dentro de casa* cada vez que se esté dentro.
- 6.3. *Instalación de videoporteros o mirillas con buena visibilidad.*
- 6.4. *Añadir luces adicionales en el exterior.*
- 6.5. *Instalación de sistemas de alarma* conectados con el centro de recepción de alarmas.

Apartado 7:

MEDIDAS DE AUTOPROTECCIÓN SOLO EN EL LUGAR DE TRABAJO

- 7.1. Informar al responsable de su situación.
- 7.2. Solicitar cambio de centro o de horario de trabajo y que se garantice la confidencialidad de los datos.
- 7.3. No utilizar las mismas rutas de traslado al centro de trabajo y regreso a casa.
- 7.4. Procurar desplazarse acompañada de personas o compañeros de confianza.
- 7.5. Mostrar en el entorno de personal de seguridad y de confianza fotografía reciente del agresor.
- 7.6. Mantener fotocopia de disposiciones judiciales de protección/alejamiento en el lugar de trabajo.

ANEXOS SOBRE RISCANVI

1. RISCANVI COMPLETO

A continuación, reproducimos la traducción al castellano del documento: Antonio Andrés Pueyo; Karin Arbach Lucioni y Santiago Redondo Illescas (2009): *RisCanvi Complet. Definició d'ítems (pre-edició)*, facilitado por la Generalitat de Catalunya en febrero de 2023.

Salvo error por nuestra parte, se trata del mismo documento que en 2024 la Generalitat ha colgado en su página web junto a los resultados del informe de fiscalización (<https://justicia.gencat.cat/ca/ambits/reinsercio_i_serveis_penitenciaris/serveis_penitenciaris/RisCanvi/>). Nos gustaría señalar que la descripción de las partidas (tanto la que recibimos en febrero de 2023 como la que actualmente está colgada en la web) es un documento fechado en 2009. Sabemos que RisCanvi se ha revisado desde entonces, por ejemplo, para añadir la nueva escala de riesgo de reincidencia general. También parece que la codificación de algunos ítems ha cambiado, ya que en los documentos de 2009 todos los ítems de RisCanvi-C se codifican con tres respuestas posibles, mientras que la plantilla de codificación de variables de uno de los últimos estudios publicados por el CEFJE tiene algunas variables con cuatro respuestas posibles. Por lo tanto, no podemos estar seguros de que la descripción actual de los ítems sea exactamente la misma que la fechada en 2009, y traducida.

FACTORES DELICTIVOS

(Relacionados con el *delito base*, el más grave por el que el sujeto está interno en el centro penitenciario, ya sea como penado o preventivo en el momento de la valoración).

1. Delito base violento: Uso de violencia física, coerción o amenazas en el momento de la realización del delito base.

Se entiende por violencia la acción que produce un daño real, o el intento o la amenaza de dañar, a una o más personas. Se trata de actos que pueden ser razonablemente considerados susceptibles de dañar a otras personas, los comportamientos amenazantes claros, las agresiones sexuales y la destrucción de objetos.

En el supuesto de que el sujeto haya cometido varios delitos, se tiene que considerar si en cualquier de estos ha estado presente la violencia.

NO: No hay pruebas de que el sujeto utilizara la violencia en la comisión del delito o delitos base.

?: El sujeto ha hecho uso de violencia moderada en la realización del delito base. Ejemplos de actos de violencia moderada podrían ser abofetear, empujar u otras conductas que no causan a la víctima lesiones serias o permanentes.

SÍ: Hay evidencia clara de que el sujeto ha hecho uso de violencia en la realización del delito (o delitos) base. Los actos de violencia grave incluirían acciones, a pesar de que no se limitarían, que hayan producido la muerte de la víctima o lesiones físicas o psicológicas muy severas. Es significativo para la calificación de este ítem el hecho de haber utilizado un arma o un objeto amenazante por la comisión del delito.

2. Edad en el momento del delito base: Edad en el momento de cometer el delito base.

Registre la edad del sujeto en el momento del delito base.

0: 22 años o menos.

1: De 23 a 28 años.

2: 29 años o más.

3. Intoxicación durante la comisión del delito base: Intoxicación por consumo de alcohol o sustancias psicotrópicas en el momento de la comisión del delito base.

El individuo había consumido una cantidad significativa de drogas o alcohol en un período de tiempo cercano a cometer el delito base (por ejemplo, entre una hora y unos minutos antes) o durante la comisión del propio delito.

NO: Las pruebas indican que el sujeto era abstinentes en el momento de cometer el delito o delitos base por los que está encarcelado en el momento de la evaluación.

?: Las pruebas sugieren la presencia de intoxicación por alcohol u otras sustancias, aunque no son del todo concluyentes.

SÍ: La evidencia de intoxicación es clara.

4. Víctimas con lesiones: Número de víctimas con lesiones físicas o psicológicas de gravedad moderada o grave, es decir, que requieren atención profesional.

En este ítem hay que considerar si hubo o no víctimas en el delito base. Si las hubo, hay que tener en cuenta aquellas con lesiones, tanto físicas como psicológicas, graves o moderadas.

Las **lesiones leves** son aquellas que no requieren atención médica por ser transitorias o clínicamente imperceptibles. Las **lesiones moderadas** son las que requieren atención profesional, aunque no pongan en peligro la vida (por ejemplo, contusiones, esguinces, estado de ansiedad). Las **lesiones graves** son aquellas que requieren atención profesional (fracturas, pérdida de piezas dentales, cortes profundos, pérdida de conciencia, trastorno de estrés postraumático, etc.) y pueden

poner en riesgo la vida de la víctima, bien por causas físicas directas, bien por el riesgo suicida asociado a determinadas psicopatologías.

- 0:** No ha habido víctimas o estas presentan lesiones físicas o psicológicas leves.
- 1:** Una víctima con heridas moderadas o graves.
- 2:** Dos o más víctimas con lesiones moderadas o graves.

5. Duración de la condena: Duración total de la pena efectiva actual, es decir, la condena que cumple en su conjunto (que podría ser el resultado de la acumulación de varias condenas).

- 0:** 2 años o menos.
- 1:** de 2 a 6 años.
- 2:** 6 años o más.

6. Tiempo ininterrumpido en prisión: Suma, en días, del tiempo transcurrido en prisión desde el último ingreso en libertad, ingreso voluntario o regreso de permiso/libertad, hasta la fecha de la evaluación.

Tiempo total transcurrido desde su último ingreso en prisión hasta el momento actual. El último ingreso puede haber sido de libertad, permisos de salida o salidas programadas.

- 0:** 1 año o menos.
- 1:** de 1 a 3 años.
- 2:** 3 años o más.

FACTORES RELACIONADOS CON EL HISTORIAL DELICTIVO Y VIOLENTO

7. Antecedentes de violencia: Historial de comportamiento violento en la comunidad antes del delito base.

La violencia se define como la acción que produce un daño real o el intento o la amenaza de causar daño a una o más personas. Se incluyen en esta definición las agresiones sexuales o físicas, los comportamientos abiertamente amenazadores y la destrucción de objetos.

La violencia previa incluye los actos violentos y los antecedentes de delitos violentos en la comunidad (es decir, fuera de la institución) antes del delito base. El ítem pretende detectar la presencia o ausencia de violencia en el pasado (hasta la comisión del delito base), independientemente de la frecuencia (por ejemplo, en una sola ocasión o en varias) o de si se ha considerado delito o no.

Cualquier comportamiento violento llevado a cabo dentro del centro penitenciario, o durante el cumplimiento de una condena o medida, no se valorará en este punto, sino en el apartado de factores relacionados con el comportamiento penitenciario del interno (puntos 10 a 14).

NO: No hay pruebas ni antecedentes de comportamiento violento antes del delito base.

?: Existen indicios moderados o parciales de antecedentes de conducta violenta anteriores al delito base, aunque no están del todo claros.

SÍ: Existen pruebas claras de un historial de comportamiento violento anterior al delito base.

8. Inicio de la actividad delictiva o violenta: Edad del sujeto en el momento del primer delito o incidente violento conocido.

Debe tenerse en cuenta la fecha del primer acto violento o delictivo conocido y no solo la fecha del primer delito identificado que haya dado lugar a una intervención policial o judicial. Aunque estos datos pueden coincidir a menudo, no siempre será así; por ejemplo, en el caso de comportamientos violentos hacia profesores o compañeros en edad escolar, que no han sido denunciados, pero de los que existe información o registros fiables. En este caso, la puntuación será «1» o «2», es decir, el comportamiento violento debe evaluarse como presente, haya sido o no denunciado o castigado como delito.

0: El primer delito o el inicio del comportamiento violento se produjo antes de los 16 años.

1: El primer delito o el inicio del comportamiento violento se produjo entre los 17 y los 30 años.

2: Las pruebas sugieren que la persona nunca ha tenido un comportamiento violento, o que ha tenido un comportamiento delictivo o violento después de los 30 años.

9. Aumento de la frecuencia, gravedad y diversidad de los delitos: Por ejemplo, comisión de un mayor número de delitos en menos tiempo, de diferentes tipos de delitos, o agravamiento de delitos no violentos a violentos.

Este ítem describe a un individuo con antecedentes penales que incluyen cargos o condenas por diferentes tipos de delitos, o cuyos antecedentes penales reflejan un aumento de la gravedad o del número de delitos. Deben tenerse en cuenta todos los delitos que aparecen en los antecedentes penales del individuo, independientemente de si era menor en el momento de cometerlos, y debe prestarse atención a la variedad de delitos, su frecuencia y gravedad.

NO: El individuo presenta una única tipología delictiva, sin aumento de la frecuencia o gravedad de las infracciones, es su primer delito o ha cometido un delito de manera aislada.

?: Existen indicadores para sospechar un aumento de la frecuencia, gravedad y diversidad de los delitos, aunque no sean concluyentes, o el aumento haya sido moderado.

SÍ: Pruebas claras de un aumento de la frecuencia, gravedad y diversidad de los delitos. Por ejemplo, más delitos en menos tiempo, diferentes tipos de delitos o evolución hacia delitos violentos.

FACTORES RELACIONADOS CON EL COMPORTAMIENTO PENITENCIARIO

10. Conflictos con otros internos: Genera discusiones o peleas, provoca o recibe provocaciones o presiones de otros internos o utiliza estrategias inadecuadas para relacionarse con otros internos, por ejemplo, amenazas o coacciones. Es víctima de acoso o extorsión, o él mismo las ejerce. *Considere el encarcelamiento actual.*

NO: El preso no ha estado anteriormente en prisión o no hay indicios de conflicto con otros presos.

?: Hay indicios de conflicto interpersonal, pero las pruebas no son concluyentes.

SÍ: Evidencia clara de conflictos interpersonales graves del sujeto con otros internos.

11. Incumplimiento de medidas penales: Ha incumplido una medida penal impuesta o la vigilancia establecida durante el cumplimiento de esta. Por ejemplo, evita acudir a entrevistas o someterse a seguimiento, realiza cambios bruscos en sus rutinas horarias (trabajo, tratamiento, entrevistas, etc.) sin previo aviso. *Para la evaluación, téngase en cuenta el último encarcelamiento.*

Este factor de riesgo se refiere a los fracasos e incumplimientos de diferentes medidas penales o de supervisión planificada, mientras el individuo cumplía una pena alternativa, en libertad condicional o vigilada, en una institución penitenciaria o en un hospital psiquiátrico.

Un incumplimiento de condena se considera grave cuando el individuo ha reincidido durante el cumplimiento de la condena y, por lo tanto, ha sido internado de nuevo en un centro penitenciario o de salud mental. Ejemplos típicos de «incumplimientos graves» son: la reincidencia durante la libertad condicional, la revocación de la libertad condicional o el incumplimiento de la atención psiquiátrica ordenada por el tribunal. Los «fracasos moderados» se refieren a infracciones o incumplimientos de las condiciones de libertad que dan lugar a una medida disciplinaria menor (por ejemplo, reprobación o pérdida de privilegios), y pueden ser: incumplimiento de las normas de conducta inherentes a un permiso de salida, causar molestias, incumplir una prescripción médica, o cualquier otra conducta que, por su gravedad o reiteración, haya dado lugar a una medida restrictiva (por ejemplo, SAM¹ nivel D, expediente disciplinario, retirada de permisos de salida durante un periodo de tiempo, o regresión de grado).²

¹ SAM son las siglas de Sistema d'Avaluació i Motivació (Sistema de Evaluación y Motivación): se trata de una evaluación aritmética en tres áreas (actividades, conducta y salidas) que clasifica a los internos en 4 niveles (A, B, C y D), cada uno de los cuales permite (o no) obtener determinadas recompensas en la vida penitenciaria (art. 263 del Reglamento Penitenciario), o incluso beneficios penitenciarios (art. 202.2 del Reglamento Penitenciario: libertad condicional anticipada o indulto especial).

² Para el significado de esta expresión, véase la nota 3.

Este ítem **excluye** las evasiones y fugas, que deberán considerarse en el ítem **13**, y los quebrantamientos de permisos, que deberán evaluarse en el ítem **15**.

NO: El individuo siempre ha cumplido las medidas que se le han aplicado de forma adecuada y sin incidentes.

?: Existen indicadores de incumplimiento menor o las pruebas no son concluyentes para codificar «Sí».

SÍ: Existen indicadores claros de incumplimiento efectivo (ya sea grave o moderado).

12. Expedientes disciplinarios: Faltas disciplinarias graves o muy graves y otros problemas de adaptación institucional en prisión. *Tener en cuenta el encarcelamiento actual o anteriores.*

Se consideran problemas de conducta penitenciaria las faltas disciplinarias, los quebrantamientos del orden penitenciario, las agresiones verbales o amenazas, o las conductas agresivas de cierta gravedad. Piénsese en las faltas graves y muy graves establecidas en los artículos 108 y 109 del Reglamento Penitenciario. En el caso de los internos en tercer grado, el factor de riesgo se refiere a la conducta del individuo en relación con la vigilancia y las medidas aplicadas en cada caso concreto.

Existe una amplia variedad de definiciones operativas para categorizar el comportamiento agresivo o violento en prisión, que van desde las amenazas de violencia hasta cualquier ataque que provoque lesiones físicas graves. Hay que tener en cuenta la violencia cometida durante el cumplimiento de la condena o medida de seguridad, así como la violencia dirigida contra el evaluador durante la entrevista, tanto si se produce dentro de la prisión como en contextos similares (por ejemplo, condenas comunitarias).

Este factor de riesgo **excluye** las conductas descritas en los ítems **11**, **13** y **15**. Es decir, los incumplimientos de medidas penales, así como los incumplimientos de permisos y las evasiones, no tienen que considerarse en este apartado.

NO: El interno es primario o no hay indicios de que haya realizado ninguna de las conductas descritas.

?: El interno tiene expedientes disciplinarios por faltas leves.

SÍ: El interno ha sido sancionado por faltas disciplinarias graves o muy graves, ha participado en disturbios o incidentes institucionales graves o ha utilizado violencia, armas o amenazas hacia otros internos o hacia los funcionarios. Incluir la información recogida en el ítem **10** siempre que los conflictos con otros internos hayan dado lugar a expedientes disciplinarios.

13. Evasiones o fugas: Evasiones o fugas de un centro penitenciario (por las que puede o no haber sido dado de baja del centro). *Considerar encarcelamiento actual o anteriores.*

Las fugas o evasiones deben registrarse tanto si constan en los archivos penitenciarios (formal o informalmente) como si son conocidas por el personal, aunque no estén registradas, así como las fugas que el sujeto comunique espontáneamente en el transcurso de una entrevista.

NO: No hay pruebas de evasión.

?: Existen indicadores de evasiones o fugas, conocidos por el personal o comunicados por el interno, pero que no constan en su expediente.

SÍ: Hay indicadores claros de alguna evasión o fuga.

14. Regresión de grado:³ Evolución negativa en relación con el tratamiento penitenciario. *Considere el encarcelamiento actual o anteriores.*

NO: El sujeto ha evolucionado de forma positiva o normal en el tratamiento penitenciario o no hay evidencias de una evolución negativa.

?: Existen evidencias moderadas o parciales de una evolución negativa del sujeto respecto a su tratamiento penitenciario.

SÍ: Existen evidencias claras de una evolución negativa del sujeto respecto a su tratamiento penitenciario.

15. Incumplimiento de permisos: El interno ha disfrutado de permisos y ha incumplido las normas que los regulan. No regresar de un permiso, por ejemplo, no regresar una noche al centro o sección de régimen abierto sin haber sido autorizado para ello. *Considerar el último encarcelamiento.*

NO: No se han tomado permisos o no hay registro de infracciones de permisos.

?: Se ha producido un incumplimiento justificado o leve de un permiso, por ejemplo, el interno ha llegado un poco más tarde de la hora establecida, pero puede justificarlo y los motivos son accidentales.

SÍ: Se han producido reiterados o graves incumplimientos de permisos: los motivos alegados por el interno para el no reingreso tienen poca credibilidad o son injustificados, existen evidencias claras de que el interno consumió drogas o alcohol, o el incumplimiento del permiso supuso la baja en el centro.

FACTORES PERSONALES Y BIOGRÁFICOS

16. Desajuste infantil: Problemas de comportamiento en la infancia o un patrón de mal comportamiento habitual (por ejemplo, peleas con compañeros

³ Los presos en España se clasifican en uno de tres grados, cada uno de los cuales corresponde a un régimen penitenciario diferente: primer grado o régimen cerrado (con una vigilancia muy intensa y actividades muy limitadas con otros internos), segundo grado o régimen ordinario (vivir dentro de la prisión compartiendo actividades con otros internos y salir ocasionalmente de permiso), y tercer grado o régimen abierto (poder salir todos los días a trabajar y sólo pasar las noches y fines de semana en prisión, o incluso permanecer toda la semana en casa con algún tipo de vigilancia electrónica, etc.). Para ser clasificado en un grado menos restrictivo el interno tiene que cumplir unos requisitos, como obedecer las normas penitenciarias y mostrar una conducta aceptable, y ser considerado no peligroso. La expresión en español para la clasificación en mejor grado es “progresar de grado”; la clasificación en peor grado se dice “regresar de grado”.

o mal comportamiento en casa). Bajo rendimiento escolar (con muchos suspensos), absentismo o abandono escolar.

El ítem describe a un individuo con problemas de conducta significativos en la infancia y adolescencia temprana (normalmente antes de los 12 años). Estos problemas pueden incluir comportamientos como mentir de forma persistente, engañar, robar, agredir, provocar incendios, abusar de sustancias (incluido el abuso de alcohol y la inhalación de pegamento), vandalismo, violencia, absentismo escolar, interrupción de las actividades en clase, acoso a compañeros, fugarse de casa y actividad sexual precoz. Estos comportamientos destacan por ser más graves que los exhibidos por la mayoría de los niños de su edad y han dado lugar a problemas como quejas de otros, suspensos reiterados, expulsión del colegio, intervención policial u otros.

NO: No hay pruebas de problemas de comportamiento en la infancia y la adolescencia.

?: Existen algunos problemas de comportamiento infantil o adolescente, pero de carácter leve.

SÍ: Se ha evidenciado un desajuste en al menos dos de los tres contextos siguientes: el hogar, la escuela y la comunidad. O la inadaptación se ha producido en un solo contexto, pero de forma grave (por ejemplo, mal comportamiento escolar grave y persistente).

17. Distancia entre la residencia habitual de la persona y el centro penitenciario: la residencia a la que la persona va a ir (con motivo de su salida en libertad, de permiso, etc.) está lejos del centro penitenciario.

En el caso de estancias prolongadas en prisión, sin permisos, considere la residencia que tendría el recluso si saliera del centro penitenciario.

0: Menos de 100 km.

1: Entre 100 y 300 km.

2: Más de 300 km.

18. Nivel educativo: Grado académica alcanzado, nivel de estudios finalizados.

La valoración de este ítem podría cambiar durante la estancia en el centro en caso de que el interno haya desarrollado actividades educativas y haya progresado en su nivel académico.

0: Bajo nivel educativo, educación primaria.

1: Nivel educativo medio, educación secundaria o formación profesional.

2: Nivel educativo alto, educación superior o universitaria.

19. Problemas relacionados con el empleo: Desempleo crónico, inestabilidad laboral o muchos empleos anteriores de corta duración. Dificultades de inserción laboral. Se niega a buscar trabajo, es despedido con frecuencia o lo abandona con facilidad. Considere también el trabajo realizado dentro de la prisión. Los

problemas laborales no están causados por una discapacidad física o una incapacidad legal. *Considere toda la vida del individuo.*

Este ítem se refiere a las posibles dificultades del sujeto en relación con el entorno laboral y las relaciones prosociales que de él se derivan. Los individuos puntuarán en este factor de riesgo si se caracterizan por: negarse a buscar un trabajo legítimo (social y legalmente aceptable), tener un historial previo de muchos trabajos en cortos periodos de tiempo, haber sido despedidos o abandonar frecuentemente el trabajo. El ítem evalúa los posibles problemas actitudinales, de responsabilidad y relacionales del sujeto con el trabajo, no su capacidad/discapacidad física o mental para encontrar un empleo.

Los cambios recientes en la situación laboral oficial deben examinarse detenidamente para ver si pueden indicar una mejora. Sin embargo, no todos los cambios formales o legales en la situación laboral deben interpretarse como positivos a efectos de este ítem. Por ejemplo, una persona con un historial previo de 15 años de problemas laborales, pero que haya obtenido recientemente una pensión de incapacidad laboral, puede seguir puntuando «Sí» (es decir, negativo) en este factor de riesgo.

Para los inmigrantes, considere si han mantenido un historial de empleo asalariado o estable, independientemente de su situación legal (incluido si tienen o no permiso de trabajo).

Preste atención también a la participación del sujeto en trabajos dentro de la prisión o en programas de formación. Tenga en cuenta que este factor de riesgo evalúa la disposición mostrada por el sujeto a participar en actividades laborales tanto dentro como fuera del centro penitenciario.

NO: Ha sido una persona activa y trabajadora, sin problemas sistemáticos relacionados con el trabajo.

?: Presencia posible/moderada de problemas relacionados con el trabajo.

SÍ: Ha sufrido problemas laborales graves o muy graves.

20. Falta de recursos económicos: Nivel económico estimado en el último año o antes del ingreso si lleva más de 12 meses en prisión.

Se refiere a los bienes y recursos económicos (ingresos, dinero) de que dispone el interno para hacer frente, con suficiente independencia, a las exigencias económicas de la vida diaria en la calle. Hay que considerar conjuntamente la situación económica actual (o anterior a su ingreso en prisión) y la que sería probable en caso de un permiso o régimen abierto. Los recursos económicos pueden proceder del individuo, de la familia o de los servicios sociales. Los factores que pueden aportar información sobre el apoyo económico procedente de la familia son: hábitos de trabajo y nivel de estabilidad laboral de los miembros de la familia, bienes o ahorros de que dispone el sujeto, nivel de endeudamiento, equilibrio entre ingresos y gastos, y entre ingresos y peculio del interno,⁴ y nivel de cobertura de las necesidades mínimas.

⁴ En el informe final (n. 24, p. 62) se ofrece una definición de este término.

NO: Dispone de recursos económicos suficientes (incluso medios o altos) para cubrir sus necesidades básicas (por ejemplo, alimentación o vivienda). Los recursos económicos no son un aspecto problemático para este sujeto.

?: Dispone de algunos recursos económicos, pero estos son insuficientes o inestables para cubrir sus necesidades básicas, convirtiendo la vida cotidiana actual del interno en una situación crítica o problemática. Los recursos económicos son un aspecto problemático para este sujeto.

SÍ: El interno carece de apoyo económico para cubrir sus necesidades básicas. Su subsistencia o la de su familia dependiente depende de la asistencia social.

21. Falta de planes viables para el futuro: Planes de futuro poco realistas o poco viables, falta de planificación a medio y largo plazo.

Incapacidad o falta de voluntad para plantearse y llevar a cabo, a medio y largo plazo, planes realistas de ocupación o integración psicosocial (trabajo, familia, desarrollo personal, ocio, etc.). Tiende a vivir exclusivamente al día y cambia sus planes con frecuencia. No se preocupa por el futuro. No le interesa tener un trabajo estable. Puede llevar una vida nómada y describirse a sí mismo como un «trotamundos».

A veces puede expresar metas futuras específicas, pero poco realistas, como ser abogado, escritor, cirujano, trabajador social, psicólogo, piloto de avión, etc., sin ser consciente o ignorando las cualificaciones requeridas para estas profesiones.

La inviabilidad de sus planes de futuro también puede deberse a la falta de apoyo de la red social (familia y otros), o a que la persona no quiere recibir esta ayuda.

NO: Planes de futuro realistas y viables a medio y largo plazo.

?: Planes de futuro poco realistas y viables.

SÍ: Planes de futuro muy poco realistas y claramente inviables.

FACTORES SOCIALES Y FAMILIARES

22. Antecedentes penales en la familia de origen: Familiares de primer o segundo grado (o análogos) que hayan realizado conductas antisociales o delictivas.

Los padres, hermanos, abuelos, tíos o primos del recluso tienen antecedentes penales o han tenido un comportamiento antisocial notorio, aunque no conste como delito. *Análogos* se refiere a los miembros de la familia extensa, como la familia política.

NO: No hay indicios de antecedentes penales en la familia de origen del sujeto.

?: Existen indicios moderados o parciales de antecedentes penales en la familia de origen del sujeto, aunque no son concluyentes.

SÍ: Hay pruebas claras de antecedentes penales en la familia de origen del sujeto.

23. Socialización o educación problemática en la familia de origen: Las relaciones familiares en la familia de origen (biológica, adoptiva, de acogida, etc.) se han caracterizado por la indiferencia, la hostilidad o el uso frecuente del castigo.

El sujeto ha sido testigo o víctima de comportamientos violentos, abusos o negligencias en el entorno familiar. Tuvo una relación negativa con la persona o personas con las que convivió más frecuentemente durante su infancia, o creció en un hogar de acogida sin vínculo afectivo con los tutores. Así como si sus padres fallecieron o se ausentaron del hogar, y el sujeto mantuvo una relación fundamentalmente negativa con la persona o personas que ejercían su tutela. También si el sujeto sufrió maltrato emocional, abandono o agresiones físicas o sexuales en el entorno familiar. *Considere la infancia y adolescencia de la persona.*

NO: El sujeto experimentó una educación cálida y satisfactoria en su familia de origen.

?: Pruebas, aunque no concluyentes, de una educación familiar indiferente, hostil y poco cariñosa.

SÍ: Evidencia clara de mucho castigo, relaciones desagradables, poca protección y mucha hostilidad o indiferencia en su crianza.

24. Falta de apoyo familiar y social: Falta de red social, de relaciones satisfactorias y de apoyo o de contacto regular con la familia y los amigos. *Considere el último año, incluido el momento de la evaluación.*

Este factor de riesgo es dinámico, ya que puede variar con el tiempo. Debe evaluarse teniendo en cuenta las interacciones sociales, familiares o conyugales actuales (en los últimos 12 meses). En el caso de las personas institucionalizadas, hay que tener en cuenta la frecuencia de las visitas y las cartas recibidas, así como los contactos telefónicos durante el periodo de evaluación. También se evalúa si el sujeto tiene a alguien a quien recurrir en caso de necesitar apoyo y si se siente cómodo haciéndolo (independientemente de si actualmente lo hace o no). Un aspecto esencial del ítem es la percepción del individuo sobre la disponibilidad de apoyo social con el que puede contar.

La presencia de familiares, amigos o grupos comunitarios (por ejemplo, religiosos, ONG, etc.) que sean pacientes, tolerantes y alentadores con el sujeto puede ser de gran ayuda en el diseño y mantenimiento de un plan de reinserción social. Debe valorarse la disponibilidad de un entorno «familiar» sano, confortable y seguro; de personas significativas en el entorno familiar y de pareja como fuentes interpersonales de refuerzo de conductas prosociales y de rechazo de conductas antisociales. La ausencia de apoyo positivo puede facilitar la futura implicación del individuo en conductas ilícitas. Los miembros de la familia o la pareja pueden actuar como modelos positivos, pero también promover activamente el comportamiento antisocial.

Es muy importante determinar exactamente de qué ayuda y apoyo se dispone y de quién o de dónde procede. Hay que valorar que no se trata sólo de «buenas intenciones» de familiares y amigos, que se utilizan para obtener la excarcelación, el tercer grado u otros privilegios.

NO: El sujeto mantiene relaciones frecuentes y satisfactorias con otras personas como padres, hermanos, abuelos, cónyuge. Afirma que tiene amigos íntimos o familiares que le ofrecen apoyo, lo cual agradece.

?: El recluso cuenta con una o varias personas que podrían proporcionarle apoyo, pero no está claro cuáles podrían ser o se siente incómodo al pedir ayuda o apoyo.

Sí: No existen relaciones positivas ni apoyo por parte de familiares, no hay amigos íntimos a los que recurrir, o dicho apoyo (emocional, económico o físico) existe, pero el sujeto no quiere aceptarlo. Indique también la presencia del ítem si las redes sociales y el apoyo del individuo están relacionados con sus actos violentos y delictivos.

25. Pertenencia a banda criminal o tiene amigos criminales: es miembro de una banda organizada implicada en actividades delictivas o está vinculado a redes criminales. *Considerar el momento actual.*

El individuo tiene amigos (con los que pasa su tiempo libre, cuyas opiniones valora, que podrían ayudarle en sus dificultades...) que tienen antecedentes penales y están implicados en actividades delictivas. Las bandas organizadas o redes delictivas se refieren a grupos de tres o más miembros que delinquen juntos. Los miembros suelen tener una relación estrecha y una ideología común. Las redes delictivas suelen tener motivaciones económicas o políticas.

NO: No hay pruebas de que el sujeto pertenezca a ninguna banda criminal o tenga amigos delincuentes.

?: Existen pruebas parciales o dudosas de que el sujeto pertenece a una banda criminal o tiene amigos criminales.

Sí: Hay pruebas claras de que el sujeto pertenece a una banda criminal o tiene amigos implicados en actividades delictivas.

26. Pertenencia a grupos sociales con riesgo de delincuencia, distintos de una banda criminal: Pertenece a grupos sociales (*distintos de una banda criminal, definida en el ítem 25*) que pueden correr el riesgo de cometer actos delictivos (por ejemplo, el recluso es un joven marginado, no tiene domicilio conocido, es traficante o distribuidor de drogas, participa en actividades relacionadas con la prostitución, etc.). *Considere la situación actual.*

NO: No pertenece a grupos sociales de riesgo.

?: Existen pruebas parciales o dudosas de pertenencia a grupos sociales de riesgo.

Sí: Hay pruebas claras de pertenencia a grupos sociales de riesgo.

27. Rol delictivo destacado: En su entorno inmediato o círculo de familiares, amigos, conocidos y vecinos, goza de gran prestigio, sus actos delictivos destacan y es respetado dentro de la subcultura delictiva. *Consideremos el momento actual.*

Está muy bien considerado en su comunidad debido a su historial delictivo y se valora a sí mismo por este destacado papel de liderazgo delictivo. Para ponderar el elemento hay que evaluar su comportamiento tanto dentro como fuera de prisión.

NO: El individuo no tiene un papel delictivo destacado y reconocido en su entorno inmediato, ni su actividad delictiva le reporta una buena reputación personal.

?: Tiene cierto «reconocimiento social» como resultado de su actividad delictiva, pero el sujeto no presta mucha atención a este reconocimiento.

SÍ: Tiene un alto «reconocimiento social» por sus actividades delictivas y, además, el sujeto valora positivamente este reconocimiento.

28. Víctima de violencia de género (en caso de mujeres): La interna es víctima de violencia física, sexual o psicológica por parte de su pareja o expareja. *Considerar el último año, incluido el momento de la evaluación.*

Se considerará violencia de género cualquier acto de agresión que haya causado sufrimiento físico, sexual o psicológico a la mujer. También se incluyen las amenazas de agresión, las coacciones o la privación arbitraria de libertad.

NO: La mujer no es víctima de violencia de género por parte de su pareja o expareja.

?: Probablemente sufre violencia de género, pero no hay pruebas claras.

SÍ: Es claramente víctima de violencia de género por parte de tu pareja o expareja.

29. Cargas familiares actuales: El sujeto es responsable del cuidado de miembros de su familia, por ejemplo, hijos menores, padres ancianos o familiares enfermos. *Considera el momento actual.*

En este ítem hay que valorar si el interno tiene familia, ya sea adquirida (hijos o cónyuge) o de origen (padres, hermanos, abuelos), y todos o algunos de sus miembros dependen socioeconómicamente de él. También debe tenerse en cuenta el caso de que el interno tenga compromisos económicos con la familia como consecuencia de una sentencia judicial o sea responsable del alquiler o hipoteca de la vivienda donde reside la familia.

NO: El sujeto no tiene actualmente responsabilidades familiares o ningún miembro de su familia está a su cargo.

?: Probablemente tiene algunas responsabilidades familiares actuales, pero no dependen exclusivamente de él, ya que pueden ser cumplidas por otro miembro de la familia.

SÍ: Evidencia clara de cargas familiares actuales que el sujeto debe atender ineludiblemente.

FACTORES CLÍNICOS

30. Abuso o dependencia de drogas: Consumo de drogas, legales o ilegales, que interfiere negativamente en diferentes aspectos de su vida (familiar, laboral y social) o en su salud física. *Considere el último año.*

El abuso grave de sustancias dificulta las interacciones y los comportamientos prosociales, y puede promover directamente comportamientos delictivos a través de la desinhibición, o de la necesidad urgente y crónica de obtener dinero para pagar el hábito. La codificación de este ítem requiere dos pasos. En primer lugar, hay que determinar si el sujeto consume drogas actualmente (en los últimos 12 meses). En segundo lugar, hay que evaluar si este consumo interfiere en su desarrollo escolar, familiar, laboral o de otro tipo. La evaluación debe ser precisa y basarse en las pruebas disponibles. También incluye el uso indebido de medicamentos con receta y el consumo de sustancias tóxicas como disolventes y pegamentos. Se excluye el consumo de nicotina y cafeína; el abuso de alcohol se trata en el punto 31.

Un procesamiento o condena por posesión o tráfico de drogas no indica necesariamente un problema grave de abuso de drogas. Algunos traficantes pueden no abusar de las drogas. El tema se centra en el abuso y la dependencia de sustancias más que en los delitos relacionados con las drogas.

El hecho de que el individuo considere el abuso de sustancias como un problema que tiene puede constituir una prueba de este factor de riesgo, pero no al revés, es decir, cuando el sujeto niega tener un problema con el consumo de sustancias.

También deben tenerse en cuenta los síntomas psiquiátricos o las alteraciones neurológicas derivadas del consumo de sustancias (trastorno psicótico inducido por sustancias, deterioro grave de la memoria, etc.). *Tenga en cuenta los últimos 12 meses.*

Cuando se evalúa a una persona que ha estado en prisión durante más de 12 meses, hay que tener en cuenta su comportamiento en relación con el consumo de drogas mientras estaba en prisión, durante los permisos y si está o ha estado vinculada a un tratamiento contra las drogas.

NO: El sujeto no consume drogas, o lo hace de forma ocasional y controlada, y el consumo de sustancias no interfiere en su comportamiento y relaciones habituales (por ejemplo, mantenimiento del empleo, ausencia de conflictos conyugales o familiares debidos al consumo de drogas, etc.). También se valorará como «No» a quienes, habiendo consumido en el pasado, se hayan mantenido abstinentes durante el último año.

?: Evidencia de consumo frecuente de sustancias con deterioro leve o esporádico del funcionamiento psicosocial.

SÍ: Evidencia de un deterioro grave y permanente de su funcionamiento psicosocial debido al consumo de sustancias (por ejemplo, centrar sus actividades sociales en el consumo de sustancias o cometer delitos debido al consumo de sustancias).

31. Abuso o dependencia del alcohol: Consumo de alcohol que interfiere negativamente en diferentes aspectos de su vida (familiar, laboral y social) o en su salud física. *Considere el último año.*

El abuso grave de alcohol dificulta las interacciones y los comportamientos prosociales; también puede fomentar comportamientos delictivos por su desinhibición o por la necesidad de obtener dinero para mantener el hábito. La codificación de este ítem requiere dos pasos. En primer lugar, hay que establecer si el sujeto consume alcohol actualmente (en los últimos 12 meses). En segundo lugar, hay que

evaluar si este consumo interfiere con la escuela, la familia, el trabajo u otras actividades. El diagnóstico de abuso o dependencia del alcohol debe evaluarse con precisión sobre la base de pruebas concretas.

Si el individuo considera que el consumo de alcohol es un problema que tiene, esto puede constituir una prueba de este factor de riesgo, pero no al revés, es decir, cuando el sujeto niega tener problemas con el consumo de alcohol.

También deben abordarse los síntomas psiquiátricos o las alteraciones neurológicas derivadas del consumo de alcohol (*delirium tremens*, trastorno psicótico inducido por el alcohol, deterioro grave de la memoria, etc.).

Cuando se evalúa a una persona que lleva más de 12 meses en prisión, hay que tener en cuenta su comportamiento en relación con el consumo de alcohol durante su estancia en prisión, si ha recibido medidas disciplinarias como consecuencia de ello (intoxicación en el momento de cometer el delito, consumo de alcohol durante un permiso, etc.) y si está o ha estado vinculada a un tratamiento de desintoxicación. En el contexto penitenciario, el consumo continuado de alcohol puede ser un indicador de la presencia del factor de riesgo.

NO: El sujeto no consume alcohol o lo hace de forma ocasional y controlada, y el consumo de alcohol no interfiere en su comportamiento y relaciones habituales (por ejemplo, mantenimiento del empleo, ausencia de conflictos conyugales o familiares debidos a la bebida, etc.). También se valorará como «No» a quienes, habiendo consumido en el pasado, se hayan mantenido abstinentes durante el último año.

?: Evidencia de consumo frecuente de alcohol con deterioro leve o esporádico del funcionamiento psicosocial.

SÍ: Evidencia de deterioro grave o permanente del funcionamiento psicosocial debido al abuso de alcohol (por ejemplo, beber hasta el punto de intoxicación, centrar las actividades sociales en la bebida o cometer delitos debido a la bebida).

32. Trastorno mental grave: Se ha diagnosticado un trastorno mental grave, en algún momento de la vida del sujeto a evaluar y según DSM-IV (APA, 1994), CIE-10 (OMS, 1992) o sus versiones actualizadas. Se consideran trastornos mentales graves o severos los trastornos psicóticos y bipolares, con una duración prolongada (por ejemplo, 2 años) de la enfermedad o de su tratamiento, y que producen una disfunción de moderada a severa del funcionamiento global de la persona.

Este factor de riesgo se codifica en función de los antecedentes y la codificación no tiene por qué variar en función de si el trastorno está actualmente activo o en remisión. Este factor de riesgo se aplica tanto a los trastornos del pensamiento (por ejemplo, trastornos del espectro psicótico) como a los de naturaleza afectiva (por ejemplo, trastorno maníaco del estado de ánimo).

Los indicadores de trastorno mental grave pueden ser: incapacidad para discernir la realidad, aislamiento grave, hiperactividad, delirios o alucinaciones, deterioro del juicio y del pensamiento racional, desorientación, desorganización conceptual, comportamiento desafiante, suspicacia, manía, retraimiento emocional, afectividad inapropiada, manierismos o posturas extrañas, etc.

También se incluyen en la evaluación de este ítem las lesiones cerebrales (por ejemplo, traumatismo craneoencefálico, contusión cerebral, epilepsia, secuelas de

infecciones u otras enfermedades, etc.). Debe prestarse especial atención a la presencia de síntomas psicóticos que invaliden o anulen la sensación de autocontrol o que amenacen la sensación de seguridad o bienestar. La ideación homicida debe seguir siendo cuidadosamente explorada.

Otros trastornos mentales, como los trastornos de ansiedad, los trastornos somatomorfos, las parafilias, los trastornos del sueño, los trastornos del control de los impulsos y los trastornos de la personalidad no tienen que codificarse en este apartado.

NO: No hay evidencia de la presencia de trastorno mental grave a lo largo de la vida del sujeto.

?: Evidencia no concluyente de trastorno mental grave a lo largo de la vida del sujeto.

SÍ: Presencia cierta o muy probable de trastorno mental grave en algún momento de la vida del sujeto.

33. Comportamiento sexual promiscuo y de riesgo, o parafilia: Promiscuidad sexual de riesgo, hipersexualidad irresponsable, comportamiento sexual violento. También desviación, anomalía, perversión sexual o trastorno de las preferencias sexuales. *Considerar toda la vida de la persona.*

El factor de riesgo evalúa dos aspectos diferentes. Por un lado, describe a un individuo que solo mantiene relaciones sexuales casuales e impersonales. Esto puede reflejarse en frecuentes encuentros sexuales casuales («rollos de una noche»), selección indiscriminada de parejas sexuales, mantenimiento de varias parejas sexuales al mismo tiempo, infidelidades repetidas, uso frecuente de la prostitución, deseo de participar en una amplia variedad de actividades sexuales o hipersexualidad. Además, el individuo puede haber coaccionado sexualmente a otras personas y puede haber sido procesado o condenado por agresiones sexuales (este supuesto es prueba suficiente de la presencia de este factor de riesgo). Por otro lado, el ítem evalúa la posible presencia de parafilia. Esta se define como una desviación, anomalía, perversión o trastorno de las preferencias sexuales. Implica fantasías, impulsos, excitación sexual intensa y repetida y comportamiento sexual en respuesta a estímulos desviados. Las principales parafilias clasificadas son: exhibicionismo, fetichismo, frotismo, masoquismo y sadismo sexual, voyerismo y pedofilia.

Para hacer la evaluación, se puede explorar su visión del sexo, su actitud hacia las parejas sexuales, el número y la naturaleza de las relaciones sexuales que mantiene el individuo, etc.

NO: Comportamiento sexual normalizado y ausencia de conductas sexuales promiscuas y de riesgo o parafilias a lo largo de la vida del sujeto.

?: Posible/moderada presencia de conductas sexuales promiscuas y de riesgo o parafilias a lo largo de la vida del sujeto o pruebas no concluyentes.

SÍ: Presencia segura o muy probable (con pruebas firmes) de comportamiento sexual promiscuo y de riesgo o parafilia.

34. Respuesta limitada al tratamiento psicológico o psiquiátrico: El sujeto ha recibido tratamiento psicológico o psiquiátrico (o farmacológico), o por abuso de sustancias, y no ha mostrado adherencia o los resultados han sido pobres. *Considerar el último año.*

La evaluación se refiere al estado del sujeto en el momento actual (últimos 12 meses) como consecuencia de un tratamiento que está recibiendo o ha recibido recientemente (en los últimos años).

Aparte de conocer el esfuerzo terapéutico de la persona, el factor crítico es si ha mejorado o no en las competencias necesarias para hacer frente a las exigencias vitales, sociales, laborales e interpersonales básicas, presentes y futuras. También es importante saber si la persona ha buscado ayuda y si la ha aceptado, si la ha descartado de plano o si ha accedido a recibir ayuda solo para causar una buena impresión ante un tribunal, una junta de revisión u otras autoridades, pero sin una verdadera motivación para el cambio.

En este apartado se incluyen los tratamientos (también farmacológicos) destinados a reducir los problemas delictivos, psiquiátricos, psicológicos, sociales o laborales. Por otra parte, los tratamientos irrelevantes para su comportamiento violento o delictivo (por ejemplo, el tratamiento del asma) no deben evaluarse aquí.

Una respuesta escasa al tratamiento puede manifestarse por los siguientes síntomas: inestabilidad emocional, estrés, nerviosismo, debilidad, preocupaciones que interfieren en el funcionamiento diario u otros tipos de disfunción o síntomas psicológicos (delirios o alucinaciones, ansiedad, obsesiones y compulsiones, anorexia, fobias, disfunción sexual, hipocondría, tartamudez, etc.).

NO: La respuesta al tratamiento recibido es adecuada y favorable, o el sujeto no ha recibido tratamiento por no presentar ninguna de las disfunciones o síntomas mencionados anteriormente.

?: Respuesta parcial al tratamiento recibido, o pruebas insuficientes de la eficacia/efectividad del tratamiento.

SÍ: El sujeto no ha mostrado respuesta al tratamiento o esta ha sido escasa; se ha negado a iniciar el tratamiento, lo ha iniciado pero lo ha abandonado, ha fingido seguir el tratamiento o lo ha cumplido sin sacarle provecho; no ha recibido tratamiento a pesar de que las disfunciones o síntomas que presenta favorecen la conducta infractora.

35. Trastorno de la personalidad relacionado con la ira, la impulsividad o la violencia: Cumple los criterios del DSM-IV o de la CIE-10 para un trastorno de la personalidad del clúster B, o presenta ira, impulsividad o comportamiento violento de forma habitual y generalizada. *Considere toda la vida del recluso.*

El diagnóstico del trastorno de la personalidad debe ajustarse a uno de los sistemas de clasificación oficiales, como el DSM-IV (APA, 1994) o la CIE-10 (OMS, 1992), o a sus versiones actualizadas. Este factor de riesgo se codifica en función de los antecedentes y la codificación no se ve afectada si el trastorno está actualmente activo o en remisión. Un patrón habitual de sentimientos de ira y comportamiento compulsivo y violento puede ser indicativo de un trastorno de este tipo, aunque debe descartarse en presencia de otros diagnósticos que puedan explicar los estados

emocionales y el comportamiento desviado (por ejemplo, demencia o trastornos psicóticos).

NO: No hay indicadores de trastornos de la personalidad u otros diagnósticos psicopatológicos.

?: Presencia posible o probable (pero no segura) de indicadores o rasgos de un trastorno de la personalidad.

SÍ: Evidencia clara de un trastorno de la personalidad del clúster B (antisocial, *borderline*, narcisista o histriónico) o un patrón habitual de ira, impulsividad o comportamiento violento.

36. Mal afrontamiento del estrés: Tendencia a experimentar un alto nivel de estrés. Falta de recursos cognitivos y conductuales para afrontar (reducir, minimizar, dominar o tolerar) las demandas internas y externas de una determinada situación estresante. *Considerar el momento presente.*

Este factor de riesgo implica anticipar qué fuentes de estrés puede encontrar el individuo y cómo reacciona y las afronta. Es necesario tener en cuenta la vulnerabilidad particular del individuo y la capacidad y los recursos de que dispone para superar y hacer frente a las dificultades. Puede ser útil disponer de algún indicador de su nivel de estrés, evaluado en función de los problemas a los que ha estado expuesto en su vida cotidiana. Deben tenerse en cuenta las habilidades de afrontamiento, ya que un alto nivel de estrés en combinación con buenas habilidades de afrontamiento puede dar lugar a una puntuación baja en este factor de riesgo.

La soledad y la alienación parecen favorecer la vulnerabilidad a los acontecimientos vitales negativos, que pueden conducir a la ruptura emocional y, en determinadas condiciones, a la acción suicida. Por lo tanto, estas situaciones deben tenerse especialmente en cuenta en la evaluación de este factor de riesgo.

NO: El individuo no experimenta estrés, el nivel de estrés experimentado es bajo o sus recursos de afrontamiento son adecuados.

?: El individuo está expuesto a un nivel de estrés de moderado a alto, pero lo ha afrontado adecuadamente en el pasado.

SÍ: El individuo experimenta un estrés de moderado a elevado y no ha sido capaz de afrontarlo satisfactoriamente.

37. Intentos o conductas autolesivos: Conductas autolíticas, como autolesiones o intentos de suicidio, que han puesto en riesgo la vida o la integridad física del individuo, independientemente de su intención. *Considerar cualquier momento de su vida.*

Antecedentes de violencia autoinfligida, tanto dentro como fuera de prisión, en cualquier momento de su vida. Intentos o amenazas de suicidio, autolesiones leves y graves o planes persistentes, explícitos y factibles de autolesión. El factor de riesgo incluye:

- Intentos manifiestos de suicidio (por ejemplo, intento de estrangulación, asfixia, autoincineración, sobredosis, ingestión voluntaria de sustancias nocivas, defenestración, etc.);
- lesión directa y deliberada del propio cuerpo que causa daños en los tejidos o deja marcas que duran más de unos minutos, y que a menudo se realiza para hacer frente a una situación extremadamente estresante;
- automutilación, comportamiento parasuicida y autolesiones no suicidas (por ejemplo, cortarse, quemarse partes limitadas del cuerpo con cerillas, mecheros o velas, golpearse, romperse los huesos, ingerir objetos o productos químicos nocivos, etc.);
- la presencia clara de un plan o ideación suicida altamente letal, es decir, con probabilidad de provocar la muerte de forma rápida e irreversible (por ejemplo, ahorcándose o pegándose un tiro).

Es importante conocer los intentos de suicidio o los comportamientos autolesivos anteriores. Esta información podría recogerse preguntando al sujeto de forma directa, pero sensible sobre dichos comportamientos (por ejemplo, «¿alguna vez ha pensado en hacerse daño?», «¿alguna vez se ha autolesionado voluntariamente?»).

NO: El individuo declara que nunca ha intentado o planeado suicidarse, y no hay evidencia para sospechar ninguna de estas conductas autolesivas o suicidas.

?: Posible o probable presencia de intentos de suicidio o conductas autolesivas en el pasado; o si el individuo no admite directamente haber realizado un intento de suicidio/autoagresión pero otras fuentes de información son indicativas de dicho comportamiento.

SÍ: El individuo admite haber intentado suicidarse o haberse autolesionado, o existen pruebas claras de autolesión.

FACTORES DE PERSONALIDAD

Estas características se refieren al comportamiento habitual del interno en cualquier situación a partir de los 16 años, tanto en libertad como en prisión.

38. Actitudes procriminales o valores antisociales. Actitudes manifiestas de naturaleza antisocial, características de subculturas delictivas que apoyan o justifican explícitamente el uso de la violencia y el comportamiento delictivo. Considerar actitudes sádicas, homicidas, paranoides, xenófobas, misóginas o sexistas, que no derivan de un trastorno mental.

Para codificar este factor de riesgo, hay que determinar en qué medida las actitudes, valores, creencias y pensamientos predominantes del individuo son antisociales o prosociales. Hay que explorar las actitudes del sujeto hacia otras personas, instituciones y servicios sociales, la ley y las autoridades. Hay que evaluar las actitudes hacia actos delictivos pasados y si el individuo muestra signos de remordimiento sincero o, por el contrario, no tiene remordimientos, es insensible y carece de empatía hacia las posibles víctimas.

No se trata de una actitud pesimista ocasional, sino de aquellas creencias y actitudes negativas y antisociales intensas y arraigadas que se convierten en actos violentos. Algunos individuos declaran, en mayor o menor medida, su propensión antisocial en función de los grupos u organizaciones sociales a los que pertenecen.

NO: El individuo valora las interacciones con otras personas, atribuye funciones positivas y beneficiosas a las actividades *convencionales* (estudiar, trabajar, hacer deporte, controlar el consumo de alcohol...). También valora favorablemente las instituciones de la sociedad (gobierno, empresa, trabajo, sistemas de servicios, hogar y familia, escuela y trabajo, etc.) y los principios que las justifican.

?: Existen pruebas dudosas o parciales de que el sujeto tiene valores prosociales. Critica el sistema de justicia penal y cree que en algunos casos el uso de la violencia puede estar justificado, pero en otros es crítico con su uso.

SÍ: El sujeto valora positivamente la utilidad de la actividad delictiva frente a las actividades prosociales. Acepta a otros delincuentes, sus valores y actividades, y expresa hostilidad hacia el sistema de justicia penal. Puede mostrar actitudes sádicas, homicidas o paranoides que no son consecuencia de un trastorno mental; así como actitudes o comportamientos discriminatorios o que justifican el uso de la violencia, actitudes racistas, xenófobas, misóginas y sexistas.

39. Baja capacidad mental e inteligencia: Rendimiento inadecuado en nuevas tareas y procesos de aprendizaje. Escaso dominio del idioma, aunque conoce la lengua que se habla en el centro. No muestra buena comprensión lectora y no es eficaz en el seguimiento de instrucciones escritas.

La persona con un bajo nivel de inteligencia tiene más dificultades para adaptarse a nuevas situaciones y es fácilmente manipulable e influenciable por otras personas. Su aprendizaje es lento, aunque esté motivado. Esta aparente incapacidad también se observa en la adquisición de nuevos hábitos sociales e interpersonales, parecen tener pocos conocimientos para su edad y experiencia biográfica. En el caso de inmigrantes con escaso dominio del catalán o castellano, hay que asegurarse de que la codificación se basa en la capacidad real del sujeto, sin que interfieran las dificultades de comunicación entre evaluador y evaluado o el desconocimiento del medio por razones culturales.

NO: Inteligencia superior a un CI de 85.

?: A pesar de tener dificultades intelectuales (por ejemplo, analfabetismo funcional) no hay pruebas concluyentes de que su capacidad intelectual sea deficiente.

SÍ: Inteligencia baja, por debajo de un CI de 85.

40. Temeridad. Disposición temperamental que se traduce en una aparente necesidad crónica de actividades de riesgo, nuevas experiencias y rechazo de la actividad rutinaria o aislada. Preferencia por las actividades en grupo y por tener compañeros similares a él/ella.

Es activo y a veces muestra una necesidad crónica y excesiva de estímulos nuevos y excitantes, así como una predisposición inusual al aburrimiento. Expresa un gran interés por las situaciones de riesgo, por «vivir la vida al límite» o «estar donde

está la acción» y por hacer cosas excitantes, arriesgadas y peligrosas. Puede probar y consumir muchos tipos de drogas. Se queja con frecuencia de que la escuela, el trabajo y las relaciones a largo plazo son aburridas y tediosas. Puede manifestar un abandono frecuente de las actividades que no puede explicarse por motivos económicos, de enfermedad o de otro tipo.

NO: No tiende a realizar actividades de riesgo y nuevas experiencias, ni a rechazar las actividades rutinarias y cotidianas.

?: Evidencia parcial o dudosa de una necesidad crónica de actividades de riesgo, nuevas experiencias, así como un marcado rechazo de las actividades rutinarias.

SÍ: Evidencia clara de una necesidad crónica de actividades arriesgadas, excitantes y peligrosas y una predisposición inusual al aburrimiento.

41. Impulsividad e inestabilidad emocional: Propensión a mostrar fluctuaciones dramáticas en el estado de ánimo o en el comportamiento en general. Propensión a reaccionar de forma exagerada, repentina y explosiva, tanto conductual como emocionalmente. Las respuestas impulsivas no se ajustan a lo que se espera del sujeto en determinadas circunstancias. Lleva un estilo de vida caracterizado por la inestabilidad en las relaciones interpersonales, el trabajo o las residencias.

La impulsividad implica cambios marcados en el estado de ánimo o en el comportamiento general que suelen ser exagerados o inadecuados para las situaciones que los generan; supone la incapacidad de mantener la compostura y el control en situaciones de estrés y acción.

La persona impulsiva muestra un comportamiento no premeditado, sin reflexión ni premeditación. Suele hacer las cosas «en el momento», porque «lo siente» o porque «se le presenta la oportunidad». No es capaz de evaluar de antemano las consecuencias de una acción para sí mismo o para los demás. Con frecuencia rompe relaciones, deja el trabajo, cambia de planes repentinamente o se traslada de un lugar a otro. Puede describirse como irascible, explosivo, brusco y desequilibrado. Tiende a responder a la frustración, los errores, la disciplina o las críticas con violencia o amenazas y agresiones verbales. Estos comportamientos suelen parecer inapropiados, dado el contexto en el que se producen.

El individuo con un alto nivel de impulsividad tiende a (sobre)reaccionar ante desaires, insultos y decepciones, ya sean reales o imaginarios. Generalmente, experimentan frustración ante los contratiempos. Les resulta difícil mantener la compostura y el control en situaciones de estrés y cansancio o fatiga importantes. Pueden mostrar respuestas «irracionales», incongruentes con su capacidad mental o nivel de formación.

NO: Muestra un estado de ánimo estable, responde adecuadamente a las situaciones, aprecia las consecuencias de sus actos, no se frustra fácilmente y mantiene la compostura y el control en situaciones de estrés y acción.

?: Existen pruebas parciales o equívocas de que el sujeto presenta una disposición a mostrar cambios marcados en el estado de ánimo o en el comportamiento, así como una tendencia a responder de forma impulsiva e irracional ante las situaciones.

SÍ: Evidencia clara de que el sujeto tiene una disposición a mostrar marcadas fluctuaciones en el estado de ánimo o el comportamiento, a responder impulsivamente a las situaciones y de forma agresiva y con frustración a los contratiempos. Las relaciones son inestables en diversos ámbitos (por ejemplo, laboral, social, familiar).

42. Hostilidad: Disposiciones y actitudes desfavorables y comportamiento agresivo hacia los demás. Generalmente, desconfiados y hostiles hacia todo lo que les rodea. Sus actitudes y comportamientos indican una profunda falta de empatía, crueldad e indiferencia hacia los sentimientos, derechos y bienestar de los demás. Se siente fácilmente atacado y se enfada o se muestra agresivo por trivialidades.

Estas personas ven el mundo de forma hostil, son suspicaces y desconfían de los demás, tienden a pensar lo peor, pueden sentirse atacadas con facilidad y se enfadan o se vuelven agresivas por nimiedades. La agresividad tiende a producirse como consecuencia frecuente o covariante de la hostilidad. El individuo tiende a racionalizar y justificar las violaciones de la ley, puede negar o minimizar la conciencia y la responsabilidad de sus actos (por ejemplo, «no fue culpa mía», «nadie resultó herido» o «la víctima me buscó, me provocó»). Le preocupan más los efectos que sus actos puedan tener sobre sí mismo que el sufrimiento que pueda causar a sus víctimas o el daño que pueda causar a la sociedad. Manifiesta un rechazo frontal o es incapaz de empatizar con los sentimientos, deseos y expectativas de los demás, incluidas las víctimas de sus delitos.

También pueden mostrar poca receptividad a las ideas, comentarios o sentimientos de los demás, por ejemplo, a la preocupación de su familia o amigos por su encarcelamiento. Son vanidosos, manipuladores y poco comprensivos. Suele tener pocos amigos, ya que en general es poco servicial con los demás.

NO: El individuo no expresa ira ni hostilidad, es sensible a las consecuencias negativas de las infracciones y acepta la responsabilidad de sus actos.

?: Tiene cierta disposición y actitudes desfavorables hacia ciertas personas, pero no en general. Rara vez es hostil, es consciente de la responsabilidad de sus actos, pero justifica algunos de ellos. Aunque tiene dificultades para empatizar con la gente, su círculo social es amplio.

SÍ: Hay pruebas claras de que el sujeto muestra una disposición y actitud desfavorables, y un comportamiento agresivo hacia los demás. Tiende a justificar sus actos, es incapaz de empatizar con los demás y se muestra hostil, cruel y manipulador. Tiene un círculo de amigos muy reducido.

43. Irresponsabilidad: Voluntad de no cumplir las obligaciones o compromisos con los demás. Escaso sentido de la responsabilidad o lealtad hacia la familia, los amigos, el trabajo, la sociedad, las ideas o las causas. Dificultad para tomar decisiones y para adquirir y cumplir compromisos y responsabilidades adecuados a su edad. Puede mostrar un estilo de vida parasitario de su entorno inmediato. No acepta la responsabilidad de sus propios actos, incluidas las actividades delictivas por las que ha sido procesado o condenado.

La irresponsabilidad puede evidenciarse en diversas áreas como: financiera (no devolver préstamos, no pagar deudas, etc.), comportamientos que ponen en peligro

a otras personas (conducir ebrio, exceso de velocidad recurrente, etc.), desempeño laboral (llegar tarde o ausentarse con frecuencia, descuidar tareas, aunque pueda hacerlas mejor, etc.), área interpersonal y familiar (no es respetuoso como padre o cónyuge, poco cuidadoso con sus hijos, etc.). Muestra una notable inmadurez personal. A pesar de ser capaz, evita el trabajo estable y remunerado. Se aprovecha de familiares, conocidos, amigos o ayudas sociales para mantenerse. También incluye a las personas que eluden la responsabilidad de sus actos ilícitos racionalizando su conducta (por ejemplo, proyectando la culpa del delito en la víctima o en las circunstancias), y minimizando sus consecuencias para los demás (por ejemplo, negando que la víctima haya sufrido consecuencias físicas, emocionales o económicas graves o directas) o negando sus actos (por ejemplo, afirmando que son inocentes, que la víctima miente o que fueron engañados).

NO: El individuo cumple con sus obligaciones o compromisos familiares, laborales, sociales, etc. El nivel de madurez, compromiso y responsabilidad es adecuado para su edad.

?: Existen evidencias dudosas o parciales del sentido de la responsabilidad del sujeto, cumpliendo únicamente con determinadas obligaciones o compromisos. El nivel de madurez, compromiso y responsabilidad no alcanza el nivel deseable para su edad.

SÍ: Hay pruebas claras de que el sujeto, en general, no cumple con sus obligaciones en muchos ámbitos (por ejemplo, social, familiar, laboral, etc.). No acepta la responsabilidad de sus actos, se aprovecha de los demás para satisfacer sus necesidades y es incapaz de mantener un trabajo estable. Muestra un nivel de madurez, compromiso y responsabilidad muy deficiente para su edad.

2. RISCANVI SCREENING

Resumen elaborado por Lucía Martínez Garay a partir de Antonio Andrés Pueyo; Karin Arbach Lucioni y Santiago Redondo Illescas (2009): *RisCanvi Versió Screening. Definició d'ítems (pre-edició)*, que nos facilitó la Generalitat de Catalunya en febrero de 2023. Salvo error, es el mismo documento que en 2024 la Generalitat ha colgado en su web junto a los resultados del informe de auditoría (<<https://links.uv.es/Ty6ovdCTy6ovdC>>).

RisCanvi Screening (RS) consta de 10 ítems. La descripción de algunos de ellos es igual o muy similar a la de los ítems que forman parte de *RisCanvi Completo* (RC), y los demás combinan en un solo ítem lo que en RC son dos factores separados. Por este motivo, no incluiremos la descripción de los factores del RS (que puede encontrarse en el documento que describe los factores del RC), sino que nos limitaremos a señalar a cuál/es de los factores del RC corresponde cada factor del RS.

Por otra parte, a diferencia del CR, el RS no da tres alternativas de respuesta para codificar cada ítem, sino sólo dos (sí o no); por tanto, describiremos la redacción de estas dos alternativas para cada uno de los 10 factores del RS.

1. Inicio de la actividad delictiva o violenta: Edad del sujeto en el momento del primer delito o incidente violento (punto 8 del CR).

NO: El primer delito o el inicio del comportamiento violento se produjo a la edad de 17 años o más.

SÍ: El primer delito o el inicio del comportamiento violento se produjo antes de los 16 años.

2. Historial de violencia: delito base violento o historial de comportamiento violento previo (puntos 1 y 7 en RC).

NO: Ni el delito base fue violento ni hay pruebas de conducta violenta anterior al delito base.

SÍ: El delito base fue violento o existen pruebas claras de antecedentes de conducta violenta anteriores al delito base.

3. Problemas de comportamiento en prisión: Faltas disciplinarias graves o muy graves y otros problemas de adaptación institucional en prisión. Considerar cualquier momento de la vida penitenciaria del sujeto (ítems 10 y 12 en RC).

NO: El interno es primario (primera vez que ingresa en prisión) o no hay indicios de mala conducta penitenciaria previa.

SÍ: El interno tiene problemas interpersonales con otros internos, ha sido sancionado por faltas disciplinarias graves o muy graves (violentas y no violentas), ha estado implicado en disturbios o incidentes institucionales graves o ha utilizado violencia, armas o amenazas contra el personal u otros internos.

4. Evasiones, quebrantamientos o incumplimientos: Antecedentes de evasiones, fugas o quebrantamientos de condena, hayan sido o no exitosos. El interno ha disfrutado de permisos penitenciarios y ha incumplido las normas que regulan dichos permisos u otras medidas penales. Considerar cualquier momento de su vida penitenciaria (puntos 11, 13 y 15 del RC).

NO: No hay constancia de que el interno se haya fugado (o intentado hacerlo) de un centro, haya quebrantado un permiso penitenciario y haya infringido las normas de conducta de una medida penal. No hay constancia de que se hayan impuesto sanciones disciplinarias por estos comportamientos.

SÍ: Existen indicadores claros de fugas, incumplimientos de permisos o quebrantamientos de medidas penales (ya sean graves o moderados).

5. Problemas con el consumo de drogas o alcohol: Consumo de drogas, legales o ilegales, o de alcohol que interfiere significativamente en diferentes aspectos de su vida (familiar, laboral o social) o en su salud física. Considere el último año (ítems 30 y 31 en RC).

NO: El sujeto no consume drogas ni alcohol, o lo hace de forma ocasional y controlada, y el consumo de sustancias o alcohol no interfiere en su comportamiento y relaciones habituales (por ejemplo, mantenimiento del empleo, ausencia de conflictos conyugales o familiares debidos al consumo de drogas o alcohol, etc.). También

se valorará con un «No» a quienes, habiendo consumido en el pasado se han mantenido abstinentes durante el último año.

SÍ: Evidencia de deterioro en su funcionamiento psicosocial debido al abuso de drogas o alcohol (por ejemplo, beber hasta intoxicarse, centrar sus actividades sociales en el consumo de sustancias, cometer delitos debido al consumo de sustancias, etc.).

6. Respuesta limitada al tratamiento psicológico o psiquiátrico: El sujeto ha recibido tratamiento psicológico o psiquiátrico (o farmacológico), o por abuso de sustancias, y no ha mostrado adherencia o los resultados han sido pobres. Considerar el último año (ítem 34 en RC).

NO: La respuesta al tratamiento recibido es adecuada y favorable, o el sujeto no ha recibido tratamiento porque no presenta ninguna de las disfunciones o síntomas mencionados anteriormente.

SÍ: El sujeto no ha mostrado respuesta al tratamiento o ésta ha sido escasa; se ha negado a iniciarlo; lo ha comenzado, pero lo ha abandonado; ha fingido seguirlo o lo ha cumplido sin ningún beneficio; no ha recibido tratamiento a pesar de que sus disfunciones o síntomas apoyan su conducta infractora.

7. Intentos o comportamientos autolesivos: Conductas autolesivas, como autolesiones o intentos de suicidio, que han puesto en riesgo la vida o la integridad física, independientemente de su intención. Considerar cualquier momento de su vida (ítem 37 en RC).

NO: El individuo declara que nunca ha intentado o planeado suicidarse y no hay evidencia de autolesiones o comportamiento suicida.

SÍ: El individuo admite haber intentado suicidarse o haberse autolesionado, o existen pruebas claras de tal intento o autolesión.

8. Falta de recursos económicos: Nivel económico estimado en el último año o antes del ingreso en el caso de haber pasado más de 12 meses en prisión (ítem 20 en RC).

NO: Dispone de recursos económicos suficientes para cubrir sus necesidades básicas (por ejemplo, alimentación o vivienda).

SÍ: El interno no dispone de apoyo económico para cubrir sus necesidades básicas. Su subsistencia o la de su familia dependen de ayudas sociales.

9. Falta de apoyo familiar y social: Falta de red social, de relaciones satisfactorias y de apoyo o de contacto regular con la familia y los amigos. Considerar el último año (ítem 24 en RC)

NO: El sujeto mantiene relaciones frecuentes y satisfactorias con otras personas como padres, hermanos, abuelos, cónyuge. Informa de que tiene amigos íntimos o familiares que le ofrecen apoyo, lo cual agradece.

SÍ: No existen relaciones positivas ni apoyo de familiares, no hay amigos íntimos o familiares a los que recurrir, o este apoyo (emocional, financiero o físico) existe, pero el sujeto no quiere aceptarlo. Indique también la presencia del ítem si las redes y el apoyo social del individuo están relacionados con sus actos violentos y delictivos.

10. Actitud hostil o valores procriminales: Actitudes manifiestas de carácter antisocial, típicas de subculturas criminales que apoyan o justifican explícitamente el uso de la violencia y el comportamiento delictivo. Falta de empatía e indiferencia hacia los sentimientos, derechos y bienestar de los demás. *Considerar el momento presente* (ítems 38 y 42 del RC).

NO: El individuo es sensible a las consecuencias negativas de los delitos y acepta la responsabilidad de sus actos, muestra empatía con la víctima y rechaza a otros delincuentes. Valora las interacciones con otras personas, atribuye funciones positivas y beneficiosas a las actividades convencionales (estudiar, trabajar, hacer deporte, controlar el consumo de alcohol...), acepta las instituciones de la sociedad (gobierno, trabajo, familia, etc.) y sus valores subyacentes.

SÍ: El sujeto valora positivamente la utilidad de la actividad delictiva frente a las actividades prosociales. Acepta a otros delincuentes, sus valores y sus actividades. No acepta o minimiza la responsabilidad de sus actos, es incapaz de empatizar con los demás y expresa hostilidad hacia el sistema de justicia penal. Muestra actitudes sádicas, homicidas o paranoides que no son consecuencia de un trastorno mental; así como actitudes o comportamientos agresivos, discriminatorios, racistas o sexistas que justifican el uso de la violencia.

ADENDA: PREGUNTAS Y RESPUESTAS

Adenda al informe *Tres enfoques policiales predictivos en España*:

Una primera versión de este documento fue revisada por una persona con experiencia en el ámbito policial y de los derechos humanos, que planteó una serie de preguntas sobre su contenido. En la medida en que se trata de preguntas pertinentes, que nos permitieron aclarar algunos aspectos importantes del informe, hemos incluido aquí algunas de esas preguntas junto con nuestras respuestas, pues creemos que pueden contribuir a una mejor comprensión de este trabajo.

1. PREGUNTAS ESPECÍFICAS SOBRE VIOGÉN

1. «Las evaluaciones de riesgos se realizan a través de dos formularios: el VPR 5.0 (valoración policial del riesgo) y el VPER 4.1 (valoración policial de la evolución del riesgo)».

COMENTARIO: Si he entendido bien, las VPR 5.0 y la 4.1 se aplican en paralelo, ¿verdad? ¿Sabemos cómo llegaron a la conclusión sobre los diferentes indicadores para la 5.0 (la evaluación de los diferentes indicadores)? Supongo que no se debió a los malos resultados de la 4.1, sino a la diferente perspectiva de la VPR 5.0: = riesgo potencial grave, por ejemplo, de homicidio, ¿correcto?

No, no se aplican en paralelo. El VPR 5.0 se utiliza para la primera evaluación del riesgo, y el VPER 4.1 (que significa «valoración policial de la evolución del riesgo») se utiliza en las evaluaciones posteriores de las mujeres, que deben hacerse periódicamente (con mayor o menor frecuencia en función del nivel de riesgo). Y el VPR 5.0 no solo evalúa el riesgo de violencia letal: evalúa ambos, el riesgo de cualquier nueva agresión y (simultáneamente) el riesgo de homicidio. VPR 5.0 tiene dos escalas que funcionan simultáneamente (la escala normal y la escala H de riesgo de homicidio, que trabaja con un subconjunto de los mismos factores de riesgo y los pondera de forma diferente).

2. «Parece que al menos algunos de los procedimientos internos de VioGén se realizan con machine learning, o al menos así se está implementando. Según una nota de prensa hecha pública el 15 de diciembre de 2020 por la empresa SAS Iberia: “[...] el nuevo ‘agente digital’, la plataforma de analítica, apoya a las fuerzas de seguridad en la ardua tarea de analizar grandes cantidades de datos, automatizando procesos que antes consumían recursos humanos, que ahora pueden destinarse a labores de protección de forma más eficiente. [...]”».

COMENTARIO: Esta argumentación (gran cantidad de datos) no es realmente válida para la violencia de género...

Depende. No sabemos exactamente con qué propósito están implementando el aprendizaje automático, pero VioGén contiene una base de datos muy grande. Aunque por ahora hay algo más de 73.000 casos activos, muchos más casos fueron evaluados en su día y ahora están inactivos, y cada mujer (cada «caso») se somete no solo a una evaluación, sino a veces a muchas más (evaluaciones periódicas con el VPER). Según un comunicado de prensa del Ministerio del Interior, «VioGén ha analizado 701.563 casos y realizado más de 5,4 millones de valoraciones policiales de riesgo desde su creación» (<<https://www.interior.gob.es/opencms/gl/detalle/articulo/VioGen-cumple-15-anos-con-mas-700.000-casos-analizados-y-54-millones-de-valoraciones-de-riesgo-realizadas/>>), por lo que el número de evaluaciones realizadas (no la cantidad de datos relativos a cada mujer, sin embargo) puede ser, en efecto, bastante grande.

3. «Junto a VioGén, los otros elementos que tienen en cuenta los jueces son las declaraciones de la víctima y del acusado, los antecedentes penales, la gravedad de los hechos, el tipo de delito cometido, la reincidencia, el apoyo social de la víctima y del presunto agresor, los informes médicos de lesiones, las declaraciones de testigos, el consumo de drogas y alcohol y la presencia de enfermedades mentales».

COMENTARIO: ¿Pero no son ya factores que forman parte de los indicadores para calcular el nivel de riesgo con VioGén?

Sí, lo son. Pero de todas formas los jueces hacen su propia valoración y dan a estos factores la relevancia que consideran oportuna. Tal vez no sean conscientes de que VioGén ya trabaja con esta información o, si lo son, siguen queriendo hacer su propio juicio personal sobre su importancia.

4. «VioGén tampoco determina si los jueces imponen al acusado un dispositivo de vigilancia telemática. Para ello, tienen en cuenta no solo el nivel de riesgo, sino también el resto de la información disponible sobre el caso. Afirman que hay circunstancias en las que no es aconsejable colocar el dispositivo que no están relacionadas con el nivel de riesgo, por ejemplo, cuando la víctima no quiere llevarlo o cuando se trata de una indigente».

COMENTARIO: ¿Por qué la víctima? ¿No se trata del acusado?

Se trata ante todo de la víctima. Por supuesto, no se impone la pulsera de vigilancia telemática si el juez considera que el riesgo no es lo suficientemente alto como para justificarlo. Pero, si el juez decide imponerla, no es necesario el consentimiento del acusado: es una medida coercitiva para él como el resto de las medidas cautelares. Lo que sí es necesario es el consentimiento de la mujer víctima: si no

está de acuerdo no se puede imponer, porque el funcionamiento del sistema exige que ella también lleve un dispositivo. Y algunas mujeres prefieren no llevarlo por varias razones: bien para proteger su intimidad (no quieren que se vea, y que se sepa que están con una orden de protección), bien para evitar los sustos, porque cuando hay control telemático ella es avisada cada vez que el acusado se acerca a una distancia inferior a la estipulada por el juez. Esto a veces ocurre por casualidad (sobre todo cuando ambos viven en pueblos pequeños) y genera muchas falsas alarmas que asustan a la mujer.

2. PREGUNTAS ESPECÍFICAS SOBRE RISCANVI

1. «El punto de partida para el desarrollo del sistema RisCanvi fue un estudio retrospectivo de casos y controles con una muestra de 643 casos y 428 controles. El estudio analizó la prevalencia de cuatro criterios de riesgo relacionados con la violencia en las prisiones e identificó los factores relevantes para cada uno de ellos».

COMENTARIO: ¿Sabéis cuáles? ¿Y cómo se identificaron?

La explicación detallada del proceso de construcción de RisCanvi figura en un informe técnico que, aunque se cita en varias publicaciones (citado como: Andrés-Pueyo, Arbach-Lucioni y Redondo: *Informe de elaboración del protocolo de valoración y gestión del riesgo de violencia en las prisiones RisCanvi*, Departamento de Justicia de la Generalitat de Catalunya, 2010), no es de acceso público, por lo que no hemos podido consultarlo para este informe. No podemos aportar más información sobre estos «cuatro criterios de riesgo».

2. «La creación de RisCanvi incluyó un análisis retrospectivo de 106 factores de riesgo de delincuencia y de violencia en casos y controles obtenidos de cuatro fuentes: otros protocolos de evaluación de riesgo penitenciario preexistentes (LSI-R, COMPAS, FOTRES, OASYS, HCR-20, etc.); [...]».

COMENTARIO: ¿Es problemático que RisCanvi se diseñara teniendo en cuenta herramientas ya existentes que se habían revelado problemáticas debido a sesgos y otros problemas (por ejemplo, se descubrió que Compas mostraba resultados muy sesgados en perjuicio de los estadounidenses de raza negra)?

Todas las herramientas actuariales de valoración de riesgo tienen el problema de los sesgos sistémicos, es decir, los datos con los que trabajan son un reflejo de la sociedad en la que se realizan las evaluaciones, y toda sociedad está afectada por una serie de desigualdades y discriminaciones estructurales derivadas de la evolución histórica y cultural y de la situación socioeconómica de cada país o región. En una sociedad puede haber minorías que son y han sido históricamente discriminadas, por ejemplo, los afroamericanos en EE. UU. o las poblaciones gitanas en Europa, o los inmigrantes en muchos países: las posibilidades de quienes pertenecen a estos grupos para acceder a empleos legítimos son menores, están sometidos a una vigilancia policial más intensa, tienen menos recursos económicos y culturales para disponer de una defensa jurídica competente, etc., y esto puede llevar a que los índices oficiales de delincuencia de estas minorías sean superiores a los de la media, o a que también lo sea su representación en términos relativos en la población reclusa.

Si este es el caso, un algoritmo que se haya construido a partir de los datos existentes sobre la delincuencia en ese país o región aprenderá que ser negro, gitano, inmigrante (o enfermo mental, o pobre, etc.) correlaciona con una mayor probabilidad de tener antecedentes penales, de pasar más tiempo en prisión o de ser detenido de nuevo y condenado, y en consecuencia considerará que la presencia de estos factores aumenta el riesgo de delinquir, y asignará a los pertenecientes a estos grupos un mayor riesgo de reincidencia.

Es lo que se conoce como sesgos sistémicos, y son un problema desde el punto de vista de los derechos fundamentales, porque asignar un mayor riesgo de reincidencia o violencia a quienes pertenecen a minorías que ya sufren desigualdad es injusto, pues supone tratar a las personas de forma diferente (y peor, porque a mayor riesgo, mayor control policial/penal) por factores por los que las normas nacionales e internacionales de derechos humanos prohíben discriminar, y que no están vinculados a su responsabilidad individual o a sus decisiones libres. Esta discriminación por sesgo sistémico puede producirse tanto directa como indirectamente: directamente cuando el algoritmo considera de manera explícita el factor problemático como factor de riesgo y aumenta la puntuación cuando está presente; e indirectamente cuando, sin que el factor aparezca como elemento en la herramienta, resulta que los resultados de las evaluaciones de riesgo perjudican de manera sistemática a unos grupos frente a otros.

El problema saltó a la opinión pública hace años de forma especialmente intensa en relación con Compas y la discriminación de los afroamericanos, cuando en 2016 Pro Publica publicó un informe en el que se señalaba que los afroamericanos que no reincidían eran, sin embargo, considerados de alto riesgo en un porcentaje muy superior al de los blancos que no reincidían. Nótese que en este caso la discriminación era indirecta: ni Compas ni ninguna otra herramienta utilizada en EE. UU. considera la raza como factor de riesgo porque, según la opinión mayoritaria, sería inconstitucional, pero los resultados para los afroamericanos eran sistemáticamente peores que para los blancos.

Estas son las cifras analizadas por ProPublica:

Blancos (tasa de reincidencia del 39 %)

	<i>Reincidencia</i>	<i>No reincidencia</i>	<i>Total</i>	
<i>Riesgo medio o alto</i>	505	349	854	59 % Valor predictivo positivo
<i>Riesgo bajo</i>	461	1.139	1.600	71 % Valor predictivo negativo
<i>Total</i>	966	1.488	2.454	
	Sensibilidad 52,27	Especificidad 76,54		
		Tasa de falsos positivos: 23,5 %.		

Afroamericanos (tasa de reincidencia del 52 %)

	<i>Reincidencia</i>	<i>No reincidencia</i>	<i>Total</i>	
<i>Riesgo medio o alto</i>	1.369	805	2.174	63 % Valor predictivo positivo
<i>Riesgo bajo</i>	532	990	1.522	65 % Valor predictivo negativo
<i>Total</i>	1.901	1.795	3.696	
	Sensibilidad 72	Especificidad 55,1		
		Tasa de falsos positivos: 44,9 %		

Fuente: <<https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>>.

Y estas son las conclusiones de ProPublica:

«La predicción falla de forma diferente para los acusados negros»

	<i>Blanco</i>	<i>Afroamericanos</i>
<i>Etiquetado como de alto riesgo, pero no reincidió</i>	23,5 %	44,9 %
<i>Etiquetado como de bajo riesgo, pero reincidió</i>	47,7 %	28.0%

En general, la herramienta de evaluación de Northpointe predice correctamente la reincidencia en el 61 % de las ocasiones. Sin embargo, los negros tienen casi el doble de probabilidades que los blancos de ser considerados de alto riesgo sin llegar a reincidir. El error es el contrario entre los blancos: Éstos tienen muchas más probabilidades que los negros de ser etiquetados como de menor riesgo, pero llegar a cometer otros delitos. *Fuente:* análisis de ProPublica de datos del condado de Broward, Florida; citado de: <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.

ProPublica señaló que entre los acusados que finalmente no reincidieron, los negros tenían más del doble de probabilidades que los blancos de ser clasificados como de riesgo medio o alto: 44,9 % negros y 23,5 % blancos. Y que entre los acusados que sí reincidieron solo el 52,27 % de los blancos habían sido clasificados de riesgo medio o alto, mientras que el 72 % de los negros lo habían sido. ProPublica argumentó que un algoritmo justo no debería cometer estos graves errores con más frecuencia para un grupo racial que para el otro.

Pero Northpointe (la empresa que había creado la herramienta) argumentó que las puntuaciones de Compas eran equitativas entre grupos raciales, porque significaban lo mismo independientemente de la raza de los acusados. Por ejemplo, entre los acusados blancos que habían obtenido una puntuación de riesgo medio o alto, el 59 % reincidía, y entre los negros que habían obtenido la misma puntuación reincidían en un porcentaje muy similar, el 63 % (valores predictivos positivos).

Ambos son criterios de equidad relevantes (mantener las tasas de error comparables entre grupos y tratar de la misma manera a las personas con las mismas puntuaciones de riesgo), y el problema es que no es posible satisfacer ambos al mismo tiempo, si cada uno de los dos grupos tiene tasas de reincidencia diferentes, como ocurre en la muestra analizada por ProPublica.

Si la tasa de reincidencia de los acusados blancos y negros es la misma dentro de cada categoría de riesgo, y si los acusados negros tienen una tasa de reincidencia global más alta, entonces una mayor proporción de acusados negros será clasificada como de alto riesgo. Y si una mayor proporción de acusados negros son clasificados como de alto riesgo, entonces, como ilustra el gráfico, una mayor proporción de acusados negros que no reinciden también serán clasificados como de alto riesgo (Corbett-Davies, Pierson, Feller y Goel: «A computer program used for bail and sentencing decisions was labelled biased against blacks. It's actually not that clear», *The Washington Post*, 17 de octubre de 2016, en línea: <<https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/>>).

El problema está muy bien ilustrado en un gráfico interactivo en Hao y Stray: «Can you make AI fairer than a judge? Play our courtroom algorithm game», *MIT Technology Review*, 2018, en línea: <<https://www.technologyreview.com/2019/10/17/75285/ai-fairer-than-judge-criminal-risk-assessment-algorithm/>> (consulta: 17/10/2019). Para mantener las tasas de error comparables entre los dos grupos, es necesario establecer un umbral diferente entre bajo y alto riesgo para cada uno de ellos, por ejemplo, que los acusados blancos se consideren de alto riesgo con una puntuación de riesgo de 7, pero que los acusados negros se consideren de alto riesgo solo con una puntuación de riesgo de 9. Pero esto, «una vez más, no parece justo. Dos personas con la misma puntuación de riesgo tienen la misma probabilidad de volver a ser detenidas, así que ¿no deberían recibir el mismo tratamiento?» (Hao y Stray, 2019).¹

Obviamente, las diferentes tasas de reincidencia entre negros y blancos se deben a la historia de EE. UU., en la que la policía comete a un control mucho más intenso a las minorías. Como explican Hao y Stray (2019),

las predicciones reflejan los datos utilizados para hacerlas, ya sea mediante algoritmos o no. Si los acusados negros son arrestados en una tasa más alta que los

¹ Véase también Chouldechova y Roth: *The Frontiers of Fairness in Machine Learning*, 2018, p. 3: «Diferentes medidas estadísticas de equidad pueden estar en desacuerdo entre sí. Por ejemplo, [...]: excepto en entornos triviales, es imposible igualar simultáneamente las tasas de falsos positivos, las tasas de falsos negativos y el valor predictivo positivo entre los grupos protegidos» (<<https://arxiv.org/pdf/1810.08810.pdf>>).

acusados blancos en el mundo real, también tendrán una tasa más alta de arrestos predichos. Esto significa que también tendrán puntuaciones de riesgo más altas por término medio, y que un porcentaje mayor de ellos serán etiquetados como de alto riesgo, tanto correcta como incorrectamente. Esto es cierto independientemente del algoritmo que se utilice, siempre que esté diseñado de forma que cada puntuación de riesgo signifique lo mismo independientemente de la raza.

En resumen, los sesgos sistémicos son un gran problema en cualquier algoritmo de valoración del riesgo. Creemos que Compas probablemente no es mejor ni peor que cualquier otro algoritmo en lo que respecta a la raza, en la medida en que las disparidades raciales que ProPublica señaló en relación con esta herramienta serán las mismas o similares en cualquier otra herramienta que tenga que trabajar con grupos que tengan diferentes tasas de reincidencia. Y que, por tanto, el hecho de que RisCanvi se haya inspirado (en parte) en Compas no es necesariamente un problema mayor que si se hubiera inspirado en otras herramientas.

No se trata de negar la importancia de los sesgos sistémicos, por supuesto. Existen diferentes soluciones técnicas que se han propuesto para evitar o tratar de reducir los resultados dispares entre los diferentes grupos (que requerirían una extensa explicación técnica),² y cabe señalar que cualquier solución tiene «costes», y normalmente reducen la exactitud de las predicciones. En cualquier caso, la cuestión relevante con respecto a RisCanvi es, en nuestra opinión, que –sabiendo, como tienen que saber, que cualquier herramienta algorítmica de evaluación de riesgos tiene sesgos sistémicos– ni sus diseñadores ni la Generalitat de Catalunya parecen haber analizado seriamente si los resultados de RisCanvi podrían ser diferentes entre grupos (por raza, por origen nacional, por estatus socioeconómico, etc.),³ lo que sería el primer paso para, a partir de ahí, empezar a pensar en posibles causas y posibles soluciones para este problema.

3. COMENTARIO: Hay muchos elementos que se prestan a interpretaciones muy subjetivas (por ejemplo, amigos delincuentes o antisociales). Incluso la situación económica puede interpretarse de forma diferente. Muchos de los criterios son en realidad criterios muy blandos, no datos duros, pero se convierten en supuestos datos duros, con el riesgo de «blanquear» gracias a la tecnología juicios muy subjetivos sobre la persona.

Compartimos su impresión de que, a pesar de todas las definiciones, ejemplos, directrices y reuniones de equipo, la información con la que RisCanvi tiene que trabajar es, por su propia naturaleza, imposible de convertir en datos puramente objetivos. Y que insistir demasiado en la objetividad de la herramienta lo olvida y puede, de hecho, equivaler a un «blanqueamiento tecnológico» del problema.

² Véase, por ejemplo, un resumen en J. R Bent: «Is Algorithmic Affirmative Action Legal?», *The Georgetown Law Journal*, (108), 2020, pp. 814-823; Chouldechova y Roth: *The Frontiers of Fairness in Machine Learning*, 2018, en línea: <<https://arxiv.org/pdf/1810.08810.pdf>>.

³ Los primeros análisis de este tipo que conocemos son los realizados por C. Castillo y cols. (<https://chato.cl/research/discrimination_bias>).

Sin embargo, hay que tener en cuenta que a RisCanvi hay que compararlo con lo que ocurría antes de su implantación, y con lo que sigue ocurriendo, por ejemplo, en el resto de España, donde la normativa de ejecución de penas es muy similar a la que se aplica en Cataluña, y por lo tanto las decisiones sobre libertad condicional, permisos, progresión de grado, etc., requieren que se hagan valoraciones similares. Quizás RisCanvi haya conseguido mejoras en el sentido de una mayor objetividad, como la homogeneización de criterios entre distintos profesionales y centros penitenciarios, por ejemplo, que antes no existieran o existieran en menor medida, y esto podría valorarse positivamente desde el punto de vista del derecho a la igualdad de trato. O, por otro lado, podría ser que RisCanvi haya empeorado la situación, porque se codifican los factores de forma mecánica por falta de tiempo o porque el personal penitenciario no confía plenamente en ellos, de manera que las evaluaciones se hacen con menos rigor que antes... En este punto contamos con la opinión de los abogados, que fueron muy críticos con RisCanvi y dijeron que, en su opinión, solo sirve para justificar y, de hecho, «blanquear» decisiones que el personal penitenciario ya ha tomado o quiere justificar de cualquier manera. Pero nos ha faltado un elemento importante para poder emitir un juicio más fundado al respecto, que hubiera sido entrevistar directamente a un número suficiente de trabajadores penitenciarios y contrastar sus opiniones no solo sobre RisCanvi, sino también sobre la comparación entre trabajar con y sin esta herramienta.

4. «En la evaluación profesional del riesgo de reincidencia violenta mediante RisCanvi, la evaluación puede complementarse con otros instrumentos más específicos en función de las características criminológicas del delito y del delincuente. En el caso de los delincuentes sexuales, es práctica común utilizar el SVR-20/RSVP junto con el RisCanvi, y en el caso de los delincuentes IPV (intimate partner violence) se suele añadir el SARA, una herramienta de evaluación más específica. Con delincuentes crónicos o delincuentes con antecedentes de trastornos mentales, se utiliza con frecuencia el HCR-20 V3, y en el caso de delincuentes juveniles con alto riesgo y delitos menos graves, el RisCanvi se combina a menudo con el PCL-SV para evitar falsos positivos de reincidencia violenta; estos delincuentes pueden tener antecedentes penales menos graves pero un alto riesgo potencial de reincidencia. El e-RisCanvi (la versión electrónica del protocolo) tiene un programa que aconseja a los usuarios complementar la evaluación del riesgo que ofrece el RisCanvi con estos protocolos, que tienen adaptaciones en español».

COMENTARIO: ¿Significa esto reconocer debilidades de RisCanvi?

Creemos que no. En el antiguo paradigma de la «peligrosidad» —que subyace tras los juicios clínicos—, el profesional emitía un juicio global e inespecífico sobre si el sujeto era peligroso o no. En cambio, una de las características de los instrumentos modernos de evaluación del riesgo es que intentan ser más específicos en cuanto al tipo de variable dependiente. Parten de la base de que existen diferentes tipos de delincuencia y de violencia, y que no todos están necesariamente asociados a los mismos factores de riesgo. Así, para un mismo sujeto, si lo que nos interesa conocer es la probabilidad de que reincida con un nuevo delito sexual, lo adecuado es aplicar un instrumento diseñado específicamente para estimar este tipo de riesgo, mientras que si lo que nos interesa conocer es la probabilidad de que reincida con la comisión

de cualquier tipo de delito, lo adecuado es aplicar una herramienta diseñada para estimar el riesgo de reincidencia genérico. El propio RisCanvi es un ejemplo de este enfoque: contiene cinco escalas diferentes porque cada una de ellas estima el riesgo de algo distinto, y, por ejemplo, no es lo mismo la escala de reincidencia genérica que la de reincidencia violenta.

Por tanto, el hecho de que RisCanvi aconseje en ocasiones aplicar otra herramienta más específica para determinados tipos de delincuentes (presos jóvenes, presos condenados por delitos sexuales, presos con patología mental, etc.) es coherente con el propio planteamiento desde el que se concibe RisCanvi y, en general, las herramientas de valoración del riesgo: que para determinados tipos de delitos/delincuentes existen factores de riesgo específicos. RisCanvi se pensó inicialmente para la delincuencia violenta y luego se extendió a la reincidencia en general, pero no incluye escalas más específicas para estimar el riesgo de reincidencia en tipologías concretas, por lo que en los casos en los que se piensa que hay algún riesgo específico que valorar, recomienda aplicar además una herramienta diseñada para estimar ese tipo de riesgo.

Otra cosa es que el propio enfoque suscite, precisamente por esta premisa de partida, algunas dudas. Si existen diferentes perfiles de delincuentes o tipos específicos de delitos con diferentes factores de riesgo, ¿tiene sentido aplicar alguna vez herramientas que estimen el riesgo global de reincidencia? Incluso dentro de categorías como «reincidencia violenta», no es lo mismo un robo con violencia que un homicidio o una agresión sexual. Y dentro de los «delitos sexuales» no es lo mismo abusar de menores, consumir pornografía infantil o agredir violentamente a mujeres. Por otro lado, a una misma persona se le pueden aplicar situaciones diferentes: ser joven, tener un trastorno mental y haber cometido delitos sexuales, por ejemplo. ¿Deben aplicarse varias herramientas específicas en estos casos? ¿Y qué hacer si no todas muestran el mismo nivel de riesgo?

5. «RisCanvi no es un instrumento bien calibrado: según los resultados de los datos de Karimi-Haghighi y Castillo, el porcentaje de sujetos clasificados como de alto riesgo en la muestra total (10,5 %) duplica al de los que realmente cometieron un nuevo delito violento en el periodo de seguimiento (4,9 %). Esta proporción sigue siendo similar para muchos de los subgrupos analizados. Eso significa que RisCanvi asigna sistemáticamente más riesgo del que realmente hay en los diferentes subgrupos considerados y en la muestra total».

COMENTARIO: Podría ser un problema general de estos sistemas. ¿Qué opináis?

El hecho de que RisCanvi no esté calibrado no es algo que ocurra por casualidad. El modelo podría ajustarse (por ejemplo, modificando los umbrales entre los distintos niveles de riesgo) para que los niveles medio y alto comprendieran una proporción que se asemejara más a la proporción de reincidentes conocidos. Pero esto tendría consecuencias: probablemente, la sensibilidad sería menor, lo que significaría que el modelo identificaría como de riesgo a menos individuos del grupo de los que realmente reinciden, y por tanto tendría más falsos negativos. Es una decisión que hay que tomar al diseñar cada herramienta, y al probar su rendimiento: a qué objetivo hay que dar prioridad, es decir, si queremos minimizar los falsos positivos o minimizar los falsos negativos.

En nuestro caso, solo pudimos comprobar el rendimiento de RisCanvi a la hora de evaluar el riesgo de un tipo muy especial de reincidencia: un nuevo delito violento que lleva a la cárcel. La prevalencia de este tipo de reincidencia es baja en las muestras (4,9 % en la muestra de Castillo; 10,9 % en la de Capdevila).

Cuando la prevalencia de un fenómeno es baja (en nuestro caso, un determinado tipo de reincidencia), es necesario fijar los umbrales de discriminación (entre bajo y alto riesgo) no demasiado altos para lograr una sensibilidad suficiente, es decir, para poder detectar un número relevante de verdaderos positivos. Pero, a cambio, el modelo tendrá que asumir también un número importante de falsos positivos y, en consecuencia, no estará bien calibrado. Cuando la prevalencia de un fenómeno es mayor, la mayoría de los verdaderos positivos pueden detectarse sin necesidad de bajar demasiado el umbral de discriminación, es decir, sin necesidad de asumir muchos falsos positivos, por lo que será más fácil tener tanto un modelo bien calibrado como una sensibilidad elevada. Es necesario destacar, no obstante, que estos son solo principios generales, que pueden cumplirse o no, en mayor o menor medida, en cada herramienta dependiendo de cómo esté diseñada, qué tipo de evento estime, con qué datos trabaje, etc.

Teniendo esto en cuenta, y volviendo al caso concreto de RisCanvi, hay tipos de reincidencia que esta herramienta también estima que tienen una tasa de prevalencia más alta: en la muestra de Capdevila, por ejemplo, la prevalencia de reincidencia general (comisión de cualquier tipo de nuevo delito) castigada con sanciones no privativas de libertad es del 29,85 % (desconocemos la cifra para la muestra de Castillo). Podría ser que respecto a este tipo de reincidencia RisCanvi esté mejor calibrado sin tener que sacrificar demasiado la sensibilidad. Pero, como no hemos podido analizar los datos de la escala general de reincidencia de RisCanvi, no podemos saberlo.

6. «RisCanvi no parece discriminar ni directa ni indirectamente a los extranjeros a la hora de evaluar el riesgo de reincidencia violenta. Clasifica como de alto riesgo a un mayor porcentaje de españoles (13,5 %) que de extranjeros (5,8 %), y la tasa de reincidencia violenta es de hecho mayor para los españoles (5,2 %) que para los extranjeros (4,3 %). De hecho, ser extranjero es un factor que computa -2 en la escala de riesgo de reincidencia violenta».

COMENTARIO: Me parece muy problemático aceptar sin más estas estadísticas: los índices de delincuencia, incluida la reincidencia, pueden ser muy fácilmente el resultado de perfiles étnicos, lo que lleva a que la policía detenga a más delincuentes de origen extranjero (inmigrantes, determinadas etnias) porque la policía se centra en ellos. ¿Podríais revisar vuestra evaluación en la sección sobre extranjeros teniendo en cuenta esta cuestión y comprobar si sigue siendo válida?

Es muy importante resaltar que todo lo que afirmamos en el texto de nuestro informe y en esta adenda se refiere únicamente a la reincidencia penitenciaria violenta, es decir, a la comisión de un nuevo delito violento que ha llevado al encarcelamiento. Desconocemos si RisCanvi discrimina o no a los extranjeros cuando evalúa el riesgo de reincidencia general, o el riesgo de ser condenado a una sanción penal que no conlleve pena de prisión, porque no hemos tenido acceso a muestras

de datos en las que se hayan estimado estos tipos de riesgo, y se hayan comparado con las tasas reales de reincidencia.⁴

Dicho esto, pensamos que nuestra conclusión (que el baremo REVI de RisCanvi no discrimina, ni directa ni indirectamente, a los extranjeros) sigue siendo válida.

Por lo que se refiere a los índices «oficiales» de delincuencia, es cierto que los índices de delincuencia, incluida la reincidencia, pueden ser muy fácilmente el resultado de la elaboración de perfiles étnicos, que hace que la policía detenga a más delincuentes de origen extranjero (inmigrantes, determinadas etnias) porque la policía se centra en ellos. Precisamente por ello, cabría esperar que las tasas oficiales de reincidencia fueran más elevadas en el caso de los extranjeros, pero no es así. O, para ser más precisos: lo son, efectivamente, para la reincidencia general, como explicamos en el texto de nuestro informe, pero no para la reincidencia violenta, que es el fenómeno que estamos analizando. Si los datos oficiales sobre tasas de reincidencia violenta fueran más elevados para los extranjeros que para los españoles, podríamos atribuirlo –al menos en parte– a una actuación policial con sesgo racial o étnico, y cuestionar así que las estadísticas oficiales reflejen las tasas reales de delincuencia. Pero como son más bajas, y no hay indicio alguno de que la policía española sea más indulgente con los extranjeros que con los nacionales (muy al contrario), pensamos que no hay razón para poner en duda que las cifras aquí reflejadas sí reflejan una tasa real de reincidencia violenta más baja para los extranjeros, y que RisCanvi es capaz de tener esto en cuenta, ya que asigna a los extranjeros un riesgo menor que a los españoles.

Esta menor tasa de reincidencia violenta de los extranjeros se mantiene en los distintos países de origen: como puede observarse en el análisis de la muestra de RisCanvi realizado por Karimi-Haghighi y Castillo (2022, tabla 3.4), no solo los nacionales de la UE tienen una tasa de reincidencia violenta inferior a la de los españoles, sino también las personas procedentes del Magreb y de América Central y del Sur. La mayoría de los inmigrantes irregulares en España proceden de estas dos últimas regiones, y hay muchas razones para temer que estas minorías sean las más vigiladas, porque en estos casos (junto a las personas procedentes de África, pero con respecto a estas el número en la muestra era demasiado pequeño para sacar conclusiones) la diferencia no es solo el país de origen, sino que también hay disparidades étnicas/raciales. Sin embargo, RisCanvi sigue asignando a estos grupos niveles de riesgo inferiores: como puede verse en el cuadro 3.4, el 75 % de las personas procedentes del Magreb y el 88,2 % de las personas procedentes del centro y Sudamérica se consideran de bajo riesgo, mientras que solo el 65,8 % de los

⁴ Hemos tenido acceso a este tipo de información (datos de reincidencia general, y datos de reincidencia con sanciones no privativas de libertad) en la muestra de Capdevila, pero –en nuestra opinión– no tiene sentido extraer conclusiones sobre la precisión o los sesgos para estos tipos de reincidencia a partir de los datos de esta muestra porque: 1) esta muestra no es representativa de la población liberada al año, sino que tiene una prevalencia muy alta de delincuentes de alto riesgo; y 2) esta muestra fue evaluada con la escala REVI, no con la nueva escala de reincidencia general, por lo que no consideramos adecuado juzgar el rendimiento de RisCanvi en la predicción de la reincidencia general con niveles de riesgo derivados de una escala que fue diseñada para estimar el riesgo de reincidencia violenta.

españoles lo son, y el 9,8 % de las personas procedentes del Magreb y el 2,9 % de las personas procedentes del centro y Sudamérica se consideran de alto riesgo, mientras que el 13,8 % de los españoles lo son.

En resumen: por supuesto, puede ser que la policía detecte un mayor porcentaje de delitos violentos perpetrados por extranjeros, y por minorías en general (que no tienen por qué ser extranjeras: España tiene un problema muy antiguo y grave de discriminación de la población gitana), porque se les vigila más intensamente. Pero RisCanvi funciona con la delincuencia detectada, tanto para nacionales como para extranjeros. Y, en este sentido, que sea capaz de asignar menor riesgo a los colectivos que (en las cifras oficiales) reinciden menos es, en nuestra opinión, un logro de la herramienta.

Paradójicamente, existe una norma, que en nuestra opinión es en sí misma una prueba de la discriminación institucionalizada de los extranjeros en España, que podría suscitar dudas sobre si estas cifras oficiales podrían estar *subestimando* las tasas reales de reincidencia violenta de los extranjeros: las deportaciones. La legislación española contiene disposiciones que permiten o incluso obligan a deportar a los inmigrantes en situación irregular, así como a los extranjeros condenados por un delito. El artículo 89 del Código Penal español establece que:

1. Las penas de prisión de más de un año impuestas a un ciudadano extranjero serán sustituidas por su expulsión del territorio español. Excepcionalmente, cuando resulte necesario para asegurar la defensa del orden jurídico y restablecer la confianza en la vigencia de la norma infringida por el delito, el juez o tribunal podrá acordar la ejecución de una parte de la pena que no podrá ser superior a dos tercios de su extensión, y la sustitución del resto por la expulsión del penado del territorio español. En todo caso, se sustituirá el resto de la pena por la expulsión del penado del territorio español cuando aquél acceda al tercer grado o le sea concedida la libertad condicional.
2. Cuando hubiera sido impuesta una pena de más de cinco años de prisión, o varias penas que excedieran de esa duración, el juez o tribunal acordará la ejecución de todo o parte de la pena, en la medida en que resulte necesario para asegurar la defensa del orden jurídico y restablecer la confianza en la vigencia de la norma infringida por el delito. En estos casos, se sustituirá la ejecución del resto de la pena por la expulsión del penado del territorio español, cuando el penado cumpla la parte de la pena que se hubiera determinado, acceda al tercer grado o se le conceda la libertad condicional. [...].

Y hay diferentes artículos en la Ley Orgánica 4/2000 (que regula los derechos y las obligaciones de los extranjeros en España) que también prevén la expulsión o deportación de los inmigrantes, por no tener la documentación requerida para una estancia legal, y también por haber sido condenados, por ejemplo, el artículo 57.2:

Asimismo, constituirá causa de expulsión, previa tramitación del correspondiente expediente, que el extranjero haya sido condenado, dentro o fuera de España, por una conducta dolosa que constituya en nuestro país delito sancionado con pena privativa de libertad superior a un año, salvo que los antecedentes penales hubieran sido cancelados.

El hecho de estar procesado o imputado en un procedimiento judicial por un delito para el que la ley prevea una pena privativa de libertad inferior a seis años

o una pena de distinta naturaleza también es causa de expulsión, según el artículo 57.7 de la Ley Orgánica 4/2000.

¿Podría ser que la tasa de delitos violentos, o la tasa de reincidencia violenta de los extranjeros fuera más alta de lo que parecen mostrar las cifras, pero que con las expulsiones/deportaciones se esté rebajando artificialmente? Es decir: que los extranjeros cometen –o cometerían– más delitos violentos, pero como son deportados lo hacen fuera de España y esos delitos no se pueden computar en las estadísticas. Creemos que no, por dos razones:

- No hay ninguna razón para suponer que los extranjeros cometen más delitos de los que muestran las estadísticas, salvo prejuicios racistas o xenófobos.
- Además, según las estadísticas oficiales, las expulsiones debidas al artículo 89 del Código Penal parecen ser poco frecuentes. Los datos oficiales son muy inconsistentes, como han denunciado numerosos académicos e investigadores, así como organizaciones civiles, pero, en cualquier caso, las expulsiones por la comisión de infracciones penales vienen oscilando entre 1.000 y 2.000 al año para toda España.⁵

No hemos encontrado cifras para Cataluña. Es la comunidad autónoma con mayor porcentaje de residentes extranjeros, por lo que, si las expulsiones tuvieran algún efecto en la tasa de reincidencia violenta de los extranjeros, debería ser más visible allí. Pero el informe de Capdevila et al. sobre la *Reincidencia en excarcelaciones* de alto riesgo (2014-2016) dice que «ha habido un notable descenso en la reincidencia por parte de extranjeros [respecto al anterior informe sobre reincidencia en excarcelaciones de alto riesgo, años 2010-2013]. No podemos saber si se debe a la aplicación de la expulsión administrativa del país como medida complementaria, dado que no sabemos a cuántos de ellos realmente se les ha aplicado, ni cuántos, si esto ha pasado, retornan de forma irregular» (Capdevila et al., 2022: 49). Como las expulsiones no han aumentado en los últimos años,⁶ esta no puede ser la razón de la disminución de la reincidencia de los extranjeros.

En resumen, pensamos que los extranjeros, y en particular los inmigrantes que difieren étnica y racialmente de la población nativa están, como han demostrado muchos estudios e informes, bajo un escrutinio policial más intenso, y sujetos a

⁵ Véanse la tabla 6 en p. 89 y la tabla 8 en p. 97 del informe de Berrio et al.: *Vulneraciones de Derechos Humanos en las deportaciones*, Iridia-Novact, 2020, con datos basados en una respuesta parlamentaria emitida por el Ministerio del Interior para este informe (<<https://iridia.cat/es/Publicaciones/vulneracions-de-drets-humans-en-les-deportacions/>>). Véase también M. Odriozola Gurrutxaga: *Expulsión penal y expulsión administrativa de personas extranjeras*, Thomson Reuters Aranzadi, 2022, pp. 190-193.

⁶ Odriozola Gurrutxaga afirma que «en el periodo comprendido entre 2009 y 2019, las expulsiones del artículo 89 nunca han supuesto más del 10 % del total de excarcelaciones de internos extranjeros. [...] Incluso si tenemos en cuenta las distintas modalidades de expulsión que implican el retorno del extranjero a su país de origen, desde 2009 nunca han superado el 30 % del total de excarcelaciones de internos extranjeros» (pp. 191 y ss.).

identificaciones y registros discriminatorios.⁷ Pero, sin embargo, no muestran una mayor tasa de reincidencia violenta y RisCanvi no les asigna como grupo (ni tampoco a diferentes grupos por región de origen) un mayor riesgo que a los españoles. Lo cual es algo que en nuestra opinión hay que reconocer como un mérito de la herramienta.

Dicho esto, sin embargo, hay que hacer una advertencia importante: las muestras con las que hemos trabajado no son muy grandes. La muestra de Castillo es $N = 2.514$ y la de Capdevila es $N = 1.072$. El hecho de que RisCanvi haya demostrado no discriminar a los extranjeros en cuanto a su riesgo de reincidencia violenta penitenciaria en estas muestras no significa que lo mismo fuera aplicable a otras muestras (por ejemplo, a la población excarcelada en años distintos a los analizados en estos estudios) o, como apuntábamos al principio, respecto a otros tipos de riesgo, como el de reincidencia violenta que no conduce a prisión, o el de reincidencia con cualquier delito (reincidencia general).

7. «La adición de más de un factor de riesgo relacionado con la situación socioeconómica en RisCanvi provoca, en ocasiones, incrementos del nivel de riesgo para los colectivos en los que esto ocurre totalmente desproporcionados respecto a su tasa real de reincidencia».

COMENTARIO: Aquí se me ocurre una pregunta importante: dado el elemento altamente subjetivo de muchos de los criterios: ¿hasta qué punto es grande el riesgo de que determinados criterios (por ejemplo, los problemas relacionados con el empleo o la gravedad de los problemas económicos) se evalúen con mayor severidad en el caso de una persona con escasa capacidad mental o de un extranjero? Esto podría implicar un problema incluso antes de que RisCanvi entre en juego, justo al nivel de los que introducen los datos, y podría incluso agravar los problemas de los que se habla aquí.

En relación con mi comentario anterior: ¿Es algo que podría tener relación con la forma en que las personas que introducen los datos perciben a una persona (el problema de la subjetividad), o lo que presentáis aquí sería una segunda capa de un problema similar que podría haberse producido en el momento de la introducción de los datos?

En cuanto a la subjetividad en la codificación de los factores de riesgo, hemos añadido en el texto una explicación bastante extensa de cómo los diseñadores de RisCanvi tratan de evitarla, estableciendo protocolos de trabajo aparentemente serios y rigurosos. Sin embargo, también hemos explicado en el texto cómo en nuestra opinión esta subjetividad es imposible de erradicar, ya que la propia descripción de los ítems obliga a realizar juicios de valor (como, por otra parte, no podía ser de otra manera teniendo en cuenta el tipo de factores de riesgo con los que trabaja RisCanvi). El problema de que en algunos casos personas con muchos factores de riesgo diferentes obtengan niveles de riesgo muchas veces superiores a los incrementos de

⁷ Por ejemplo, Martínez Escamilla y Sánchez Tomás: «Los requerimientos de identificación racistas y detenciones en el control migratorio», *Revista Crítica Penal y Poder* 18, 2019, en línea: <<https://revistes.ub.edu/index.php/CriticaPenalPoder/article/view/30410/30683>>.

las tasas de reincidencia puede estar, efectivamente, relacionado con el problema de la subjetividad, en el sentido de que quienes valoran la presencia o ausencia de los ítems pueden percibir cada factor como más problemático cuando lo ven concurrir con muchos otros. Pero también podría estar relacionado con el problema de la discriminación interseccional, es decir, «cuando dos o múltiples motivos de discriminación operan simultáneamente e interactúan de forma inseparable, produciendo formas distintas y específicas de discriminación».⁸ La combinación de muchos factores (factores de riesgo en el caso de RisCanvi) puede crear una sinergia negativa, de modo que el efecto global no puede entenderse plenamente como la suma de los criterios por sí solos, sino que acaba multiplicando los niveles de riesgo de forma desproporcionada.

Téngase en cuenta que aquí solo sugerimos posibles explicaciones para los efectos multiplicadores que se observan en los datos. Pero no podemos estar seguros de que alguna de ellas (o quizá otras) sean las causas reales o principales de los problemas. Para hacer un análisis mejor fundamentado hubiera sido necesario, por un lado, disponer directamente de una muestra de datos representativa del total de la población excarcelada cada año, para poder intentar construir distintas combinaciones entre los grupos y comprobar cuándo se producen en mayor medida estos efectos multiplicadores. Al tener que recurrir a cálculos que otros hacían por nosotros (profesor Castillo), nos hemos visto limitados en lo que podíamos pedir. En segundo lugar, también hubiera sido necesario poder realizar un estudio cualitativo, hablando con las personas que realizan las evaluaciones a diario, para conocer su opinión sobre estos casos y qué explicaciones dan a estos efectos.

8. «La Generalitat de Catalunya tiene algunos datos abiertos publicados en su web que nos han permitido –y permiten a cualquier investigador interesado– hacer algunos análisis; ha cedido previamente muestras a otros investigadores (el equipo del profesor Castillo), y ha encargado una auditoría externa de RisCanvi, cuyos resultados se harán públicos, o al menos eso nos han dicho. A la vista de todo ello, y aunque la herramienta no es transparente, hay más transparencia sobre RisCanvi, y probablemente la habrá, que sobre el resto de las herramientas que hemos estudiado en este informe, e incluso que sobre otros algoritmos utilizados por las administraciones públicas en España que trabajan con datos mucho menos sensibles que el riesgo de reincidencia de los presos».

COMENTARIO: Es una ventaja muy relativa que no cambia nada del problema general.

No estamos totalmente de acuerdo con este comentario. Pensamos que una política de datos abiertos que pone a disposición de cualquier persona interesada muestras de datos anonimizados es, efectivamente, una muy buena política de transparencia. Es cierto que la Generalitat ha aplicado esta política de datos abiertos solo a algunas de las investigaciones que ha llevado a cabo, pero –al menos en el contexto español y comparado con lo que hacen otras administraciones públicas– creemos

⁸ Definición extraída de la página web del Consejo de Europa: <<https://www.coe.int/en/web/gender-matters/intersectionality-and-multiple-discrimination#>>.

que merece un juicio muy favorable. La política de datos abiertos es mejor que poner los datos a disposición de académicos o investigadores concretos que presentan una petición o solicitud individual, porque en estos últimos casos los investigadores suelen estar sujetos a un acuerdo de confidencialidad, de manera que no pueden compartir los datos y las investigaciones que realizan no pueden ser objeto de revisión o verificación por parte de otros investigadores independientes. Por supuesto, esto no quiere decir que la Generalitat esté siendo totalmente transparente con RisCanvi; ya hemos subrayado en el texto que nuestra solicitud de información en particular no ha sido atendida hasta la fecha y que esto ha tenido un impacto negativo en nuestra investigación, limitando mucho el alcance de esta. Además, la falta de transparencia de RisCanvi para con las personas directamente afectadas por sus evaluaciones es también un problema muy grave. Pero nada de esto debe impedirnos reconocer que el esfuerzo que está haciendo la Generalitat con su política de datos abiertos respecto a –al menos algunos de– los estudios sobre población reclusa, en lo que a RisCanvi se refiere, es una refrescante novedad en el contexto español que, en nuestra opinión, debe ser aplaudida y bienvenida.

3. PREGUNTAS ESPECÍFICAS SOBRE VERIPOL

1. COMENTARIO: ¿Puede afectar a la precisión del algoritmo el hecho de que más de la mitad de los datos de entrenamiento fueran de informes falsos confirmados (538 frente a 522), lo que parece ser un porcentaje mucho mayor que en la realidad? Es decir, el algoritmo tuvo menos posibilidades de entrenarse con datos relacionados con denuncias verdaderas. ¿Quizás con un mayor número de informes verdaderos ciertos «patrones» identificados en los informes falsos también habrían empezado a surgir en los informes verdaderos (al menos con más frecuencia), reduciendo así la probabilidad de que una determinada característica sea un indicador de un informe falso?

Según los matemáticos de nuestro equipo, si se sabe (o se supone) que el porcentaje de denuncias verdaderas es mucho mayor que el porcentaje de denuncias falsas, es razonable «sobremuestrear» la categoría infrarrepresentada (denuncias falsas) al ajustar un modelo. De lo contrario, una muestra aleatoria de informes podría llevar a considerar un número excesivamente pequeño de informes falsos, con la consiguiente pérdida de potencia en el análisis estadístico. Si ponemos, por ejemplo, un 90/10 % de informes verdaderos y falsos en los datos de entrenamiento porque entendemos que esa es la proporción real, y el algoritmo clasificará el 100 % de los informes como verdaderos, estaría dando un «rendimiento» fantástico de forma artificial, ya que solo se equivocaría en el 10 % de los casos, pero a costa de no poder detectar el problema que se supone que debe descubrir.

Por lo general, no es necesario equilibrar artificialmente las proporciones de ejemplos positivos y negativos en los datos de entrenamiento para un algoritmo de aprendizaje automático que solo tiene dos salidas posibles (verdadero y falso). El algoritmo aprenderá a predecir la clase más frecuente («verdadero» en este caso) con mayor precisión, pero también debería ser capaz de predecir la clase menos

frecuente («falso» en este caso) hasta cierto punto, siempre que los datos de entrenamiento contengan suficientes ejemplos de ambas clases.

Sin embargo, si los datos de entrenamiento están muy desequilibrados, con un número muy pequeño de ejemplos negativos, el rendimiento del algoritmo en la clase menos frecuente puede ser pobre. En este caso, puede ser beneficioso equilibrar los datos de entrenamiento sobremuestreando la clase minoritaria (añadiendo más ejemplos de la clase «falsa») o submuestreando la clase mayoritaria (eliminando algunos ejemplos de la clase «verdadera») para lograr un mejor equilibrio. Esto puede ayudar al algoritmo a aprender a predecir la clase minoritaria con mayor precisión.

Lo que podría ser un problema en el caso de VeriPol es que el número de denuncias con las que se ha entrenado el modelo (1.122) es pequeño comparado con el número de denuncias por robos violentos que se presentan cada año (unos 30.000 robos con intimidación en la vía pública al año). Serían necesarios más datos para entrenar adecuadamente el algoritmo.

Desarrollo territorial *papers*



¿Cuál es la capacidad predictiva de los algoritmos diseñados para prevenir la delincuencia? ¿Son eficaces? ¿Son lo suficientemente transparentes? ¿Tienen sesgos? Para responder a estas preguntas, este estudio analiza tres herramientas algorítmicas utilizadas por la Administración española para hacer estimaciones de riesgo en los ámbitos policial y penitenciario (VioGén, RisCanvi y VeriPol, esta última con inteligencia artificial) con la finalidad de evaluar su funcionamiento, en particular en términos de su respeto a los derechos fundamentales.

El estudio combina métodos cuantitativos y cualitativos, y ha sido realizado por profesores de la Universitat de València y de la Universitat Politècnica de València, expertos en distintas ramas del derecho (penal, administrativo y constitucional), así como en estadística y matemáticas. Los objetivos principales son analizar la capacidad predictiva de los tres sistemas y su eficacia para alcanzar el objetivo que pretenden (reducción de la delincuencia); valorar su grado de transparencia, tanto para la ciudadanía en general como para las personas directamente afectadas por su uso; identificar y discutir posibles sesgos, y comprobar si respetan el marco legal europeo y español vigente, prestando especial atención a la normativa sobre protección de datos personales y a la regulación del uso de sistemas automatizados de toma de decisiones.

También se analiza si los costes para los derechos fundamentales que conlleva el uso de estas herramientas exceden lo que sería legítimo, necesario y/o proporcionado en relación con el beneficio actual o probable en términos de prevención del delito. El estudio subraya la importancia de la transparencia para hacer posible el necesario debate público que debe existir sobre herramientas que, como las tres analizadas, son utilizadas por los poderes públicos como apoyo para tomar decisiones que pueden redundar en restricciones de derechos fundamentales.